



Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Estudios de Postgrado

Maestría en Estadística Aplicada

**MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA
DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y
ESTRATÉGICAS**

Ing. Sergio Eduardo Segura Cordero

Asesorado por el Mtro. Edwin Adalberto Bracamonte Orozco

Guatemala, septiembre de 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA
DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y
ESTRATÉGICAS**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

ING. SERGIO EDUARDO SEGURA CORDERO
ASESORADO POR EL MTRO. EDWIN BRACAMONTE

AL CONFERÍRSELE EL TÍTULO DE

MAESTRO EN ESTADÍSTICA APLICADA

GUATEMALA, SEPTIEMBRE DE 202

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Vladimir Cruz Lorente
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANA	Inga. Aurelia Anabela Cordova Estrada
DIRECTOR	Mtro. Ing. Edgar Darío Alvarez Cotí
EXAMINADOR	Mtro. Ing. Edwin Adalberto Bracamonte Orozco
EXAMINADOR	Mtro. Ing. William Eduardo Fagiani Cruz
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA
DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y
ESTRATÉGICAS**

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Postgrado, con fecha 28 de enero de 2020.

A handwritten signature in black ink, appearing to be 'S. Segura', with a stylized, cursive script.

Ing. Sergio Eduardo Segura Cordero

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Estudios de Posgrado, al Trabajo de Graduación titulado: **MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y ESTRATÉGICAS**, presentado por: **Sergio Eduardo Segura Cordero**, que pertenece al programa de Maestría en artes en Estadística aplicada después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:



Inga. Aurelia Anabela Cordova Estrada

Decana

Guatemala, septiembre de 2022



Guatemala, septiembre de 2022

LNG.EEP.OI.638.2022

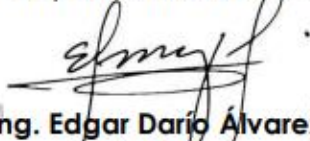
En mi calidad de Director de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor, verificar la aprobación del Coordinador de Maestría y la aprobación del Área de Lingüística al trabajo de graduación titulado:

"MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y ESTRATÉGICAS"

presentado por **Sergio Eduardo Segura Cordero** correspondiente al programa de **Maestría en artes en Estadística aplicada**; apruebo y autorizo el mismo.

Atentamente,

"Id y Enseñad a Todos"


Mtro. Ing. Edgar Darío Álvarez Cotí
Director
Escuela de Estudios de Postgrado
Facultad de Ingeniería



Ref. Coordinación de Estadística Aplicada

Guatemala, 27 de enero de 2020

Maestro
Edgar Dario Álvarez Coti
Director de la Escuela de Estudios de Postgrados
Presente.

Estimado Mtro. Álvarez:

Atentamente le informo que he revisado y aprobado el Diseño de Investigación titulado: "MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y ESTRATÉGICAS", del estudiante Sergio Eduardo Segura Cordero, quien se identifica con carné 201021118, alumno del programa de maestría en Artes en "Estadística Aplicada".

Agradeciendo su atención y deseándoles éxitos en sus actividades me suscribo.

Cordialmente,

"Id y Enseñad a Todos"


Mtro. Edwin Adalberto Bracamonte
Coordinador de Maestría
Estadística Aplicada
Escuela de Estudios de Postgrado
Facultad de Ingeniería



Guatemala 15 de agosto 2021.

M.A. Edgar Darío Álvarez Cotí
Director
Escuela de Estudios de Postgrado
Presente

M.A. Ingeniero Álvarez Cotí:

Por este medio informo que en mi calidad de asesor, he revisado y aprobado el Informe Final del trabajo de graduación titulado **"MODELO DE SERIE TEMPORAL PARA EL PRONÓSTICO DE VENTAS EN UNA CADENA DE RESTAURANTES PARA LA CORRECTA TOMA DE DECISIONES OPERATIVAS Y ESTRATÉGICAS"** del estudiante **Sergio Eduardo Segura Cordero** quien se identifica con número de carné **999001054** del programa de Maestría en Estadística Aplicada.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.

Atentamente,


MSc. Ing. Edwin Adalberto Bracamonte Orozco
Colegiado 2856
Asesor
Edwin Adalberto Bracamonte Orozco
INGENIERO INDUSTRIAL
Colegiado No. 2856

ACTO QUE DEDICO A:

Dios	Por su bendita misericordia y por permitir este acto.
Mis padres	Elías Segura y Nidia Cordero de Segura. Por su amor incondicional
Hermanos	Oscar, Luis, Mónica y Fabiola Segura, por su apoyo en momentos difíciles.
Hermanito	Andrés Segura, por ser mi fuente de inspiración

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Por ser mi <i>alma mater</i> y un lugar de constante aprendizaje.
Facultad de Ingeniería	Por el alimento intelectual.
Mis amigos de la Facultad	A toda la cohorte de la maestría.
Dra. Mayra Castillo	Por el apoyo incondicional y por ser una excelente catedrática.
MSc. Edwin Bracamonte	Por su guía y apoyo desde siempre, una excelente persona.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
LISTA DE SÍMBOLOS	VII
GLOSARIO	IX
RESUMEN	XI
PLANTEAMIENTO DEL PROBLEMA.....	XIII
OBJETIVOS.....	XVII
RESUMEN DEL MARCO METODOLÓGICO	XIX
INTRODUCCIÓN	XXV
1. MARCO REFERENCIAL	1
1.1 Estudios previos sobre pronósticos de ventas.....	1
2. MARCO TEÓRICO.....	9
2.1 Fundamentos estadísticos.....	9
2.1.1 Series de tiempo.....	9
2.1.2 Componentes de una serie de tiempo.....	10
2.1.2.1 Tendencia.	10
2.1.2.2 Estacionalidad.....	11
2.1.2.3 Componente cíclica	11
2.1.2.4 Componente aleatoria o irregular.....	11
2.1.3 Series de tiempo estacionarias y no estacionarias .	12
2.1.4 Coeficientes de autocorrelación y correlogramas ...	12
2.1.4.1 Coeficiente de autocorrelación	13
2.1.4.2 Correlograma.....	13
2.1.5. Prueba de Dickey-Fuller	14

