

PONTIFICIA UNIVERSIDAD  
CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS SOCIALES



Modelo de predicción de demanda y abandono de clientes

Trabajo de Suficiencia Profesional para optar el título de Licenciado en  
Economía presentado por:

Wagner Wagner, Herbert

Asesor:

Orihuela Paredes, Jose Carlos

Lima, 2022

## Informe de Similitud

Yo, Jose Carlos Orihuela Paredes, docente de la Facultad de Ciencias Sociales de la Pontificia Universidad Católica del Perú, asesor del trabajo de investigación titulado Modelo de predicción de demanda y abandono de clientes, del autor Herbert Wagner Wagner.

dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 23%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 06/12/2022.
- 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: Lima, 07 de Diciembre de 2022

Apellidos y nombres del asesor: Orihuela Paredes, Jose Carlos	
DNI: 06785398	
ORCID: 0000-0002-2999-5561	
Firma	

## Agradecimientos

En primer lugar, a mi madre por apoyarme a lo largo de mi desarrollo profesional, así como en mi formación personal brindándome principios, valores y su amor incondicional y, en segundo lugar, a mi hermano, por sus consejos, conocimientos y confianza para alcanzar mis metas. Les agradezco infinitamente por ser los dos pilares en mi vida sobre los cuales pude forjar la persona en la que me convertí.



## **Resumen**

La experiencia profesional que sustentó, en el presente trabajo de suficiencia profesional, es sobre la puesta en práctica de los conceptos y criterios que aprendí en el curso de Econometría II, la cual desarrollé al desempeñar el cargo de analista de pricing en el área de finanzas de OPP Film Perú S.A. La teoría está enfocada en los modelos de series de tiempo tomando en cuenta el método Holt-Winters y el modelo de abandono (Customer Churn), a través de regresión logística.

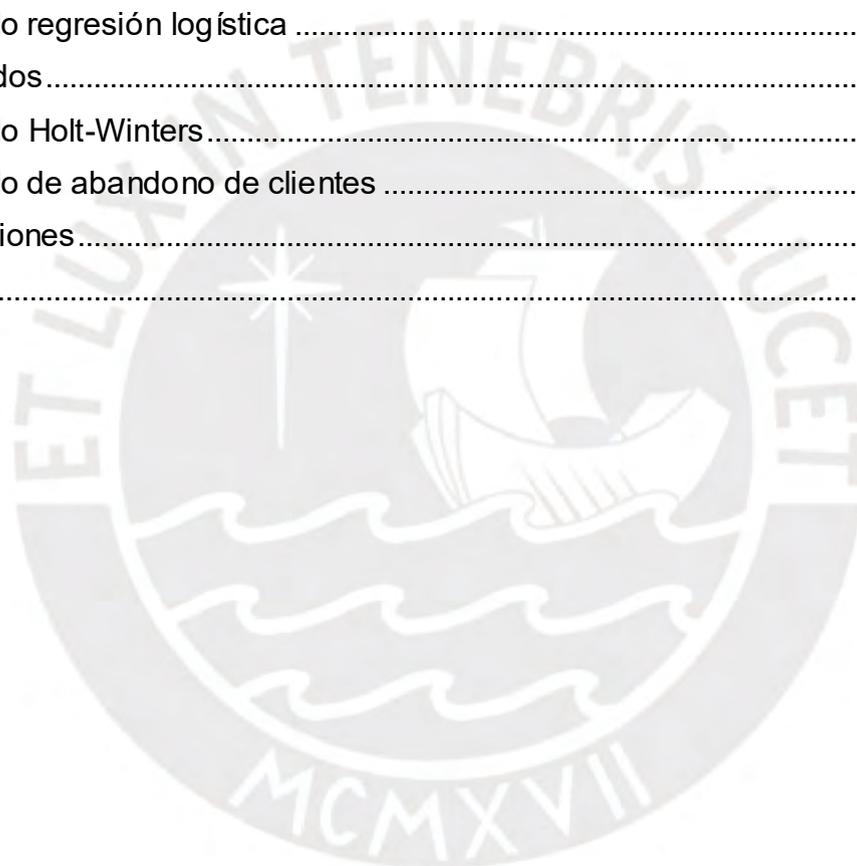
En el primer modelo, se desarrolla el análisis con base en la optimización de proyección de ventas, considerando el nivel, tendencia y estacionalidad de la serie; y en el segundo, con base en la predicción de abandono de clientes mediante los comportamientos y atributos precisos del cliente que señalan el riesgo y el momento del abandono de este. La importancia de ambos análisis radica en generar herramientas de información con datos cuantitativos y cualitativos que permitan visualizar oportunidades para el beneficio de la empresa.

## **Palabras Clave**

Pricing, churn, Holt-Winters, abandono, clientes

## Índice de trabajo de suficiencia profesional

Introducción .....	1
1. Contexto profesional .....	3
2. Literatura de referencia .....	4
2.1. Modelo de predicción de series.....	4
2.2. Modelo de abandono de clientes .....	5
3. Base de datos .....	7
4. Metodología.....	8
4.1. Modelo Holt-Winters.....	8
4.2. Modelo regresión logística .....	9
5. Resultados.....	10
5.1. Modelo Holt-Winters.....	10
5.2. Modelo de abandono de clientes .....	19
6. Conclusiones.....	22
Bibliografía .....	23



## Índice de tablas

Tabla 1 Argentina: Variación TM 2017-2018(p).....	16
Tabla 2 Chile: Variación TM 2017-2018(p).....	17
Tabla 3 Perú: Variación TM 2017-2018(p) .....	18



## Índice de figuras

Figura 1 Iniciando el Modelo Argentina.....	10
Figura 2 Films Bopp : Argentina Demanda Des-estacionalizada 2015-2016 .....	10
Figura 3 Films Bopp : Argentina Demanda Des-estacionalizada 2015-2017 .....	10
Figura 4 Argentina: serie, nivel y tendencia -2017 .....	11
Figura 5 Argentina: serie, nivel y tendencia -2018 .....	11
Figura 6 Bopp : Argentina Ventas Históricas (2015-2017) y Estimadas (2018) .....	11
Figura 7 Iniciando el Modelo Chile.....	12
Figura 8 Films Bopp : Argentina Demanda Des-estacionalizada 2015-2017 .....	12
Figura 9 Films Bopp : Argentina Demanda Des-estacionalizada 2015-2018 .....	12
Figura 10 Chile: serie, nivel y tendencia -2017 .....	13
Figura 11 Chile: serie, nivel y tendencia -2018.....	13
Figura 12 Argentina Ventas Históricas (2015-2017) y Estimadas (2018) .....	13
Figura 13 Iniciando el Modelo Perú .....	14
Figura 14 Films Bopp : Perú Demanda Des-estacionalizada 2015-2017.....	14
Figura 15 Films Bopp : Perú Demanda Des-estacionalizada 2015-2018.....	14
Figura 16 Perú: serie, nivel y tendencia -2017 .....	15
Figura 17 Perú: serie, nivel y tendencia -2018 .....	15
Figura 18 Bopp : Perú Ventas Históricas (2015-2017) y Estimadas (2018) .....	15
Figura 19 Cuadro estadístico .....	19
Figura 20 Cuadro resultado Churn total .....	19
Figura 21 Cuadro resultado Churn por países.....	20
Figura 22 Matriz de confusión y métricas asociadas.....	20

## **Introducción**

Mi experiencia en el presente trabajo de suficiencia profesional, es sobre la puesta en práctica de los modelos de series de tiempo tomando en cuenta el método Holt-Winters y el modelo de abandono (Customer Churn), a través de regresión logística. El problema que se analiza, radica en la diferenciación de precios con valor agregado para los bienes de consumo masivo dentro de un mercado de competencia perfecta, de manera que se genere un mayor beneficio sin aumentar los costos;

En los últimos años, el área de pricing se ha abierto campo en las empresas de consumo masivo, puesto que los mercados han generado alta competitividad y las empresas se han visto en la necesidad de implementar un área de pricing, debido a que esta tiene como función principal integrar información de diferentes áreas dentro de una empresa, ello conlleva a una mejor toma de decisiones de precios, los cuales van a permitir maximizar mayores beneficios.

Dentro del área de pricing de OPP Film, las funciones fueron de desarrollo y facilitación de mejoras de procesos y modelos de pricing entre ellos la predicción de demanda y el abandono del cliente. Cabe resaltar que se desarrollaron modelos cuantitativos, tal como el método de Holt-Winters Model y por otro lado modelos cualitativos, tal como el Customer Churn Model.

