

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

**MARCO TEÓRICO Y ESTUDIOS DE CASO PARA LA MEJORA EN LA OPTIMIZACIÓN DE LA RED DE
AGENCIAS DE UNA EMPRESA BANCARIA EN LIMA METROPOLITANA**

**Trabajo de investigación para obtener el grado de BACHILLER EN CIENCIAS CON MENCIÓN EN
INGENIERÍA INDUSTRIAL**

AUTOR

Fernando David Briones Gallegos

ASESOR:

Eduardo Carbajal López

Lima, mayo, 2021

Resumen

La investigación toma sustento debido al proceso importante de transformación digital que están afrontando los bancos, lo cual implica una nueva estrategia de canales y educar a sus clientes a usar más aplicativos digitales. Esto es clave si estas organizaciones desean mantener una supervivencia en el mediano plazo debido a que hoy están saliendo nuevos competidores en el mercado. El objetivo de la investigación es identificar las fuentes teóricas que ayuden a plantear la mejor solución para la problemática identificada al momento de realizar un diagnóstico de los procesos en el Banco ABC: mejora del proceso de optimización de canales físicos usando marketing analytics y minería de datos. Como sustentos teóricos, toma como base algoritmos de machine learning de clustering relacionados a los modelos k-means y regresión multivariada.

El procedimiento consiste en investigar en distintas fuentes académicas herramientas de diagnóstico de procesos, herramientas de la propuesta de mejora como conceptos de marketing analytics y minería de datos o algoritmos como regresiones y clustering. Finalmente, se analiza 3 casos que plantean problemáticas similares a la que se desea abordar en distintas industrias para poder comparar metodologías a seguir. Como resultados, se pudo consolidar una lista completa de conceptos sólidos del marco teórico que ayuden a sustentar la solución planteada, además, en los 3 casos planteados se identificó que existe un procedimiento claro de cómo abordar un problema de clustering. Como conclusión principal, se resume en que hoy existe mucha información sobre estos temas y casos prácticos como los que se abordan para poder sustentar cualquier propuesta de marketing analytics para una problemática en específica. Se sugiere a los lectores manejar conceptos teóricos previos de estadística aplicada y algoritmos más sencillos como regresiones lineales para que pueda ser fácilmente entendible la teoría abordada al momento de buscar información de este tipo.

Índice

Introducción	1
Capítulo 1. Marco Teórico	2
Herramientas de diagnóstico	2
Mapeo de procesos	2
Matriz QDF	6
Flujograma	7
Matriz de priorización y ponderaciones pareadas	9
Fichas indicador	12
Diagrama de Pareto	12
Diagramas causa-efecto	13
Diagrama de Cinco Por qué	15
Matris FACTIS	16
Herramientas de la propuesta de mejora	17
Conceptos generales de marketing analytics y minería de datos	17
Regresión lineal	20
Regresión lineal múltiple	20
Regresión logística	21
Clustering	22
Análisis conjunto	23
Técnicas de muestreo	24
Estudio de casos	29
Primer caso	30
Segundo caso	36
Tercer caso	41
Resumen metodológico	45
Bibliografía	47
Anexos	51

Introducción

Actualmente los bancos se encuentran en un contexto de transformación digital, el cual busca mejorar la experiencia de los clientes con el fin de aumentar las tasas de retención en un mercado de alta competencia que no ofrece productos diferenciados y donde la barrera de entrada se vuelve cada vez más débil debido al ingreso de nuevos competidores como las ‘fintech’ o ‘bigtech’. La banca peruana no es ajena a esta tendencia y es uno de los principales rubros del país que está afrontando cambios disruptivos en cuanto a cómo interactúan los clientes con las empresas. Según, la Asociación de Bancos del Perú, el 2018 fue el primer año donde las transacciones en el canal banca móvil (58.06 millones) superó a las transacciones realizadas en las ventanillas de las sucursales físicas (56.63 millones). Además, las transacciones en banca móvil, banca por internet, POS, cajeros automáticos y agentes han crecido en un 107.4%, 18.4%, 21.52%, 9.2%, 16.36% a comparación del 2017¹. Esto refleja el hecho de que cada vez son más los clientes que dejan de usar los canales tradicionales y optan por canales alternativos y/o digitales. Cabe resaltar que, en promedio, un canal digital permite un ahorro de costos de 80% de lo que cuesta un canal físico, por lo que cada vez son más las empresas que apuestan por canales más costo-eficientes.²

A partir de lo mencionado, surge la necesidad de replantear la estrategia de oferta de canales del Banco ABC para mejorar los ratios de eficiencia y experiencia del cliente. Por ello, el presente trabajo de investigación realiza un diagnóstico y propuesta de mejora en la optimización de la red de agencias de la empresa usando herramientas de marketing analytics y minería de datos con el

¹ ASBANC “Preferencia por canales virtuales sigue creciendo entre usuarios bancarios” (2018). Lima, Perú. Febrero 2018. Consulta: 07/04/2019. Véase:

<<https://www.asbanc.com.pe/Paginas/Noticias/DetalleNoticia.aspx?ItemID=762>>

² PERÚ 21. “Uso de canales digitales ahorran hasta 80% en la interacción entre empresa y cliente” (2018). Perú 21. Septiembre 2018. Consulta: 07/04/2019. Véase:

<<https://peru21.pe/peru/canales-digitales-ahorran-80- interaccion-empresa-cliente-428049>>

fin de establecer una lógica omnicanal y visión a largo plazo que permita conocer mejor el comportamiento de los clientes y con ello, plantear las estrategias adecuadas.

En este primer capítulo de marco teórico, se describe cada herramienta o concepto usado en la presente investigación de forma concisa y con un ejemplo para un mayor entendimiento. Asimismo, se realizó un breve resumen de tres casos de estudios similares al tema abordado. Finalmente, se concluye con un resumen metodológico.

Marco Teórico

En el presente capítulo, en base a cada bibliografía consultada, se detallarán los conceptos teóricos y herramientas usadas en la presente investigación. En primer lugar, se describirán las herramientas que se usarán para el diagnóstico de la empresa y sus procesos. Luego, se describirán los conceptos teóricos, herramientas y software a usar en la propuesta de mejora planteada. Finalmente, se presentaran 3 casos de estudio a fines al tema de investigación y un resumen breve de la metodología a usar en cada capítulo.

1.1 Herramientas de diagnóstico

A continuación, se describirá las herramientas usadas para diagnosticar los procesos de la empresa y poder determinar la o las contramedidas a usar en base a las causas raíces encontradas. Es importante resaltar la importancia del adecuado uso de cada herramienta para poder llegar a la solución óptima ya que todas son dependientes entre sí.

1.1.1 Mapeo de procesos.

Según Cadena (2015), los procesos son un conjunto de actividades interrelacionadas que utilizan los elementos de entrada para transformarlos en elementos de salida que otorgan valor a los productos o servicios ofrecidos por las empresas. Asimismo, estos tienen los siguientes elementos:

- Entradas: Recursos que ingresan al sistema para ser transformados.

- Mecanismos o recursos: Herramientas o máquinas que permiten la transformación de las entradas en salidas.
- Salidas: Resultado de la actividad de un proceso.
- Controles: Regulan el funcionamiento del proceso.
- Límites del proceso: Determinan el inicio y el fin del proceso.

A continuación, se muestra la jerarquía de procesos, la cual depende de que tan complejo son los procesos del sistema estudiado. En la imagen 1.1, se puede observar una representación de esta jerarquía. Cabe resaltar que, en base a esto último, se decidirá hasta que nivel realizar el desglose de los procesos para no caer en redundancias.

- Procesos nivel 1 o macroprocesos: Son una representación global de los procesos principales de la empresa.
- Procesos nivel 2 o procesos: Son el desglose de un proceso nivel 1.
- Procesos nivel 3 o subprocessos: Son el desglose y detalle de un proceso nivel 2.
- Actividades: Detallan todas las actividades involucradas en cada uno de los procesos nivel 3.

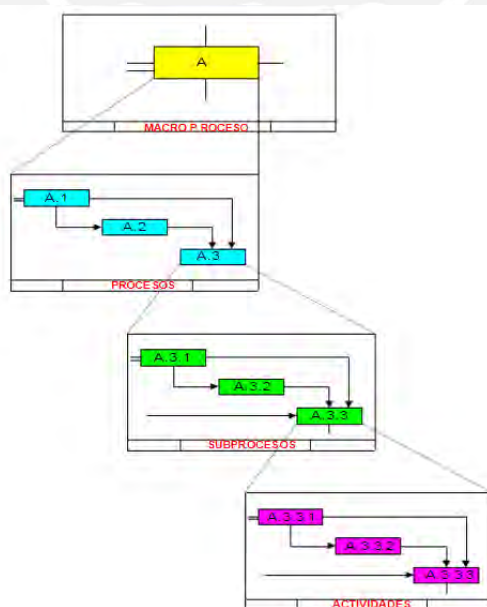


Imagen 1.1. Jerarquía de procesos de una organización. Fuente: Cadena (2015).

Por otro lado, Cadena también indica que estos procesos se clasifican en tres tipos, los cuales son: procesos estratégicos, procesos operativos o misionales y procesos auxiliares. En primer lugar, los procesos estratégicos son los responsables de que la organización pueda alcanzar sus objetivos a largo plazo y visión. En segundo lugar, los procesos operativos o misionales son los procesos responsables de realizar la acción de transforman las entradas en salidas (la razón de ser de la empresa). Por último, los procesos auxiliares brindan apoyo y soporte para que los procesos operativos puedan darse sin dificultad alguna y exista un adecuado funcionamiento del sistema. En la imagen 1.2 puede observarse un esquema general del mapa de procesos.

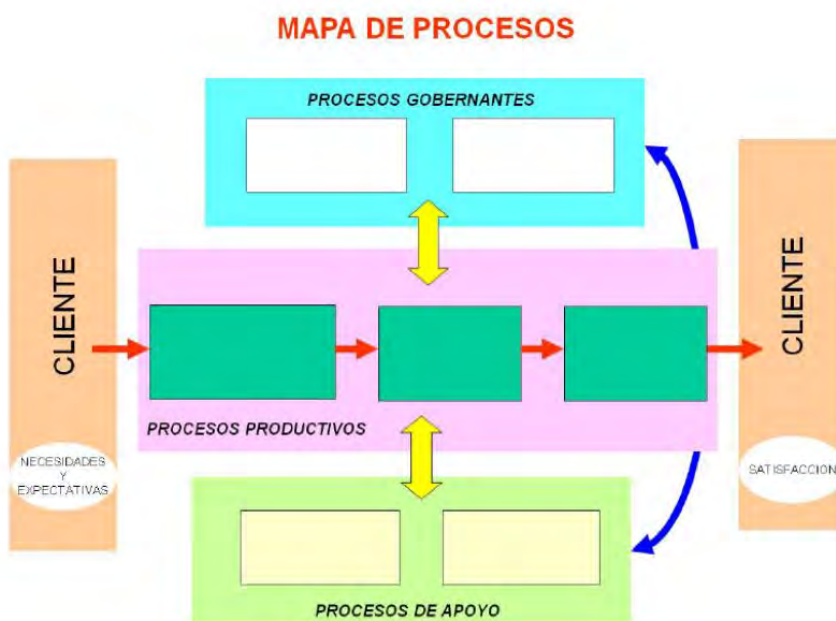


Imagen 1.2. Mapa de macroprocesos. Fuente: Cadena (2015).

Finalmente, en la imagen 1.3 puede observar un ejemplo de mapeo de macroprocesos, realizado por ESSALUD. Podemos observar que los procesos misionales, tales como la gestión de aseguramiento de salud y prestaciones transforman los requerimientos del gobierno, ciudadanos y grupos de interés en clientes y grupos de interés atendidos gracias a la dirección estratégica de los procesos estratégicos, como la gestión de calidad, que brinda se encarga de realizar las planificaciones necesarias para lograr que el servicio cumpla los estándares de calidad definidos por la organización, y gracias a los procesos de soporte, como la gestión logística, que permite que los medicamentos puedan llegar a las instalaciones que lo requieran.

