

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



**ANÁLISIS Y PROPUESTAS DE MEJORA DE PROCESO APLICANDO
HERRAMIENTAS DE ANALYTICS EN UNA EMPRESA DE SERVICIOS
AEROPORTUARIOS**

Tesis para obtener el título profesional de Ingeniera Industrial

AUTORAS:

Romina Gonzales Soto

Alexandra Natividad Herrera Tafur

ASESOR:

Renzo Alejandro Benavente Sotelo


Lima, 2026

Informe de Similitud

Yo, Renzo Alejandro Benavente Sotelo, docente de la Facultad de Ciencias e Ingeniería de la Pontificia Universidad Católica del Perú, asesor de la tesis titulada
ANÁLISIS Y PROPUESTAS DE MEJORA DE PROCESO APLICANDO HERRAMIENTAS DE ANALYTICS EN UNA EMPRESA DE SERVICIOS AEROPORTUARIOS,
de las autoras Romina Gonzales Soto y Alexandra Natividad Herrera Tafur,
dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 6%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 24/03/2026.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de Suficiencia Profesional, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha: San Miguel, 24/03/2026

Apellidos y nombres del asesor: <u>Benavente Sotelo, Renzo Alejandro</u>	
DNI: 70007275	
ORCID: 0000-0002-3192-2249	
Firma	



AGRADECIMIENTOS

Una gran mención de agradecimiento al ingeniero Renzo Benavente, por su asesoría y aporte de su experiencia profesional y observaciones críticas fueron claves para el desarrollo de esta investigación. Asimismo, agradecemos a la empresa aeroportuaria donde trabajamos por facilitar el acceso a la información necesaria, permitiendo que esta investigación se consolide con éxito. Por último, agradecemos a la Pontificia Universidad Católica del Perú por los años de formación académica y brindarnos una base sólida de aprendizaje para nuestro crecimiento profesional a lo largo de nuestras carreras.

(Alexandra Herrera – Romina Gonzales)

Quiero expresar mi profundo agradecimiento en primer lugar a Dios por guiar mis pasos en este proceso tan arduo. A mi familia, por su amor incondicional, su apoyo constante y por brindarme la oportunidad de llegar tan lejos en mi formación académica. A mi querido Brandon, mi fiel amigo perruno, cuya presencia llenó de calma y amor cada jornada. A mis mejores amigos, Priscila y John, que me acompañaron en este proceso con mucho ánimo y motivación. A Ernesto, por caminar a mi lado, por su paciencia y comprensión durante la culminación de esta tesis. A todos los amigos y familiares que me acompañaron de una u otra forma en las distintas etapas de este proceso. A Romi, mi compañera de tesis, por todo su compromiso y dedicación en esta tesis, agradezco los años de amistad y el poder cerrar este proyecto juntas, le deseo, además, los mejores éxitos en su vida personal y laboral.

(Alexandra Herrera)

Expreso mi más profundo agradecimiento a mis padres, por ser mi motor, constante apoyo y la motivación principal durante todo este proceso. Les dedico esta tesis como símbolo de mi crecimiento y éxito profesional. Agradezco a Dios por darme fuerzas para continuar con este proceso de vital importancia en mi crecimiento profesional y permitirme seguir avanzando en mi carrera. Un agradecimiento especial al Gremio de Aventureras, Ale y Kathy; gracias por quince años de amistad desde el colegio y por siempre brindarme los ánimos para perseverar hasta el final. A Hana, mi querida gata, por su fiel compañía durante las largas jornadas de trabajo. A Anita y Fati, mis amigas desde la infancia hasta ahora, quienes siempre estuvieron apoyándome con gran ánimo para culminar la tesis. A Piero, por acompañarme en esta última etapa del proceso y brindarme apoyo emocional para alcanzar esta meta. A mi familia y seres queridos que estuvieron pendientes por saber cómo iba con este proceso hasta el final. Finalmente, quiero expresar mi gratitud a Ale, mi compañera de tesis, por su dedicación, esfuerzo y por el gran equipo que formamos durante estos últimos años para culminar juntas este proyecto, a quien deseo los mejores éxitos en su siguiente etapa profesional.

(Romina Gonzales)

RESUMEN

El estudio de tesis desarrollado tiene como objeto de análisis a una empresa proveedora de servicios terrestres a las aerolíneas que operan en el aeropuerto de Lima, con enfoque en el área de Planificación para la atención de vuelos de la unidad de negocio Rampa. En ese sentido, la empresa actualmente enfrenta una problemática relacionada con el incumplimiento de metas en las métricas de servicio al cliente tales como demoras de vuelos y nivel de servicio; las cuales generan penalidades e impactan negativamente en su desempeño; así mismo, incurre en altos costos de planilla y movilidad, distorsionando así los ratios financieros. Además, existe una excesiva rotación de personal y un control ineficiente de horas, los cuales repercuten en reprocesos para el área de Reclutamiento de nuevas contrataciones y horas extras del personal que afectan la productividad respectivamente.

El objetivo fundamental es desarrollar propuestas de mejora empleando herramientas de Data Analytics aplicado a los procesos de la compañía. En ese sentido, se realizó el diagnóstico y se aportaron las siguientes tres soluciones: En primer lugar, el diseño de un modelo de predicción de demanda con herramientas de aprendizaje automático en el software R; en segundo lugar, el diseño de un modelo de programación lineal para optimizar la cantidad personal y turnos; y por último, el diseño de un escenario de simulación sobre adelantos y retrasos de vuelos en el software Arena, para validar que la cantidad de grupos de personal obtenida en la propuesta de optimización cumpla con atender la demanda tomando en cuenta las variaciones de la realidad operativa como demoras y adelantos de vuelos.

En línea con los modelos diseñados, se procedió a validar cada uno de ellos, a fin de garantizar su viabilidad en el contexto del estudio. Respecto a la primera solución se comprobó que el modelo SARIMA, comparado al modelo de regresión lineal y al de ARIMA, presentaba una opción más acorde en tendencia de crecimiento y patrón de comportamiento respecto a los datos reales, esto reflejado en las métricas AIC, MAE y RSME. Para la siguiente parte de la propuesta, se obtuvo un resultado de requerimiento de 42 grupos de personal, cuyos costos equivalen a una reducción del 30% comparado al requerimiento actual (52 grupos de personal). Finalmente, se ejecutó las simulaciones que demostraron que la optimización de los turnos y grupos de personal mejoró indicadores como utilización, productividad y redujo las horas muertas; sin embargo, generó una degradación en el indicador de demoras. A fin de evitar esto, se planteó como alternativa integrar dos grupos adicionales que no reduzcan mucho los ahorros generados. Los resultados de los modelos antes planteados arrojaron un ahorro significativo de S/66,481 mensuales. En conclusión, esta propuesta integral generaría gran ahorro a la empresa mediante el reclutamiento de personal anticipado, la optimización de turnos y la mejora de los indicadores operativos y comerciales. Es importante mencionar que, estas soluciones requieren cierto mantenimiento periódico para garantizar que los modelos permanezcan actualizados respecto al contexto real, que puede verse afectado por variables externas.

ÍNDICE GENERAL

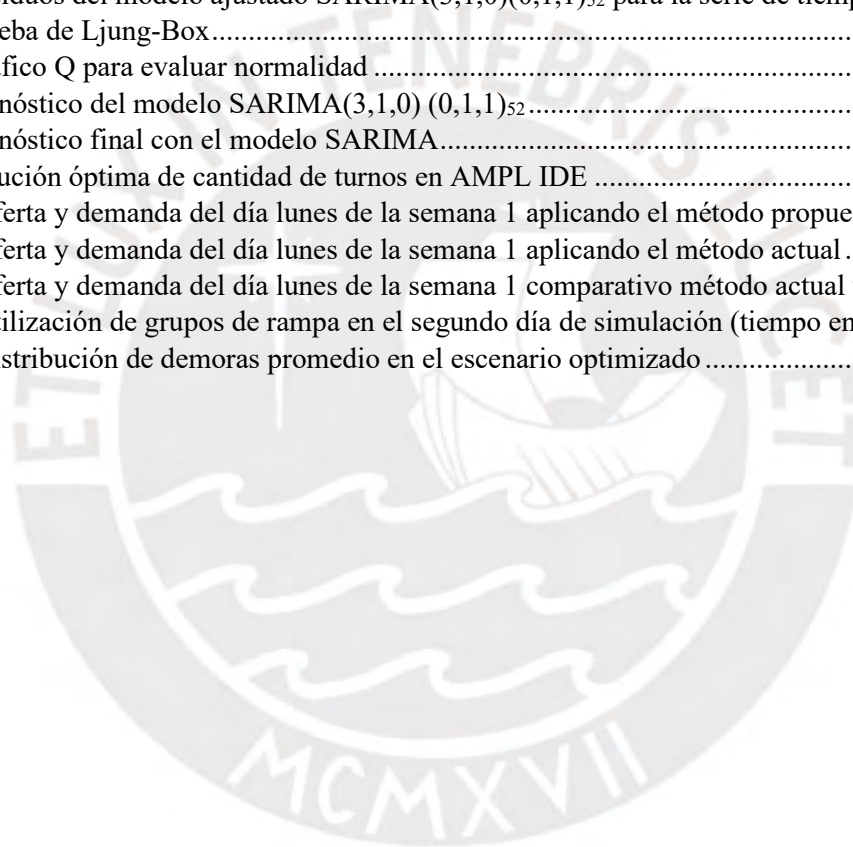
ÍNDICE DE FIGURAS.....	iv
ÍNDICE DE TABLAS	vi
ÍNDICE DE ANEXOS	viii
INTRODUCCIÓN	1
CAPITULO 1. MARCO TEÓRICO	3
1.1 Conceptos preliminares <i>Machine Learning</i>	3
1.1.1 Conceptos generales.....	3
1.1.2 Tipos de aprendizaje	4
1.1.3 Modelos de Machine Learning.....	6
1.1.4 <i>Workflow</i> : Ciclo de la analítica	18
1.2 Modelos de Optimización.....	19
1.2.1 Conceptos Generales.....	19
1.2.2 Formulación de un modelo de programación lineal.....	20
1.2.3 Modelo de programación de la capacidad del recurso humano	21
1.3 Simulación.....	22
1.3.1 Conceptos previos	22
1.3.2 Definición.....	25
1.3.3 Elementos de un modelo de simulación.....	25
1.3.4 Modelos de simulación	25
1.3.5 Análisis de datos	26
1.4 Estado del arte	28
1.4.1 Caso 1: Long-Term And Short-Term Forecasting Techniques For Regional Airport Planning	28
1.4.2 Caso 2: A workforce planning and allocation model for the outbound baggage loading area at Santiago International Airport.....	30
1.4.3 Caso 3: Propuesta de mejora de procesos en la atención de aeronaves en tránsito de la flota Airbus de LAN Perú, para incrementar la productividad en el área de mantenimiento	33
CAPITULO 2. LA EMPRESA	36
2.1 Descripción de la empresa.....	36
2.1.1 Descripción general.....	36
2.1.2 Organigrama.....	37
2.1.3 Proceso de Planificación	37
2.2 Análisis de la situación actual y diagnóstico	43
2.2.1 Situación actual	43
2.2.2 Diagnóstico	45
CAPITULO 3. PROPUESTA DE MEJORA	50

3.1	Metodología del problema.....	50
3.2	Diseño de un modelo de predicción de demanda con herramientas de aprendizaje automático.....	53
3.2.1	Recopilación de data histórica de vuelos	53
3.2.2	Análisis exploratorio de la data.....	54
3.2.3	Preprocesamiento de la data.....	54
3.2.4	Planteamiento de herramientas de forecasting.....	59
3.2.5	Distribución del forecasting de vuelos	70
3.3	Diseño de un modelo de programación lineal para la optimización de horarios y personal.....	74
3.3.1	Preparación de los datos.....	75
3.3.2	Planteamiento del modelo de programación de la capacidad del recurso humano	75
3.4	Diseño de escenarios de simulación de la situación actual y la propuesta de mejora	79
3.4.1	Componentes del modelo.....	79
3.4.2	Descripción del modelo.....	80
3.4.3	Recolección de los datos	81
3.4.4	Análisis de los datos.....	81
3.4.5	Modelamiento del sistema de simulación	87
CAPITULO 4. RESULTADOS		89
4.1	Predicción de demanda mediante el aprendizaje automático	89
4.2	Programación optimizada de la capacidad del recurso humano	94
4.3	Simulación del escenario actual y la propuesta de mejora	97
4.3.1	Validación de resultados del modelo	97
4.3.2	Comparación de indicadores de desempeño	98
4.4	Evaluación de la propuesta de mejora	100
CAPITULO 5. CONCLUSIONES Y RECOMENDACIONES		101
5.1	Conclusiones.....	101
5.2	Recomendaciones	102
GLOSARIO DE TÉRMINOS.....		103
BIBLIOGRAFÍA		104
ANEXOS		107

ÍNDICE DE FIGURAS

Figura 1.1 Esquema general de investigación de forecasting basada en big data.....	4
Figura 1.2 Tipos de aprendizaje.....	5
Figura 1.3 Series de tiempo mensual.....	7
Figura 1.4 Esquema de la Metodología Box-Jenkins para modelo ARIMA.....	9
Figura 1.5 PACF de una serie de tiempo.....	11
Figura 1.6 ACF de una serie de tiempo.....	11
Figura 1.7 Comportamiento evolutivo de los residuos del modelo ARIMA.....	13
Figura 1.8 ACF de los residuos del modelo ARIMA.....	13
Figura 1.9 Normalidad de los residuos del modelo ARIMA.....	14
Figura 1.10 K Vecinos más cercanos.....	15
Figura 1.11 Árboles de decisión.....	15
Figura 1.12 Redes de alimentación directa de una sola capa.....	17
Figura 1.13 Redes de alimentación directa multicapa.....	17
Figura 1.14 Redes recurrentes.....	18
Figura 1.15 Ciclo del business analytics.....	18
Figura 1.16 Nivel de significancia del 5%.....	24
Figura 1.17 Rendimiento del modelo de series de tiempo.....	29
Figura 1.18 Predicción de la demanda de pasajeros por trimestre.....	30
Figura 1.19 Ejemplos de distribución de llegadas de maletas durante el período de preparación del vuelo anterior a la salida de un vuelo nacional e internacional.....	31
Figura 1.20 Congestión de bolsas en carruseles: modelo versus real. Carruseles M3-M6 durante un día representativo.....	33
Figura 1.21 Resultados de la corrida del proceso actual en el software Arena.....	34
Figura 1.22 Resultados de la corrida del proceso propuesto en el software Arena.....	35
Figura 2.1 Ejemplo de curva de demanda.....	42
Figura 2.2 Ejemplo de curva de demanda con oferta.....	42
Figura 3.1 Esquema de metodología del problema.....	50
Figura 3.2 Histórico de vuelos del 2015 al 2023.....	54
Figura 3.3 Vuelos ajustados por factor pandemia.....	55
Figura 3.4 Evolutivo de demanda histórica final.....	59
Figura 3.5 Serie de tiempo de vuelos semanales.....	60
Figura 3.6 Evolutivo del train de vuelos semanales.....	60
Figura 3.7 Resumen del modelo de Regresión Lineal.....	61
Figura 3.8 Prueba de Dickey-Fuller Aumentada para la serie de tiempo de vuelos.....	62
Figura 3.9 Serie diferenciada de primer orden.....	63
Figura 3.10 Prueba de Dickey-Fuller Aumentada para la serie diferenciada.....	63
Figura 3.11 Autocorrelación parcial de la serie diferenciada.....	64
Figura 3.12 Autocorrelograma de la serie diferenciada.....	64
Figura 3.13 Resumen del modelo ARIMA(2,1,2).....	65
Figura 3.14 Descomposición aditiva de series de tiempo.....	66
Figura 3.15 Serie diferenciada estacionalmente.....	67
Figura 3.16 Prueba de Dickey Fuller Aumentada para la serie estacionalmente diferenciada.....	67
Figura 3.17 Serie doblemente diferenciada.....	68
Figura 3.18 Prueba de Dickey Fuller Aumentada para la serie doblemente diferenciada.....	68
Figura 3.19 Autocorrelación parcial en serie doblemente diferenciada.....	69
Figura 3.20 Autocorrelación en serie doblemente diferenciada.....	69
Figura 3.21 Resumen del modelo ARIMA(3,1,0)(1,1,0) _{s2}	69
Figura 3.22 Demanda de vuelos por día de una semana de muestra.....	70
Figura 3.23 Distribución de demanda por día de una semana de muestra.....	70

Figura 3.24 Demanda de personal por franja horaria de un día de muestra.....	73
Figura 3.25 Distribución de demanda por franja horaria de un día de muestra	73
Figura 3.26 Planteamiento del modelo de programación de la capacidad del recurso humano.....	79
Figura 3.27 Análisis de varianza de dos factores para determinar la dependencia entre el tiempo de atención de aeronaves con referencia al fuselaje y tipo de atención.....	82
Figura 3.28 Análisis de varianza de dos factores para determinar la dependencia entre el tiempo de llegadas entre aeronaves con referencia al día de semana y franja horaria.....	83
Figura 3.29 Definición del elemento “Schedule” en el modelo de simulación para la atención de aeronaves.....	87
Figura 3.30 Definición del elemento “Resources” en el modelo de simulación para la atención de aeronaves.....	87
Figura 4.1 Pronóstico del modelo de regresión lineal.....	89
Figura 4.2 Pronóstico del modelo ARIMA(2,1,2)	90
Figura 4.3 Resumen del modelo SARIMA(3,1,0)(0,1,1) ₅₂	90
Figura 4.4 Residuos del modelo ajustado SARIMA(3,1,0)(0,1,1) ₅₂ para la serie de tiempo	91
Figura 4.5 Prueba de Ljung-Box.....	91
Figura 4.6 Gráfico Q para evaluar normalidad	92
Figura 4.7 Pronóstico del modelo SARIMA(3,1,0) (0,1,1) ₅₂	92
Figura 4.8 Pronóstico final con el modelo SARIMA.....	93
Figura 4.9 Solución óptima de cantidad de turnos en AMPL IDE	94
Figura 4.10 Oferta y demanda del día lunes de la semana 1 aplicando el método propuesto.....	95
Figura 4.11 Oferta y demanda del día lunes de la semana 1 aplicando el método actual	96
Figura 4.12 Oferta y demanda del día lunes de la semana 1 comparativo método actual y propuesto.....	96
Figura 4.13 Utilización de grupos de rampa en el segundo día de simulación (tiempo en minutos)....	98
Figura 4.14 Distribución de demoras promedio en el escenario optimizado	99



ÍNDICE DE TABLAS

Tabla 1.1 Condiciones para seleccionar el tipo de distribución de probabilidad.....	24
Tabla 1.2 Validación cruzada para modelos de hora de la semana.....	30
Tabla 1.3 Resumen de resultados.....	33
Tabla 2.1 Ejemplo de proyección de demanda inicial abril 2023	38
Tabla 2.2 Ejemplo de cumplimiento histórico de demanda 2022	38
Tabla 2.3 Ejemplo de proyección de demanda final abril 2023.....	39
Tabla 2.4 Ejemplo de ratio oferta-demanda del primer trimestre 2023	40
Tabla 2.5 Ejemplo de proyección de la oferta abril 2023	40
Tabla 2.6 Ejemplo de itinerario de la Aerolínea A	41
Tabla 2.7 Ejemplo de parámetro para la aerolínea A.....	41
Tabla 2.8 Ejemplo de tiempos de atención	42
Tabla 2.9 Ejemplo de turnos diarios	42
Tabla 2.10 Problemática de situación actual de la empresa.....	45
Tabla 2.11 Propuestas de solución sobre herramientas de Analytics.....	49
Tabla 3.1 Descripción de variables de la data.....	53
Tabla 3.2 Muestra de la data histórica de vuelos	53
Tabla 3.3 Data de vuelos afectados por la tercera ola del COVID-19	56
Tabla 3.4 Data ajustada de vuelos afectados por la tercera ola del COVID-19.....	56
Tabla 3.5 Data de vuelos afectados por accidente aeroportuario	57
Tabla 3.6 Data ajustada de vuelos afectados por accidente aeroportuario.....	57
Tabla 3.7 Data de vuelos afectados por manifestaciones nacionales.....	57
Tabla 3.8 Data ajustada de vuelos afectados por manifestaciones nacionales.....	58
Tabla 3.9 Ejemplo de la nueva versión de la data.....	58
Tabla 3.10 Separación de datos.....	59
Tabla 3.11 Muestra de vuelos de las primeras diez semanas del 2023	71
Tabla 3.12 Distribución de vuelos de las primeras diez semanas del 2023	71
Tabla 3.13 Resultados de pruebas de hipótesis para cada día de la semana	72
Tabla 3.14 Resumen de distribución de demanda por día de semana.....	72
Tabla 3.15 Resumen de distribución de demanda por hora del día	74
Tabla 3.16 Turnos de trabajo	76
Tabla 3.17 Ejemplo de turno con días de descanso	76
Tabla 3.18 Ejemplo de turnos con refrigerio	77
Tabla 3.19 Tipos de fuselaje	82
Tabla 3.20 Definición de variables por tipo de fuselaje y tipo de atención	82
Tabla 3.21 Proporción de demanda Setiembre 2023 por día de semana.....	83
Tabla 3.22 Grupos por días de semana	83
Tabla 3.23 Proporción de demanda Setiembre 2023 por franja horaria.....	84
Tabla 3.24 Grupos por franja horaria.....	84
Tabla 3.25 Tamaño de población de las variables para el tiempo de atención de aeronaves.....	85
Tabla 3.26 Tamaño de población de las variables para el tiempo entre llegada de aeronaves	85
Tabla 3.27 Tamaño de muestra de las variables de tiempo de atención de la aeronave	86
Tabla 3.28 Tamaño de muestra de las variables de tiempo entre llegada de aeronaves	86
Tabla 3.29 Cálculo del error de la muestra preliminar.....	88
Tabla 3.30 Cálculo del número de réplicas para la simulación.....	88
Tabla 4.1 Matriz de métricas de los modelos.....	93
Tabla 4.2 Resumen de cantidad de grupos y costos del método propuesto	95
Tabla 4.3 Resumen de cantidad de grupos y costos del método actual	96
Tabla 4.4 Ahorro generado entre el método propuesto y actual	96

Tabla 4.5 Comparativo de indicadores de utilización y horas muertas entre el método actual y propuesto.....	97
Tabla 4.6 Comparativo de indicadores entre el método actual y optimizado	98
Tabla 4.7 Penalidades de demora semanales por rango	99
Tabla 4.8 Comparativo de indicadores entre el escenario actual, optimizado y optimizado ajustado	100
Tabla 4.9 Ahorro total del escenario optimizado vs el actual sobre las propuestas de Pronóstico, Optimización y Simulación de la atención de vuelos	100



ÍNDICE DE ANEXOS

Anexo 1 Ejemplo de programación lineal.....	107
Anexo 2 Modelos de demanda por día de semana: Método de promedio simple y método LOESS..	109
Anexo 3 Modelo de generación de tareas	110
Anexo 4 Curva comparativa propuesta vs real	112
Anexo 5 Flujo de actividades del técnico en el proceso en el software Arena: Actual y propuesto...	113
Anexo 6 Organigrama de la empresa.....	114
Anexo 7 Organigrama de la unidad de negocio	115
Anexo 8 Proceso de Planificación de la demanda y oferta	116
Anexo 9 Evolutivo de demoras de la aerolínea A, nivel de SNC y costos de personal contratado 2022	117
Anexo 10 Evolutivo de costos de movilidad, nivel de rotación de personal y productividad 2022 ...	118
Anexo 11 Diagrama de Pareto de la problemática actual	119
Anexo 12 Diagramas de Ishikawa: Altos costos de personal y movilidad	120
Anexo 13 Diagramas de Ishikawa: Incumplimiento del nivel de servicio y excesiva rotación de personal.....	121
Anexo 14 Clasificación ABC de causas	122
Anexo 15 Gráfica estacional de vuelos por años	123
Anexo 16 Código en RStudio para la aplicación de pruebas Kolmogorov-Smirnov por día de semana	124
Anexo 17 Muestra de demanda de personal por hora y distribución de demanda por hora en puntos porcentuales	125
Anexo 18 Resultados de pruebas de hipótesis para cada hora del día	126
Anexo 19 Histogramas de Datos para las franjas horarias de: 01:00, 02:00 y 03:00 h.....	127
Anexo 20 Resumen de demanda de la semana 1 por franja horaria y día de semana	128
Anexo 21Detalle de turnos.....	129
Anexo 22 Costos por turnos y horarios.....	133
Anexo 23 Datos de demanda, costo y turnos ingresados al software AMPL	134
Anexo 24 Clasificación de variables aleatorias del tiempo entre llegadas de aeronaves.....	135
Anexo 25 Pruebas de bondad de ajuste para las variables de tiempo entre llegadas de aeronaves y tiempo de atención de aeronaves	136
Anexo 26 Resumen de distribuciones por variable aleatoria.....	144
Anexo 27 Bloques del modelo de simulación para la atención de aeronaves.....	145
Anexo 28 Sistema de simulación propuesto para la atención de aeronaves en el aeropuerto Jorge Chávez	146
Anexo 29 Muestra inicial de 15 réplicas sobre la demanda semanal de atenciones	147
Anexo 30 Detalle de turnos de personal para la semana 1	148
Anexo 31 Oferta y demanda de la semana 1 aplicando el método propuesto.....	149
Anexo 32 Oferta y demanda de la semana 1 aplicando el método actual	151
Anexo 33 Validación de resultados de la variable tiempo de entre llegadas de aeronaves	153

INTRODUCCIÓN

Siendo la aviación uno de los sectores clave para el transporte de pasajeros y carga a nivel mundial, la demanda de vuelos de pasajeros en 2021 alcanzó un 47% del nivel de la operación durante el 2019 de acuerdo con la Asociación Internacional de Transporte Aéreo (IATA), puesto que fue afectada por la pandemia iniciada en 2020. Así mismo, la IATA estimó una recuperación global en 2023 de 94% respecto a niveles pre-covid del 2019, donde el negocio internacional alcanzaría un 82% y el doméstico un 103% con un escenario más optimista. Adicionalmente, en cuanto a los vuelos operados dentro de Sudamérica, la recuperación proyectada para los vuelos es de 97% para el mismo año. Esta recuperación permitirá la reactivación y crecimiento de otros sectores relacionados con la aviación como el turismo, hotelería y logística y transporte de carga. Por otro lado, las aerolíneas y empresas proveedoras de servicios terrestres dentro de un aeropuerto toman como referencia el denominado Acuerdo Estándar de Servicios de Asistencia en Tierra (SGHA) definido para entablar negociaciones de un contrato entre ambas partes; puesto que este acuerdo consiste en determinar condiciones de los procesos que se van a ejecutar por la empresa proveedora. En ese sentido, la SGHA contiene artículos de temas como provisión de servicios, buenas prácticas, estándar de trabajo, subcontratación de servicios, entre otros. Así mismo, el SGHA se complementa con el Acuerdo de Nivel de Servicio (SLA), el cual plantea cuáles serán los indicadores claves de rendimiento (KPIs), cómo se realizan las mediciones de los KPIs y qué acciones tomar de acuerdo con los resultados de estos. En consecuencia, un acuerdo transparente permite una relación comercial sólida entre ambas partes con el fin de mantener el vínculo contractual a largo plazo.

Por otra parte, la industria aeroportuaria inició una recuperación operativa en el año 2021 y alcanzó un 86.48% a nivel global, donde diversas aerolíneas lograron una mejora en cuanto al desempeño de este KPI con respecto al año anterior. Se resalta la importancia de medir la puntualidad puesto que los resultados permiten analizar las áreas críticas de operación e identificar optimización de recursos como personal, flota de aviones, entre otros. Así mismo, cabe destacar que los vuelos puntuales contribuyen con incrementar la satisfacción y fidelidad del pasajero; por lo tanto, se disminuye la tasa de reclamos y pérdidas de conexiones. Finalmente, se resalta que la puntualidad es un KPI estratégico que permite posicionar la marca de las aerolíneas dentro de la misma región y destacar su fiabilidad con sus stakeholders.

Ciertamente, el desempeño de las operaciones dentro de un aeropuerto es el reflejo de la gestión de todos los stakeholders y, en la cadena de procesos de atención terrestre, la planificación es esencial. Por ello, es fundamental estimar una proyección precisa sobre la demanda de vuelos, tal como el caso del aeropuerto de Bolonia en Italia, donde se destaca la importancia del pronóstico de demanda de pasajeros aplicando técnicas de series de tiempo a corto y largo plazo. Además de ello, otra estimación importante

es la oferta del recurso humano, donde se desarrollaron modelos de programación lineal entera para definir la cantidad de personal óptima y turnos para atender la zona de equipajes en el aeropuerto internacional Arturo Merino Benitez de Chile. Por último, otra herramienta empleada en la aviación es la simulación, donde se aplica al proceso de mantenimiento a una aeronave en tierra con el fin de incrementar la productividad del área.

El objetivo general es desarrollar el diagnóstico de la situación actual de la empresa aeroportuaria en estudio y posteriormente, proponer alternativas de mejora empleando herramientas de Data Analytics, lo cual implica investigar sobre modelos de Machine Learning, Optimización y Simulación aplicados al contexto aeroportuario.

El presente trabajo se segmenta en cinco capítulos. En el primer capítulo, se establece el marco teórico con los conceptos que abarcan las herramientas de Machine Learning, Optimización y Simulación. Así mismo, se expondrán casos de referencia para cada tema relacionado a un contexto aeroportuario. En el segundo capítulo, se desarrolla la descripción sobre la empresa en estudio y cuál es el panorama de la situación actual que presenta dentro del aeropuerto de Lima como parte del diagnóstico. En el tercer capítulo, se estarán planteando las propuestas que contribuyan a mejorar la planificación de la demanda de vuelos y oferta de personal en la empresa, utilizando herramientas de Machine Learning, Optimización y Simulación. En el cuarto capítulo, tras finalizar el diseño de las propuestas y realizar las corridas de cada herramienta, se estarán comparando los resultados de cada escenario actual y optimizado respectivamente con el fin de evaluar el impacto de cada mejora. En el quinto capítulo, se presentan las conclusiones sobre los resultados operativo y financiero de las propuestas planteadas; así como las recomendaciones que debe seguir la empresa para garantizar su sostenibilidad en el tiempo.

CAPITULO 1. MARCO TEÓRICO

El presente capítulo abarca el desarrollo conceptual de tres temas pilares de este estudio, los cuales son *Machine Learning*, Modelos de Optimización y Simulación; así como el estado del arte, donde se describen casos representativos de cada tema.

1.1 Conceptos preliminares *Machine Learning*

En primer lugar, se desarrollan las definiciones generales de *Machine Learning*, así como la clasificación de los tipos de aprendizaje. Posteriormente, se explica acerca de los diversos modelos que lo conforma y, por último, se describe el ciclo de la analítica.

1.1.1 Conceptos generales

Las definiciones que se detallarán respectivamente son: *Big Data*, *Analytics* y *Machine Learning*

Big Data

De acuerdo con los autores Tang et al. (2022), se presentan tres pasos claves que se aplican en el desarrollo de investigaciones sobre *forecasting* para *big data*: colección de la data, procesamiento de la data y mejora de la predicción. A continuación, en la Figura 1.1 se ilustra la secuencia de pasos a seguir mencionados para big data.

A continuación, se describen cada uno de los tres pasos mencionados:

- Colección de la data: Obtener la big data de distintas fuentes asociadas al tema de estudio
- Procesamiento de datos: En primer lugar, se debe realizar un preprocesamiento que limpie la data recolectada; en segundo lugar, se procede a generar data representativa y finalmente se realiza una selección de la data para extraer conocimiento predictivo
- Mejora de la predicción: En este último paso, la data preprocesada se emplea para la aplicación de modelos de *forecasting* y evaluar su predicción.

Definición de *Analytics*

Según Cooper (2012), el término *analytics* es el proceso que abarca el planteamiento de un problema en base al desarrollo de *actionable insights*, emplea modelos estadísticos y análisis de data a través del tiempo, ya sea actual o futura

La expresión *actionable insights* se refiere a lo que son acciones prácticas que parten en base a un conocimiento profundo, y que la conclusión del proceso de analítica debe orientar a las personas a una toma de decisiones por diversas perspectivas que implique valores y factores externos al análisis, como por ejemplo: la data que no es posible obtenerse; así mismo, también considera que las conclusiones están sustentadas de forma válida con medidas estadísticas tales como el nivel de confianza o el sesgo

En continuación con lo anterior, el autor describe que existen varios tipos de analítica, donde se destaca el *business analytics*, y acorde con Lutkevich (2010), es aquella práctica que emplea *insights* que impactan en la toma de decisiones del negocio y se pueden aplicar para la automatización y mejora de los procesos de este, considerando la exploración y análisis estadístico de la data organizacional a nivel iterativo y metódico.

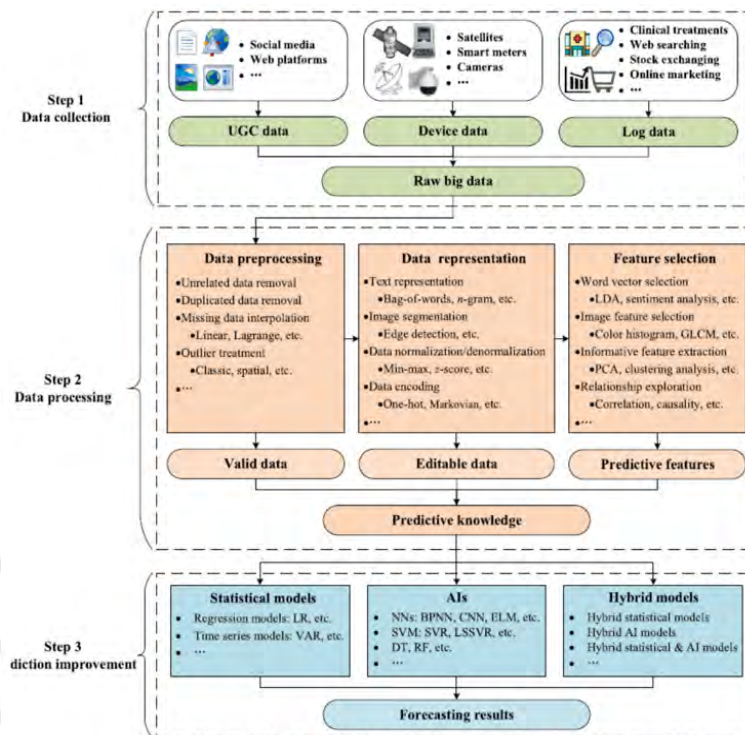


Figura 1.1 Esquema general de investigación de forecasting basada en big data

Fuente: Tang et al. (2022)

Definición de *Machine Learning*

De acuerdo con Martin Sewell (2007), el aprendizaje automático, comúnmente denominado *Machine Learning*, es una rama de la inteligencia artificial que se encarga de que los sistemas informáticos puedan aprender automáticamente, mediante la experiencia.

En línea con lo mencionado, Zhi-Hua Zhou (2021) afirma que, para estos sistemas la experiencia está representada en forma de datos; estos alimentan los algoritmos de aprendizaje para obtener un modelo que, posteriormente pueda realizar predicciones en base a nuevas observaciones o datos.

1.1.2 Tipos de aprendizaje

Se puede considerar tres tipos de aprendizaje de algoritmos de *machine learning*, de acuerdo con los datos que se dispongan y el objetivo planteado, estos son: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo. En la Figura 1.2, se muestra un resumen de la división por tipo de aprendizaje.

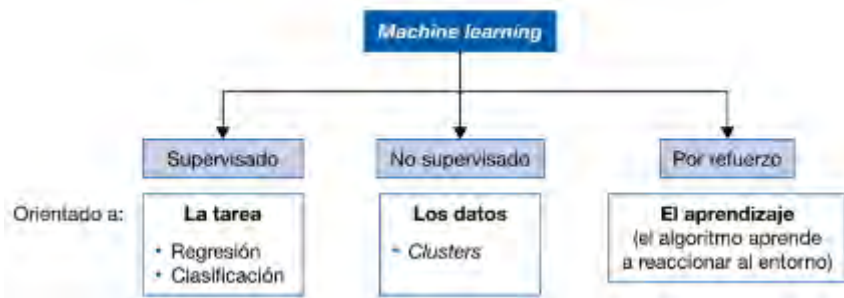


Figura 1.2 Tipos de aprendizaje

Fuente: Beunza et al., 2019

Aprendizaje supervisado

Los datos utilizados para entrenar estos modelos son conocidos, incluyendo los resultados finales. A partir de este entrenamiento, el modelo ajusta sus parámetros interiores para adaptarse a los datos y, finalmente, pueda realizar predicciones coherentes para nuevos datos, cuyos resultados de salida sean desconocidos (Beunza et al., 2019).

Como modelos de aprendizaje supervisado tenemos a los árboles de decisión, *K-Nearest Neighbor*, *Support Vector Regression* (SVR), entre otros.

Los modelos comprendidos en el aprendizaje supervisado se clasifican en dos tipos:

Modelos de clasificación: Aquellos que, a partir de datos conocidos, predicen como resultados a una etiqueta discreta o clase, la cual se encuentra dentro de un conjunto finito de etiquetas (Ministerio de asuntos económicos y transformación digital, 2020).

Modelos de regresión: Aquellos que, a partir de datos conocidos, predicen como resultados variables continuas (Beunza et al., 2019).

Aprendizaje No supervisado

A diferencia del aprendizaje supervisado, en el aprendizaje no supervisado no se predice una clase de salida determinada ni se tiene un *output* predeterminado, sino que, las aplicaciones de este tipo de aprendizaje se enfocan en el agrupamiento o *clustering* de datos (Paredes, 2020).

Beunza et al. (2019) afirman que “la labor del algoritmo es la de encontrar patrones de agrupación (*clusters*) para sujetos que por sus características en todas las variables introducidas en el modelo son parecidos” (p. 37).

Por ejemplo, estos algoritmos pueden ser utilizados para la segmentación de clientes en grupos homogéneos. Como modelos de aprendizaje no supervisado tenemos al agrupamiento por k-medias, agrupamiento jerárquico, entre otros.

Aprendizaje por refuerzo

Este tipo de aprendizaje consiste en recompensar los comportamientos que se buscan y penalizar los no deseados, de forma que, el algoritmo aprende mediante la prueba y error, con un objetivo fijado a largo plazo (Ministerio de asuntos económicos y transformación digital, 2020). Las aplicaciones del aprendizaje por refuerzo van desde los juegos hasta la robótica y sistemas de control.

1.1.3 Modelos de Machine Learning

En esta sección se describen las diferentes herramientas de Machine Learning conformadas por: regresión lineal, series de tiempo, K vecinos más cercanos, árboles de decisión, *random forest* y redes neuronales.

Regresión Lineal

Según Mohammed y Wagner(2020), se considera un conjunto de atributos o variables X_1, X_2, \dots, X_d llamado el predictor o variables independientes, y dado un atributo de interés o variable Y denominada como respuesta o variable dependiente, el objetivo de la regresión es el de predecir la variable respuesta basada en las variables independientes. A continuación, se presenta la formulación para una función de regresión lineal (1):

$$f(\mathbf{X}) = \beta + \omega_1 X_1 + \omega_2 X_2 + \dots + \omega_d X_d = \beta + \sum_{i=1}^d \omega_i X_i = \beta + \omega^T \mathbf{X} \quad (1)$$

Donde, el parámetro β es el sesgo, el parámetro ω es el coeficiente de regresión o peso del atributo X_i , y $\omega = (\omega_1, \omega_2, \dots, \omega_d)^T$ es el vector de peso para una dimensión d . Así mismo, los coeficientes de regresión y el sesgo son desconocidos, por lo que se deben estimar desde la data de entrenamiento D comprendiendo n puntos $x_i \in \mathbb{R}^d$ en un espacio dimensional d , y los valores correspondientes de respuesta $y_i \in \mathbb{R}$, para $i = 1, 2, \dots, n$. Se denota b como el valor estimado de β y w_i es denotado como el valor estimado para el coeficiente de regresión ω_i , para $i = 1, 2, \dots, d$. Además, se denota $\mathbf{w} = (\omega_1, \omega_2, \dots, \omega_d)^T$ como el vector de los pesos estimados. Una vez que se tienen el sesgo y los valores de los pesos estimados, se predice la variable respuesta para cualquier punto $\mathbf{x} = (x_1, x_2, \dots, x_d)^T$, como sigue (2):

$$\hat{y} = b + w_1 x_1 + \dots + w_d x_d = \mathbf{b} + \mathbf{w}^T \mathbf{x} \quad (2)$$

Debido al error aleatorio, generalmente el valor predicho \hat{y} no coincide con el valor respuesta observado. En ese sentido, la diferencia entre el valor observado y el predicho se llama error residual y es dado como (3):

$$\epsilon = y - \hat{y} = y - b - \mathbf{w}^T \mathbf{x} \quad (3)$$

Para predecir el sesgo y los coeficientes de regresión, se aplica con frecuencia el método de los mínimos cuadrados. Es decir, dada la data de entrenamiento D con puntos \mathbf{x} y valores de respuesta y_i (para $i=1,$

..., n), se buscan valores b y w, con el fin de minimizar la suma de errores cuadrados representado en (4):

$$SSE = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b - w^T x_i)^2 \quad (4)$$

El valor de SSE indica el nivel de variación en Y que no puede ser explicado por el modelo lineal. Se puede comparar este valor con la dispersión total, también llamado suma de cuadrados totales, para la variable dependiente Y, definido como (5):

$$TSS = \sum_{i=1}^n (y_i - \mu_Y)^2 \quad (5)$$

Así mismo, la fracción de la variación que no se explica por el modelo es dado por el ratio $\frac{SSE}{TSS}$. Por otro lado, la fracción de la variación que es explicada por el modelo se denomina coeficiente de determinación o simplemente estadístico R^2 , dado como (6):

$$R^2 = \frac{TSS - SSE}{TSS} = 1 - \frac{SSE}{TSS} = \frac{RSS}{TSS} \quad (6)$$

Mientras el valor del estadístico R^2 sea más alto, la estimación del modelo será mejor, con $R^2 \in [0,1]$

Series de Tiempo

De acuerdo con Prabhakaran (2019), una serie de tiempo es una secuencia de observaciones en intervalos de tiempo regulares y , dependiendo de su frecuencia, puede ser horaria, diaria, semanal, mensual, trimestral y anual. A continuación, se ilustra un ejemplo en la Figura 1.3.

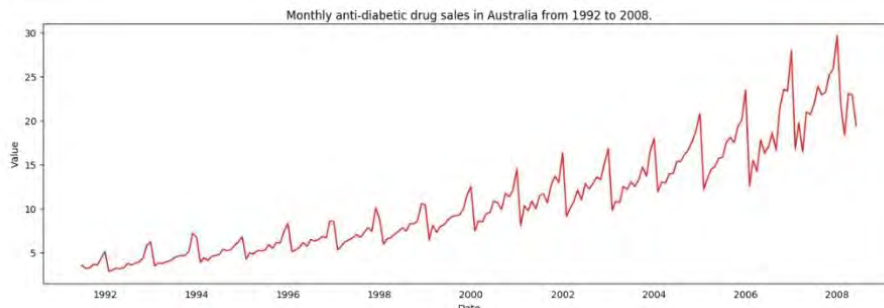


Figura 1.3 Series de tiempo mensual

Fuente: Prabhakaran, 2019

Las series temporales pueden presentar ciertas características como tendencia, estacionalidad o ambos tal como lo representado en la Figura 1.3. En primer lugar, la tendencia se puede observar a través de una pendiente creciente o decreciente a lo largo de la serie. En segundo lugar, la estacionalidad se aprecia cuando existe un patrón repetido distinto entre periodos regulares. Además, una serie puede ser cíclica en caso se observe que las fluctuaciones no ocurran con una frecuencia fija.

a) Autoregresión

Según Prabhakaran (2021), un modelo Auto Regresivo (AR) es aquel donde el valor a predecir depende de sus propios rezagos, es decir, sus valores pasados. En la siguiente función (7) se representa el modelo AR:

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_1 \quad (7)$$

Donde:

α : Constante

β_1 : Coeficiente del rezago 1 de la serie

Y_{t-1} : Rezago 1 de la serie

ϵ_1 : Error de la serie Y_t

Así mismo, el término p del componente Y_{t-p} , se refiere a la cantidad de rezagos de la serie.

La ecuación autoregresiva de Y_t no es más que la regresión lineal de Y con sus propios retrasos como predictores

b) ARIMA

De acuerdo con Prabhakaran (2021), el modelo ARIMA es una abreviatura de *Auto Regressive Integrated Moving Average* que explica una serie de tiempo determinada en función de sus propios valores pasados, es decir, sus propios retrasos y los errores de pronóstico retrasados, por lo que se puede usar para pronosticar valores futuros.

Este modelo se caracteriza por los parámetros p, d y q ; donde p representa el orden de la componente de Auto Regresión (AR) explicada en líneas anteriores, d es el número de diferenciaciones requeridas para que la serie sea de naturaleza estacionaria (los valores de la serie no dependen del tiempo) y, por último, q representa el orden de la componente de Media Móvil (MA).

Con respecto a la componente MA, es un modelo donde el valor futuro depende de los errores predecidos de los rezagos representado en (8):

$$Y_t = \alpha + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} \quad (8)$$

Donde:

α : Constante

ϕ_1 : Coeficiente del error 1 de la serie

ϵ_{t-1} : Error del rezago 1 de la serie

ϵ_t : Error de la serie Y_t

Por lo tanto, la función para el modelo ARIMA es la siguiente (9):

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p} + \epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q} \quad (9)$$

Donde:

α : Constante

$\beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_p Y_{t-p}$: Componente AR

$\epsilon_t + \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + \dots + \phi_q \epsilon_{t-q}$: Componente MA

Según Gómez (2020), para construir el modelo ARIMA se sigue el siguiente esquema basado en la metodología Box-Jenkins

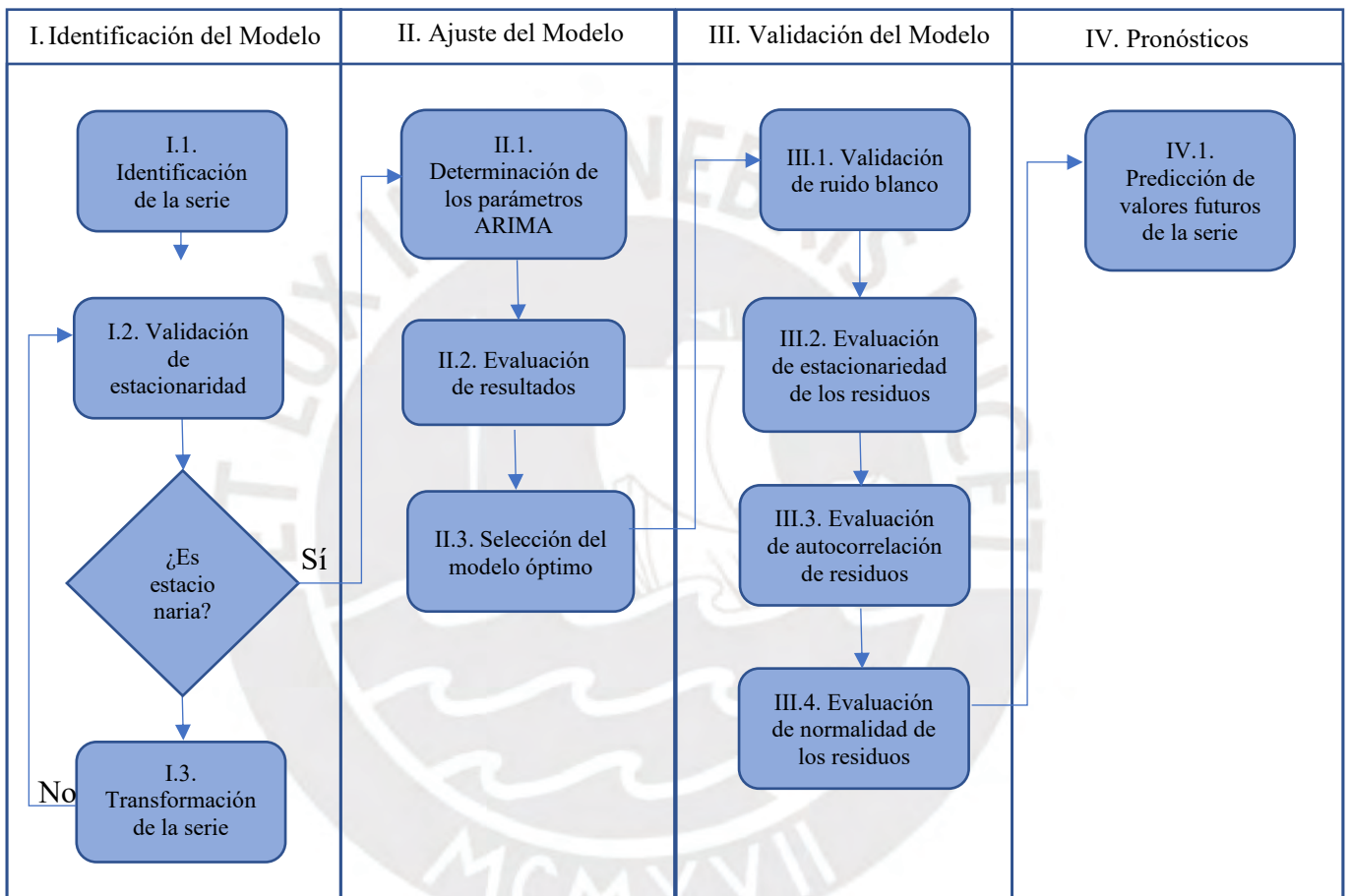


Figura 1.4 Esquema de la Metodología Box-Jenkins para modelo ARIMA

Fuente: Gómez, 2020

Como se aprecia en la Figura 1.4, el desarrollo del modelo ARIMA está conformado por cuatro etapas. A continuación, se describe en qué consiste cada una:

I. Identificación del Modelo

En esta primera fase, se identifica el comportamiento de la serie y se valida si es a través de la prueba de Dickey-Fuller. En caso resulte lo contrario, se debe aplicar una transformación.

I.1. Identificación de la serie

En base al supuesto que el modelo debe estar compuesto por una serie de tipo estacionaria, el cual se refiere a que tiene propiedades que no dependen del tiempo, el primer paso es graficar la data e identificar su naturaleza.

I.2. Validación de estacionaridad

Para validar que una serie sea estacionaria (10), se aplica la prueba de Dickey Fuller Aumentada (ADF), la cual tiene como premisa en la hipótesis nula que la serie posea una raíz unitaria, es decir, que la serie sea no estacionaria, mientras que la hipótesis alternativa refiere que la serie sea estacionaria.

$$H_0 : \alpha = 1$$

$$H_1 : \alpha < 1$$

$$Y_t = c + \beta_t + \alpha Y_{t-1} + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \dots + \phi_p \Delta Y_{t-p} + e_t \quad (10)$$

Donde:

Y_{t-1} : Rezago 1 de la serie

ΔY_{t-1} : Primera diferencia de la serie en el tiempo (t-1)

I.3. Transformación de la serie

En caso se rechace la hipótesis nula de la prueba mencionada anteriormente, el valor del parámetro d sería 0; de lo contrario, para transformar la serie a una de tipo estacionaria, se puede aplicar lo siguiente:

- Diferenciación de la serie (una o más)
- Tomar el logaritmo natural de la serie
- Sacar la raíz enésima de la serie
- Combinación de los métodos anteriores

El método más común de los mencionados es el de la diferenciación de la serie al menos una vez hasta que se convierta aproximadamente en estacionaria y consiste en restar los valores de la serie de periodos consecutivos.

II. Ajuste del Modelo

En esta segunda etapa, se procede a hallar cada uno de los parámetros del modelo para determinar un modelo ARIMA inicial. Después, se evalúan los resultados de este primer modelo y luego en base al indicador AIC, se seleccionará el óptimo.

II.1. Determinación de los parámetros ARIMA

En primer lugar, el parámetro d se define en base al paso anterior debido a que es la cantidad de diferencias de la serie para lograr que sea estacionaria. se procede a determinar p y q . En ese sentido, para identificar la componente AR del modelo (11) se debe determinar el valor p y evaluarlo a través de la autocorrelación parcial (PACF), la cual indica la correlación pura entre la serie y el rezago k exceptuando la correlación entre los rezagos intermedios.

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \epsilon_t \quad (11)$$

Como se muestra en la ecuación de ejemplo, si Y_t es la serie actual y Y_{t-1} es el rezago 1 de la serie, entonces su autocorrelación parcial es α_1 .

Cuando se genere el correlograma de PACF de la serie estacionaria (ver Figura 1.5), inicialmente se debe seleccionar los rezagos con los valores más significativos del modelo, es decir, los que se encuentren fuera del intervalo de confianza.