2.1.6	Selección del método de pronóstico.....	15
2.1.6.1	Técnicas de pronóstico.....	16
2.1.6.1.1	Promedio móvil simple	16
2.1.6.1.2	Promedio móvil doble.....	17
2.1.6.1.3	Suavizamiento exponencial	17
2.1.6.1.4	Exponencial Holt.....	18
2.1.6.1.5	Métodos de Winters.....	19
2.1.6.2	Regresiones lineales	19
2.1.6.2.1	Objetivo	20
2.1.6.2.2	Notación.....	20
2.1.6.2.3	Variable.....	20
2.1.7.	Índices estacionales	21
2.1.7.1	Promedio móvil para medir la variación estacional.....	21
2.1.7.2	Pasos para calcular el índice estacional.....	21
2.1.8.	Error cuadrático medio	22
2.1.9.	Modelo de Holt-Winter estacional aditivo	22
2.2	Toma por medio de datos de ventas.....	23
2.2.1.	Ventas	23
2.2.2	Importancia de un pronóstico de ventas.....	24
2.2.3.	Alcance del pronóstico de ventas.....	25
3.	PRESENTACIÓN DE RESULTADOS.....	27
3.1.	Modelo de mejor ajuste a la tendencia de ventas	27
3.2.	Variables categóricas en el comportamiento de ventas.	43
3.3.	Ciclos más periódicos de ventas.	44

3.4.	Determinar el índice estacional óptimo por medio del análisis de series de tiempo, para proyectar las ventas a 4 años o menos.....	44
4.	DISCUSIÓN DE RESULTADOS	47
4.1	Modelo lineal de mejor ajuste a la tendencia de ventas	47
4.2	Variables categóricas que afectan el comportamiento de ventas.....	51
4.3	Ciclos con más periodicidad de ventas en el año.....	52
4.4	Índices estacionales de la serie de tiempo	52
	CONCLUSIONES	55
	RECOMENDACIONES.....	57
	REFERENCIAS	59

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Serie de tiempo, ventas de restaurantes 2015-2018	27
2.	Descomposición de la serie de tiempo	28
3.	Función de autocorrelación para la serie de tiempo de ventas.....	29
4.	Gráfica del pronóstico promedio móvil doble y venta real	32
5.	Gráfica pronóstico promedio móvil simple	33
6.	Gráfica de pronóstico suavizamiento exponencial doble	35
7.	Gráfica pronóstico método Holt	36
8.	Gráfica de pronóstico método Holt-Winter	38
9.	Gráfica del pronóstico con regresión lineal	39
10.	Gráfica de pronóstico de ventas Holt-Winter aditivo	43

TABLAS

I.	Identificación de variables	XX
II.	Criterios de series temporales.....	XXII
III.	Técnicas de previsión según criterios de estacionalidad y tendencia	15
IV.	Volumen de ventas de restaurantes 2015-2018.....	28
V.	Prueba Dickey-Fuller.....	30
VI.	Resultados del pronóstico con promedio móvil doble	31
VII.	Resultados del pronóstico con promedio móvil simple.....	32
VIII.	Resultados del método suavizamiento exponencial doble	34

IX.	Resultados del pronóstico con método Holt.....	35
X.	Resultados del pronóstico con método Holt-Winter	36
XI.	Resultados del pronóstico con una regresión lineal.....	38
XII.	Pronósticos de las ventas en la cadena de restaurantes.....	39
XIII.	Parámetros utilizados completamente al azar	40
XIV.	Indicadores de la regresión lineal de los datos	41
XV.	Coeficientes de intercepto y pendiente de la regresión lineal	41
XVI.	Longitud del ciclo y pendiente de la serie de datos	41
XVII.	Pronóstico Holt-Winter aditivo año 2019 y 2020	42
XVIII.	Pronósticos 2019 y 2021	43
XIX.	Índices estacionales de la serie de tiempo	45
XX.	Índices estacionales corregidos.....	46

LISTA DE SÍMBOLOS

Símbolo	Significado
R^2	Coeficiente de determinación
$L1$	Longitud de la serie de tiempo
$T1$	Pendiente de la serie de tiempo
K	Periodo de tiempo de la serie de tiempo

GLOSARIO

ACF	Autocorrelación
AIC	Criterio de información de Akaike
BIC	Criterio de información Bayesiano
EMC	Error cuadrático medio
ETS	Algoritmo para suavizamiento exponencial
Aprovisionamiento	Es la relación que consiste en recibir mercancías y optimizar los procesos de recepción.
Componente	Se entiende por aquello que forma parte de la composición de un conjunto de datos.
Periodicidad	Se refiere a la frecuencia con la que aparece, sucede o se realiza una cosa repetitiva.
Productividad	Es la relación entre la cantidad de productos obtenida por un sistema productivo y los recursos utilizados para obtener dicha producción.
Pronóstico	Es el proceso de estimación de un valor en situaciones de incertidumbre.

Solve

Herramienta de Microsoft Excel

Tendencia

Se entiende como la inclinación positiva o negativa de las posibles ventas futuras en una organización.

RESUMEN

El siguiente estudio surgió de la necesidad de tener una estimación de las ventas en los restaurantes de comida rápida, para predecir cuáles serán los recursos adecuados para abastecer a la demanda de clientes. Actualmente la cadena de restaurantes utiliza únicamente su información histórica de ventas para tomar decisiones empíricas tomando como referencias las ventas del año anterior.

Por lo cual se identificó una necesidad establecer un modelo de pronóstico estadístico, por medio de series de tiempo y análisis de regresión, que permita predecir los niveles de venta en los restaurantes de la ciudad de Guatemala, con el objetivo de tomar decisiones estratégicas más acertadas.

Para el desarrollo de la investigación se empleó el uso de técnicas básicas para poner a prueba varios modelos, de los cuales se seleccionó el de mejor ajuste para pronosticar los niveles de venta a 4 años o menos.

Estadísticamente hablando y con un nivel de confiabilidad del 95 % se observó que los niveles de venta serán 35 % más que el año anterior, por lo cual se recomienda considerar tener el abastecimiento adecuado para lograr atender a la mayor demanda posible.

PLANTEAMIENTO DEL PROBLEMA

- Contexto general

En una cadena de restaurantes de comida rápida especializada en pollo frito, con más de 40 años de trayectoria en la ciudad de Guatemala, se ha observado la necesidad de implementar nuevos métodos que permitan anticiparse a la demanda de ventas. Por lo que se necesita diseñar un modelo de predicción que ayude a tomar correctas decisiones operativas y estratégicas; tomando como base información histórica y variables externas que afecten los niveles de venta.