Por un lado, si bien es cierto, dentro de las series de tiempo existen una amplia variedad de modelos de predicción, los autorregresivos, los de medias móviles o ambos, entre otros, para ello se debe analizar según sea el caso.

La predicción es una herramienta que permite estimar los valores dentro de una serie, sujeto ciertas variables y parámetros preestablecidos. Sin embargo, la amplia gama de métodos de predicción genera una dificultad en la elección de este.

Una categoría de métodos muy conocida y usada en predicciones de series es la del método Holt-Winters, esta es una ampliación perfeccionada del enfoque de la suavización exponencial, ya que consiste en encontrar un grupo de observaciones con características observables y similares al grupo que recibió el tratamiento. Este método se adecua a una serie con variables de patrones determinísticos y generalmente es usado para ventas.

En el caso del modelo de abandono, se utiliza un modelo de regresión logística múltiple, que permite predecir (tomando en cuenta variables cualitativas) el abandono de clientes debido a las características de dichas variables.

Con el objetivo de organizar la información del presente informe, se presenta de acuerdo con las siguientes secciones (1) Contexto profesional (2) Literatura de referencia; (3) Base de datos; (4) Metodología; (5) Resultados; (6) Conclusión y Bibliografía.



## 1. Contexto profesional

Las dificultades que las empresas deben enfrentar, de acuerdo con el tipo de mercado en el que se encuentran, son diversas; por lo cual se ven en la necesidad de crear estrategias específicas con la finalidad de maximizar el beneficio del producto o servicio que se oferta.

Las empresas latinoamericanas se han venido adaptando a las nuevas herramientas en lo que corresponde a la gestión de ingresos y costos, desde su aparición en 1985 de la mano de American Airlines (Cross, 1997; Smith et al, 1992; Castelló 2016); estas herramientas se han ido desarrollando en las áreas de revenue management y pricing.

Para Cheraghi et al (2010), el revenue management es definido como la ciencia que usa el pasado histórico y los niveles actuales de volumen de actividad de una empresa para prever la demanda de la manera precisa, con el objetivo de poder adaptarse a las decisiones de precio y de disponibilidad de productos, en los diferentes canales de venta, y de esta forma maximizar la rentabilidad.

Por ello, el reto dentro de las áreas de revenue management y pricing es tratar de usar la información de datos cuantitativa y cualitativa para poder reaccionar anticipadamente ante los comportamientos de los clientes, y de esta forma tomar acciones correspondientes, para esto es que se presenta tanto el modelo de predicción de demanda como el modelo de abandono de clientes.

## 2. Literatura de referencia

A continuación, se presentarán conceptos y evidencias de desarrollo de modelos de predicción de ventas y de abandono de clientes, tomando como referencias algunos trabajos que han analizado estos temas.

### 2.1. Modelo de predicción de series

Para entender el desarrollo y el uso del modelo de predicción de series, tomamos en consideración la siguiente afirmación:

Para Nojek et al: “La elección e implementación de un método adecuado de pronósticos siempre ha sido un tema de gran importancia para las empresas. Se utilizan los pronósticos en el área de compras, marketing, ventas, etc. Un error significativo en el pronóstico de ventas podría dejar a una empresa sin la materia prima o insumos necesarios para su producción, o podría generarle un inventario demasiado grande. En ambos casos, el pronóstico erróneo disminuye las utilidades de la empresa” (2003, p.1).

Partiendo de lo afirmado anteriormente por Nojek, se debe agregar, según Sapag Chain (2000), que la estimación del comportamiento futuro de algunas variables puede realizarse utilizando diversas técnicas de pronóstico. Cada una de las técnicas de proyección tiene una aplicación de carácter especial que hace de su selección un problema de decisión influido por diversos factores, como la validez y disponibilidad de los datos históricos, la precisión deseada del pronóstico, los periodos futuros que se desee pronosticar (como se citó en Nojek, 2003, p.1).

Ahora bien, existen varios modelos de predicción de acuerdo con el tipo de serie que se desea analizar; en ese sentido, para el presente informe se toman referencias de los siguientes trabajos:

Angel y Marquez (2018), en su trabajo para resolver el problema de establecer un adecuado pronóstico de demanda de los productos para exportación, plantea desarrollar un modelo matemático y de gestión que permita generar una combinación de pronósticos. Para lo cual se centra en los modelos de Holt Winters, ARIMA y Redes Neuronales; de los resultados obtenidos se pudo demostrar que se logró una reducción del error de pronóstico

comparada con la anterior metodología que la empresa llevaba.

Por otro lado, Nojek et al (2003), en su trabajo de investigación plantea usar un método de pronóstico no tradicional conocido como redes neuronales para obtener un pronóstico de ventas que resulte tener mayor precisión en comparación a los resultados que se pudieran obtener usando un método de pronóstico estadístico clásico. Para ello desarrolla 12 métodos de pronóstico clásicos y un método de Redes neuronales usando el software Neural Network, los resultados le permitieron concluir que con un nivel de confianza del 82.29%, las redes neuronales realizaron un pronóstico más acertado en comparación a los 12 métodos estadísticos utilizados.

En resumen, de los métodos de predicción de demanda se puede inferir una característica notable, pues su precisión esta correlacionada al plazo de predicción; por lo que, el peso de los datos históricos más recientes es mayor, asimismo la predicción de corto plazo tiende a ser más precisa que la de largo plazo.

## **2.2. Modelo de abandono de clientes**

Para entender el desarrollo y el uso del modelo de abandono de clientes, tomamos en consideración la siguiente afirmación:

El índice de rotación clientes es uno de los problemas que enfrentan las empresas, pues existen costos asociados a generar nuevos clientes, los cuales son relativamente mayores que mantener los ya existentes. El resultado, como parte de un plan estratégico de consumo dirigido a disminuir la rotación de clientes, el descubrimiento y la exploración para los clientes con un fuerte deseo de dejar la organización, o la predicción de abandono de clientes, es esencial (Blattberg et al. 2008)] (Guliyev & Yerdelen, 2021, p.2).

Es necesario resaltar el objetivo a lograr cuando se desarrollan herramientas de este tipo, ya que “en cualquier negocio donde los costos de captación son elevados y los márgenes de ganancia del producto o servicio son pequeños, la manera más eficiente para generar rentabilidad es con base al consumo recurrente de los productos o servicios de la compañía. Por tanto, minimizar la tasa de abandono de los clientes adecuados se vuelve un pilar de

supervivencia de las compañías” (Castro & Pérez, 2020, p. 612).

De las afirmaciones de Guliyev y Castro, resulta clara la necesidad de las empresas por desarrollar un modelo de predicción de abandono de cliente, en el cual se pueda actualizar y moldear según las características del mercado.

Por un lado, se tiene el método de minería de datos, según lo afirmado por Jélvez et al: “la aplicación técnica de minería de datos a una empresa le permite identificar factores que faciliten clasificar clientes nuevos y antiguos como potencialmente fugables por medio de conglomerados, además, de generar un modelo predictor que estime la probabilidad de fuga como función de otras variables. En este sentido, la minería de datos permite descubrir información en forma de patrones, cambios, asociaciones y estructuras significativas de grandes cantidades de datos almacenados” (2014, p. 2)

Por otro lado, Jain et al (2020), en su trabajo de investigación, utilizando regresión Logística y Logit Boost, pretendió mejorar la precisión aumentando medidas de rendimiento. Los resultados del análisis concluyen lo siguiente: “This model was measured using many performance measures. These performance measures showed us that both the techniques outperformed. There was not much difference in the results of both techniques. Logistic regression had an accuracy 85.2385% where Logit Boost also had an accuracy 85.1785%. This model used two standalone techniques that performed well but standalone techniques cannot have all features that can raise accuracy.” [Este modelo se midió utilizando muchas medidas de rendimiento, dichas medidas de rendimiento nos mostraron que ambas técnicas se desempeñaron mejor. No hubo mucha diferencia en los resultados de ambas técnicas. La regresión logística tuvo una precisión del 85,2385 %, y la Logit Boost también tuvo una precisión del 85,1785 %. Este modelo usó dos técnicas independientes que funcionaron bien, pero las técnicas independientes no pueden tener todas las características para aumentar la precisión necesariamente.] (Jain et al, 2020, p.111)

Como se observó en los trabajos anteriormente expuestos, la variedad de modelos para desarrollar la predicción de abandono de clientes es amplia; sin embargo, cada uno de estos se aplica según las características e información que se tenga de la empresa.

### 3. Base de datos

La base de datos que se utilizó para los análisis tanto para la predicción de demanda como para el modelo de abandono, son de la misma empresa OPP film.