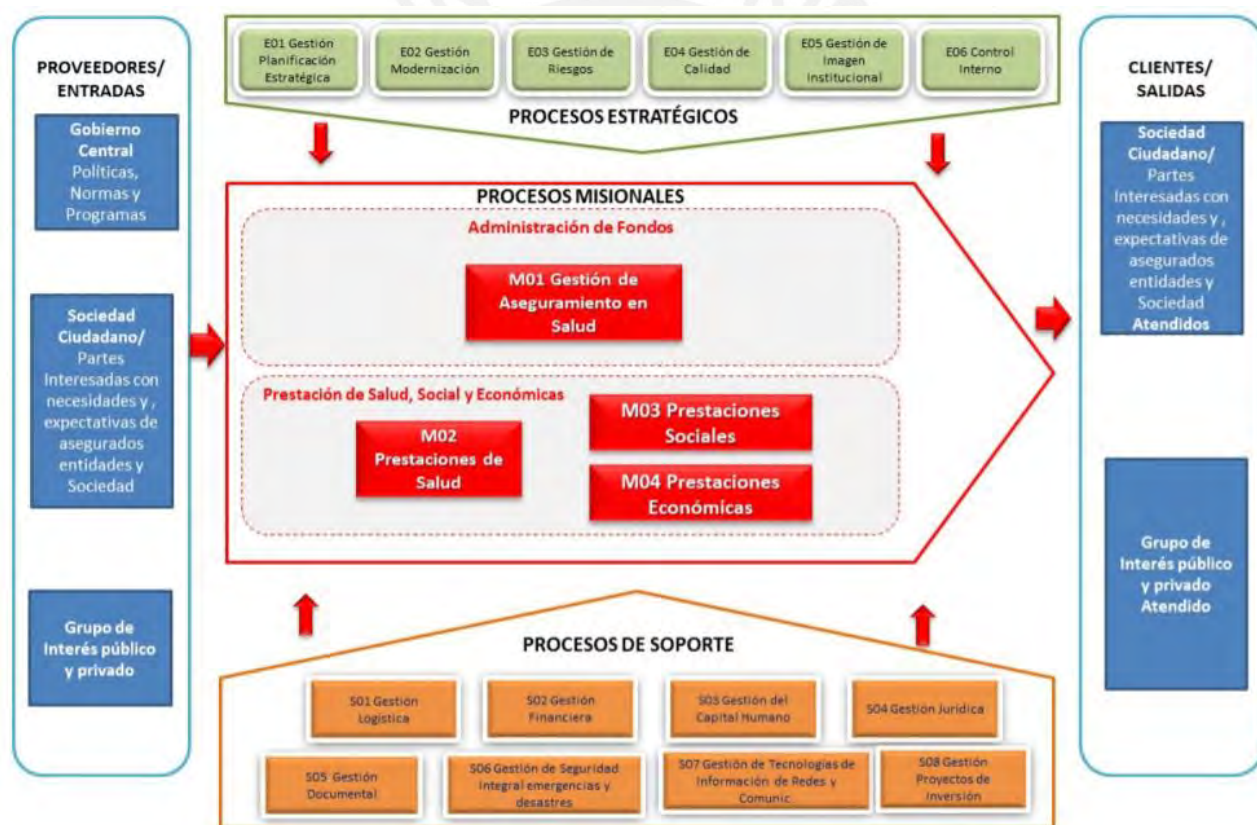


Imagen 1.3. Mapa de macroprocesos de ESSALUD. Fuente: Mapa de macroprocesos ESSALUD (2019).

1.1.2 Matriz QDF.

Según Bernal (2009), “*Quality Function Deployment (QFD) es un sistema con el objetivo de traducir y planificar la voz del cliente en las características de calidad de los productos, procesos y servicios para lograr su satisfacción*”. Como se observa en la imagen 1.4, esta tiene varios componentes, los cuales se describirán a continuación:

- *VOC*: Hace referencia a los requerimientos del cliente. Para determinar las variables a analizar debe realizarse la pregunta ¿qué es lo que desea el cliente?
- *Design Requirements*: Hace referencia a el cómo los requisitos del cliente (*VOC*) serán satisfechos por el diseño del servicio.
- *Relations matrix*: Hace referencia a que tan ligados están los requerimientos del cliente con los diseños del servicio en una escala que será usada para cada par comparado.
- *Benchmarking*: Hace referencia a qué tanto los competidores satisfacen las necesidades del cliente.
- *Importance Level*: Indica que tan relevante son requerimientos del diseño en base a los requerimientos del cliente (suma de resultados obtenidos en la matriz de relación para cada uno).
- *Correlation matrix*: Indica que tanta correlación existe los diseños del sistema. Esto significa que de solo mejorarse uno de dos diseños que estén estrechamente correlacionados, ambos tendrán una mejora significativa.

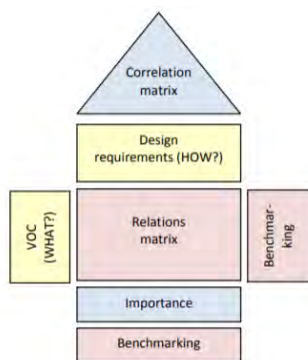


Imagen 1.4. Mapa de macroprocesos de ESSALUD. Fuente: Mapa de macroprocesos ESSALUD (2019).

En la imagen 1.5, se muestra un ejemplo donde se selecciona el proceso o “diseño” más importante en el servicio de seguridad que ofrece el estado a los ciudadanos.

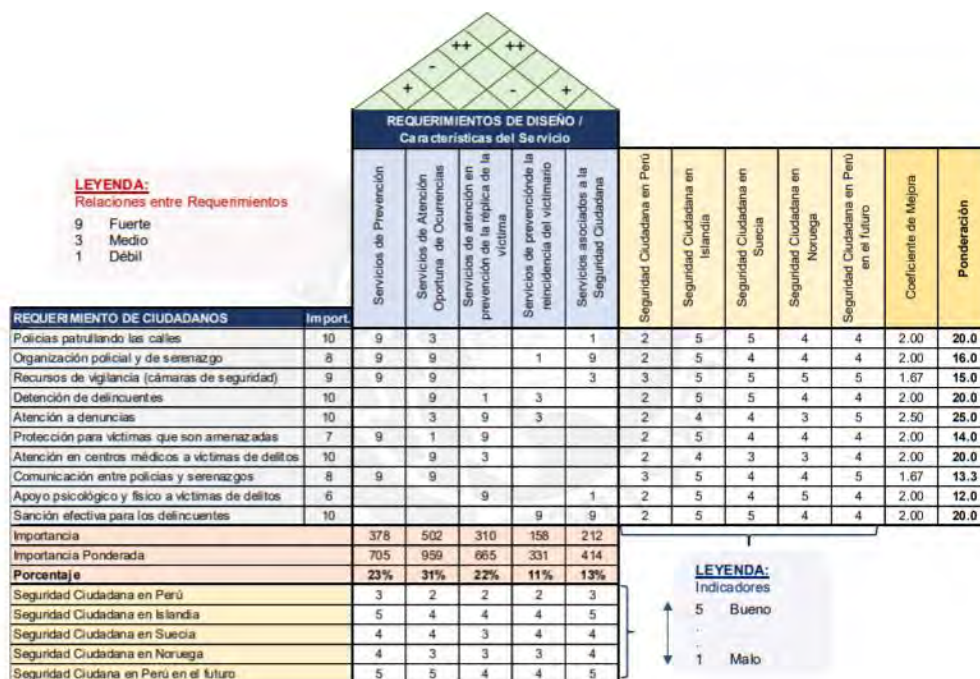


Imagen 1.5. QDF del proceso principal del servicio de seguridad. Fuente: Sotelo (2019).

1.1.3 Flujograma.

Según Pardo (2012), “la representación gráfica de un proceso, como un ente individual, se puede realizar mediante una herramienta denominada diagrama de flujo o flujograma”. Los principales símbolos para elaborar un flujograma son los óvalos, rectángulos, rombo, flecha, símbolos de entrada y salida y los conectores. El detalle de cada uno de estos se muestra a continuación en la imagen 1.6.



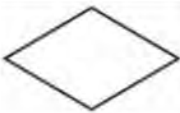


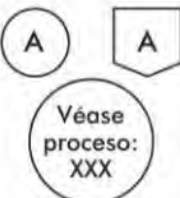
Símbolo	Nombre	Descripción
	Elipse u óvalo	Indica el inicio y el final del diagrama de flujo. Está reservado a la primera y a la última actividad. Un proceso puede tener varios inicios y varios finales
	Rectángulo o caja	Se utiliza para definir cada actividad o tarea. Debe incluir siempre un verbo de acción. Las cajas se pueden numerar
	Rombo	Aparece cuando es necesario tomar una decisión. Incluye siempre una pregunta
	Flecha	Utilizada para unir el resto de símbolos entre sí, indicando la dirección secuencial de las actividades
	Símbolos de entrada y salida	Se utilizan para representar entradas necesarias para ejecutar actividades del proceso, o para recoger salidas generadas durante el desarrollo del mismo
	Conectores	Usados para representar conexiones con otras partes del flujograma o con otros procesos. Si el proceso es largo y el diagrama de flujo no cabe en una hoja, se suele utilizar algún símbolo para conectar una hoja con otra. Una letra o un número en el interior del símbolo indican que la secuencia enlaza con un símbolo equivalente. También se pueden utilizar para vincular el proceso que estamos dibujando con otro proceso relacionado

Imagen 1.6. Simbología usada en los flujogramas. Fuente: Pardo (2012).

Algunas de las ventajas que menciona el autor de elaborar un flujograma son las siguientes:

- Útil al momento de establecer indicadores.
- Permite observar de forma visual el proceso.
- Permite identificar problemas y oportunidades de mejora en conjunto con otras personas.
- Permite definir de forma clara las responsabilidades de cada integrante de un equipo.
- Buena alternativa para la documentación de procesos.

Finalmente, en la imagen 1.7. se observa un ejemplo de un flujograma que el autor brinda sobre el proceso de contratación de nuevo personal para un proyecto, el cual empieza en la necesidad de contratar un equipo de proyecto y culmina en la firma del contrato.

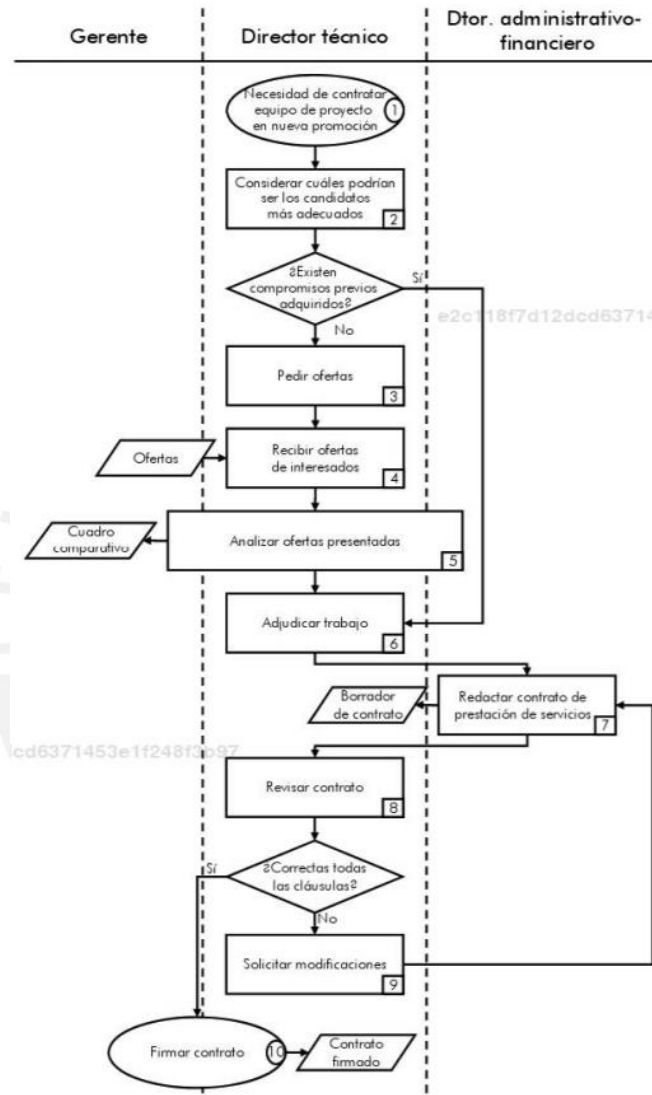


Imagen 1.7. Flujograma del proceso de contratación de personal de un equipo de proyecto.

Fuente: Pardo (2012).

1.1.4 Matriz de priorización y matriz de ponderación pareada.

Según AITECO (2019), “la matriz de priorización es una herramienta que permite la selección de opciones sobre la base de la ponderación y aplicación de criterios”. Esta permite evaluar distintas alternativas en base a la ponderación que se otorga a cada factor planteado para cuantificar cada una y determinar la mejor opción para tener un sustento en la toma de decisiones de una empresa.

a) Matriz de ponderación pareada

En primer lugar, se debe realizar una matriz de ponderación pareada para no caer en métodos subjetivos para determinar el peso que tendrá cada uno de los factores. Un ejemplo de esto puede observarse en la imagen 1.8, donde se seleccionó 5 criterios a evaluar y se realiza una comparación de todos contra todos, como es el caso del factor personal motivado y el impacto social, en la que se considera una relación de 5 a 1 (por ello, se otorga un puntaje de 1/5 en la casilla de posición espejo). Una vez evaluado cada pareja de factor, se procede a determinar el factor de ponderación en base al puntaje total asignado.

PONDERACIÓN DE LOS CRITERIOS	Impacto social	Procesos clave	Personal motivado hacia la mejora	Imagen de la institución	Madurez organizativa	TOTAL	Ponderación del Criterio
1 Impacto social		1,00	5,00	5,00	5,00	16,00	0,29
2 Procesos clave	1,00		5,00	10,00	0,20	16,20	0,29
3 Personal motivado hacia la mejora	0,20	0,20		0,20	1,00	1,60	0,03
4 Imagen de la institución	0,20	0,10	5,00		10,00	15,30	0,28
5 Madurez organizativa	0,20	5,00	1,00	0,10		6,30	0,11
TOTALES	1,60	6,30	16,00	15,30	16,20	55,40	1,00

Imagen 1.8. Matriz de ponderación pareada. Fuente: AITECO (2019).

b) Matriz de priorización.

En segundo lugar, realiza una matriz de priorización donde se visualiza las opciones a evaluar y se asigna un criterio en una escala definida. Estos valores serán multiplicados por cada criterio definido en la matriz pareada para finalmente, asignar el puntaje total a cada opción evaluada. En la imagen 1.9, se muestra un ejemplo donde se tiene 5 factores en columnas y 5 opciones en filas, resultando finalmente el servicio de atención al ciudadano como la mejor opción.