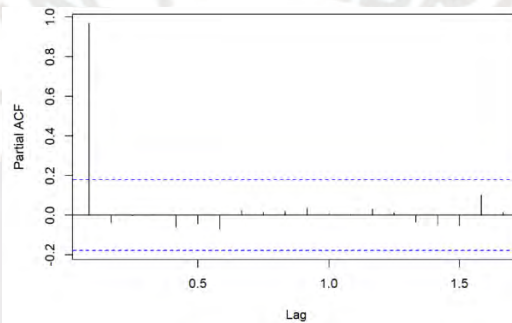


Figura 1.5 PACF de una serie de tiempo

Fuente: Gómez, 2020

Así mismo, para desarrollar la componente MA, se debe hallar el parámetro q a través del gráfico de autocorrelación (ACF), el cual indica cuántos términos MA se requieren para eliminar cualquier autocorrelación en la serie estacionaria. Al igual que en el correlograma de PACF, se identifican los valores de los rezagos más significativos para determinar el valor de q (Ver Figura 1.6).

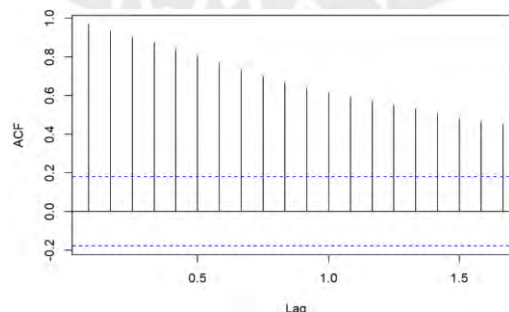


Figura 1.6 ACF de una serie de tiempo

Fuente: Gómez, 2020

II.2. Evaluación de resultados

Al terminar de hallar los tres parámetros, se desarrolla el modelo y se procede a revisar los siguientes resultados:

- Los valores de los coeficientes de cada componente AR y MA y sus respectivos p value.
- Identificar el performance del indicador AIC (Akaike Information Criterion)

En caso se identifique alguna observación en una de las componentes AR o MA, se debe reconstruir esa parte del modelo y comparar los valores de los resultados con respecto al modelo anterior.

Según Profillidis y Botzoris (2019), el AIC es una medida estadística empleada para realizar una comparación entre modelos de series temporales y su valor indica una estimación de la información perdida del proceso construido a partir de una data determinada. En ese sentido, el AIC permite identificar la bondad de ajuste del modelo que se analiza.

II.3. Seleccionar modelo óptimo

Al terminar la evaluación de resultados del modelo inicial, se procede a generar todos los posibles modelos y el criterio para seleccionar el óptimo será considerar el menor valor del indicador AIC.

III. Validación del Modelo

En esta etapa, se analiza el comportamiento de los residuos del modelo seleccionado en la fase anterior.

III.1. Validar ruido blanco

Se realiza la prueba de Portmanteu para validar si hay ruido blanco, es decir, se refiere a que el error debe cumplir las siguientes premisas:

- La media debe ser 0
- La varianza debe ser constante
- La media y la varianza no deben estar correlacionadas

La premisa en la hipótesis nula refiere que exista ruido blanco, mientras que la hipótesis indica lo contrario; por lo que se debe evaluar si el *p-value* para validar si se rechaza o no la hipótesis nula.

De validar la existencia de ruido blanco en los errores residuales de la serie, se procede a buscar el modelo ARIMA óptimo realizando una comparación gráfica entre los valores predichos y la data real y evaluando el performance del AIC. En ese sentido, se debe elegir el menor valor de AIC para encontrar el mejor modelo ARIMA.

III.2. Evaluar estacionariedad de los residuos

En este criterio se debe validar que el comportamiento de los valores sea de tipo regular en torno a una media de 0. A continuación, se ilustra un ejemplo en la Figura 1.7:

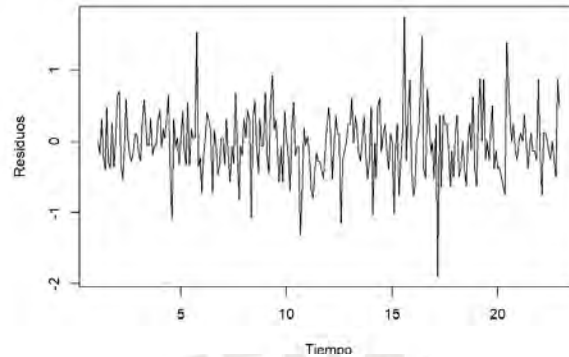


Figura 1.7 Comportamiento evolutivo de los residuos del modelo ARIMA

Fuente: Gómez, 2020

III.3. Evaluar autocorrelación de residuos

En este criterio se debe validar que la autocorrelación de los residuos en el tiempo no sea significativa. A continuación, se ilustra un ejemplo en la Figura 1.8:

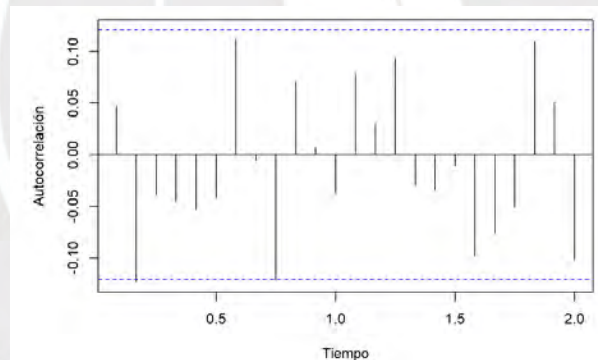


Figura 1.8 ACF de los residuos del modelo ARIMA

Fuente: Gómez, 2020

III.4. Evaluar normalidad de los residuos

En este criterio se debe validar que la distribución de los residuos se encuentre concentrada en la línea de normalidad o cercana a ella. A continuación, se ilustra un ejemplo en la Figura 1.9:

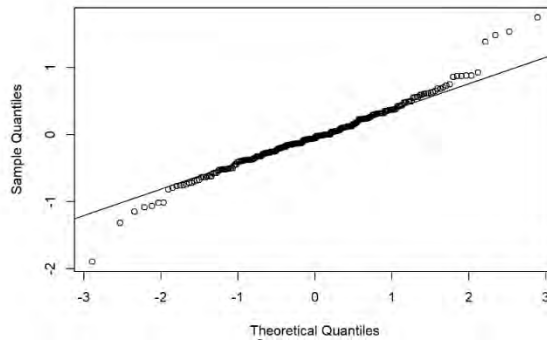


Figura 1.9 Normalidad de los residuos del modelo ARIMA

Fuente: Gómez, 2020

IV. Pronósticos

En esta última etapa se procede directamente con la predicción de los valores futuros de la serie en estudio.

IV.1. Predecir valores futuros de la serie

Una vez que se ha validado el modelo ARIMA final, este ya se encuentra apto para predecir los periodos futuros a determinar de la serie

Una variante del modelo ARIMA, en caso la serie sea de naturaleza estacional, se denomina SARIMA (*Seasonal Auto Regressive Integrated Moving Average*) siendo representado por los parámetros $(p,d,q) \times (P,D,Q)$; es decir, está conformado por dos componentes: (P,D,Q) de carácter estacional mientras que los factores no estacionales son (p,d,q) .

K Vecinos más cercanos

Como método de aprendizaje supervisado, K Vecinos más cercanos (kNN) es uno de los algoritmos de clasificación más simples, en el que, a partir de una muestra, se localiza las “k” muestras de entrenamiento más cercanas y se utilizan para realizar predicciones. Asimismo, es posible ponderar la muestra basado en la distancia a la que se encuentra, asignándole un peso más alto a las muestras más cercanas (Zhou, 2021).

A continuación, en la Figura 1.10, se muestra como la predicción varía dependiendo del valor elegido para “k”. Por ejemplo, en el caso $k=1$, el primer vecino es un signo “+”, por lo que la muestra de prueba tomaría también el valor de “+”. Estas predicciones también pueden variar de acuerdo con los cálculos de distancia que se utilicen.

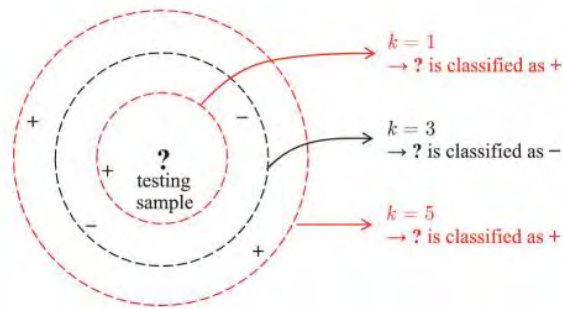


Figura 1.10 K Vecinos más cercanos

Fuente: Zhou, 2021

Árboles de Decisión

Según Mohammed y Wagner (2020), un árbol de decisión se compone de nodos internos que representan las decisiones correspondientes para los hiperplanos o puntos de división, y nodos de hoja que representan regiones o particiones de cada espacio de datos, el cual está etiquetado con la clase mayoritaria. Así mismo, una región se caracteriza por el subconjunto de puntos de la data que yacen en cada una. En la Figura 1.11, se ilustra un ejemplo del modelo:

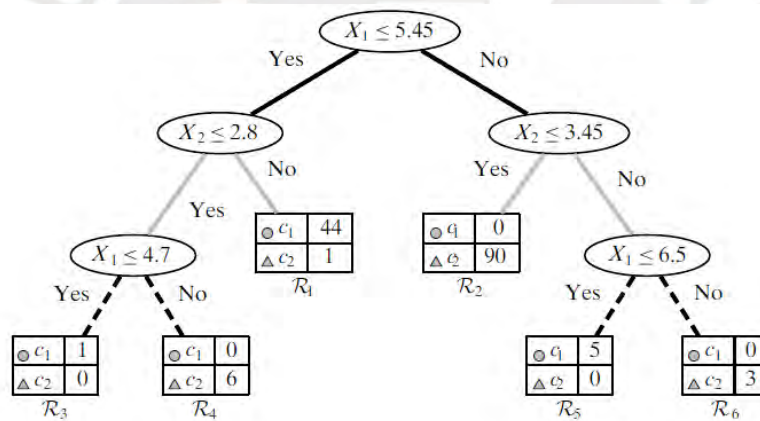


Figura 1.11 Árboles de decisión

Fuente: Mohammed y Wagner, 2020

Así mismo, una medida de evaluación para los puntos de división es la entropía, la cual en general mide el grado de desorden para un sistema. En caso del modelo de árbol de decisión, si la mayoría de los puntos dentro de una partición pertenecen a una misma clase, entonces existe una baja entropía. Por otro lado, la entropía será mayor si existen clases mixtas dentro de una misma partición.

Random Forest

De acuerdo con Breiman(2001), los bosques aleatorios son una combinación de predicción de árboles donde cada uno depende de los valores de un vector aleatorio muestreado de forma independiente y con la misma distribución para todos los árboles del bosque. Cada árbol se desarrolla en base a una muestra usando *bootstrapping*; la cual, según Hastie, Tibshirani y Friedman (2008), es una herramienta que se usa para estimar la distribución de muestreo a través de réplicas de múltiples muestras generadas a partir

de una muestra de entrenamiento. Luego se van generando un conjunto de atributos aleatorios ,y posteriormente, se debe seleccionar el mejor de estos para la división. Finalmente, el modelo se basa en el voto mayoritario de cada árbol individual del bosque.

El error de generalización para bosques converge hasta un límite a medida que el número de árboles en el bosque va aumentando; así mismo, depende de la fuerza de los árboles individuales en el bosque y la correlación entre ellos.

Además, este modelo es una herramienta eficaz en la predicción pues no ocurre el sobreajuste, y de considerar el tipo correcto de aleatoriedad, brinda precisión en los resultados de los valores predichos.

Redes Neuronales

Haykin (2009) sostiene que las redes neuronales artificiales están inspiradas en la forma de procesamiento del cerebro humano, el cual tiene la capacidad de organizar sus componentes estructurales (neuronas) para realizar cálculos, como reconocimiento de patrones, percepción, entre otros.

Del mismo modo, Haykin (2009) define a la red neuronal como un procesador distribuido en paralelo de forma masiva, compuesto por unidades de procesamiento simples que tiene una propensión natural a almacenar conocimiento proveniente de la experiencia y ponerlo a disposición para su uso. Al igual que un cerebro humano, la red neuronal artificial adquiere conocimiento de su entorno a través de un proceso de aprendizaje y, el conocimiento adquirido se almacena en las fuerzas de interconexión entre neuronas, denominado pesos sinápticos. Las unidades de procesamiento simples, denominadas neuronas, se conectan en una estructura de capas para formar la red neuronal.

En una red neuronal, se establecen tres tipos de capas: capa de entrada, con campos de entrada para la información externa; capas ocultas, cuya cantidad depende de la arquitectura de la red; y capa de salida, donde se representan los campos de destino, es decir, los resultados finales (IBM Corporation, 2021).

Haykin (2009) señala que se pueden identificar tres tipos de arquitecturas de red:

i. Redes de alimentación directa de una sola capa

En este tipo de redes, solo se cuenta con una capa de entrada de nodos de origen que se conecta en una dirección a la capa de salida de neuronas, en las que se realizan los cálculos para obtener los resultados finales. En la Figura 1.12, se muestra un ejemplo de red con cuatro nodos en las capas de entrada y salida.

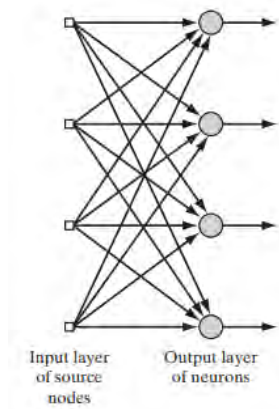


Figura 1.12 Redes de alimentación directa de una sola capa

Fuente: Haykin, 2009

ii. Redes de alimentación directa multicapa

Cuenta con la presencia de una o más capas ocultas en la red. La función de las neuronas ocultas, nodos de las capas ocultas, es de intervenir de forma útil a información que se traslada de la capa de entrada a la de salida. Se muestra en la Figura 1.13 un ejemplo de este tipo de red, con diez nodos de entrada, cuatro ocultos y dos de salida.

iii. Redes recurrentes

A diferencia de los tipos de redes antes descritos, estas redes poseen al menos un circuito de retroalimentación, en el cual la salida de una neurona retroalimenta a su propia entrada, esto se muestra en la Figura 1.14. Asimismo, estas redes pueden o no poseer capas ocultas.

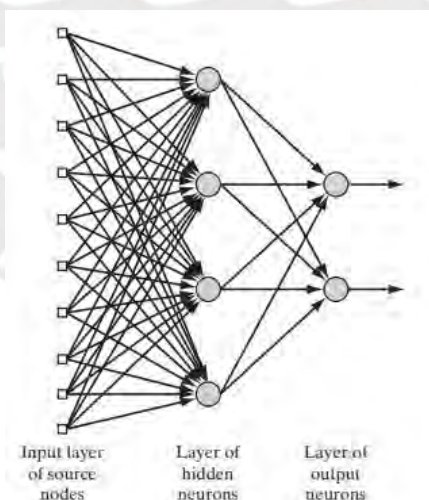


Figura 1.13 Redes de alimentación directa multicapa

Fuente: Haykin, 2009

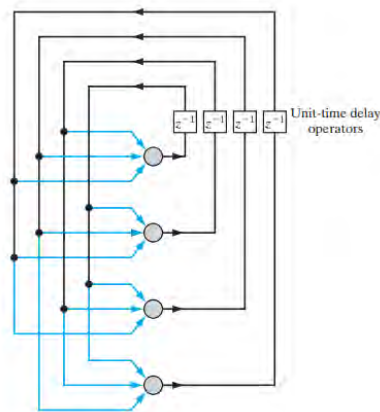


Figura 1.14 Redes recurrentes

Fuente: Haykin, 2009

En conclusión, cada modelo antes mencionado tiene fortalezas y debilidades dependiendo de los tipos de datos y el problema que se quiere resolver. En los capítulos posteriores, se probará algunos de estos modelos, ajustando sus parámetros hasta encontrar el óptimo para la problemática del presente estudio.

1.1.4 Workflow: Ciclo de la analítica

Según la Universidad Estatal de Michigan (2019), para la solución de un problema de negocio, se debe tomar en cuenta las 6 etapas del ciclo de *business analytics* que se ilustran en la Figura 1.15:

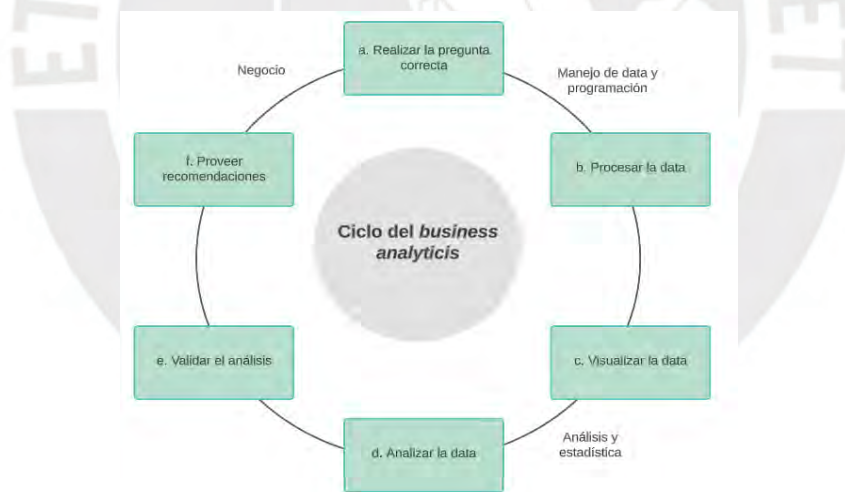


Figura 1.15 Ciclo del business analytics

Fuente: Universidad Estatal de Michigan, 2019

A continuación, se describen los seis pasos del ciclo de la analítica:

a. Realizar la pregunta correcta

El primer paso está orientado a formular la pregunta correcta para identificar el problema que se quiere resolver, la meta que se quiere alcanzar o qué tipo de *insight* se quiere desarrollar y consiste en construir un modelo para recolectar la data relevante.

b. Procesar la data en respuesta a la pregunta

Inicialmente la data se encuentra concentrada en una gran base de datos del negocio, la cual requiere ser procesada para extraer solo la información relevante, y para ello, se requiere considerar lo siguiente:

- Recolectar la data usando programas (SQL, Power BI, entre otros) que transformen y preparen la data para elaborar un modelo de analítica .
- Limpiar la data de las inconsistencias o valores atípicos que pueden influir en la precisión de esta y su funcionalidad.

c. Visualizar la data

Una vez concluido el paso anterior, se debe representar la data a través de la estadística descriptiva como histogramas, gráficos de dispersión, entre otros dependiendo de la naturaleza de datos a trabajar.

d. Analizar la data

En esta etapa del ciclo del *business analytics*, se pueden emplear herramientas tales como el análisis de correlación, regresiones o técnicas predictivas que nos brinden una solución en base al histórico de la data.

e. Validar el análisis

Una vez aplicadas las herramientas en el paso anterior, se deben desarrollar escenarios para comprobar efectivamente que la data sea precisa y apropiada aplicando análisis estadístico y validando los parámetros objetivo con el fin de lograr una solución óptima y cercana a la meta del negocio.

f. Aplicación al negocio

Finalmente, con el modelo ya validado, se proponen las recomendaciones basadas en los *insights* descubiertos.

1.2 Modelos de Optimización

En este acápite se desarrollan los conceptos generales acerca de la investigación operativa, así como la formulación de un modelo de programación lineal y se detalla el modelo de programación de la capacidad del recurso humano.

1.2.1 Conceptos Generales

La investigación operativa busca, mediante la aplicación del método científico, la resolución óptima de problemas donde están involucrados sistemas de hombres y/o máquinas (Ramos et al., 2010, p. 5). Para

este fin, se plantean modelos matemáticos que traten de representar la realidad de la problemática y proporcionen resultados que aporten a la toma de decisiones (Pérez, 2019). Entre las principales aplicaciones, se encuentran la resolución de problemas de planificación de materiales, programación de rutas y asignación de recursos y personal, presentes en una amplia variedad de campos como en los servicios de cuidado de la salud, transporte terrestre, transporte aéreo, líneas de producción, servicios financieros, entre otros.

En un problema de optimización se encontrarán los siguientes elementos: función objetivo, variables de decisión, restricciones y parámetros. De esta forma, las decisiones a tomar corresponden a los valores que tomen las variables, con el fin que se maximice o minimice la función objetivo, en el marco de las restricciones y parámetros establecidos por el problema (Hillier y Lieberman, 2010).

Los modelos de optimización pueden clasificarse en dos tipos: determinísticos y probabilísticos, o también llamados estocásticos. En los modelos determinísticos, se busca valores puntuales para las variables que satisfacen las restricciones en un conjunto de ecuaciones y, optimizan la función objetivo (Pérez, 2019). Según Renard et al. (2013), los modelos determinísticos, a diferencia de los estocásticos, están limitados por no considerar la incertidumbre, apartándose así de la representación real de la problemática y dificultando la confiabilidad total en sus resultados. Los parámetros de un problema determinístico son datos constantes, mientras que, en un problema estocástico, estos son desconocidos y se representan mediante distribuciones estadísticas (Renard et al., 2013).

De acuerdo con Pérez (2019), la investigación operativa determinística se apoya en una serie de métodos de la programación lineal para encontrar la solución óptima, dependiendo del tipo de problema que se busca resolver.

1.2.2 Formulación de un modelo de programación lineal

En línea con lo descrito por Pérez (2019), para definir la estructura matemática de un modelo, se requiere definir los siguientes elementos:

- Variables de decisión (x_i): Aquellas que representan la solución del planteamiento matemático de un problema. Dependiendo de este, se considera la cantidad de variables necesarias (x_1, x_2, \dots, x_n).
- Función Objetivo (F.O.): Aquella función formulada mediante la interrelación de las variables de decisión del problema. En el planteamiento del problema, se desea optimizar el resultado de esta función, ya sea al maximizar o minimizar el objetivo.
- Restricciones: En estas se representa las limitaciones del contexto en el que se encuentra el problema, como la disponibilidad, límites de uso de los recursos y preferencias impuestas por el negocio o los recursos. Estas pueden ser de igualdad o desigualdad y mostrar conexiones entre las variables.

- Rango de existencia: Conjunto de valores que pueden tomar las variables sin violar las restricciones y limitaciones del problema. En este, puede considerarse la condición de no negatividad para algunos problemas, donde todas las variables del modelo deben ser mayores o iguales a cero.

Asimismo, tanto la función objetivo como las restricciones deben ser lineales, lo cual implica que las mismas variables de la expresión no se multipliquen entre sí; es decir, no contenga términos cuadráticos. De esta forma, el modelo puede presentarse, en líneas generales, de la siguiente forma (12):

$$\text{F.O.} \quad \text{Max o Min } f(x) = \sum_{j=1}^n c_j x_j \quad (12)$$

Sujeto a las siguientes restricciones:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n \{ \leq, \geq, = \} b_1 \quad (13)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n \{ \leq, \geq, = \} b_2 \quad (14)$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{mn}x_n \{ \leq, \geq, = \} b_m \quad (15)$$

Donde $x_j \geq 0$, para $i = 1 \dots m; j = 1 \dots n$

1.2.3 Modelo de programación de la capacidad del recurso humano

De acuerdo con Ernst, Jiang, Krishnamoorthy y Sier (2004), *Employee Scheduling, Staff Scheduling* o programación de la capacidad del recurso humano es el proceso de estructurar horarios y asignar el personal idóneo para estos, de forma que se pueda satisfacer la demanda de tareas del negocio en lapsos de tiempo determinados; tomando en cuenta las condiciones laborales, regulaciones y preferencias de los empleados.

El problema de la programación de la capacidad del recurso humano en cualquier organización abarca los siguientes retos: determinar los requerimientos de la demanda, diseñar los bloques básicos de trabajo más adecuados (turnos, funciones, asignaciones, etc.), ordenar esos bloques en líneas de trabajo u horarios y asignar el personal a los horarios (da Silva Rocha, 2013).

Asimismo, Ernst, Jiang, Krishnamoorthy y Sier (2004) describen estos tres componentes característicos de la siguiente forma:

- Variables de decisión (x_i): En este caso son los empleados o personal que debe ser asignado a los turnos, de forma que puedan satisfacer la demanda. La asignación de los empleados a los turnos puede ir de acuerdo con restricciones, políticas o reglas de la empresa, como el tipo de contrato, el cargo que ocupa o sus preferencias.

- Restricciones: Para este caso, las restricciones se estructuran a partir de la demanda y los turnos u horarios a considerar. La demanda se refiere al trabajo requerido por la empresa, este puede mostrarse como una lista de tareas específicas o como requerimientos de personal para determinados lapsos de tiempo, dependiendo del contexto de la problemática. Los turnos u horarios se refieren a los periodos de tiempo en el que un empleado se encuentra laborando. La estructura y duración de estos horarios depende de las condiciones que se planteen en el problema. Además, los horarios pueden considerar refrigerios o descansos, de acuerdo a políticas o requerimientos de la empresa, de forma que se pueda representar un modelo más cercano a la realidad.

Una restricción puede ser expresada como $x_1 + x_2 + x_3 + \dots + x_n \{\leq \geq\} b_j$; donde x_i : Cantidad de empleados en el turno i , b_j : Demanda de empleados en el rango horario j .

- Función Objetivo (F.O.): En este tipo de problemas se busca minimizar la cantidad de personal requerido para satisfacer la demanda (16):

$$\text{Min } f(x) = \sum_{j=1}^n c_j x_j \quad (16)$$

Con el fin de ilustrar la metodología antes descrita, en el Anexo 1 se muestra un problema de planificación de personal de policía descrito por Hasan y Arefin (2017), en el que se busca minimizar el número total de oficiales necesarios para cumplir con todos los requisitos de turnos.

1.3 Simulación

En este acápite se describirán algunas nociones estadísticas previas junto con conceptos acerca de qué es una simulación y el proceso de análisis de los datos que conlleva desde la clasificación, el muestreo y las pruebas de bondad de ajuste

1.3.1 Conceptos previos

Para todo desarrollo de un estudio sobre simulación, es de importancia definir los siguientes conceptos estadísticos:

- Población y muestra

En primer lugar, según Arias (2006), la población es aquel conjunto de elementos que presentan características en común y serán acotadas de acuerdo con el enfoque del estudio a desarrollar. Así mismo, las poblaciones con una cantidad muy grande de elementos se clasifican como poblaciones infinitas; mientras que aquellas con cantidad de elementos limitada conforman parte de una población finita.

En segundo lugar, de acuerdo con el autor Tamayo, T. Y Tamayo, M (1997), la muestra es el grupo de individuos que se extrae de la población con el fin de estudiar un fenómeno estadístico pues a

partir de esta, es posible identificar la naturaleza de una problemática y desviaciones de un determinado proceso.

- Media muestral, Varianza muestral y Proporción esperada

En continuación con lo antes mencionado respecto a la muestra, según Córdova (2006) se describen sus siguientes características:

Media muestral

La media, que se denota como \bar{x} , es aquel valor promedio de todos los elementos que conforman un conjunto determinado, que en este caso aplicaría para los valores de la muestra.

Varianza muestral

Se describe varianza como la medida de dispersión de los valores de una variable aleatoria X con respecto a su media, y es denotada como $V(x)$

La proporción esperada

Se refiere a la cantidad de ocurrencias de un valor dentro de una muestra

- Pruebas de hipótesis

Según Levin y Rubin (2004), se formula/plantea una suposición acerca de un parámetro de la población a estudiar con el fin de probar su validez, esta se define como la hipótesis nula (denotada con H_0). Por ejemplo, se desea probar que la media de la productividad de una empresa de servicios es igual a 32; en ese sentido, la hipótesis nula es la siguiente:

$$H_0: \mu = 32$$

Se lee como “la hipótesis nula es que la media de la productividad es igual a 32”

En continuación con lo anterior, si los resultados de la muestra seleccionada no avalan lo mencionado en la hipótesis planteada, entonces se debe rechazar la suposición, concluyendo así que no se rechaza la hipótesis alternativa (denotada con H_1), la cual plantea una suposición diferente a la nula; en este caso se presentan tres hipótesis alternativas posibles:

$H_1 : \mu \neq 32$ ← La hipótesis alternativa que la media de la productividad no es igual a 200

$H_1 : \mu > 32$ ← La hipótesis alternativa que la media de la productividad es mayor a 200

$H_1 : \mu < 32$ ← La hipótesis alternativa que la media de la productividad es menor a 200

Una vez establecidas las hipótesis nula y alternativa, se procede a seleccionar un nivel de significancia, el cual se refiere a la proporción de medias muestrales que exceden límites determinados e indicará el criterio de aceptación o rechazo de la hipótesis nula. En ese sentido, si se fija un nivel de significancia del 5%, se puede observar en la Figura 1.16 que dentro de la región del

95% del área bajo la curva no existe una diferencia significativa entre el estadístico en evaluación y el valor supuesto en la hipótesis nula por lo que esta sería la región de aceptación, es decir, no hay evidencia estadística suficiente para rechazar la hipótesis nula.

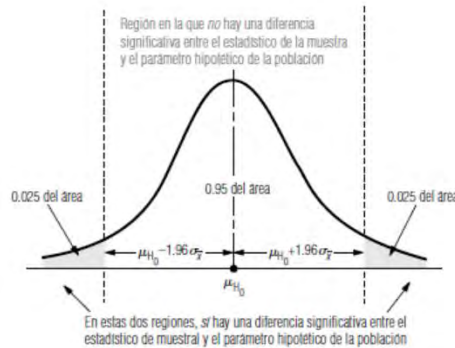


Figura 1.16 Nivel de significancia del 5%

Fuente: Levin y Rubin, 2004

Posteriormente, el siguiente paso es la definición del tipo de distribución de probabilidad a usar, con lo cual se tienen dos alternativas a seleccionar: la distribución normal y la distribución t. En la Tabla 1.1, se describen las condiciones a tomar en cuenta para la selección del tipo de distribución:

Tabla 1.1 Condiciones para seleccionar el tipo de distribución de probabilidad

	Cuando se conoce la desviación estándar de la población	Cuando no se conoce la desviación estándar de la población
El tamaño de muestra n es mayor que 30	Distribución normal, tabla z	Distribución normal, tabla z
El tamaño de muestra n es 30 o menos y suponemos que la población es normal o aproximadamente normal	Distribución normal, tabla z	Distribución t, tabla t

Fuente: Levin y Rubin, 2004

Así mismo, se tienen dos tipos de pruebas de hipótesis: de dos colas y una cola. En cuanto a la primera, se emplea en caso de que la hipótesis nula contenga un supuesto de igualdad sobre un valor determinado; sin embargo, la segunda es adecuada cuando el supuesto considere un valor mayor o menor al estadístico planteado.

- Intervalos de confianza

Acorde a Levin y Rubin (2004), el intervalo de confianza es el rango que se estima para el parámetro de una población determinada, siendo los niveles de estimación más usados los de 90,95 y 99 %. Por ejemplo, si se tiene un nivel de confianza del 90% que la media de la población de pasajeros por cada vuelo está entre 170 a 180, se interpreta que el rango 170-180 es el intervalo de confianza a considerarse. Así mismo, la expresión para los límites del intervalo de confianza se muestra a continuación:

$$\bar{x} + 1.64\sigma\bar{x} = \text{l\u00edmite superior del intervalo de confianza (LSC)}$$

$$\bar{x} - 1.64\sigma\bar{x} = \text{l\u00edmite inferior del intervalo de confianza (LIC)}$$

As\u00ed mismo, los autores comentan que, si en la pr\u00e1ctica se considera altos niveles de confianza, estos generan que los intervalos sean m\u00e1s extensos y, en consecuencia, las estimaciones pierden precisi\u00f3n

1.3.2 Definici\u00f3n

Seg\u00fan Shannon (1975), simulaci\u00f3n es aquel proceso de desarrollo de un modelo dise\u00f1ado en base a un sistema real con el fin de analizar el funcionamiento del sistema o sopesar estrategia para el sistema que se est\u00e1 analizando

1.3.3 Elementos de un modelo de simulaci\u00f3n

Seg\u00fan Leonelli (2021), hay dos tipos de objetos usados para un modelo de simulaci\u00f3n:

- Entidades: Son aquellos elementos individuales del sistema que se est\u00e1n simulando, las cuales poseen un comportamiento que se est\u00e1 observando de forma expl\u00edcita
- Recursos: Al igual que las entidades, son elementos individuales del sistema, pero no se modelan de forma individual y son considerados como elementos contables. As\u00ed mismo, no existe un rastreo del comportamiento de estos elementos

Con respecto a la organizaci\u00f3n de las entidades y recursos dentro de un sistema de simulaci\u00f3n, se describen los siguientes conceptos:

- Atributos: Se refieren a las propiedades que poseen los objetos del sistema y se usan para controlar el comportamiento del objeto
- Estado: Representa una colecci\u00f3n de variables necesarias para describir el sistema en cualquier instante del tiempo.
- Lista: Es un conjunto de entidades o recursos ordenados de alguna forma l\u00f3gica.

As\u00ed mismo, durante un estudio de simulaci\u00f3n las entidades y los recursos van a interactuar y pueden cambiar de estado. A continuaci\u00f3n, se detallan los siguientes conceptos referentes a ello:

- Evento: Se refiere al instante de tiempo cuando cambia el estado del sistema.
- Actividad: Es un per\u00edodo de tiempo de duraci\u00f3n espec\u00edfica y se conoce cuando empieza
- Demora: Es aquel periodo de tiempo de duraci\u00f3n no espec\u00edfica y no se conoce hasta que finaliza
- Reloj: Es aquella variable que representa el tiempo simulado

1.3.4 Modelos de simulaci\u00f3n

Seg\u00fan Taha (2004), se tienen dos modelos de simulaci\u00f3n descritos a continuaci\u00f3n:

- Continuo: Es aquel que estudia el comportamiento de un sistema cuando cambia de forma continua en el tiempo

- Discreto: En este modelo, el comportamiento de un sistema cambia según eventos que ocurren en hitos específicos del tiempo

1.3.5 Análisis de datos

En este apartado se procede a desarrollar la clasificación de los datos de entrada, definición del tamaño de muestra y las pruebas de bondad de ajuste

- Clasificación de datos

Los datos pueden ser de dos tipos: determinísticos, proporciones o variables aleatorias

Datos determinísticos

De acuerdo con Taha (2004), se refiere a aquellos datos concretos y conocidos con certeza, por ejemplo: un avión que arriba al aeropuerto con una frecuencia diaria.

Proporciones o probabilidades

Según Hamdy Taha (2004), se refiere a aquellos datos que ocurren de forma aleatoria, por ejemplo, la llegada de los pasajeros para la salida de un vuelo comercial.

Variables aleatorias

Según Levin y Rubin (2004), se considera una variable aleatoria a aquella que toma distintos valores generados por un experimento aleatorio y se puede clasificar en discreta o continua. En caso de las variables aleatorias de tipo discreta, solo toman una cantidad fija de valores; mientras que las de tipo continua pueden tomar valores ilimitados dentro de un rango determinado

- Muestreo

Según Córdova (2003), se presentan las estimaciones del tamaño de muestra de la media y la proporción:

Estimación de la media

“Se puede determinar que tan grande debe ser el tamaño de la muestra, n , de manera que si μ se estima por \bar{x} ”, el error de estimación no sea mayor que un valor dado e . En efecto, el valor de n se obtiene de” (p. 397) ; siendo el valor de n representado en (17)

$$Z_{1-\frac{\alpha}{2}} \sigma_{\bar{x}} \leq e$$

“Entonces, si \bar{x} estima a μ , entonces, se tiene una confianza del $(1 - \alpha) \times 100\%$ de que el error no será mayor que el valor dado e cuando el tamaño de la muestra sea” (p. 397):

$$n = \frac{(Z_{1-\frac{\alpha}{2}} \sigma)^2}{e^2} \quad (17)$$

“Si la población es finita de tamaño N y el muestreo es sin sustitución, el error estándar es $\sigma_{\bar{x}} = \left(\frac{\sigma}{\sqrt{n}}\right)\sqrt{(N-n)/(N-1)}$ y el valor de n se calcula por” (p. 397), donde n se representa en (18):

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2 N}{Z_{1-\frac{\alpha}{2}}^2 \sigma^2 + e^2(N-1)} \quad (18)$$

Estimación de la proporción

“Dado el error máximo e de la estimación de p con confianza de $(1 - \alpha) \times 100\%$ el tamaño n de la muestra” (p.420), donde e se puede determinar de la siguiente forma (19):

$$e = Z_{1-\frac{\alpha}{2}} \sqrt{\bar{p}(1-\bar{p})/n} \quad (19)$$

y n resulta como en (20),

$$n = \frac{(Z_{1-\frac{\alpha}{2}})^2 \bar{p}(1-\bar{p})}{e^2} \quad (20)$$

- Pruebas de bondad de ajuste

De acuerdo con Levin y Rubin (1998), se refiere a aquellas pruebas donde se verifica si existe una diferencia significativa entre la distribución teórica y la distribución de frecuencias observadas de los datos. Así mismo, existen dos tipos de prueba: Chi Cuadrado y Kolmogorov Smirnov.

1. Chi Cuadrado

Según Levin y Rubin (1998), esta prueba es de utilidad para evaluar si una distribución de probabilidad es la que mejor representa los datos que se están analizando, es decir, se puede determinar qué tan significativa es la distribución teórica con la observada y así definir la bondad de ajuste entre ambas. Además, para su aplicación el tamaño de muestra debe ser mayor o igual a 90. Para ello se deben plantear las pruebas de hipótesis siguientes:

Ho: La distribución de una variable aleatoria se ajusta a los datos observados

H1: La distribución de una variable aleatoria no se ajusta a los datos observados

Así mismo, para calcular el estadístico de prueba, se aplica la siguiente fórmula (21):

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e} \quad (21)$$

f_o : Una frecuencia observada

f_e : Una frecuencia esperada

2. Kolmogorov Smirnov

Según Levin y Rubin (1998), esta prueba de una cola para variables aleatorias de tipo continua es considerada más sencilla de aplicar a comparación de la chi-cuadrado debido a que no es necesario que los datos sean agrupados. En esta prueba se usa el estadístico D_n para comprobar la proximidad entre la distribución de frecuencias observada y la esperada pues la distribución de probabilidad del estadístico D_n depende del tamaño de muestra n ; sin embargo, no depende de la distribución de frecuencias esperada. Para realizar el cálculo del estadístico, se aplica la siguiente fórmula (22):

$$D_n = \max |F_e - F_o| \quad (22)$$

Donde:

F_e : Función de distribución acumulada esperada

F_o : Función de distribución acumulada observada

1.4 Estado del arte

En esta sección se describen casos representativos de los temas de *Machine Learning*, Modelos de Optimización y Simulación respectivamente, de los cuales se tomarán como referencia para el desarrollo de este estudio

1.4.1 Caso 1: Long-Term And Short-Term Forecasting Techniques For Regional Airport Planning

El artículo elaborado por Danesi et al. (2017) resalta el crecimiento sustancial del número anual de pasajeros que viajan en el transporte aéreo en los últimos años, especialmente desde los aeropuertos regionales, tanto en el número de vuelos como en el factor de ocupación medio de cada vuelo. Este crecimiento se debe en parte a factores macroeconómicos, como el aumento de la globalización y el cambio en el comportamiento de los viajes que surge por los cambios demográficos durante los auges económicos, en particular en las economías de Asia y Europa del Este. Señalan que, a nivel de la gestión de un aeropuerto regional, el rápido crecimiento del número de pasajeros ejerce presión sobre una planificación eficaz de la capacidad del aeropuerto. Las mejoras de capacidad en la infraestructura aeroportuaria representan grandes inversiones de capital y, por lo tanto, las previsiones a largo plazo de los volúmenes de pasajeros y los volúmenes en las horas pico son de gran importancia.

En este artículo, se exploran técnicas para pronosticar la demanda de pasajeros en un aeropuerto regional a corto y largo plazo. Primero, aplican un modelo de series temporales ARIMA (Modelo autorregresivo integrado de media móvil) estacional para una previsión de la demanda de pasajeros a largo plazo a nivel trimestral. Asimismo, para la demanda de pasajeros por día de la semana establecen un modelo predictivo no paramétrico con la técnica de regresión local (LOESS) y la técnica de valor promedio simple. Las metodologías mencionadas se aplican en datos del Aeropuerto Guglielmo Marconi de Bolonia, un aeropuerto en la región de Emilia Romagna de Italia que manejó 6,9 millones de pasajeros en 2015.

El modelo de serie de tiempo $ARIMA(p;d;q) \times (P;D;Q)_4$, utilizado para pronosticar la demanda trimestral de pasajeros de 2016 a 2026, incluye demoras estacionales en su metodología y, además, los valores óptimos para p , P , q y Q fueron determinados mediante el criterio de Información de Akaike (AIC) (Ver Figura 1.17). Por otro lado, tanto el modelo que utiliza el método LOESS como el de promedio simple buscan encontrar la demanda relativa de pasajeros que toman vuelos en el aeropuerto de estudio por hora en cada día de la semana. Los datos considerados para los modelos incluyen la cantidad de pasajeros de todos vuelos de 2013 a 2015 en intervalos de quince minutos.

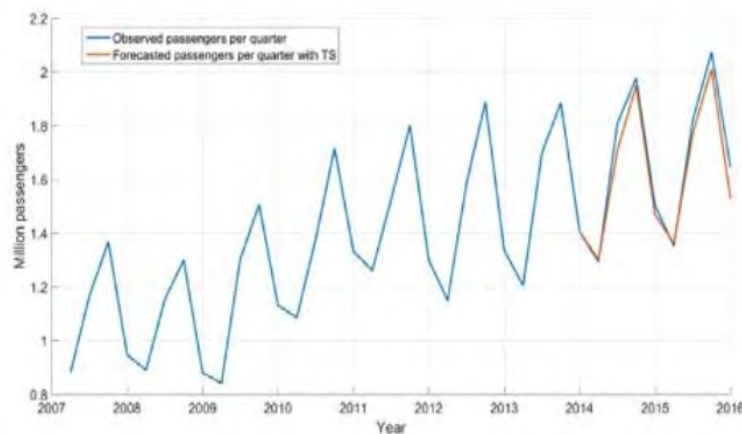


Figura 1.17 Rendimiento del modelo de series de tiempo

Fuente: Danesi et al., 2017

Por otro lado, tanto el modelo que utiliza el método LOESS como el de promedio simple buscan encontrar la demanda relativa de pasajeros que toman vuelos en el aeropuerto de estudio por hora en cada día de la semana. Los datos considerados para los modelos incluyen la cantidad de pasajeros de todos vuelos de 2013 a 2015 en intervalos de quince minutos. En el Anexo 2, se muestran los modelos antes mencionados por día de semana, comparados a los datos reales de demanda relativa de pasajeros por hora.

Se pudo concluir que, los modelos de series de tiempo establecidos en el estudio pronostican un crecimiento significativo del volumen de pasajeros en el aeropuerto de estudio. Se prevé que el número

de pasajeros anuales supere los 11 millones en 2026, con una tasa media anual de crecimiento de 4,77% (Ver Figura 1.18).

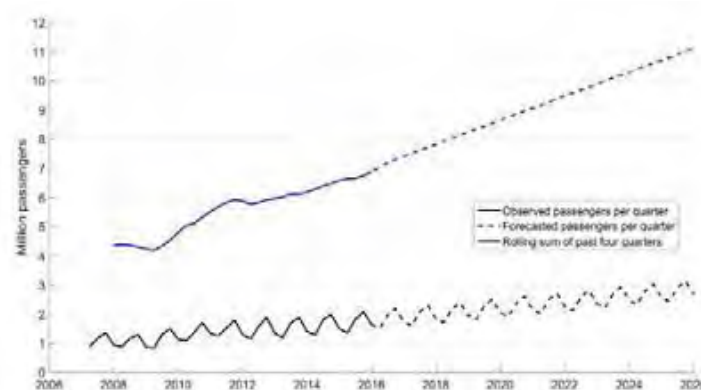


Figura 1.18 Predicción de la demanda de pasajeros por trimestre

Fuente: Danesi et al., 2017

Con respecto a las curvas de tiempo por día de semana, se concluyó que el método LOESS no demuestra ninguna ventaja en precisión sobre el método de promedio simple (Ver Tabla 1.2).

Tabla 1.2 Validación cruzada para modelos de hora de la semana

Method	Cross validation
Método LOESS	2.97212×10^{-5}
Método promedio simple	2.96291×10^{-5}

Fuente: Danesi et al., 2017

1.4.2 Caso 2: A workforce planning and allocation model for the outbound baggage loading area at Santiago International Airport

Dada la creciente demanda de vuelos observada en los últimos años en el aeropuerto internacional Arturo Merino Benitez de Chile, Cavada et al. (2020) detectaron que el sistema de equipajes estaba alcanzando una operación cercana al límite de capacidad física de la infraestructura del aeropuerto, lo que vuelve más complejo el proceso y atención de manejo de equipajes para el proveedor de servicios terrestres Andes Airport Services. Debido a que los itinerarios de las aerolíneas están diseñados en línea a criterios comerciales estándar para la satisfacción de los clientes, estos toman poca consideración a las condiciones operativas del aeropuerto. En consecuencia, la actividad en el aeropuerto fluctúa ampliamente, con periodos de alta demanda cuando el sistema está sobrecargado y periodos de baja demanda cuando el sistema está infrautilizado. Estas variaciones en los niveles de demanda generan desafíos complejos para la planificación del personal de manejo de equipajes.

De esta forma, los autores plantean una metodología que incluye la generación de tareas y el modelo de asignación de turnos para asegurar la dotación diaria del personal respectivo al área de equipajes del aeropuerto. El modelo propuesto está planteado en dos etapas que incluyen dos modelos de programación lineal entera que se resuelven secuencialmente. El primer modelo de generación de tareas define la cantidad de personal necesario para atender la demanda que surge de la llegada de maletas al carrusel de carga en función de los horarios de los vuelos. Un ejemplo de la demanda mencionada se muestra en la Figura 1.19, en un rango de 4 horas donde se observa que fluctúa ampliamente. El segundo modelo de asignación de turnos utiliza los resultados generados por el primer modelo, agrupados en intervalos de media hora, para determinar el número óptimo de personal necesario por turno por día. Este modelo está limitado por una serie de restricciones como la duración del turno, los períodos de descanso elegibles, la duración del descanso, el tiempo de trabajo mínimo asignable, el número máximo asignable de tareas diferentes y las combinaciones de tareas permitidas.

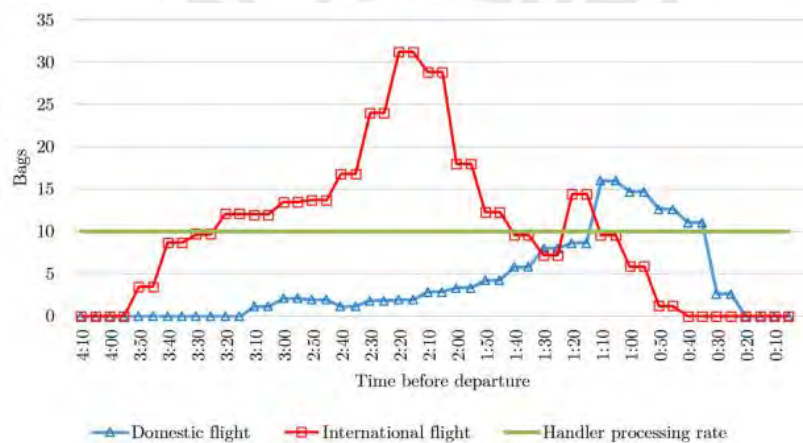


Figura 1.19 Ejemplos de distribución de llegadas de maletas durante el período de preparación del vuelo anterior a la salida de un vuelo nacional e internacional

Fuente: Cavada, 2020

En el primer caso, el modelo sigue a grandes rasgos la estructura definida a continuación, en la que se busca minimizar la suma del número requerido de personal en cada carrusel y la congestión en los carruseles como se expresa en la ecuación (23), con el fin de que no se supere el umbral de congestión y a la vez, el personal sea lo más eficiente permitido. El detalle de las restricciones se encuentra en el Anexo 3.

Modelo

$$\min \beta_1 \sum_{t \in T} \sum_{c \in C} x_{tc} + \beta_2 \sum_{t \in T} \sum_{c \in C} y_{tc} \quad (23)$$

Variables

x_{tc} : Número requerido de manipuladores en el período t en cada carrusel c .

f_{tv} : Número de maletas del vuelo v procesadas en el carrusel asociado al mismo en el periodo t .

s_{tv} : Número de maletas del vuelo v que quedan por procesar al final del período t (inventario).

y_{tc} : Número de bolsas por encima del umbral de congestión al final del período t en cada carrusel c .

En el segundo caso, el modelo planteado garantiza que se satisfaga la demanda de personal por turno. La función objetivo busca minimizar el número de turnos utilizados, otorgando cierta preferencia o penalización a algunos turnos de acuerdo con políticas de la empresa o preferencias del empleador.

Modelo:

F.O.
$$\min \sum_{i \in I} c_i z_i \tag{24}$$

Variables:

z_i : número de turnos tipo i utilizados.

Parámetros:

R_{jp} : Necesidad de personal para el puesto j en el bloque p .

α_{ijp} : Igual a 1 si el turno i cubre el trabajo j durante el bloque p , de lo contrario, 0.

c_i : Penalización por usar el tipo de turno i .

Restricciones:

$$\sum_i \alpha_{ijp} z_i \geq R_{jp} \quad \forall j, \forall p \tag{25}$$

$$z_i \in \mathbb{Z}^+ \quad \forall i \tag{26}$$

Se probaron los modelos antes mencionados de forma secuencial basado en datos recopilados en el aeropuerto de diciembre a febrero. Al realizar las comparaciones entre los resultados finales de la metodología propuesta y la planificación real de personal utilizada por el proveedor en el Aeropuerto Internacional de Santiago en Chile se obtuvo que, la solución del modelo propuesto reduce la necesidad de personal y turnos asignados de manera óptima (Ver Tabla 1.3).

Tabla 1.3 Resumen de resultados

Period	Avg. daily number of flights	Avg. shifts assigned		Difference		
		Model	Benchmark	Min.	Max.	Avg.
December	204	55	69	10.3%	35.2%	20.9%
January	287	93	98	1.0%	11.5%	5.5%
February	297	95	103	2.9%	16.0%	7.5%

Fuente: Cavada, 2020

Asimismo, la utilización del modelo representaría una reducción de personal de manejo de equipajes de 47 a 42 en promedio diario y, a la vez, reduce la congestión de maletas en los carruseles al solo alcanzar dos veces al día niveles críticos de congestión (50 maletas) en contraste con la planificación actual, donde se alcanza alrededor de doce veces al día (Ver Figura 1.20). En el Anexo 4 se muestra como los modelos presentan una mejor performance de cantidad de maletas cargadas en el tiempo contra la realidad. Finalmente, como resultado se obtiene una operación de manejo de equipaje más ordenada y adecuada a la demanda.

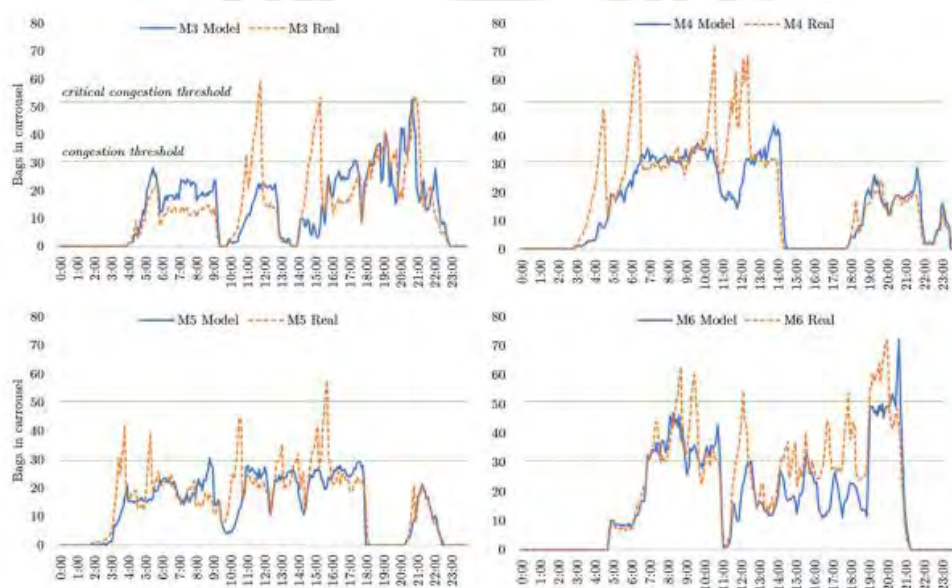


Figura 1.20 Congestión de bolsas en carruseles: modelo versus real. Carruseles M3-M6 durante un día representativo

Fuente: Cavada, 2020

1.4.3 Caso 3: Propuesta de mejora de procesos en la atención de aeronaves en tránsito de la flota Airbus de LAN Perú, para incrementar la productividad en el área de mantenimiento

De acuerdo con el estudio de esta tesis, el autor Luis Andrés Gamarra Sánchez (2018) se centra en la optimización del proceso de atención de un vuelo en tránsito para una aeronave modelo Airbus por parte de un técnico de la aerolínea LAN Perú y enfoca la problemática en el área de Mantenimiento, pues se presenta un promedio de 30% de cancelaciones de las órdenes de trabajo programadas por motivos tales como la falta de horas hombre de técnicos, falta de repuestos, entre otros; sobrepasando así el ratio límite permitido de 15% a nivel compañía.