Los datos cualitativos son: ventas en toneladas métricas mensualizadas del 2015 al 2017, Nombre del cliente, código de país, flete, tiempo de demora en entrega, descuento, años de servicio en el mercado, calidad de servicio y abandono del cliente.

Para el caso del modelo de abandono de clientes, se generó una encuesta VOC (Voice of Consumer), esta consiste en desarrollar una encuesta en la que se pueda plasmar la experiencia del cliente con nosotros en ciertas características cualitativas.



#### 4. Metodología

##### 4.1. Modelo Holt-Winters

Debido a que el objetivo es demostrar el uso del modelo de predicción de demanda, se usarán datos de ventas mensuales de 2015-2017. Asimismo, se utilizará la minimización de errores al cuadrado para optimizar los parámetros  $\alpha$ ,  $\beta$ , y  $\gamma$ , con ello se estimará los valores mensuales de venta para el mercado de Argentina, Chile y Perú para el periodo del 2018.

##### Modelo Holt-Winters

$$S_t = \alpha \left( \frac{d_t}{C_{t-L}} \right) + (1 - \alpha)(S_{t-1} + B_{t-1})$$

$$C_t = \gamma \left( \frac{d_t}{S_t} \right) + (1 - \gamma) C_{t-L}$$

$$B_t = \beta(S_t + S_{t-1}) + (1 - \beta)(B_{t-1})$$

$$F_{t+k} = (S_t + kB_t) + (C_{t+k-L})$$

Donde:

$S_t$  = Valor suavizado para el periodo t

$C_t$  = Estacionalidad estimada

$B_t$  = Tendencia estimada

$d_t$  = Valor real de la demanda en el periodo t

$S_{t-1}$  = Pronóstico suavizado para el periodo t - 1

$B_{t-1}$  = Tendencia estimada en el periodo t - 1

$\alpha$  = Constante de suavizamiento para la serie

$\beta$  = Constante de suavizamiento para la tendencia estimada

$\gamma$  = Constante de suavizamiento para la estacionalidad estimada

$F_{t+k}$  = Pronóstico para k periodos futuros

$L$  = Longitud de la estacionalidad

Puesto que el objetivo es demostrar el uso del Modelo de abandono de clientes (churn model) a través de machine learning (PYTHON), se usarán datos mensuales de 2015-2017 cualitativos generados a partir de una encuesta realizada por la empresa OPP Film Perú (Voice of Consumer).

## 4.2. Modelo regresión logística

### Modelo Regresión Logística

$$p(y = 1 | (x_1, x_2, \dots, x_n)) = f(y)$$

$$f(y) = \frac{1}{(1 + e^{-y})}$$

Donde:

$y$  una variable dependiente binaria con dos posibles valores: 0 y 1 (cliente activo o churn) y un conjunto de variables independientes,  $(x_1, x_2, \dots, x_n)$  que son capaces de predecir el valor de  $y$ .

Donde:

$X_1 = \text{flete}$

$X_2 = \text{tiempo de demora en entrega}$

$X_3 = \text{descuento}$

$X_4 = \text{años de servicio en el mercado}$

$X_5 = \text{calidad de servicio}$

$y = \text{churn o non churn}$

La variable churn se plantea como respuesta del modelo, de tal manera que:

$$y_i = \begin{cases} 0 & \text{si el usuario esta en estado Activo} \\ 1 & \text{si el usuario está en estado Churn} \end{cases}$$

Se observa a  $y_i$  como la generación de una variable aleatoria  $y_i$  que puede tomar los valores uno y cero con probabilidades  $\rho$  y  $1 - \rho$ , respectivamente.

## 5. Resultados

### 5.1. Modelo Holt-Winters

En los resultados se puede observar que los pronósticos de ventas responden a distintos factores que afectan al nivel, tendencia y estacionalidad de las series de cada país, según se muestra en las siguientes figuras:

Figura 1

#### Iniciando el Modelo Argentina

2015 - 2016	a	634.9
	b	3.2

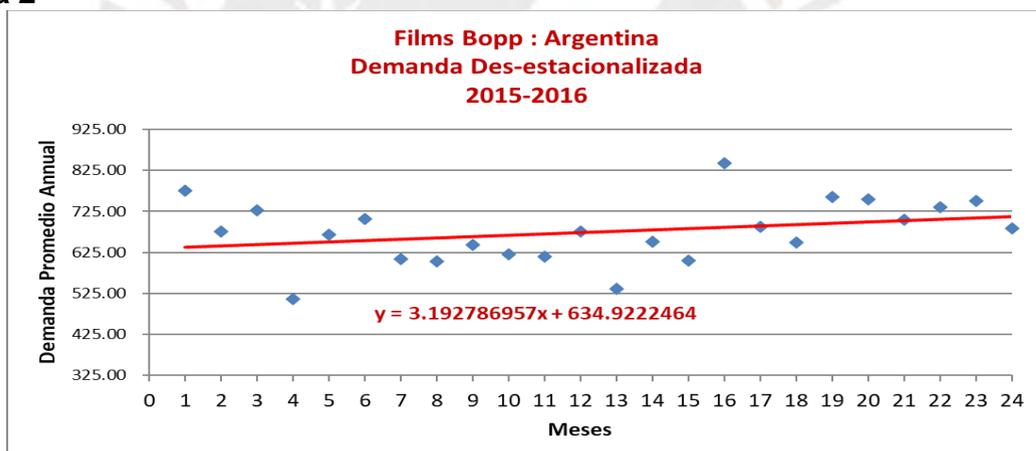
Alfa	0.3192
Beta	0.0243
Gamma	0.5180

$$a = \frac{\sum X^2 \sum Y - \sum X \sum XY}{N \sum X^2 - (\sum X)^2}$$

$$b = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

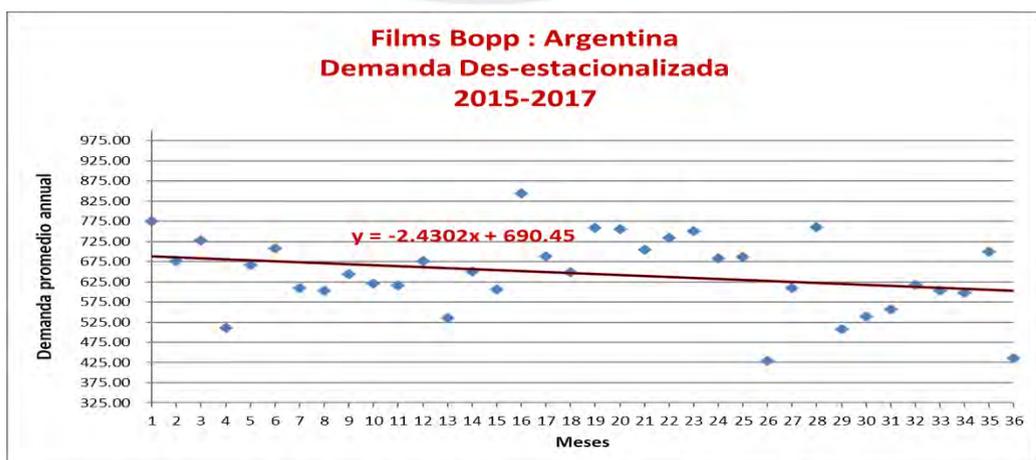
Fuente: Elaboración Propia

Figura 2



Fuente: Elaboración Propia

Figura 3



Fuente: Elaboración Propia

Figura 4

## Argentina: serie, nivel y tendencia -2017

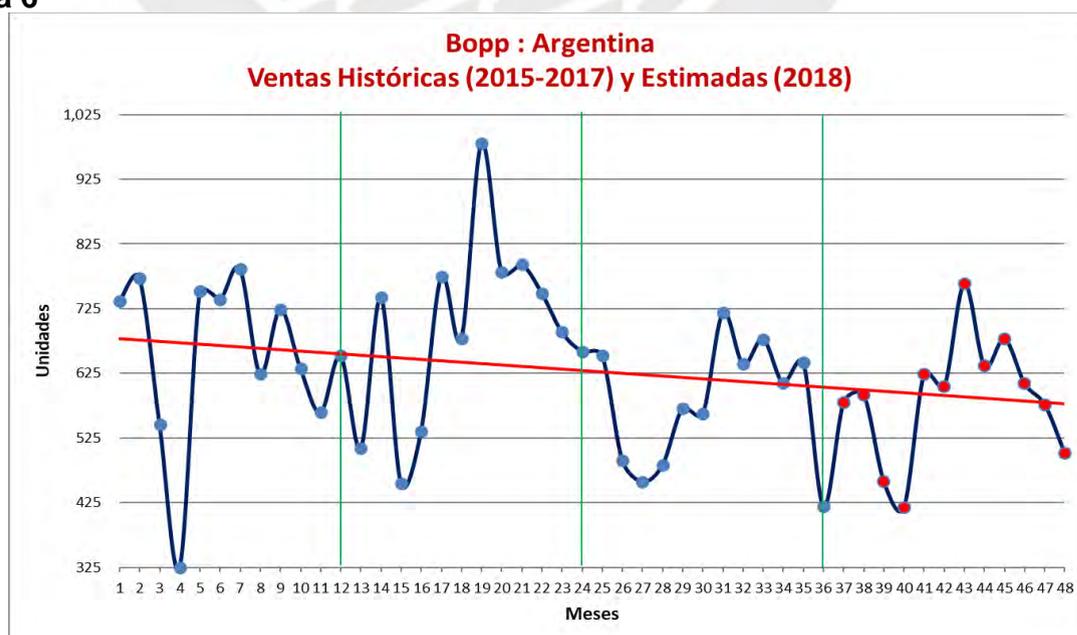
Año	Mes		Ventas históricas	Nivel	Tendencia	Factor de estacionalidad ajustado	Pronóstico 2017
1	2	3	4	5	6	7	8
	24			<b>711.55</b>	<b>3.1928</b>		
2017	25	Enero	653	705.92	2.9785	0.93657	<b>679</b>
	26	Febrero	490	619.80	0.8141	0.95964	<b>809</b>
	27	Marzo	457	617.18	0.7307	0.74506	<b>465</b>
	28	Abril	483	663.50	1.8382	0.68346	<b>393</b>
	29	Mayo	571	614.89	0.6127	1.02430	<b>749</b>
	30	Junio	562	590.86	0.0141	0.99667	<b>643</b>
	31	Julio	719	579.84	(0.2539)	1.26469	<b>763</b>
	32	Agosto	639	591.69	0.0401	1.05894	<b>600</b>
	33	Septiembre	678	595.08	0.1215	1.13247	<b>666</b>
	34	Octubre	609	596.03	0.1416	1.02100	<b>607</b>
	35	Noviembre	641	629.11	0.9417	0.97031	<b>547</b>
	36	Diciembre	420	567.89	(0.5683)	0.84726	<b>607</b>
			<b>6,923</b>			<b>11.64037</b>	<b>7,528</b>

Figura 5

## Argentina: serie, nivel y tendencia -2018

Año	Mes		Ventas históricas	Nivel estimado	Tendencia	Factor de estacionalidad normalizado	Ventas - pronóstico 2018
	36			<b>602.96</b>	<b>(2.4302)</b>		
2018	37	Enero		600.53	(2.4302)	0.96551	<b>580</b>
	38	Febrero		598.10	(2.4302)	0.98929	<b>592</b>
	39	Marzo		595.67	(2.4302)	0.76808	<b>458</b>
	40	Abril		593.24	(2.4302)	0.70458	<b>418</b>
	41	Mayo		590.81	(2.4302)	1.05595	<b>624</b>
	42	Junio		588.38	(2.4302)	1.02746	<b>605</b>
	43	Julio		585.95	(2.4302)	1.30376	<b>764</b>
	44	Agosto		583.52	(2.4302)	1.09166	<b>637</b>
	45	Septiembre		581.09	(2.4302)	1.16746	<b>678</b>
	46	Octubre		578.66	(2.4302)	1.05254	<b>609</b>
	47	Noviembre		576.23	(2.4302)	1.00029	<b>576</b>
	48	Diciembre		573.79	(2.4302)	0.87344	<b>501</b>
					<b>12.00002</b>	<b>7,041</b>	

Figura 6



Fuente: Elaboración Propia

Figura 7

Iniciando el Modelo Chile

2015 - 2016

a	699.7
b	(8.9)