- Descripción del problema

En esta época de alta competencia, revolución tecnológica y transformación digital acelerada, campos como la inteligencia de negocios enfocada en modelos estadísticos, surge con un alto potencial para que las empresas puedan obtener bases concretas para la toma de decisiones operativas y estratégicas. Partiendo de lo anterior, se desea explotar el uso de la estadística en la empresa para establecer un modelo de predicción que permita conocer con alto grado de confiabilidad los niveles de venta en un periodo determinado; esto permitió realizar una estimación sobre cuánto será la inversión económica necesaria para las siguientes actividades: producción, abastecimiento de insumos, mano de obra, costos directos e indirectos, y pago de servicios. Con esto se desea lograr que la compañía sea parte de esta revolución tecnológica para que pueda aumentar sus niveles de eficiencia y productividad.

Dentro de los aspectos requeridos para la elaboración del modelo de predicción se encuentran, información histórica de ventas, realizar un análisis de serie de tiempo para identificar las componentes de estacionalidad, tendencia, ciclicidad, y aleatoriedad; aplicar métodos de suavización a la serie de tiempo, y por último aplicar una regresión lineal para encontrar una recta de mejor ajuste en los datos.

- Formulación del problema

¿Cuál es el modelo estadístico de mejor ajuste que permite pronosticar con mayor exactitud las ventas realizadas por los restaurantes de la ciudad de Guatemala, en un periodo de 5 años?

- Preguntas auxiliares
 - ¿Cuál es el modelo lineal que mejor se ajusta a la tendencia de las ventas históricas de 4 años de antigüedad?
 - ¿Cuáles son las variables categóricas que afectan los comportamientos de venta en los restaurantes de la ciudad de Guatemala?
 - ¿Cuáles son los periodos de año en donde las ventas son más significativas en los restaurantes de la ciudad de Guatemala?
 - ¿Cuáles son los índices estacionales óptimos para realizar la mejor proyección de ventas a 4 años o menos?

- Delimitación del problema

Se realizó un análisis predictivo para las ventas que generan los 70 restaurantes que operan actualmente en la ciudad de Guatemala. Para este análisis se utilizó la base de datos proporcionada, la cual contiene información histórica de cuatro (4) años de antigüedad. Se incluyeron los siguientes canales de venta que utilizan los restaurantes: despacho en mesas, despachos para llevar, autoservicio y entrega a domicilio.

OBJETIVOS

General

Establecer un modelo de pronóstico estadístico, por medio de series de tiempo y análisis de regresión, que permita predecir los niveles de venta en los restaurantes de la ciudad de Guatemala.

Específicos

1. Establecer el modelo lineal de mejor ajuste a la tendencia de ventas, por medio de un análisis de regresión lineal, que permita explicar el comportamiento de ventas que realizan los restaurantes en estudio.
2. Determinar las variables categóricas que afecten el comportamiento de ventas por medio de una técnica de observación, para incluirlas en el análisis de dependencia en el modelo predictivo.
3. Determinar los ciclos más periódicos de ventas, por medio de un análisis de descomposición de series de tiempo, para conocer en que estaciones del año las ventas son más significativas.
4. Determinar el índice estacional óptimo por medio del análisis de series de tiempo, para proyectar las ventas a 5 años o menos.

RESUMEN DEL MARCO METODOLÓGICO

- **Características del estudio**

El enfoque del estudio propuesto es cuantitativo, porque realizó un pronóstico de los niveles de venta de un periodo en específico, es decir que, si el modelo estadístico pronostica un crecimiento porcentual en las ventas, entonces el área de producción debe tener la capacidad de cumplir con esta demanda. Este estudio contribuyó con la obtención de cuota de mercado introduciendo nuevos productos o retirar otros que causen una crisis.

El alcance es correlacional dado que se asocian las variables de ventas y tiempo mediante patrones identificados en una serie temporal, lo que permite comprender tendencias, estacionalidades, ciclos; y con base en estos análisis poder predecir con cierto grado de confiabilidad las ventas en un periodo en específico.

El diseño adoptado es no experimental, pues la información histórica utilizada para la elaboración del modelo estadístico que pronostique las ventas se analizó en su estado original sin ninguna manipulación; además es transversal pues se estudió la tendencia de ventas en la cadena restaurantes que operan en la ciudad de Guatemala.

- **Unidades de análisis**

La población en estudio son los restaurantes que operaron en el año 2019 en la ciudad de Guatemala; estos restaurantes se encuentran divididos en subpoblaciones conocidos como canales de distribución, es decir, autoservicio,

a domicilio, despacho en la caja y despacho en mesas; de estas subpoblaciones se extendió muestras de forma estratificada, que serán estudiadas en su totalidad.

- Variables

Las variables en estudio se describen a continuación:

Tabla I. Identificación de variables

Variable	Definición teórica	Definición operativa
Ventas	Es la acción entre una empresa y un consumidor, que se genera al vender un bien o un servicio a cambio de dinero.	Se utilizó la venta mensual neta en quetzales. La cual se encuentra registrada en la base de datos de la cadena de restaurantes que opera en la ciudad de Guatemala.
Periodo	Es un tiempo determinado en el cual se realiza una acción o se desarrolla un acontecimiento.	Se estudió los periodos del año donde existan días festivos y vacaciones.

Fuente: elaboración propia.

- Fases del estudio

Fase 1. Revisión de información, la cual consiste en consultar fuentes como tesis de postgrados, artículos, revistas y material que trate sobre series temporales, pruebas de tendencia, pruebas de estacionalidad y regresiones lineales. Con base en la revisión se extrajeron ideas, conceptos y técnicas que permitieron reforzar el estudio acerca del modelo estadístico predictivo.

Fase 2. Gestión de datos, se solicitó la base de datos de la cadena de restaurantes en la cual se encuentra registrado el histórico de ventas que comprende el periodo de 2014 al 2019. Esta base de datos es generada desde la base de datos transaccional de ventas, en la cual se registra la información regional de la cadena de restaurantes.

Fase 3. Análisis de información, se depuró la información del histórico de ventas para trabajar únicamente con restaurantes que operan en la ciudad de Guatemala. Las ventas se encuentran registradas mensualmente en quetzales; no se realizaron conversiones de tasa de cambio durante el estudio.