En ese sentido, el autor determina que las causas principales son la falta de horas hombre de mecánicos y especialistas en estructuras; en consecuencia, la aerolínea debía incurrir en costos adicionales como el pago de horas extras al personal y el costo por hora de trabajo no realizado en el Hangar. A continuación, el autor procede a analizar el proceso actual a cargo del técnico e identifica actividades que generan tiempos muertos, por lo que emplea la Ingeniería de Métodos para reducir estas tareas y proponer un proceso mejorado. Posteriormente, aplica la simulación como herramienta de validación de su propuesta y, a través del software Arena, genera las corridas del escenario actual y mejorado, donde limita la ventana horaria en un rango de atención desde las 07:00 a 19:00 h considerando una muestra de 1000 arribos para ambos casos.

Previamente al desarrollo de las corridas en Arena, el autor usa la herramienta Input Analyzer para identificar las distribuciones de probabilidad de las variables evaluadas (llegada de aeronaves y aeronave en tránsito). Una vez realizado ello, el autor comienza ingresando los datos de las variables y las distribuciones resultantes del paso previo para la corrida del primer escenario en Arena y configura los tiempos de cada tarea del técnico para el proceso actual (Ver Anexo 5).

Resource				
Usage				
Instantaneous Utilization				
	Average	Half Width	Minimum Value	Maximum Value
TECNICO	0.4092	(Insufficient)	0.00	0.7692
Number Busy				
	Average	Half Width	Minimum Value	Maximum Value
TECNICO	5.3194	(Insufficient)	0.00	10.0000
Number Scheduled				
	Average	Half Width	Minimum Value	Maximum Value
TECNICO	13.0000	(Insufficient)	13.0000	13.0000

Figura 1.21 Resultados de la corrida del proceso actual en el software Arena

Fuente: Gamarra, 2018

Al finalizar la corrida, los resultados obtenidos (Ver Figura 1.21) indicaron que la cantidad de aeronaves atendidas fueron un promedio de 58 y un requerimiento promedio de 13 técnicos dentro del rango horario establecido.

Después de terminar con el primer escenario, procede con el desarrollo del segundo escenario y en este caso solo toma en consideración la variable de llegada de aeronaves puesto que la de aeronave en tránsito ya no es relevante debido a la reestructuración del proceso mejorado, el cual posee menor cantidad de tareas del técnico y redistribuye las otras tareas al personal operativo de la aerolínea y el proveedor (Ver Anexo 5). Al término de la corrida, los resultados indicaron que se atendieron 54 aeronaves y se emplearon un promedio de 2 técnicos en el rango de 07:00 a 19:00 h (Ver Figura 1.22).

Resource				
Usage				
Instantaneous Utilization				
	Average	Half Width	Minimum Value	Maximum Value
OPERADOR	0.8038	(Insufficient)	0.00	1.0000
TECNICO	0.5649	(Insufficient)	0.00	1.0000
Number Busy				
	Average	Half Width	Minimum Value	Maximum Value
OPERADOR	4.0188	(Insufficient)	0.00	5.0000
TECNICO	1.1297	(Insufficient)	0.00	2.0000
Number Scheduled				
	Average	Half Width	Minimum Value	Maximum Value
OPERADOR	5.0000	(Insufficient)	5.0000	5.0000
TECNICO	2.0000	(Insufficient)	2.0000	2.0000

Figura 1.22 Resultados de la corrida del proceso propuesto en el software Arena

Fuente: Gamarra, 2018

En ese sentido, las bondades de utilizar la simulación en el software Arena son las siguientes:

- Permite modelar el proceso analizado en base a la cantidad de aviones que demanda el negocio para la ventana horaria de tiempo correspondiente
- Contrastar el proceso actual en comparación con el propuesto para validar que genere ahorro en recurso de mano de obra y en base a eso, cuantificar su impacto económico
- Al validar la propuesta que el autor hace para la situación actual, se soluciona el problema principal sobre la cancelación de trabajos de mantenimiento.
- Se genera un ahorro en pago de horas extras a técnicos valorizado en S/. 217,435.25

En conclusión, en base al desarrollo de las corridas de la simulación realizada, el autor valida efectivamente que la propuesta planteada del proceso mejorado es del uso de 2 técnicos de Mantenimiento para la atención de una aeronave Airbus en la franja horaria de 07:00 a 19:00 h y a comparación del proceso actual, reduce un total de 11 técnicos.

CAPITULO 2. LA EMPRESA

En este capítulo, el desarrollo consiste en la descripción de la empresa en estudio, cómo se encuentra organizada por áreas y jerarquía, limitando así el alcance a nivel de unidad de negocio y profundizando el flujo del proceso de planificación de la compañía. Posteriormente, se explican los indicadores principales que maneja la empresa y se describe la problemática actual en base al performance de los resultados históricos 2022. En base a ello, se emplea la herramienta de Pareto con el fin de identificar los problemas más relevantes en cuanto a impacto económico y, a partir de ello, analizar las causas raíces para luego seleccionar las más de mayor prioridad a solucionar.

Abarca el desarrollo conceptual de tres temas pilares de este estudio, los cuales son *Machine Learning*, Modelos de Optimización y Simulación; así como el estado del arte, donde se describen casos representativos de cada tema.

2.1 Descripción de la empresa

En esta sección se estará desarrollando inicialmente una descripción general de la empresa, seguida de la explicación de su estructura según el organigrama y, por último, el detalle del funcionamiento del proceso de planificación.

2.1.1 Descripción general

El presente estudio hace referencia a una empresa del sector aeroportuario encargada de prestar servicios en tierra con cobertura en países de Latinoamérica, siendo líder en el rubro por más de 30 años. La empresa está conformada en las siguientes unidades de negocio:

- Carga: Servicios de traslado de carga aérea
- Rampa: Servicios de atención de vuelos en tierra
- Pasajeros: Servicios de atención al pasajero
- OMA: Servicios de mantenimiento de aeronaves

Cabe indicar que el enfoque de esta investigación es el negocio de Rampa del Aeropuerto Internacional Jorge Chávez de Lima.

Misión:

Garantizar la eficiencia, seguridad y calidad en la propuesta de atención a los servicios aeroportuarios para los clientes y dentro de todo nivel de la organización.

Visión:

Ser líder en el negocio de servicios aeroportuarios a nivel Latinoamérica y consolidar altos estándares de seguridad, calidad y operaciones.

2.1.2 Organigrama

La empresa de estudio tiene distribuidas sus áreas en cuatro gerencias diferenciadas: la gerencia de la unidad de negocios rampa y pasajeros, la gerencia de la unidad de negocios carga, la gerencia de la unidad de negocios OMA y la gerencia de soporte. En el Anexo 6, se muestra el organigrama general de la empresa.

Asimismo, dentro de la unidad de negocio de rampa y pasajeros se cuenta con tres áreas principales: *planning*, control de operaciones (CCO) y *rostering* (Ver Anexo 7). El presente trabajo se enfocará en los procesos del área de *planning* y como estos se relacionan a las demás áreas de soporte.

2.1.3 Proceso de Planificación

En el área de Planificación se elaboran dos métodos para la estimación de la demanda y oferta enfocado en el recurso mano de obra, donde el primero se determina a nivel macro y el segundo a través de una curva gráfica para mayor precisión. De acuerdo con el método que se aplique, se interactúa con distintas áreas para que finalmente, el planeamiento se refleje en el área de Operaciones. De esta manera, las etapas del proceso se describen en el Anexo 8.

I. Proyección de la oferta

En esta primera fase del proceso de planificación, se calcula a nivel macro la demanda de vuelos mensual de un periodo determinado y la oferta de personal que se requiere para ese nivel de demanda, la cual consta de las siguientes tres etapas:

a. Consolidación de itinerarios

En primer lugar, el área de Planificación realiza mensualmente la cuantificación de la demanda de vuelos que se atenderán a tres meses del tiempo presente, la cual se obtiene a partir de los itinerarios preliminares enviados por las aerolíneas de forma trimestral. Para ello, se procede a consolidar la información que se recibe para homogenizarla en un solo formato.

b. Cálculo de la demanda

En segundo lugar, para la estimación de la demanda (28), se realiza la sumatoria de vuelos por cada aerolínea de la información previamente consolidada (27) y posteriormente, se aplica un ratio mensual de cumplimiento histórico de vuelos reales (29) con el fin de no estimar el escenario más optimista.

$$VT = \sum_{i=1}^n Vi \quad (27)$$

Donde:

- VT: Cantidad de vuelos totales itinerario
- Vi: Cantidad de vuelos itinerario por aerolínea

- n: Cantidad de aerolíneas

$$DV = \sum_{i=1}^n Vi * RHn \quad (28)$$

$$RHn = VR / VTh \quad (29)$$

Donde:

- DV: Demanda de vuelos proyectada del periodo a evaluarse
- RHn: Ratio histórico de cumplimiento de vuelos itinerario vs real por aerolínea
- VTh: Cantidad de vuelos totales itinerario de periodos históricos
- VR: Cantidad de vuelos totales reales de periodos históricos

Así mismo, a continuación, se ilustra un ejemplo de cómo se realiza la proyección a futuro de personal del mes de abril-2023 (Ver Tabla 2.1):

Tabla 2.1 Ejemplo de proyección de demanda inicial abril 2023

Demanda vuelos 2023				
Cliente	Planificado			Proyección
	Enero	Febrero	Marzo	Abril
Aerolínea A	1550	1260	1705	1800
Aerolínea B	465	392	527	510
Aerolínea C	155	112	155	150
Total vuelos	2170	1764	2387	2460

Como se puede apreciar, se consolida la cantidad de vuelos brindados por las aerolíneas en el mes de abril siendo inicialmente un total de 2460 vuelos al mes. Sin embargo, como se mencionó en líneas anteriores, se realiza un ajuste puesto que históricamente se mide el cumplimiento de la operación de estos vuelos de itinerario. En ese sentido, se aplica un promedio del último semestre histórico tal como se muestra en la Tabla 2.2:

Tabla 2.2 Ejemplo de cumplimiento histórico de demanda 2022

Cumplimiento histórico 2do semestre 2022							
Cliente	Julio	Agosto	Setiembre	Octubre	Noviembre	Diciembre	Ratio promedio
Aerolínea A	97%	90%	87%	85%	90%	95%	91%
Aerolínea B	95%	97%	98%	100%	100%	100%	98%
Aerolínea C	98%	98%	98%	100%	100%	100%	99%

Posterior a la aplicación de estas ratios promedio por aerolínea, la cantidad de vuelos proyectado como demanda para abril-2023 sería de 2283 vuelos (Ver Tabla 2.3):

Tabla 2.3 Ejemplo de proyección de demanda final abril 2023

Demanda vuelos 2023				
Cliente	Planificado			Proyección
	Enero	Febrero	Marzo	Abril
Aerolínea A	1550	1260	1705	1632
Aerolínea B	465	392	527	502
Aerolínea C	155	112	155	149
Total vuelos	2170	1764	2387	2283

c. Cálculo de la oferta

Una vez obtenida la cantidad total mensual de vuelos a atender en ese periodo, se procede a calcular la oferta (31), es decir, la cantidad de personal que se requerirá para esa demanda; para lo cual, se toma en cuenta un ratio histórico mensual de vuelos-personal (30) por área operativa de atención en base a los meses que cumplieron o superaron la meta del indicador de productividad del negocio. Seguido de ello, se realizan cálculos de capacitación, ausentismo, ceses y vacaciones del personal para cuantificar el total de trabajadores en planilla (32). Al obtener este total, se evalúa la variación de las cantidades mensuales del mes a proyectar y el anterior para identificar el valor que se solicitará al área de Selección (33).

$$R_p = HCh / VR \quad (30)$$

$$HCo = DV * R_p \quad (31)$$

Donde:

- R_p : Ratio que mide el promedio de la cantidad de personal que atendió un vuelo
- HCh : Cantidad de personal que atendió la demanda de vuelos en un periodo pasado
- VR : Cantidad de vuelos totales reales de un periodo pasado
- HCo : Cantidad de personal requerido para atender la demanda de vuelos proyectados
- DV : Demanda de vuelos proyectada del periodo a evaluarse

$$HCr = HCo + HCa + HCp + HCv \quad (32)$$

$$HCs = HCr - HC_{r-1} \quad (33)$$

Donde:

- HCr : Cantidad de personal total por rol requerido para atender la demanda de vuelos proyectados en un periodo determinado
- HCo : Cantidad de personal para atender la demanda de vuelos proyectados
- HCa : Cantidad de personal adicional para cobertura de inasistencia
- HCp : Cantidad de personal adicional que releva a los trabajadores que reciben capacitación

- HCv: Cantidad de personal adicional que cobertura las vacaciones de los trabajadores de todas las áreas
- HCs: Cantidad de personal que se solicita al área de Selección
- HCr-1: Cantidad de personal total por rol requerido para atender la demanda de vuelos proyectados del periodo anterior

En continuación con el ejemplo anterior de la tabla, una vez que se obtiene la demanda de vuelos del periodo a proyectarse, se aplica un ratio promedio de HC/Vuelos como se observa en la tabla para obtener la cantidad de personal de la oferta (Ver Tabla 2.4):

Tabla 2.4 Ejemplo de ratio oferta-demanda del primer trimestre 2023

Ratio Oferta-Demanda 2023				
Mes	Enero	Febrero	Marzo	Ratio promedio (Ene-Mar)
HC/Vuelos	0.08	0.10	0.07	0.08

Al calcular el ratio de HC/Vuelos con la demanda proyectada en la tabla, se obtiene un total de HC oferta de 184 para el mes de abril-2023. A su vez, se consideran otros factores como el HC ausentismo, capacitación y vacaciones, los cuales incrementan el HC requerido a 250 personas como se observa en la Tabla 2.5:

Tabla 2.5 Ejemplo de proyección de la oferta abril 2023

Oferta HC 2023				
Mes	Enero	Febrero	Marzo	Abril (Proyección HC)
HC oferta	175	170	156	184
HC ausentismo	35	34	31	37
HC capacitación	10	10	10	10
HC vacaciones	18	18	16	19
HC requerimiento	238	232	213	250

Como se aprecia, la diferencia entre el requerimiento de abril a diferencia de marzo es de 37 personas, por lo que este sería el valor que se debe solicitar al área de Selección para su respectivo reclutamiento.

II. Reclutamiento

Posteriormente, Selección recibe el requerimiento mensual de la cantidad de personal que debe reclutar y luego empieza la captación de los posibles candidatos según el perfil correspondiente. Al término de esta etapa, el área identifica la cantidad de personas seleccionadas y envía el número al área de Planificación confirmando si se cumplió o no con el requerimiento inicial.

III. Elaboración de *capacity*

El proceso de elaboración de *capacity* tiene como objetivo definir los turnos del personal por hora de forma que, se cubra óptimamente todas las atenciones y servicios requeridos por los clientes. Este proceso consta de cuatro etapas que se detallan a continuación:

- a. Consolidación de itinerarios: Los clientes, en este caso las aerolíneas, comparten la información de sus vuelos de forma mensual. Esta información consta del detalle de cada vuelo: fecha, modelo de aeronave, matrícula de la aeronave, tipo de vuelo, origen del vuelo, destino del vuelo, hora de salida, hora de llegada. En la Tabla 2.6, se muestra un ejemplo de vuelo tránsito que conecta en Lima. Una vez que se tienen los itinerarios de todos los clientes, estos se consolidan en un solo repositorio.

Tabla 2.6 Ejemplo de itinerario de la Aerolínea A

Cliente	Fecha	Aeronave	Matrícula	Tipo de vuelo	Origen	Destino	ETA	ETD
Aerolínea A	11/12/2022	A-320	CX-BFF	Tránsito	Cuzco	Cajamarca	18:15	19:00

- b. Parametrización: Para cada tipo de vuelo se tiene definido una cantidad específica de personal necesario para atenderlo y los tiempos de inicio y fin de atención de vuelo respecto al itinerario de ese tipo de vuelo. A continuación, se muestra un ejemplo de parámetros para el proceso de Rampa en la Tabla 2.7. En este caso, se considera 1 Operador Líder, 1 Operador 1 y 3 Auxiliares, lo que da un total de 5 personas para atender un vuelo de este tipo.

Tabla 2.7 Ejemplo de parámetro para la aerolínea A

Cliente	Tipo de aeronave	Tipo de vuelo	Tipo de destino	Cantidad de personal	Hito 1	Tiempo antes	Hito 2	Tiempo después
Aerolínea A	NB	Tránsito	Nacional	5	ETA	-18	ETD	10

- c. Elaboración de la curva de demanda: En este paso, los itinerarios de vuelo por aerolínea se transforman en tiempos de inicio y fin de atención, mediante el maestro de parámetros. Por ejemplo, basado en la Tabla 2.6 y Tabla 2.7, se modela la Tabla 2.8, donde se obtiene los tiempos de atención para ese vuelo. Al acumular todos los tiempos de inicio y fin de atención de todos los vuelos de un día, se dibujará en una gráfica la curva de demanda de personal, en el eje Y se mostrará la cantidad de personas y en el eje X la hora del día en las que se necesitan. Por ejemplo, en la Figura 2.1 se presenta que, a las 6 a.m. sería necesario contar con 20 personas para atender los vuelos.

Tabla 2.8 Ejemplo de tiempos de atención

ID Itinerario	ETA	ETD	Cantidad de personal	Tiempo inicio de atención	Tiempo fin de atención
Vuelo 1	18:15	19:00	5	17:57	19:10

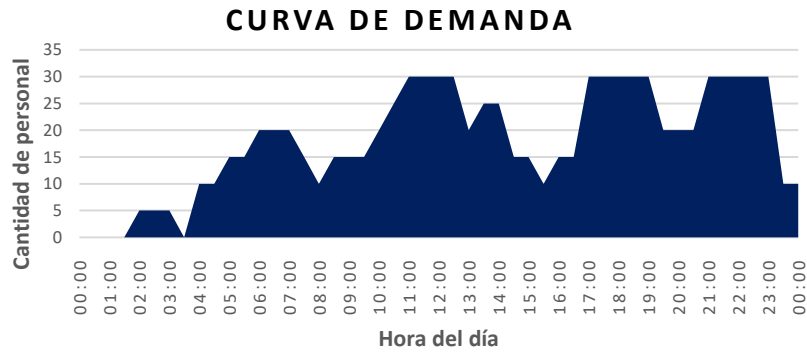


Figura 2.1 Ejemplo de curva de demanda

- d. Delimitación de la oferta: Finalmente, a partir de la demanda dibujada en la gráfica, se define la cantidad de personal por turnos que serían necesarios. En general, estos turnos son de 9 horas, incluyendo 1 hora de refrigerio (Ver Tabla 2.9). Al juntar todos los turnos del día, se obtiene la cantidad de personal disponible por hora y este resultado se dibuja como oferta en el gráfico. Como se puede observar en la Figura 2.2, debido a que se elige los turnos y cantidad de personal de forma manual, muchas veces se generan tiempos muertos para el personal, volviendo el proceso menos eficiente.

Tabla 2.9 Ejemplo de turnos diarios

Turno	Hora Inicio	Hora Fin	Cantidad de personal
Turno 1	01:00	10:00	5
Turno 2	03:00	12:00	5

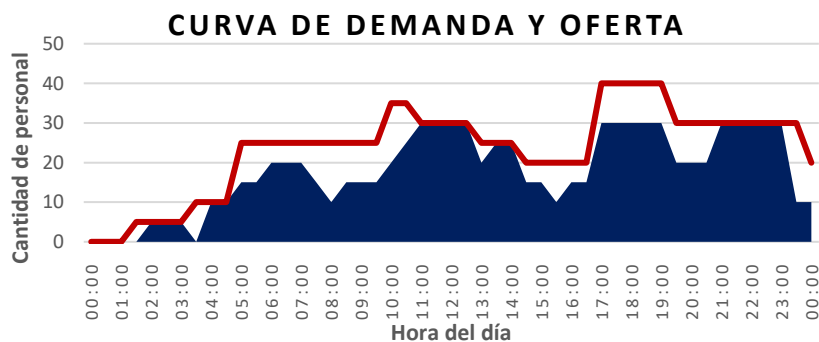


Figura 2.2 Ejemplo de curva de demanda con oferta

IV. Armado de rol

El área de Rostering se encarga de asignar las personas en base a la cantidad y horarios que definió Planificación por cada área operativa y cargo, considerando turnos de trabajos rotativos y descansos de forma grupal o individual. Posteriormente, se encarga de subir la información a una plataforma online y al término de la carga, procede con la publicación del rol para comunicar a los trabajadores en qué turnos y áreas les corresponde laborar.

V. Programación operativa

El área del Centro de Control de Operaciones (CCO) se encarga de revisar la programación de vuelos diaria y accede al rol online de personal por horarios en base a la información cargada por el área de Rostering mencionada en líneas anteriores. Posterior a ello, revisa las personas que estarán de turno ese día para asignarlas directamente a la atención de cada vuelo. En caso del proceso principal operativo, el cual es Rampa, CCO se encarga de seleccionar un grupo de cinco personas conformado por un operador líder, un operador 1 y tres auxiliares e informar al líder del grupo cuál será el número de vuelo a atender, en qué PEA se encuentra y a qué aerolínea pertenece.

VI. Atención del vuelo

Una vez que CCO ha comunicado y confirmado el grupo de Rampa que atenderá el vuelo, el líder y su operador se dirigen a buscar los equipos necesarios que se encuentran dentro de la misma plataforma del aeropuerto. Después, todo el grupo se reúne, dependiendo del tipo de atención, para iniciar con la colocación de equipos o esperar a la llegada de la aeronave.

2.2 Análisis de la situación actual y diagnóstico

En este apartado se narra la situación actual de la empresa objeto de estudio, así como la elaboración del diagnóstico a partir de la identificación de las problemáticas que presenta.

2.2.1 Situación actual

Este acápite abarca la identificación de la situación actual de la empresa con respecto al proceso de planificación previamente expuesto. En ese sentido, se describen y presentan los resultados de los principales KPI's de gestión:

KPI's de servicio al cliente:

- Demoras de la aerolínea A: Se mide en base a la cantidad de vuelos que salieron posterior a la hora planificada de su itinerario por responsabilidad de la empresa, dividida entre la cantidad de vuelos totales que operó la aerolínea. Es uno de los indicadores de mayor importancia pues esta aerolínea abarca el 70% de la industria aeroportuaria en vuelos comerciales. A continuación, en el Anexo 9 se observa el performance mensual 2022 del indicador:

Como se puede apreciar en el Anexo 9, los valores del indicador estuvieron fuera de la meta de 2.5% durante 9 meses del 2022 siendo abril el de peor desempeño con 4.97% de demoras. Es decir, la meta de puntualidad de la aerolínea A estuvo en un 97.5% de acuerdo con el contrato acordado. Sin embargo, debido a la nueva coyuntura de recuperación de la industria aeroportuaria, se actualizó la meta de demoras a 5%, lo cual equivale a un 96% de nivel de puntualidad a partir del 2023.

- SNC: Se mide en base a la cantidad de vuelos que recibieron una calificación positiva en todos los criterios de servicios y cuánto representa esta de la cantidad total de vuelos atendidos por la empresa.

Como se puede observar en el Anexo 9, los valores del indicador estuvieron fuera de la meta de 96% durante todo el 2022 siendo abril el mes con el desempeño más bajo con 80% de cumplimiento en el nivel de servicio.

KPI's financieros:

- Costos de personal contratado: Se mide en base a los costos relacionados al personal entre la venta percibida por atención a aeronaves. Estos costos abarcan los costos por planilla (sueldos básicos y beneficios por ley), nocturnidad, horas extra, contratación, movilidad, beneficios de la empresa, entre otros. Es muy importante monitorear este indicador, dado que, indirectamente puede mostrar si el negocio está siendo rentable o no, y si la distribución de costos es la adecuada para la industria. En el Anexo 9, se observa el performance mensual 2022 del indicador. Se visualiza que los valores del indicador estuvieron fuera de la meta en seis de los doce meses de ese año.
- Costos de movilidad: Se mide los costos de movilidad entre la venta percibida por atención a aeronaves. El beneficio de movilidad se le brinda al personal con horarios que incluyan hora de entrada o salida en el rango de 00:00 a 04:00. En el Anexo 10, se observa el performance mensual 2022 de este indicador. Se visualiza que los valores del indicador estuvieron fuera de la meta en la mayoría de los meses de ese año.

KPI's operativos:

- Rotación de personal: Se mide de acuerdo con la cantidad de personal que sale de la empresa con respecto a la cantidad total de personal que conforma la planilla operativa. En el Anexo 10 se observa el performance mensual 2022 del indicador.

Como se puede observar en el anexo, los valores del indicador estuvieron fuera de la meta de 6% durante todo el 2022 siendo enero el mes con el desempeño más bajo con 10.53% de rotación del personal.

- Productividad: Se mide la demanda representada por vuelos, en contraste con las horas trabajadas del personal dentro de un periodo determinado. A continuación, en el Anexo 10 se observa el performance mensual 2022 del indicador:

Como se puede apreciar en el Anexo 10, los valores del indicador estuvieron fuera de la meta de 23 puntos de productividad durante los meses de febrero a abril y diciembre del 2022, siendo febrero el mes con el menor desempeño de 18.2 puntos de productividad.

2.2.2 Diagnóstico

Con base en los indicadores previamente presentados, se ha observado que el performance anual de los resultados no está acorde a lo esperado por la compañía, por lo que se identifica la problemática de cada uno y cómo se percibe el impacto de los costos adicionales que generan a la empresa (Ver Tabla 2.10):

Tabla 2.10 Problemática de situación actual de la empresa

KPI	Problema	Impacto
Demoras de la aerolínea A	Incumplimiento de puntualidad de la aerolínea A	Si el vuelo no cumple con la puntualidad meta de la aerolínea A, se cobra una penalidad por cada demora
Servicios no conformes	Incumplimiento del nivel de servicio	Si algún criterio de nivel de servicio no se cumple en la atención del vuelo, se genera una penalidad por cada inconformidad
Productividad	Control ineficiente de horas	Si se generan exceso horas de turnos full time u horas extras fuera de la planificación que afecten el nivel de horas hombre según meta, se incurren en horas pagadas adicionales
Rotación de personal	Excesiva rotación de personal	La no cobertura de personal que cesa por fuera de lo previsto genera que se deba invertir más en consultoras para contratación de nuevo personal
%Costo de movilidad/vuelo	Altos costos de movilidad	Se generan costos de movilidad no contemplados en el presupuesto anual, lo cual altera los ratios financieros
% Planilla/venta	Altos costos de personal	Se generan costos del personal no contemplados en el presupuesto anual, lo cual altera los ratios financieros

Diagrama de Pareto

En el Anexo 11, se ilustra el diagrama de Pareto de los problemas enlistados en la tabla y su equivalente valorizado(S/) para identificar los de mayor relevancia.

Como se puede apreciar en el Anexo 11, el 80% de la problemática sobre la situación actual de la empresa está enfocado en altos costos de personal, altos costos de movilidad, incumplimiento del nivel de servicio y excesiva rotación de personal respectivamente. En ese sentido, se realizará el análisis por cada problema empleando el diagrama de Ishikawa para examinar las causas respectivas.

Diagramas de Ishikawa

A continuación, se presentan los diagramas de Ishikawa que permitirán identificar las causas de cada problema representativo presentado en el diagrama de Pareto:

- Causas de los altos costos de personal

Este problema abarca la mayor cantidad de costos presentados en el diagrama de Pareto. Como se observa en el Anexo 12, las causas se han analizado bajo los siguientes factores:

- a. Materiales: En este factor se presenta como causa a los altos costos por equipos de protección personal, uniformes e implementos, necesarios para el tipo de servicio que brinda la empresa.
- b. Métodos: En este factor se indican tres causas: extenso proceso de formación para ascender, extenso programa de capacitación y talleres de seguridad, y altos costos de contratación.
- c. Políticas del aeropuerto: Esta componente se refiere a las papeletas impuestas a los trabajadores por haber incumplido alguna norma del aeropuerto.
- d. Personal: En este último factor se indican costos relacionados a horarios nocturnos, pago de horas extras, requerimiento de personal *backup*, y principalmente, el exceso de oferta en rango horario con baja concentración de vuelos.

- Causas de los altos costos de movilidad

En este último problema, como se observa en el Anexo 12, las causas se han analizado bajo los siguientes dos factores:

- a. Métodos: Esta componente se refiere básicamente a los costos del proveedor de movilidad por tema de rutas, generación de falsos fletes y nivel de ocupación de las unidades de transporte
- b. Personal: En este factor se indican causas internas como activación de personal por tema operativo y nocturnidad de horarios. Así mismo, se menciona una causa externa debido a que la ubicación entre los puntos de recojo de los trabajadores es distante.

- Causas del Incumplimiento del nivel de servicio

En este caso, como se observa en el Anexo 13, las causas se han analizado bajo los siguientes cuatro factores:

- a. Estiba/Desestiba: En este criterio se agrupan dos causas, las cuales son daño de equipajes y demora en la estiba/desestiba. Con respecto a la primera, esta sucede debido a que el personal no manipula correctamente las maletas generando una no conformidad en el vuelo para la empresa. En cuanto a la segunda, la demora del proceso puede ocurrir por una búsqueda prolongada de equipajes en caso el vuelo sea una salida, llegada tarde de la carga o llegada tarde de los equipajes desde el *Counter*, el Patio de Equipajes o hacia las Aduanas.

- b. Personal: En este factor se indican las causas de falta de personal y personal llega tarde. En cuanto a la primera, esto sucede principalmente por el ausentismo del personal y por una planificación poco anticipada pues genera un rol incompleto de personal; en cuanto a la segunda, los motivos principales suceden por la alta cantidad de horarios nocturnos y el cambio de itinerario de vuelos.
 - c. Proceso: En este componente se mencionan las causas de demora en limpieza y error de procedimiento. Con referencia a la primera, sucede porque la política operativa define que el personal nuevo ingrese al área de Cabina y luego rote al área de Rampa. En cuanto a la segunda causa, el personal puede cometer un error por mala comunicación interna del grupo o debido a que la información recibida no se ha interpretado de forma correcta.
 - d. Equipos: En este último elemento, las causas pueden agruparse en falla de equipo, equipo llega tarde y falta de equipo. Con respecto a la primera, puede suceder debido a la mala maniobra del operador ya sea por falta experiencia o factores externos, y falla mecánica por tema de obsolescencia. En cuanto a la segunda, se presentan los mismos motivos que el factor de personal además de la causa de demora de gaseo. Por último, es posible que el equipo quede inoperativo en el taller e indisponible para su uso así como un error en la asignación del equipo para un servicio determinado.
- Causas de la excesiva rotación

En este segundo problema, como se observa en el Anexo 13, las causas se han analizado bajo los siguientes cuatro factores:

 - a. Clima laboral: En este factor se presentan dos causas, las cuales son alta cantidad de horarios nocturnos y beneficios por horario. Con respecto a la primera, la programación de horarios se determina en la elaboración de *capacity* en el proceso de planificación y según la curva de demanda se pueden considerar horarios nocturnos que se ajusten mejor. En cuanto a la segunda, la empresa brinda beneficios extra a horarios diurnos y nocturnos pues el personal puede salir o ingresar en horas complicadas para encontrar movilidad.
 - b. Operación: En esta componente se indican tres causas, las cuales son turnos rotativos, excesivos cambios de horarios y horas extras no planificadas. En cuanto al primer motivo, se debe a que la demanda de vuelos es variable por rango horario y día de la semana. Además, es posible que su programación de horarios sufra excesivos cambios en el mes. Por último, las horas extras no planificadas se originan por los cambios de itinerario y debido a que no se consideraron en el *capacity* por la poca anticipación en la planificación.
 - c. Personal: En este elemento se mencionan la causa de ausentismo reiterativo e incumplimiento de políticas internas, siendo esta última debido a algún tipo de falta por parte del trabajador al reglamento que rige la compañía.
 - d. Captación de personal: En este último criterio, se indica la causa de déficit de nuevo personal debido a que el mismo trabajador cuando inicia su etapa de inducción en la

empresa no continua o por motivo que el reclutamiento de personal por el lado del área de Recursos Humanos no se concreta a tiempo ya que el área de Planificación no envía con anticipación el requerimiento de la cantidad de personal.

Al término del análisis de causas de los problemas mencionados, se ha elaborado una clasificación ABC (Ver Anexo 14). Para priorizar aquellas de mayor relevancia, donde se realizó la asignación de un puntaje de impacto con valoración del 1 al 5 en base al criterio de experto de los especialistas y jefe del área de Planificación. Así mismo, se tomó en consideración la cantidad de repeticiones de las causas indicadas en los cuatro diagramas de Ishikawa, la cual equivale al valor de la frecuencia.

Como se observa en el Anexo 14, las causas más representativas con 38% de frecuencia acumulada en la categoría A son: Cambio de itinerario, poca anticipación en la planificación y alta cantidad de horarios nocturnos. En primer lugar, el cambio de itinerario impacta al nivel de servicio puesto que el personal no llega a tiempo a la atención del vuelo, retrasando así el proceso en la cadena operativa; en cuanto a los altos costos de personal, al existir variaciones en la hora planificada del vuelo, se busca que el personal cubra las atenciones con horas extras; a su vez, otra medida que toma la empresa para las contingencias operativas es la de activar al personal en su día de descanso programado con la condición de transportar al personal en movilidad.

En segundo lugar, la poca anticipación en la planificación afecta al nivel de servicio ya que genera subdotación de personal reflejándose en falta de personal en la operación; así mismo, es un punto crítico en la dotación de personal, ya que con una sola consultora no se alcanza el requerimiento necesario de la planificación por lo que se recurre a otras consultoras y así, la compañía debe invertir más en este aspecto. Además, en caso de la excesiva rotación, una demora en la planificación genera que al personal nuevo se le asigna más carga laboral pues no se cuenta con el *capacity* completo, lo que provoca descontento en el personal.

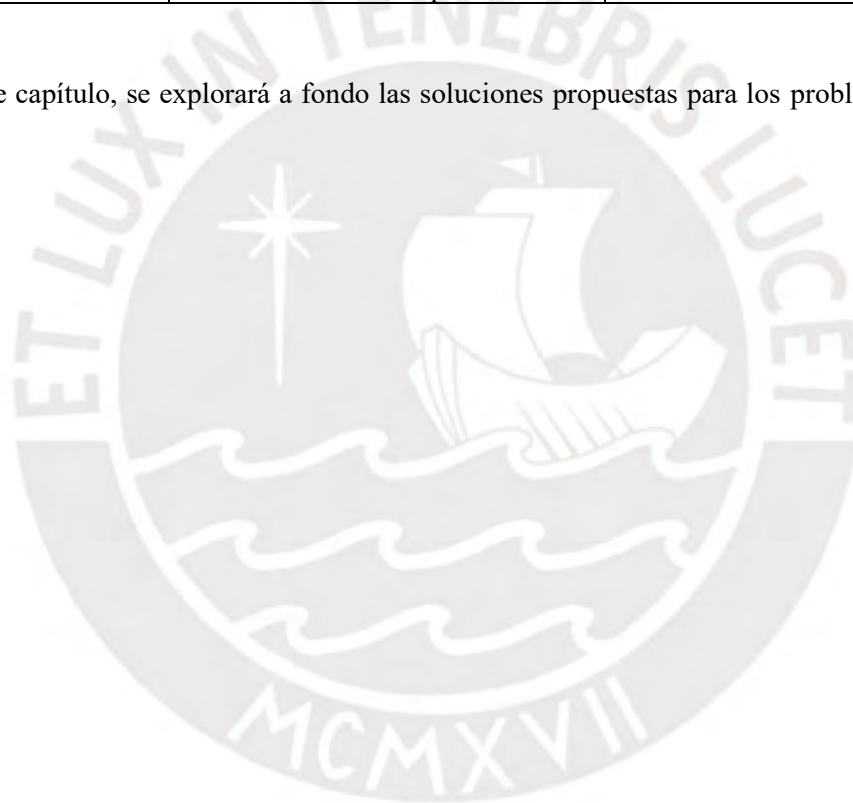
Por último, la alta cantidad de horarios nocturnos es un factor de gran impacto en los altos costos puesto que la nocturnidad representa un pago adicional al trabajador por cada hora nocturna. Así mismo, por política de la empresa, todo trabajador recibirá el beneficio de movilidad siempre que se encuentre en un horario nocturno. En ese sentido, la baja ocupación de capacidad en las unidades de transporte está vinculada a que los trabajadores viven en lugares dispersos y el uso de las unidades de transporte varían por horario, lo cual genera mayores costos en el proveedor de movilidad. Además, en caso de la alta rotación de personal, el clima laboral se afecta pues los trabajadores que se encuentran en otros horarios no reciben el beneficio de nocturnidad y movilidad. Finalmente, el nivel de servicio puede verse afectado por trabajadores en horario nocturnos que llegan tarde a la atención del vuelo.

En conclusión, al evaluar los problemas de la situación actual y seleccionar las causas de mayor relevancia, en la Tabla 2.11 se presentan las posibles soluciones.

Tabla 2.11 Propuestas de solución sobre herramientas de Analytics

Causa	Problema	Solución
Cambio de itinerario	Incumplimiento del nivel de servicio	Diseño de escenario de simulación sobre adelantos y retrasos de vuelos
	Altos costos de personal	
	Altos costos de movilidad	
Poca anticipación en la planificación	Incumplimiento del nivel de servicio	Diseño de un modelo de predicción de demanda con herramientas de aprendizaje automático
	Altos costos de personal	
	Excesiva rotación de personal	
Alta cantidad de horarios nocturnos	Incumplimiento del nivel de servicio	Diseño de un modelo de programación lineal para la optimización de horarios y personal
	Altos costos de personal	
	Altos costos de movilidad	
	Excesiva rotación de personal	

En el siguiente capítulo, se explorará a fondo las soluciones propuestas para los problemas de mayor criticidad.



CAPITULO 3. PROPUESTA DE MEJORA

En este capítulo se desarrolla la formulación de los métodos de solución planteados frente a la problemática indicada en el capítulo anterior, con el fin de resolver los temas de costos y nivel de servicio a causa de la poca anticipación a la planificación, la alta cantidad de horarios nocturnos y el cambio de itinerario. En primer lugar, se explica la metodología a seguir para cada modelo de solución y en los siguientes acápite, se desarrolla cada planteamiento en la siguiente secuencia: Diseño de un modelo de predicción de demanda con herramientas de aprendizaje automático, Diseño de un modelo de programación lineal para la optimización de horarios y personal; y Diseño de escenario de simulación sobre adelantos y retrasos de vuelos.

3.1 Metodología del problema

A continuación, en la Figura 3.1 se ilustran los pasos a seguir de las propuestas de solución

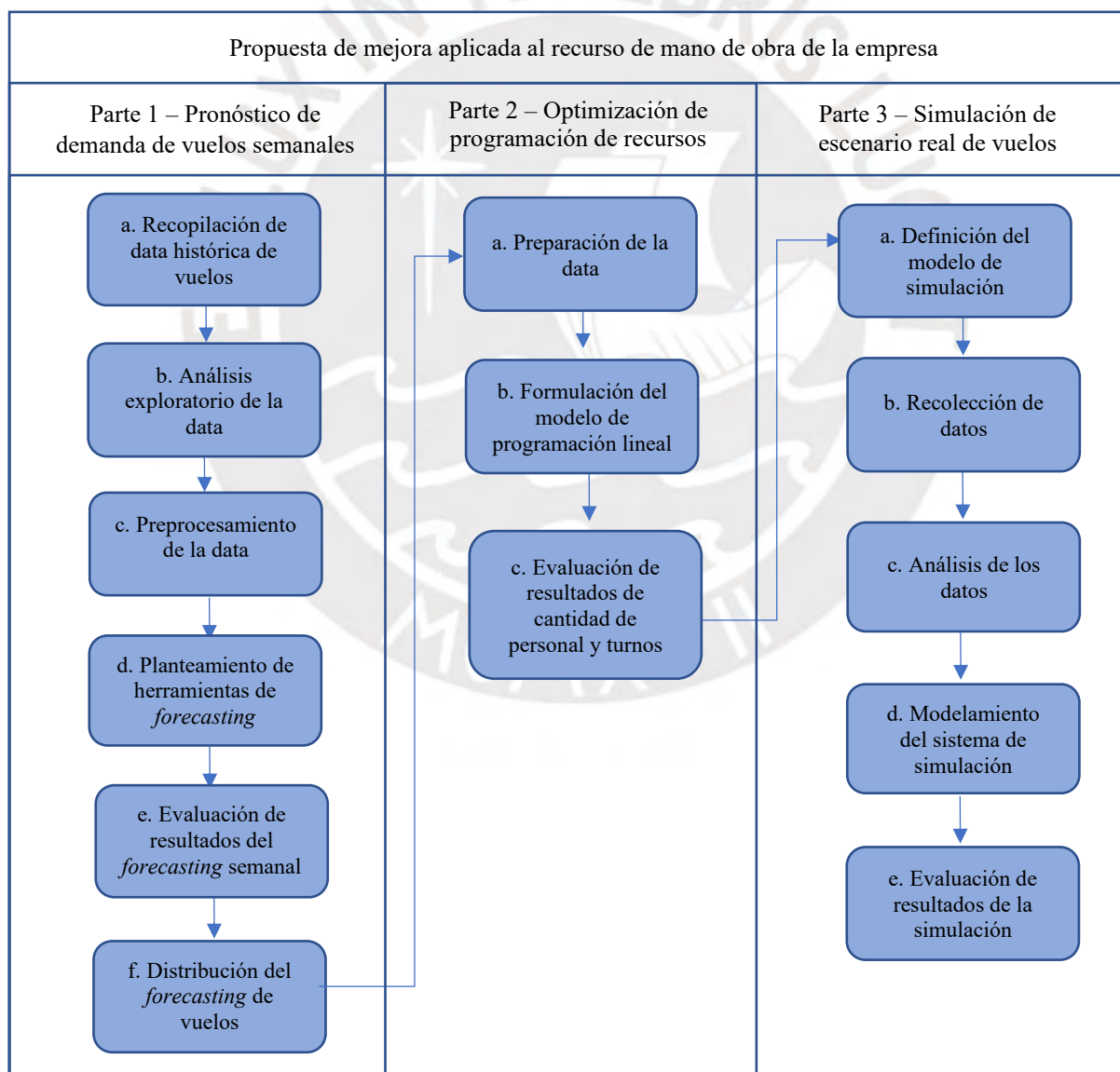


Figura 3.1 Esquema de metodología del problema

Parte 1 – Pronóstico de demanda de vuelos semanales

a. Recopilación de data histórica de vuelos

En primer lugar, se consolidará el reporte de vuelos del sistema de la empresa y se definirá el volumen de data a considerar para el diseño del modelo de predicción.

b. Análisis exploratorio de la data

En segundo lugar, se desarrollará una analítica descriptiva de la data para observar su comportamiento, tendencias de las diferentes variables relacionadas con la demanda de vuelos.

c. Preprocesamiento de la data

En base a lo observado en la exploración de la data, se identifican las variables que requieran un tratamiento previo a la aplicación de modelos de aprendizaje automático y se procede con la aplicación de las siguientes técnicas de Preprocesamiento de acuerdo con lo requerido para cada variable: Eliminación de *outliers*, completar datos faltantes, normalización, estandarización o reducción dimensional.

d. Planteamiento de herramientas de *forecasting*

Una vez que se tiene la data preparada, se determinará el tipo de aprendizaje acorde al modelo que se tiene como objetivo diseñar y, a partir de ello, se aplicarán las técnicas de predicción correspondientes.

e. Evaluación de resultados del *forecasting* semanal

Al realizar las corridas de los modelos de aprendizaje automático, se interpretarán los resultados de cada uno para compararlos y seleccionar el de mejor ajuste para la predicción.

f. Distribución del *forecasting* de vuelos

Al seleccionar el mejor modelo de predicción del volumen de vuelos por semana, se procederá a aplicar una distribución de vuelos por día y, posteriormente, por franja horaria de la curva de demanda de personal con el fin usarlo como input para el diseño de la optimización de horarios y personal. Con esta propuesta de aprendizaje automático, se espera que el modelo de predicción de la demanda de vuelos permita realizar proyecciones precisas y confiables dentro del plazo acordado, y así, evitar que la empresa se vea afectada por el nivel de servicio y sobrecostos de los que actualmente incurre.

Parte 2 – Optimización de programación de recursos

a. Preparación de la data

En primer lugar, en base a las distribuciones de vuelos por día y por franja horaria halladas en el punto anterior, se procede a aplicarlas para transformar la predicción de cantidad de vuelos en términos de demanda equivalente a grupos de personal.

b. Formulación del modelo de programación lineal

Se formulará el modelo matemático considerando los siguientes elementos: variables de decisión, función objetivo, restricciones y rango de existencia. Estos serán alimentados por la información de cantidad de vuelos obtenida en la propuesta anterior. Además de considerar la información de costos de personal por franja horaria y turnos del personal. Seguidamente, se ejecutará el modelo en la plataforma AMPL, con la información antes mencionada.

c. Evaluación de resultados de cantidad de personal y turnos

El modelo arrojará una serie de resultados óptimos para generar ahorro y ser lo más eficiente posible. Con esta propuesta de optimización, se espera que el modelo de programación lineal genere una gama de horarios que se ajusten a la curva de demanda sin ser nocturnos pues genera sobrecosto por tema de movilidad, bono de nocturnidad y afecta al nivel de servicio.

Parte 3 – Simulación de escenario de vuelos con itinerario real

a. Definición del modelo de simulación

Inicialmente, se describirá el sistema a simular junto con sus elementos tales como entidades, recursos, variables de estado y atributos. Así mismo, se estarán detallando las presunciones a considerar para la elaboración del modelo.

b. Recolección de datos

Se consolidará la base de datos de vuelos tomando en cuenta las horas reales de llegada y salida, según el proceso a seleccionar y se identificarán las variables que serán parte del sistema a simular.

c. Análisis de los datos

Una vez definidas las variables, se procederá a calcular los tamaños de muestra y con ello, analizar la data en el programa Input Analyzer para identificar cuál es la distribución que siguen los valores.

d. Modelamiento del sistema de simulación

Se construirá el flujo del proceso en el programa Arena considerando todos los elementos y las distribuciones de las variables descritas y obtenidas en los pasos previos.

e. Evaluación de resultados de la simulación

Al término de la corrida del modelo planteado, se obtendrá la cantidad de recursos necesarios en un escenario que refleja la realidad operativa ante las variaciones que se pueden presentar en el itinerario inicial programado.

Con esta propuesta de simulación, se espera validar un modelo de *capacity* de personal que considera adelantos y demoras de los vuelos pues es un factor de la realidad operativa que no se tiene reflejado al realizar la planificación actual y genera sobrecostos de personal.

3.2 Diseño de un modelo de predicción de demanda con herramientas de aprendizaje automático

Esta propuesta de solución se compone principalmente de la elaboración de un modelo para el pronóstico de la demanda de vuelos semanales y, adicional a ello, se realizará un análisis de la distribución de la demanda por día de semana y franja horaria, con el fin de considerarlo para el desarrollo del modelo de programación lineal.

3.2.1 Recopilación de data histórica de vuelos

Este primer paso inicia con la consolidación de la data de vuelos histórica que se extrae desde el sistema de la empresa, la cual abarca un periodo mayor a 8 años que va de enero 2015 a mayo 2023. A continuación, en la Tabla 3.1 se describen las variables que presenta la data:

Tabla 3.1 Descripción de variables de la data

Variable	Descripción
Semana	Esta variable representa el número de semana de los vuelos por año
Fecha	Esta variable representa la fecha en la que operaron los vuelos
Vuelos	Esta variable representa el volumen de vuelos reales diarios
DiaSem	Esta variable representa el día de la semana en números del 1 al 7

Así mismo, se muestra en la Tabla 3.2, parte de la data histórica con la demanda de vuelos recopilada por semanas de enero 2015:

Tabla 3.2 Muestra de la data histórica de vuelos

Semana	Fecha	Vuelos	DiaSem
1	1/01/2015	100	4
1	2/01/2015	99	5
1	3/01/2015	99	6
1	4/01/2015	104	7
2	5/01/2015	102	1
2	6/01/2015	89	2
2	7/01/2015	97	3
2	8/01/2015	94	4
2	9/01/2015	100	5
2	10/01/2015	92	6
2	11/01/2015	98	7
3	12/01/2015	97	1
3	13/01/2015	91	2
3	14/01/2015	97	3
3	15/01/2015	99	4
3	16/01/2015	99	5
3	17/01/2015	98	6
3	18/01/2015	97	7

3.2.2 Análisis exploratorio de la data

En este segundo paso, se procede con el análisis de tendencias históricas de vuelos en función de las demás variables. Como se puede observar en la Figura 3.2, se aprecia una tendencia estable y creciente de vuelos desde enero 2015 hasta febrero del 2020, puesto que en marzo de ese año se visualiza una caída abrupta de la demanda. Es de conocimiento que, a partir de ese periodo se desarrolló una coyuntura atípica de pandemia, la cual generó el cierre del aeropuerto de Lima afectando así la operación de vuelos a nivel nacional e internacional. En ese sentido, se observa que la recuperación de la demanda en la industria aeroportuaria empezó desde la segunda mitad del 2020, con algunas caídas debido a olas de COVID, conflictos sociales y otros factores externos.

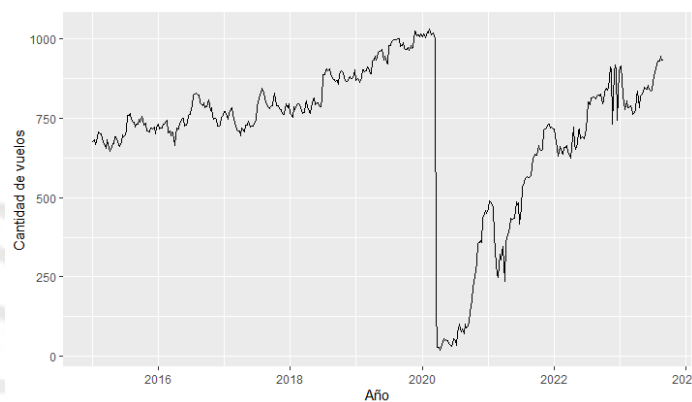


Figura 3.2 Histórico de vuelos del 2015 al 2023

3.2.3 Preprocesamiento de la data

En base a lo expuesto en la etapa anterior de exploración, se observa que se debe realizar un preprocesamiento al conjunto de datos atípicos del 2020 al 2021 con el fin de no sesgar los resultados del modelo de predicción a desarrollar. A continuación, se detallan los factores que afectaron en la demanda histórica de vuelos:

Factor Pandemia

El factor principal que afectó los vuelos a nivel global desde el año 2020 hasta la actualidad es la pandemia por COVID-19, aún al día de hoy la industria aeronáutica no se ha recuperado por completo de lo que supuso esta crisis.

De acuerdo con el informe de desempeño de la OSITRAN (2021), el tráfico anual de aeronaves de los años 2020 y 2021 comparados al año 2019, presentan un decrecimiento del 63% y 48%, respectivamente. Estas caídas se evidencian en la Figura 3.2. Con el fin de evitar sesgos en la información de vuelos, se resuelve no considerar los años 2020 y 2021, dado que estos presentan la mayor caída en las cifras históricas.

Asimismo, acorde a lo indicado en la memoria integrada 2021 de LATAM Airlines Group S.A., aerolínea de mayor volumen de vuelos en Perú, en diciembre 2021 logró alcanzar el 70.1% de su capacidad comparado con 2019.

Con el objetivo de ajustar los datos previos a pandemia (2015-2019) a la misma escala que los datos posteriores a pandemia (2022-2023), se utilizó el indicador antes mencionado. Los datos ajustados se muestran en la Figura 3.3. Se observa que la tendencia se ve mejor; sin embargo, aún se presentan semanas atípicas por otros factores que se detallan posteriormente.