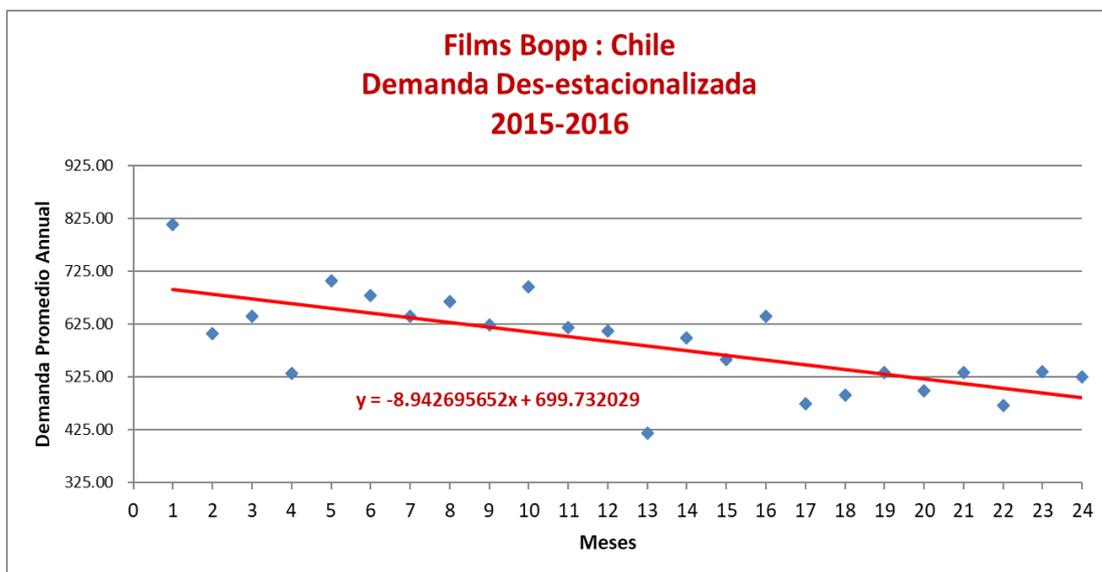
Alfa	0.4741
Beta	0.0000
Gamma	0.1428

$$a = \frac{\sum X^2 \sum Y - \sum X \sum XY}{N \sum X^2 - (\sum X)^2}$$

$$b = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

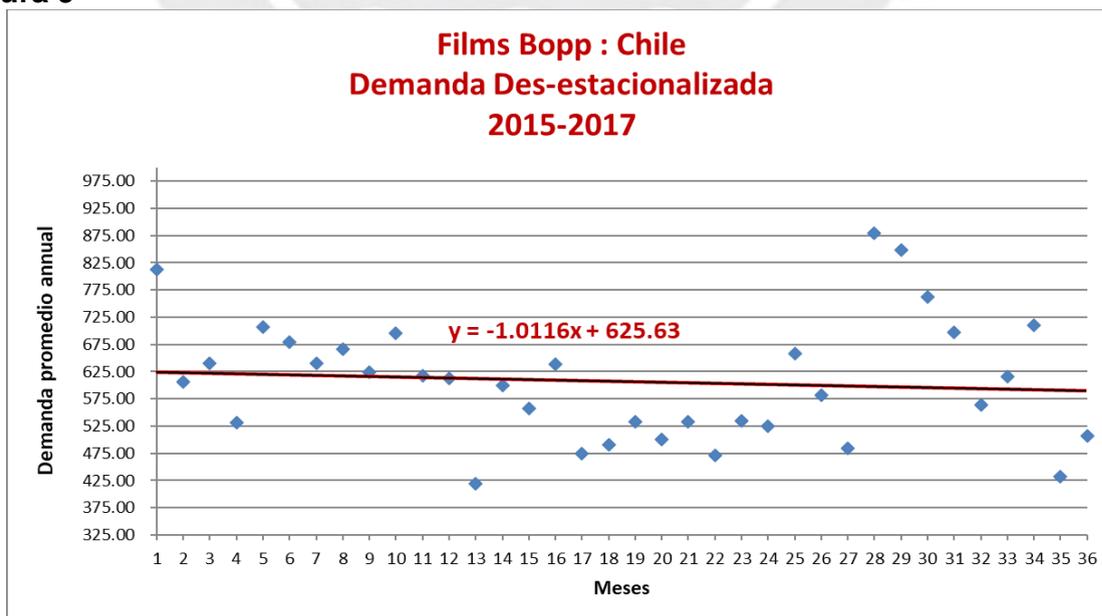
Fuente: Elaboración Propia

Figura 8



Fuente: Elaboración Propia

Figura 9



Fuente: Elaboración Propia

Figura 10

Chile: serie, nivel y tendencia -2017

Año	Mes	Ventas históricas	Nivel	Tendencia	Factor de estacionalidad ajustado	Pronóstico 2017	
1	2	3	4	5	6	7	
	24		<b>485.11</b>	<b>(8.9427)</b>			
2017	25	Enero	459	562.66	(8.9425)	0.71430	<b>332</b>
	26	Febrero	626	566.80	(8.9425)	1.08079	<b>596</b>
	27	Marzo	593	522.72	(8.9426)	1.21347	<b>684</b>
	28	Abril	818	687.18	(8.9423)	0.96670	<b>478</b>
	29	Mayo	726	758.60	(8.9421)	0.87085	<b>581</b>
	30	Junio	762	755.84	(8.9421)	1.00060	<b>749</b>
	31	Julio	704	723.29	(8.9421)	1.00495	<b>754</b>
	32	Agosto	621	642.64	(8.9422)	1.08278	<b>787</b>
	33	Septiembre	660	625.59	(8.9422)	1.06875	<b>679</b>
	34	Octubre	683	660.77	(8.9421)	0.97184	<b>593</b>
	35	Noviembre	509	547.37	(8.9423)	1.14482	<b>769</b>
	36	Diciembre	450	523.23	(8.9423)	0.88434	<b>478</b>
		<b>7,611</b>				<b>12.00419</b>	<b>7,480</b>

Fuente: Elaboración Propia

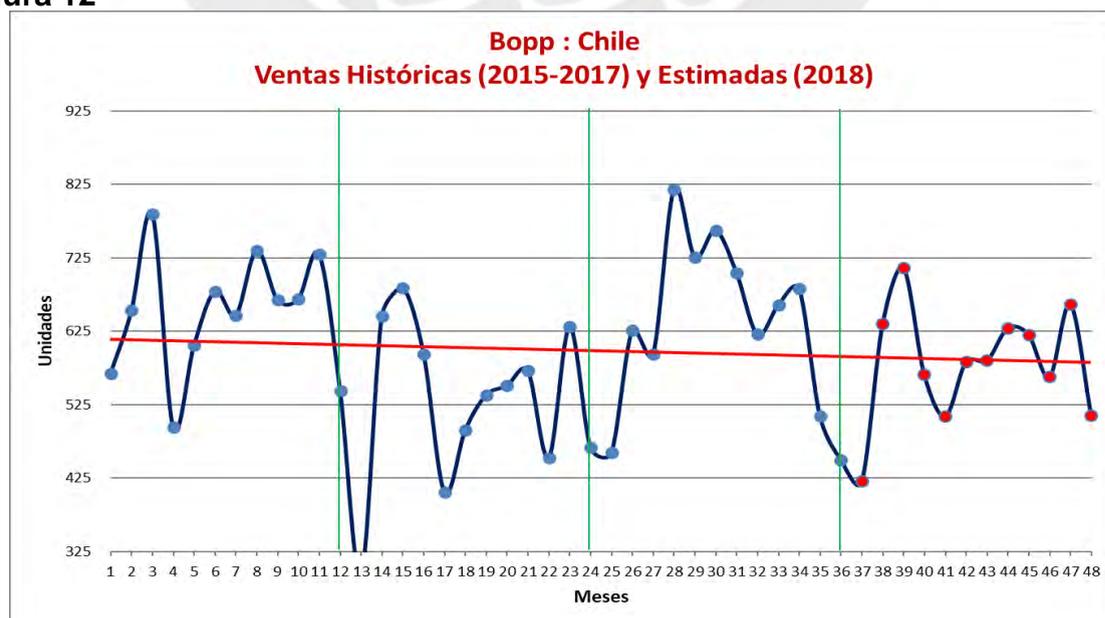
Figura 11

Chile: serie, nivel y tendencia -2018

Año	Mes	Ventas históricas	Nivel estimado	Tendencia	Factor de estacionalidad normalizado	Ventas - pronóstico 2018
	36		<b>589.21</b>	<b>(1.0116)</b>		
2018	37	Enero	588.20	(1.0116)	0.71405	<b>420</b>
	38	Febrero	587.18	(1.0116)	1.08041	<b>634</b>
	39	Marzo	586.17	(1.0116)	1.21305	<b>711</b>
	40	Abril	585.16	(1.0116)	0.96636	<b>565</b>
	41	Mayo	584.15	(1.0116)	0.87055	<b>509</b>
	42	Junio	583.14	(1.0116)	1.00025	<b>583</b>
	43	Julio	582.13	(1.0116)	1.00460	<b>585</b>
	44	Agosto	581.11	(1.0116)	1.08240	<b>629</b>
	45	Septiembre	580.10	(1.0116)	1.06838	<b>620</b>
	46	Octubre	579.09	(1.0116)	0.97150	<b>563</b>
	47	Noviembre	578.08	(1.0116)	1.14442	<b>662</b>
	48	Diciembre	577.07	(1.0116)	0.88403	<b>510</b>
					<b>12.00000</b>	<b>6,991</b>

Fuente: Elaboración Propia

Figura 12



Fuente: Elaboración Propia

Figura 13

Iniciando el Modelo Perú

2015 - 2016

a	989.9
b	(6.3)