La herramienta que se utilizó para el análisis de la información fue el software estadístico R-Studio, debido a su precisión para realizar cálculos, y porque utiliza código abierto y paquetes que permiten realizar análisis estadísticos avanzados. Otra característica importante de este software es que es bastante ágil para el tratamiento de grandes volúmenes de datos y para la automatización de cálculos y gráficas. R-Studio cuenta con varias librerías que permiten descomponer y graficar las componentes de una serie temporal, así como aplicar pruebas de estacionalidad y funciones de autocorrelación.

Fase 4. Interpretación de información, a partir del análisis de la información se construyeron gráficas y tablas para mostrar los resultados. Esto permitió modelar la información a manera que pueda ser de fácil interpretación. Las gráficas mostraron el comportamiento de la serie temporal, su tendencia y estacionalidad. Mientras que las tablas fueron utilizadas para visualizar los valores que definieron cuál es la recta de mejor ajuste del modelo predictivo.

- Técnicas de análisis de información

Serie de tiempo: se graficaron los datos de las ventas mensuales para obtener un panorama de las componentes de la serie de tiempo. Cabe mencionar que este primer análisis es subjetivo y que es necesario utilizar herramientas sofisticadas que permitan decidir si la serie presenta componentes de tendencia, estacionalidad, ciclos y aleatoriedad de los datos.

Correlograma: se aplicó la función de autocorrelación (Ec.2) con su respectiva gráfica. En el correlograma se observó si existe correlación en los valores de k, lo cual permitió verificar si existe tendencia en los datos. Otra característica de una serie con tendencia es cuando los valores de los coeficientes de autocorrelación disminuyen conforme aumenta el valor k; esto también se consideró en el análisis.

Prueba de Dickey-Fuller: se aplicó esta prueba para determinar si la serie temporal es estacionaria o no. Para determinar esto se obtuvo el p-Valor, el cual fue menor al nivel de significancia, por lo que se rechazó la hipótesis nula (H_0 : No es estacionario).

Métodos de previsión: se utilizaron los siguientes 7 métodos para determinar cuál de ellos presenta el menor el error cuadrático medio.

- Promedio móvil simple
- Promedio móvil doble
- Suavización exponencial
- Holt
- Holt-Winters
- Regresión lineal

Los criterios para aplicar los métodos anteriores fueron los siguiente:

Tabla II. Criterios de series temporales

	Sin estacionalidad	Con estacionalidad
Sin tendencia	Promedio móvil simple y Suavización exponencial simple	Estacional aditivo y estacional multiplicativo
Con tendencia	Promedio móvil doble y exponencial doble	Holt-Winter Aditivo y Multiplicativo

Fuente: elaboración propia.

El modelo que se obtuvo de este análisis fue el que modeló de mejor manera los datos de la serie de tiempo y el que se utilizó para pronosticar las ventas mensuales de la cadena restaurantes en un periodo no mayor de 4 años.

INTRODUCCIÓN

La transformación digital que se vive en la actualidad está dando oportunidad al uso de modelos estadísticos en áreas como la inteligencia de negocios para elaborar modelos predictivos. El siguiente trabajo consiste en la elaboración de un modelo estadístico para pronosticar las ventas en una cadena de restaurantes, utilizando series temporales y regresiones lineales. En la actualidad la empresa toma decisiones de manera empírica y subjetiva para la planificación de las ventas.

Para la construcción de este modelo se utilizará la información histórica de ventas de 4 años atrás, se delimitará el análisis a restaurantes solo de la ciudad de Guatemala, y se emplearán métodos estadísticos tales como: series de tiempo para identificar patrones y tendencias; regresiones lineales para encontrar una recta de mejor ajuste, pruebas de hipótesis para determinar si existe aleatoriedad, estacionalidad y tendencia en los datos; todo esto con el objetivo de elaborar un modelo de predicción que permita conocer la proyección de ventas en el futuro.

La línea de investigación sobre la cual se realizó este trabajo de investigación es pronóstica, debido a que en la actualidad los pronósticos proporcionan beneficios como anticipación ante fenómenos no planificados o irregulares que puedan afectar de manera negativa a la operatividad de la cadena de restaurantes.

Este informe consta de cuatro capítulos, el primero es el marco referencial, el cual abarca fundamentos estadísticos donde se explican las herramientas

estadísticas que se emplearán para el estudio, tales como los componentes de una serie de tiempo, correlación, correlograma y pruebas de normalidad.

El segundo es el marco teórico donde se coloca una definición puntual acerca de lo que son las ventas y la importancia de los pronósticos en la toma de decisiones por parte de las gerencias.

En el tercer capítulo se encuentra la presentación de resultados, en el cual se da a conocer el resultado de la investigación y el modelo que mejor se ajusta al pronóstico. Por último, en el cuarto capítulo, se presenta la discusión de resultados, donde se expone los beneficios de la investigación y los aportes que contribuyen a la rentabilidad de la operación.

1. MARCO REFERENCIAL

Durante el desarrollo de este trabajo de investigación fueron consultadas fuentes bibliográficas dedicadas al pronóstico de ventas en sectores que no necesariamente son semejantes a los del tema principal de este trabajo. Se observaron fuentes que tratan sobre la implementación de series de tiempo y análisis de regresión para el desarrollo de métodos predictivos.

1.1. Estudios previos sobre pronósticos de ventas

El avance que ha tenido la revolución tecnológica en las compañías ha impulsado a que ciencias como la Estadística, tengan una participación notable en el área de inteligencia de negocio; esto se debe a que esta ciencia ofrece modelos para pronosticar o predecir con cierto grado de confiabilidad, comportamientos y tendencias en operaciones como inventarios, ventas, compras, presupuestos, entre otros.

Frías (2006) menciona que la estadística permite trabajar con variables aleatorias o estocásticas, es decir, variables que tienen distribuciones probabilísticas; con base en esto es posible obtener modelos que se puedan usar para calcular probabilidades de ventas futuras, las cuales estarán entre límites específicos de confiabilidad.

Este autor también indica que variables como los niveles de precio, descuentos y promociones entre otros; son factores que influyen en la venta mensual de productos o servicios, consideradas como variables explicativas.

Estas variables hacen que la predicción de ventas no sea totalmente exacta, debido a los errores que se presentan en la medición de estas.