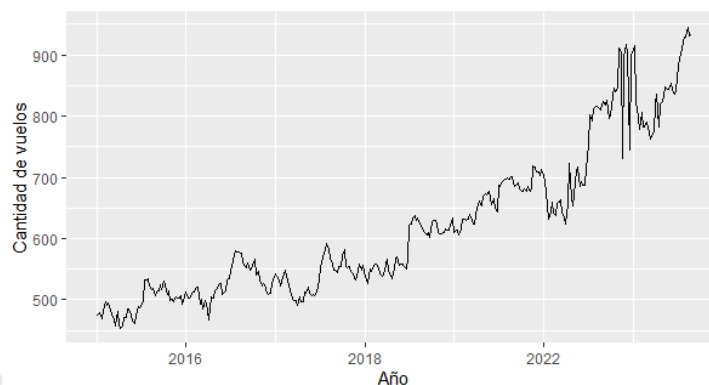


Figura 3.3 Vuelos ajustados por factor pandemia

Factores externos de coyuntura nacional

Al observar la curva de vuelos histórica en el año 2022, se puede apreciar tres periodos atípicos en la gráfica, los cuales se presentan a continuación:

a) Periodo del 10.01 al 30.01

De acuerdo con la información del MINSA (2022), en enero de ese año ocurrió el acontecimiento de la tercera ola por crecimiento de casos de contagio de COVID-19, puesto que la cifra a nivel nacional se duplicó con respecto a las registradas en diciembre del año anterior, afectando así la demanda en caída durante las semanas 3 a la 5 tal como se muestra en la Tabla 3.3. En ese sentido, se realizó un ajuste en los vuelos diarios en base a los vuelos confirmados por las aerolíneas considerando aplicar un factor de 96% de cumplimiento histórico del mes de enero de ese año previo a la tercera ola. Como resultado del ajuste, estos son los nuevos valores de las semanas de enero 2022 (ver Tabla 3.4).

Tabla 3.3 Data de vuelos afectados por la tercera ola del COVID-19

Semana	Fecha	Vuelos	DiaSem
3	10/01/2022	97	1
3	11/01/2022	88	2
3	12/01/2022	89	3
3	13/01/2022	97	4
3	14/01/2022	93	5
3	15/01/2022	92	6
3	16/01/2022	100	7
4	17/01/2022	93	1
4	18/01/2022	89	2
4	19/01/2022	90	3
4	20/01/2022	93	4
4	21/01/2022	93	5
4	22/01/2022	84	6
4	23/01/2022	89	7
5	24/01/2022	92	1
5	25/01/2022	85	2
5	26/01/2022	90	3
5	27/01/2022	95	4
5	28/01/2022	98	5
5	29/01/2022	90	6
5	30/01/2022	93	7

Tabla 3.4 Data ajustada de vuelos afectados por la tercera ola del COVID-19

Semana	Fecha	Vuelos	DiaSem
3	10/01/2022	100	1
3	11/01/2022	94	2
3	12/01/2022	98	3
3	13/01/2022	102	4
3	14/01/2022	103	5
3	15/01/2022	96	6
3	16/01/2022	103	7
4	17/01/2022	100	1
4	18/01/2022	92	2
4	19/01/2022	96	3
4	20/01/2022	100	4
4	21/01/2022	98	5
4	22/01/2022	97	6
4	23/01/2022	100	7
5	24/01/2022	100	1
5	25/01/2022	92	2
5	26/01/2022	96	3
5	27/01/2022	100	4
5	28/01/2022	98	5
5	29/01/2022	97	6
5	30/01/2022	100	7

b) Periodo del 18.11 al 20.11

En base a lo extraído según el diario La República (2022), el 18 de noviembre de ese año se produjo la colisión de un avión de la aerolínea LATAM con un camión de bomberos dentro de la pista de aterrizaje del aeropuerto Jorge Chávez, por lo que hubo cierre del mismo incluso hasta las 00:00 h del domingo 20 de noviembre de acuerdo con lo mencionado por el diario El Comercio (2022). Este suceso afectó negativamente en la operación aeroportuaria por lo que hubo cancelación de vuelos tal como se ve en la Tabla 3.5.

Tabla 3.5 Data de vuelos afectados por accidente aeroportuario

Semana	Fecha	Vuelos	DiaSem
47	18/11/2022	93	5
47	19/11/2022	1	6
47	20/11/2022	109	7

Sin embargo, en caso de un escenario regular de operación con apertura de aeropuerto, se considera aplicar un factor histórico de 98% del mes de diciembre. En consecuencia, estos son los nuevos valores en el mes de noviembre 2022 (ver Tabla 3.6):

Tabla 3.6 Data ajustada de vuelos afectados por accidente aeroportuario

Semana	Fecha	Vuelos	DiaSem
47	18/11/2022	126	5
47	19/11/2022	129	6
47	20/11/2022	128	7

c) Periodo del 12.12 al 18.12

Según información del diario El Comercio(2022), en diciembre de ese año se produjo el cierre de los aeropuertos de Cusco, Arequipa, Juliaca y Ayacucho a causa de las violentas protestas por coyuntura política del país y, así como en los factores anteriores, se tuvo una caída de los vuelos para esos destinos afectando de esta forma la demanda total (ver Tabla 3.7).

Tabla 3.7 Data de vuelos afectados por manifestaciones nacionales

Semana	Fecha	Vuelos	DiaSem
51	12/12/2022	132	1
51	13/12/2022	97	2
51	14/12/2022	85	3
51	15/12/2022	97	4
51	16/12/2022	106	5
51	17/12/2022	115	6
51	18/12/2022	113	7

Por esta razón, se realiza un ajuste aplicando un factor de 98% histórico del mes de diciembre 2022 al total de vuelos para evitar incluir una demanda de comportamiento atípico (ver Tabla 3.8).

Tabla 3.8 Data ajustada de vuelos afectados por manifestaciones nacionales

Semana	Fecha	Vuelos	DiaSem
51	12/12/2022	137	1
51	13/12/2022	132	2
51	14/12/2022	123	3
51	15/12/2022	136	4
51	16/12/2022	135	5
51	17/12/2022	133	6
51	18/12/2022	135	7

Posteriormente, para analizar la data como serie de tiempo, se agrega una columna denominada “Semana 2”, la cual indica el número de semanas de forma secuencial como se observa en la Tabla 3.9

Tabla 3.9 Ejemplo de la nueva versión de la data

Semana1	Semana2	Fecha	Vuelos	DiaSem
53	53	28/12/2015	73	1
53	53	29/12/2015	72	2
53	53	30/12/2015	73	3
53	53	31/12/2015	70	4
1	53	1/01/2016	69	5
1	53	2/01/2016	74	6
1	53	3/01/2016	73	7
2	54	4/01/2016	76	1
2	54	5/01/2016	76	2
2	54	6/01/2016	71	3
2	54	7/01/2016	74	4
2	54	8/01/2016	73	5
2	54	9/01/2016	70	6
2	54	10/01/2016	72	7
3	55	11/01/2016	74	1
3	55	12/01/2016	69	2
3	55	13/01/2016	68	3
3	55	14/01/2016	72	4
3	55	15/01/2016	75	5
3	55	16/01/2016	72	6
3	55	17/01/2016	72	7

Seguidamente, al finalizar los ajustes de la data según lo mencionado sobre el factor pandemia atípico entre el 2020 y 2021 y los demás factores de coyuntura interna ocurridos en el 2022, se puede visualizar la demanda final que se usará para la serie de tiempo en la Figura 3.4.

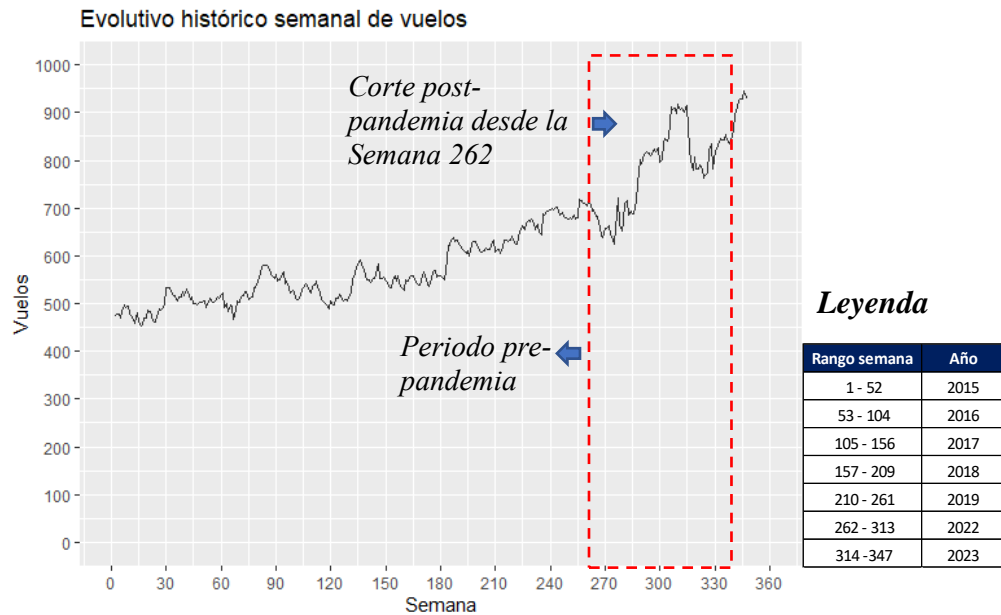


Figura 3.4 Evolutivo de demanda histórica final

Como se aprecia, la fecha de corte pre-pandemia es el domingo 05 de enero del 2020 y luego pasa a la fecha de recuperación post-pandemia, la cual es el lunes 03 de enero del 2022.

Por ende, la data de análisis para construir el modelo ahora abarcará un periodo mayor a 6 años; es decir, un total de 347 semanas desde enero 2015 hasta agosto 2023.

3.2.4 Planteamiento de herramientas de forecasting

Una vez que se tiene la data modificada, se determinará el tipo de aprendizaje acorde al modelo que se tiene como objetivo diseñar y, a partir de ello, se aplicarán las técnicas de predicción correspondientes. De esta manera, se realiza la división de las 347 semana de la data para segmentar el *train* y *test* como se detalla en la Tabla 3.10.

Posteriormente, al identificar que en esta primera propuesta del presente estudio se tiene como objetivo pronosticar la demanda de vuelos semanales, se deberá emplear técnicas para modelos de regresión, de los cuales se considera optar por la regresión lineal como primer modelo básico y por el uso de los métodos ARIMA y SARIMA, puesto que se está empleando una serie de tiempo como análisis para los pronósticos.

Tabla 3.10 Separación de datos

Data	División	Cantidad de semanas	Hito de Semana	Nº Semana	Rango fechas
Train	80%	295	Inicio	1	05.01.2015 - 11.01.2015
			Fin	295	22.08.2022 - 28.08.2022
Test	20%	52	Inicio	296	29.08.2022 - 04.09.2022
			Fin	347	21.08.2023 - 27.08.2023

A continuación, se presentan los modelos mencionados que fueron desarrollados en el software RStudio:

Regresión Lineal

Previamente a la elaboración del modelo, en la Figura 3.5 se muestra parte de los valores de la serie de tiempo de vuelos semanales.

Vuelos	
[1,]	473
[2,]	477
[3,]	479
[4,]	469
[5,]	483
[6,]	496
[7,]	491
[8,]	494
[9,]	482
[10,]	476
[11,]	468
[12,]	458
[13,]	480
[14,]	461

Figura 3.5 Serie de tiempo de vuelos semanales

Cabe indicar que en el comando del código del software RStudio, se está tomando en cuenta que la duración del año es equivalente a 52 semanas.

En la Figura 3.6 se muestra la gráfica de la serie de tiempo del *train* mencionada en líneas anteriores. Como observación de la gráfica, se puede apreciar que la curva presenta oscilaciones con periodos picos y valles que van en tendencia creciente a lo largo del tiempo.

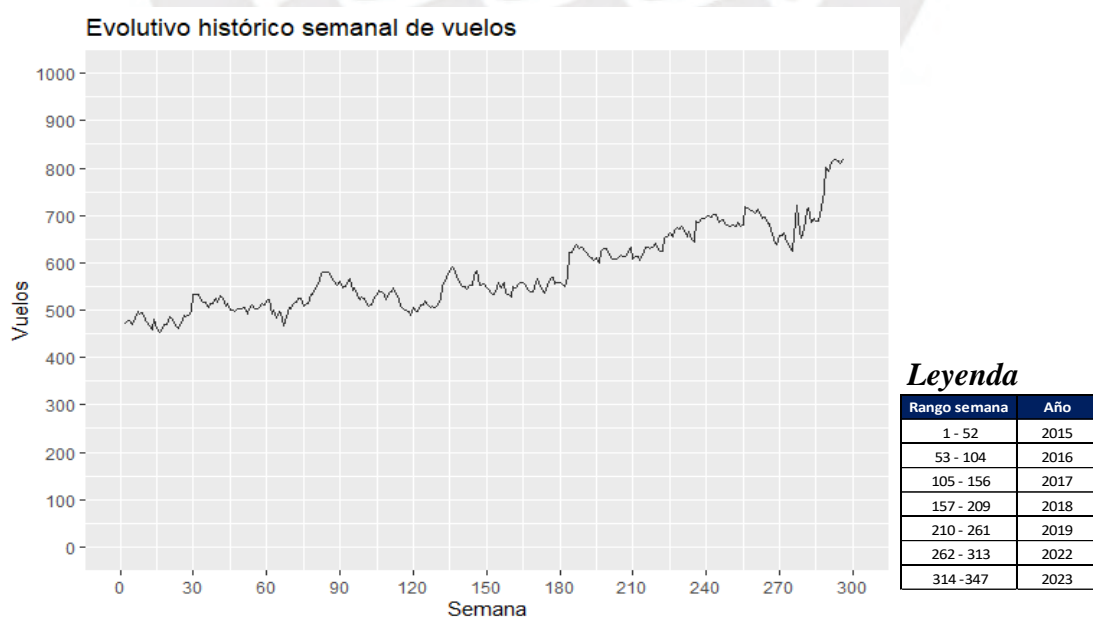


Figura 3.6 Evolutivo del *train* de vuelos semanales

Una vez que se tiene definida la serie, se procede a desarrollar el modelo de regresión lineal considerando lo siguiente:

- Variable dependiente (y): Vuelos
- Variable independiente (x): Semana

A continuación, en la Figura 3.7 se presenta el resumen del modelo calculado en RStudio :

```
Call:
lm(formula = Vuelos ~ Semana, data = train_rl, na.action = na.exclude)

Residuals:
    Min       1Q   Median       3Q      Max
-66.272 -22.976   1.146  20.716 112.374

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 451.01896   3.99060  113.02  <2e-16 ***
Semana       0.86852    0.02337   37.16  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 34.18 on 293 degrees of freedom
Multiple R-squared:  0.825,    Adjusted R-squared:  0.8244
F-statistic: 1381 on 1 and 293 DF,  p-value: < 2.2e-16
```

Figura 3.7 Resumen del modelo de Regresión Lineal

Para el análisis del modelo de regresión aplicado en la serie de tiempo para pronosticar la variable “Vuelos”, se puede observar que el coeficiente del intercepto es de 451.01896, el cual se muestra significativamente distinto de cero puesto que el valor p es en extremo pequeño ($<2e^{-16}$), siendo este mucho menor que un nivel de significancia de 0.05. En ese sentido, esto indica que el intercepto es un valor relevante para explicar el impacto de las variaciones de demanda de vuelos debido al transcurso de cada semana.

Así mismo, por cada semana transcurrida en el tiempo, esta se relaciona con el crecimiento en la cantidad de vuelos en 0.86852; es decir, en 1 vuelo; el cual, es posible que suceda en caso la demanda se mantenga estable semana a semana.

Además, el modelo presenta un valor de R-cuadrado ajustado de 0.8244, donde se interpreta que el 82% de la variabilidad en la demanda de vuelos puede explicarse por la variable semana.

ARIMA

En esta sección, se desarrollará el modelo ARIMA considerando las siguientes fases de Identificación del Modelo, Ajustar el Modelo, Validación del Modelo y Pronósticos ilustrados en la Figura 1.4.

I. Identificación del Modelo

En esta primera etapa se incluye la data de la serie de tiempo que usará para el entrenamiento del modelo y la validación de la misma para probar si es o no estacionaria.

I.1. Identificación de la serie

La serie de tiempo a usar para el modelo ARIMA es el que ya se mencionó en líneas anteriores para el modelo de regresión lineal.

I.2. Validar estacionaridad

Ahora se procede a aplicar la prueba de Dickey Fuller Aumentada para validar la estacionaridad de la serie; es decir, que esta presente propiedades estadísticas que no varíen en el tiempo como por ejemplo media y varianza constantes y ausencia de tendencia. De esta manera, se plantean las siguientes hipótesis:

H_0 : La serie de tiempo es estacionaria.

H_1 : La serie de tiempo no es estacionaria.

La evaluación si se rechaza o no la prueba consistirá en comparar el p-value resultante en contraste con un nivel de significancia (α) equivalente a 0.05. Entonces, en caso el p-value obtenido es menor o igual al nivel de significancia (α), no existe fundamento para rechazar la prueba y validaría la estacionaridad de la serie; de lo contrario, se rechaza la premisa de la hipótesis descartando así la estacionaridad de la serie.

Seguidamente, se realiza la prueba en el software RStudio y estos son los resultados obtenidos: (Ver Figura 3.8).

```
Augmented Dickey-Fuller Test
data: ts_train
Dickey-Fuller = -2.0903, Lag order = 6, p-value = 0.5382
alternative hypothesis: stationary
```

Figura 3.8 Prueba de Dickey-Fuller Aumentada para la serie de tiempo de vuelos

Donde “ts_train” es la serie de tiempo usada para el entrenamiento del modelo.

Se puede observar que el resultado del p-value en la prueba fue de 0.5382, el cual es un valor mayor a 0.05; por lo tanto, la hipótesis nula sobre si la serie presenta estacionaridad se rechaza.

I.3. Transformar la serie

Debido a que se comprobó por la prueba de Dickey Fuller Aumentada que la serie no es estacionaria, entonces se aplicará una transformación de la primera diferencia para evaluar si ya cumple las condiciones de estacionaridad. Así mismo, en la Figura 3.9, se ilustra la gráfica de esta primera diferencia aplicada a la serie y se puede observar que las oscilaciones de los valores se mueven en torno a 0.

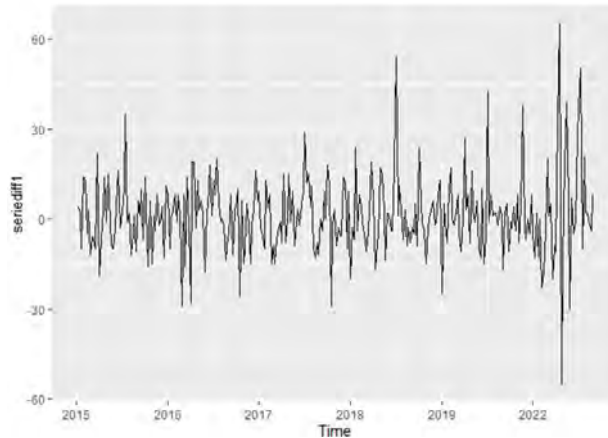


Figura 3.9 Serie diferenciada de primer orden

El siguiente paso es comprobar a través de la prueba de Dickey Fuller Aumentada que esta primera diferencia de la serie presenta estacionaridad. En esa línea, se exponen las siguientes pruebas de hipótesis:

H_0 : La primera diferencia de la serie de tiempo es estacionaria.

H_1 : La primera diferencia de la serie de tiempo no es estacionaria.

Tal como el caso anterior, se tomará la decisión de rechazar o no la hipótesis nula en base al p-value que se genere al aplicar la prueba. En otras palabras, si el valor del p-value resulta menor o igual al nivel de significancia (α) de 0.05, no hay sustento para rechazar la hipótesis nula que implica la estacionaridad de la serie; de lo contrario, se tomaría la decisión de rechazar la hipótesis nula y se tendría que aplicar una segunda diferencia a la serie de tiempo y volver a repetir la prueba de Dickey Fuller Aumentada hasta que se compruebe que alguna diferencia de la serie de tiempo presente estacionaridad.

```

Augmented Dickey-Fuller Test
data: seriediff1
Dickey-Fuller = -7.2617, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary

```

Figura 3.10 Prueba de Dickey-Fuller Aumentada para la serie diferenciada

Donde “seriediff1” es la primera diferencia de la serie de tiempo usada para el entrenamiento del modelo.

En la Figura 3.10 se presenta el resultado de la prueba de Dickey Fuller Aumentada para la primera diferencia siendo el valor del p-value menor al nivel de significancia (α) de 0.05; por esa razón, la hipótesis nula acerca de la estacionaridad de la serie no se rechaza.

II. Ajustar el Modelo

Una vez concluida la validación de la estacionaridad de la serie y se ha comprobado que, con una diferencia, la serie que se está analizando logra esta condición, se procede a determinar los parámetros para el modelo ARIMA.

II.1. Determinar los parámetros ARIMA

En primer lugar, el parámetro d del modelo sería igual a 1 puesto que es la cantidad de diferenciaciones para que la serie sea considerada estacionaria. Posterior a ello, se continúa hallando los parámetros p y q en el siguiente orden:

Componente AR

Para hallar el parámetro p acerca de la cantidad de valores autoregresivos, en la Figura 3.11 se muestra la gráfica de la función de autocorrelación parcial de la serie. Tal como se aprecia en la figura, al analizarse solo el correlograma de PACF, se observa que hay dos rezagos por fuera del rango de confianza hasta antes de las primeras 52 semanas, por lo que se considera que el valor del parámetro p sería igual a 2 y se tendría un modelo ARIMA (2,1,0).

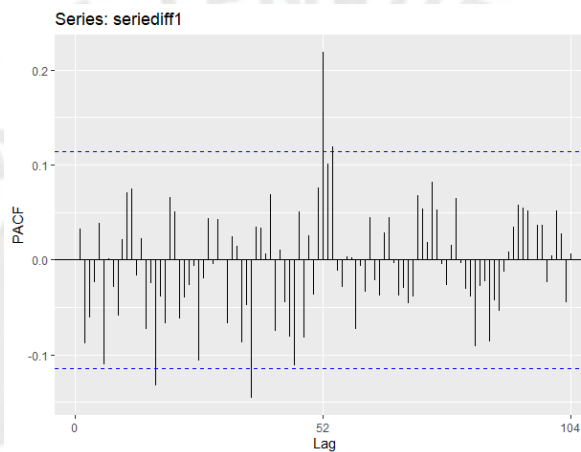


Figura 3.11 Autocorrelación parcial de la serie diferenciada

Componente MA

Para determinar el parámetro q acerca de la cantidad de medias móviles de la serie, se presenta a continuación la Figura 3.12 sobre la gráfica de la función de autocorrelación:

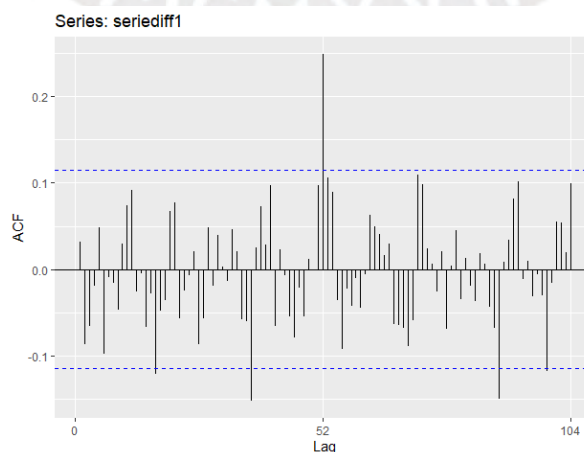


Figura 3.12 Autocorrelograma de la serie diferenciada

Como se puede observar en la Figura 3.12, si el análisis se enfoca en el correlograma ACF, se observan dos rezagos por fuera del rango de confianza hasta antes de las primeras 52 semanas; por lo tanto, se considera que el parámetro q tome el valor de 2. Por lo tanto, se obtendría un modelo ARIMA(0,1,2).

Por otro lado, si se analiza conjuntamente la componente AR y MA, entonces se formaría el modelo ARIMA(2,1,2).

De esta manera, se seleccionará un primer modelo tomando en ambas componentes AR y MA, por lo que el modelo a desarrollar será el ARIMA(2,1,2).

II.2. Evaluar del modelo

Ahora que ya se tienen los parámetros de este primer modelo ARIMA(2,1,2), en la Figura 3.13 se muestra un resumen del modelo en el software RStudio:

```
Call:
arima(x = ts_train, order = c(2, 1, 2))

Coefficients:
      ar1      ar2      ma1      ma2
  0.5753 -0.1027 -0.5376  0.0019
s.e.  0.5221  0.4940  0.5251  0.4935

sigma^2 estimated as 176.4:  log likelihood = -1177.56,  aic = 2365.12
```

Figura 3.13 Resumen del modelo ARIMA(2,1,2)

A continuación, se detalla la interpretación de los resultados del modelo analizado en la Figura 3.13:

- Con referencia al primer coeficiente para la componente autoregresiva, el valor calculado es de 0.5753, el cual implica que existe una dependencia lineal con peso positivo entre la observación presente y la inmediata anterior.
- Con referencia al segundo coeficiente para la componente autoregresiva, el valor calculado es de -0.1027, el cual implica que existe una dependencia lineal con peso negativo entre la observación presente y la de dos periodos anteriores.
- El valor obtenido del primer coeficiente de la componente de media móvil es -0.5376, el cual indica que existe una dependencia lineal con peso negativo entre la observación presente y el error del periodo anterior.
- Con respecto al segundo coeficiente para la componente de media móvil, el valor calculado es de 0.0019, donde se interpreta que existe una dependencia lineal con peso positivo entre la observación actual y el error de dos periodos pasados.
- En cuanto a las desviaciones estándar, se observa que los valores para los coeficientes de la componente AR y MA son de 0.5221, 0.4940, 0.5251 y 0.4935 respectivamente, siendo este último el menor error pues es más cercano a 0.

- Se observa que el AIC del modelo tiene un valor de 2365.12, el cual debe ser comparado con el AIC de otro modelo puesto que es una medida de contraste.

II.3. Seleccionar modelo óptimo

Una vez que se tiene un primer modelo ARIMA hallado en la sección anterior, se procede a encontrar el modelo que posea el mejor ajuste para la serie de tiempo que está analizando, el cual se detallará más adelante en el capítulo 4.

SARIMA

En esta sección, se desarrollará el modelo SARIMA o también llamado ARIMA estacional, considerando las mismas fases que el modelo ARIMA: Identificación del Modelo, Ajuste del Modelo, Validación del Modelo y Pronósticos. Las dos últimas fases se detallan en el capítulo 4.

I. Identificación del Modelo

En esta primera etapa se incluye la identificación de la serie, validación de estacionaridad y transformación de la serie.

I.1. Identificación de la serie

Se utiliza la misma serie de tiempo que en ARIMA, la cual se detalló anteriormente. Esta consta de 295 registros por semana, desde el 2015 hasta agosto de 2023, saltando los años 2020 y 2021.

En la Figura 3.14, se muestra la descomposición de la serie, con el objetivo de mostrar gráficamente si existe tendencia o estacionalidad. En este caso, podemos observar que existe una tendencia de crecimiento, con un pequeño valle hacia el final. Asimismo, se observa un patrón estacional que se repite cada año, lo cual podría sugerir que un modelo SARIMA sería adecuado.

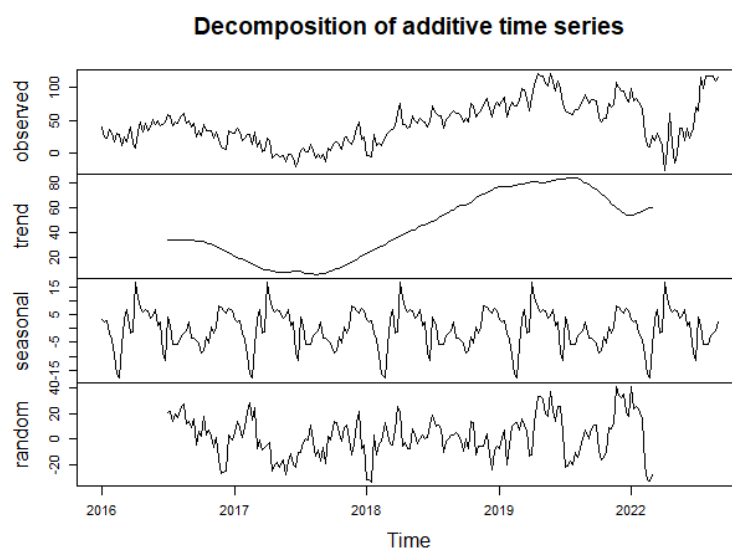


Figura 3.14 Descomposición aditiva de series de tiempo

En el Anexo 15, se muestra la serie de tiempo en una gráfica estacional dividida por años. En esta también se puede visualizar la presencia de un posible patrón estacional. Por ejemplo, se observa que en la mayoría de los años existe un crecimiento en la cantidad de vuelos a mitad de año, alrededor de la semana 27, en el mes de Julio. Este crecimiento podría deberse a la cantidad de feriados de ese mes, que fomentan los viajes dentro del país.

I.2. Validación de estacionaridad

Dado que se evalúa la misma serie que para ARIMA, ya se tiene los resultados de la validación de la estacionaridad de la serie. Para este caso, la serie temporal no pasó la prueba de Dickey Fuller Aumentada, por lo que esta no sería estacionaria. Para que lo sea, se tendría que transformar la serie, lo cual se detalla en el siguiente punto.

I.3. Transformación de la serie

Debido a que la serie no es estacionaria y presentan cierta estacionalidad, se procederá a aplicar una diferencia estacional, considerando un periodo de 52, por la cantidad de semanas que consideramos por año. En la Figura 3.15, se muestra la serie estacionalmente diferenciada, la cual gráficamente no parece ser estacionaria, para verificarlo se aplicará nuevamente la prueba de Dickey Fuller Aumentada.

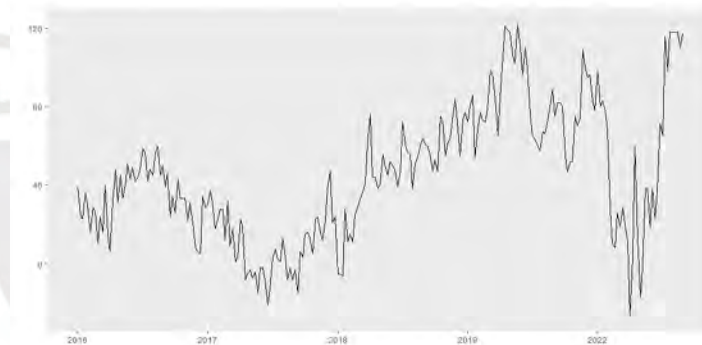


Figura 3.15 Serie diferenciada estacionalmente

Para la prueba, se consideran las siguientes hipótesis:

H_0 : La primera diferencia de la serie de tiempo es estacionaria.

H_1 : La primera diferencia de la serie de tiempo no es estacionaria.

El p-value determinará si los datos pasan la prueba, en caso su valor sea menor o igual al nivel de significancia (α) de 0.05, no habría sustento para rechazar la hipótesis nula, lo que supondría que la serie es estacionaria.

```
Augmented Dickey-Fuller Test
data: sarima_diff
Dickey-Fuller = -2.5302, Lag order = 6, p-value = 0.3528
alternative hypothesis: stationary
```

Figura 3.16 Prueba de Dickey Fuller Aumentada para la serie estacionalmente diferenciada

Donde “sarima_diff” es la primera diferencia estacional de la serie de tiempo usada para el entrenamiento del modelo.

De acuerdo con lo descrito en la Figura 3.16, el valor de *p-value* es mayor a 0.05, por lo que se rechaza la hipótesis nula y tal como se observó gráficamente, la serie no sería estacionaria. Por lo tanto, será necesario volver a realizar una diferencia para que la serie se vuelva estacionaria. La serie doblemente diferenciada se muestra en la Figura 3.17, gráficamente se observa que los valores oscilan en torno a 0, por lo que se podría suponer que es estacionaria, para validarlo se volverá a aplicar la prueba de Dickey Fuller Aumentada. De acuerdo con los resultados de la prueba, el valor de *p-value* es menor que el nivel de significancia, por lo que se acepta la hipótesis nula, en consecuencia, se comprueba que la serie es estacionaria (Ver Figura 3.18).

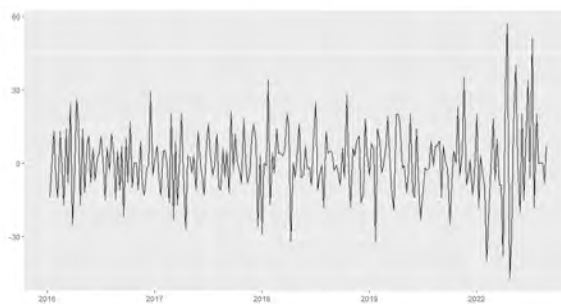


Figura 3.17 Serie doblemente diferenciada

```
Augmented Dickey-Fuller Test
data: sarima_diff2
Dickey-Fuller = -6.0223, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

Figura 3.18 Prueba de Dickey Fuller Aumentada para la serie doblemente diferenciada

II. Ajuste del Modelo

Luego de transformar la serie, de forma que sea estacionaria, el siguiente paso es determinar los parámetros base para el modelo SARIMA.

I.1. Determinar los parámetros SARIMA

Con el fin de determinar los parámetros adecuados para el modelo base de SARIMA se utilizan los gráficos de autocorrelación y autocorrelación parcial de la serie de tiempo doblemente diferenciada (Ver Figura 3.19 y Figura 3.20).

Se puede observar en la Figura 3.19, tres picos significativos en los rezagos 1, 2 y 3, lo que sugiere un componente AR(3) no estacional y a su vez, el rezago significativo 52 sugiere un componente AR(1) estacional. Asimismo, dado que la serie fue diferenciada al menos una vez, se considera para el parámetro *d* y *D* el valor de 1. En consecuencia, se sugiere como modelo base: ARIMA(3,1,0)(1,1,0)₅₂.

Por otro lado, si se analizaba el gráfico ACF, también se habría podido sugerir un componente MA(2) no estacional, al tener dos rezagos significativos, y un componente MA(1) estacional, al tener un pico significativo en el rezago 52, obteniéndose de esta forma un modelo base: $ARIMA(0,1,2)(0,1,1)_{52}$.

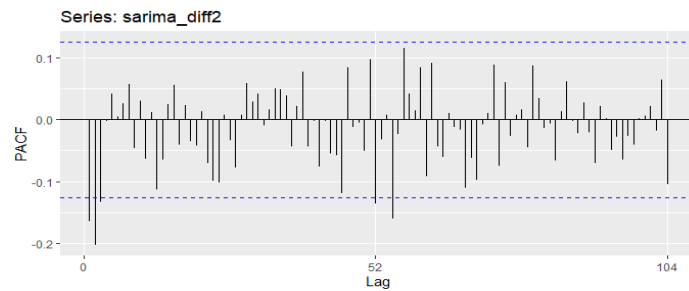


Figura 3.19 Autocorrelación parcial en serie doblemente diferenciada

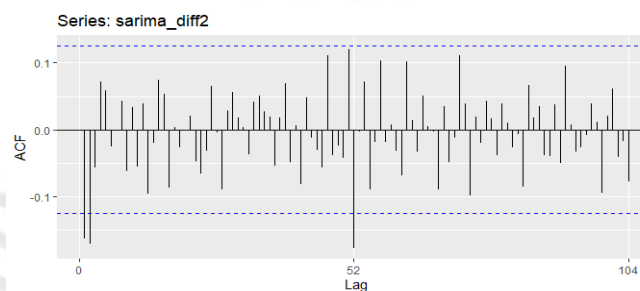


Figura 3.20 Autocorrelación en serie doblemente diferenciada

I.2. Evaluación del modelo

De acuerdo con lo determinado en la sección anterior, se muestra los resultados del modelo base en la Figura 3.21, hallados mediante el software RStudio.

```
Series: ts_train
ARIMA(3,1,0)(1,1,0)[52]

Coefficients:
      ar1      ar2      ar3      sar1
    -0.2072 -0.2352 -0.1707 -0.3109
s.e.    0.0635  0.0628  0.0641  0.0778

sigma^2 = 182.5:  log likelihood = -974.09
AIC=1958.18  AICC=1958.43  BIC=1975.62
```

Figura 3.21 Resumen del modelo $ARIMA(3,1,0)(1,1,0)_{52}$

- Con respecto al primer coeficiente para la componente autorregresiva ($ar1$), el valor calculado es de -0.2072 , lo cual implica que existe una dependencia lineal con peso negativo entre la observación presente y la inmediata anterior. Asimismo, respecto al segundo y tercer coeficiente de la componente autorregresiva ($ar2$ y $ar3$), sus valores negativos implican que existe una dependencia lineal con peso negativo entre la observación presente y las observaciones de dos y tres periodos pasados.
- Con respecto al coeficiente para la componente autorregresiva estacional ($sar1$), el valor calculado es de -0.3109 , lo cual implica que existe una dependencia lineal con peso negativo entre la observación presente y la observación de 52 periodos pasados.

- El AIC del modelo tiene un valor de 1958.18, el cual debe ser comparado con el AIC de otro modelo, con el fin de elegir el más adecuado.

I.3. Selección del modelo óptimo

A partir del modelo base SARIMA delimitado en la sección anterior, se procederá a buscar el modelo de mejor ajuste para la serie de tiempo, este proceso se detalla en el capítulo 4.

3.2.5 Distribución del forecasting de vuelos

En este acápite se desarrolla cómo está conformada la distribución semanal y por franja horaria de la demanda de vuelos y cantidad de personal respectivamente, puesto que servirá posteriormente como parte del input para la propuesta del diseño de modelo de programación lineal.

Distribución por día de semana

Las variaciones en la cantidad de vuelos en el tiempo, observadas previamente al realizar el modelo predictivo, no solo se presentan por semana, sino que también serán analizadas por día para determinar su distribución. A continuación, en la Figura 3.22 se ilustra la demanda de vuelos por día de una semana de muestra extraída del 2023. A partir de esta semana de muestra, se calcula la distribución porcentual de demanda por día de semana mediante la división de la demanda de vuelos de cada día entre el total de vuelos para esa semana determinada (Ver Figura 3.23).



Figura 3.22 Demanda de vuelos por día de una semana de muestra



Figura 3.23 Distribución de demanda por día de una semana de muestra

Para obtener una distribución porcentual representativa se consideró data semanal de los meses de enero a junio 2023, los cuales conforman 27 semanas. Así mismo, en la Tabla 3.11 se presenta la demanda por día de semana de las primeras 10 semanas del 2023. Posteriormente, en la Tabla 3.12, se muestra la conversión de la demanda a puntos porcentuales.

Tabla 3.11 Muestra de vuelos de las primeras diez semanas del 2023

Número de semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	TOTAL
1	134	131	125	135	133	127	131	916
2	122	121	115	118	103	121	124	824
3	124	121	115	116	105	106	112	799
4	109	111	100	113	111	116	118	778
5	121	120	107	115	113	115	115	806
6	117	114	106	109	110	110	116	782
7	113	114	104	112	109	114	117	783
8	116	113	110	111	113	114	113	790
9	116	115	106	114	110	106	115	782
10	108	112	106	113	108	104	112	763

Tabla 3.12 Distribución de vuelos de las primeras diez semanas del 2023

Número de semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo	TOTAL
1	14.6	14.3	13.6	14.7	14.5	13.9	14.3	916
2	14.8	14.7	14.0	14.3	12.5	14.7	15.0	824
3	15.5	15.1	14.4	14.5	13.1	13.3	14.0	799
4	14.0	14.3	12.9	14.5	14.3	14.9	15.2	778
5	15.0	14.9	13.3	14.3	14.0	14.3	14.3	806
6	15.0	14.6	13.6	13.9	14.1	14.1	14.8	782
7	14.4	14.6	13.3	14.3	13.9	14.6	14.9	783
8	14.7	14.3	13.9	14.1	14.3	14.4	14.3	790
9	14.8	14.7	13.6	14.6	14.1	13.6	14.7	782
10	14.2	14.7	13.9	14.8	14.2	13.6	14.7	763

Para la muestra extraída, se requiere determinar si la media de cada día es una medida representativa. Por ello, se aplicará la prueba de bondad de ajuste de Kolmogorov-Smirnov (KS), con el fin de validar si los datos de la muestra siguen una distribución normal. En ese sentido, se plantean las siguientes hipótesis para día de la semana:

H_0 : Los datos de muestra por día de semana (i) siguen una distribución normal.

H_1 : Los datos de muestra por día de semana (i) no siguen una distribución normal.

Siendo (i) el día de semana, desde el lunes al domingo.

Para hallar si la hipótesis nula se rechaza o no, se toma en cuenta el estadístico de contraste D, el cual representa la mayor diferencia entre la frecuencia acumulada observada y la teórica. En este caso, el análisis será evaluado en base al *p-value*, el cual está relacionado al estadístico antes mencionado.

$$p\text{-value} = P(D > D_{\text{obs}} / H_0 \text{ es cierta})$$

En consecuencia, para un nivel de significancia (α) de 0.05, si el *p-value* resulta ser mayor a este valor, implica que el estadístico observado *D* era esperable, y, por lo tanto, la hipótesis nula no se rechaza. Sin embargo, para un *p-value* menor a α , no se espera obtener el estadístico observado, por lo que se rechaza la hipótesis nula. A continuación, se sigue esta regla de decisión para un nivel de significancia (α):

$$\text{Si } p - \text{value} \geq \alpha \Rightarrow \text{Aceptar } H_0$$

$$\text{Si } p - \text{value} < \alpha \Rightarrow \text{Rechazar } H_0$$

Así mismo, cabe indicar que la prueba fue desarrollada en el software RStudio(Ver Anexo 16) y los resultados para cada día se resumen en la Tabla 3.13:

Tabla 3.13 Resultados de pruebas de hipótesis para cada día de la semana

Día de semana	p-value	Comentario
Lunes	0.84	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Martes	0.97	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Miércoles	0.81	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Jueves	0.50	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Viernes	0.66	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Sábado	0.14	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.
Domingo	0.63	Como $p\text{-value} > \alpha$ (0.05), no se rechaza la prueba de hipótesis.

Como se aprecia en la Tabla 3.13, todos los días de la semana pasaron la prueba y, en consecuencia, se va a considerar la media como medida representativa para calcular la distribución por día de semana

Por último, la distribución final de demanda por día de semana en puntos porcentuales se muestra en la Tabla 3.14.

Tabla 3.14 Resumen de distribución de demanda por día de semana

Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
14.4	14.3	13.8	14.4	14.4	14.1	14.7

Distribución por franja horaria

La cantidad de vuelos varía durante las veinticuatro horas del día, por ende, la cantidad de personal que los atiende también debe ser variable durante el día. En la Figura 3.24, se muestra la necesidad de personal en base a los vuelos de un día de muestra tomado del 2023. A partir de este día de muestra, se calcula la distribución porcentual de demanda por hora mediante la división de la demanda de cada hora entre el área total de demanda del día (Ver Figura 3.25).

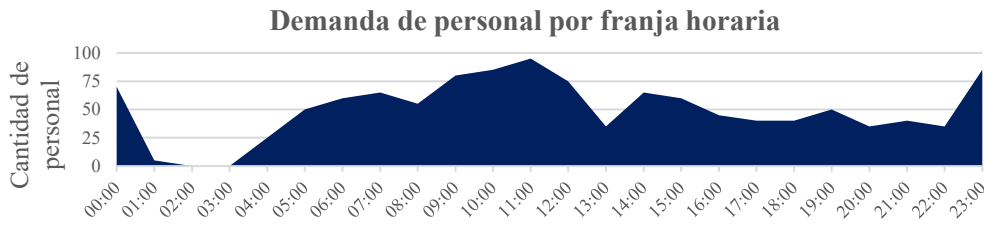


Figura 3.24 Demanda de personal por franja horaria de un día de muestra

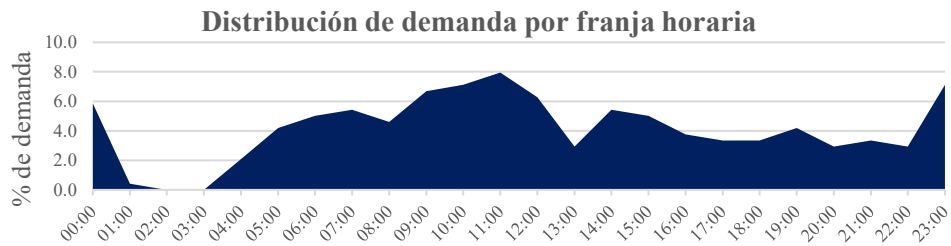


Figura 3.25 Distribución de demanda por franja horaria de un día de muestra

Para obtener una distribución porcentual representativa se consideró información diaria de los meses de enero a junio 2023, los cuales conforman 181 días. En el Anexo 17, se muestra la demanda diaria por hora de los primeros 10 días de enero 2023. A su vez, también se muestra como la información de demanda fue convertida en puntos porcentuales.

Con la muestra tomada, es necesario analizar si la media de cada hora es una medida representativa, para esto, se utilizará la prueba de bondad de ajuste de Kolmogorov-Smirnov (KS), con el objetivo de verificar si los datos de muestra siguen una distribución normal. De esta forma, se plantean las siguientes hipótesis para cada hora:

H_0 : Los datos de muestra de la hora (i) siguen una distribución normal.

H_1 : Los datos de muestra de la hora (i) no siguen una distribución normal.

Siendo (i) la hora del día, desde las 0 horas a las 23 horas.

Para decidir si aceptar o rechazar la hipótesis nula, se analiza el estadístico de contraste D, el cual es la mayor diferencia entre la frecuencia acumulada observada y la teórica. En este caso, la toma de decisión será mediante el uso del *p-value*, que está asociado al estadístico antes mencionado.

$$p\text{-value} = P(D > D_{\text{obs}} / H_0 \text{ es cierta})$$

Por ende, si el *p-value* es elevado, el valor del estadístico observado D era esperable, con lo que se concluye que no hay razones para descartar la hipótesis nula. En caso contrario, para un *p-value* bajo, es improbable que se obtenga el estadístico observado, con lo que se rechaza la hipótesis nula. Para un nivel de significancia (α), la regla de decisión es la siguiente:

$$\text{Si } p\text{-value} \geq \alpha \Rightarrow \text{Aceptar } H_0$$

$$\text{Si } p - \text{value} < \alpha \Rightarrow \text{Rechazar } H_0$$

Se calculó los p-value de cada conjunto de datos mediante el uso de RStudio, estos se muestran en el Anexo 18.

Para los casos que no pasaron la prueba, se detalla las consideraciones:

i. Hora 1

Debido a la distribución de frecuencias de la muestra (Ver Anexo 19), se considera que la mediana sería una medida más representativa para el conjunto de datos tomado debido a que esta es menos sensible a los valores anómalos observados.

ii. Hora 2

Al observar la distribución de frecuencias de la muestra (Anexo 19), se define que la moda sería la medida más representativa para el conjunto de datos tomado, en este caso, el valor de 0.

iii. Hora 3

En el Anexo 19, se muestra un acumulado mayor de frecuencias en el rango 0.0-0.2, por lo que, se considera que la moda sería la medida más representativa para el conjunto de datos tomado, en este caso, el valor de 0.

Finalmente, como distribución final de demanda por hora en puntos porcentuales se tiene lo siguiente:

Tabla 3.15 Resumen de distribución de demanda por hora del día

00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
5.9	0.8	0.0	0.0	2.5	4.8	4.5	3.9	5.4	6.7	6.9	7.0	6.3	3.5	4.8	4.4	3.8	2.8	4.2	4.2	3.0	3.4	4.5	6.7

La información antes descrita servirá como *input* del modelo de optimización, para fijar la demanda con la que se van a iterar las corridas en el programa AMPL, y así, hallar la cantidad de grupos requeridos de la oferta. En ese sentido, el desarrollo de este segmento continuará en el siguiente apartado.

3.3 Diseño de un modelo de programación lineal para la optimización de horarios y personal

Esta segunda propuesta de solución abarca el desarrollo del modelo de programación de la capacidad del recurso humano basado en la demanda previamente proyectada y su distribución por franja horaria, de modo que la cantidad de turnos para atender la operación de los vuelos sea la óptima considerando minimizar los costos de los horarios.

3.3.1 Preparación de los datos

La demanda se construye primordialmente a partir de la cantidad de vuelos semanal que se proyecta en el modelo predictivo SARIMA; dado que, a partir de estos se generan los servicios o tareas que el personal debe atender.

Previamente a la transformación de los vuelos en demanda por rango horario, es necesario convertir los vuelos en atenciones de servicio. Las atenciones, a diferencia de los vuelos, contemplan también las recepciones de aeronaves, no solo las salidas. Para este caso, se realiza la conversión al multiplicar la cantidad de vuelos por un factor de 1.5219, basado en data histórica del sector.

Esta cantidad de atenciones se distribuye por día de semana y por franja horaria, según la Tabla 3.14 y Tabla 3.15. Se obtiene como resultado el requerimiento de grupos de personal por hora y día de semana, los siete días de la semana.

En el Anexo 20, se muestra el detalle de la cantidad de grupos de personas segregado por franja horaria y día de semana, considerando que los datos fueron redondeados al mayor superior.

3.3.2 Planteamiento del modelo de programación de la capacidad del recurso humano

En este apartado se desarrollará la descripción acerca de las condiciones actuales que ha establecido la empresa para la gestión de horarios del personal operativo. Posteriormente, en base a estos lineamientos, se realizará el planteamiento del problema a resolver.

En primer lugar, se enlistan las condiciones determinadas por la organización:

- Turnos: Los turnos que se han definido abarcan una cobertura de 24 horas puesto que el aeropuerto de Lima está habilitado para operar en todo el día; sin embargo, la empresa ha establecido no asignar turnos que inicien entre 23:00 h a 02:00 am puesto que, optan por cubrir la demanda de este rango con turnos que inicien más temprano ya que históricamente, se ha identificado una menor concentración de vuelos de acuerdo con lo analizado en la sección anterior. Así mismo, todos los turnos comprenden una duración de 9 horas considerando 8 horas laborales más 1 hora de refrigerio y son asignados por cada hora. A continuación, en la Tabla 3.16 se muestran todos los turnos por hora que pueden ser asignados:

Tabla 3.16 Turnos de trabajo

Turno	Inicio	Fin
1	03:00	12:00
2	04:00	13:00
3	05:00	14:00
4	06:00	15:00
5	07:00	16:00
6	08:00	17:00
7	09:00	18:00
8	10:00	19:00
9	11:00	20:00
10	12:00	21:00
11	13:00	22:00
12	14:00	23:00
13	15:00	00:00
14	16:00	01:00
15	17:00	02:00
16	18:00	03:00
17	19:00	04:00
18	20:00	05:00
19	21:00	06:00
20	22:00	07:00

- Rotación: La política de rotación para los turnos laborales se ha estandarizado en 6x1, es decir, una jornada de 6 días de trabajo continuos seguido de 1 día de descanso. En ese sentido, no se tienen restricciones en la asignación del día de descanso, es decir, cada turno tiene como posible descanso todos los días de la semana como se muestra en el ejemplo de la Tabla 3.17.

Tabla 3.17 Ejemplo de turno con días de descanso

Turno	Inicio	Fin	Descanso
1	03:00	12:00	Lunes
2	03:00	12:00	Martes
3	03:00	12:00	Miércoles
4	03:00	12:00	Jueves
5	03:00	12:00	Viernes
6	03:00	12:00	Sábado
7	03:00	12:00	Domingo

- Refrigerio: Cada horario posee dos opciones de refrigerio que inician entre 2 a 7 horas posterior a su hora de ingreso. Así mismo, las horas de refrigerio se han determinado en base a la disponibilidad del comedor. En la Tabla 3.18, se muestra un ejemplo del turno con ingreso a las 14:00 h con dos posibles horas de refrigerio.

Tabla 3.18 Ejemplo de turnos con refrigerio

Turno	Inicio	Fin	Inicio Refrigerio	Fin Refrigerio
1	14:00	23:00	18:00	19:00
2	14:00	23:00	19:00	20:00

En resumen, en el Anexo 21 se presentan todos los turnos posibles a ser asignados con sus respectivos refrigerios y días de descanso

- Costos por turno: Los costos que implican los turnos asignados se aplican por los siguientes conceptos:
 - o Nocturnidad: De acuerdo con el Ministerio de Trabajo y Promoción del Empleo (2023), en caso los trabajadores se encuentren de turno durante el rango de 22:00 a 06:00 h, se considera que tienen un horario nocturno y por ese motivo, la remuneración mínima que deben percibir es de S/1,383.75. En caso de la empresa, se tienen los cargos de auxiliar y operador 1 los que aplican para este tipo de pago nocturno, puesto que al cargo de líder le corresponde un sueldo mayor a los S/1,383.75. En ese sentido, el auxiliar gana S/1.49 por hora nocturna, mientras que el operador 1 gana S/1.10 respectivamente.
 - o Movilidad: Para aquellos turnos que poseen una salida o llegada entre las 00:00 a 04:00 am, se considera traslado de personal con movilidad, el cual tiene un costo diario de S/12.67 por persona.

En el Anexo 22 se presenta el detalle de los costos por los conceptos de nocturnidad y movilidad agrupados por turnos y horarios

Para definir el modelo de programación de la capacidad del recurso humano, es necesario precisar los siguientes componentes:

- Variables de decisión (x_i): Para este caso, se define como la cantidad de grupos de personal por cada turno establecido bajo las condiciones de trabajo antes mencionadas, con el objetivo de cumplir con la demanda.

Por lo que, x_i el número de grupos de personas por turno i ($i = 1, 2, \dots, 280$).

Asimismo, se define que las variables deben ser enteras y positivas (34).

$$x_i \geq 0 \quad (34)$$

- Restricciones: Se establecen de acuerdo con la demanda y los turnos establecidos. Cada turno solo cubre cierto rango temporal, por lo que, cada franja horaria (del 1 al 24) debe ser cubierta por la suma de ciertos turnos definidos. Para ejemplificar este caso, se arma una restricción de

ejemplo: en el rango de las 5:00 del día lunes de la semana 1, se requiere 10 personas (ver Anexo 20) y, para cubrir específicamente esa hora se tienen los turnos del 1 al 42 y del 253 al 280 (ver Anexo 21); la restricción se expresaría como:

$$\sum_{i=1}^{42} x_i * T_{(5,1,i)} + \sum_{i=253}^{280} x_i * T_{(5,1,i)} \geq 10 \quad (35)$$

Donde x_i : Cantidad de grupos de personas en el turno i

$T_{(5,1,i)}$: Factor de asignación de turno i , con cobertura a las 5:00hh, el día 1 (lunes).

Asimismo, como consideración adicional se tiene que, los turnos también deben cubrir el ausentismo, este representa el 8% del personal en planilla de este proceso según información histórica. Por lo que, cada dato de demanda debe ser multiplicado por un factor de 1.08.

En general, se tendría el siguiente conjunto de restricciones:

$$\sum_{h=1}^{24} \sum_{d=1}^7 \sum_{i=1}^{280} x_i * T_{(h,d,i)} \geq 1.08 * \sum_{h=1}^{24} \sum_{d=1}^7 D_{(h,d)} \quad (36)$$

Donde x_i : Cantidad de grupos de personas en el turno i ,

$T_{(h,d,i)}$: Factor de asignación del turno i , en la franja horaria h , en el día d ,

$D_{(h,d)}$: Demanda de grupos de personal en la franja horaria h , en el día d .

- Función objetivo: En esta problemática se buscará reducir los costos por personal, sin dejar de cumplir los requerimientos de la demanda antes descritos.

$$\text{Minimizar } f(x) = \sum_{i=1}^{280} c_i x_i \quad (37)$$

Donde x_i : Cantidad de grupos de personas en el turno i

c_i : Costo por turno i .

Finalmente, los componentes se detallan en el programa AMPL IDE, para dar una solución óptima al modelo de programación de la capacidad del recurso humano. En la Figura 3.26, se define el planteamiento del modelo; y, los datos ingresados de demanda, costos y turnos se muestran en el Anexo 22 y Anexo 23.

Los resultados de las soluciones óptimas de las cuatro semanas se detallan en el Capítulo 4.