MATRIZ RESUMEN	SERVICIO DE URBANISMO	SERVICIO DE ATENCIÓN AL CIUDADANO	SERVICIO DE MANTENIMIENTO URBANO	SERVICIO DE DEPORTES	SERVICIO ECONÓMICOS
1 Impacto social	0,13	0,08	0,05	0,03	0,00
2 Procesos clave	0,13	0,09	0,02	0,05	0,00
3 Personal motivado hacia la mejora	0,00	0,01	0,00	0,01	0,01
4 Imagen de la institución	0,03	0,12	0,08	0,05	0,00
5 Madurez organizativa	0,00	0,05	0,00	0,04	0,02
TOTALES	0,29	0,35	0,15	0,18	0,03

Imagen 1.9. Matriz de priorización. Fuente: AITECO (2019).

c) *Matriz de evaluación de riesgos*

Según SUSALUD (2016), otro tipo de matriz de priorización, es la matriz en la que se califica a cada alternativa por las variables probabilidad e impacto en una misma escala. El procedimiento a realizar consiste en la multiplicación de ambos criterios para obtener un puntaje. A mayor sea el puntaje asignado, mayor será la valorización de la alternativa.

En la imagen 1.10, se puede observar una evaluación que realizó una entidad de salud, evaluando cada componente entre varios factores. En este caso, los autores realizaron un promedio simple de los impactos para calcular el “Nivel de Impacto²”. Finalmente, multiplicaron este impacto con la probabilidad de

ocurrencia, obteniendo la columna “Nivel de exposición al Riesgo2”, donde se tiene un empate técnico entre la primera y tercera opción.

I. Formalización Sanitaria								
Verificador	Impacto Legal2	Impacto sobre la salud del Asegurado2	Impacto Económico2	Impacto Operacional2	Impacto de Imagen2	Nivel de impacto2	Probabilidad de Ocurrencia del Riesgo2	Nivel de exposición al Riesgo2
1.1	3	4	3	3	3	3	2	6
1.2	3	3	2	2	2	2	2	4
1.3	3	2	3	3	3	3	2	6

Imagen 1.10. Matriz de priorización. Fuente: SUSALUD (2016).

1.1.5 Fichas Indicador.

Las fichas indicador son una herramienta que permite visualizar de forma gráfica la evolución de un indicador y a su vez, poder determinar en qué rango de cumplimiento respecto a su meta se encuentra, lo cuál es visualizado por el color de la gráfica donde se encuentra la curva del indicador. En el anexo 1.1, se puede visualizar un ejemplo de esta herramienta.

1.1.6 Diagrama de Pareto.

Según Kume (1985), el diagrama de Pareto permite solucionar la problemática de la priorización de problemas con eficiencia a través de la técnica “pocos vitales y muchos triviales”. Esto puede representarse usando una curva de frecuencia acumulada de una variable, por ejemplo, cantidad de errores o costo monetario total por cada problema, para separar los problemas que traen el 80% del total de la variable evaluada, los cuales vendrían a ser los “pocos vitales” del sistema analizado. Sin embargo, una vez hecha la separación, siempre debe priorizarse elegir una poca cantidad de problemas a priorizar a pesar de que sean muchos los que se encuentren dentro del 80% (principio de “pocos vitales”). A continuación, se menciona los pasos que deben realizarse para un diagrama de Pareto según el autor:

- Decidir qué problemas se investigarán y que datos
- Registre los datos y calcule totales

- Elaborar lista de frecuencias individuales y acumuladas
- Organice los ítems en orden en base a su frecuencia acumulada
- Dibuje un gráfico donde se pueda observar un diagrama de barras de frecuencia relativa y una curva de frecuencia acumulada (curva de Pareto) por cada ítem.

En la imagen 1.11 se puede observar un ejemplo de dos curvas de Pareto las cuales son antes y después de haber realizado la mejora del proceso. En ambas curvas, se puede apreciar que se muestra en el eje vertical de la izquierda, la cantidad de errores por cada problema y en el eje vertical de la derecha el porcentaje de la frecuencia acumulada de la curva de Pareto. En un inicio, siguiendo los principios de 80-20% y pocos vitales, los problemas A y B son los más críticos, y posteriormente, se observa que al eliminar el problema A, los problemas B y C se convierten en los problemas más críticos del sistema.

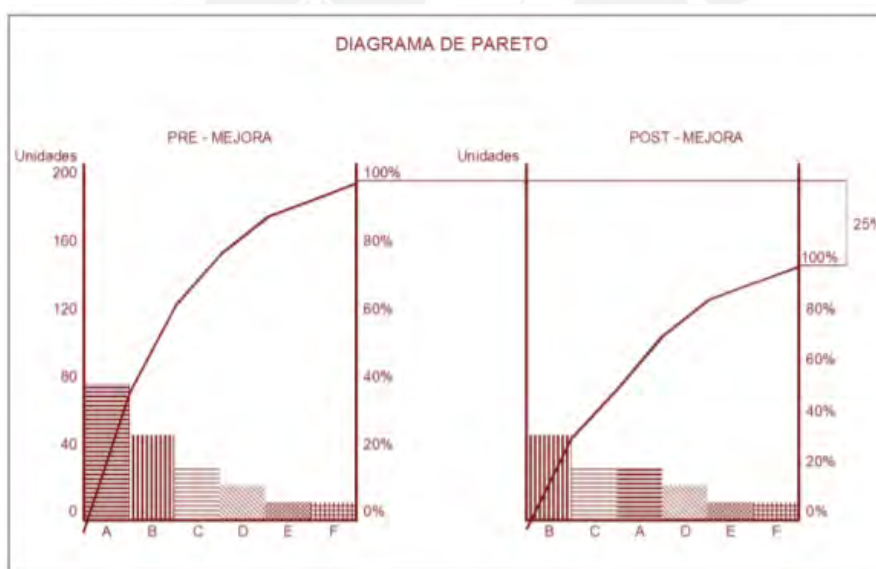


Imagen 1.11. Diagrama de Pareto de la cantidad de errores por problema antes y después de la mejora de un proceso. Fuente: Paz, Aragón (2018).

1.1.7 Diagrama causa-efecto

Según Kume (1985), los diagrama causa-efecto son “*diagramas que muestra la relación entre una característica de calidad y los factores*”. Este método permite analizar resultados complejos de un proceso realizando cadenas sistemáticas entre las causas y efectos involucrados con el fin de determinar las causas más relevantes que generan resultados no deseados en el atributo de calidad evaluado. El autor menciona el siguiente procedimiento para realizar un diagrama causa-efecto:

- Seleccionar atributo de calidad.
- Escribir todas las causas que puedan afectar la característica de la calidad.
- Agrupar causas que tengan afinidad entre sí, sean primarias o secundarias, y conectar las que tengan un impacto significativo en la característica de la calidad.
- Asignar un valor de importancia a cada causa.
- Escribir cualquier información que pueda ser de utilidad.

En la imagen 1.12 se puede observar un ejemplo de un diagrama causa-efecto, en el cual se desea evaluar las causas de los errores en los informes de ensayo de un proceso (atributo de calidad) y en base a las causas descritas, se realizó una agrupación en materiales, personal, método y equipo. También puede apreciarse, por ejemplo, que en método se asignó la causa primaria “falta de control” que tiene como causa secundaria a “falta de personal”. Este análisis permite visualizar gráficamente las principales causas para evaluar cada causa individualmente y tomar acciones de corrección.

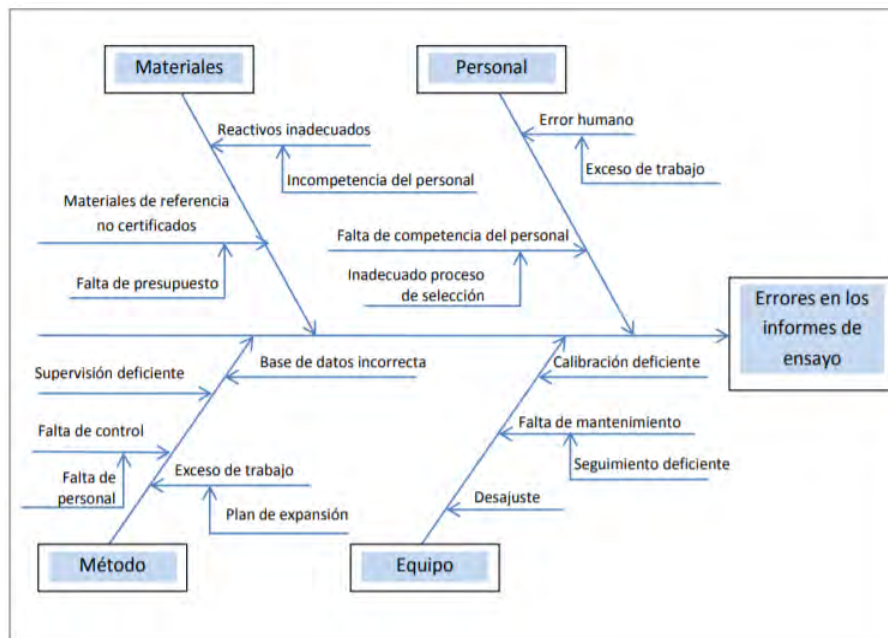


Imagen 1.12. Diagrama de causa-efecto de la cantidad de errores por problema antes y después de la mejora de un proceso. Fuente: Paz, Aragón (2018).

1.1.8 Cinco Porqués.

La herramienta de cinco porqués permite hallar las causas raíces de un problema. Según Serrat (2009), “cuando se busca resolver un problema, es útil comenzar con el resultado final, reflexionar sobre lo que causó eso, y cuestionar la respuesta cinco veces”. Al autor resalta la importancia de responder la pregunta “¿Por qué?” a cada causa analizada durante 5 veces de forma consecutiva y honesta o hasta donde no se caiga en redundancia. La estructura de la metodología puede observarse en la imagen 1.13, la cual sigue los siguientes pasos:

- Definir el problema de forma completa.
- Realizar las 5 preguntas “¿Por qué?”.
- Discutir los resultados para llegar a un consenso sobre cuál sería la ruta del problema y la causa raíz ante distintos posibles resultados.

- Finalmente, realizar acciones para eliminar la causa raíz determinada.

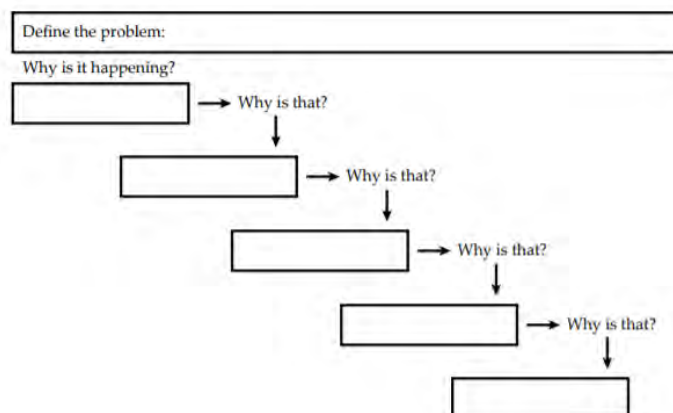


Imagen 1.13. Five Whys Worksheet. Fuente: Serrat (2009).

Asimismo, en el anexo 1.2, se puede visualizar un ejemplo de esta herramienta.

1.1.9 Matriz FACTIS.

La matriz FACTIS es una matriz de priorización que se basa en 5 factores que se adaptan a diversos sistemas y que permiten determinar la solución óptima en base a criterios cuantificables, dando así soporte a la toma de decisiones de la empresa. Los factores que evalúa se manejan en una escala definida por los que evalúan el proceso (no necesariamente la misma escala para cada factor). A continuación, se muestran los factores que se usan en la matriz FACTIS.

- F: Indica que tan fácil se puede implementar la propuesta
- A: Indica que tanto afecta su implementación a otras áreas.
- C: Indica cuanto mejora la calidad del proceso analizado.
- T: Indica que tanto tiempo de implementación que cada opción requiere en caso de realizarse.
- I: Indica que tanta inversión requerida.

- S: Indica que tanta seguridad ofrece la opción analizada a comparación del estado actual.

En la imagen 1.13 se puede observar una matriz FACTIS realizada por la empresa Aceros Arequipa, para evaluar qué problema es el principal en base a los cinco criterios mencionados anteriormente. Cabe resaltar que previamente debe calcularse los factores de ponderación algún un método como el de matriz de comparación pareada.

MATRIZ FACTIS						
N°	PUNTAJE	PRIORIDAD	NOMBRE DE LOS PROBLEMAS A EVALUAR			
A	51	1	ONDULADO EN BARRAS CORRUGADAS DE 6 mm.			
B	27	2	BAJO RITMO DE PRODUCCIÓN.			
C	23	3	BAJO RENDIMIENTO.			

CRITERIOS DE SELECCIÓN		F. P.							
F	Facilidad para solucionarlo 1. Muy difícil 2. Difícil 3. Fácil	2							
A	Afecta a otras áreas su implementación 1. Sí 3. Algo 5. Nada	1							
C	Mejora la calidad 1. Poco 3. Medio 5. Mucho	5							
T	Tiempo que implica solucionarlo 1. L. Plazo 2. M. Plazo 3. C. Plazo	2							
I	Requiere inversión 1. Alta 3. Media 5. Poca	3							
S	Mejora la seguridad industrial 1. Poco 2. Medio 5. Mucho	2							

PROB.	F	A	C	T	I	S	Total
A	6	5	15	6	9	10	51
B	4	3	5	4	9	2	27
C	4	5	5	4	3	2	23

Imagen 1.13. Matriz FACTIS de problemas evaluados en el ondulado de barras corrugadas.