Los pronósticos dependen fuertemente de las series temporales. De acuerdo con el estudio realizado por Gonzáles (2018) una serie temporal es capaz de determinar cómo se comporta el movimiento de un producto o servicio, por medio de patrones observables en un periodo de tiempo determinado. A partir de lo mencionado “Se puede plantear un modelo de pronóstico que se ajuste a la serie y pueda predecir con cierto grado de exactitud, el comportamiento futuro” (p.31).

Un aspecto importante, que hay que tomar en cuenta antes de realizar un pronóstico, es la información disponible. Frausto (2009) menciona que la calidad de un modelo de pronósticos está limitada por la cantidad y calidad de la información disponible. “Todo pronóstico depende de la información con la que se crea, sea en forma explícita o implícita” (p.28). Asegurar la calidad de la información histórica de ventas contribuyó con el diseño del modelo predictivo. La calidad de la información de ventas se podrá validar por medio de un análisis de compatibilidad, de datos faltantes y verificando que la información sea completa y precisa.

Cerreño (2012) aplicó una regresión lineal para ajustar datos de venta en una empresa distribuidora de maquinaria pesada, en la cual determinó una variabilidad significativa utilizando los últimos doce meses del histórico de ventas y obteniendo un coeficiente de correlación de Pearson menor a 0.7. Este parámetro permitió tener una referencia acerca de que tan confiable es el modelo estadístico para el pronóstico de ventas.

De acuerdo con el trabajo realizado por Roque (2016), en el cual menciona que, dentro del análisis descriptivo de una serie de temporal, se encuentra identificar el comportamiento estacional en los datos de ventas. “Para lo cual como primer paso será graficarla, esto permitirá identificar su tendencia, estacionalidad y variaciones irregulares” (p.93). Este aporte permite conocer cuál es el primer paso que se debe realizar para analizar una serie temporal y las componentes que se utilizan para diseñar el modelo de pronósticos.

Este autor recomienda que para utilizar algoritmos de series de tiempo como el de Holt-Winters, Holt y ETS se debe utilizar como mínimo 25 meses de información histórica. En su trabajo determinó que el algoritmo ETS es muy bueno para predecir valores cercanos al presente. Esta recomendación permite conocer algunos algoritmos que son aplicables al análisis de series de temporales, tomando en cuenta siempre la cantidad de información histórica recomendada para utilizarlos.

De acuerdo con el estudio realizado por González (2018) las componentes de una serie de tiempo que hay que considerar son tendencia, la cual indica un comportamiento creciente o decreciente en los datos; también se encuentra la estacionalidad, la cual indica “cambios regulares que se dan en la serie cada determinado periodo de tiempo” (p.14). Las definiciones de estas componentes permitieron conocer el comportamiento de los datos y cómo deben ser tratados cuando se realice un modelo predictivo.

Como tercer componente se encuentra la cíclica, esta indica variaciones que se presentan en periodos mayores a un año, las cuales pueden ser producidas en muchas ocasiones por periodos de abundancia o escasez económica. Esta componente aportó información acerca de los niveles de venta en periodos de días festivos.

Y, por último, la componente aleatoria o irregular, indica una variabilidad con aleatoriedad generada por factores que no estaban contemplados y que no son frecuentes. Esta componente permitió conocer si existen variables cualitativas que afecten de manera significativa el comportamiento de los datos. Con esto se logró comprender cuales son las componentes principales de un análisis de series de tiempo y ver cómo estas influyen en el comportamiento de los datos.

De acuerdo con la investigación realizada por Gil (2016), en la cual determinó que el mejor modelo que se ajusta al pronóstico de la demanda de energía eléctrica en Colombia es el modelo cuadrático, debido a que presentó un mayor R cuadrado ajustado, un menor criterio de información Akaike (AIC) y un menor criterio de información Bayesiano (BIC). Estos indicadores estadísticos contribuyeron al análisis de regresión lineal que se aplicó a los datos, para determinar una recta de mejor ajuste que permita predecir los niveles de venta.

Madrigal (2011) en uno de sus dos modelos propuestos menciona que, para pronosticar una serie de tiempo, se necesita suponer que existe un modelo matemático capaz de explicar el proceso que se está estudiando. El modelo consiste en “ajustar una recta para simular el efecto de tendencia lineal y cuya pendiente variará en cada estación para simular el efecto de estacionalidad creciente” (p.28).

Melo (2016) realizó un histograma de la serie de números de defunciones intrahospitalarias en el cual observó que sigue una distribución normal, por lo que “afirmó que los valores individuales de la serie giran en torno a su valor medio, por lo tanto, la serie es estacionaria en la media” (p.80). Lo anterior indica que una serie temporal sigue una distribución normal cuando la media y la varianza son constantes en un periodo de tiempo determinado.

El histograma de los datos históricos de ventas permitió conocer si existe normalidad en los datos y si la serie es estacionaria.

González (2015) en su investigación, calculó la media, la desviación estándar y el coeficiente de variación para los datos de la demanda de una pieza para automóviles. Con estos datos determinó que la variabilidad se encontraba un 55 % más arriba que la media. Por lo que concluyó que existen pedidos muy grandes y pequeños en el comportamiento de los datos en la serie de tiempo de 38 meses. La media, la desviación estándar y el coeficiente de variación, permitió conocer la existencia de valores atípicos en el comportamiento de la serie de tiempo.

González en su investigación desarrolló cuatro modelos clásicos de pronósticos basados en datos históricos, los cuales fueron: último valor, promedio, promedios móviles, y suavizado exponencial. La exactitud de los modelos anteriores se midió por medio de la desviación absoluta media (DAM). El mejor resultado de pronóstico para la pieza 190R lo obtuvo con el suavizado exponencial; la DAM obtenida en el estudio fue de 582.67 unidades. En esta investigación se aplicaron estos métodos de suavización y también se calculó el error cuadrático medio, para determinar cuál es el modelo predictivo de mejor ajuste para pronosticar las ventas de la compañía.

Gil (2016) utilizó el software estadístico R para el desarrollo de tres modelos de series de tiempo: lineal, cuadrático y cúbico; para el pronóstico de la demanda de energía eléctrica en Colombia, utilizando datos históricos comprendidos entre 2008 y 2014. Para este análisis de series de tiempo se utilizó el software estadístico R y Microsoft Excel, en el cual se implementaron, regresiones lineales y métodos de suavización.