```

set turnos :=1..280;

set franja:=1..24;
set dia:=1..7;
set diaxfranja:={franja,dia};
set matriz:={franja,dia,turnos};

param Turnos{matriz};
param Demanda{diaxfranja};
param Costo{turnos};

var X {i in turnos}>=0 integer;

minimize costos:
sum {i in turnos} X[i]*Costo[i];

subject to asignacion{h in franja, d in dia}:
sum{i in turnos} Turnos[h,d,i]*X[i]>=1.08*Demanda[h,d];

```

Figura 3.26 Planteamiento del modelo de programación de la capacidad del recurso humano

3.4 Diseño de escenarios de simulación de la situación actual y la propuesta de mejora

Esta tercera propuesta abarca el diseño del modelo de simulación basado en data real de la demanda del mes de setiembre 2023, en el que se busca realizar comparaciones de dos escenarios, uno considerando los recursos de la situación actual y otro con los recursos de la propuesta de la programación optimizada del apartado anterior.

3.4.1 Componentes del modelo

El modelo de simulación a desarrollar abarca los siguientes componentes:

- a) Entidades: La aeronave es el objeto de estudio del presente modelo puesto que cada vez que llega a una PEA, esta debe atenderse en tierra por la célula de trabajo de la empresa.
- b) Atributos: Se consideran los siguientes dos atributos en el modelo:
 - Tipo Aeronave: Describe el tipo de aeronave según su tamaño:
 - Tipo Aeronave==1: Si la aeronave que ingresa a la PEA para ser atendida es NB
 - Tipo Aeronave==2: Si la aeronave que ingresa a la PEA para ser atendida es WB
 - Tipo Destino: Describe el tipo de aeronave según el destino de salida:
 - Tipo Destino==1: Si la aeronave está asignada a un vuelo doméstico
 - Tipo Destino==2: Si la aeronave está asignada a un vuelo internacional
- c) Recursos: La célula de trabajo es el recurso para este modelo y está conformada por 5 personas (1 líder, 1 operador y 3 auxiliares)
- d) Horarios: La programación de turnos está sujeta a la distribución de la demanda por franja horaria, la cual se desarrolló en el apartado anterior en la propuesta de optimización. En ese

sentido, se estará empleando los turnos actuales como primer escenario, y, como segundo escenario a plantear, los turnos de la propuesta optimizada.

- e) Colas: Son 3 colas que se presentan en el modelo de acuerdo con el tipo de aeronave y destino.
- f) Contadores: Son 3 contadores que se presentan en el modelo de acuerdo con el tipo de aeronave y destino.
- g) Indicadores de desempeño: Los indicadores a tomar en cuenta para el modelo son los siguientes.
 - Utilización del recurso: Se mide a través de cuánto representa las horas productivas del personal con respecto a las horas totales que se encuentran en turno con frecuencia semanal.

$$\text{Utilización del recurso} = \frac{\text{Horas productivas}}{\text{Horas totales}}$$

- Productividad: Se determina considerando las horas de la demanda atendida con respecto a las horas totales del personal programado por día.

$$\text{Productividad} = \frac{\text{Volumen de atenciones}}{\text{Cantidad de grupos programados}}$$

- Horas muertas: Representan las horas donde el personal se encuentra disponible sin asignación de vuelo con frecuencia semanal.

$$\text{Horas muertas} = \text{Horas totales} - \text{Horas productivas}$$

- % Demoras: # demoras Son aquellas atenciones que finalizan posterior a la hora inicial de salida y se quedan en cola dentro del sistema a simular con frecuencia semanal.

$$\% \text{ Demoras} = \frac{\text{Volumen de demoras}}{\text{Volumen de vuelos}}$$

3.4.2 Descripción del modelo

A continuación, se definen los supuestos del modelo a plantear:

- El aeropuerto opera durante las 24 horas del día.
- Toda aeronave que ingresa a una PEA debe ser atendida por una célula de trabajo.
- La célula de trabajo se mantiene constante con la misma cantidad de 5 personas por cada atención.
- Las aeronaves no siempre cumplen el itinerario de la aerolínea debido a factores internos o externos que puedan impactar en lo planificado.
- Si el recurso se encuentra ocupado en una atención, no es posible que inicie una nueva hasta que termine con la actual.

- Se considera que, durante las horas muertas de atención, el personal puede movilizarse para ir a los servicios higiénicos o tomar un descanso.
- El tiempo de refrigerio ya está considerado en la cantidad de personal disponible de los turnos de trabajo tanto de la situación actual como en la propuesta optimizada.
- El ratio de supervisores que se debe mantener en la operación por grupo de Rampa es de 1 a 7.
- Por protocolo de seguridad aeroportuario, siempre se debe tener un supervisor en turno para mantener comunicación las 24 horas del día.
- Los tiempos de atención están sujetos a variación del tipo de aeronave y tipo de destino.
- Si se genera una demora en la salida de la aeronave por responsabilidad de la empresa, se aplica una penalidad por afectar el nivel de servicio de puntualidad.

3.4.3 Recolección de los datos

En esta sección, se extrae la data real de demanda del periodo de setiembre 2023, el cual se consideró en las propuestas anteriores de Machine Learning y Optimización. La data contiene campos de fecha y hora de arribo y salida de la atención tanto real como programada, destino y modelo de fuselaje, los cuales serán relevantes para reflejar las demoras y adelantos de aeronaves en el sistema a simular con el objetivo de analizar la demanda bajo un carácter probabilístico.

3.4.4 Análisis de los datos

En este acápite, se presenta el desarrollo de la clasificación de los datos, así como la definición del conjunto de variables aleatorias a tomar en cuenta. Seguidamente, se realiza el cálculo para el tamaño de la población y la muestra. Finalmente, las muestras se ingresan al software Input Analyzer para la aplicación de las pruebas de bondad de ajuste:

a) Clasificación de los datos

Previamente a la definición de variables, se realiza la clasificación de los datos para el sistema a simular acerca de la operación de vuelos en el aeropuerto de Lima.

- Tiempo de atención de la aeronave: Este tiempo es el requerido donde el grupo de trabajo realiza la atención en tierra desde el parqueo en la PEA hasta la salida del *push back* previo a su despegue. Así mismo, se observa que esta variable está relacionada con el tipo de fuselaje y el tipo de atención. En ese sentido, se formula una hipótesis acerca de esta relación y, para comprobarlo, se ha empleado el software MINITAB como se muestra en la Figura 3.27.

Información del factor			
Factor	Tipo	Niveles	Valores
Tipo atención	Fijo	3	O; R; T
Fuselaje	Fijo	2	F1; F2

Análisis de Varianza					
Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Tipo atención	2	964363	482181	1198.50	0.000
Fuselaje	1	380840	380840	946.61	0.000
Tipo atención*Fuselaje	2	75729	37864	94.12	0.000
Error	5822	2342307	402		
Total	5827	4708233			

Figura 3.27 Análisis de varianza de dos factores para determinar la dependencia entre el tiempo de atención de aeronaves con referencia al fuselaje y tipo de atención

Como se aprecia en la Figura 3.27, los valores de p-Value tanto para el factor de día de semana como en el de franja horaria son menores a 0.05, por lo tanto, se afirma que existe una dependencia entre el tiempo de atención de aeronaves con respecto al fuselaje y tipo de atención. En ese sentido, se le asigna el tipo de fuselaje por variable como se observa en la Tabla 3.19 y, se agrupan el tipo de fuselaje con el tipo de atención en una sola variable como se muestra en la Tabla 3.20.

Tabla 3.19 Tipos de fuselaje

Variable	Fuselaje
F1	NB
F2	WB

Tabla 3.20 Definición de variables por tipo de fuselaje y tipo de atención

Variable	Descripción
F1-O	Tiempo de atención de una salida para una aeronave de fuselaje NB
F1-R	Tiempo de atención de una recepción para una aeronave de fuselaje NB
F1-T	Tiempo de atención de un tránsito para una aeronave de fuselaje NB
F2-O	Tiempo de atención de una salida para una aeronave de fuselaje WB
F2-R	Tiempo de atención de una recepción para una aeronave de fuselaje WB
F2-T	Tiempo de atención de un tránsito para una aeronave de fuselaje WB

- Tiempo de llegadas entre aeronaves: En base a lo observado en la data de atenciones, el tiempo de llegadas entre aeronaves está relacionado con los días de la semana y la franja horaria. A partir de ello, se elabora la hipótesis sobre esta relación y, con el fin de comprobarla, se ha empleado el software MINITAB.

Información del factor

Factor	Tipo	Niveles	Valores
Día de semana	Fijo	7	1, 2, 3, 4, 5, 6, 7
Hora ATA	Fijo	24	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Día de semana	6	2992	498.6	4.64	0.000
Hora ATA	23	1610809	70035.2	651.68	0.000
Día de semana*Hora ATA	138	31816	230.5	2.15	0.000
Error	37758	4057813	107.5		
Total	37925	5707719			

Figura 3.28 Análisis de varianza de dos factores para determinar la dependencia entre el tiempo de llegadas entre aeronaves con referencia al día de semana y franja horaria

Como se observa en la Figura 3.28, los valores de p-Value tanto para el factor de día de semana como en el de franja horaria son menores a 0.05, por lo tanto, se afirma que existe una dependencia entre el tiempo entre llegadas de aeronaves con respecto al día de semana y franja horaria.

A continuación, se procede a agrupar la data de acuerdo con los factores de día de semana y franja horaria.

Con respecto al primer factor, se realizó el cálculo de demanda promedio por día tal como se presenta en la Tabla 3.21.

Tabla 3.21 Proporción de demanda Setiembre 2023 por día de semana

Día de semana	% Demanda Setiembre 2023
Lunes	14.2%
Martes	14.1%
Miércoles	14.1%
Jueves	14.1%
Viernes	14.4%
Sábado	14.4%
Domingo	14.2%

Como se puede observar en la Tabla 3.21, las proporciones de demanda son de alrededor de 14% entre cada día, y el criterio a considerar será agrupar los días de mayor proporción; es decir, viernes, sábado, domingo y lunes en un primer conjunto. Por lo tanto, el segundo conjunto será el martes, miércoles y jueves ya que representan la menor proporción como se puede observar en la **¡Error! La autoreferencia al marcador no es válida..**

Tabla 3.22 Grupos por días de semana

Grupo	Día de semana
D1	domingo, lunes, viernes y sábado
D2	martes, miércoles y jueves

Así mismo, para el factor de franja horaria, se calcularon las proporciones de demanda en base a cada franja en intervalos de 1 hora como se muestran a continuación en la Tabla 3.23.

Tabla 3.23 Proporción de demanda Setiembre 2023 por franja horaria

Franja Horaria	% Demanda Setiembre 2023
00:00	18.5%
01:00	17.2%
02:00	12.8%
03:00	4.2%
04:00	2.4%
05:00	2.1%
06:00	2.4%
07:00	2.3%
08:00	1.7%
09:00	2.0%
10:00	1.3%
11:00	2.0%
12:00	2.5%
13:00	3.4%
14:00	2.5%
15:00	2.6%
16:00	3.4%
17:00	3.5%
18:00	2.5%
19:00	2.9%
20:00	2.3%
21:00	1.8%
22:00	1.6%
23:00	2.2%

Como se puede observar en la Tabla 3.23, se tomará el mismo criterio de agrupación como en el factor anterior, por lo que los conjuntos quedarían de la siguiente manera:

Tabla 3.24 Grupos por franja horaria

Grupo	Franja horaria
R1	00:00 - 02:00 h
R2	03:00 - 07:00 h
R3	08:00 - 12:00 h
R4	13:00 - 17:00 h
R5	18:00 - 21:00 h
R6	22:00 - 23:00 h

Finalmente, en el Anexo 24, se presentan la combinación de ambos factores y las variables que se emplearán para el análisis de datos.

b) Tamaño de la población

Como se mencionó en líneas anteriores, el periodo de análisis a simular es el de setiembre 2023, el cual abarca un total de 5829 atenciones, siendo este un volumen menor a 200000 atenciones. En ese sentido, la población se define como finita y, como se observa en la Tabla 3.25 y la Tabla 3.26, las poblaciones de cada variable también cumplen con esta condición.

Tabla 3.25 Tamaño de población de las variables para el tiempo de atención de aeronaves

Variable	Población
F1-O	1580
F1-R	1586
F1-T	2122
F2-O	222
F2-R	222
F2-T	96
Total	5828

Tabla 3.26 Tamaño de población de las variables para el tiempo entre llegada de aeronaves

Variable	Población
R1-D1	62
R1-D2	40
R2-D1	724
R2-D2	472
R3-D1	1008
R3-D2	644
R4-D1	604
R4-D2	433
R5-D1	625
R5-D2	432
R6-D1	469
R6-D2	315
Total	5828

c) Muestreo

Una vez definida la población de la data, se procede a realizar el cálculo para los tamaños de muestra por la población de las diez variables aleatorias establecidas. Para ello, se considera lo siguiente:

- El valor de Z se calcula en aproximadamente 1.96
- El valor para el nivel de significancia α se fija en 0.05
- El error porcentual se define en 5%

- El valor de la proporción p se fija en 50%

Al realizar el cálculo, se obtuvieron los siguientes tamaños de muestra de la Tabla 3.27 y Tabla 3.28.

Tabla 3.27 Tamaño de muestra de las variables de tiempo de atención de la aeronave

Variable	Muestra
F1-O	309
F1-R	309
F1-T	325
F2-O	141
F2-R	141
F2-T	77

Tabla 3.28 Tamaño de muestra de las variables de tiempo entre llegada de aeronaves

Variable	Muestra
R1-D1	54
R1-D2	36
R2-D1	251
R2-D2	212
R3-D1	278
R3-D2	241
R4-D1	235
R4-D2	204
R5-D1	238
R5-D2	204
R6-D1	211
R6-D2	173

d) Pruebas de bondad de ajuste

Una vez que se han calculado los tamaños de muestra de las variables, se procedió a realizar las pruebas de bondad de ajuste con el fin de determinar las distribuciones de cada una. Para ello, se empleó el software Input Analyzer y se siguieron los criterios de selección de las pruebas. En ese sentido, se identificó que todas las variables son de tipo continua, por lo que aquellas con un tamaño de muestra menor a 90, se les aplicó la prueba Kolmogorov Smirnov ; de lo contrario, se aplicaron tanto la prueba Chi Cuadrado como la Kolmogorov Smirnov, comprobando que el p value sea mayor a 0.05 en cualquiera de los casos. Así mismo, para cada muestra ingresada, se aplicó la opción Fit All para encontrar la distribución de mejor ajuste y la opción Fit All Summary para observar el error cuadrado por cada distribución.

En el Anexo 25 se detallará los resultados de las pruebas que se realizaron para cada variable.

En resumen, en el Anexo 26 se presenta el consolidado de las variables mencionadas junto a la expresión de la distribución obtenida del Input Analyzer.

3.4.5 Modelamiento del sistema de simulación

Al terminar de hallar las distribuciones que mejor ajustan a las variables aleatorias, se realizó el modelamiento del sistema a simular en el software Arena Simulation, donde se definió a la aeronave como única entidad, los atributos llamados “TipoAeronave” y “TipoDestino” y el recurso configurado como “Grupos de rampa”. Así mismo, se armó el flujo del sistema a simular de acuerdo con el Anexo 27.

Además de los bloques que forman el flujo del proceso a simular en el Anexo 27, se ha definido el elemento “Schedule” como se muestra en la Figura 3.29.

	Name	Type	Time Units	Scale Factor	File Name	Recordset	Comment
1 ▶	Horario1	Capacity	Hours	1.0	File 2	Horario actual	
2	Horario2	Capacity	Hours	1.0	File 4	Horario	

Figura 3.29 Definición del elemento “Schedule” en el modelo de simulación para la atención de aeronaves

Como se observa en la Figura 3.29, se ingresaron dos registros de tipo “Schedule”, los cuales corresponden a los horarios y grupos para los dos escenarios a simular. Es decir, el primer escenario acerca del input “Horario1”, se modela en base a la situación actual de la empresa; mientras que el segundo que tiene por input “Horario2”, se consideró la cantidad de grupos y horarios obtenidos del software AMPL de la propuesta mejorada en el segmento de Optimización.

Posteriormente, se procedió a asignar los recursos del flujo a simular agregando los inputs “Grupos de rampa 2”, al cual se le asignó el *Schedule* “Horario1” para el primer escenario y “Grupos de rampa”, donde se asignó el *Schedule* “Horario2” tal como se muestra en la Figura 3.30:

	Name	Type	Schedule Name	Schedule Rule	Busy / Hour	Idle / Hour	Per Use	StateSet Name	Failures	Report Statistics	Comment
1 ▶	Grupos de rampa actual	Based on Schedule	Horario1	Wait	0.0	0.0	0.0		0 rows	<input checked="" type="checkbox"/>	
2	Grupos de rampa	Based on Schedule	Horario2	Wait	0.0	0.0	0.0		0 rows	<input checked="" type="checkbox"/>	

Figura 3.30 Definición del elemento “Resources” en el modelo de simulación para la atención de aeronaves

En el Anexo 28, se ilustra el sistema de simulación de la atención de una aeronave en base a los bloques descritos en el Anexo 27.

Así mismo, debido a que el presente sistema a simular es de tipo terminal puesto que se ha delimitado evaluar la demanda de atenciones la primera semana de setiembre del 2023, se ha realizado el cálculo de réplicas en base a los siguientes pasos:

En primer lugar, se ha elaborado un muestreo inicial de 15 réplicas con los datos que se muestran en el Anexo 29.

Se procede a calcular el error de esta muestra preliminar con la siguiente fórmula:

$$e = \frac{\sigma t}{\sqrt{n}} \quad (38)$$

Donde:

e: error de la muestra

σ : desviación de la muestra

t: estadístico t

n: tamaño de muestra

A continuación, en la Tabla 3.29 se muestran los valores que se han considerado para el cálculo:

Tabla 3.29 Cálculo del error de la muestra preliminar

Dato	Valor
n	15
σ	30.91
t	2.1447867
e	17.1171357

De la Tabla 3.29, se observa que el error para la muestra de tamaño 15, es de 17 atenciones a la semana.

Seguidamente, en base a los valores hallados como el estadístico t , la desviación estándar σ y considerando un error de 7 atenciones por semana, se realiza el cálculo para el tamaño de muestra equivalente al número de réplicas que se requieren obtener para la corrida en la simulación. A continuación, se presenta el cálculo en la Tabla 3.30.

Tabla 3.30 Cálculo del número de réplicas para la simulación

Dato	Valor
σ	30.91
t	2.1447867
e	7
n	89.7

Finalmente, se obtiene que la cantidad de réplicas a considerar para la corrida es de 90.

En suma, en el presente capítulo se ha formulado como primera propuesta el uso de técnicas de Machine Learning en RStudio empleando la Regresión Lineal, ARIMA y SARIMA con el fin de identificar el más adecuado para el pronóstico de la demanda. Seguidamente, se planteó el modelo de programación lineal en AMPL para definir el *capacity* de personal óptimo. Por último, se desarrolló el sistema de simulación basado en el flujo de demanda de vuelos en Arena considerando adelantos y retrasos. En ese sentido, en el capítulo siguiente se presentarán los resultados obtenidos de las ejecuciones de cada modelo junto a su análisis comparativo con respecto a la situación actual correspondiente.

CAPITULO 4. RESULTADOS

En este capítulo se desarrolla la interpretación de los resultados de los modelos planteados previamente y se divide en tres segmentos de la misma manera que el capítulo anterior acorde a los siguientes temas: Aprendizaje automatizado, Optimización y Simulación.

4.1 Predicción de demanda mediante el aprendizaje automático

En este apartado, se procede a determinar el modelo óptimo SARIMA y realizar el testeo con la data del 20% restante para contrastar los modelos de regresión lineal, ARIMA y SARIMA.

Pronóstico 1: Modelo de regresión lineal

En la Figura 4.1, se presenta el pronóstico calculado del modelo de regresión lineal en el software RStudio.

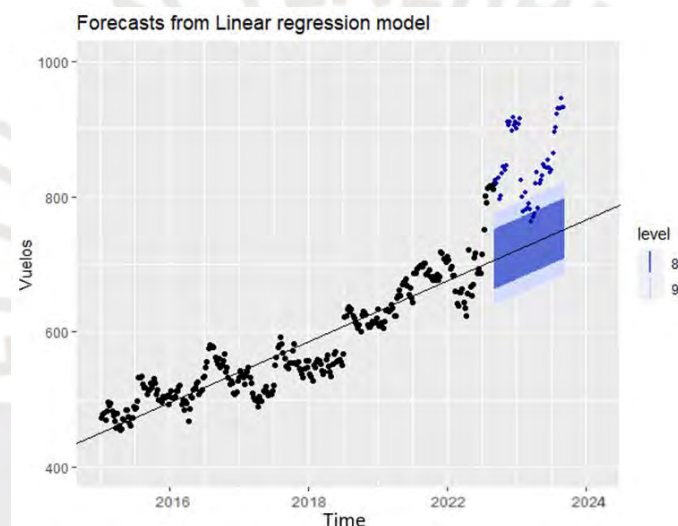


Figura 4.1 Pronóstico del modelo de regresión lineal

Como se puede apreciar, los valores de dispersión en color azul son los datos reales de demanda que pertenecen al *test* y la recta de regresión que se encuentra entre el área sombreada azul y celeste, contiene los valores predecidos del modelo. En ese sentido, se observa que los valores reales están alejados de los valores predecidos y si bien es cierto, el modelo refleja la tendencia de crecimiento de la demanda, pero no el patrón de comportamiento.

Pronóstico 2: Modelo ARIMA(2,1,2)

En la Figura 4.2, se presenta el pronóstico calculado del modelo ARIMA(1,1,2) en el software RStudio

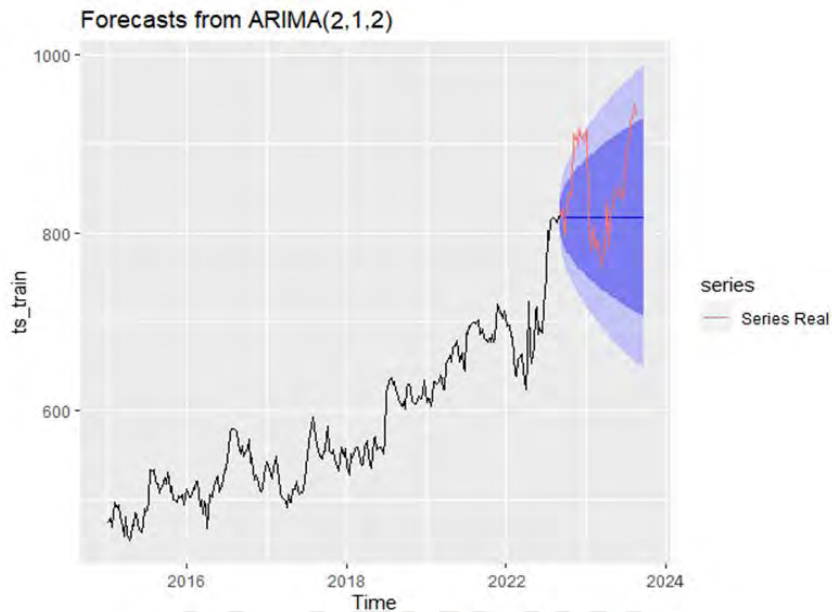


Figura 4.2 Pronóstico del modelo ARIMA(2,1,2)

Como se puede observar, la línea azul pertenece a los valores calculados del pronóstico, mientras que la línea roja muestra los valores históricos de la demanda. De esta manera, se aprecia que los valores pronosticados no varían en todo el rango de las fechas de testeo a diferencia de lo que se esperaba en los valores reales y, por lo tanto, no captura la tendencia de la serie ni el patrón de comportamiento.

Pronóstico 3: Modelo SARIMA

A. Selección de modelo óptimo

Con el fin de seleccionar el modelo óptimo SARIMA para la serie de tiempo, se utiliza las funciones de RStudio, en las que se iteran diferentes modelos hasta encontrar el de mejor ajuste. Luego de las iteraciones respectivas, se obtiene como modelo de mejor ajuste: SARIMA(3,1,0)(0,1,1)₅₂, el detalle se encuentra en la Figura 4.3.

```
Series: ts_train
ARIMA(3,1,0)(0,1,1)[52]

Coefficients:
      ar1      ar2      ar3      sma1
    -0.1960  -0.2408  -0.1818  -0.4738
s.e.   0.0636   0.0625   0.0638   0.0926

sigma^2 = 170.8: log likelihood = -970.07
AIC=1950.14  AICc=1950.39  BIC=1967.58
```

Figura 4.3 Resumen del modelo SARIMA(3,1,0)(0,1,1)₅₂

- Respecto al primer coeficiente para la componente autorregresiva (ar1), el valor es de -0.196, lo que significa que existe una dependencia lineal con peso negativo entre la observación presente y la inmediata anterior. Asimismo, respecto al segundo y tercer coeficiente de la componente autorregresiva (ar2 y ar3), sus valores negativos también suponen que existe una dependencia

lineal con peso negativo entre la observación presente y las observaciones de dos y tres periodos pasados.

- Con respecto al coeficiente para la componente de media móvil estacional (sma1), el valor calculado es de -0.4738, lo que significa que existe una dependencia lineal con peso negativo entre la observación presente y el error de 52 periodos pasados.
- Al contrastarlo con el modelo base mostrado en el capítulo 3: SARIMA(3,1,0)(1,1,0)₅₂, se observa que el AIC del modelo base 1958.18 es ligeramente mayor que el AIC del nuevo modelo 1950.14, por lo que se comprueba que el modelo óptimo sería ARIMA(3,1,0)(0,1,1)₅₂.

B. Validación del Modelo

En la Figura 4.4 se muestran los residuos del modelo ajustado, así como la prueba de Ljung-Box para la serie en la Figura 4.5. A pesar de que se observan algunos picos significativos en el ACF, el modelo pasa la prueba de Ljung-Box, al tener un p-value mayor al nivel de significancia (α) 0.05. Mediante esta última prueba, se demuestra que los residuos no tienen autocorrelación.

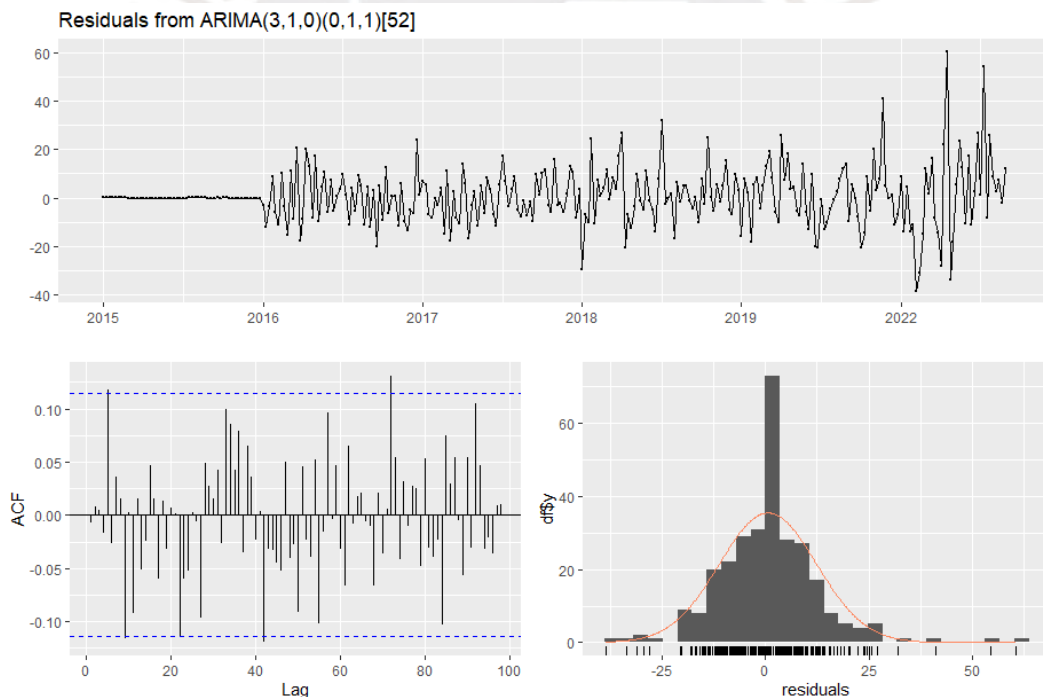


Figura 4.4 Residuos del modelo ajustado SARIMA(3,1,0)(0,1,1)₅₂ para la serie de tiempo

Ljung-Box test

```
data: Residuals from ARIMA(3,1,0)(0,1,1)[52]
Q* = 59.905, df = 55, p-value = 0.3024

Model df: 4. Total lags used: 59
```

Figura 4.5 Prueba de Ljung-Box

Se muestra en la Figura 4.6, el gráfico Q-Q de los residuos del modelo $ARIMA(3,1,0)(0,1,1)_{52}$. Se observa que en su mayoría los puntos están cerca de la línea recta, aunque se desvía ligeramente en los extremos. Gráficamente, no se podría rechazar la normalidad de los residuos.

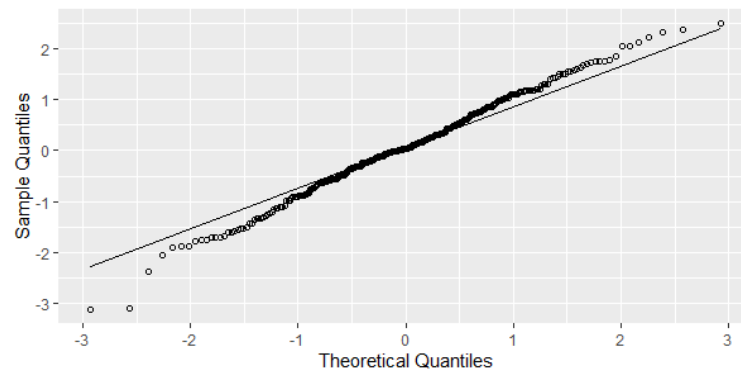


Figura 4.6 Gráfico Q para evaluar normalidad

C. Forecast

A partir del modelo planteado, se genera un primer pronóstico solo en base a los datos de entrenamiento. En la Figura 4.7, la línea azul representa los valores pronosticados por el modelo y la línea roja los valores reales, que son los datos de testeo.

Se aprecia que los primeros valores reales escapan del pronóstico del modelo y puede deberse a un factor atípico. Posterior a ello, el pronóstico se encuentra bastante cercano a los valores reales y parece capturar un comportamiento estacional.

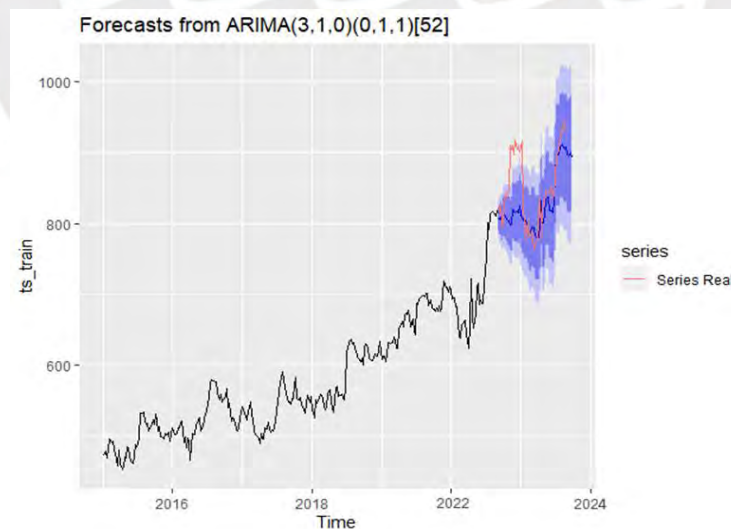


Figura 4.7 Pronóstico del modelo $SARIMA(3,1,0) (0,1,1)_{52}$

Comparativo de métricas de modelos

Con el objetivo de seleccionar el modelo final para el pronóstico de vuelos, a continuación, se muestra un resumen por cada modelo sobre las métricas obtenidas de AIC, MAE y RMSE basadas en la data de testeo a tomar en cuenta en la Tabla 4.1:

Tabla 4.1 Matriz de métricas de los modelos

Modelo	AIC	MAE	RMSE
1. Regresión Lineal	2924.89	120.5608	130.8712
2. ARIMA (2,1,2)	2365.12	49.087521	61.83335
3. ARIMA (3,1,0)(0,1,1) ₅₂	1950.14	33.746839	46.81049

- Con respecto al AIC, se puede observar que el modelo de menor valor es el ARIMA (3,1,0)(0,1,1)₅₂ con 1950.14, el cual indica que presenta un mejor ajuste a los valores de la serie de tiempo.
- En cuanto al MAE, el modelo que muestra el menor error es el ARIMA (3,1,0)(0,1,1)₅₂ con un valor de 33.746839; es decir, que en promedio los valores pronosticados de la demanda tienden a desviarse alrededor de 34 vuelos en base a la cantidad real por semana.
- Con referencia al RMSE, la cual es una medida que enfatiza los errores más grandes de las predicciones a comparación del MAE, se puede apreciar que el modelo de menor error es el SARIMA (3,1,0) (0,1,1)₅₂ con un valor de 46.81049; el cual refiere a que, en promedio, las predicciones se alejan de los valores reales en aproximadamente 47 vuelos semanales.

Por último, en base al contraste de las métricas analizadas para los tres modelos en líneas anteriores, se toma la decisión de elegir al SARIMA (3,1,0) (0,1,1)₅₂ como el mejor modelo para el pronóstico de la demanda de vuelos semanales (Ver Figura 4.8).

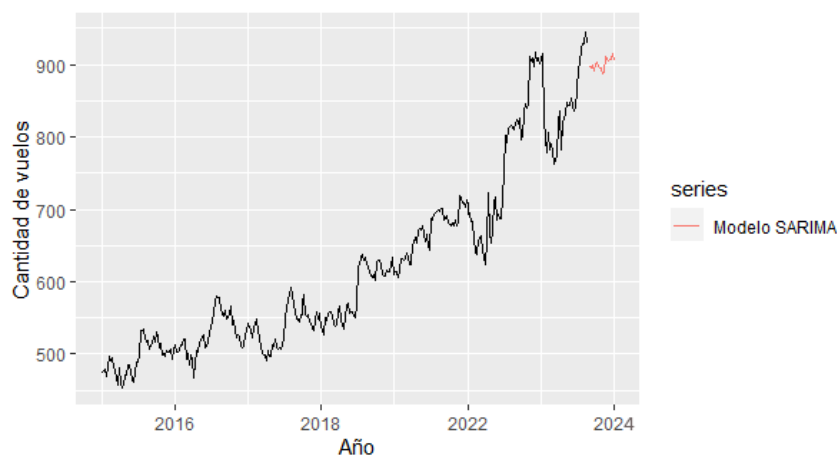


Figura 4.8 Pronóstico final con el modelo SARIMA

Cabe resaltar que, con un modelo de pronóstico definido, ya no será necesario realizar requerimiento de personal extra fuera del tiempo establecido al área de Reclutamiento. En ese sentido, el costo de reclutamiento por persona es de S/500; sin embargo, si se realiza el pedido a destiempo, el costo se penaliza y asciende a S/750. Adicionalmente, de manera histórica se tiene el dato que 20% del requerimiento necesario se realiza fuera del plazo acordado. En la siguiente sección se estará abordando a detalle este costo una vez que se obtengan los resultados de la propuesta de Optimización.

4.2 Programación optimizada de la capacidad del recurso humano

En este apartado, se muestra la solución óptima resultante del modelo previamente planteado en el capítulo 3 sobre la programación de la capacidad del recurso humano para las cuatro semanas de proyección del mes de setiembre 2024, obtenida mediante el uso del programa AMPL IDE (Ver Figura 4.9).