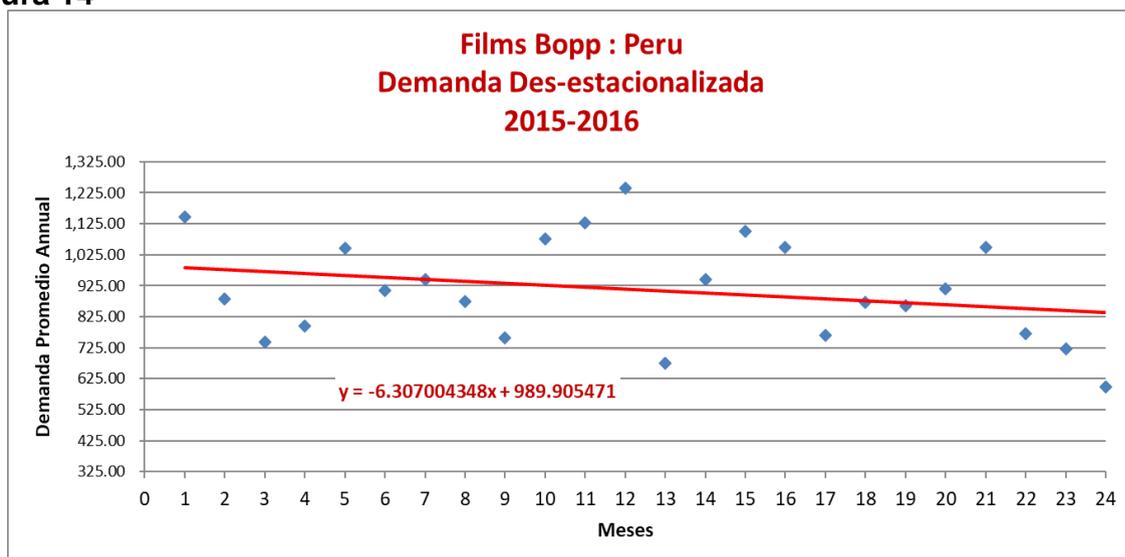
Alfa	0.3980
Beta	0.0000
Gamma	0.5606

$$a = \frac{\sum X^2 \sum Y - \sum X \sum XY}{N \sum X^2 - (\sum X)^2}$$

$$b = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

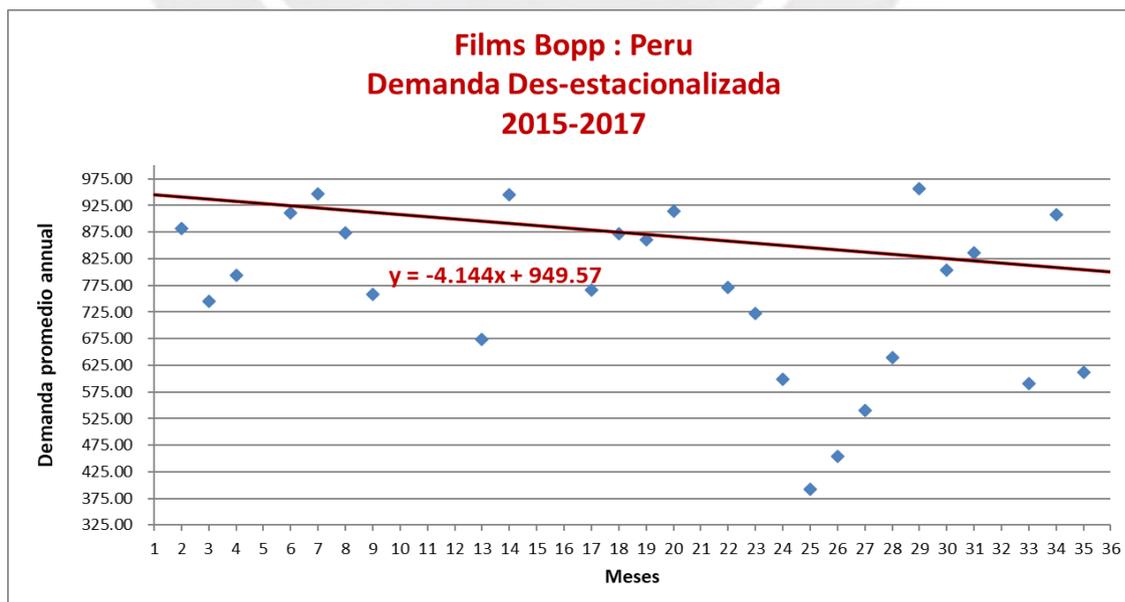
Fuente: Elaboración Propia

Figura 14



Fuente: Elaboración Propia

Figura 15



Fuente: Elaboración Propia

Figura 16

## Perú: serie, nivel y tendencia -2017

Año	Mes		Ventas históricas	Nivel	Tendencia	Factor de estacionalidad ajustado	Pronóstico 2017
1	2	3	4	5	6	7	8
	24			<b>838.54</b>	<b>(6.3070)</b>		
2017	25	Enero	412	656.94	(6.3095)	0.81329	<b>875</b>
	26	Febrero	604	572.53	(6.3106)	1.17494	<b>864</b>
	27	Marzo	670	555.87	(6.3108)	1.22039	<b>702</b>
	28	Abril	733	585.60	(6.3103)	1.20562	<b>630</b>
	29	Mayo	965	729.29	(6.3081)	1.18575	<b>585</b>
	30	Junio	730	755.12	(6.3076)	0.94054	<b>656</b>
	31	Julio	863	783.27	(6.3071)	1.07103	<b>773</b>
	32	Agosto	1,391	1,084.75	(6.3026)	1.11298	<b>697</b>
	33	Septiembre	561	884.59	(6.3054)	0.77185	<b>1,022</b>
	34	Octubre	875	890.21	(6.3052)	0.97440	<b>846</b>
	35	Noviembre	489	775.69	(6.3068)	0.70465	<b>706</b>
	36	Diciembre	863	970.66	(6.3039)	0.79556	<b>521</b>
			<b>9,155</b>			<b>11.97100</b>	<b>8,877</b>

Fuente: Elaboración Propia

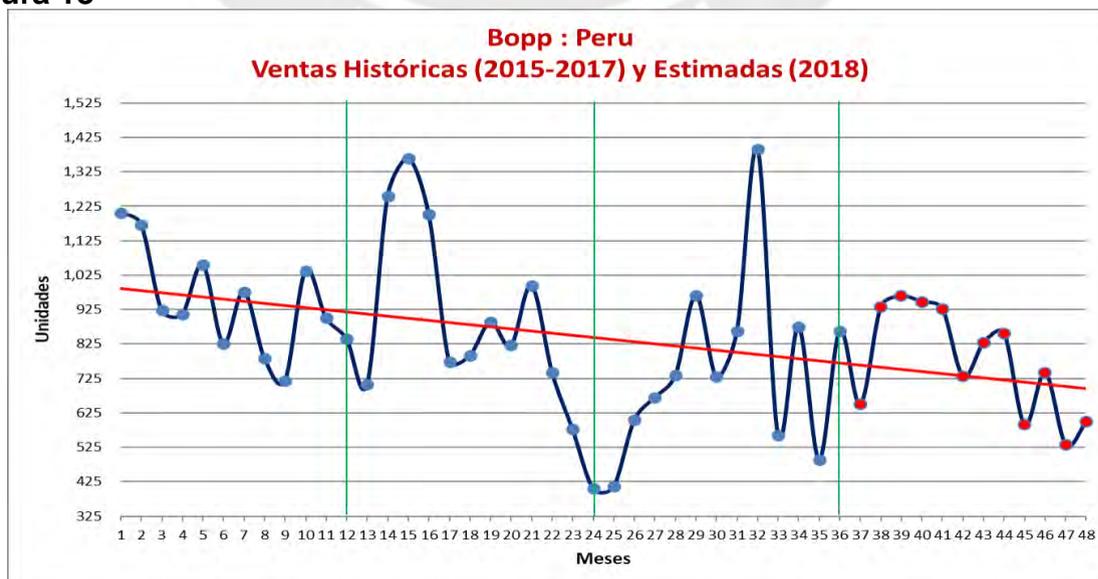
Figura 17

## Perú: serie, nivel y tendencia -2018

Año	Mes		Ventas históricas	Nivel estimado	Tendencia	Factor de estacionalidad normalizado	Ventas - pronóstico 2018
	36			<b>800.38</b>	<b>(4.1440)</b>		
2018	37	Enero		796.24	(4.1440)	0.81526	<b>649</b>
	38	Febrero		792.09	(4.1440)	1.17779	<b>933</b>
	39	Marzo		787.95	(4.1440)	1.22335	<b>964</b>
	40	Abril		783.81	(4.1440)	1.20854	<b>947</b>
	41	Mayo		779.66	(4.1440)	1.18862	<b>927</b>
	42	Junio		775.52	(4.1440)	0.94282	<b>731</b>
	43	Julio		771.37	(4.1440)	1.07362	<b>828</b>
	44	Agosto		767.23	(4.1440)	1.11568	<b>856</b>
	45	Septiembre		763.09	(4.1440)	0.77372	<b>590</b>
	46	Octubre		758.94	(4.1440)	0.97676	<b>741</b>
	47	Noviembre		754.80	(4.1440)	0.70636	<b>533</b>
	48	Diciembre		750.66	(4.1440)	0.79749	<b>599</b>
					<b>12.00001</b>	<b>9,299</b>	

Fuente: Elaboración Propia

Figura 18



Fuente: Elaboración Propia

En el caso de Argentina se observa lo siguiente:

**Tabla 1**  
*Argentina: Variación TM 2017-2018(p)*

Meses	2017	2018	% Var
Enero	653	580	-11.1%
Febrero	490	592	20.7%
Marzo	457	458	0.1%
Abril	483	418	-13.5%
Mayo	571	624	9.2%
Junio	562	605	7.5%
Julio	719	764	6.3%
Agosto	639	637	-0.4%
Septiembre	678	678	0.1%
Octubre	609	609	-0.1%
Noviembre	641	576	-10.1%
Diciembre	420	501	19.5%
<b>TOTAL</b>	<b>8940</b>	<b>9059</b>	<b>1.3%</b>

Fuente: Elaboración Propia

Según la Tabla 1, en el mes de enero se pronostica un total de 580 TM de ventas para el 2018, comparadas con los 653 TM de ventas del 2017 se refleja un caída del 11.1%, en abril se pronostica un total de 418 TM de ventas para el 2018, comparadas con los 483 TM de ventas del 2017 se refleja un caída del 13.5% y en noviembre se pronostica un total de 576 TM de ventas para el 2018, comparadas con los 641 TM de ventas del 2017 se refleja un caída del 10.1%, así como de un aumento de ventas en los demás meses siendo a destacar que en febrero el pronóstico es de un total de 592 TM de ventas para el 2018, comparadas con los 490 TM de ventas del 2017 se refleja un aumento del 20.7% y en diciembre el pronóstico es de un total de 501 TM de ventas para el 2018, comparadas con los 420 TM de ventas del 2017 se refleja un aumento del 19.5%.

En conclusión, se prevé un total de 9059 TM de ventas durante el 2018, frente a las 8940 TM del año anterior; se considera que al final del año el incremento de ventas será de un 1.3% equivalentes a 119 TM.

Además, de acuerdo con el Gráfico 4, se puede observar que el promedio de errores al cuadro es de 0.04; asimismo, los valores de los parámetros Alpha, Beta y Gamma, fueron obtenidos de la minimización del promedio de errores al cuadro para la estimación del 2017, siendo los valores Alpha:0.3192, Beta:0.0243 y Gamma:0.5180.