Fuente: CORPORACIÓN ACEROS AREQUIPA (2015)

1.2 Herramientas de la propuesta de mejora

En base a las causas raíces que se encuentren, se debe seleccionar la contramedida más adecuada que permita minimizar su ocurrencia. En este trabajo de investigación, se usará herramientas de marketing analytics y minería de datos como base de la propuesta de mejora. A continuación, se describirá las herramientas y conceptos que se usarán para concretar la propuesta a lo largo del capítulo 3.

1.2.1 Conceptos generales de minería de datos y marketing analytics.

Big Data. Según la revista Harvard Business Review (2014), este concepto hace referencia a los grandes volúmenes y distintos tipos de información que las empresas hoy en día pueden recolectar gracias al uso de sistemas de bases de datos. Esta puede ser información que proviene desde la misma empresa o desde fuentes externas. Asimismo, los datos pueden ser del tipo estructurado (cantidad de clientes que compran un producto determinado) como datos no estructurados (comentarios recopilados en una encuesta de satisfacción). Cabe resaltar que es totalmente necesario que la información sea colocada en formatos estructurados para que se pueda realizar un correcto análisis de esta.

Analytics. Por otro lado, la revista mencionada previamente, define este concepto como la mezcla entre herramientas de matemática y estadística que permiten analizar la información bruta obtenida por la empresa y convertirla en información que pueda agregar valor a la toma de decisiones. Existen tres tipos de *analytics*:

- *Descriptive Analytics:* Permite describir que pasó en el pasado a través de indicadores clave o KPI's. Por ejemplo: "La cantidad de clientes que asistió a un supermercado fue 10,000 personas mensuales durante el 2019".
- *Predictive Analytics:* Permite describir que pasará en el futuro usando información histórica y modelos predictivos. Por ejemplo: "La cantidad de clientes que asistirá a un supermercado será de 20,000 personas mensuales durante el 2025".
- *Prescriptive Analytics:* Permite determinar, de forma óptima cuál es la mejor decisión que la empresa puede tomar para tener el mejor beneficio. Por ejemplo: "La cantidad óptima de clientes que la empresa debe buscar que asistan a su supermercado principal debe ser de 15,000 personas mensuales durante el 2020 con el fin de maximizar su rentabilidad y minimizar el impacto en la satisfacción por las largas colas generadas".

Marketing analytics y Data mining.

Según la Real Academia Española (1992), el marketing se define como el *“conjunto de principios y prácticas que buscan el aumento del comercio, especialmente de la demanda”*.

Según Maturana (2018), quien es doctor de la Universidad de California, *“podemos definir marketing analytics como los procesos y tecnologías que les permiten a las empresas evaluar la efectividad de sus iniciativas de marketing”*. Asimismo, el autor también hace énfasis en que estas técnicas antes solo eran usadas por especialistas en marketing digital, sin embargo, actualmente se usa cada vez más en distintos rubros debido a su efectividad y el buen aprovechamiento de la información de las empresas.

Según Veliz (2018), quien es doctor en ciencias y en ingeniería industrial por la Pontificia Universidad Católica del Perú, *“la minería de datos es un proceso que conjuga la estadística, la inteligencia artificial y la informática para extraer información de grandes cantidades de datos de manera automática y semiautomática. Esta información se traduce en patrones útiles, reglas y asociaciones, lo que permite el conocimiento de la realidad y la ayuda en la toma de decisiones”*. El autor resalta que las etapas del proceso de minería de datos son las siguientes:

- Entendimiento de los objetivos y del problema a desarrollar
- Conocimiento y recolección de los datos
- Preparación de los datos
- Modelación de los datos
- Evaluación de los modelos aplicados
- Consolidación del conocimiento.

Por otro lado, el autor también realiza la diferencia entre dos tipos de modelos de aprendizaje automático, los cuales son los más comunes a usar en minería de datos. Estos se describen a continuación.

- Modelos de aprendizaje supervisado: Se explican variables dependientes en términos de otras variables independientes. En caso la variable dependiente es del tipo categórica, se trata de un

modelo de clasificación, mientras que si esta es del tipo continua, se trata de un modelo de regresión.

- Modelos de aprendizaje no supervisado: En vez de explicar variables dependientes en términos de otras variables independientes, se realiza la búsqueda de relaciones entre las variables independientes.

1.2.2 Regresión lineal.

El modelo de regresión lineal, según Veliz (2018), “*relaciona los valores x e y de dos variables numéricas X e Y , respectivamente*”. Es un tipo de modelo de aprendizaje supervisado de regresión. Este puede definirse a través de la siguiente ecuación:

$$Y = C_0 + C_1 * X_1 + \epsilon$$

Donde cada variable y parámetro tienen el siguiente significado:

- Y: variable dependiente.
- C₀: intercepto.
- C₁: pendiente de la recta y coeficiente de la variable X₁
- X₁: variable independiente.
- ε: error de medida o valor de variables no consideradas.

1.2.3 Regresión lineal múltiple.

El modelo de regresión lineal múltiple, según Veliz (2018), “*es una generalización del modelo de regresión lineal simple*”. Es un tipo de modelo de aprendizaje supervisado de regresión. Un modelo de regresión lineal múltiple de n variables independientes se presenta en la siguiente ecuación:

$$Y = C_0 + C_1 * X_1 + C_2 * X_2 + \dots + C_n * X_n + \epsilon$$

Donde cada variable o parámetro tienen el siguiente significado:

- Y: variable dependiente.
- C0: intercepto.
- Ci: coeficiente de la variable Xi.
- Xi: variable independiente.
- E: error de medida o valor de variables no consideradas.

1.2.4 Regresión logística.

Según Veliz (2018), en contraste con el modelo de regresión lineal, este modelo busca explicar los valores de una variable de carácter categórico a través de variables que pueden ser numéricas o no numéricas. A pesar de tener el la palabra “regresión” en su nombre, es modelo probabilístico de aprendizaje supervisado de clasificación.

Por otro lado, según Chapman (2015), diversos problemas y decisiones de marketing se basan en la decisión de estimar cuál es la probabilidad asociada a la ocurrencia de eventos, la cual puede ser representada a través de unos y ceros (forma binaria). El autor menciona que la regresión logística se adecua bastante a estos problemas debido a dos razones principales:

- Modela directamente la probabilidad que se busca estimar
- A diferencia del modelo lineal, limita el valor de la variable estimada en un rango de [0,1]

La ecuación del modelo en mención se define en su forma más básica en la siguiente ecuación:

$$p(Y) = \frac{e^x}{e^x + 1}$$

donde x toma la forma $C_0 + C_1 * X_1 + C_2 * X + \dots + C_n * X_n$

- Y: variable dependiente.
- C0: intercepto.
- Ci: coeficiente de la variable Xi.

- X_i : variable independiente.

Es importante resaltar que los valores de X_i no necesariamente deben ser variables lineales.

Algunos ejemplos donde se puede usar el modelo de regresión logística se describen a continuación:

- Determinar la probabilidad de que un cliente abandone un banco en un periodo definido en base a su segmento, edad, sueldo y cantidad de productos con la empresa.
- Determinar la probabilidad de ocurrencia de un sismo en base al año, día, lugar y cantidad de tiempo desde la ocurrencia del último sismo.

1.2.5 *Clustering* o análisis de conglomerados.

Según Veliz (2018), “*el análisis de conglomerados comprende una serie de técnicas descriptivas del análisis multivariado que tiene por objetivo la clasificación o agrupamiento de individuos*”. El resultado final de este modelo son las agrupaciones de cada elemento del sistema en base a los atributos evaluados. Cabe resaltar que estas agrupaciones no se conocen, sino que son descubiertas una vez obtenido los resultados del modelo. A continuación, se describirá brevemente los modelos más conocidos de este tipo.

Técnica jerárquica aglomerativa.

La técnica se resume en los siguientes 4 pasos, según Veliz:

- Se parte de tantos conglomerados como elementos existan
- Se calculan las distancias entre los elementos iniciales
- Se forma un nuevo grupo con los elementos x e y más próximos. Posteriormente, la matriz de distancias se vuelve a calcular eliminando los elementos x e y tomándolos en cuenta como un solo elemento $[x,y]$.
- Se procede a iterar este procedimiento hasta obtener un solo grupo con todos los elementos.

Existen tres formas de calcular la distancia dos los elementos, las cuales pueden ser a través del enlace del vecino más cercano ($f = \text{máximo}$), del vecino más lejano ($f = \text{mínimo}$) o del enlace promedio ($f = \text{promedio}$). Esto puede representarse a través de la ecuación que se muestra a continuación.

$$d(x, y) = f(d(x, y)) \quad (1.4)$$

Donde las variables a considerar son las siguientes:

x: primer elemento

y: segundo elemento

f: función que puede adoptar los valores de máximo, mínimo o promedio, según se decida.

Técnicas no jerárquicas: El método de K-medias.

Según Veliz (2018), “*las técnicas no jerárquicas permiten la formación de conglomerados cuyo número K es previamente fijado*”. El objetivo del modelo es que la suma de las distancias al cuadrado definida entre cada elemento y su centroide sea mínima. Esto se puede representar a través de la siguiente ecuación.

$$\operatorname{argmin} S = \sum_{j=1}^K \sum_{xi \in S_j} \operatorname{abs}(xi - mj)^2$$

Los pasos a seguir para realizar el modelo, son los siguientes:

- Formar K grupos de forma aleatoria y calcular para cada uno su centroide.
- Calcular la distancia de cada punto a cada centroide formado. En caso de tener una distancia menor a un centroide al que no fue asignado, se vuelve a asignar el grupo.
- El proceso puede terminar si la distancia entre los centroides iniciales y centroides actuales es pequeña o en caso de haber definido un máximo número de iteraciones.

1.2.6 Análisis conjunto.

En esta sección, se describirá la forma más común del análisis conjunto. Según Chapman (2015), el análisis conjunto determina la importancia relativa entre varios atributos (variables independientes como tiempo de espera, seguridad, comodidad, entre otras) para cuantificar que tanto influye cada uno de estos en una variable (variable dependiente como la satisfacción del cliente). Gracias a este modelo, se puede determinar cuáles son los atributos que más valoran los clientes para que la empresa realice estrategias que permitan maximizar sus ingresos y aumentar sus tasas de retención. Algunas de las principales ventajas se muestra a continuación:

- Permite conocer las preferencias del cliente en base a data histórica.
- Permite cuantificar el efecto de mover una variable, la cuál puede traducirse en una variación de costos e ingresos a la vez.
- Permite conocer la combinación óptima de características a ofrecer en un servicio o producto. Esto se obtiene gracias a que el modelo permite simular cuál de todas será la que tenga mayor aceptación por el mercado sin tener que probarlo en la vida real.

Para determinar la influencia de cada variable, se usará la técnica de regresión lineal multivariada, la cual se describió en la sección 1.2.3. de la presente investigación.

1.2.7 Técnicas de muestreo.

A continuación, se describirá brevemente las técnicas de muestreo más usadas y sus respectivos estimadores de media. Es importante determinar, en cada uno de estos muestreos, el error estándar de cada estimación pues este permite conocer que tan precisa y confiable son los resultados que el estudio arroja.

Técnica simple de muestreo aleatorio.

Según Scheaffer (2007), en un muestreo aleatorio simple (MAS) se define principalmente porque cada elemento de la población tiene la misma probabilidad de ser elegido (debido a aleatoriedad en la que se obtiene la muestra). Los tipos de muestreo simple pueden ser los siguientes:

- Con reemplazamiento: Al elegir un elemento, este se vuelve a reintegrar a la población, por lo que existe la probabilidad de que pueda volver a ser elegido.
- Sin reemplazamiento: Al elegir un elemento, este no se vuelve a reintegrar a la población, por lo la probabilidad de que pueda volver a ser elegido es nula.

A continuación, se presenta el estimador de la media poblacional y su error estándar de estimación para el muestreo aleatorio simple.

Estimador de media poblacional (μ'):

$$\mu' = \frac{\sum_{i=1}^n y_i}{n}$$

Varianza estimada de (μ'):

$$V'(\mu') = \frac{S^2 (N - n)}{n N}$$

Donde las variables a considerar son las siguientes:

n = cantidad de elementos en la muestra.
 N = cantidad de elementos de la población
 y_i = cantidad observada del elemento i
 S^2 = desviación estándar muestral

Técnica de muestreo estratificado.

Según Scheaffer (2007), en un MAE (muestreo aleatorio estratificado) se define principalmente porque se conocen subconjuntos de la población (estratos), lo cual permite realizar una segmentación para luego realizar un MAS en cada uno de estos segmentos definidos antes de realizar el muestreo. Según Valdivieso (2019), las principales características de este muestreo son las siguientes:

- Se puede determinar características de la población de forma más precisa.
- Se puede encontrar diferencias entre grupos distintos.

- Estimaciones más precisas
- Requiere un mayor costo y cantidad de muestra.

A continuación, se presenta el estimador de la media poblacional y su error estándar de estimación para el muestreo aleatorio estratificado.

Estimador de media poblacional (μ'):

$$\mu' = y' = \frac{\sum_{i=1}^L Ni * yi'}{N}$$

Varianza estimada de (μ'):

$$V'(\mu') = \frac{\sum_{i=1}^L Ni^2 (Ni - ni)}{N^2} \left(\frac{Si^2}{ni} \right)$$

Donde las variables a considerar son las siguientes:

n_i = cantidad de elementos en la muestra del estrato i
 L = cantidad de estratos definidos
 N_i = cantidad de elementos en el estrato i
 N = cantidad de elementos de la población
 y_i' = media muestral observada del estrato i
 S_i = desviación estándar muestral del estrato i

Técnica de muestreo por conglomerados.