```

CPLEX 22.1.1.0:
<BREAK> (cplex)
CPLEX solution status 13 with fixed integers:
    aborted in phase II
aborted, integer solution exists; objective 62490
106520529 MIP simplex iterations
6082129 branch-and-bound nodes
absmipgap = 807.901, relmipgap = 0.0129285
No basis.
X [*] :=
 1 0 29 1 57 0 85 1 113 0 141 0 169 0 197 0 225 0 253 0
 2 0 30 1 58 0 86 0 114 0 142 0 170 0 198 2 226 0 254 0
 3 0 31 1 59 0 87 0 115 0 143 0 171 0 199 0 227 0 255 0
 4 0 32 0 60 0 88 0 116 0 144 0 172 0 200 0 228 0 256 0
 5 0 33 1 61 0 89 1 117 0 145 0 173 0 201 0 229 0 257 0
 6 1 34 0 62 2 90 0 118 0 146 0 174 1 202 1 230 0 258 0
 7 0 35 0 63 0 91 0 119 0 147 0 175 0 203 2 231 0 259 0
 8 0 36 1 64 0 92 0 120 0 148 0 176 0 204 0 232 0 260 0
 9 0 37 0 65 0 93 0 121 0 149 0 177 0 205 0 233 0 261 0
10 0 38 2 66 0 94 0 122 0 150 0 178 1 206 0 234 0 262 0
11 0 39 0 67 2 95 0 123 0 151 0 179 0 207 0 235 0 263 0
12 0 40 0 68 1 96 0 124 0 152 0 180 0 208 0 236 0 264 0
13 0 41 0 69 0 97 0 125 0 153 0 181 0 209 0 237 0 265 0
14 0 42 0 70 2 98 0 126 0 154 0 182 0 210 0 238 0 266 0
15 0 43 0 71 0 99 0 127 0 155 0 183 1 211 0 239 0 267 1
16 0 44 0 72 2 100 0 128 0 156 0 184 0 212 0 240 0 268 0
17 0 45 0 73 1 101 0 129 0 157 0 185 0 213 0 241 0 269 2
18 0 46 0 74 0 102 0 130 0 158 0 186 0 214 0 242 0 270 1
19 0 47 0 75 0 103 0 131 0 159 0 187 2 215 0 243 0 271 1
20 0 48 0 76 0 104 0 132 0 160 0 188 0 216 0 244 0 272 0
21 0 49 0 77 0 105 0 133 0 161 0 189 0 217 0 245 0 273 0
22 0 50 0 78 0 106 1 134 0 162 0 190 1 218 0 246 0 274 0
23 0 51 0 79 0 107 0 135 0 163 0 191 0 219 0 247 0 275 0
24 0 52 0 80 0 108 0 136 0 164 0 192 0 220 0 248 0 276 0
25 0 53 0 81 1 109 0 137 0 165 0 193 2 221 0 249 0 277 0
26 0 54 0 82 0 110 0 138 0 166 0 194 0 222 0 250 0 278 0
27 0 55 0 83 0 111 0 139 0 167 0 195 0 223 0 251 0 279 1
28 1 56 0 84 0 112 0 140 0 168 0 196 0 224 0 252 0 280 0

```

Figura 4.9 Solución óptima de cantidad de turnos en AMPL IDE

Para la semana 1, se obtuvo como resultado un requerimiento de 42 grupos para cubrir adecuadamente la necesidad de personal por hora para atender los vuelos proyectados en el punto anterior. Así mismo, en cuanto a los costos de planilla de la solución, el monto a incurrir para la asignación de turnos de los grupos sería de S/ 62,490.

El detalle de los turnos que fueron seleccionados para la semana 1 se muestra en el Anexo 30.

Con el fin de validar la solución óptima, se recrea una gráfica de oferta y demanda de un día específico, teniendo como premisa que la oferta debería ser mayor a la demanda en todos los rangos horarios (Ver Figura 4.10).

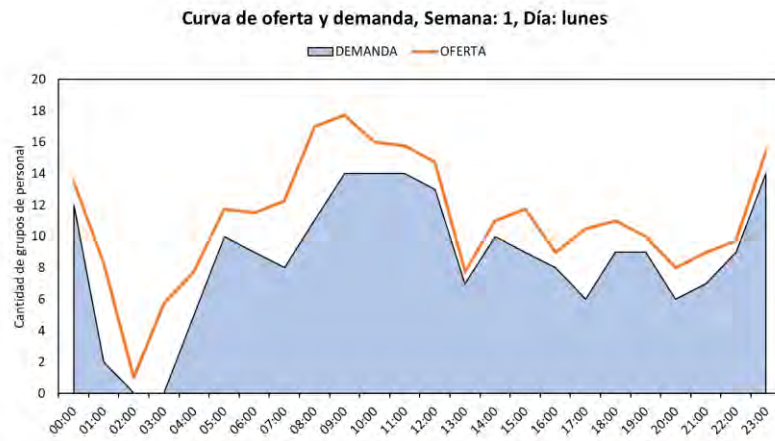


Figura 4.10 Oferta y demanda del día lunes de la semana 1 aplicando el método propuesto

En efecto, como se puede apreciar en la Figura 4.10, la curva de oferta cumple con el supuesto esperado de atender el requerimiento de demanda en su totalidad dentro de las 24 horas del día.

En el Anexo 31 se presentan las gráficas de oferta y demanda de los demás días de la semana 1.

Posteriormente, se iteró el modelo con la demanda de las tres semanas restantes a proyectar; de esta forma, se obtiene el resumen de resultados de la optimización en la Tabla 4.2, donde el costo total de los grupos en las cuatro semanas asciende a S/ 249 877, manteniendo una cantidad de 42 grupos semanales.

Tabla 4.2 Resumen de cantidad de grupos y costos del método propuesto

Semana	Vuelos proyectados	Nº Grupos	Costos
1	897	42	S/62,490
2	896	42	S/62,497
3	900	42	S/62,497
4	892	42	S/62,393

En contraste con el método actual, el resultado del cálculo para la cantidad de grupos por turno, considerando el mismo volumen de vuelos proyectados, fue de 52 grupos; el cual equivale en costos a un valorizado de S/326,607 para las cuatro semanas (Ver Tabla 4.2).

Así mismo, en la Figura 4.11 se ilustra como ejemplo la gráfica de oferta y demanda calculada de acuerdo con este método para el lunes de la semana 1, tomando en cuenta la cobertura completa de las atenciones.

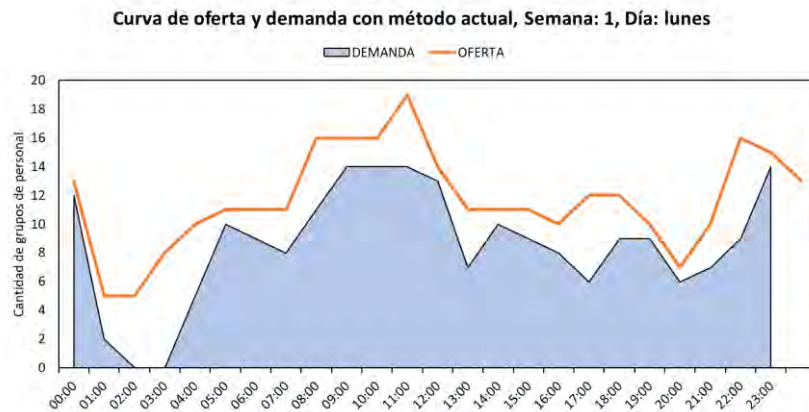


Figura 4.11 Oferta y demanda del día lunes de la semana 1 aplicando el método actual

En cuanto a los demás días de la semana, las gráficas se presentan en el Anexo 32.

Tabla 4.3 Resumen de cantidad de grupos y costos del método actual

Semana	Vuelos proyectados	Nº Grupos	Costos
1	897	52	S/81,652
2	896	52	S/81,652
3	900	52	S/81,652
4	892	52	S/81,652

Con el fin de realizar una mejor comparativa, se grafican ambas ofertas para la demanda del día lunes de la semana 1. Como se observa en la Figura 4.12, ambos modelos cubren por completo la demanda; sin embargo, el método propuesto se muestra más ajustado.

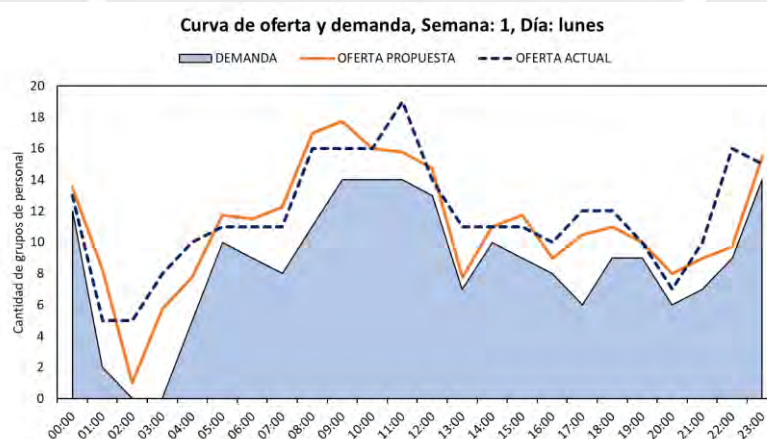


Figura 4.12 Oferta y demanda del día lunes de la semana 1 comparativo método actual y propuesto

Finalmente, al realizar la comparación de ambos métodos, basado en el resumen de la Tabla 4.2 y la Tabla 4.3, se determina que el método optimizado obtiene un ahorro de 10 grupos y S/76,731 en costos con respecto al acumulado mensual (Ver Tabla 4.4).

Tabla 4.4 Ahorro generado entre el método propuesto y actual

Semana	Vuelos proyectados	Método actual		Método propuesto		Ahorro
		Nº Grupos	Costos	Nº Grupos	Costos	
1	897	52	S/81,652	42	S/62,490	S/19,162
2	896	52	S/81,652	42	S/62,497	S/19,155
3	900	52	S/81,652	42	S/62,497	S/19,155
4	892	52	S/81,652	42	S/62,393	S/19,259
Total Ahorro						S/76,731

Asimismo, en la Tabla 4.5 se puede observar el comparativo de los indicadores de utilización y horas muertas de las 4 semanas, donde los resultados de utilización del escenario propuesto en cada semana son superiores en contraste con los de la actual respectivamente. Además, se observa una reducción en la cantidad de horas muertas por cada semana, lo cual refleja un mejor uso del recurso humano.

Tabla 4.5 Comparativo de indicadores de utilización y horas muertas entre el método actual y propuesto

Indicador	Semana 1		Semana 2		Semana 3		Semana 4	
	Propuesta	Actual	Propuesta	Actual	Propuesta	Actual	Propuesta	Actual
Utilización	76%	73%	76%	73%	76%	74%	76%	73%
Horas muertas	454	524	456	526	449	519	460	530

Por otro lado, de acuerdo con el escenario actual, la cantidad de grupos obtenida fue de 52 grupos, de los cuales 3 grupos fueron pedidos a fuera del plazo establecido al área de Reclutamiento. Por lo tanto, eso refleja un costo adicional de S/3,750. Sin embargo, a comparación del escenario optimizado, la empresa ya no incurriría en este costo puesto que no existirían pedidos extemporáneos.

4.3 Simulación del escenario actual y la propuesta de mejora

En este apartado, se muestra los resultados obtenidos en las simulaciones respectivas al escenario actual y a la propuesta de mejora; específicamente, se compara los indicadores obtenidos para cada simulación.

4.3.1 Validación de resultados del modelo

Con el fin de comprobar que los valores de la variable tiempo entre llegadas de aeronaves que se modelan en las réplicas, se simulen dentro de un rango que se asemeje a la realidad operativa, se ha realizado el cálculo de intervalos de confianza para cada franja horaria de esta variable como se muestra en el Anexo 32.

Tal como se aprecia en el Anexo 33, al realizar los cálculos de cada variable para las 90 réplicas basadas en los números generados por las distribuciones ingresadas del software Input Analyzer, se observa que

los valores reales del tiempo entre llegadas se encuentran dentro de los intervalos de confianza, por lo que se comprueba la validación de los resultados de las corridas simuladas.

4.3.2 Comparación de indicadores de desempeño

Al terminar de generar las corridas de la simulación para los escenarios de la situación actual y el optimizado, se obtienen los resultados que se presentan en la Tabla 4.6.

Tabla 4.6 Comparativo de indicadores entre el método actual y optimizado

Indicador	Escenario Actual	Escenario Optimizado	Variación
% Utilización	53.1%	57.2%	4.1%
Productividad	3.71	4.57	0.86
Horas muertas	1170	863	-307
% Demoras	0.2%	8.8%	-8.6%

Con respecto al % de Utilización, se observa que se obtiene un mejor resultado en el escenario optimizado, el cual es de 57.2%, puesto que es mayor al cálculo obtenido del resultado para el escenario actual, siendo este de 53,1%. Así mismo, los resultados para este indicador se generaron partir del reporte de Arena sobre el *schedule utilization* y se ha ejemplificado en la gráfica de tiempo real como se muestra en la Figura 4.13.

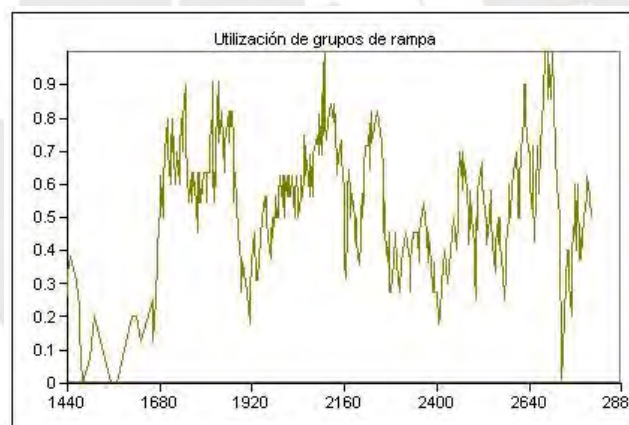


Figura 4.13 Utilización de grupos de rampa en el segundo día de simulación (tiempo en minutos)

Así mismo, al evaluar el indicador de Productividad, se obtiene que, en el escenario optimizado, un grupo de rampa puede atender en promedio 4.57 aeronaves por día; mientras que, en el escenario actual, el grupo solo puede atender 3.71 aeronaves. Es decir, que se obtiene un mejor resultado en el escenario optimizado puesto que la capacidad de atención es mayor comparado al escenario actual. Además, en cuanto al indicador de horas muertas, en el escenario optimizado se obtuvieron 863 horas a la semana; mientras que, en el actual, unas 1170 horas. Esto significa que con un mejor % de Utilización del recurso mano de obra, se generan 307 horas muertas menos a la semana, el cual equivale a 6 grupo de rampa y

un ahorro en costos de S/8,302 a la semana; por lo que el resultado en el escenario optimizado es mejor comparado al actual.

Finalmente, con respecto al indicador de % Demoras, el resultado para el escenario actual fue de 0.2% en la semana, el cual equivale a 2; mientras que, para el optimizado, el % de demoras es 8.8%, el cual corresponde a 81 demoras. Es decir, con mayor cantidad de grupos de rampa, se pueden atender más aeronaves dentro del itinerario indicado por la aerolínea; por lo que, el resultado es mejor en el actual comparado al optimizado. Así mismo, la aerolínea ha fijado una política de penalidades aplicado a demoras de no cumplir con lo acordado en el contrato, puesto que la meta a partir del 2023 sería de 5% tal como se mencionó en el capítulo 2 del presente estudio. En ese sentido, las penalidades se aplican de acuerdo con lo presentado en la Tabla 4.7.

Tabla 4.7 Penalidades de demora semanales por rango

Nivel de rango	Rango % demoras por semana	Penalidad por demora fuera de meta
1	[0;5>	S/ 0
2	[5;10>	S/ 100
3	10≥	S/ 150

Al evaluar la penalidad de demoras para el escenario actual, no se estaría incurriendo en ningún costo ya que está cumpliendo con la meta de 5%; mientras que, para el escenario optimizado, la penalidad sería aplicada en el segundo rango, la cual es de S/100 por demora fuera de meta; es decir, del total de 81 demoras, serían 35 a las que se le estaría aplicando la penalidad del rango, lo cual genera un costo total de S/3,500 por penalidades en demoras a la semana.

Cabe indicar que estos últimos tres indicadores fueron calculados a partir del reporte generado en Arena.

Por otro lado, a fin de que se cumpla la meta de demoras de 5%, se realizó un ejercicio adicional considerando las franjas horarias donde se generan más demoras, las cuales son de 13:00 a 14:00 h y de 20:00 a 21:00 h tal como se muestra a continuación en la Figura 4.14.



Figura 4.14 Distribución de demoras promedio en el escenario optimizado

En este último escenario, se agregó 2 grupos de respaldo operativo para minimizar el impacto de las demoras del rango horario mencionado y se obtuvo los siguientes resultados:

Tabla 4.8 Comparativo de indicadores entre el escenario actual, optimizado y optimizado ajustado

Indicador	Escenario Actual	Escenario Optimizado	Escenario Optimizado ajustado
% Utilización	53.1%	57.2%	54.7%
Productividad	3.71	4.57	4.38
Horas muertas	1170	863	957
% Demoras	0.2%	8.8%	3.6%

De acuerdo con la Tabla 4.8, el escenario optimizado ajustado alcanza un 3.6% de demoras al incrementar los grupos extras mencionados previamente, lo cual representa una mejora significativa en el cumplimiento de la meta establecida del 5%.

4.4 Evaluación de la propuesta de mejora

En resumen, tal como se ha desarrollado en el presente acápite, los resultados se han evaluado para el segmento de Predicción, donde se seleccionó al modelo SARIMA (3,1,0) (0,1,1)_{S2} como el mejor para pronosticar la demanda de vuelos semanales, el cual nos genera un ahorro de S/3,750 por mejorar el tiempo de anticipación en la planificación y así evitar el costo adicional por reclutamiento fuera de tiempo; para el segmento de Optimización, donde se obtuvo un resultado de requerimiento de personal de un total de 42 grupos de rampa, el cual genera un ahorro mensual de S/76,731 en costos debido a la reducción de 10 grupos comparado a la planilla actual. Por último, en el segmento de Simulación, se validó que la propuesta de Optimización cumpla con los niveles de servicio en la atención de vuelos; sin embargo, el resultado del indicador de puntualidad se obtuvo como fuera de meta, generando un costo de S/14,000 en penalidades.

Tabla 4.9 Ahorro total del escenario optimizado vs el actual sobre las propuestas de Pronóstico, Optimización y Simulación de la atención de vuelos

Segmento	Escenario actual	Escenario optimizado	Ahorro mensual
Predicción	S/3,750	S/0	S/3,750
Optimización	S/326,608	S/249,877	S/76,731
Simulación	S/0	S/14,000	-S/14,000
Total Ahorro	S/330,358	S/263,877	S/66,481

Como se puede apreciar en la Tabla 4.9 de todas las propuestas de mejora, el ahorro total sería de S/66,481.

CAPITULO 5. CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

- Mediante la herramienta de *Machine Learning* empleando el modelo SARIMA (3,1,0)(0,1,1) desarrollado en el software RStudio para el pronóstico de la demanda de vuelos, se obtuvo que los resultados de la proyección generada pueden variar con respecto a la demanda real de vuelos en 34 vuelos semanales, el cual equivale en promedio a 5 vuelos diarios, siendo este un margen de error manejable operativamente. Por lo tanto, este modelo se considera adecuado para implementarse en la empresa con el fin de solucionar la poca anticipación en la planificación puesto que se calcularía una demanda más precisa y ya no sería necesario esperar al envío del itinerario de la aerolínea. En consecuencia, se evitarían costos extras mensuales de S/3,750 en el reclutamiento de personal.
- Con la aplicación de la Programación Lineal para la optimización de horarios del personal encargado del proceso de atención de aeronaves, en base a los resultados obtenidos en el software AMPL IDE, se logró atender el mismo volumen de demanda proyectada empleando la herramienta previa de *Machine Learning* con 10 grupos de Rampa menos que el recurso humano actual de la empresa, el cual representa un ahorro de personal de S/76,731 al mes. De esta manera, la empresa podría implementar este modelo para reducir los costos de recurso humano y movilidad debido a la alta cantidad de horarios nocturnos actuales. Además, la reducción de turnos nocturnos estaría contribuyendo a reducir la excesiva rotación del personal.
- Al aplicar la herramienta de Simulación al proceso de atención de aeronaves mediante el software Arena, se reflejó una variación de adelantos y retrasos con respecto al itinerario de la aerolínea. Estas desviaciones permiten asemejar el escenario estudiado a la realidad operativa del aeropuerto, ya que, históricamente, no siempre se cumple la hora programada de llegada o salida del vuelo por diversos factores tanto internos como externos a la aerolínea. En ese sentido, los resultados alcanzados indican que, si la empresa cuenta con la cantidad de grupos de Rampa y les asigna los turnos obtenidos en la propuesta de Optimización, se observa una mejora en el desempeño de los indicadores: incremento del 4.1% en la utilización, aumento de 0.86 puntos de Productividad y 307 horas muertas menos comparado a la situación actual. No obstante, este escenario genera un 8.8% de demoras, por lo que se incurre en un costo de S/14,000 al mes por no alcanzar la meta del cliente. En contraste con ello, si la empresa decide contratar 2 grupos de Rampa adicionales a la propuesta de Optimización en rangos horarios críticos, el % de Demoras se reduce a 3.6%, mejorando así el nivel de servicio con el cliente, donde se estaría incurriendo en un costo de S/12,211 al mes por ambos grupos y así se evitaría el costo de penalidades por demoras. Esto confirma la viabilidad de la propuesta de Optimización, la cual contribuye con la reducción de los costos de personal previamente mencionados.

- Al implementar de forma conjunta las tres propuestas de mejora, la empresa podría generar un ahorro mensual de S/66,481, al desarrollar una planificación con un requerimiento de personal más preciso y anticipado, contribuyendo con una optimización en la utilización del recurso humano y la mejora del nivel de servicio acordado, estableciendo una relación comercial más sólida con la aerolínea y reforzando la competitividad de la empresa.

5.2 Recomendaciones

- Una vez aplicados los modelos propuestos, es necesario mantener una validación continua de los datos, dado que, el contexto analizado puede verse afectado por nuevas variables externas tales como las condiciones de infraestructura y ajustes de procedimientos aeroportuarios que modifiquen el comportamiento de la demanda estudiada, y, de este modo, se deban replantear los modelos ya diseñados y se alteren los resultados obtenidos de la presente tesis.
- Se recomienda que parte del ahorro generado por la implementación de las propuestas optimizadas sea destinada como inversión para mejorar el indicador de demoras de la aerolínea con la contratación de 2 grupos de Rampa como respaldo operativo para ciertos rangos horarios críticos y así alcanzar un 3.6% para cumplir un desempeño dentro de meta. Esta recomendación, y el análisis final de los resultados debe ser acompañado por opinión experta de las áreas de Planificación y Operaciones, con el fin de validar que estén consideradas todas las variables correspondientes.
- En cuanto a la escalabilidad de la propuesta integral, se recomienda que el proyecto sea liderado por un analista senior de Planeamiento en conjunto con el equipo del área y con validación por parte de los jefes de Operaciones. Asimismo, se sugiere el uso de las herramientas RStudio, AMPL y Arena en sus versiones gratuitas. Con respecto al modelo SARIMA, la frecuencia de uso para pronosticar la demanda de vuelos debe ser trimestral, a fin de anticipar la planificación de *Capacity*, el requerimiento de personal al área de Reclutamiento y la elaboración del presupuesto. En cuanto al modelo de programación lineal, se recomienda una ejecución mensual en base a la cantidad de vuelos pronosticados obtenidos en la propuesta de mejora y, de requerir modificación de turnos, estarían sujetas a lineamientos financieros y continuidad operativa durante las 24 horas del día. Por último, se plantea el uso de la simulación para escenarios críticos con cambios de infraestructura que impacten negativamente los niveles de servicio afectados por demoras en el software Arena.
- Se sugiere evaluar la aplicación de los modelos desarrollados en la presente tesis en otros terminales de mayor demanda gestionados por la empresa, tal como el aeropuerto de Cuzco, tomando en consideración las condiciones particulares y recursos de ese aeropuerto, ya que es el segundo destino con mayor afluencia de pasajeros del país después de Lima. Esta expansión permitiría optimizar la planificación en el segundo aeropuerto clave para la empresa, lo cual elevaría más su posicionamiento en el *ground handling* aeroportuario a nivel Perú.

GLOSARIO DE TÉRMINOS

- Vuelos del itinerario: Son aquellos vuelos que se consideran en la programación inicial enviada por las aerolíneas
- Vuelos reales: Son aquellos vuelos que fueron operados en el aeropuerto
- ETA: Hora estimada de arribo del vuelo
- ETD: Hora estimada de salida del vuelo
- ATA: Hora real de arribo del vuelo
- ATD: Hora real de salida del vuelo
- PEA: Puesto de estacionamiento aéreo
- Recepción: Cuando la aeronave aterriza, se realiza el desembarque de pasajeros que arribaron a su destino y permanece en tierra en un aeropuerto determinado
- Tránsito: Cuando la aeronave aterriza, se realiza el desembarque de pasajeros que arribaron a su destino y, a su vez, se embarcan pasajeros para transportarlos hacia un destino diferente. Aquí la aeronave permanece en tierra un tiempo menor a dos horas para volver a volar.
- Salida: Cuando la aeronave ya se encuentra en tierra previamente y se embarcan pasajeros para transportarlos hacia un destino diferente
- NB: Tipo de aeronave de fuselaje angosto
- Capacity: Cantidad de personal considerando descansos que conforma la planificación a partir de la curva de demanda.

BIBLIOGRAFÍA

- Arias, F. (2006). El proyecto de investigación: Introducción a la metodología científica. (5º. ed.) Caracas - Venezuela: Episteme. P.81
- Beunza, J. J., Puertas, E., & Condés, E. (2019). Manual práctico de inteligencia artificial en Entornos Sanitarios. Elsevier Health Sciences.
- Breiman, L. (2001). *Random Forests*
- Cavada, J. P., Cortés, C. E., & Rey, P. A. (2020). , 58(3), 537-559.
- Córdova, M. (2006) Estadística Aplicada.
- Cooper, A. (2012) Vol 1, No 5. What is Analytics? Definition and Essential Characteristic. <http://publications.cetis.org.uk/wp-content/uploads/2012/11/What-is-Analytics-Vol1-No-5.pdf>
- Corporación Peruana de Aeropuertos y Aviación Civil – CORPAC S. A., C. P. (n.d.). *Informe de Desempeño. Ositran*. <https://www.ositran.gob.pe/anterior/wp-content/uploads/2017/12/id-corpac-2021.pdf>
- Da Silva Rocha, M. S. F. (2013). *The staff scheduling problem: a general model and applications* [Tesis de doctorado, Universidade do Porto]. Repositorio Abierto - Universidade do Porto.
- Danesi, A., Mantecchini, L., & Paganelli, F. (2017). *Long-term and short-term forecasting techniques for regional airport planning. ARPN Journal of Engineering and Applied Sciences*, 12(2), 739-745.
- El Comercio. (18 de diciembre de 2022). *Aeropuertos cerrados: ¿Cuándo abrirán los terminales de Cusco, Arequipa, Ayacucho y Juliaca? El MTC responde. El Comercio Perú*. <https://elcomercio.pe/peru/aeropuertos-cerrados-cuando-abriran-los-terminales-de-cusco-arequipa-ayacucho-y-juliaca-el-mtc-responde-noticia/?ref=ecr>
- El Comercio. (19 de noviembre de 2022). *Accidente en el Jorge Chávez: Latam canceló 261 vuelos tras extensión del cierre del aeropuerto. El Comercio Perú*. <https://elcomercio.pe/lima/accidente-en-jorge-chavez-latam-cancelo-261-vuelos-tras-extension-del-cierre-del-aeropuerto-jorge-chavez-bomberos-fallecidos-accidente-callao-lima-airport-partners-rmmn-noticia/?ref=ecr>
- Ernst, A. T., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). *Staff scheduling and rostering: A review of applications, methods and models. European journal of operational research*, 153(1), 3-27.
- Gamarra Sánchez, L. A., & Eyzaguirre Munarriz, J. C. (2018). *Propuesta de mejora de procesos en la atención de aeronaves en tránsito de la flota Airbus de LAN Perú, para incrementar la*

- productividad en el área de mantenimiento* [Tesis de licenciatura, Universidad Peruana de Ciencias Aplicadas]. <https://doi.org/10.19083/tesis/624201>
- Gómez, Brian (2020). Metodología de Box-Jenkins
- Hasan, M. M., & Arefin, M. R. (2017). *Application of linear programming in scheduling problem*. *Dhaka University Journal of Science*, 65(2), 145-150.
- Hastie T., Tibshirani R., Friedman J.(2008) *The Elements of Statistical Learning Data Mining, Inference, and Prediction Second Edition*
- Haykin, S. (2009). *Neural networks and learning machines*. (3^a ed.). Pearson Education.
- Hillier, F. y Lieberman, G. (2010). *Investigación de operaciones*. McGraw-Hill Education.
- IBM Corporation. (17 de agosto de 2021). *El modelo de redes neuronales*. <https://www.ibm.com/docs/es/spss-modeler/saas?topic=networks-neural-model>
- LATAM Airlines Group. (2022). *Memoria Integrada 2021*. [Latamairlinesgroup.net. https://www.latamairlinesgroup.net/static-files/ed33af2b-c466-4313-b29e-1419390ed436](https://www.latamairlinesgroup.net/static-files/ed33af2b-c466-4313-b29e-1419390ed436)
- Levin R.& Rubin D. (2004) *Estadística para Administración y Economía*, 7ma. Edición
- Lutkevich, B.(2010). *Business Analytics* <http://searchbusinessanalytics.techtarget.com/definition/business-analytics-BA>
- Manuele,L (2021) *Simulation and Modelling to Understand Change*
- Ministerio de Asuntos Económicos y Transformación Digital. (31 de agosto de 2020). *¿Cómo aprenden las máquinas? Machine Learning y sus diferentes tipos*. <https://datos.gob.es/es/blog/como-aprenden-las-maquinas-machine-learning-y-sus-diferentes-tipos>
- Ministerio de Trabajo y Promoción del Empleo. (7 de diciembre de 2023). *Guía para el trabajador en horario nocturno*. Gob.Pe. <https://www.gob.pe/institucion/mtpe/informes-publicaciones/4415453-guia-para-el-trabajador-en-horario-nocturno>
- Michigan State University. (2019). *Developing Business Solutions with the Business Analytics Cycle*. <https://www.michiganstateuniversityonline.com/resources/business-analytics/business-solutions-with-business-analytics-cycle/>
- Mohammed J. Zaki, Wagner Meira, Jr.(2020) *Data Mining and Machine Learning: Fundamental Concepts and Algorithms*, 2nd Edition
- Paredes, D. (2020). *Data Science con R. Análisis de Datos y algoritmos de predicción con R*. Editorial Daniel Paredes. <https://bookdown.org/dparedesi/data-science-con-r/>
- Pérez, R. (2019). *Introducción a los modelos de optimización*. Universidad Piloto de Colombia.
- Prabhakaran, S. (2019). *Time Series Analysis in Python – A Comprehensive Guide with Examples* <https://www.machinelearningplus.com/time-series/time-series-analysis-python/>
- Prabhakaran, S. (2021). *ARIMA Model – Complete Guide to Time Series Forecasting in Python* <https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/>

- Profillidis, V. A., & Botzoris, G. N. (2019). *Trend projection and time series methods*. En *Modeling of Transport Demand* (pp. 225–270). Elsevier
- Ramos, A., Sánchez, P., Ferrer, J. M., Barquín, J., & Linares, P. (2010). *Modelos matemáticos de optimización*. Escuela Técnica Superior de Ingeniería, Departamento de Organización Industrial. https://www.gams.com/docs/contributed/modelado_en_gams.pdf
- Renard, P., Alcolea, A., & Ginsbourger, D. (2013). *Stochastic versus deterministic approaches*. *Environmental modelling: Finding simplicity in complexity*, 133-149.
- Sociedad L R. (18 de noviembre de 2022). *Aeropuerto Jorge Chávez suspendió vuelos hasta la 1 p. m. de este sábado por accidente de avión de Latam*. *La República*. <https://larepublica.pe/sociedad/2022/11/18/aeropuerto-jorge-chavez-suspende-operaciones-por-accidente-de-avion-de-latam>
- Tamayo y Tamayo, M. (1997). *El Proceso de la Investigación científica*. Editorial Limusa S.A. México.1997.
- Tang L., Li J., Du H., Ling L. Wang W. (2022) *Big Data in Forecasting Research: A Literature Review* *Big Data Research*, Volume 27, <https://www.sciencedirect.com/science/article/pii/S2214579621001064>
- Zhou, Z. (2021). *Machine learning*. Springer Nature

ANEXOS

Anexo 1 Ejemplo de programación lineal

El departamento de policía programa a los oficiales en turnos de 8 horas de duración. Los horarios son los siguientes: 8 a.m. a 4 p.m., 12 p.m. a 8 p.m., 4 p.m. a 12 a.m., 8 p.m. a 4 a.m., 12 a.m. a 8 a.m. y 4 a.m. a 12 p.m. La dotación de personal en servicio requerido en el departamento de policía varía según la hora. En la siguiente tabla se detallan los requerimientos por rango horario.

Hora del día	Mínima cantidad de oficiales de turno
8:00 a.m. - 12:00 p.m.	5
12:00 p.m. - 4:00 p.m.	6
4:00 p.m. - 8:00 p.m.	10
8:00 p.m. - 12:00 a.m.	7
12:00 a.m. - 4:00 a.m.	4
4:00 a.m. - 8:00 a.m.	6

Para la solución de este problema, los autores elaboran un modelo de programación lineal que determinará la cantidad de policías programados por turno de ocho horas. Este modelo sigue la siguiente estructura:

VARIABLES DE DECISIÓN (x_i):

Siendo x_i el número de oficiales por turno i ($i = 1, 2, 3, 4, 5, 6$), tenemos que,

x_1 = cantidad de oficiales que trabajan en el turno 1: desde las 8 a.m. a las 4 p.m.

x_2 = cantidad de oficiales que trabajan en el turno 2: desde las 12 p.m. a las 8 p.m.

x_3 = cantidad de oficiales que trabajan en el turno 3: desde las 4 p.m. a las 12 a.m.

x_4 = cantidad de oficiales que trabajan en el turno 4: desde las 8 p.m. a las 4 a.m.

x_5 = cantidad de oficiales que trabajan en el turno 5: desde las 12 a.m. a las 8 a.m.

x_6 = cantidad de oficiales que trabajan en el turno 6: desde las 4 a.m. a las 12 p.m.

Restricciones:

Considerando que en el rango horario de 8:00 a.m. - 12:00 p.m. se requieren por lo menos 5 oficiales, la suma de oficiales que laboren en el turno 1 y 6 debe ser al menos 5. De esta forma, la primera restricción sería

$$x_6 + x_1 \geq 5$$

Siguiendo la misma lógica, se elaboran las siguientes restricciones:

$$x_1 + x_2 \geq 6$$

$$x_2 + x_3 \geq 10$$

$$x_3 + x_4 \geq 7$$

$$x_4 + x_5 \geq 4$$

$$x_5 + x_6 \geq 6$$

Condición de no negatividad:

Se agregan las restricciones de no negatividad:

$$x_1 \geq 0; x_2 \geq 0; x_3 \geq 0; x_4 \geq 0; x_5 \geq 0; x_6 \geq 0.$$

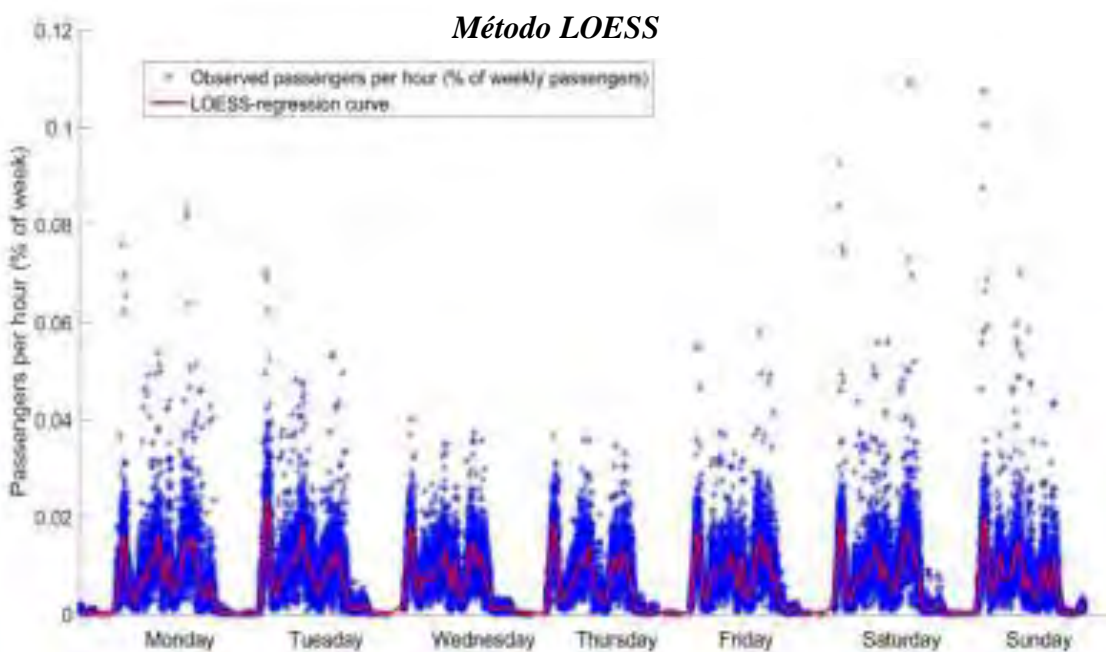
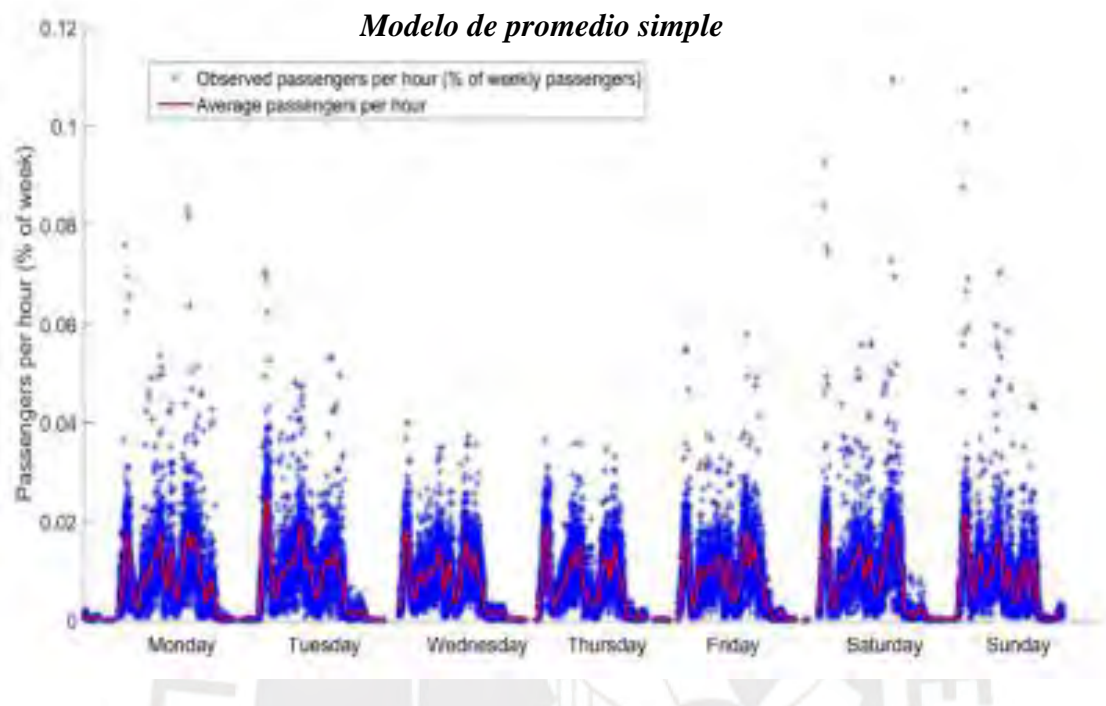
Función Objetivo (F.O.):

Se busca minimizar el número total de oficiales, cumpliendo con todos los requisitos de turno antes mencionados (39).

$$\text{Min } z = \sum_{i=1}^6 x_i \quad (39)$$

Los autores utilizan el software MATHEMATICA para realizar las iteraciones necesarias al modelo para conseguir la solución óptima. Se obtuvo que, el número mínimo de oficiales requerido es 19, siendo $x_1 = 3$, $x_2 = 3$, $x_3 = 7$, $x_4 = 0$, $x_5 = 4$, $x_6 = 2$

Anexo 2 Modelos de demanda por día de semana: Método de promedio simple y método LOESS



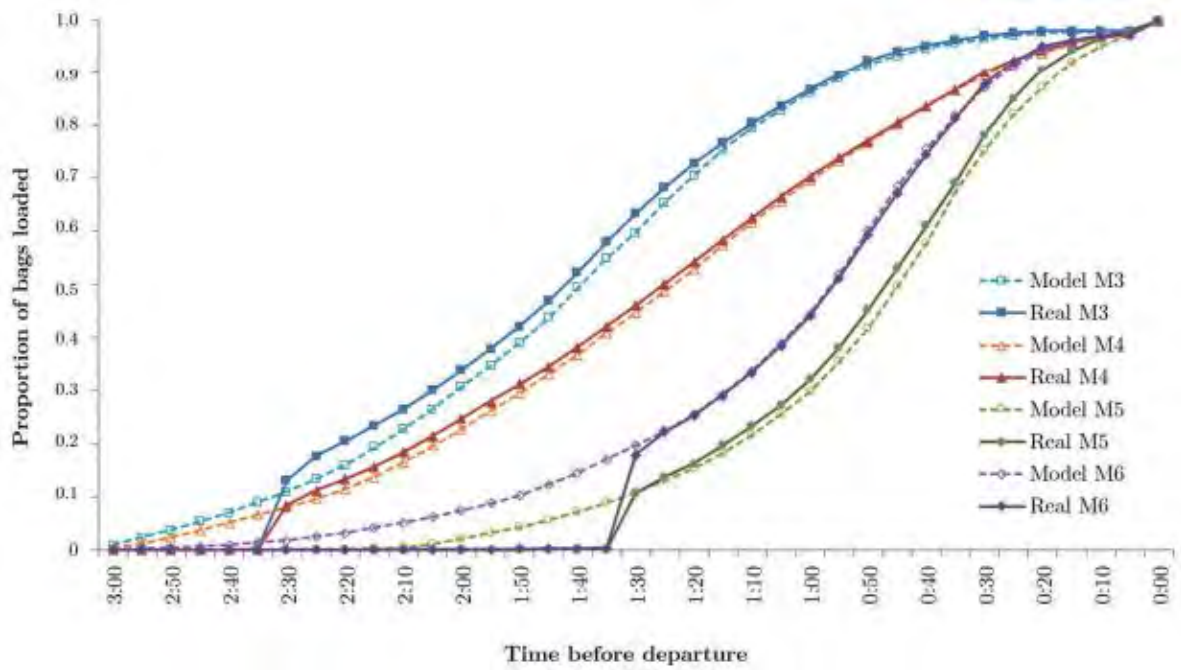
Anexo 3 Modelo de generación de tareas

Modelo
$\min \beta_1 \sum_{t \in T} \sum_{c \in C} x_{tc} + \beta_2 \sum_{t \in T} \sum_{c \in C} y_{tc} \quad (40)$
Variables
<p>x_{tc}: Número requerido de manipuladores en el período t en cada carrusel c.</p> <p>f_{tv}: Número de maletas del vuelo v procesadas en el carrusel asociado al mismo en el periodo t.</p> <p>s_{tv}: Número de maletas del vuelo v que quedan por procesar al final del período t (inventario).</p> <p>y_{tc}: Número de bolsas por encima del umbral de congestión al final del período t en cada carrusel c.</p> <p>F_{tv}: Número de maletas para el vuelo v procesadas en el carrusel M9 en el período t</p> <p>\hat{s}_{tv}: Número de maletas del vuelo v que quedan en el carrusel M9 para ser procesadas al final del período t.</p>
Restricciones

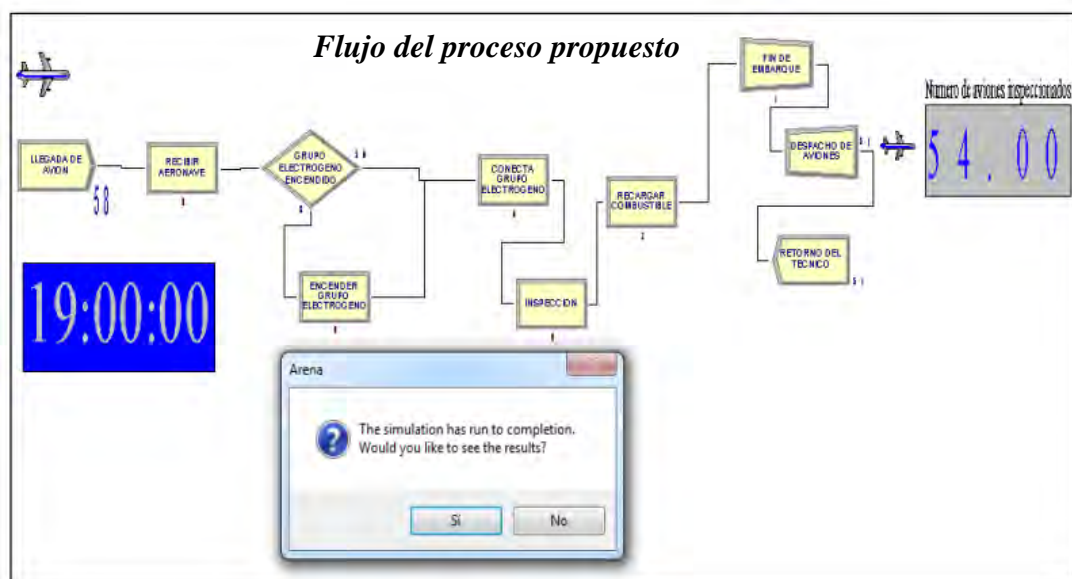
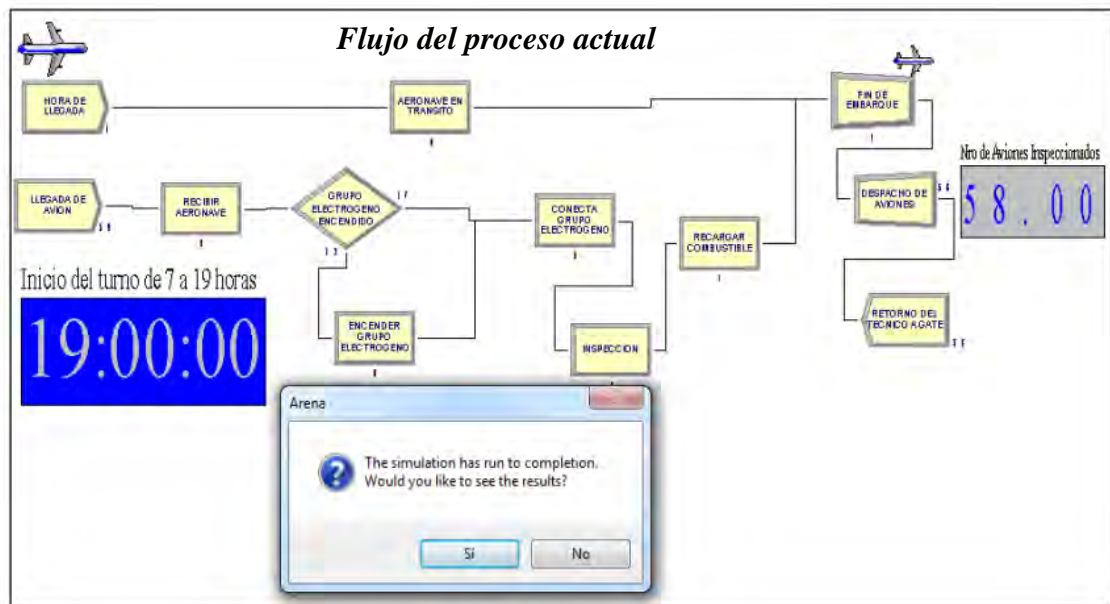
$s_{(t+1)v} = s_{tv} - f_{tv} + \delta_{tv} + \tilde{f}_{(t-1)v}$	$\forall v, \forall t : 1 \leq t \leq T-1$
$\tilde{s}_{(t+1)v} = \tilde{s}_{tv} - \tilde{f}_{tv} + \tilde{\delta}_{tv}$	$\forall v, \forall t : t \leq T-1$
$s_{tv} = 0$	$\forall v, t \in \{0, T\}$
$\tilde{s}_{tv} = 0$	$\forall v, t \in \{0, T\}$
$\sigma_{tc} + \sum_{v \in V} \gamma_{vc} s_{tv} \leq M_c$	$\forall t, \forall c \neq M9$
$\sigma_{tc} + \sum_{v \in V} \tilde{s}_{tv} \leq M_c$	$\forall t, c = M9$
$y_{tc} \geq \sum_{v \in V} \gamma_{vc} s_{tv} - L_c$	$\forall t, \forall c$
$\lambda x_{tc} \geq \sum_{v \in V} \gamma_{vc} f_{tv}$	$\forall c : c \neq M9, \forall t$
$\tilde{\lambda} x_{tc} \geq \sum_{v \in V} \tilde{f}_{tv}$	$c = M9, \forall t$
$x_{tc} + \theta_{tc} \leq H_c$	$\forall t, \forall c$
$-\Delta_t \leq x_{tc} - x_{(t-1)c} \leq \Delta_t$	$\forall c, \forall t \geq 1$
$x_{tc} \in \mathbb{Z}^+$	$\forall t, \forall c$
$y_{tc} \geq 0$	$\forall t, \forall c$
$f_{tv}, \tilde{f}_{tv}, s_{tv}, \tilde{s}_{tv} \geq 0$	$\forall t, \forall v$



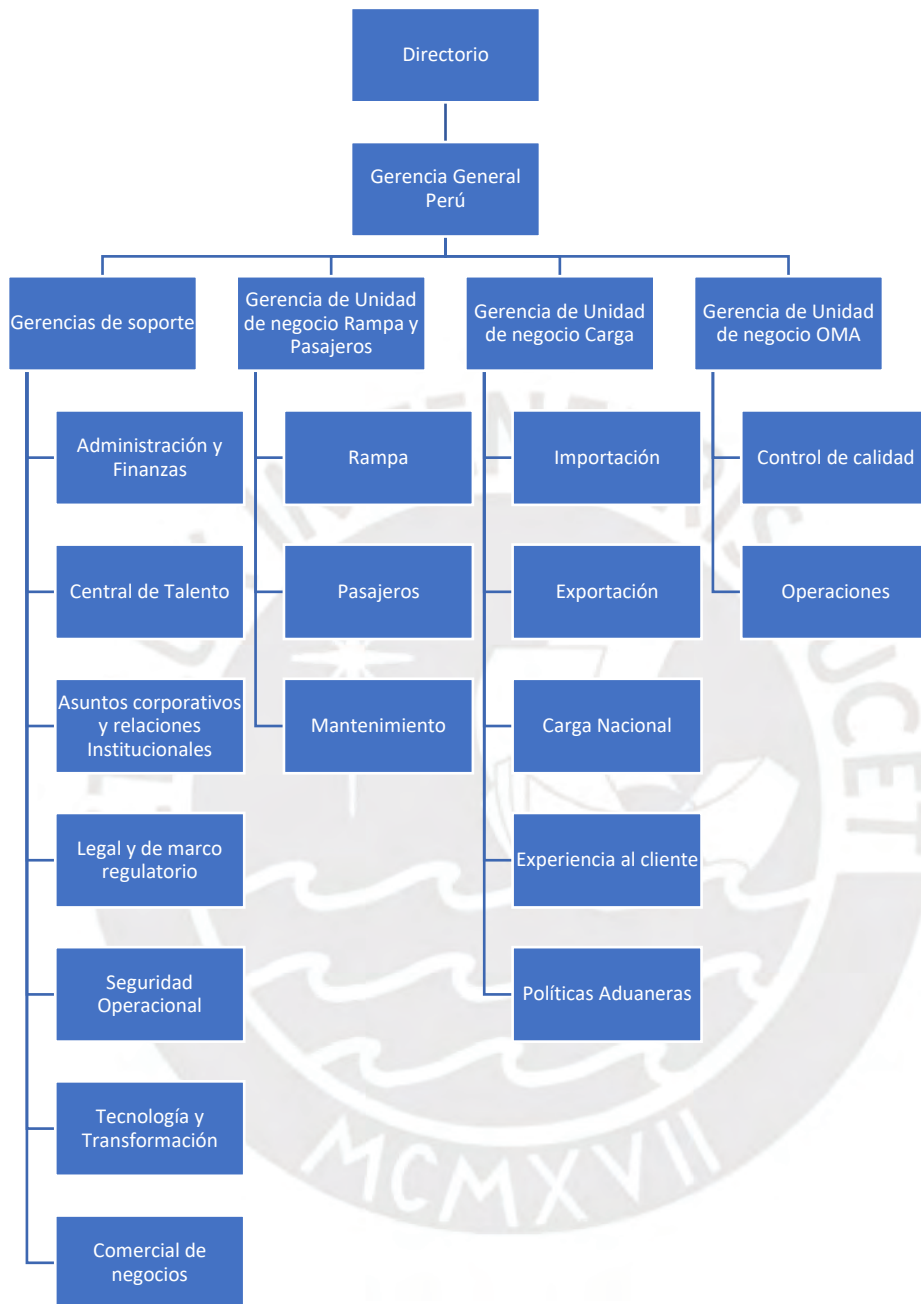
Anexo 4 Curva comparativa propuesta vs real



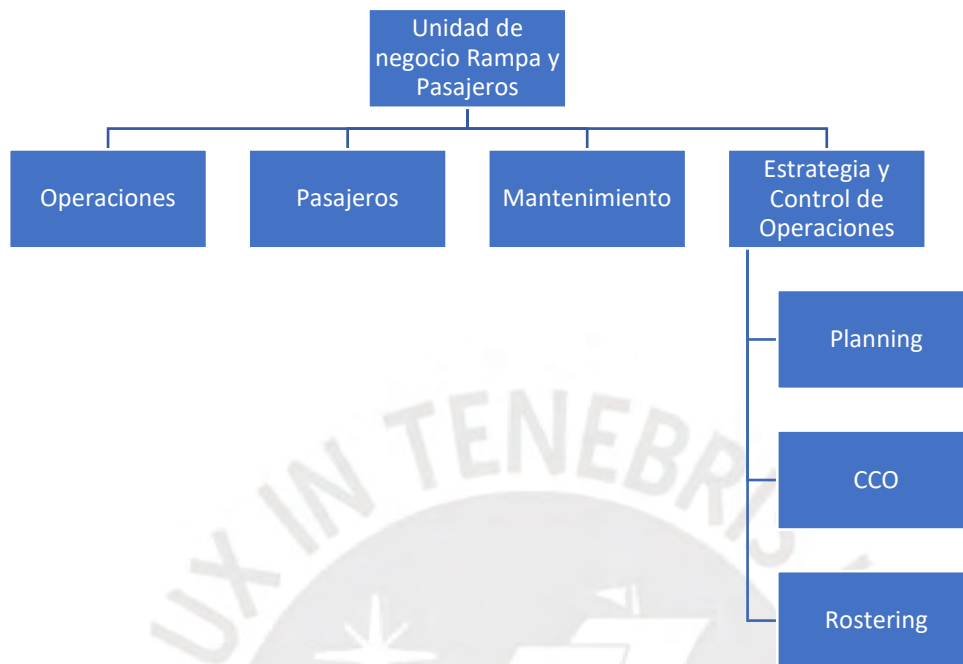
Anexo 5 Flujo de actividades del técnico en el proceso en el software Arena: Actual y propuesto



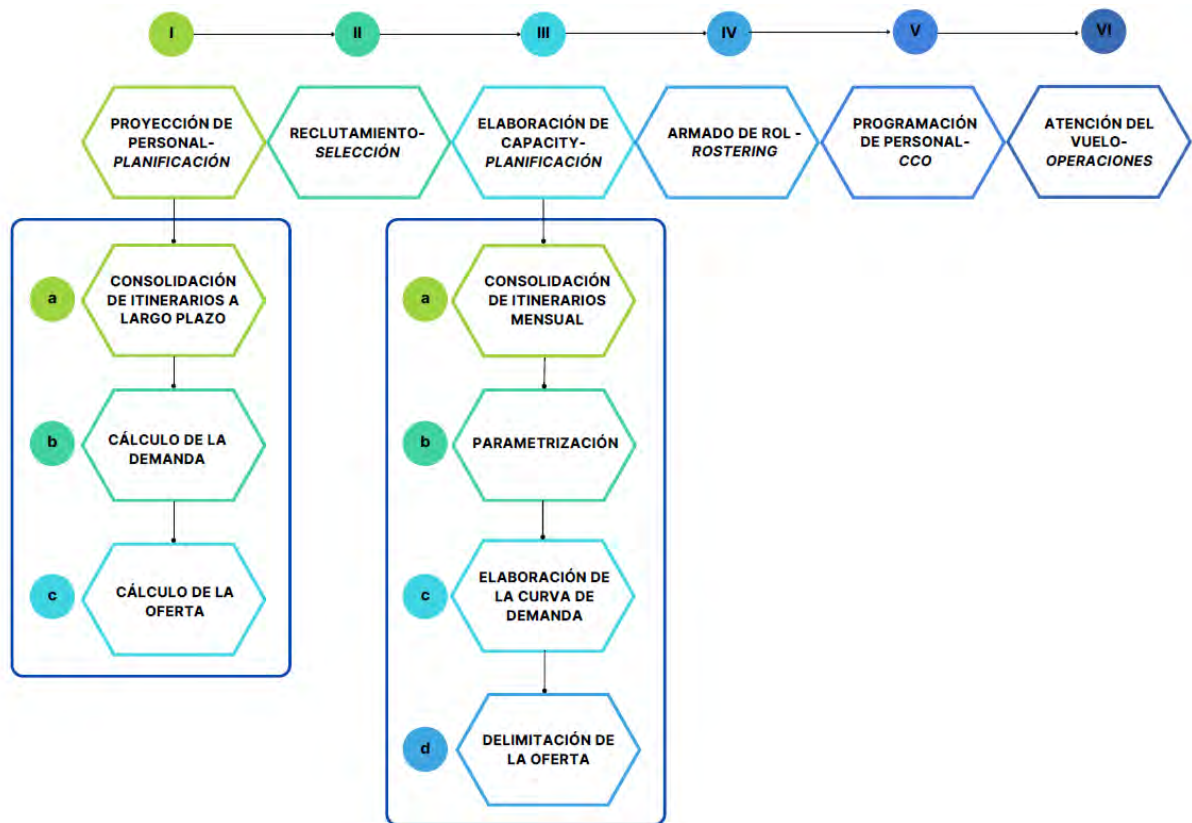
Anexo 6 Organigrama de la empresa



Anexo 7 Organigrama de la unidad de negocio

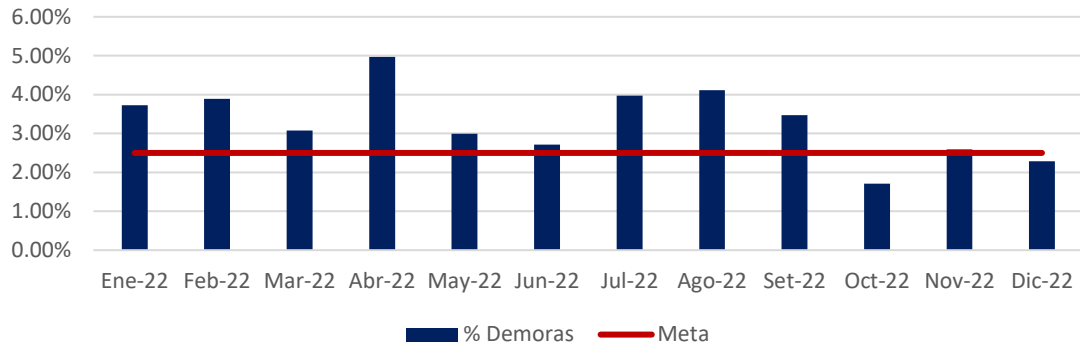


Anexo 8 Proceso de Planificación de la demanda y oferta

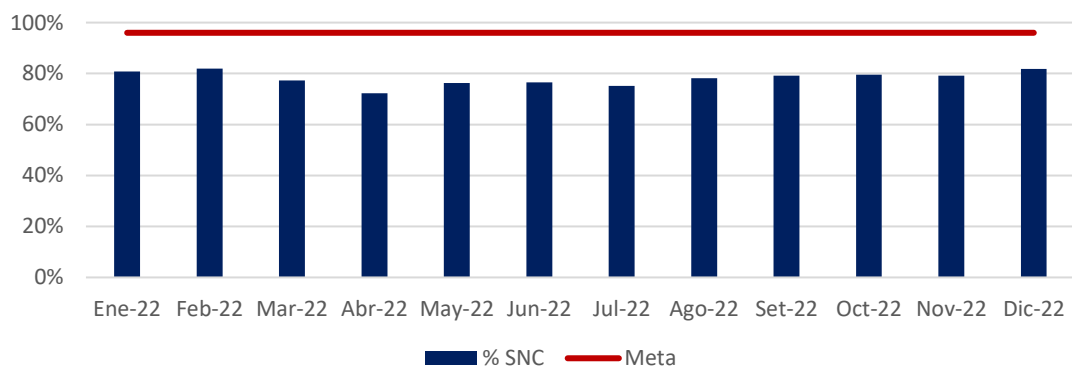


Anexo 9 Evolutivo de demoras de la aerolínea A, nivel de SNC y costos de personal contratado 2022

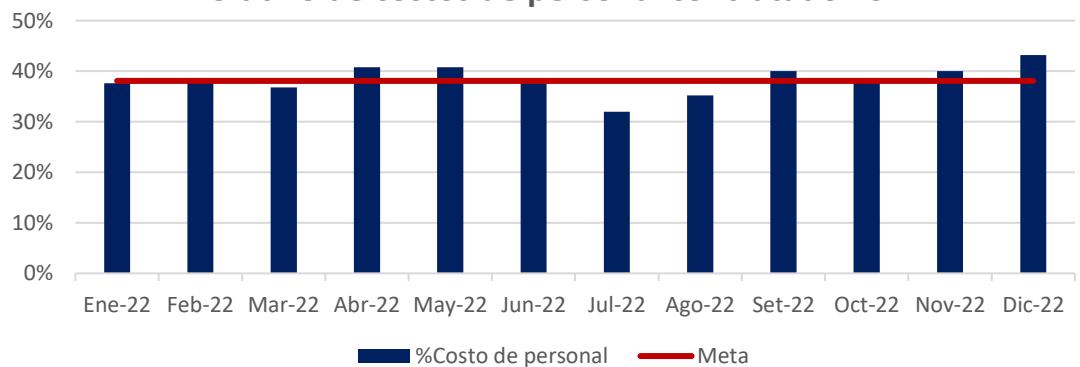
Evolutivo de demoras de la aerolínea A 2022



Evolutivo del nivel de servicios no conformes(SNC) 2022

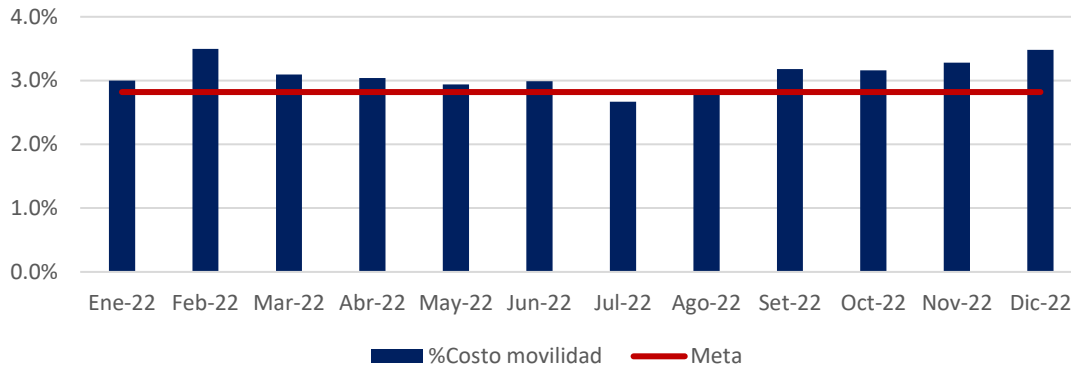


Evolutivo de costos de personal contratado 2022

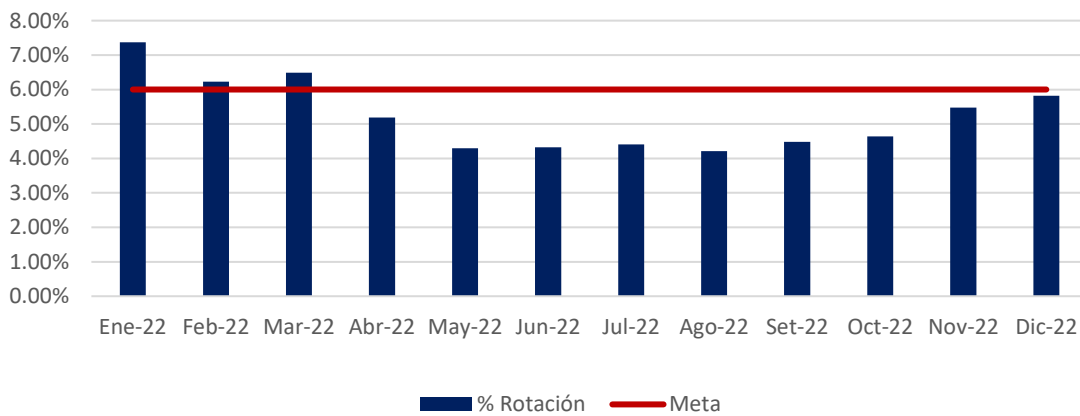


Anexo 10 Evolutivo de costos de movilidad, nivel de rotación de personal y productividad 2022

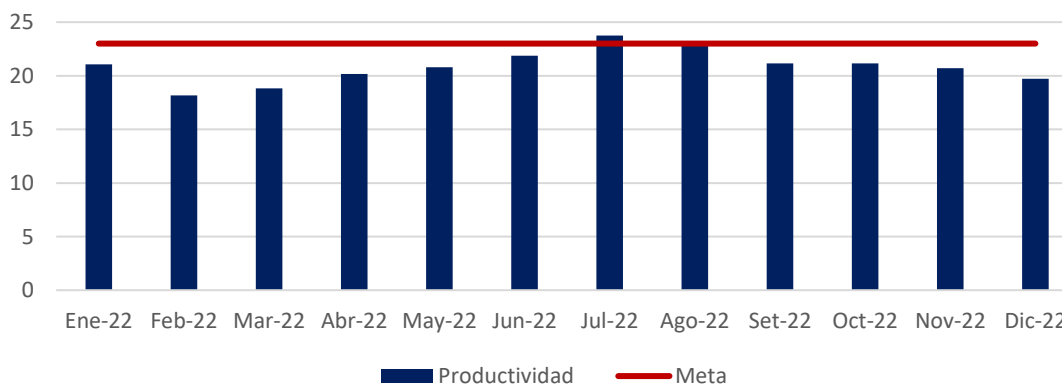
Evolutivo de costos de movilidad 2022



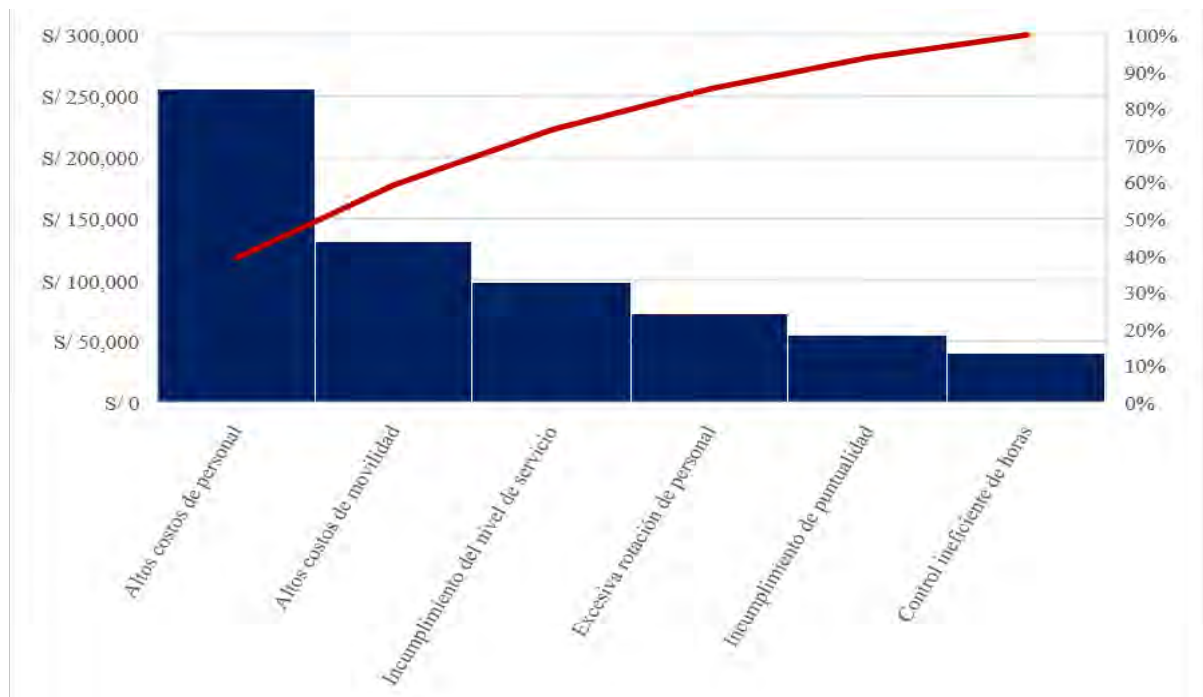
Evolutivo del nivel de rotación de personal 2022



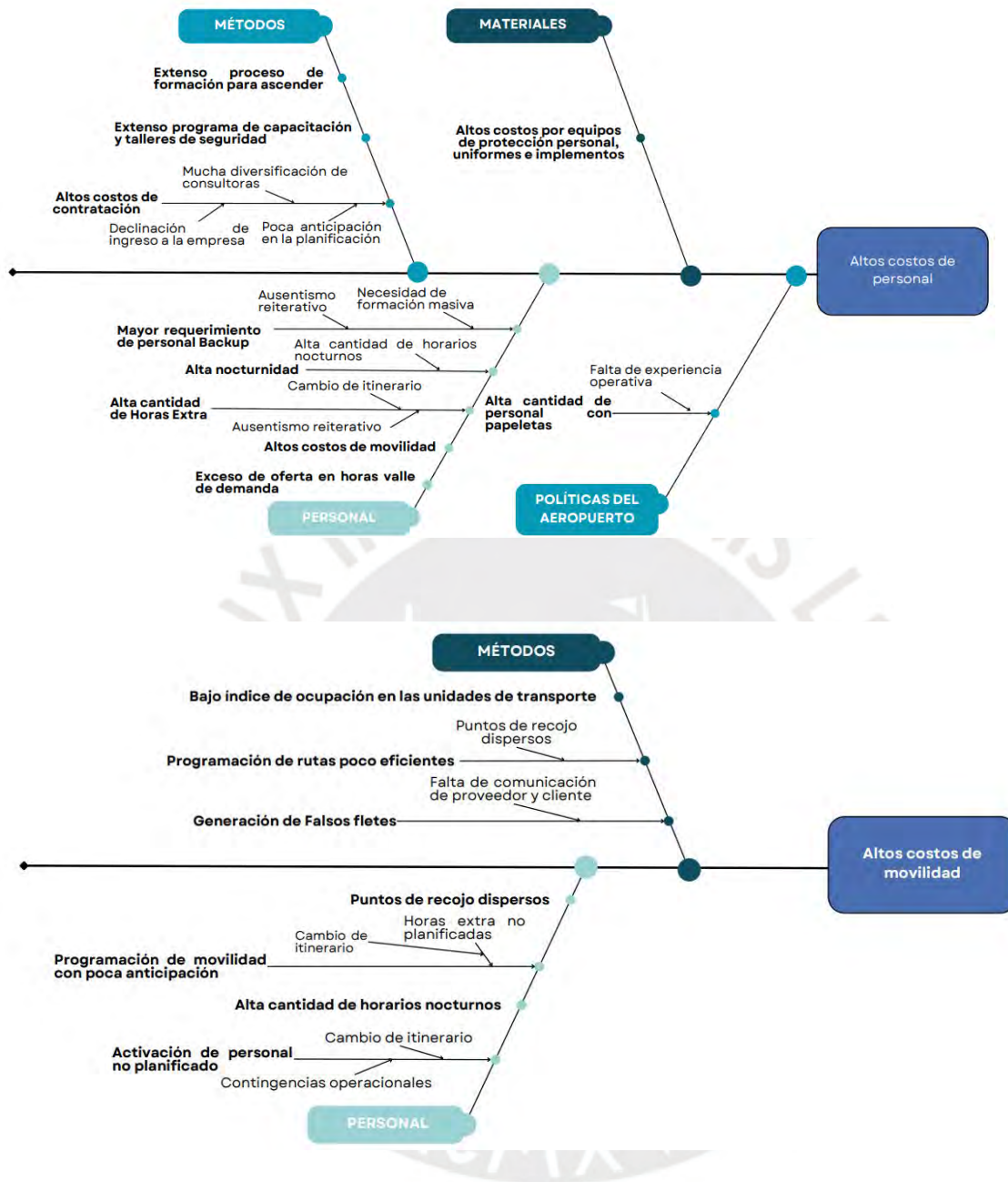
Evolutivo de la productividad 2022



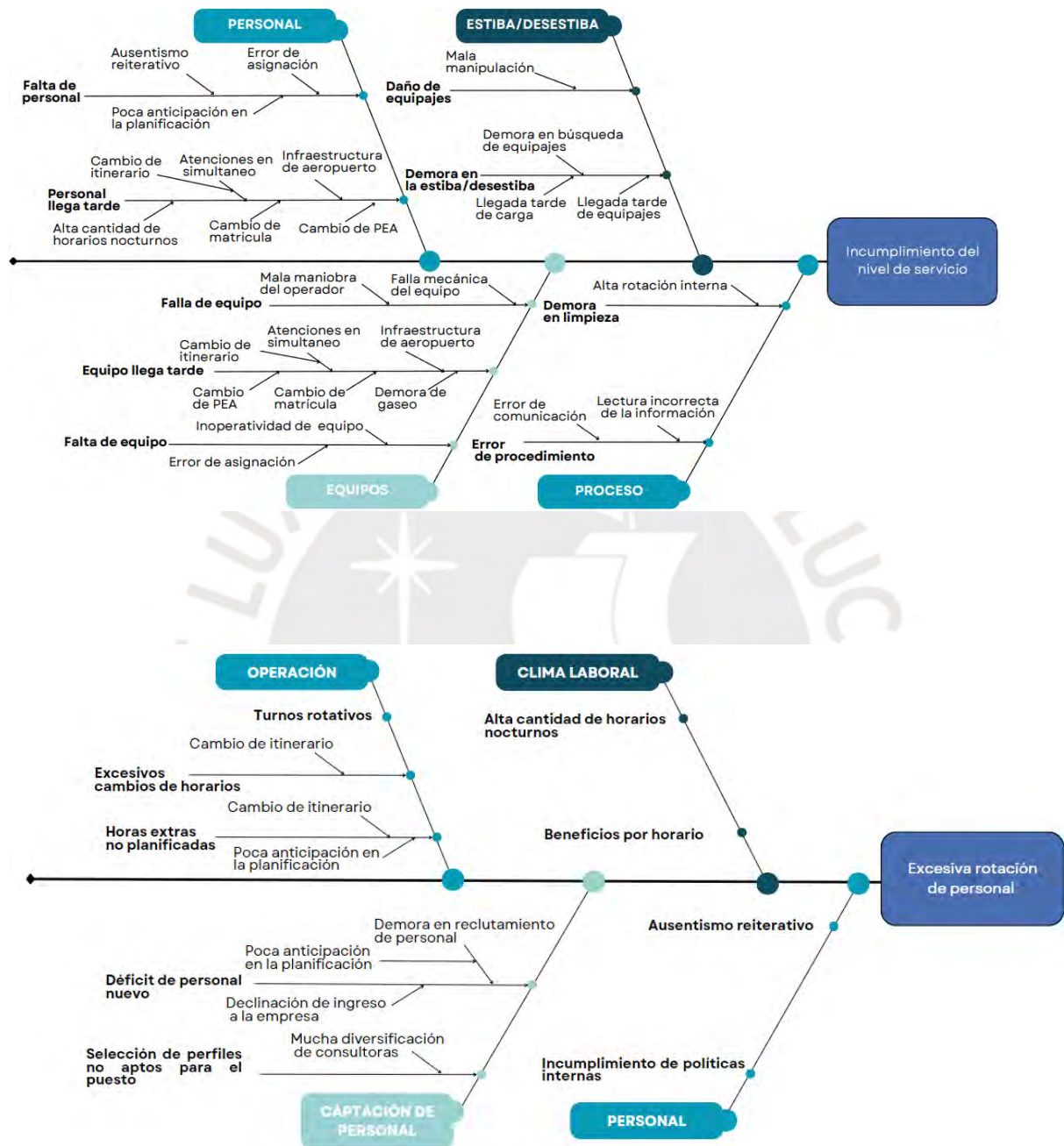
Anexo 11 Diagrama de Pareto de la problemática actual



Anexo 12 Diagramas de Ishikawa: Altos costos de personal y movilidad



Anexo 13 Diagramas de Ishikawa: Incumplimiento del nivel de servicio y excesiva rotación de personal

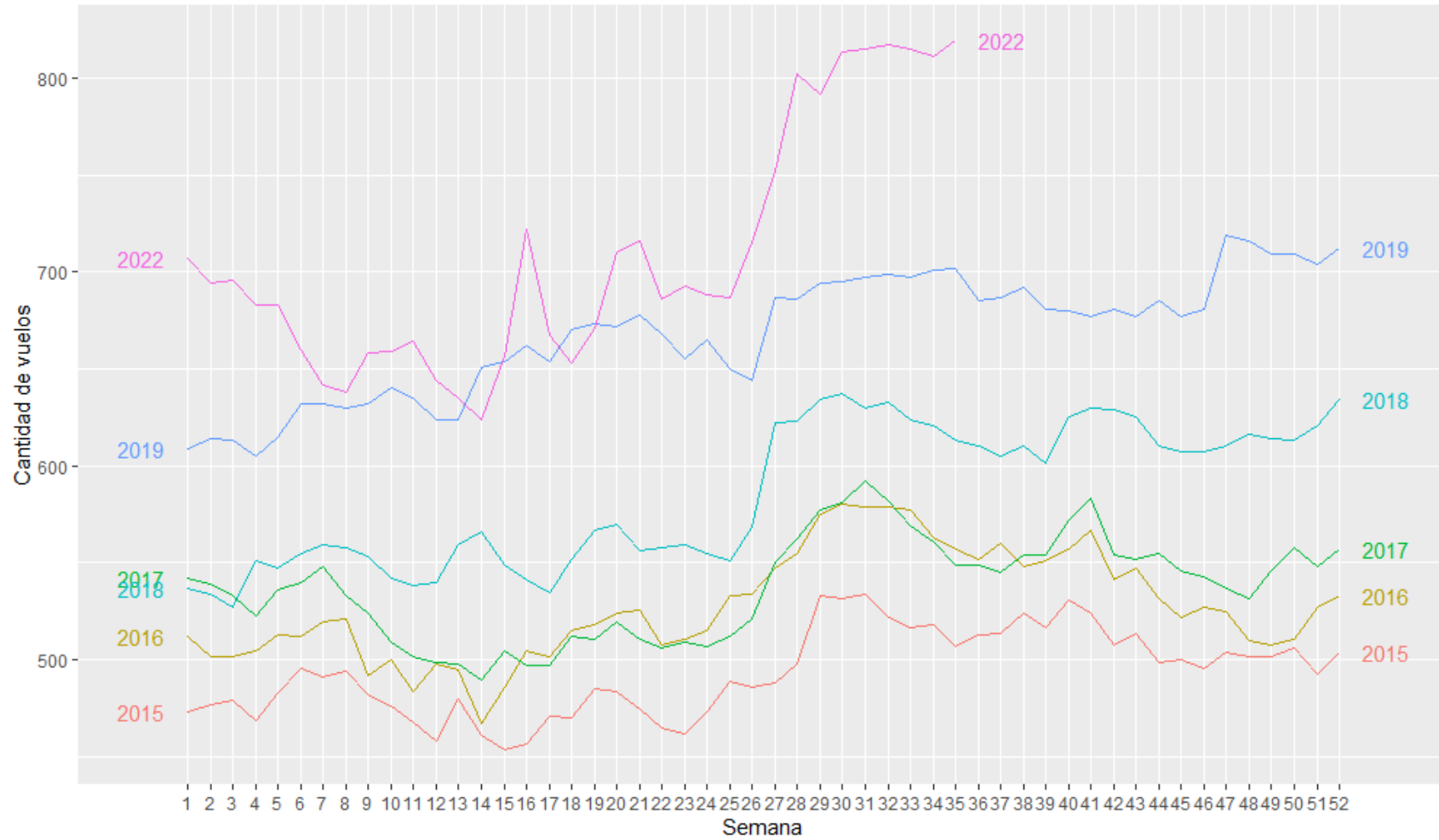


Anexo 14 Clasificación ABC de causas

N°	Causa	Frecuencia	Impacto	Puntaje	% Frecuencia	% Acumulado	Clasificación
1	Cambio de itinerario	7	4	28	18%	18%	A
2	Poca anticipación en la planificación	4	4	16	10%	28%	A
3	Alta cantidad de horarios nocturnos	4	4	16	10%	38%	A
4	Ausentismo reiterativo	4	4	16	10%	48%	B
5	Altos costos de movilidad	2	4	8	5%	53%	B
6	Declinación de ingreso a la empresa	2	3	6	4%	57%	B
7	Mucha diversificación de consultoras	2	3	6	4%	60%	B
8	Exceso de oferta en horas valle de demanda	1	5	5	3%	64%	B
9	Infraestructura de aeropuerto	2	2	4	3%	66%	B
10	Necesidad de formación masiva	2	2	4	3%	69%	B
11	Falta de experiencia operativa	1	4	4	3%	71%	B
12	Alta rotación interna	1	3	3	2%	73%	C
13	Puntos de recojo dispersos	1	3	3	2%	75%	C
14	Bajo índice de ocupación en las unidades de transporte	1	3	3	2%	77%	C
15	Falta de personal en relevo de refrigerio	1	3	3	2%	79%	C
16	Contingencias operacionales	1	2	2	1%	80%	C
17	Demora en búsqueda de equipajes	1	2	2	1%	81%	C
18	Mala maniobra del operador	1	2	2	1%	82%	C
19	Cambio de matrícula	2	1	2	1%	84%	C
20	Cambio de PEA	2	1	2	1%	85%	C
21	Inoperatividad	1	2	2	1%	86%	C
22	Falta de comunicación de proveedor y cliente	1	2	2	1%	87%	C
23	Altos costos por equipos de protección personal, uniformes e implementos	1	2	2	1%	89%	C
24	Error de asignación	2	1	2	1%	90%	C
25	Incumplimiento de políticas internas	1	2	2	1%	91%	C
26	Beneficios por horario	1	2	2	1%	92%	C
27	Falla mecánica	1	2	2	1%	94%	C
28	Turnos rotativos	1	2	2	1%	95%	C
29	Extenso proceso de formación para ascender	1	2	2	1%	96%	C
30	Extenso programa de capacitación y talleres de seguridad	1	2	2	1%	97%	C
31	Mala manipulación	1	1	1	1%	98%	C
32	Llegada tarde de equipajes	1	1	1	1%	99%	C
33	Demora de gaseo	1	1	1	1%	99%	C
34	Llegada tarde de carga	1	1	1	1%	100%	C

Anexo 15 Gráfica estacional de vuelos por años

Gráfica estacional: Vuelos semanales



Anexo 16 Código en RStudio para la aplicación de pruebas Kolmogorov-Smirnov por día de semana