Alvarado (2016) menciona que primero se debe determinar si una señal es estacionaria o no para conocer la complejidad de la predicción. “Una de las formas posibles para determinar si la serie de tiempo es estacionaria o no es observando su función de auto-correlación (ACF); si la gráfica de ACF decae rápidamente la serie es estacionaria, si la gráfica decae suavemente la serie es no estacionaria” (p.12).

La complejidad de la predicción depende fuertemente de la estacionalidad de la serie de tiempo, entre menos estacionaria sea, más complejo será determinar el modelo. Debido a esto se aplicaron métodos para aumentar esfuerzos e identificar cómo es la estacionalidad de los datos.

Con base en las fuentes de información consultadas es posible concretar que para la elaboración de un buen modelo de predicción es necesario conocer las herramientas que brinda la Estadística como, por ejemplo, las series de tiempo, estacionalidad, aleatoriedad y el análisis de regresión. Y también considerar los tipos de variables que se deben analizar, así como la calidad y cantidad de la información disponible, con el fin de lograr un buen nivel de predicción.

La administración de la cadena de restaurantes utiliza la información histórica de ventas para analizar lo que ya pasó, y para tomar medidas correctivas, mientras que con la misma información se puede analizar lo que puede pasar con cierto grado de confiabilidad.

La necesidad identificada es la elaboración de un modelo estadístico predictivo que contribuya a la toma de decisiones operativas y estratégicas de la administración.

Para satisfacer esta necesidad se plantean los siguientes pasos:

- Obtener la información histórica de ventas del 2015 a inicios de 2019.
- Delimitar el problema a restaurantes que operan únicamente en la ciudad de Guatemala.
- Analizar la información por medio de series de tiempo, para buscar patrones y tendencias. Desarrollo de una línea de mejor ajuste al comportamiento de la serie de tiempo, por medio de regresiones lineales.
- Por último, determinar el mejor modelo de predicción.

2. MARCO TEÓRICO

Los comportamientos de ventas se pueden determinar por medio de un análisis histórico de datos, identificando tendencias son útiles para realizar proyecciones, tomando en cuenta que existen factores económicos, culturales y sociales que son susceptibles al cambio, los cuales afectan el grado de confiabilidad del pronóstico.

Este capítulo muestra temas conceptuales como: series de tiempo y sus componentes, regresiones lineales, modelos lineales, y optimización de métodos clásicos de predicción.

2.1 Fundamentos estadísticos

A continuación, se detallan los métodos estadísticos para el desarrollo de los modelos predictivos

2.1.1 Series de tiempo

De acuerdo con Verdoy, Beltran y Peris (2015), una serie temporal consiste en una secuencia de observaciones cuantitativas distribuidas en un tiempo determinado. Las cuales son recolectadas por medio de información histórica. Este método es utilizado en áreas como telecomunicaciones, inventarios, *marketing*, economía, física entre otras.

Las series de tiempo permiten identificar tendencias por medio del análisis de datos en un periodo de tiempo determinado. De acuerdo con el tipo de comportamiento que se identifique se puede proponer un modelo estadístico predictivo que se ajuste a los datos y pueda así pronosticar con cierto grado de confiabilidad, comportamientos en un plazo determinado como días, meses, trimestres, bimestres o años.

En el presente trabajo de investigación se dan a conocer métodos estadísticos de pronósticos, los cuales pretenden hacer una descomposición en los datos, para identificar componentes en una serie de tiempo, tales como: tendencia, estacionalidad, ciclo e irregularidades. (Hanke y Wichern, 2006).

2.1.2 Componentes de una serie de tiempo

El análisis de las series de tiempo contempla cuatro factores: tendencia, estacionalidad, variaciones cíclicas, y aleatoriedad o irregulares (Verdoy, *et al.*, 2015). Las cuales se describen a continuación.

2.1.2.1 Tendencia

El estudio de Hanke y Wichern (2006) muestra la tendencia como el crecimiento o decrecimiento del valor de los datos en un periodo determinado. Se puede definir también como la variación de la media de los datos a lo largo de un tiempo definido.

2.1.2.2 Estacionalidad

La estacional indica los cambios que son considerados regulares que se observan en el conjunto de datos. Por dar un ejemplo en un lapso de 5 años, las ventas de los restaurantes aumentan cada fin de año, debido a que coincide con un mes festivo.

De acuerdo con Arellano (2001), la estacionalidad representa un movimiento periódico que puede ser afectado por condiciones de tiempo como: invierno, verano u otras.

2.1.2.3 Componente cíclica

La componente cíclica se considera como variaciones que se generan en un periodo no menor a un año. Esta componente puede oscilar alrededor de la tendencia en forma de onda. (Hanke y Wichern, 2006)

Se genera por lo general por condiciones económicas como abundancia (crestas) o escasez (valles) de poder adquisitivo. También muestra comportamientos recurrentes, aunque no necesariamente periódicos, su amplitud y periodicidad se presentan con cierta irregularidad.

2.1.2.4 Componente aleatoria o irregular

Esta mide el grado de variabilidad aleatoria que es ocasionada regularmente por factores no contemplados y no recurrentes. Cuando existe una componente aleatoria en una serie de temporal, es porque ocurren cambios inusuales o comportamientos atípicos en la serie de datos. (Arango, *et al.*, 2013)

2.1.3 Series de tiempo estacionarias y no estacionarias

Por lo general, una serie de tiempo se puede clasificar como estacionaria o no estacionaria. Las series estacionarias consisten en valores que oscilan alrededor de un valor medio de la serie (la media de la fuente de datos).

Las series que no tienen estacionalidad indican un cierto grado de tendencia, es decir, que los valores de la media y la variabilidad cambia a lo largo del tiempo. Estos cambios en la media determinan una tendencia creciente o decreciente, lo que indica que la serie no oscila alrededor de un valor constante. (Botero y Cano, 2008)

2.1.4 Coeficientes de autocorrelación y correlogramas

Con el objetivo de conocer si la serie de temporal conformada con la data histórica de la empresa, presentan componentes mencionadas anteriormente como: tendencias, estacionariedad, y aleatoriedad, es indispensable examinar la serie temporal utilizando la autocorrelación. Esta es una función que proporciona información acerca de la correlación entre los datos históricos.

De acuerdo con Arango, *et al.*, (2013), al aplicar la autocorrelación se puede determinar los siguientes valores: “longitud del periodo estacional (I). Los valores iniciales de los factores estacionales de cada período (I_t), el valor inicial del intercepto (A_0) y de la tendencia (T_0)” (p.744). Derivado del análisis de autocorrelación, es posible identificar los patrones de datos que se presentan en una serie de tiempo.