Según Valdivieso (2019), los muestreos por conglomerados consisten en dividir a la población en subconjuntos que, en el caso ideal, deben ser heterogéneos por dentro, pero homogéneos por fuera. Estos pueden ser de distintos tipos, sin embargo, el primer paso del método es definir los conglomerados o subconjuntos de la población. En la presente investigación se tratarán los dos más usados:

- Muestreo por conglomerados de una etapa: Se seleccionarán unas unidades primarias (UPM) de una población mediante un MAS para luego encuestar a todas las unidades secundarias (USM) de cada conglomerado.

- Muestreo por conglomerados de dos etapas: Se seleccionarán unas unidades primarias (UPM) de una población mediante un MAS para luego realizar otro MAS dentro de cada UPM y así seleccionar a las unidades secundarias (USM) de cada conglomerado.

Según Valdivieso (2019), las principales características de este tipo de muestreo se mencionan a continuación.

- Ofrece un menor costo a comparación de un MAS y un MAE
- Permite realizar muestreos para poblaciones que se encuentran dispersas geográficamente.
- Sencillez en el trabajo de campo a comparación de un MAS y un MAE.
- Estimaciones menos precisas que un MAS y un MAE.

A continuación, se presentará las ecuaciones para calcular el estimador de la media poblacional y su error estándar de estimación para el muestreo por conglomerados de una etapa, según Scheaffer (2007).

Estimador de media poblacional (μ'):

$$\mu' = y' = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n M_i}$$

Varianza estimada de (μ'):

$$V'(\mu') = \left(\frac{N-n}{NnM^2}\right)Sr^2$$

Donde

$$Sr^2 = \sum_{i=1}^n \frac{(y_i - y' m_i)^2}{n-1}$$

Donde las variables a considerar son las siguientes:

n = cantidad de conglomerados seleccionados para encuestar

N = cantidad de conglomerados en la población
 $M_i = m_i$ = cantidad de elementos en el conglomerado i
 M = cantidad de elementos en la población
 y_i = suma de cantidades observadas en el conglomerado i
 y' = estimador de media poblacional
 S_i = desviación estándar muestral del estrato i

Nota: M' puede estimarse por m' en caso de desconocerse M .

Por otro lado, en las siguientes líneas se presentará las ecuaciones para calcular el estimador de la media poblacional y su error estándar de estimación para el muestreo por conglomerados de dos etapas, según Scheaffer (2007).

Estimador de media poblacional (μ'):

$$\mu' = y' = \frac{\sum_{i=1}^n M_i y_i'}{M' * n}$$

Varianza estimada de (μ'):

$$V'(\mu') = \left(\frac{N-n}{N}\right) \left(\frac{1}{nM^2}\right) S_b^2 + \frac{1}{nNM^2} \sum_{i=1}^n M_i^2 \left(\frac{M_i - m_i}{M_i}\right) \left(\frac{S_i^2}{m_i}\right)$$

Donde la varianza muestral entre conglomerados es

$$S_b^2 = \sum_{i=1}^n \frac{(M_i y_i' - M' y')^2}{n-1}$$

Donde la varianza muestral el conglomerado i es

$$S_i^2 = \sum_{j=1}^{m_i} \frac{(y_{ij} - y_i')^2}{m_i - 1}$$

Donde las variables a considerar son las siguientes:

n = cantidad de conglomerados seleccionados para encuestar

N = cantidad de conglomerados en la población
 m_i = cantidad de elementos seleccionados, producto de un MAS, en el conglomerado i
 M_i = cantidad de elementos en el conglomerado i
 M = cantidad de elementos en la población
 M' = cantidad promedio muestral de elementos en un conglomerado
 y_i' = promedio muestral de las cantidades observadas en el conglomerado i
 y_{ij} = observación j en el conglomerado i
 y' = estimador de media poblacional

Nota: M' puede estimarse por m' en caso de desconocerse M .

1.2.8 Software.

John Chambers creó en 1998 un software llamado R, el cual es descrito por su propia empresa como “un lenguaje y entorno para la computación estadística y gráficos” (Chapman, 2015). Además, R es un lenguaje de programación orientado a objetos. Es un software libre muy usado gracias a los 6,000 paquetes libres que ofrece a cualquier usuario, los cuáles han sido validados por estadísticos antes de ser publicados (lo cuál otorga cierto nivel de confianza, que usualmente no se tiene con los softwares de carácter colaborativo).

Las ventajas principales de usar R se muestran a continuación:

- Ofrece la más larga y diversa set de herramientas estadísticas y analíticas.
- El código que se usa puede ser reusado nuevamente en análisis similares.
- Funciona para la gran mayoría de sistemas operativos.
- Aumenta la productividad de los procesos.
- Permite automatizar diversos procesos que pueden ser engorrosos al realizarse en otros softwares como Excel.

Los principales paquetes, librerías o funciones a usar para cada modelo que se usarán en el capítulo 3 se obtuvieron de la página *R Documentation*³. Estas se mencionarán a continuación:

- Regresión lineal y regresión lineal múltiple: `lm` (función).
- Regresión logística: `glm` (función).
- Clustering: `hclust` (función), `kmeans` (función).
- Análisis conjunto: `lm` (función).
- Técnicas de muestreo complejas: `survey` (paquete).

1.3 Casos de estudio.

Para tener un mejor entendimiento de las herramientas a usar en proyectos similares, conocer la experiencia de otros investigadores y los resultados que obtuvieron, se realizó un análisis de tres casos de estudio a fines al tema de investigación. Se seleccionó dos *papers* y una tesis de posgrado. Cabe resaltar que los dos primeros trabajos de investigación usan herramientas similares a la propuesta en el mismo rubro (banca), sin embargo, el tercer caso de estudio un distinto rubro (automotriz) pero fue seleccionado por su similitud con las estrategias de retención que la contramedida seleccionada busca plantear en los canales de la empresa. A continuación, se describirá cada una de estas usando la siguiente estructura: contexto y problemática, metodología y herramientas, resultados y conclusiones.

1.3.1 Primer caso: *A New Application of RFM Clustering for Guild Segmentation to Mine the Pattern of Using Banks' e-Payment Services.*

Contexto, problemática.

La investigación se centra en un banco de Irán, quien tiene la problemática de no tener una estrategia clara de cómo distribuir sus dispositivos de transacciones electrónicas, llamados POS, en los comerciantes del

³ *R documentation. Fecha de consulta: 25 de junio del 2019. Véase: <<https://www.rdocumentation.org/>>*

país dado que estos tienen un costo significativo para la empresa. Por ello, se propone realizar una segmentación de los gremios, los cuales vienen a ser grupos de comerciantes, para poder distinguir patrones de comportamiento y realizar estrategias eficientes de distribución de recursos con foco en maximizar sus ingresos usando herramientas de minería de datos.

Una de las razones por la cual el autor opta por esta metodología, es por la revisión de bibliografía, de la cual concluye que muchos bancos están empezando a usar la segmentación de sus clientes en base a su comportamiento transaccional para obtener mayores ingresos. Según el autor, *"la técnica de agrupación es una herramienta clave en la minería de datos y el reconocimiento de patrones se puede considerar como la técnica de aprendizaje no supervisado más importante"* (Khobzi, 2014)

Metodología y herramientas.

Dado que se tienen muchos clientes, para facilitar el análisis el autor agrupó a los clientes en gremios, los cuales permitirían reducir la cantidad de puntos a clusterizar (147 en total) a comparación de analizar a todos los comerciantes del banco.

Por un lado, el autor usa como principal variable de análisis el valor de *RFM (Recency, Frequency, Money)* de los clientes analizados. Esto le permite conocer los siguientes tres atributos que considera relevantes:

- *Recency*: Recencia. Indica el tiempo que ha pasado desde la última transacción de un gremio en un POS hasta el periodo en que se analiza la información.
- *Frequency*: Frecuencia. Indica la cantidad de veces que el gremio realizó una compra en el periodo evaluado.
- *Money*: Monto de dinero. Indica el monto total que el gremio transaccionó por el POS en el periodo evaluado.

Sin embargo, dado que los gremios no tienen la misma cantidad de comerciantes, el autor opta por definir dos variables en base a las 3 mencionadas anteriormente, las cuales permitirán para poder hacer un análisis comparable entre gremios y a su vez, poder encontrar similitudes en estos.

- **RF1F2M**: Se usan las tres variables que se definieron en líneas anteriores, añadiendo F2, que viene a ser la cantidad de comerciantes en cada uno de los gremios.
- **RF*M***: De forma similar, que las variables iniciales, con la diferencia de que F* viene a ser el valor de frecuencia promedio obtenido por gremio entre F2, la cantidad de gremios. Por otro lado, de forma similar M* viene a ser el valor de flujo monetario promedio obtenido por gremio entre la cantidad de gremios.

A continuación, se mencionarán las herramientas usadas por el autor para poder realizar la segmentación que propone en la introducción.

- *Clustering*. Realiza la diferencia entre los métodos jerárquicos y particionales, dando énfasis en que los particionales son los que tienen menor complejidad al momento de trabajar con grandes volúmenes de datos. Por esta razón, decide usar el algoritmo de K-means.
- *Dunn Index* Dado que el problema del algoritmo de K-means se da en que no existe una metodología para definir la solución óptima, el autor elige este indicador para poder determinar el número de clústers óptimo. Este indicador se mide en una escala del -1 al 1 y determina que tan bien están asignados los puntos a un clústers en base a la comparación de la suma de distancias promedio de los puntos al centroide de su mismo clúster comparado con la suma de distancias promedio hacia otros centroides de clústers vecinos.

- *CLV*. El autor resalta la importancia de poder medir el valor de los clientes en valor futuro y traerlo a valor presente, por ello usa un indicador de valorización del cliente, el cuál se muestra en la imagen 1.13.

$$CLV_{Ci} = NR_{Ci} \times W_R + NF_{Ci} \times W_F + NM_{Ci} \times W_M \quad (4)$$

Imagen 1.13. Indicador de Customer Life Value. Fuente: Khobzi (2014)

Un esquema de la metodología usada por el autor puede observarse en la imagen 1.14. Esta se resume en primer lugar, a conocer y entender el funcionamiento del negocio, recolectar y preparar la información a usar, calcular los valores de las variables definidas para la clusterización, definir el K óptimo en base al *Dunn Index*, normalizar los datos para poder hacerlos comparables en el cálculo del CLV y realizar un ranking de cada clúster en base al CLV calculado (esto a través de una normalización de método infinito, dado que una de las variables afectaba de forma negativa al indicador). Finalmente, se planteará las estrategias de marketing en base a los resultados.

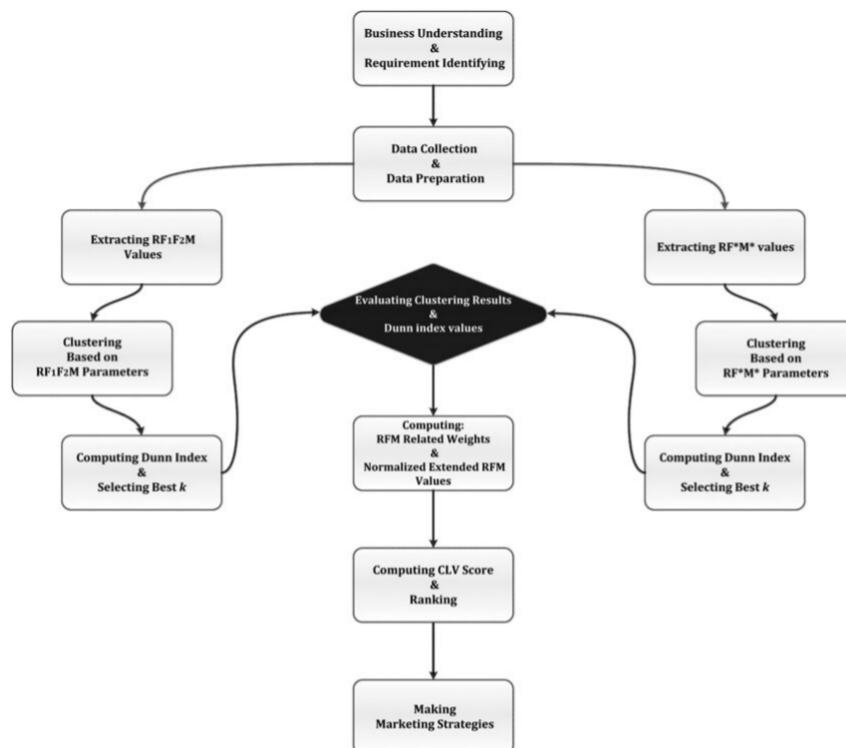


Imagen 1.14. Esquema de metodología empleada. Fuente: Khobzi (2014)

Resultados

Los resultados obtenidos para ambas clusterizaciones pueden verse en la imagen 1.15, de la cual se concluye que los k óptimos son $k=8$ y $k=4$ respectivamente para el grupo de variables definidos, RF1F2M y RF*M*. En estos se observa que existe una gran diferencia del indicador de los k óptimos a otros valores de k .