```
1 #Carga de la data
2
3
4 dfdiasem= read.csv2("C:\\Users\\Hogwarts\\Desktop\\TESIS\\Df_diasem2023.csv")
5
6 library(ggplot2)
7
8
9 # Convirtiendo variables tipo caracter a numéricas
10
11 dfdiasem$Lun <- as.numeric(dfdiasem$Lun)
12 dfdiasem$Mar <- as.numeric(dfdiasem$Mar)
13 dfdiasem$Mie <- as.numeric(dfdiasem$Mie)
14 dfdiasem$Jue <- as.numeric(dfdiasem$Jue)
15 dfdiasem$Vie <- as.numeric(dfdiasem$Vie)
16 dfdiasem$Sab <- as.numeric(dfdiasem$Sab)
17 dfdiasem$Dom <- as.numeric(dfdiasem$Dom)
18
19 # Creando un histograma por cada variable
20 hist(dfdiasem$Lun, main="Histograma de distribuciones 2023 del día lunes", xlab="% Distribución", ylab="Frecuencia", border="black")
21 hist(dfdiasem$Mar, main="Histograma de distribuciones 2023 del día martes", xlab="% Distribución", ylab="Frecuencia", border="black")
22 hist(dfdiasem$Mie, main="Histograma de distribuciones 2023 del día miércoles", xlab="% Distribución", ylab="Frecuencia", border="black")
23 hist(dfdiasem$Jue, main="Histograma de distribuciones 2023 del día jueves", xlab="% Distribución", ylab="Frecuencia", border="black")
24 hist(dfdiasem$Vie, main="Histograma de distribuciones 2023 del día viernes", xlab="% Distribución", ylab="Frecuencia", border="black")
25 hist(dfdiasem$Sab, main="Histograma de distribuciones 2023 del día sábado", xlab="% Distribución", ylab="Frecuencia", border="black")
26 hist(dfdiasem$Dom, main="Histograma de distribuciones 2023 del día domingo", xlab="% Distribución", ylab="Frecuencia", border="black")
27
28 #Aplicando pruebas de K-S
29
30 ks.test(dfdiasem$Lun, "pnorm", mean(dfdiasem$Lun), sd(dfdiasem$Lun))
31
32 ks.test(dfdiasem$Mar, "pnorm", mean(dfdiasem$Mar), sd(dfdiasem$Mar))
33
34 ks.test(dfdiasem$Mie, "pnorm", mean(dfdiasem$Mie), sd(dfdiasem$Mie))
35
36 ks.test(dfdiasem$Jue, "pnorm", mean(dfdiasem$Jue), sd(dfdiasem$Jue))
37
38 ks.test(dfdiasem$Vie, "pnorm", mean(dfdiasem$Vie), sd(dfdiasem$Vie))
39
40 ks.test(dfdiasem$Sab, "pnorm", mean(dfdiasem$Sab), sd(dfdiasem$Sab))
41
42 ks.test(dfdiasem$Dom, "pnorm", mean(dfdiasem$Dom), sd(dfdiasem$Dom))
43
```

Anexo 17 Muestra de demanda de personal por hora y distribución de demanda por hora en puntos porcentuales

Día	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	TOTAL
1/01/2023	65	25	0	0	30	50	50	55	60	65	70	95	60	35	35	30	40	35	30	50	50	50	70	100	1150
2/01/2023	80	15	10	5	30	65	55	45	70	70	65	80	70	25	55	70	55	35	35	60	50	45	45	95	1230
3/01/2023	90	0	0	5	35	60	45	35	65	80	85	65	65	50	70	40	40	40	65	40	35	45	65	80	1200
4/01/2023	75	5	10	5	30	75	55	40	60	70	70	75	75	45	60	45	50	30	50	70	35	40	55	80	1205
5/01/2023	70	5	5	5	35	65	50	40	80	80	80	50	60	55	80	50	50	35	55	55	35	55	65	85	1245
6/01/2023	80	5	0	0	30	65	60	40	55	80	75	85	90	35	45	50	50	45	55	55	40	45	45	110	1240
7/01/2023	65	10	0	0	25	65	45	35	75	80	90	70	60	35	45	65	30	35	65	45	50	45	65	60	1160
8/01/2023	70	5	0	0	20	50	50	50	65	65	75	90	60	35	45	55	40	40	60	50	40	30	65	90	1150
9/01/2023	85	10	0	0	30	60	45	35	75	85	70	70	70	30	55	75	35	25	55	60	35	35	40	80	1160
10/01/2023	65	5	0	5	35	50	40	40	90	75	90	60	85	45	50	35	40	25	65	45	30	45	80	80	1180

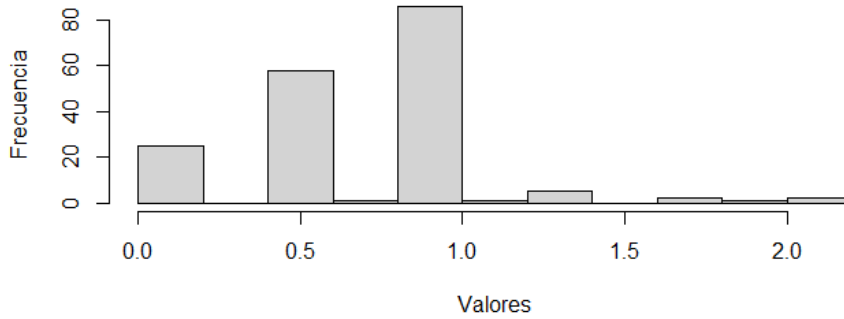
Día	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00	TOTAL
1/01/2023	5.7	2.2	0.0	0.0	2.6	4.3	4.3	4.8	5.2	5.7	6.1	8.3	5.2	3.0	3.0	2.6	3.5	3.0	2.6	4.3	4.3	4.3	6.1	8.7	100.0
2/01/2023	6.5	1.2	0.8	0.4	2.4	5.3	4.5	3.7	5.7	5.7	5.3	6.5	5.7	2.0	4.5	5.7	4.5	2.8	2.8	4.9	4.1	3.7	3.7	7.7	100.0
3/01/2023	7.5	0.0	0.0	0.4	2.9	5.0	3.8	2.9	5.4	6.7	7.1	5.4	5.4	4.2	5.8	3.3	3.3	3.3	5.4	3.3	2.9	3.8	5.4	6.7	100.0
4/01/2023	6.2	0.4	0.8	0.4	2.5	6.2	4.6	3.3	5.0	5.8	5.8	6.2	6.2	3.7	5.0	3.7	4.1	2.5	4.1	5.8	2.9	3.3	4.6	6.6	100.0
5/01/2023	5.6	0.4	0.4	0.4	2.8	5.2	4.0	3.2	6.4	6.4	6.4	4.0	4.8	4.4	6.4	4.0	4.0	2.8	4.4	4.4	2.8	4.4	5.2	6.8	100.0
6/01/2023	6.5	0.4	0.0	0.0	2.4	5.2	4.8	3.2	4.4	6.5	6.0	6.9	7.3	2.8	3.6	4.0	4.0	3.6	4.4	4.4	3.2	3.6	3.6	8.9	100.0
7/01/2023	5.6	0.9	0.0	0.0	2.2	5.6	3.9	3.0	6.5	6.9	7.8	6.0	5.2	3.0	3.9	5.6	2.6	3.0	5.6	3.9	4.3	3.9	5.6	5.2	100.0
8/01/2023	6.1	0.4	0.0	0.0	1.7	4.3	4.3	4.3	5.7	5.7	6.5	7.8	5.2	3.0	3.9	4.8	3.5	3.5	5.2	4.3	3.5	2.6	5.7	7.8	100.0
9/01/2023	7.3	0.9	0.0	0.0	2.6	5.2	3.9	3.0	6.5	7.3	6.0	6.0	6.0	2.6	4.7	6.5	3.0	2.2	4.7	5.2	3.0	3.0	3.4	6.9	100.0
10/01/2023	5.5	0.4	0.0	0.4	3.0	4.2	3.4	3.4	7.6	6.4	7.6	5.1	7.2	3.8	4.2	3.0	3.4	2.1	5.5	3.8	2.5	3.8	6.8	6.8	100.0

Anexo 18 Resultados de pruebas de hipótesis para cada hora del día

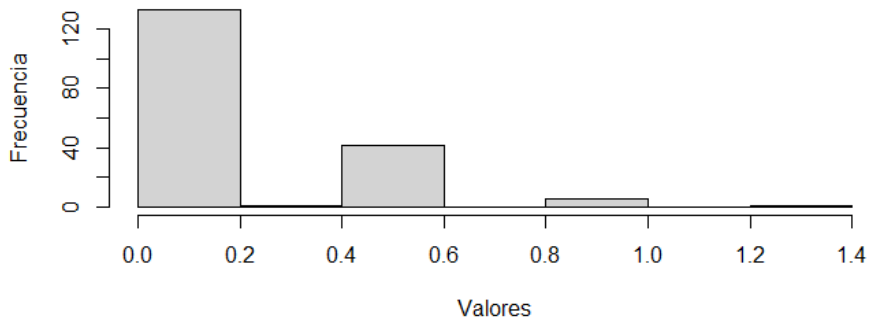
Hora	p-value	Comentario
0	0.08	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
1	4.99E-06	Como $p\text{-value} < \alpha (0.05)$, se rechaza la prueba de hipótesis.
2	2.20E-16	Como $p\text{-value} < \alpha (0.05)$, se rechaza la prueba de hipótesis.
3	1.52E-14	Como $p\text{-value} < \alpha (0.05)$, se rechaza la prueba de hipótesis.
4	0.28	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
5	0.18	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
6	0.34	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
7	0.16	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
8	0.38	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
9	0.76	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
10	0.63	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
11	0.96	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
12	0.59	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
13	0.68	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
14	0.62	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
15	0.78	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
16	0.97	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
17	0.22	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
18	0.24	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
19	0.30	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
20	0.63	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
21	0.73	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
22	0.83	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.
23	0.97	Como $p\text{-value} > \alpha (0.05)$, no se rechaza la prueba de hipótesis.

Anexo 19 Histogramas de Datos para las franjas horarias de: 01:00, 02:00 y 03:00 h

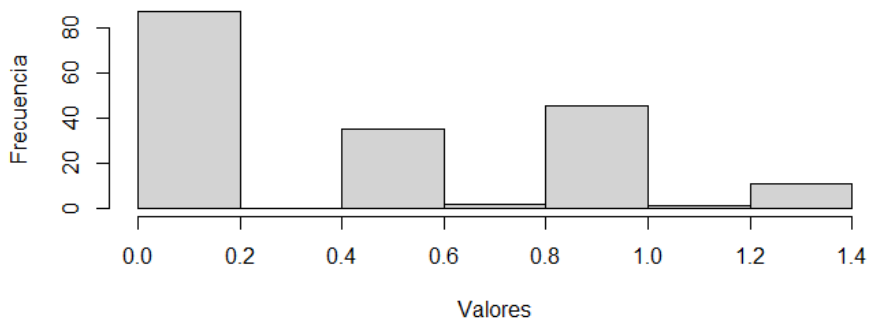
Histograma de Datos para la franja horaria 01:00 h



Histograma de Datos para la franja horaria 02:00 h



Histograma de Datos para la franja horaria 03:00 h



Anexo 20 Resumen de demanda de la semana 1 por franja horaria y día de semana

Hora del día	Día de semana						
	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
00:00	12	12	12	12	12	12	12
01:00	2	2	2	2	2	2	2
02:00	0	0	0	0	0	0	0
03:00	0	0	0	0	0	0	0
04:00	5	5	5	5	5	5	5
05:00	10	10	10	10	10	10	10
06:00	9	9	9	9	9	9	10
07:00	8	8	8	8	8	8	8
08:00	11	11	11	11	11	11	11
09:00	14	13	13	14	14	13	14
10:00	14	14	14	14	14	14	14
11:00	14	14	14	14	14	14	14
12:00	13	13	12	13	13	13	13
13:00	7	7	7	7	7	7	8
14:00	10	10	10	10	10	10	10
15:00	9	9	9	9	9	9	9
16:00	8	8	8	8	8	8	8
17:00	6	6	6	6	6	6	6
18:00	9	9	8	9	9	9	9
19:00	9	9	8	9	9	9	9
20:00	6	6	6	6	6	6	6
21:00	7	7	7	7	7	7	7
22:00	9	9	9	9	9	9	10
23:00	14	13	13	14	14	13	14



Anexo 21Detalle de turnos

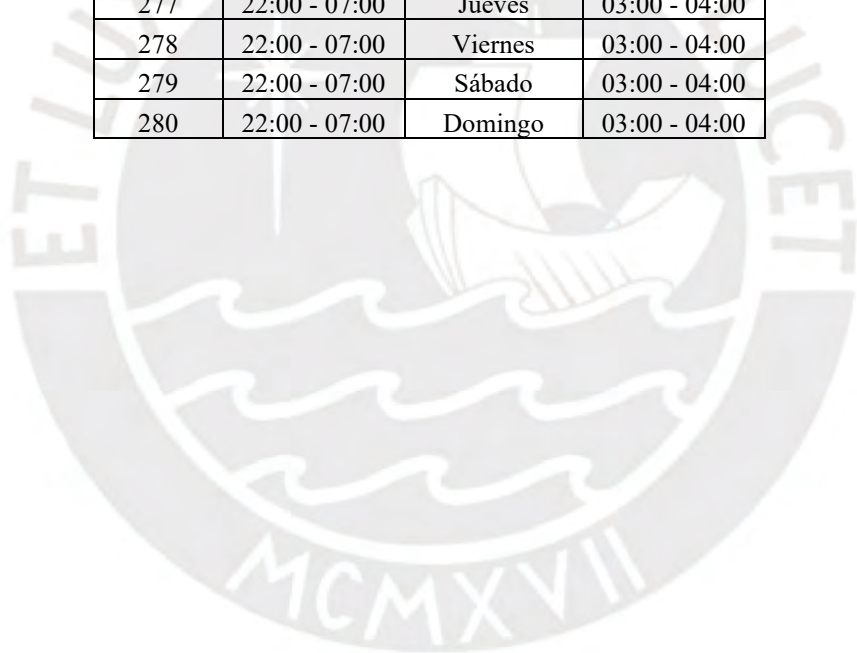
Turno	Horario	Día Descanso	Refrigerio	Turno	Horario	Día Descanso	Refrigerio
1	03:00 - 12:00	Lunes	09:00 - 10:00	44	06:00 - 15:00	Martes	11:00 - 12:00
2	03:00 - 12:00	Martes	09:00 - 10:00	45	06:00 - 15:00	Miércoles	11:00 - 12:00
3	03:00 - 12:00	Miércoles	09:00 - 10:00	46	06:00 - 15:00	Jueves	11:00 - 12:00
4	03:00 - 12:00	Jueves	09:00 - 10:00	47	06:00 - 15:00	Viernes	11:00 - 12:00
5	03:00 - 12:00	Viernes	09:00 - 10:00	48	06:00 - 15:00	Sábado	11:00 - 12:00
6	03:00 - 12:00	Sábado	09:00 - 10:00	49	06:00 - 15:00	Domingo	11:00 - 12:00
7	03:00 - 12:00	Domingo	09:00 - 10:00	50	06:00 - 15:00	Lunes	12:00 - 13:00
8	03:00 - 12:00	Lunes	10:00 - 11:00	51	06:00 - 15:00	Martes	12:00 - 13:00
9	03:00 - 12:00	Martes	10:00 - 11:00	52	06:00 - 15:00	Miércoles	12:00 - 13:00
10	03:00 - 12:00	Miércoles	10:00 - 11:00	53	06:00 - 15:00	Jueves	12:00 - 13:00
11	03:00 - 12:00	Jueves	10:00 - 11:00	54	06:00 - 15:00	Viernes	12:00 - 13:00
12	03:00 - 12:00	Viernes	10:00 - 11:00	55	06:00 - 15:00	Sábado	12:00 - 13:00
13	03:00 - 12:00	Sábado	10:00 - 11:00	56	06:00 - 15:00	Domingo	12:00 - 13:00
14	03:00 - 12:00	Domingo	10:00 - 11:00	57	07:00 - 16:00	Lunes	12:00 - 13:00
15	04:00 - 13:00	Lunes	10:00 - 11:00	58	07:00 - 16:00	Martes	12:00 - 13:00
16	04:00 - 13:00	Martes	10:00 - 11:00	59	07:00 - 16:00	Miércoles	12:00 - 13:00
17	04:00 - 13:00	Miércoles	10:00 - 11:00	60	07:00 - 16:00	Jueves	12:00 - 13:00
18	04:00 - 13:00	Jueves	10:00 - 11:00	61	07:00 - 16:00	Viernes	12:00 - 13:00
19	04:00 - 13:00	Viernes	10:00 - 11:00	62	07:00 - 16:00	Sábado	12:00 - 13:00
20	04:00 - 13:00	Sábado	10:00 - 11:00	63	07:00 - 16:00	Domingo	12:00 - 13:00
21	04:00 - 13:00	Domingo	10:00 - 11:00	64	07:00 - 16:00	Lunes	13:00 - 14:00
22	04:00 - 13:00	Lunes	11:00 - 12:00	65	07:00 - 16:00	Martes	13:00 - 14:00
23	04:00 - 13:00	Martes	11:00 - 12:00	66	07:00 - 16:00	Miércoles	13:00 - 14:00
24	04:00 - 13:00	Miércoles	11:00 - 12:00	67	07:00 - 16:00	Jueves	13:00 - 14:00
25	04:00 - 13:00	Jueves	11:00 - 12:00	68	07:00 - 16:00	Viernes	13:00 - 14:00
26	04:00 - 13:00	Viernes	11:00 - 12:00	69	07:00 - 16:00	Sábado	13:00 - 14:00
27	04:00 - 13:00	Sábado	11:00 - 12:00	70	07:00 - 16:00	Domingo	13:00 - 14:00
28	04:00 - 13:00	Domingo	11:00 - 12:00	71	08:00 - 17:00	Lunes	13:00 - 14:00
29	05:00 - 14:00	Lunes	10:00 - 11:00	72	08:00 - 17:00	Martes	13:00 - 14:00
30	05:00 - 14:00	Martes	10:00 - 11:00	73	08:00 - 17:00	Miércoles	13:00 - 14:00
31	05:00 - 14:00	Miércoles	10:00 - 11:00	74	08:00 - 17:00	Jueves	13:00 - 14:00
32	05:00 - 14:00	Jueves	10:00 - 11:00	75	08:00 - 17:00	Viernes	13:00 - 14:00
33	05:00 - 14:00	Viernes	10:00 - 11:00	76	08:00 - 17:00	Sábado	13:00 - 14:00
34	05:00 - 14:00	Sábado	10:00 - 11:00	77	08:00 - 17:00	Domingo	13:00 - 14:00
35	05:00 - 14:00	Domingo	10:00 - 11:00	78	08:00 - 17:00	Lunes	14:00 - 15:00
36	05:00 - 14:00	Lunes	11:00 - 12:00	79	08:00 - 17:00	Martes	14:00 - 15:00
37	05:00 - 14:00	Martes	11:00 - 12:00	80	08:00 - 17:00	Miércoles	14:00 - 15:00
38	05:00 - 14:00	Miércoles	11:00 - 12:00	81	08:00 - 17:00	Jueves	14:00 - 15:00
39	05:00 - 14:00	Jueves	11:00 - 12:00	82	08:00 - 17:00	Viernes	14:00 - 15:00
40	05:00 - 14:00	Viernes	11:00 - 12:00	83	08:00 - 17:00	Sábado	14:00 - 15:00
41	05:00 - 14:00	Sábado	11:00 - 12:00	84	08:00 - 17:00	Domingo	14:00 - 15:00
42	05:00 - 14:00	Domingo	11:00 - 12:00	85	09:00 - 18:00	Lunes	12:00 - 13:00
43	06:00 - 15:00	Lunes	11:00 - 12:00	86	09:00 - 18:00	Martes	12:00 - 13:00

Turno	Horario	Día Descanso	Refrigerio	Turno	Horario	Día Descanso	Refrigerio
87	09:00 - 18:00	Miércoles	12:00 - 13:00	131	12:00 - 21:00	Viernes	14:00 - 15:00
88	09:00 - 18:00	Jueves	12:00 - 13:00	132	12:00 - 21:00	Sábado	14:00 - 15:00
89	09:00 - 18:00	Viernes	12:00 - 13:00	133	12:00 - 21:00	Domingo	14:00 - 15:00
90	09:00 - 18:00	Sábado	12:00 - 13:00	134	12:00 - 21:00	Lunes	15:00 - 16:00
91	09:00 - 18:00	Domingo	12:00 - 13:00	135	12:00 - 21:00	Martes	15:00 - 16:00
92	09:00 - 18:00	Lunes	13:00 - 14:00	136	12:00 - 21:00	Miércoles	15:00 - 16:00
93	09:00 - 18:00	Martes	13:00 - 14:00	137	12:00 - 21:00	Jueves	15:00 - 16:00
94	09:00 - 18:00	Miércoles	13:00 - 14:00	138	12:00 - 21:00	Viernes	15:00 - 16:00
95	09:00 - 18:00	Jueves	13:00 - 14:00	139	12:00 - 21:00	Sábado	15:00 - 16:00
96	09:00 - 18:00	Viernes	13:00 - 14:00	140	12:00 - 21:00	Domingo	15:00 - 16:00
97	09:00 - 18:00	Sábado	13:00 - 14:00	141	13:00 - 22:00	Lunes	17:00 - 18:00
98	09:00 - 18:00	Domingo	13:00 - 14:00	142	13:00 - 22:00	Martes	17:00 - 18:00
99	10:00 - 19:00	Lunes	12:00 - 13:00	143	13:00 - 22:00	Miércoles	17:00 - 18:00
100	10:00 - 19:00	Martes	12:00 - 13:00	144	13:00 - 22:00	Jueves	17:00 - 18:00
101	10:00 - 19:00	Miércoles	12:00 - 13:00	145	13:00 - 22:00	Viernes	17:00 - 18:00
102	10:00 - 19:00	Jueves	12:00 - 13:00	146	13:00 - 22:00	Sábado	17:00 - 18:00
103	10:00 - 19:00	Viernes	12:00 - 13:00	147	13:00 - 22:00	Domingo	17:00 - 18:00
104	10:00 - 19:00	Sábado	12:00 - 13:00	148	13:00 - 22:00	Lunes	18:00 - 19:00
105	10:00 - 19:00	Domingo	12:00 - 13:00	149	13:00 - 22:00	Martes	18:00 - 19:00
106	10:00 - 19:00	Lunes	13:00 - 14:00	150	13:00 - 22:00	Miércoles	18:00 - 19:00
107	10:00 - 19:00	Martes	13:00 - 14:00	151	13:00 - 22:00	Jueves	18:00 - 19:00
108	10:00 - 19:00	Miércoles	13:00 - 14:00	152	13:00 - 22:00	Viernes	18:00 - 19:00
109	10:00 - 19:00	Jueves	13:00 - 14:00	153	13:00 - 22:00	Sábado	18:00 - 19:00
110	10:00 - 19:00	Viernes	13:00 - 14:00	154	13:00 - 22:00	Domingo	18:00 - 19:00
111	10:00 - 19:00	Sábado	13:00 - 14:00	155	14:00 - 23:00	Lunes	18:00 - 19:00
112	10:00 - 19:00	Domingo	13:00 - 14:00	156	14:00 - 23:00	Martes	18:00 - 19:00
113	11:00 - 20:00	Lunes	13:00 - 14:00	157	14:00 - 23:00	Miércoles	18:00 - 19:00
114	11:00 - 20:00	Martes	13:00 - 14:00	158	14:00 - 23:00	Jueves	18:00 - 19:00
115	11:00 - 20:00	Miércoles	13:00 - 14:00	159	14:00 - 23:00	Viernes	18:00 - 19:00
116	11:00 - 20:00	Jueves	13:00 - 14:00	160	14:00 - 23:00	Sábado	18:00 - 19:00
117	11:00 - 20:00	Viernes	13:00 - 14:00	161	14:00 - 23:00	Domingo	18:00 - 19:00
118	11:00 - 20:00	Sábado	13:00 - 14:00	162	14:00 - 23:00	Lunes	19:00 - 20:00
119	11:00 - 20:00	Domingo	13:00 - 14:00	163	14:00 - 23:00	Martes	19:00 - 20:00
120	11:00 - 20:00	Lunes	14:00 - 15:00	164	14:00 - 23:00	Miércoles	19:00 - 20:00
121	11:00 - 20:00	Martes	14:00 - 15:00	165	14:00 - 23:00	Jueves	19:00 - 20:00
122	11:00 - 20:00	Miércoles	14:00 - 15:00	166	14:00 - 23:00	Viernes	19:00 - 20:00
123	11:00 - 20:00	Jueves	14:00 - 15:00	167	14:00 - 23:00	Sábado	19:00 - 20:00
124	11:00 - 20:00	Viernes	14:00 - 15:00	168	14:00 - 23:00	Domingo	19:00 - 20:00
125	11:00 - 20:00	Sábado	14:00 - 15:00	169	15:00 - 00:00	Lunes	19:00 - 20:00
126	11:00 - 20:00	Domingo	14:00 - 15:00	170	15:00 - 00:00	Martes	19:00 - 20:00
127	12:00 - 21:00	Lunes	14:00 - 15:00	171	15:00 - 00:00	Miércoles	19:00 - 20:00
128	12:00 - 21:00	Martes	14:00 - 15:00	172	15:00 - 00:00	Jueves	19:00 - 20:00
129	12:00 - 21:00	Miércoles	14:00 - 15:00	173	15:00 - 00:00	Viernes	19:00 - 20:00
130	12:00 - 21:00	Jueves	14:00 - 15:00	174	15:00 - 00:00	Sábado	19:00 - 20:00

Turno	Horario	Día Descanso	Refrigerio
175	15:00 - 00:00	Domingo	19:00 - 20:00
176	15:00 - 00:00	Lunes	20:00 - 21:00
177	15:00 - 00:00	Martes	20:00 - 21:00
178	15:00 - 00:00	Miércoles	20:00 - 21:00
179	15:00 - 00:00	Jueves	20:00 - 21:00
180	15:00 - 00:00	Viernes	20:00 - 21:00
181	15:00 - 00:00	Sábado	20:00 - 21:00
182	15:00 - 00:00	Domingo	20:00 - 21:00
183	16:00 - 01:00	Lunes	20:00 - 21:00
184	16:00 - 01:00	Martes	20:00 - 21:00
185	16:00 - 01:00	Miércoles	20:00 - 21:00
186	16:00 - 01:00	Jueves	20:00 - 21:00
187	16:00 - 01:00	Viernes	20:00 - 21:00
188	16:00 - 01:00	Sábado	20:00 - 21:00
189	16:00 - 01:00	Domingo	20:00 - 21:00
190	16:00 - 01:00	Lunes	21:00 - 22:00
191	16:00 - 01:00	Martes	21:00 - 22:00
192	16:00 - 01:00	Miércoles	21:00 - 22:00
193	16:00 - 01:00	Jueves	21:00 - 22:00
194	16:00 - 01:00	Viernes	21:00 - 22:00
195	16:00 - 01:00	Sábado	21:00 - 22:00
196	16:00 - 01:00	Domingo	21:00 - 22:00
197	17:00 - 02:00	Lunes	22:00 - 23:00
198	17:00 - 02:00	Martes	22:00 - 23:00
199	17:00 - 02:00	Miércoles	22:00 - 23:00
200	17:00 - 02:00	Jueves	22:00 - 23:00
201	17:00 - 02:00	Viernes	22:00 - 23:00
202	17:00 - 02:00	Sábado	22:00 - 23:00
203	17:00 - 02:00	Domingo	22:00 - 23:00
204	17:00 - 02:00	Lunes	23:00 - 00:00
205	17:00 - 02:00	Martes	23:00 - 00:00
206	17:00 - 02:00	Miércoles	23:00 - 00:00
207	17:00 - 02:00	Jueves	23:00 - 00:00
208	17:00 - 02:00	Viernes	23:00 - 00:00
209	17:00 - 02:00	Sábado	23:00 - 00:00
210	17:00 - 02:00	Domingo	23:00 - 00:00
211	18:00 - 03:00	Lunes	22:00 - 23:00
212	18:00 - 03:00	Martes	22:00 - 23:00
213	18:00 - 03:00	Miércoles	22:00 - 23:00
214	18:00 - 03:00	Jueves	22:00 - 23:00
215	18:00 - 03:00	Viernes	22:00 - 23:00
216	18:00 - 03:00	Sábado	22:00 - 23:00
217	18:00 - 03:00	Domingo	22:00 - 23:00
218	18:00 - 03:00	Lunes	23:00 - 00:00

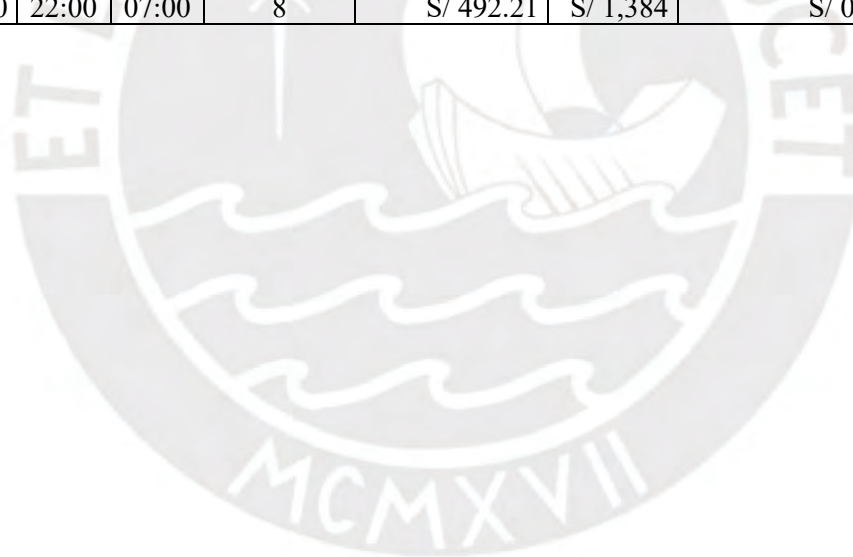
Turno	Horario	Día Descanso	Refrigerio
219	18:00 - 03:00	Martes	23:00 - 00:00
220	18:00 - 03:00	Miércoles	23:00 - 00:00
221	18:00 - 03:00	Jueves	23:00 - 00:00
222	18:00 - 03:00	Viernes	23:00 - 00:00
223	18:00 - 03:00	Sábado	23:00 - 00:00
224	18:00 - 03:00	Domingo	23:00 - 00:00
225	19:00 - 04:00	Lunes	23:00 - 00:00
226	19:00 - 04:00	Martes	23:00 - 00:00
227	19:00 - 04:00	Miércoles	23:00 - 00:00
228	19:00 - 04:00	Jueves	23:00 - 00:00
229	19:00 - 04:00	Viernes	23:00 - 00:00
230	19:00 - 04:00	Sábado	23:00 - 00:00
231	19:00 - 04:00	Domingo	23:00 - 00:00
232	19:00 - 04:00	Lunes	00:00 - 01:00
233	19:00 - 04:00	Martes	00:00 - 01:00
234	19:00 - 04:00	Miércoles	00:00 - 01:00
235	19:00 - 04:00	Jueves	00:00 - 01:00
236	19:00 - 04:00	Viernes	00:00 - 01:00
237	19:00 - 04:00	Sábado	00:00 - 01:00
238	19:00 - 04:00	Domingo	00:00 - 01:00
239	20:00 - 05:00	Lunes	00:00 - 01:00
240	20:00 - 05:00	Martes	00:00 - 01:00
241	20:00 - 05:00	Miércoles	00:00 - 01:00
242	20:00 - 05:00	Jueves	00:00 - 01:00
243	20:00 - 05:00	Viernes	00:00 - 01:00
244	20:00 - 05:00	Sábado	00:00 - 01:00
245	20:00 - 05:00	Domingo	00:00 - 01:00
246	20:00 - 05:00	Lunes	01:00 - 02:00
247	20:00 - 05:00	Martes	01:00 - 02:00
248	20:00 - 05:00	Miércoles	01:00 - 02:00
249	20:00 - 05:00	Jueves	01:00 - 02:00
250	20:00 - 05:00	Viernes	01:00 - 02:00
251	20:00 - 05:00	Sábado	01:00 - 02:00
252	20:00 - 05:00	Domingo	01:00 - 02:00
253	21:00 - 06:00	Lunes	01:00 - 02:00
254	21:00 - 06:00	Martes	01:00 - 02:00
255	21:00 - 06:00	Miércoles	01:00 - 02:00
256	21:00 - 06:00	Jueves	01:00 - 02:00
257	21:00 - 06:00	Viernes	01:00 - 02:00
258	21:00 - 06:00	Sábado	01:00 - 02:00
259	21:00 - 06:00	Domingo	01:00 - 02:00
260	21:00 - 06:00	Lunes	02:00 - 03:00
261	21:00 - 06:00	Martes	02:00 - 03:00
262	21:00 - 06:00	Miércoles	02:00 - 03:00

Turno	Horario	Día Descanso	Refrigerio
263	21:00 - 06:00	Jueves	02:00 - 03:00
264	21:00 - 06:00	Viernes	02:00 - 03:00
265	21:00 - 06:00	Sábado	02:00 - 03:00
266	21:00 - 06:00	Domingo	02:00 - 03:00
267	22:00 - 07:00	Lunes	02:00 - 03:00
268	22:00 - 07:00	Martes	02:00 - 03:00
269	22:00 - 07:00	Miércoles	02:00 - 03:00
270	22:00 - 07:00	Jueves	02:00 - 03:00
271	22:00 - 07:00	Viernes	02:00 - 03:00
272	22:00 - 07:00	Sábado	02:00 - 03:00
273	22:00 - 07:00	Domingo	02:00 - 03:00
274	22:00 - 07:00	Lunes	03:00 - 04:00
275	22:00 - 07:00	Martes	03:00 - 04:00
276	22:00 - 07:00	Miércoles	03:00 - 04:00
277	22:00 - 07:00	Jueves	03:00 - 04:00
278	22:00 - 07:00	Viernes	03:00 - 04:00
279	22:00 - 07:00	Sábado	03:00 - 04:00
280	22:00 - 07:00	Domingo	03:00 - 04:00



Anexo 22 Costos por turnos y horarios

Turnos	Inicio	Fin	Horas nocturnas	Costo por nocturnidad	Costo Planilla	Costo Movilidad	Costo Total
1 - 14	03:00	12:00	3	S/ 184.58	S/ 1,384	S/ 63	S/ 1,631.58
15 - 28	04:00	13:00	2	S/ 123.05	S/ 1,384	S/ 63	S/ 1,570.05
29 - 42	05:00	14:00	1	S/ 61.53	S/ 1,384	S/ 0	S/ 1,445.19
43 - 56	06:00	15:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
57 - 70	07:00	16:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
71 - 84	08:00	17:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
85 - 98	09:00	18:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
99 - 112	10:00	19:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
113 - 126	11:00	20:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
127 - 140	12:00	21:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
141 - 154	13:00	22:00	0	S/ 0.00	S/ 1,384	S/ 0	S/ 1,383.67
155 - 168	14:00	23:00	1	S/ 61.53	S/ 1,384	S/ 0	S/ 1,445.19
169 - 182	15:00	00:00	2	S/ 123.05	S/ 1,384	S/ 63	S/ 1,570.05
183 - 196	16:00	01:00	3	S/ 184.58	S/ 1,384	S/ 63	S/ 1,631.58
197 - 210	17:00	02:00	4	S/ 246.11	S/ 1,384	S/ 63	S/ 1,693.11
211 - 224	18:00	03:00	5	S/ 307.63	S/ 1,384	S/ 63	S/ 1,754.63
225 - 238	19:00	04:00	6	S/ 369.16	S/ 1,384	S/ 63	S/ 1,816.16
239 - 252	20:00	05:00	7	S/ 430.69	S/ 1,384	S/ 0	S/ 1,814.35
253 - 266	21:00	06:00	8	S/ 492.21	S/ 1,384	S/ 0	S/ 1,875.88
267 - 280	22:00	07:00	8	S/ 492.21	S/ 1,384	S/ 0	S/ 1,875.88



Anexo 23 Datos de demanda, costo y turnos ingresados al software AMPL