En el caso de Chile se observa lo siguiente:

**Tabla 2**

*Chile: Variación TM 2017-2018(p)*

Meses	2017	2018	% Var
Enero	459	420	-8.5%
Febrero	626	634	1.3%
Marzo	593	711	19.8%
Abril	818	565	-30.8%
Mayo	726	509	-30.0%
Junio	762	583	-23.5%
Julio	704	585	-17.0%
Agosto	621	629	1.3%
Septiembre	660	620	-6.1%
Octubre	683	563	-17.6%
Noviembre	509	662	29.9%
Diciembre	450	510	13.4%
<b>TOTAL</b>	<b>9628</b>	<b>9009</b>	<b>-6.4%</b>

Fuente: Elaboración Propia

Según la Tabla 2, en el mes de enero se pronostica un total de 420 TM de ventas para el 2018, comparadas con los 4593 TM de ventas del 2017 se refleja un caída del 8.5%, en abril se pronostica un total de 565 TM de ventas para el 2018, comparadas con los 818 TM de ventas del 2017 se refleja un caída del 30.8%, en mayo se pronostica un total de 509 TM de ventas para el 2018, comparadas con los 726 TM de ventas del 2017 se refleja un caída del 30%, en junio se pronostica un total de 583 TM de ventas para el 2018, comparadas con los 762 TM de ventas del 2017 se refleja un caída del 23.5%, en julio se pronostica un total de 585 TM de ventas para el 2018, comparadas con los 704 TM de ventas del 2017 se refleja un caída del 17% y en octubre se pronostica un total de 563 TM de ventas para el 2018, comparadas con los 683 TM de ventas del 2017 se refleja un caída del 17.6%, así como de un aumento de ventas en los demás meses siendo a destacar que en marzo el pronóstico es de un total de 711 TM de ventas para el 2018, comparadas con los 593 TM de ventas del 2017 se refleja un aumento del 19.8%, en noviembre el pronóstico es de un total de 662 TM de ventas para el 2018, comparadas con los 509 TM de ventas del 2017 se refleja un aumento del 29.9% y en diciembre el pronóstico es de un total de 510 TM de ventas para el 2018, comparadas con los 450 TM de ventas del 2017 se refleja un aumento del 13.4%.

En conclusión, se prevé un total de 9009 TM de ventas durante el 2018,

frente a las 9628 TM del año anterior; se considera que al final del año la caída de ventas será de un 6.4% equivalentes a 620 TM.

Además, de acuerdo con el Gráfico 10, se puede observar que el promedio de errores al cuadro es de 0.06; asimismo, los valores de los parámetros Alpha, Beta y Gamma, fueron obtenidos de la minimización del promedio de errores al cuadro para la estimación del 2017, siendo los valores Alpha:0.4741, Beta:0.0000 y Gamma:0.1428.

En el caso de Perú se observa lo siguiente:

**Tabla 3**

*Perú: Variación TM 2017-2018(p)*

Meses	2017	2018	% Var
Enero	412	649	57.6%
Febrero	604	933	54.5%
Marzo	670	964	43.9%
Abril	733	947	29.2%
Mayo	965	927	-4.0%
Junio	730	731	0.2%
Julio	863	828	-4.0%
Agosto	1391	856	-38.5%
Septiembre	561	590	5.3%
Octubre	875	741	-15.3%
Noviembre	489	533	9.0%
Diciembre	863	599	-30.6%
<b>TOTAL</b>	<b>11172</b>	<b>11317</b>	<b>1.3%</b>

Fuente: Elaboración Propia

Según la Tabla 3, en el mes de agosto se pronostica un total de 856 TM de ventas para el 2018, comparadas con los 1391 TM de ventas del 2017 se refleja un caída del 38.5%, en octubre se pronostica un total de 741 TM de ventas para el 2018, comparadas con los 875 TM de ventas del 2017 se refleja un caída del 15.3% y en diciembre se pronostica un total de 599 TM de ventas para el 2018, comparadas con los 863 TM de ventas del 2017 se refleja un caída del 30.6%, así como de un aumento de ventas en los demás meses siendo a destacar que en enero el pronóstico es de un total de 592 TM de ventas para el 2018, comparadas con los 412 TM de ventas del 2017 se refleja un aumento del 57.6%, en febrero el pronóstico es de un total de 933 TM de ventas para el 2018, comparadas con los 604 TM de ventas del 2017 se refleja un aumento del 54.5%, en marzo el pronóstico es de un total de 964 TM de ventas para el

2018, comparadas con los 670 TM de ventas del 2017 se refleja un aumento del 43.9% y en abril el pronóstico es de un total de 947 TM de ventas para el 2018, comparadas con los 733 TM de ventas del 2017 se refleja un aumento del 29.2%.

En conclusión, se prevé un total de 11317 TM de ventas durante el 2018, frente a las 11172 TM del año anterior; se considera que al final del año el incremento de ventas será de un 1.3% equivalentes a 145 TM.

Además, de acuerdo con el Gráfico 16, se puede observar que el promedio de errores al cuadro es de 0.16; asimismo, los valores de los parámetros Alpha, Beta y Gamma, fueron obtenidos de la minimización del promedio de errores al cuadro para la estimación del 2017, siendo los valores Alpha:0.3980, Beta:0.0000 y Gamma:0.5606.

## 5.2. Modelo de abandono de clientes

Según los gráficos 19, 20 y 21, se pueden observar los resultados de la regresión logística, tomando en consideración las siguientes características:

**Figura 19**

Cuadro estadístico

	TM	YEAR	MARKET	CustServ	Qual
<b>count</b>	362.000000	362.000000	362.000000	362.000000	362.000000
<b>mean</b>	557.051758	15.01105	15.01105	1.629834	1.629834
<b>std</b>	1839.645819	14.42873	14.42873	1.327733	1.327733
<b>min</b>	0.248800	2.00000	2.00000	0.000000	0.000000
<b>25%</b>	24.370250	9.00000	9.00000	1.000000	1.000000
<b>50%</b>	86.144890	11.00000	11.00000	1.000000	1.000000
<b>75%</b>	385.091738	15.00000	15.00000	3.000000	3.000000
<b>max</b>	24227.886760	161.00000	161.00000	7.000000	7.000000

Cliente	0
COD	0
TM	0
FLETE	0
LEAD TIME	0
DISCOUNT	0
Log Oper	0
YEAR MARKET	0
CustServ Qual	0
Churn	0
dtype:	int64

Fuente: Elaboración Propia

Figura 20

Cuadro resultado Churn total

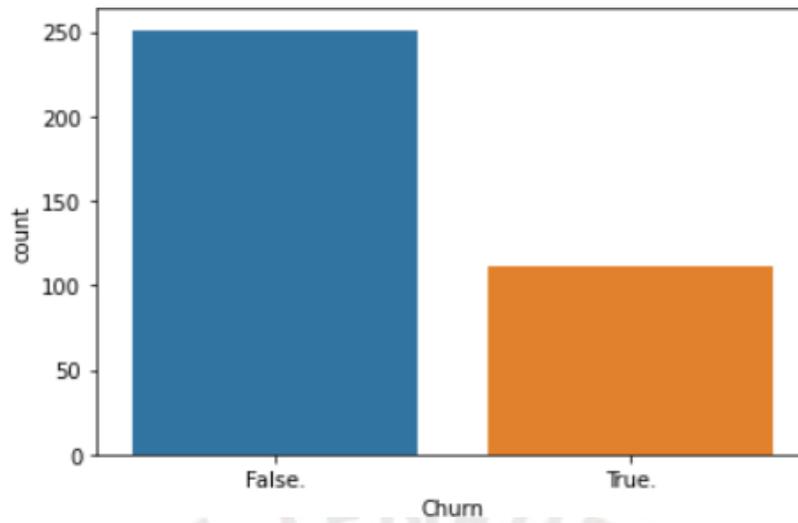
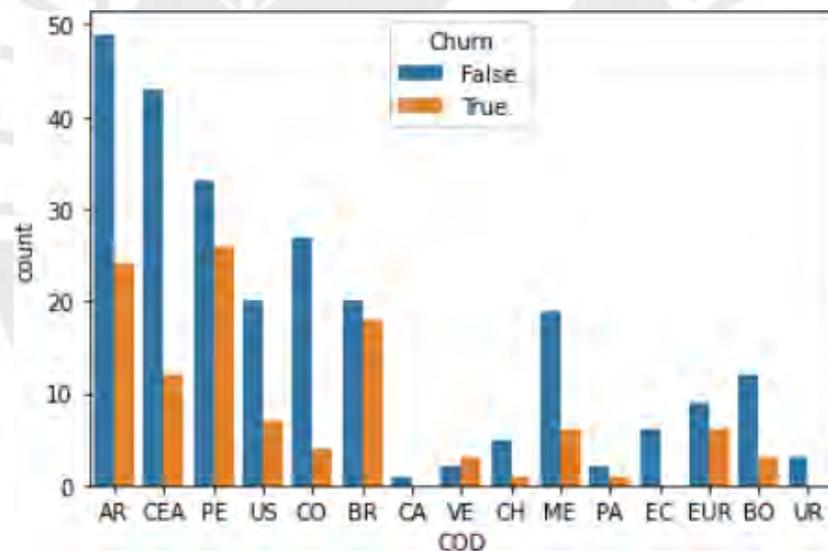