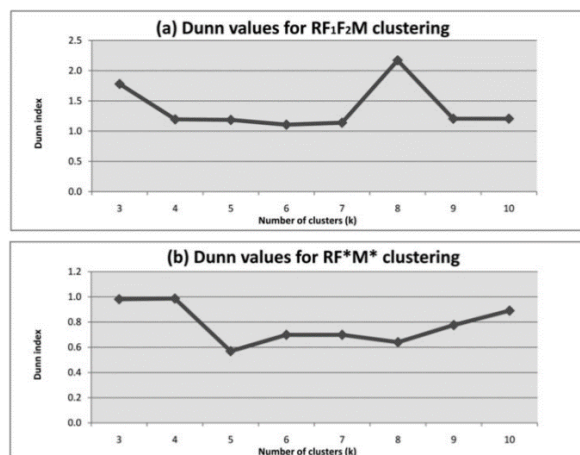


Imagen 1.15. Resultados del Indicador de Dunn para definir el K óptimo. Fuente: Khobzi (2014)

Por otro lado, en la imagen 1.16 se puede observar los resultados de la primera clusterización (con los datos ya normalizados), en base a la cual se calculó el CLV de los clientes.

Approach	Cluster	Percentage of guilds	Normalized R	Normalized F ₁	Normalized M	CLV score	Rank
RF ₁ F ₂ M	C1	52	0.128	0.007	0.028	0.029	6
	C2	1	0.106	0.401	0.132	0.215	2
	C3	4	0.170	0.002	0.002	0.016	8
	C4	1	0.106	0.215	0.047	0.105	4
	C5	6	0.149	0.006	0.008	0.019	7
	C6	34	0.106	0.024	0.059	0.052	5
	C7	1	0.106	0.182	0.133	0.146	3
	C8	1	0.128	0.164	0.591	0.418	1

Imagen 1.16. Resultados de la primera clusterización. Fuente: Khobzi (2014)

Analizando los resultados, el autor consultó con un experto del sector el cuál sugirió agrupar C8, C2, C7 y C4 como un solo grupo, dado que la cantidad de gremios era muy poca y además tienden a tener resultados relativamente altos con otros grupos. Con esta agrupación se obtuvo 5 segmentos de gremios, los cuales pudieron ser diferenciados por ser negocios de distintos rubros en cada grupo (esto permitió poder determinar, más adelante, cuales podrían ser los clientes con más potencial para el banco). Los segmentos fueron definidos por el autor definió de la siguiente forma.

- *The Main Target.* Vienen a ser los clientes más rentables (por tener un mayor CLV).
- *The Lucrative Professions.* Resultaron ser el grupo que traía mas ingresos, después del primero.
- *The Ordinary Professions.* Son grupos que generan ingresos significativos, pero que no llegan a cumplir las expectativas de la empresa, sin embargo, ayudan a tener una posición de mercado pues abarca casi el 50% de los gremios evaluados.
- *The Neutral Professions.* Se define como el grupo que tiende a no tener buenos ojos para el banco por los ingresos que generan, pero que puede tener chances de sobrevivir.

- *The Inutile Professions*. Son los puestos de comerciantes donde los POS no son usados con alta frecuencia como la empresa espera.

Finalmente, el autor menciona que propuso la segmentación a la empresa, con el fin de plantear estrategias de distribución de recursos eficientes y con justificación, esto a partir de la priorización de poder determinar cuáles son las características de los gremios que tienen mayor potencial. Los principales beneficios que obtuvo la empresa después de implementar estas estrategias fueron las siguientes:

- Entender patrones de compra para la toma de decisiones eficiente, a pesar de que su población crece, la empresa puede plantear estrategias diferenciadas para grupos genéricos ya definidos.
- Aumento del ticket promedio de compra en un 18%
- Número de comerciantes con terminales con poco uso y poco potencial de ingresos, se pudo reducir en un 13%.

Conclusiones

Se usó el método de clusterización mediante el algoritmo k-means, para definir diferencias entre gremios de comerciantes, obteniendo 5 grupos distintos para los cuales la empresa planteó diversas estrategias que trajeron un conjunto de beneficios. El autor resalta que este método puede realizarse en otros bancos y en otros casos que puedan ser de interés.

1.3.2 Segundo caso: *Quality of service and customer satisfaction: a conjoint analysis for the Nigerian bank customers.*

Contexto y problemática.

En un contexto de fuerte competencia entre distintas empresas del sector bancario en Nigeria, la satisfacción viene a ser una medición muy importante para los bancos, pues permite cuantificar si el servicio ofrecido por la empresa permitió satisfacer las necesidades del cliente, lo cual finalmente puede traducirse en qué tan fidelizados se encuentran. Según el autor, los beneficios de mejorar la calidad del servicio ofrecido

correlacionan con el aumento de los ingresos y aumentan la oportunidad de supervivencia de la empresa en el largo plazo. Algunos de estos beneficios son las siguientes:

- Mejora la imagen corporativa
- Mejor oportunidad de ventas cruzadas (esto es, concretar la venta de un producto adicional a lo que normalmente un asesor de ventas vendería en un proceso de venta común).
- Disminución de clientes que desertan de la empresa.
- Mayor probabilidad de que la empresa sea recomendada por sus mismos clientes.
- Mantener buenas relaciones a largo plazo con sus clientes

El autor menciona diversos problemas que existen en su red de canales, tales como problemas con cajeros automáticos, largos tiempos de espera, entre otros, es de interés para la empresa poder determinar que problemas priorizar en base a las variables de calidad de atención que más impactan en la percepción del cliente. Por ello, se propone realizar un análisis conjunto, el cual permitirá determinar la importancia de diferentes atributos en distintas medidas para poder ofrecer un servicio óptimo en base a los recursos de la empresa, pues como se mencionó en el primer caso analizado, la optimización de costos también viene a ser una variable importante a tomar en cuenta en la mejora de los procesos. Asimismo, el autor resalta que una de las ventajas principales de la metodología es poder estimar los impactos de realizar distintas variaciones antes de tomar la decisión en la vida real.

Metodología y herramientas.

El autor obtuvo de dos fuentes relevantes del sector bancario de su país cuales son las dimensiones más comunes que se usan para definir el servicio de calidad (9 en total). En la imagen 1.17 se puede observar las dimensiones encontradas, mientras que en la segunda columna se puede observar un atributo definido de forma más específica, el cual fue seleccionado luego de reunirse con expertos del sector.

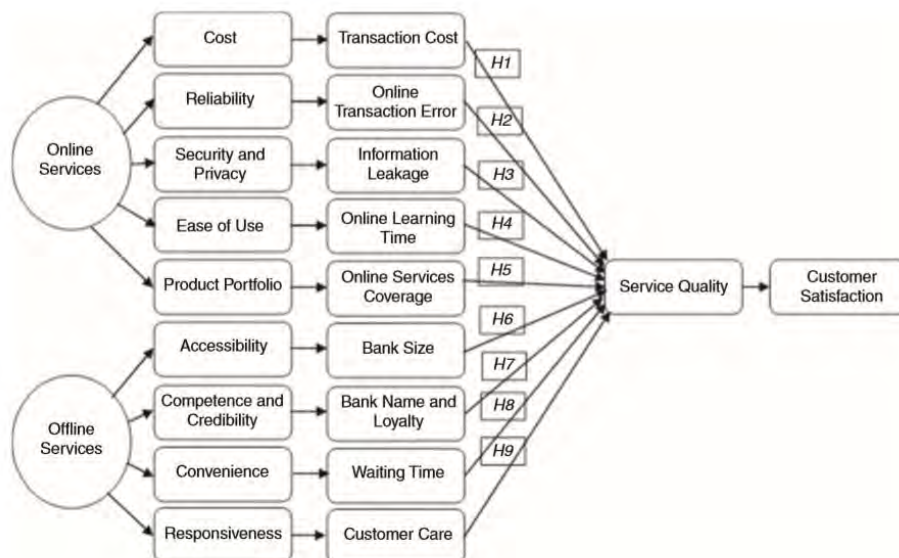


Imagen 1.17. Atributos evaluados en el cuestionario. Fuente: Yusuf (2015)

En este estudio, se usó un modelo de utilidad, el cual es una medición de la calidad ofrecida por el servicio a los clientes. Se asume que los servicios con mayor indicador de utilidad serán los que tienen mayor probabilidad de ser elegidos por los consumidores. Este modelo se puede representar a través de la siguiente ecuación, en la imagen 1.18, donde el vector X_{nj} representa la puntuación que cada cliente otorga a los atributos, el vector B_n representa los coeficientes de cada atributo y la última variable representa el error que trae cada una de estas estimaciones.

$$U_{nj} = X_{nj}\beta_n + \varepsilon_{nj}$$

Imagen 1.18. Modelo de utilidad obtenida a través de atributos. Fuente: Yusuf (2015)

Una vez definidas los atributos a evaluar, el autor recolectó la información envió 1,500 encuestas, ofreciendo incentivos para poder obtener una participación activa de los evaluados, sin embargo, solo se validaron 1,245 de las encuestas. En la imagen 1.19 se puede observar un resumen del análisis exploratorio que se realizó antes de evaluar el impacto de cada atributo. De este gráfico se puede concluir que la

población evaluada es representativa para toda la población pues fue de distintas edades, sexos, grados académicos, ingresos y pertenecientes a distintos bancos (sin embargo, se sugeriría al autor incluir el error y más detalles del muestreo en la investigación para difundir una mayor confiabilidad de la base de su estudio, en este caso, las encuestas realizadas).

Characteristics	Percentage	Characteristics	Percentage	Characteristics	Percentage	Characteristics	Percentage
Age range (18-61 yrs)		Monthly income (US\$)		Major bank		Credit card usage	
Age1 (< 30 yrs)	40.16	Low (0-315)	45.22	Bank A	20.57	0	54.30
Age2 (≥ 30 yrs)	59.84	High (≥ 316)	54.78	Bank B	31.41	1	20.56
Gender		Marital status		Bank C	30.04	≥ 2	25.14
Female	47.95	Single	46.02	Other banks	17.99	Debit card usage	
Male	52.05	Married	49.40	Deposit size		0	9.24
Education		Others	4.58	Low (0-315)	86.18	1	37.27
Edu1 (≤ diploma)	45.86	Family size		High (≥ 316)	13.82	≥ 2	53.49
Edu2 (≥ bachelor)	54.14	Small (1-5)	52.77	Transaction frequency		Bank cheque usage	
Occupation		Big (≥ 6)	47.23	1-2/week	76.14	0/week	52.13
Unemployed	14.30	Insurance		≥ 3	23.86	1/week	35.82
Civil servant	41.20	No	69.16	No. of credit cards		≥ 2	12.53
Businessman	24.74	Yes	30.84	0	53.57	Stock Investment	
Academic	12.05	Insurance Type		1	29.88	No	94.06
Others	7.71	Health	28.31	≥ 2	16.55	Yes	5.94
Working experience		Life	16.94	No. of debit cards		Stock size	
Exp1 (0-5 yrs)	50.04	Vehicle	34.80	0	8.67	Low (0-315)	70.27
Exp2 (6-10 yrs)	24.34	General	11.37	1	65.78	High (≥ 316)	29.73
Exp3 (≥ 11 yrs)	25.62	Business	8.58	≥ 2	25.54		

Note: Characteristics of study subjects $n = 1,245$

Imagen 1.19. Resumen de análisis exploratorio. Fuente: Yusuf (2015).

Resultados.

El autor realiza un modelo de coeficientes aleatorios, el cual le permite obtener medias y varianzas de cada coeficiente en base a su distribución estimada. En este caso, según el autor, se usó una estimación de probabilidad máxima con un tamaño de muestra de 33,615, las cuales se muestran en la imagen 1.20.

Variables		Mean (b) of B			Variance (w) of B			
		Mean	SE	t -value	SD	Variance	SE	t -value
Error frequency	ERFRQ	-0.0588***	0.0143	-4.13	0.0548	0.0030*	0.0288	1.90
Comprehensiveness	CMPON	-0.0031***	0.0008	-3.99	0.0008	6.4E-07	0.0023	0.34
Loyalty (Bank A)	LYBA	-0.1663***	0.0367	-4.53	0.8786	0.7719***	0.0409	21.50
Loyalty (Bank B)	LYBB	0.0098	0.0421	0.23	0.9453	0.8936***	0.0412	22.93
No. of branch offices	BRNW	0.0036***	0.0008	4.62	0.0001	1.0E-08	0.0012	0.09
Waiting time	WTIME	-0.0086***	0.0011	-7.55	-0.0014	2.0E-06	0.0024	-0.59
Extent of customer care	CUSCARE	0.0344	0.0279	1.23	-0.0115	0.0001	0.0499	-0.23
Initial learning time	LNTIME	-0.0565***	0.0176	-3.21	-0.1201	0.0144**	0.0485	-2.48
Information leakage	INFOLEAK	-0.0290**	0.0093	-3.13	-0.0078	0.0001	0.0196	-0.40
Transaction cost	COST ^a	-1.3665***	0.3262	-4.19	-1.2595	1.5863***	0.2350	-5.36

Notes: $n = 33,615$. The LYBC was used as the base in terms of loyalty to bank name; LR $\chi^2(10) = 899.20$, log likelihood = -11,759.138, LRI: $\rho = 0.414$. ^aCOST is negative COST variable so that the distribution of the coefficient with transformed distribution of COST everywhere is positive. ***, **, *Significance at 1 percent ($|t| > 2.58$), 5 percent ($|t| > 1.96$) and 10 percent ($|t| > 1.65$), respectively

Imagen 1.20. Resumen de resultados del modelo de coeficientes aleatorios. Fuente: Yusuf (2015).

Los principales hallazgos se nombran a continuación. Estos pueden ser visualizados gráficamente en la imagen 1.21.