```

param Demanda:
1 2 3 4 5 6 7:=
1 12 12 12 12 12 12 12
2 2 2 2 2 2 2 2
3 0 0 0 0 0 0 0
4 0 0 0 0 0 0 0
5 5 5 5 5 5 5 5
6 10 10 10 10 10 10 10
7 9 9 9 9 9 9 10
8 8 8 8 8 8 8 8
9 11 11 11 11 11 11 11
10 14 13 13 14 14 13 14
11 14 14 14 14 14 14 14
12 14 14 14 14 14 14 14
13 13 13 12 13 13 13 13
14 7 7 7 7 7 7 8
15 10 10 10 10 10 10 10
16 9 9 9 9 9 9 9
17 8 8 8 8 8 8 8
18 6 6 6 6 6 6 6
19 9 9 8 9 9 9 9
20 9 9 8 9 9 9 9
21 6 6 6 6 6 6 6
22 7 7 7 7 7 7 7
23 9 9 9 9 9 9 10
24 14 13 13 14 14 13 14
;

param Costo:=
1 1548
2 1548
3 1548
4 1548
5 1548
6 1548
7 1548
8 1548
9 1548
10 1548
11 1548
12 1548
13 1548
14 1548
15 1514
16 1514
17 1514
18 1514
19 1514
20 1514
21 1514
22 1514
23 1514
24 1514
25 1514
26 1514
...

param Turnos:=
[*,*,1]: 1 2 3 4 5 6 7 :=
1 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0
4 0 0.75 0.75 0.75 0.75 0.75 0.75
5 0 1 1 1 1 1 1
6 0 1 1 1 1 1 1
7 0 1 1 1 1 1 1
8 0 1 1 1 1 1 1
9 0 1 1 1 1 1 1
10 0 0 0 0 0 0 0
11 0 1 1 1 1 1 1
12 0 0.75 0.75 0.75 0.75 0.75 0.75
13 0 0 0 0 0 0 0
14 0 0 0 0 0 0 0
15 0 0 0 0 0 0 0
16 0 0 0 0 0 0 0
17 0 0 0 0 0 0 0
18 0 0 0 0 0 0 0
19 0 0 0 0 0 0 0
20 0 0 0 0 0 0 0
21 0 0 0 0 0 0 0
22 0 0 0 0 0 0 0
23 0 0 0 0 0 0 0
24 0 0 0 0 0 0 0

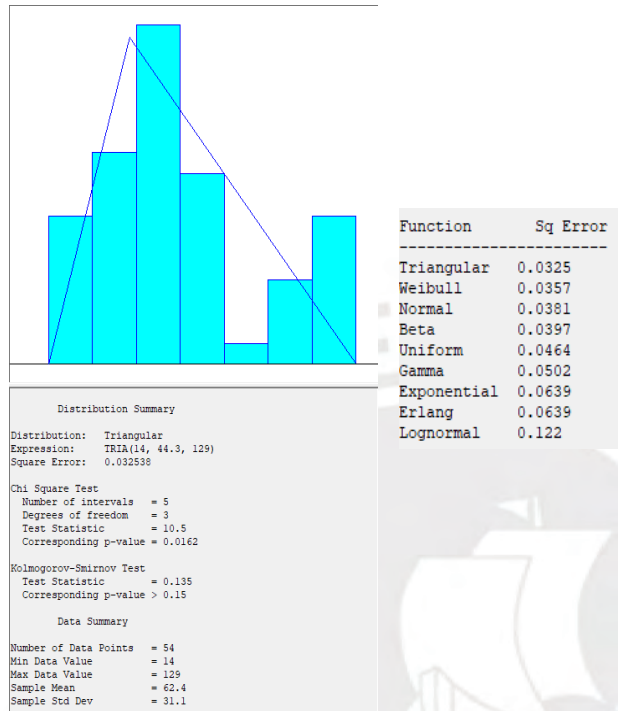
```

Anexo 24 Clasificación de variables aleatorias del tiempo entre llegadas de aeronaves

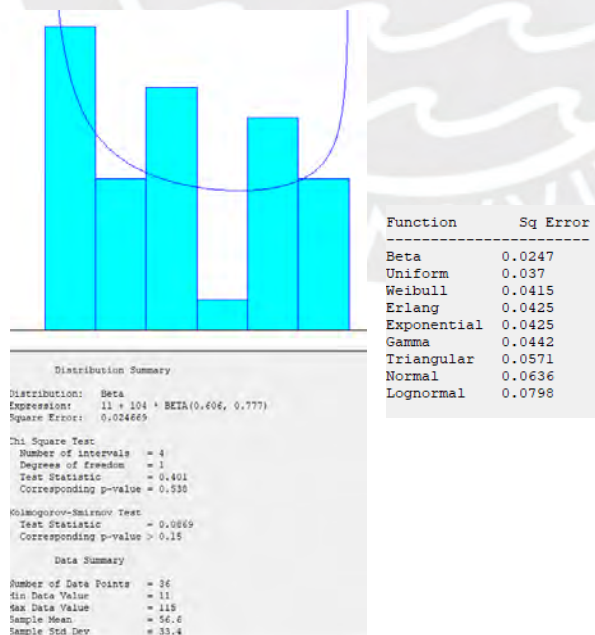
Variable	Descripción
R1-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 00:00 a 02:00 h
R1-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 00:00 a 02:00 h
R2-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 03:00 a 07:00 h
R2-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 03:00 a 07:00 h
R3-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 08:00 a 12:00 h
R3-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 08:00 a 12:00 h
R4-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 13:00 a 17:00 h
R4-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 13:00 a 17:00 h
R5-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 18:00 a 21:00 h
R5-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 18:00 a 21:00 h
R6-D1	Tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 22:00 a 23:00 h
R6-D2	Tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 22:00 a 23:00 h

Anexo 25 Pruebas de bondad de ajuste para las variables de tiempo entre llegadas de aeronaves y tiempo de atención de aeronaves

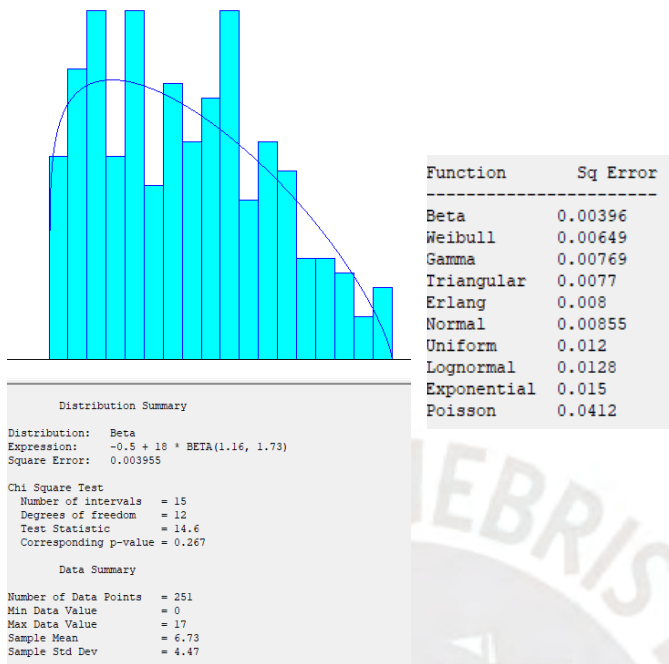
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 00:00 a 02:00 h (R1-D1)



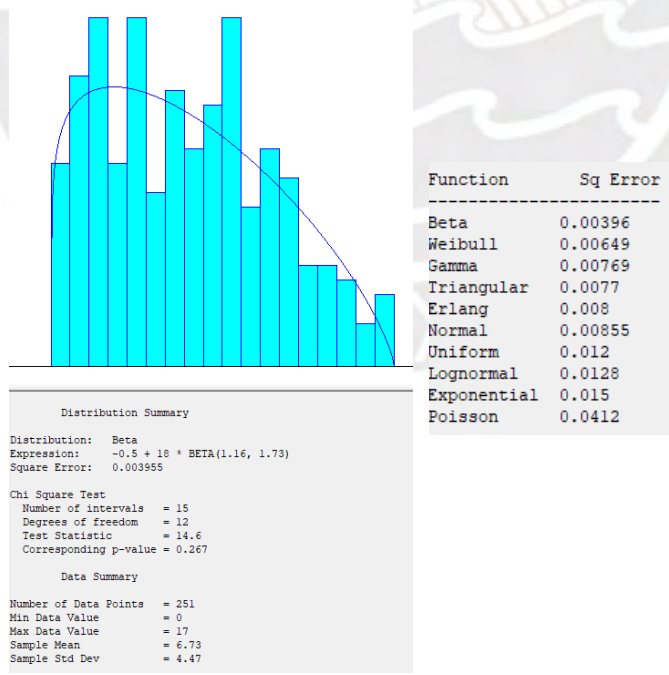
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 00:00 a 02:00 h (R1-D2)



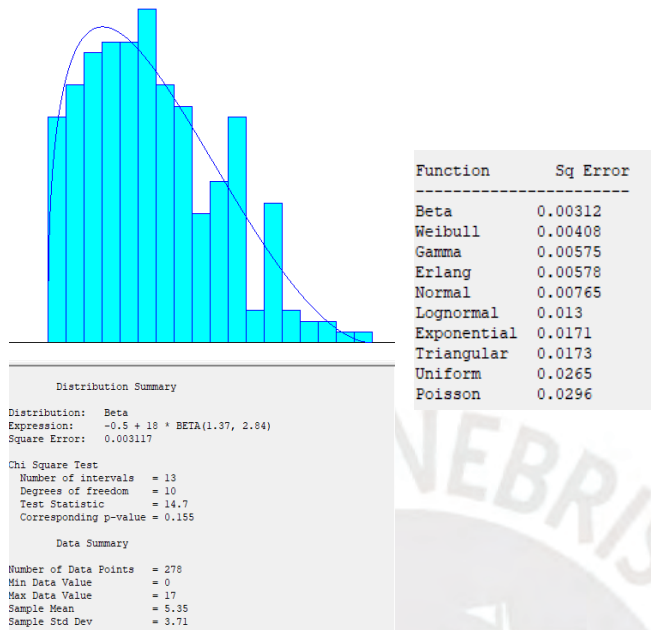
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 03:00 a 07:00 h (R2-D1)



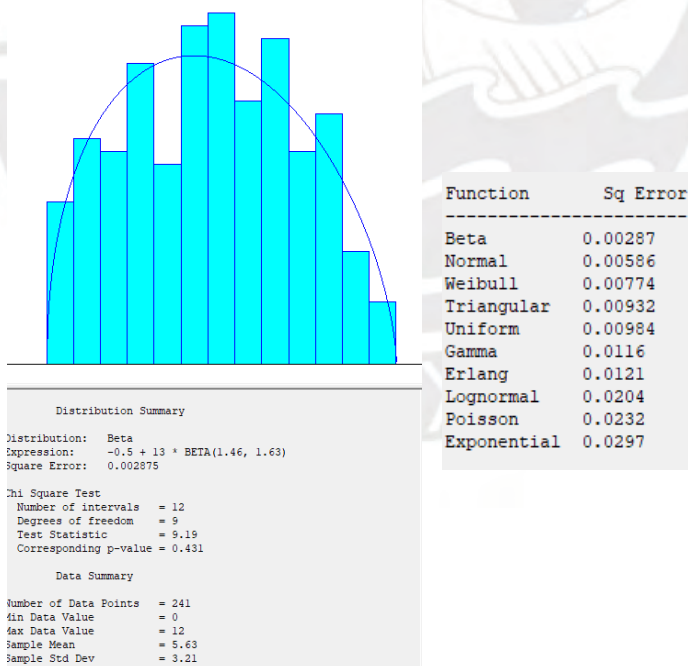
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 03:00 a 07:00 h (R2-D2)



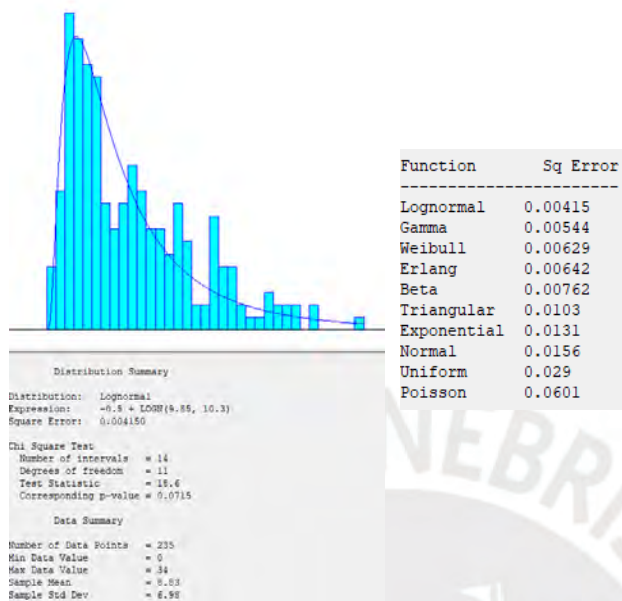
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 08:00 a 12:00 h (R3-D1)



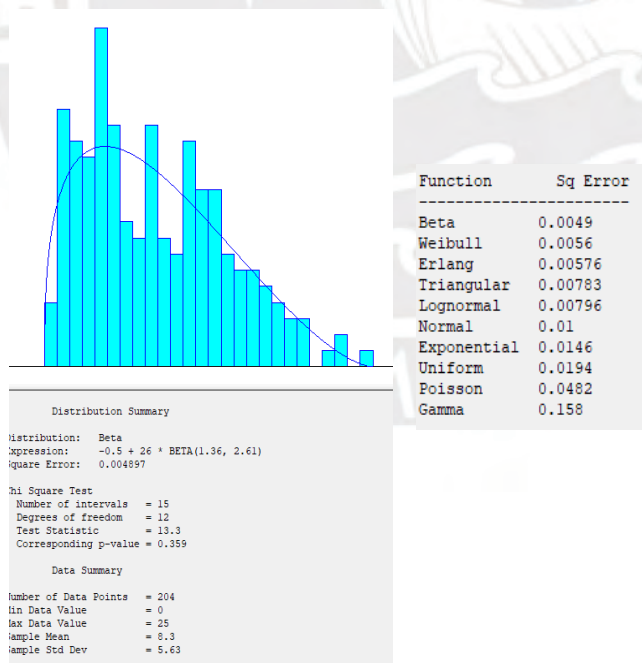
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 08:00 a 12:00 h (R3-D2)



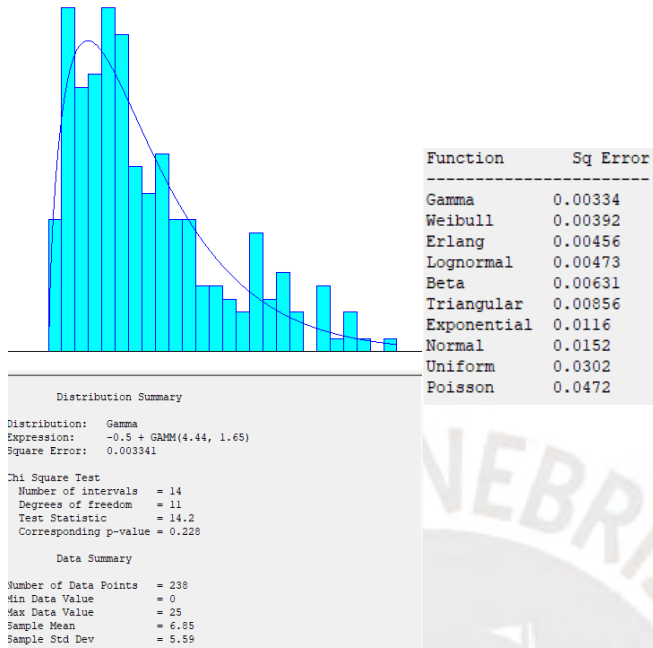
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 13:00 a 17:00 h (R4-D1)



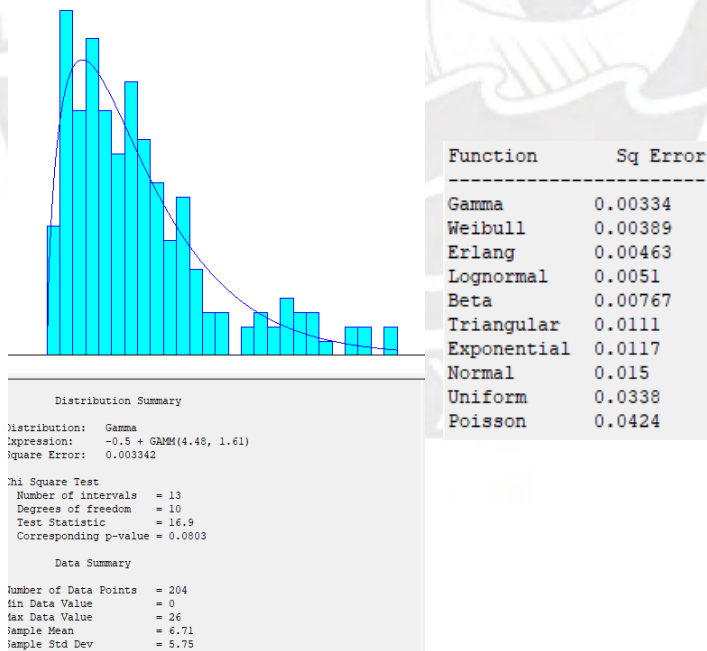
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 13:00 a 17:00 h (R4-D2)



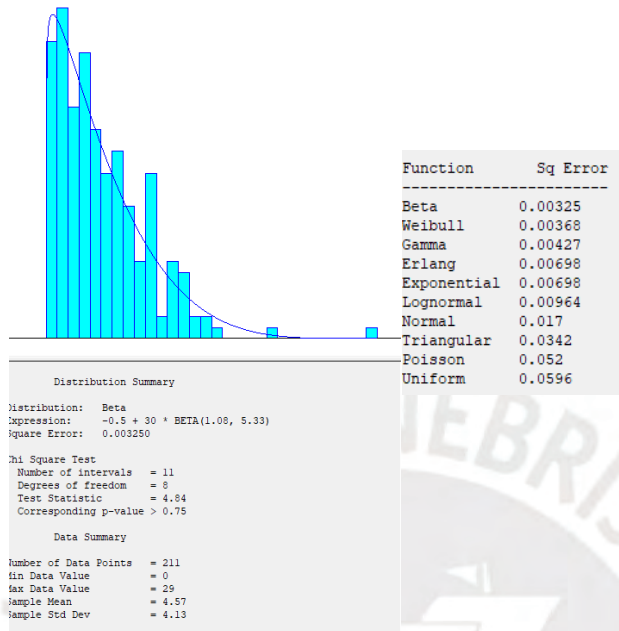
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 18:00 a 21:00 h (R5-D1)



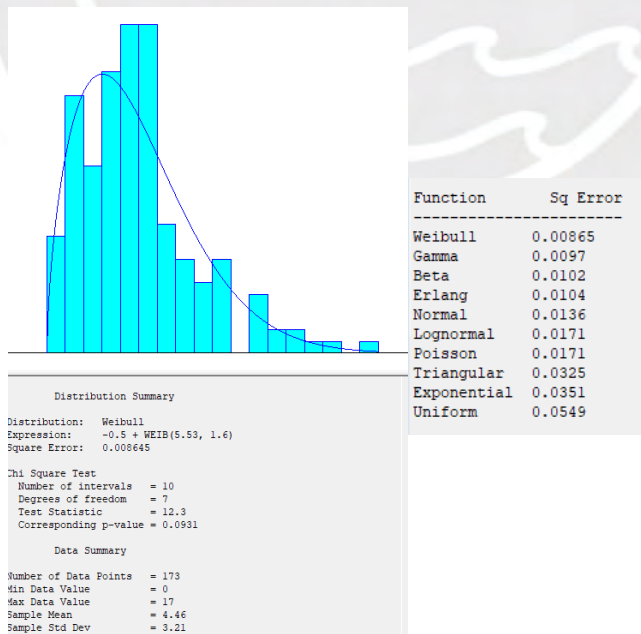
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 18:00 a 21:00 h (R5-D2)



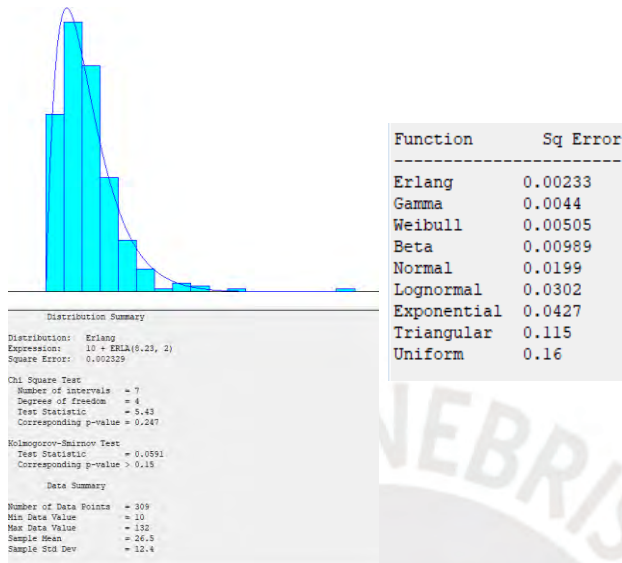
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días domingo, lunes, viernes y sábado durante las franjas horarias de 22:00 a 23:00 h (R6-D1)



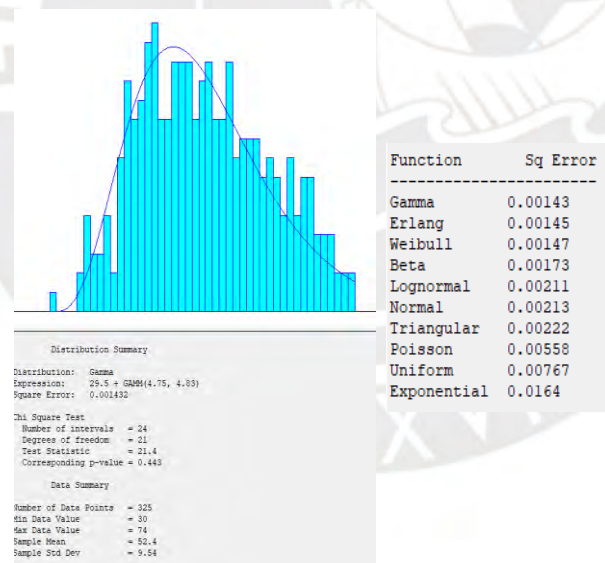
Prueba de bondad de ajuste para el tiempo entre llegadas de aeronaves los días martes, miércoles y jueves durante las franjas horarias de 22:00 a 23:00 h (R6-D2)



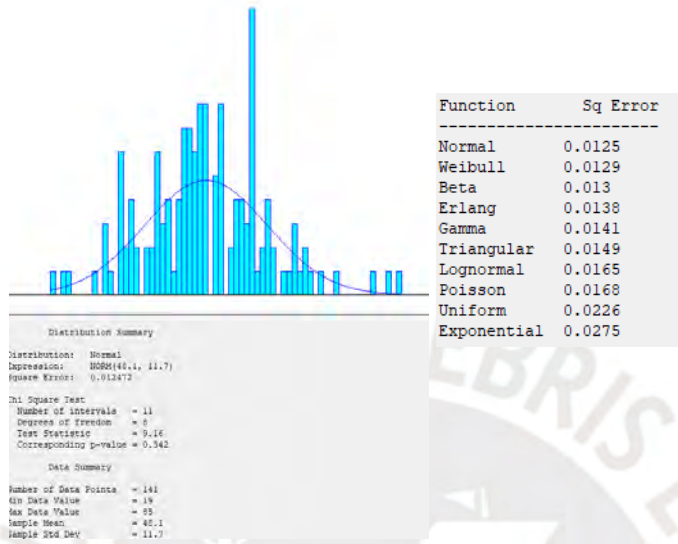
Prueba de bondad de ajuste para el tiempo de atención de una recepción para una aeronave de fuselaje NB (F1-R)



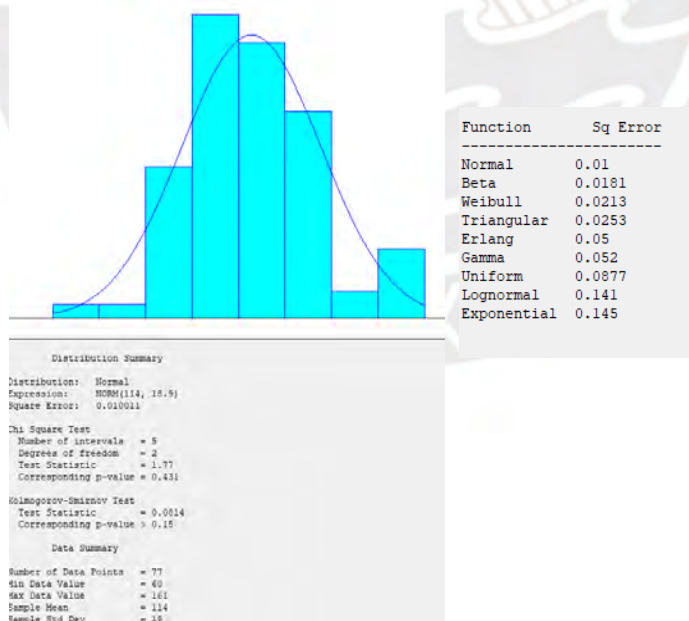
Prueba de bondad de ajuste para el tiempo de atención de una recepción para una aeronave de fuselaje NB (F1-T)



Prueba de bondad de ajuste para el tiempo de atención de una recepción para una aeronave de fuselaje WB (F2-R)



Prueba de bondad de ajuste para el tiempo de atención de un tránsito para una aeronave de fuselaje WB (F2-T)



Anexo 26 Resumen de distribuciones por variable aleatoria

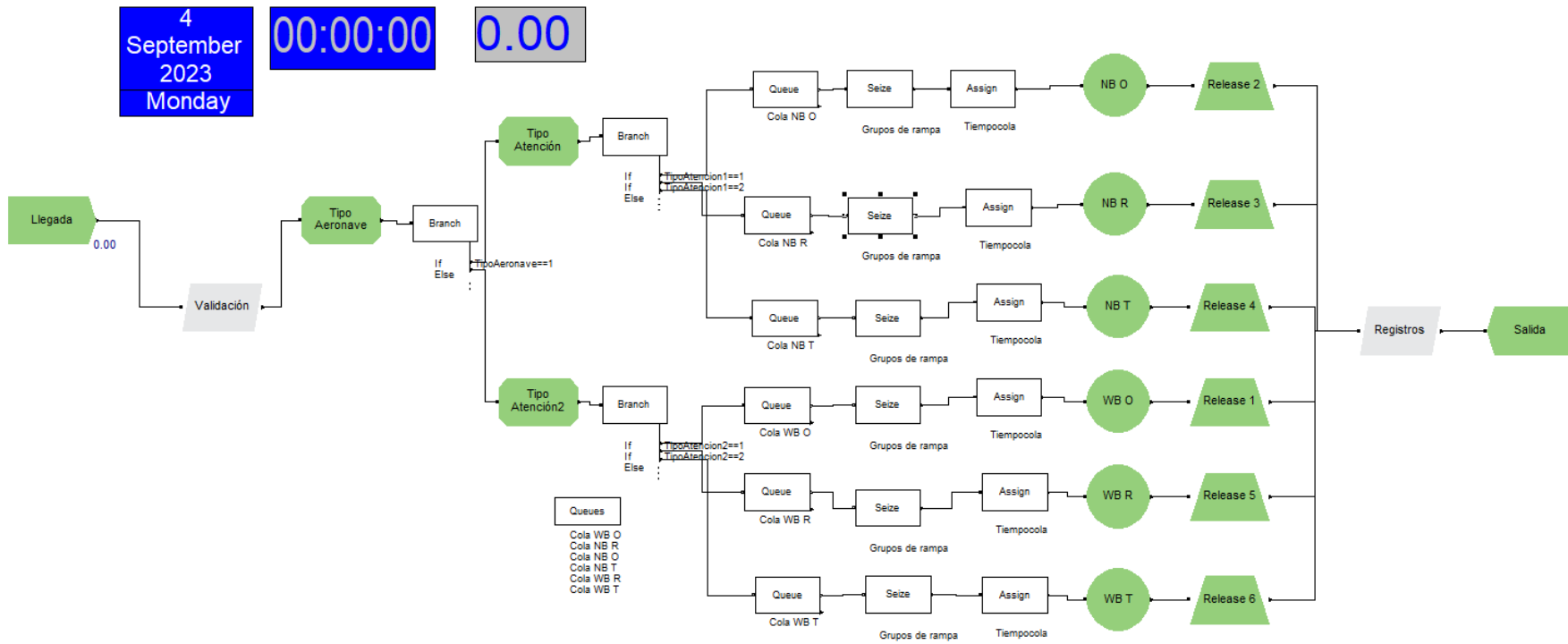
Variable	Expresión
F1-O	55
F1-R	$10 + \text{ERLA}(8.23, 2)$
F1-T	$29.5 + 52 * \text{BETA}(2.79, 3.54)$
F2-O	85
F2-R	$\text{NORM}(48.1, 11.7)$
F2-T	$\text{NORM}(114, 18.9)$
F1-O	$\text{NORM}(114, 18.9)$
R1-D1	$\text{TRIA}(14, 44.3, 129)$
R1-D2	$11 + 104 * \text{BETA}(0.606, 0.777)$
R2-D1	$-0.5 + 18 * \text{BETA}(1.16, 1.73)$
R2-D2	$-0.5 + \text{WEIB}(8.15, 1.97)$
R3-D1	$-0.5 + 18 * \text{BETA}(1.37, 2.84)$
R3-D2	$-0.5 + 13 * \text{BETA}(1.46, 1.63)$
R4-D1	$-0.5 + \text{LOGN}(9.85, 10.3)$
R4-D2	$-0.5 + 26 * \text{BETA}(1.36, 2.61)$
R5-D1	$-0.5 + \text{GAMM}(4.44, 1.65)$
R5-D2	$-0.5 + \text{GAMM}(4.48, 1.61)$
R6-D1	$-0.5 + 30 * \text{BETA}(1.08, 5.33)$
R6-D2	$-0.5 + \text{WEIB}(5.53, 1.6)$



Anexo 27 Bloques del modelo de simulación para la atención de aeronaves

Bloque	Nombre	Función
Create	Llegada	Es el inicio del flujo de la simulación y se formuló con la expresión consolidada de las distribuciones relacionadas de la variable de tiempo entre llegada de aeronaves
Assign 1	Tipo Aeronave	Se le asigna el atributo “TipoAeronave” según el fuselaje de la entidad en base a la proporción histórica entre NB y WB
Branch	Branch 1	El proceso continúa de acuerdo con el tipo de aeronave previamente configurado, es decir, si el tipo de aeronave es NB, la entidad seguirá un flujo distinto a si es de tipo WB
Assign 2	Tipo Destino	En caso la aeronave cumpla con la condición del Branch 1 de tipo NB, entonces se asigna el atributo “TipoDestino” de acuerdo con la proporción histórica entre atenciones nacionales e internacionales
Queue 2	Cola NB DOM	En base al atributo de “TipoDestino”, se agregó una cola si la aeronave es NB y tendrá una atención con salida nacional previo a la asignación del recurso
Queue 3	Cola NB INTER	En base al atributo de “TipoDestino”, se agregó una cola si la aeronave es NB y tendrá una atención con salida internacional previo a la asignación del recurso
Queue 1	Cola WB	En base al atributo de “TipoAeronave”, se agregó una cola si la aeronave es WB y tendrá una atención con salida internacional previo a la asignación del recurso
Seize	Seize 1	Se agregó este bloque para el inicio de la atención de una aeronave WB con destino internacional considerando 1 recurso denominado “Grupos de rampa”
Seize	Seize 2	Se agregó este bloque para el inicio de la atención de una aeronave NB con destino nacional considerando 1 recurso denominado “Grupos de rampa”
Seize	Seize 3	Se agregó este bloque para el inicio de la atención de una aeronave NB con destino internacional considerando 1 recurso denominado “Grupos de rampa”
Delay 1	NB DOM	En este bloque se formuló la expresión de la distribución de la variable de tiempo de atención de aeronaves de tipo de fuselaje NB con destino nacional
Delay 2	NB INT	En este bloque se formuló la expresión de la distribución de la variable de tiempo de atención de aeronaves de tipo de fuselaje NB con destino internacional
Delay 3	WB INTER	En este bloque se formuló la expresión de la distribución de la variable de tiempo de atención de aeronaves de tipo de fuselaje WB con destino internacional
Release	Release 1	Este bloque cumple la función de liberar al recurso “Grupo de rampa” para la atención de una aeronave WB con salida internacional
Release	Release 2	Este bloque cumple la función de liberar al recurso “Grupo de rampa” para la atención de una aeronave NB con salida nacional
Release	Release 3	Este bloque cumple la función de liberar al recurso “Grupo de rampa” para la atención de una aeronave NB con salida internacional
Dispose	Dispose 1	Fin del proceso de simulación

Anexo 28 Sistema de simulación propuesto para la atención de aeronaves en el aeropuerto Jorge Chávez



Anexo 29 Muestra inicial de 15 réplicas sobre la demanda semanal de atenciones

Nº réplica	Nº atenciones
1	1320
2	1356
3	1312
4	1407
5	1409
6	1353
7	1371
8	1381
9	1395
10	1371
11	1329
12	1364
13	1362
14	1407
15	1347

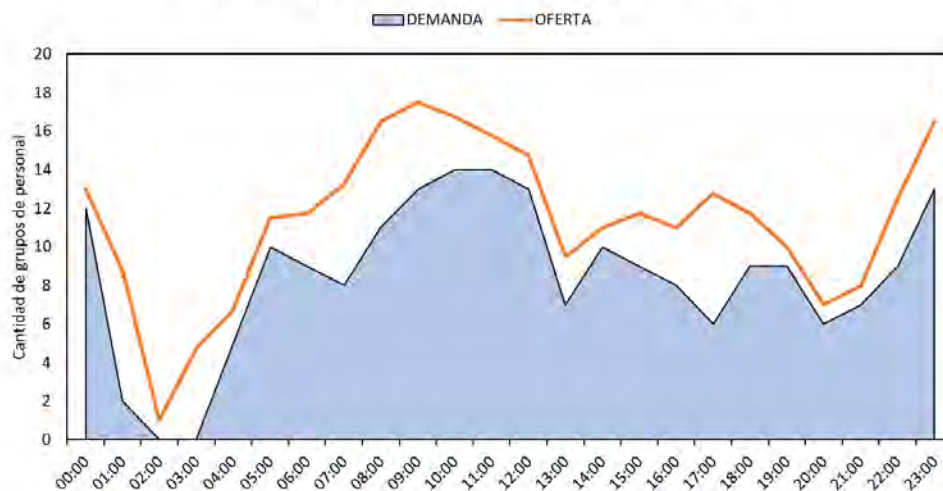


Anexo 30 Detalle de turnos de personal para la semana 1

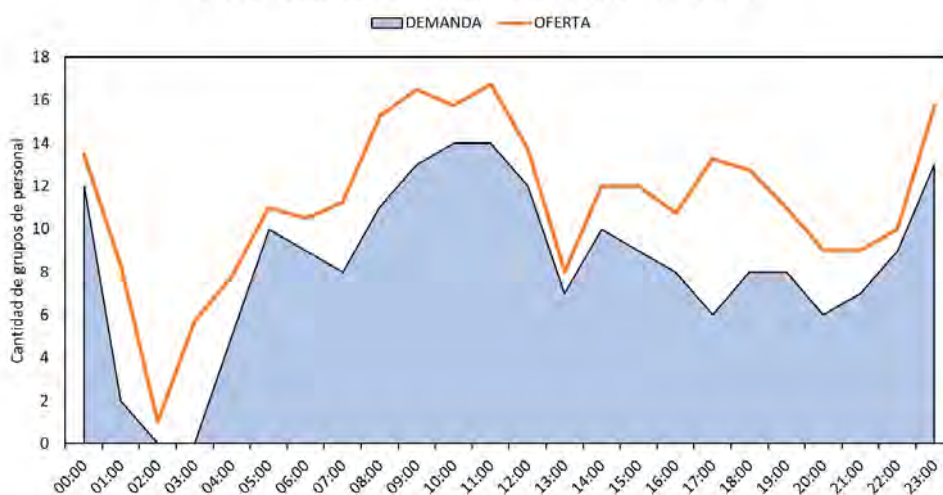
N° Turno	Grupos	Horario	Día Descanso	Hora refrigerio
6	1	03:00 - 12:00	Sábado	09:00 - 10:00
28	1	04:00 - 13:00	Domingo	11:00 - 12:00
29	1	05:00 - 14:00	Lunes	10:00 - 11:00
30	1	05:00 - 14:00	Martes	10:00 - 11:00
31	1	05:00 - 14:00	Miércoles	10:00 - 11:00
33	1	05:00 - 14:00	Viernes	10:00 - 11:00
36	1	05:00 - 14:00	Lunes	11:00 - 12:00
38	2	05:00 - 14:00	Miércoles	11:00 - 12:00
62	2	07:00 - 16:00	Sábado	12:00 - 13:00
67	2	07:00 - 16:00	Jueves	13:00 - 14:00
68	1	07:00 - 16:00	Viernes	13:00 - 14:00
70	2	07:00 - 16:00	Domingo	13:00 - 14:00
72	2	08:00 - 17:00	Martes	13:00 - 14:00
73	1	08:00 - 17:00	Miércoles	13:00 - 14:00
81	1	08:00 - 17:00	Jueves	14:00 - 15:00
85	1	09:00 - 18:00	Lunes	12:00 - 13:00
89	1	09:00 - 18:00	Viernes	12:00 - 13:00
106	1	10:00 - 19:00	Lunes	13:00 - 14:00
174	1	15:00 - 00:00	Sábado	19:00 - 20:00
178	1	15:00 - 00:00	Miércoles	20:00 - 21:00
183	1	16:00 - 01:00	Lunes	20:00 - 21:00
187	2	16:00 - 01:00	Viernes	20:00 - 21:00
190	1	16:00 - 01:00	Lunes	21:00 - 22:00
193	2	16:00 - 01:00	Jueves	21:00 - 22:00
198	2	17:00 - 02:00	Martes	22:00 - 23:00
202	1	17:00 - 02:00	Sábado	22:00 - 23:00
203	2	17:00 - 02:00	Domingo	22:00 - 23:00
267	1	22:00 - 07:00	Lunes	02:00 - 03:00
269	2	22:00 - 07:00	Miércoles	02:00 - 03:00
270	1	22:00 - 07:00	Jueves	02:00 - 03:00
271	1	22:00 - 07:00	Viernes	02:00 - 03:00
279	1	22:00 - 07:00	Sábado	03:00 - 04:00

Anexo 31 Oferta y demanda de la semana 1 aplicando el método propuesto

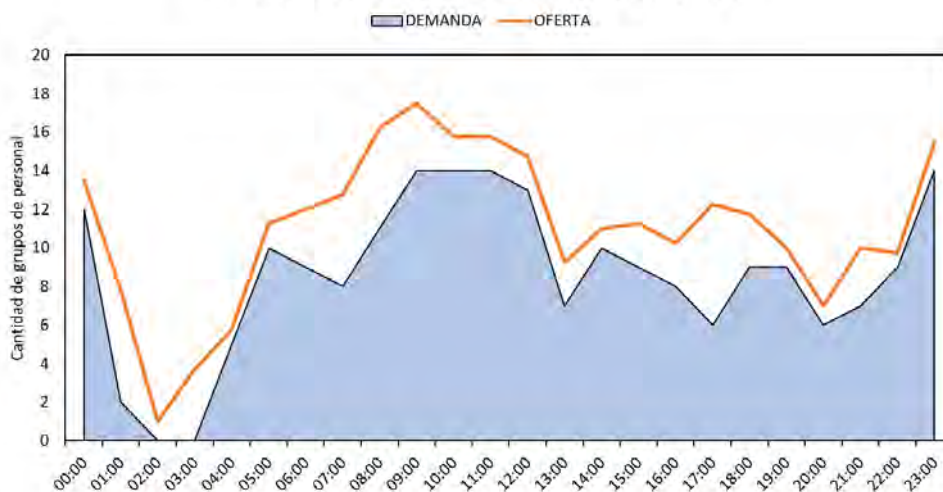
Curva de oferta y demanda, Semana: 1, Día: martes



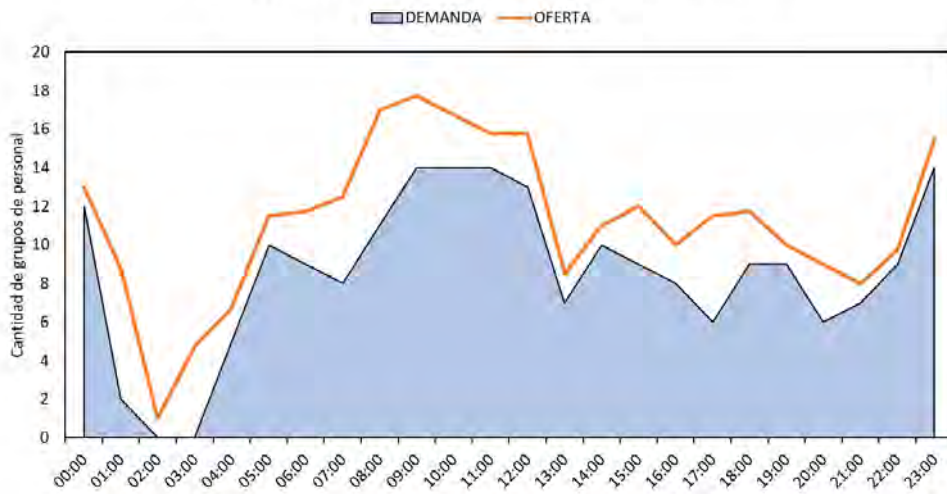
Curva de oferta y demanda, Semana: 1, Día: miércoles



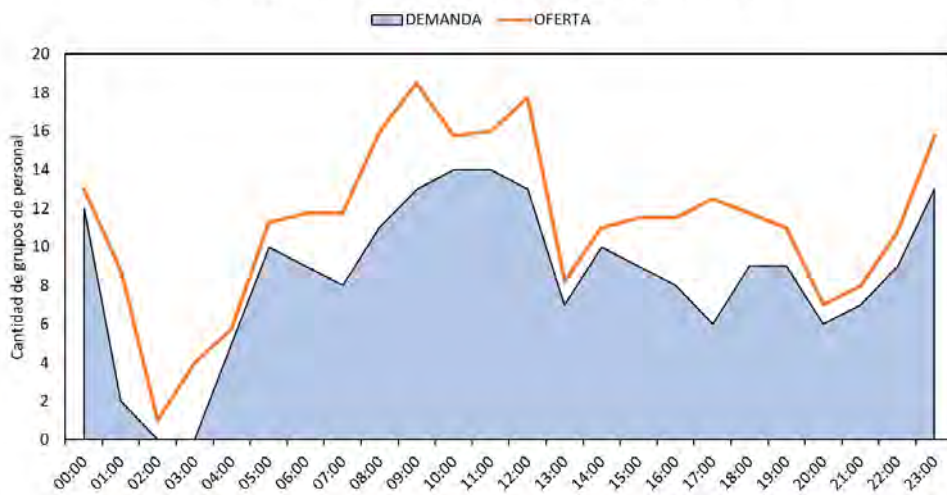
Curva de oferta y demanda, Semana: 1, Día: jueves



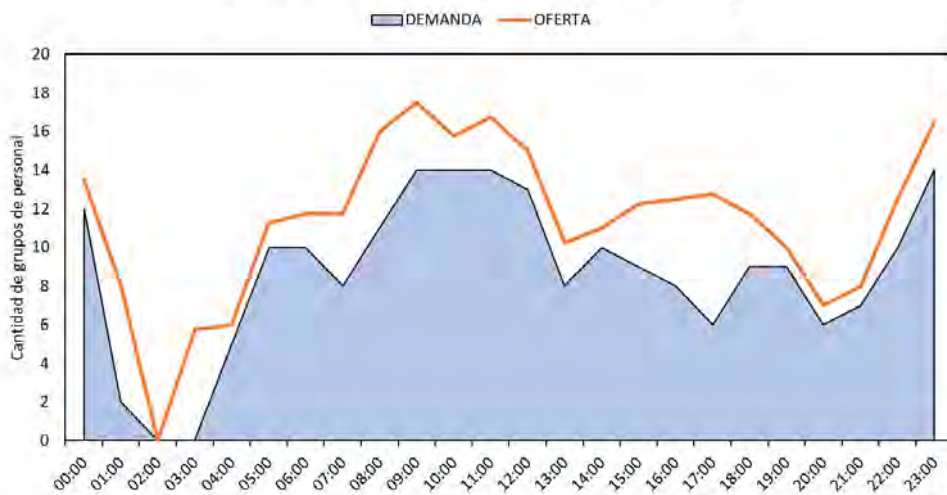
Curva de oferta y demanda, Semana: 1, Día: viernes



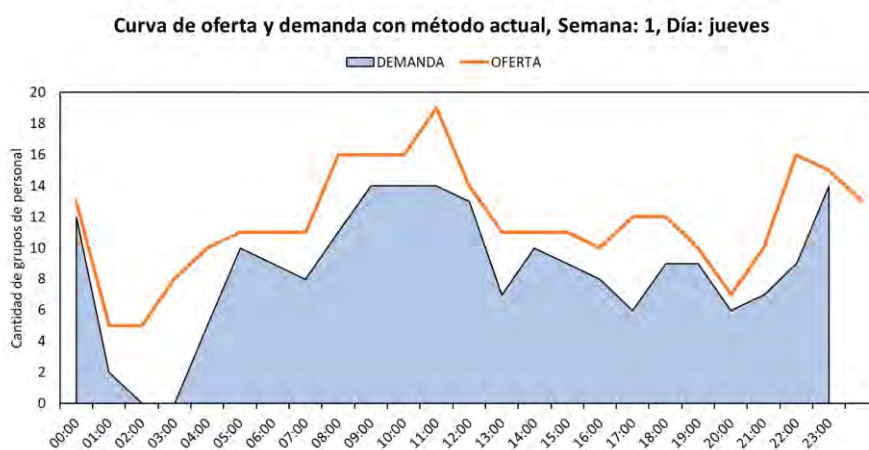
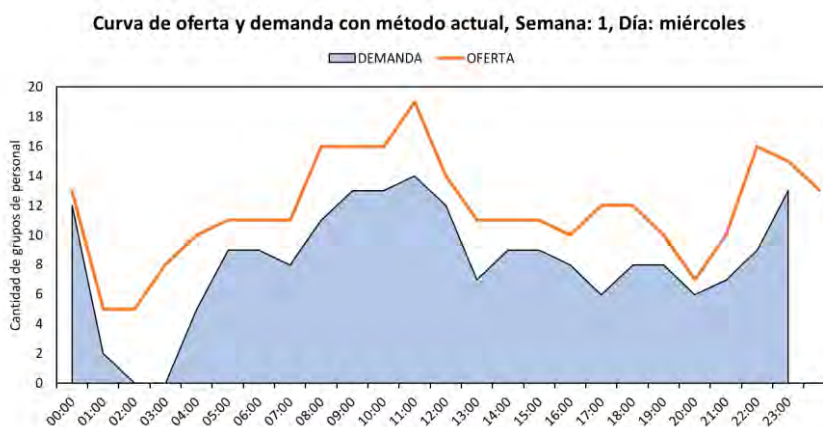
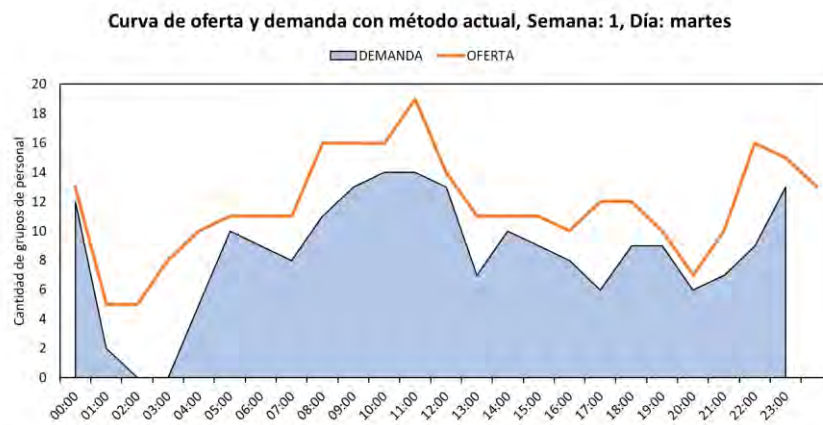
Curva de oferta y demanda, Semana: 1, Día: sábado



Curva de oferta y demanda, Semana: 1, Día: domingo



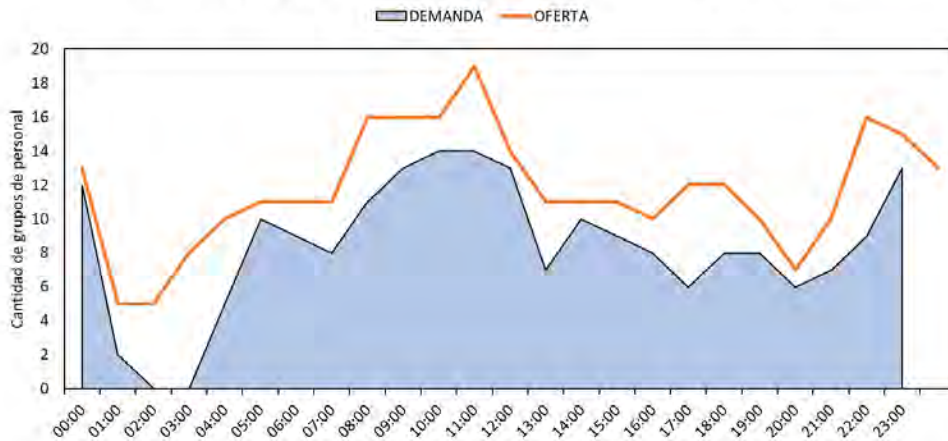
Anexo 32 Oferta y demanda de la semana 1 aplicando el método actual



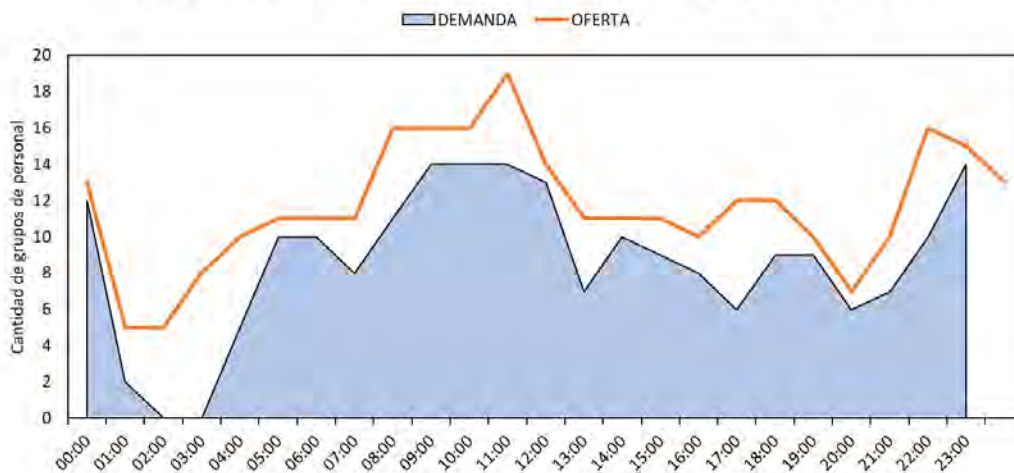
Curva de oferta y demanda con método actual, Semana: 1, Día: viernes



Curva de oferta y demanda con método actual, Semana: 1, Día: sábado



Curva de oferta y demanda con método actual, Semana: 1, Día: domingo



Anexo 33 Validación de resultados de la variable tiempo de entre llegadas de aeronaves

Variable	Promedio	Desv. Estandar	n	Límite inferior	Límite superior	Valor real	¿Cumple?
R1-D1	3.41	0.73	90	3.26	3.56	3.44	Sí
R1-D2	3.80	1.11	90	3.57	4.03	3.64	Sí
R2-D1	40.23	5.43	90	39.11	41.35	40.22	Sí
R2-D2	39.54	5.48	90	38.41	40.67	39.33	Sí
R3-D1	56.69	5.15	90	55.63	57.76	56.00	Sí
R3-D2	53.39	4.33	90	52.49	54.28	53.67	Sí
R4-D1	32.27	6.23	90	30.98	33.56	33.56	Sí
R4-D2	36.14	4.03	90	35.31	36.98	36.08	Sí
R5-D1	34.52	5.12	90	33.46	35.58	34.72	Sí
R5-D2	35.41	5.21	90	34.34	36.49	36.00	Sí
R6-D1	26.03	4.78	90	25.05	27.02	26.06	Sí
R6-D2	26.17	3.64	90	25.42	26.92	26.25	Sí