2.1.4.1 Coeficiente de autocorrelación

La ecuación 1 es utilizada para obtener los coeficientes de la autocorrelación, entre los datos y desfasados en periodos.

$$r_k = \frac{\sum_{t=1}^{N-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^N (X_t - \bar{X})^2} \quad (\text{Ec.1})$$

donde:

r_k : es el coeficiente de autocorrelación

X_t : es la observación en el periodo t

X_{t+1} : es la observación desplazada k periodos

2.1.4.2 Correlograma

Un correlograma consiste en una gráfica con k periodos. El análisis de este permite identificar la existencia de correlación entre los datos de una serie temporal. Esta correlación muestra valores que ayudan a identificar tendencias, estacionalidades y aleatoriedades. (Macías 2015)

De acuerdo con este mismo autor, si la serie identificada tiene aleatoriedad, existe independencia entre los datos de la serie temporal, y los valores de la correlación para t y $t+k$ para cualquier desfase en los datos de k son próximos a cero (0).

Si la serie temporal tiene tendencia, los datos continuos estarán muy correlacionados y los valores de correlación serán relevantes (cerca de 1) para los primeros periodos de tiempo, reduciendo paulatinamente con el incremento del valor de k .

Por último, si la serie temporal posee estacionalidad, el valor de correlación será representativamente alto para periodos $k > 1$. Es decir, la estacionalidad se repetirá cada k periodos.

2.1.5. Prueba de Dickey-Fuller

Como parte de la validación, se utilizan estadísticos de pruebas que permiten validar si una serie temporal presenta estacionariedad. Una de estas pruebas estadísticas es conocida como test de Dickey-Fuller. Esta prueba utiliza un gráfico llamado correlograma, la cual brinda información acerca del comportamiento de los datos en el tiempo y de las componentes de una serie temporal. (Sánchez, 2008)

Según el estudio realizado por este mismo autor la autoregresión de orden 1, muestra que el valor de una serie de tiempo en un momento preciso depende del valor previo del conjunto de datos de la serie:

$$y_t = \rho * y_{t-1} + \epsilon_t \quad (\text{Ec.2})$$

Donde:

y_t : Representa la variable de interés,

ρ : coeficiente del primer rezago

ϵ_t : es un término de error que siempre estará presente en la serie.

Se dice que hay una raíz unitaria si $\rho=1$, lo que indica que la serie temporal contiene una componente aleatoria y no es estacionaria.

Cuando se realiza la prueba de Dickey-Fuller se puede verificar si la serie es estacionaria o no. La hipótesis nula H_0 se establece como: la serie de datos contiene una raíz unitaria, esto quiere decir que la serie no es estacionaria.

La hipótesis alternativa es todo lo contrario. Si el P-valor es mayor al nivel de significancia, se acepta la hipótesis nula.

2.1.6 Selección del método de pronóstico

En este trabajo se emplearon métodos predictivos tradicionales: promedios móviles, tratamiento exponencial y extensiones de suavizado exponencial. Existen métodos que son considerados más complejos, y que se trabajan en herramientas con R-Studio como los son SARIMA, ARIMA o Box-Jenkins. En esta oportunidad se trabajaron los métodos tradicionales mencionados para el análisis.

Para escoger el método óptimo predictivo, es importante identificar y tener claro cuáles son las componentes de serie de tiempo, tendencia, estacionalidad, y aleatoriedad. (Hanke y Wichern, 2006) Debido a que, estas componentes son la base para diseñar el modelo predictivo de mejor ajuste.

Tabla III. **Técnicas de previsión según criterios de estacionalidad y tendencia**

	Sin estacionalidad		Con estacionalidad	
Sin tendencia	Promedio móvil simple	Suavización exponencial simple	Estacional aditivo	Estacional multiplicativo
Con tendencia	Promedio móvil doble	Suavización exponencial doble	Holt-Winters' Aditivo	Holt-Winters' multiplicativo

Fuente. González (2015). *Pronóstico de demanda de refacciones de automóvil, mediante series de tiempo, redes neuronales artificiales y modelos híbridos.*

Para seleccionar la mejor técnica de pronóstico que sea capaz de predecir el comportamiento de los datos en el futuro determinado, es importante identificar y tener claro cuáles son las componentes de la serie temporal, tendencia, estacionalidad, y aleatoriedad. Debido a que, estas componentes son la base para diseñar el modelo predictivo de mejor ajuste.

2.1.6.1 Técnicas de pronóstico

Si durante el análisis se identifica que el valor de la media es constante en el periodo del análisis, se considera que el comportamiento de los datos son estacionarios, y se recomienda que los métodos de predicción adecuadas son: promedios simples, móviles, estacional aditivo y multiplicativo.

Cuando la serie es analizada y se observa cierto grado de tendencia, se recomienda utilizar los métodos siguientes de predicción: promedios móviles, suavizamiento exponencial y regresión simple. Y cuando se identifica estacionalidad en la serie, es recomendable utilizar los métodos de suavización exponencial de Winters o regresión múltiple. (Anderson, Sweeney, y Williams, 2008)

2.1.6.1.1 Promedio móvil simple

Es un método utilizado cuando la serie de datos presenta estacionalidad, o bien, cuando la serie muestra ser constante en un lapso. De acuerdo con Anderson, Sweeney, y Williams (2008), el término móvil es utilizado debido a que cada vez que se realiza una nueva observación en la serie, ésta reemplaza al valor previo y se toma una nueva medición del promedio.

La ecuación para pronosticar un valor futuro, haciendo uso de promedios móviles es la siguiente:

$$X_{t+1} = \frac{X_t + X_{t-1} + \dots + X_{t-k+1}}{k} \quad (\text{Ec.3})$$

Donde:

X_{t+1} : es el valor pronosticado para el siguiente periodo

X_t : valor real en el periodo t

k: es el número de términos en el promedio móvil.

2.1.6.1.2 Promedio móvil doble

Es un método recomendado cuando la serie de tiempo muestra una tendencia con cierta linealidad. Este método consiste en calcular dos veces el promedio móvil, es decir, que se calcula el promedio móvil con orden k empleando la ecuación (Ec.3) descrita anteriormente. (Sweeney, y Williams, 2008) Luego se aplica la siguiente ecuación para calcular el segundo promedio móvil.

$$M'_t = \frac{M_t + M_{t-1} + \dots + M_{t-k+1}}{k} \quad (\text{Ec.4})$$

Donde:

$M_t = X_{t+1}$: es el promedio móvil de orden k.