- La percepción de los clientes por el servicio ofrecido se indica más impactada por el tiempo de espera y el precio, pues obtienen un 18% y 26% de importancia relativa.
- Los atributos menos importantes son la atención al cliente (0.16%) y el nombre y lealtad al banco (1.56%).

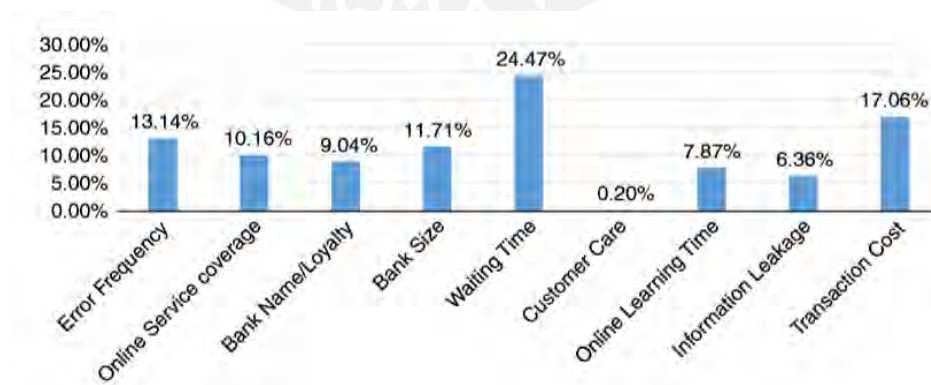


Imagen 1.21. Importancia de los coeficientes del análisis conjunto. Fuente: Yusuf (2015).

Conclusiones.

Del análisis obtenido, se concluye que la empresa debe enfocarse en solucionar los problemas que generan altos tiempos de espera y evaluar alternativas para ofrecer transacciones a un costo más bajo. De esta forma, se comprueba que la empresa puede obtener el impacto de cada atributo en sus clientes y en base a ello, plantear las estrategias adecuadas que le ayuden a fidelizar y por ende, a mejorar sus ingresos en el largo plazo.

1.3.3 Tercer caso: Diseño de Estrategias de Retención de Posventa en una Empresa Automotriz de Vehículos Livianos Aplicando Marketing Analytics.

Contexto y problemática.

En el sector automotriz, es importante aumentar la participación de mercado y fidelizar cada vez más a los clientes que se va adquiriendo, dado que los servicios postventa que se ofrecen son los que más ingresos generan a la empresa. Esto último es la clave que permite a las empresas del sector tener un mejor posicionamiento. Una forma de medir los resultados de post venta es a través de la cantidad de veces que un cliente realiza un mantenimiento (del tipo preventivo o correctivo).

El presente caso se centró en un análisis de una empresa automotriz con 22 distribuidores o ‘dealers’ a nivel nacional en la cual se planteó mejorar el indicador mencionado anteriormente para generar mejores relaciones con el cliente sostenibles en el largo plazo.

Metodología y herramientas.

La primera parte del caso empieza con un diagnóstico de los procesos y principales indicadores de la empresa. Posteriormente, se centra en el uso de herramientas de marketing analytics tal como se observa en las imágenes 1.22 y 1.23. En la primera imagen, se observa el proceso y las etapas que debe seguirse en un análisis usando minería de datos, mientras que, en la segunda imagen, se observa como interactúan

diversos procesos dentro de la empresa al momento de usar este tipo de herramientas en el área de marketing.

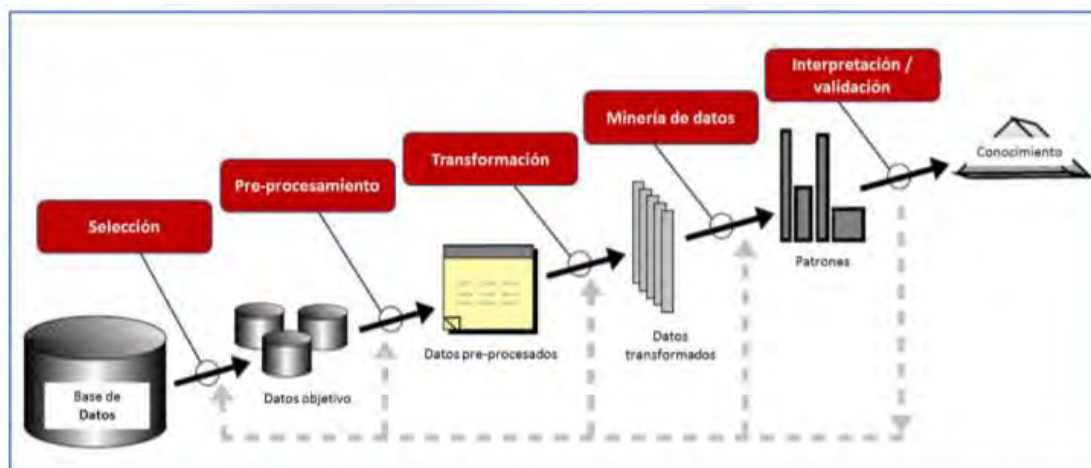


Imagen 1.22 Etapas del proceso KDD. Fuente: Torrejón (2018).

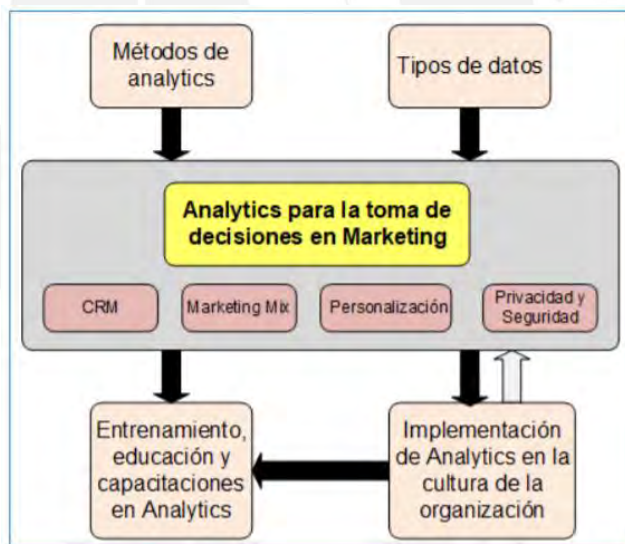


Imagen 1.23 Estructura de analytics aplicado en marketing. Fuente: Torrejón (2018).

Asimismo, los algoritmos que fueron evaluados usar por el autor, se muestran a continuación en base a los objetivos que se planteó al inicio de la investigación con el fin de determinar cuales se adecuan mejor al

problema que se desea solucionar. Esto puede observarse en la imagen 1.24. Finalmente, se decidió optar por un algoritmo de clustering y una regresión logística, dado la tipología del problema y los datos.

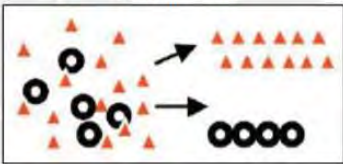
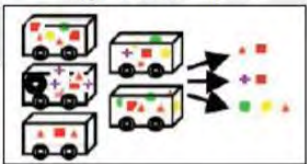

Objetivo	Técnica	Algoritmos
Predecir comportamiento del cliente	<p>Clasificación</p> 	<ul style="list-style-type: none"> - Árboles de decisión. - Redes neuronales. - Clasificación bayesiana.
Determinar hechos en común dentro de un conjunto de clientes	<p>Reglas de asociación</p> 	<ul style="list-style-type: none"> - Algoritmo Apriori. - Partition -Eclat
Dividir a clientes según características y comportamientos	<p>Agrupaciones o clustering</p> 	<ul style="list-style-type: none"> - Algoritmo K-medias. - Algoritmo K-means.

Imagen 1.24 Resumen de herramientas evaluadas. Fuente: Torrejón (2018).

Por otro lado, el software usado por el autor fue R, el cual se decidió después de realizar una comparación entre el mismo y Python. En resumen, ambos softwares son muy buenos para aplicar los algoritmos mencionados anteriormente, sin embargo, debido al desarrollo y flexibilidad que el autor buscaba, decidió optar por R como software estadístico.

	R	Python
Ventajas	<ul style="list-style-type: none"> - Muy bueno para realizar prototipos. - Muy bueno para realizar análisis estadísticos. - Gran entorno de desarrollo interactivo. 	<ul style="list-style-type: none"> - Muy bueno para automatizar las fuentes de información de minería de datos. - Se integra fácilmente a un flujo de producción. - Se puede utilizar en diferentes partes del software del equipo de ingeniería.
Desventajas	<ul style="list-style-type: none"> - Sintaxis compleja. - Los documentos de la biblioteca no son siempre amigables con los usuarios. - Difícil de integrar a un flujo de producción. 	<ul style="list-style-type: none"> - No es tan fácil de realizar análisis estadísticos como R. - La curva de aprendizaje de Python es más larga que la de R.

Imagen 1.25. Comparación entre softwares estadísticos. Fuente: Torrejón (2018).

Una vez definido el software, algoritmos a usar y data input, el autor definió que aplicaría estos modelos en un dealer representativo para poder testear resultados y luego poder escalar la iniciativa en el resto de dealers. Se realizó el proceso de minería de datos con los siguientes fines en cada algoritmo.

- Clusterización. Agrupar clientes según uso de vehículo y tipo de asistencia para poder determinar comportamientos diferenciados y por ende, soluciones personalizadas en base a sus necesidades.
- Algoritmo para definir próximo servicio. Bajo el uso de ratios y algunas condiciones de normas de mantenimiento para definir cuando el cliente necesita un servicio post venta.
- Regresión logística. Definir cuáles son los atributos que más valora un cliente en el servicio post venta.

Por último, en base a los resultados de lo mencionado anteriormente, el autor planteó distintas acciones y estrategias para mejorar el journey del cliente en el servicio post venta para poder lograr una mayor fidelización y nivel de satisfacción (brindando información que otorgue valor al cliente y comunicación para dar seguridad). Por otro lado, el autor desarrolló estrategias y planes de acción personalizados según cada perfil de cliente identificado y data sobre sus últimos mantenimientos para poder anticiparse a las necesidades del próximo servicio que este requiera de forma proactiva.

Resultados.

Gracias a los análisis previamente realizados y a la definición de estrategias de retención en el servicio postventa, se aumentó el promedio de la tasa de retención de clientes en un 7.5%, lo cual se puede traducir en un incremento de ingresos promedio mensual de \$13,600 y un nivel de servicio al cliente superior al 85%, sobrepasando las metas planteadas.

Conclusiones.

Con un conocimiento adecuado del cliente al usar la data que va dejando en las interacciones que tiene con la empresa para predecir futuros comportamientos (en este caso, anticiparse a la necesidad de un servicio postventa), se pueden realizar las adecuadas estrategias de marketing y comunicación para ofrecer servicios que el cliente realmente necesita en el momento adecuado. Finalmente, esto puede traducirse en clientes fidelizados, y por ende, mayores ingresos.

1.4 Resumen Metodológico

Una vez descrito las herramientas y conceptos que usará el presente trabajo de investigación, se resumirá en las siguientes líneas la metodología empleada en cada uno de los siguientes capítulos.

CAPITULO 2: Descripción y diagnóstico del problema

Este capítulo empezará con una descripción de la empresa, sus clientes, proveedores, organización y la situación actual en la que se encuentra a comparación de sus principales competidores. Luego, se procederá a diagnosticar sus procesos, empezando con un mapeo de macroprocesos, una matriz de ponderación pareada para determinar los factores de ponderación de cada variable y una matriz QDF que permitirá determinar el más crítico. De manera similar, se realizará con los procesos y subprocesos del macroproceso elegido, priorizando los más críticos mediante matrices de priorización. Una vez seleccionado el proceso o subproceso crítico, se procederá a realizar un diagrama de flujo y la selección de al menos un indicador para cada actividad. Cada uno de estos indicadores seleccionados será evaluado en fichas indicador para determinar el cumplimiento que tienen respecto a las metas planteadas por la organización. Luego, para cada indicador que no se encuentre en un cumplimiento aceptable, se procederá a atribuirle uno o más problemas y a cuantificar monetariamente el impacto del costo de oportunidad que conlleva cada uno de estos para poder realizar una priorización mediante una curva de Pareto. Posteriormente, para cada problema priorizado, se realizará un análisis de causa-efecto y un análisis de Cinco Porqués para sus causas más relevantes, que permitirá encontrar las causas raíces que los originan. Finalmente, se planteará al menos

una contramedida para cada causa raíz y se realizará una matriz FACTIS para determinar la contramedida más conveniente.

CAPITULO 3: Propuesta de mejora

La contramedida seleccionada, la cual minimizará las causas raíces encontradas en el capítulo 2, se basará en el uso de herramientas de marketing analytics y minería de datos. El tercer capítulo del trabajo de investigación empezará seleccionando las variables a usar y posteriormente, se realizará un análisis exploratorio de éstas para tener un mejor entendimiento del estado actual del proceso seleccionado a mayor escala. Por consiguiente, se desarrollará una breve descripción de los objetivos e importancia que tiene cada herramienta descrita en el marco teórico relacionándolos con el objetivo central de la propuesta. Una vez realizada esta descripción, se realizará una limpieza de datos y se procederá a desarrollar los modelos en el software RStudio. El capítulo concluirá con un análisis de los resultados obtenidos en cada uno de los modelos, así como los resultados que se pueden obtener en conjunto, construyendo finalmente la contramedida planteada.