Figura 21

Cuadro resultado Churn por países



Fuente: Elaboración Propia

Figura 22

Matriz de confusión y métricas asociadas

	precision	recall	f1-score	support
0	0.75	0.87	0.80	54
1	0.30	0.16	0.21	19
accuracy			0.68	73
macro avg	0.52	0.51	0.51	73
weighted avg	0.63	0.68	0.65	73

Fuente: Elaboración Propia

En el caso de la precisión:

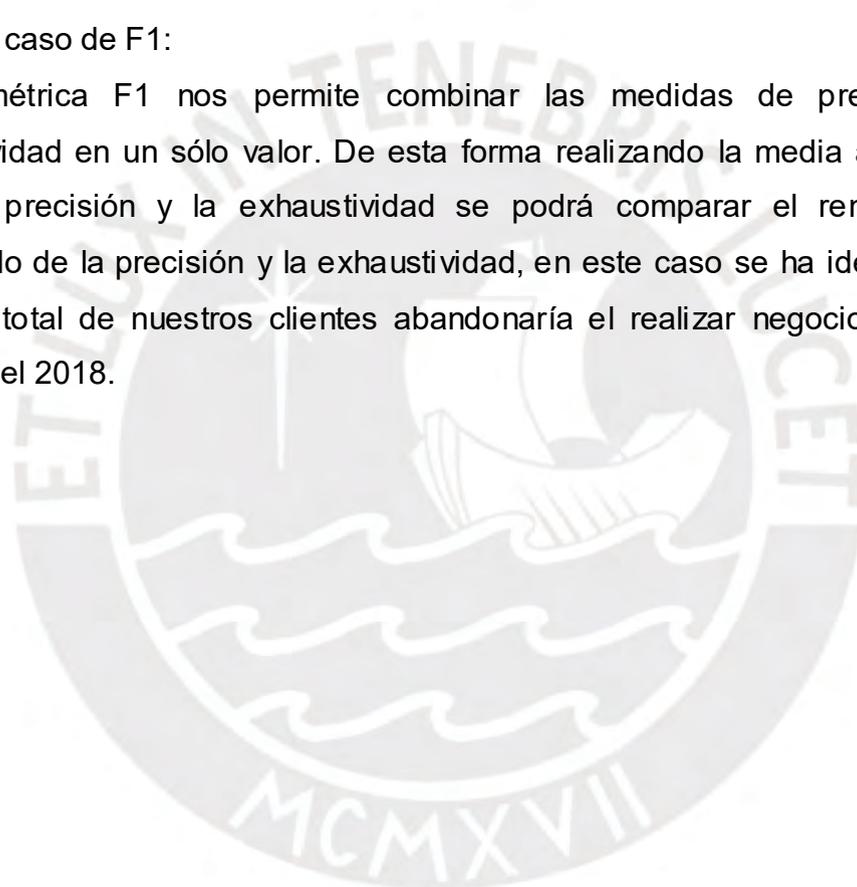
Con la métrica de precisión se mide la calidad del modelo de machine learning, en este caso sólo un 30% de los clientes abandonarían en el año 2018 realizar negocio con la empresa.

En el caso de la Exhaustividad:

La métrica de exhaustividad nos indica la cantidad que el modelo de machine learning es capaz de identificar, en este caso el modelo sólo es capaz de identificar que de los clientes que abandonarían el realizar negocios con la empresa el 2018. se ha identificado un 16%.

En el caso de F1:

La métrica F1 nos permite combinar las medidas de precisión y exhaustividad en un sólo valor. De esta forma realizando la media armónica entre la precisión y la exhaustividad se podrá comparar el rendimiento combinado de la precisión y la exhaustividad, en este caso se ha identificado 21% del total de nuestros clientes abandonarían el realizar negocios con la empresa el 2018.



## 6. Conclusiones

En resumen, de lo expuesto en el presente trabajo se puede realizar las siguientes conclusiones:

Con respecto al modelo Holt-Winters, del análisis de los resultados se puede observar que los pronósticos de ventas anuales para el 2018 tanto para Argentina, Chile y Perú, se ajustan a la realidad de las ventas en cada mercado dado su nivel, tendencia y estacionalidad, además de considerar idónea la parametrización minimizando el promedio de errores al cuadrado de los pronósticos correspondientes.

Con respecto al modelo de abandono de clientes, del análisis de los resultados se concluye que el pronóstico de abandono para el 2018 resultó ser de un 30%, tomando en cuenta las 5 variables utilizadas; dicho esto, es pertinente concluir que el pronóstico puede ser más acertado si al modelo se le agregan variables explicativas pertinentes; sin embargo, se debe resaltar que se trabajó con toda la información que la empresa brindó para el análisis.

Por ello, considero que desde mi posición como analista pricing en OPP film, he podido aplicar mis conocimientos adquiridos en econometría para el despliegue de los modelos anteriormente explicados utilizando los datos cuantitativos y cualitativos para el beneficio de la empresa.

Finalmente, considero que, al haber aplicado mis conocimientos y conceptos de la carrera de economía, concluyo que el presente trabajo de suficiencia profesional cuenta con los criterios para optar por el título de Licenciada en Economía.

## **Bibliografía**

Angel, Nelson & Márquez, Diana (2018). Modelo de gestión para la proyección de demanda de productos perecederos utilizando combinación de pronósticos por series de tiempo. Universidad de Sergio Arboleda.

Amstrong, J., Kumar V. & Morwitz, V. (2000). Sales Forecasts for Existing Consumer Products and Services: Do Purchase Intentions Contribute to Accuracy? *International Journal of Forecasting* (16), 383-397.

Castelló Emma (2016). Impacto del revenue management en los sistemas de control de gestión. *Revista Facultad de Ciencias Económicas: Investigación y Reflexión* (24), 85-101.

Castro, J. & Pérez, E. (2020). Evaluación del abandono de clientes de una compañía de telecomunicaciones por medio de cuatro modelos de aprendizaje máquina. *Research in Computing Science* (149), 611-624.

Cheraghi, S. H., Dadashzadeh, M. & Venkitachalam, P. (2010). Revenue Management in Manufacturing: A Research Landscape. *Journal of Business & Economics Research* (8), 63-72.

Devrient, F., Verrevoets, J. & Verbeke, W. (2021). Why you should stop predicting customer churn and start using uplift models. *Information Sciences*. (548), 495-515.

Falla, Jesús (2021). Predicción de abandono de clientes en telecomunicaciones mediante el aprendizaje automático. Universidad de Bogota Jorge Tadeo Lozano.

Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior*. Reading, MA: Addison-Wesley.

Gardner, J. (2006). Exponential smoothing: The state of the art-part II. *International Journal of Forecasting* (22), 637-666.

Guliyev, Hasraddin; Yervelen, Ferda. (2021). Customer churn analysis in banking sector: Evidence from explainable machine learning models. *Journal of Applied Microeconometrics*. (1), 85-99.

Hyndman, R.J., Koehler, A.B., Snyder, R.D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, (18), 439-454.

Jain, H., khunteta, A. & Srivastava, S. (2020). Churn Prediction in Telecommunication using Logistic Regression and Logit Boost. *Procedia Computer Science*, (167), 101-112.

Jélvez Caamaño, Arnaldo, Moreno Echeverría, Mauricio, Ovalle Retamal, Víctor, Torres Navarro, Carlos, & Troncoso Espinosa, Fredy (2014). Modelo predictivo de fuga de clientes utilizando minería de datos para una empresa de telecomunicaciones en Chile. *Universidad, Ciencia y Tecnología*, (18), 100-109.

Nojek S., Britos P., Rossi, B. & Garcia R. (2003). Pronóstico de Ventas: Comparación de Predicción basada en Redes Neuronales versus Método Estadístico. Departamento de Ingeniería Industrial. Instituto Tecnológico de Buenos Aires, (2), 1-18.

Smith, B.C.; Leimkuhler, J.M. & Darrow, R.M. (1992). Yield management at American Airlines. Interfaces, (22), 8-31.

Spiliotis, E.; Assimakopoulos, V. (2021). Product sales probabilistic forecasting using the M5 competition data. International Journal of Production Economics, (240), 1-37.