CAPITULO 4: Evaluación técnica y económica

Con el fin de determinar la viabilidad de la propuesta de mejora, se analizarán tres aspectos importantes: viabilidad de los clientes externos, viabilidad de los clientes internos y viabilidad económica para la empresa. Para los clientes externos, se determinará los riesgos y posibles impactos que puedan existir en el servicio que la empresa les ofrece en base a los resultados obtenidos en el capítulo 3. Para el caso de los clientes internos, se determinará la viabilidad de adaptar sus procesos, *mindset* y estrategias en base a la propuesta de mejora. Finalmente, se calculará los ahorros proyectados para determinar indicadores de rentabilidad como la Tasa Interna de Retorno (TIR) y el Valor Actual Neto (VAN) y se realizará un análisis de sensibilidad multivariado que permitirá definir los resultados del proyecto en base a distintos escenarios.

CAPITULO 5: Conclusiones y recomendaciones

En base a los principales hallazgos en cada capítulo, se procederá a realizar conclusiones y recomendaciones relacionados a cada capítulo. Por consiguiente, unificando todos los resultados, se describirá las conclusiones y recomendaciones generales planteadas por el autor para la empresa. Finalmente, se plantearán una lista de recomendaciones sobre que otras herramientas podrían utilizarse para obtener, probablemente, mejores resultados o un mayor alcance en el presente trabajo de investigación.

Bibliografía

ACOSTA, Carlos

2016 “Gestión del sistema de suministro de productos farmacéuticos, dispositivos médicos y productos sanitarios”. En *SUSALUD*. Consulta: 07 de abril de 2019.

http://www.digemid.minsa.gob.pe/UpLoad/UpLoaded/PDF/EAccMed/ReunionesTecnicas/PONENCIAS/MARZO_2016/PONENCIA_DIA3/04_SUSALUD_GEST_SIST_SUMIN.pdf

AITECO CONSULTORES, SL.

s/f “Matriz de Priorización”. Consulta: 25 de junio de 2019.

<https://www.aiteco.com/matriz-de-priorizacion/>

AMÉRICA ECONOMÍA

2018 “El dilema de generar y retener clientes, ¿son posible ambas?”. En *América Economía*. Consulta: 07 de abril de 2019.

<https://mba.americaeconomia.com/articulos/reportajes/el-dilema-de-generar-y-retener-clientes-son-posible-ambas>

AMÉRICA RETAIL

2018 “Estudios: Las 5 tendencias que marcarán a la Banca en los próximos tres años”. En *América Retail*. Consulta: 07 de abril de 2019.

<https://www.america-retail.com/estudios-consumidores/estudios-las-5-tendencias-que-marcaran-a-la-banca-en-los-proximos-tres-anos/>

ASOCIACIÓN DE BANCOS DEL PERÚ

2018 “Preferencia por canales virtuales sigue creciendo entre usuarios bancarios”. En *ASBANC*. Consulta: 07 de abril de 2019.

<https://www.asbanc.com.pe/Paginas/Noticias/DetalleNoticia.aspx?ItemID=762>

ASOCIACIÓN DE BANCOS DEL PERÚ

2019 “Servicios financieros y tendencias digitales en América Latina”. *ASBANC Semanal*. Lima, 2019, número 308. Consulta: 07 de abril de 2019.

<https://www.asbanc.com.pe/Publicaciones/ASBANC%20SEMANAL%20308.pdf>

CADENA, Jaime

2016 *Guía para el diseño y documentación de procesos*. En *Revista YURA*.

CASTILLO, Pablo y Luis, CERRÓN

2015 *Diagnóstico y propuestas de mejoras para el rediseño de los procesos, redistribución del almacén central, y el cálculo de la proyección de la demanda en una empresa comercializadora retail de productos deportivos*. Tesis para optar el título de Ingeniero Industrial. Lima: Pontificia Universidad Católica del Perú, Facultad de Ciencias e Ingeniería.

CHAPMAN, Chris y Elea MCDONELL

2015 *Use R!* Suiza: Springer Heidelberg New York Dordrecht London

CORPORACIÓN ACEROS AREQUIPA S.A.

2015 *POSTULACIÓN AL RECONOCIMIENTO A LA GESTIÓN DE PROYECTOS DE MEJORA*. En *Centro de Desarrollo Industrial*. Consulta: 25 de junio de 2019.

http://www.cdi.org.pe/pdf/PNC_2015/RGPM/ACEROS%20AREQUIPA%20IP.pdf

GALLEY, Mark

2018 “Cómo aplican 5 porques en Toyota y cómo mejorarlo”. En *Think Reliability*. Consulta: 25 de junio de 2019.

<https://es-thinkreliability.com/2018/05/03/como-aplican-5-porques-en-toyota-y-como-mejorarlo/>

GESTIÓN

2018 “Bancos atienden cada vez más por canales digitales, ante ello ¿reducirán el personal en sus agencias?”. En *Gestión*. Consulta: 07 de abril de 2019.

<https://gestion.pe/economia/bancos-atienden-vez-canales-digitales-reduciran-personal-agencias-249228>

HARVARD BUSINESS REVIEW

2014 “The Explainer: Big Data and Analytics”. Consulta: 25 de junio de 2019.

<https://hbr.org/video/3633937151001/the-explainer-big-data-and-analytics>

INEI

2019 “Producto Bruto Interno Trimestral según Actividad Económica (Nivel 14) 2007-2018-IV (Valores a precios constantes de 2007)” En *INEI*. Consulta: 07 de abril de 2019.

<https://www.inei.gob.pe/estadisticas/indice-tematico/economia/>

KATSARAS, Nikolaos, Paul, WOLFSON, Jean, KINSEY, Ben, SENAUER

2001 “Data Mining: A segmentation analysis of U.S. Grocery Shoppers”. En *The Retail Food Industry Center University of Minnesota*.

KHOBZI, Hamid, Elham, AKHONDZADEH-NOUGHABI y Behrouz, MINAEI-BIDGOLI

2014 “A New Application of RFM Clustering for Guild Segmentation to Mine the Pattern of Using Banks’ e-Payment Service”. *Journal of Global Marketing*. 2014, volumen 27, pp. 178-190. Consulta: 07 de abril de 2019.

<https://www.tandfonline.com/doi/abs/10.1080/08911762.2013.878428>

KUME Hitoshi

2002 *Herramientas estadísticas básicas para el mejoramiento de la calidad*. Bogotá: Editorial Norma.

MENDENHALL, William, Richard SCHEAFFER y Ott LYMAN

2007 *Elementos de muestreo*. España. Thomson editores.

PARDO, José

2012 *Configuración y usos de un mapa de procesos*. AENOR ediciones

PAZ, Sandro

2018 *Diapositivas de clase de Control Integral de Calidad*. Consulta: 07 de abril de 2019.

PERÚ 21

2018 “Uso de canales digitales ahorran hasta 80% en la interacción entre empresa y cliente”. En *Perú 21*. Consulta: 07 de abril de 2019.

<https://peru21.pe/peru/canales-digitales-ahorran-80-interaccion-empresa-cliente-428049>

R

2019 “R Documentation”. Consulta: 25 de junio de 2019.

<https://www.rdocumentation.org/>

RAO, Vithala

2014 *Applied Conjoint Analysis*. Berlin: Springer Heidelberg New York Dordrecht London.

REAL ACADEMIA ESPAÑOLA

1992 Significado de la palabra “Marketing”.

<https://dle.rae.es/?id=OyavUPb>

SEGURO SOCIAL DE SALUD –ESSALUD

2019 *Mapa de Macroprocesos*. Consulta: 25 de junio de 2019.

http://www.essalud.gob.pe/transparencia/procesos_procedimientos/Mapa_de_Procesos_2019.pdf

SERRAT, Olivier

2009 *The Five Whys Technique*. Consulta: 25 de junio de 2019.

<https://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=1200&context=intl>

SOTELO, Marjorie

2017 *diagnóstico y mejora para el servicio de la seguridad ciudadana en el distrito de San Martín de Porres*. Tesis para optar el título de Ingeniero Industrial. Lima: Pontificia Universidad Católica del Perú, Facultad de Ciencias e Ingeniería.

SUPERINTENDENCIA DE BANCA Y SEGUROS

2018 *Reporte de indicadores de inclusión financiera de los sistemas financieros, de seguros y de pensiones*. Consulta: 07 de abril de 2019.

<https://intranet2.sbs.gob.pe/estadistica/financiera/2018/Junio/CIIF-0001-jn2018.PDF>

TORREJÓN, Steven

2018 *Diseño de Estrategias de Retención de Posventa en una Empresa Automotriz de Vehículos Livianos aplicando Marketing Analytics*. Tesis para optar el grado académico de Magíster en Ingeniería Industrial con Mención en Gestión de Operaciones. Lima: Pontificia Universidad Católica del Perú, Facultad de Ciencias e Ingeniería.

UNIVERSITÄT LEIPZIG

2009 “Quality function deployment (QFD) for services”. En International SEPT Program. Leipzig, 2009.

VALDIVIESO, Luis

2019 “Técnicas de muestreo: Notas de clase”. En Pontificia Universidad Católica del Perú Escuela de Posgrado: Maestría en Estadística.

WINSTON, Wayne

2014 *Marketing Analytics: Data-Driven Techniques with Microsoft Excel*. Indianapolis: John Wiley & Sons, Inc.

VÉLIZ, Carlos

2018 *Aprendizaje automático. Análisis para la minería de datos y big data*. En Pontificia Universidad Católica del Perú. Sección Matemáticas. Departamento académico de ciencias.

YUSUF, Samson y Jongsu LEE

2015 “Quality of service and customer satisfaction: a conjoint analysis for the Nigerian bank customers”. En *International Journal of Bank Marketing*. 2016, volumen 34, número 6, pp. 841-867.

<https://www.emeraldinsight.com/loi/ijbm>

Anexos

1.1 Ejemplo de la herramienta Ficha indicador

Ficha de Indicador:

- C1.a Costos generados por la venta de colchones

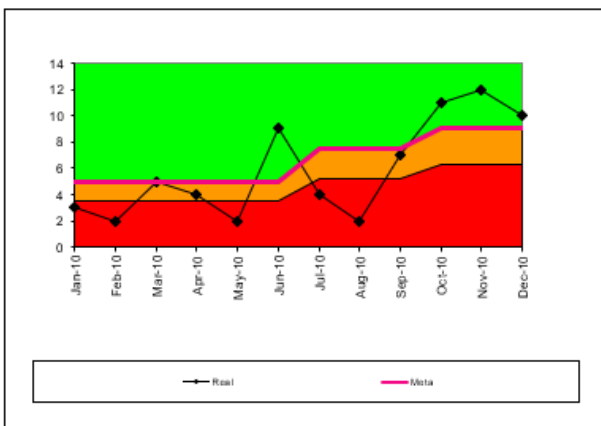
MD: No Definido
MF: Rango Mal Definido

Definiciones Generales

Fórmula / Cálculo:	Costos generados por los nuevos productos en millones de soles			
Responsable:	PS	Tipo:	C	Uaidad: %
Fuente / Procesamiento:	Informes de Venta mensuales, aprobados por contabilidad			
Frecuencia de Medición:	Mensual	Oportunidad:	Último día de cada mes	
Glosario	Nuevo producto = Aquel lanzado hace 6 meses o menos			

Fecha Cuatrimestre	Mar-10
Real	5.0
Meta	5.0
Verde	5.0
Rojos	3.5
100%	100%

Definiciones Específicas



Serie Principal				
Fecha	Real	Meta	Verde	Rojos
Jan-10	3.00	5.0	5.0	3.5
Feb-10	2.00	5.0	5.0	3.5
Mar-10	5.00	5.0	5.0	3.5
Apr-10	4.00	5.0	5.0	3.5
May-10	2.00	5.0	5.0	3.5
Jun-10	3.00	5.0	5.0	3.5
Jul-10	4.00	7.5	7.5	5.3
Aug-10	2.00	7.5	7.5	5.3
Sep-10	7.00	7.5	7.5	5.3
Oct-10	11.00	3.0	3.0	6.3
Nov-10	12.00	3.0	3.0	6.3
Dec-10	10.00	3.0	3.0	6.3
Jan-11		3.0	3.0	6.3
Feb-11		3.0	3.0	6.3
Mar-11		3.0	3.0	6.3
Apr-11		3.0	3.0	6.3
May-11		3.0	3.0	6.3
Jun-11		3.0	3.0	6.3

Información de Control

Comentarios (Adjuntar Documentación Sustentatoria de Causas y de Acciones)					
Fecha de Control	Riesgos / Problemas	Acciones Preventivas / Correctivas	Resp.	Fecha	Terminado
Mar-10	Se tendrá un nuevo competidor ingresante al mercado	Fidelizar a los clientes con ofertas en los nuevos productos	JP	31-May	T

Imagen 1.1. Ejemplo de ficha indicador. Fuente: Elaboración propia.

1.2 Ejemplo de la herramienta Cinco Por qué.

5 por qué en Mapa de Causas

Básico - Lineal (35 palabras)

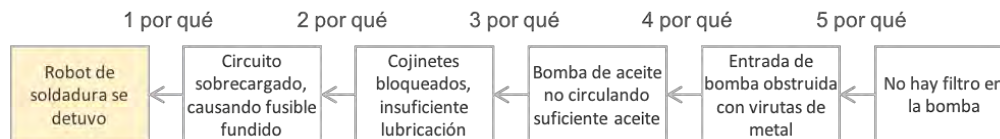


Imagen 1.2. Ejemplo de ficha indicador. Fuente: Galley (2018).