M'_t : es el segundo promedio móvil

k: es el número de términos en el segundo promedio móvil.

2.1.6.1.3 Suavizamiento exponencial

Este método utiliza un promedio ponderado para los valores actuales o más recientes. El cálculo se realiza de manera automática y se van reduciendo cuando los datos son más antiguos (Anderson, Sweeney, y Williams, 2008)

$$F_{t+1} = \alpha X_t + (1 - \alpha)F_t \quad (\text{Ec.5})$$

Donde:

M_{t+1} : pronóstico para el periodo t+1 de la serie de tiempo

X_t : valor real en el periodo t de la serie de tiempo.

F_t : pronóstico para el periodo t de la serie de tiempo.

α : constante de suavizamiento.

De acuerdo con estos autores “el pronóstico obtenido mediante suavizamiento exponencial para cualquier periodo es un promedio ponderado de todos los valores reales anteriores de la serie de tiempo”. (Anderson, Sweeney, y Williams, 2008, p.774). Con este método se revisa de forma continua un pronóstico usando las experiencias más recientes (Hanke y Wichern, 2006).

2.1.6.1.4 Exponencial Holt

Es un método que se emplea al momento que la serie de datos tiene una tendencia pronunciada. Utiliza un ajuste exponencial a la tendencia

Las ecuaciones para pronosticar los valores futuros de la serie de tiempo por este método son las siguientes:

Nivel: $M_t = \alpha * X_t + (1 - \alpha) * (M_{t-1} + T_{t-1})$ (Ec. 6)

Tendencia: $T_t = \beta * (M_t - M_{t-1}) + (1 - \beta) * T_{t-1}$ (Ec. 7)

Previsión: $F_{t+n} = M_t + n * T_t$ (Ec.8)

Donde:

M_t : estimado del nivel actual

T_t : estimado de la tendencia

F_{t+n} : pronóstico para el periodo n en el futuro

n : periodo a pronosticar en el futuro

t : periodo actual

α : constante de suavizamiento para el nivel $0 < \alpha < 1$

β : constante de suavizamiento para el estimado de tendencia $0 < \beta < 1$.

2.1.6.1.5 Método de Winter

Este método es muy útil si la serie de datos además de tener tendencia también presenta estacionalidad.; se ajusta a las variaciones de tendencia y estacionales. Este método utiliza cuatro ecuaciones para realizar el pronóstico, las cuales son:

Nivel: $M_t = \alpha X_t S_{t-s} + (1 - \alpha) * (M_{t-1} + T_{t-1})$ (Ec.9)

Tendencia: $T_t = \beta * (M_t - M_{t-1}) + (1 - \beta) * T_{t-1}$ (Ec.10)

Estacionalidad: $S_t = \gamma X_t M_t + (1 - \gamma)(S_{t-s})$ (Ec.11)

Previsión: $F_{t+n} = (M_T + nT_t)S_{t-s+n}$ (Ec.12)

Donde:

M_t : Estimado del nivel actual

T_t : Estimado de la tendencia

F_{t+n} : Pronóstico para el periodo n en el futuro

n : Número del periodo futuro a pronosticarse

t : Periodo actual s Longitud de la estacionalidad.

α : Constante de suavizamiento para el nivel $0 < \alpha < 1$

β : Constante de suavizamiento para el estimado de tendencia $0 < \beta < 1$

γ : Constante de suavizamiento para el estimado de estacionalidad $0 < \gamma < 1$

2.1.6.2 Regresiones lineales

A continuación, se detallan las características de una regresión lineal.

2.1.6.2.1 Objetivo

El objetivo principal de las regresiones lineales es explicar situaciones o fenómenos que han sido observados en muchas ocasiones y que cuentan con varias variables, considerando que existe una relación entre ellas. La relación de estas variables se diseña por medio de un modelo matemático, el cual pueda explicar el comportamiento y grado de asociación. (Tusell 2013) De acuerdo con este autor una regresión lineal permite realizar predicciones sobre las variables dependientes, logrando explicar con cierto grado de confiabilidad, el comportamiento del fenómeno y la relación entre las variables.

2.1.6.2.2 Notación

Considerando una variable aleatoria Y (variable regresora) la cual está dada por:

$$Y = \beta_0 X_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1} + \varepsilon \quad (\text{Ec.13})$$

Donde:

$\beta_0, \dots, \beta_{p-1}$: parámetros fijos desconocidos.

$X_0, \dots, X_{p-1}$: variables explicativas no estocásticas, regresores

ε : variable aleatoria inobservable.

2.1.6.2.3 Variable

Es una función que arroja posibles valores de un fenómeno en estudio. De acuerdo con Tusell (2013), la ecuación (Ec.4) "Indica que la variable aleatoria Y se genera como combinación lineal de las variables explicativas, salvo en una perturbación aleatoria" (p.245).

2.1.7. Índices estacionales

La estacionariedad en las series temporales es un comportamiento que se manifiesta en las cercanías de la tendencia que transcurre en un periodo de un año o menos, el cual es predecible y repetible. Levin y Rubin (2004) expresan que para detectar la variación estacional los intervalos de tiempo deben medirse en unidades pequeñas, como días, semanas, meses o trimestres.

2.1.7.1 Promedio móvil para medir la variación estacional

Según lo establecido por Levin y Rubin, (2004) “El método de promedio móvil proporciona un índice que describe el grado de variación estacional. El índice está basado en una media de 100, con el grado de estacionalidad medido por las variaciones respecto a la base” (p.692).

2.1.7.2 Pasos para calcular el índice estacional

De acuerdo con Mason, *et al.*, (2002) los pasos para calcular los índices estacionales son los siguientes.

- Calcular el promedio móvil de cada trimestre, mes, día, entre otros.
- Centrar el promedio móvil.
- Calcular el porcentaje del valor real respecto al valor del promedio móvil.

$$\frac{Real}{Promedio\ móvil} * 100 \quad (Ec.14)$$

- Reunir las repuestas del paso anterior y calcule la medida modificada.
- Aplicar un factor de corrección a cada valor medio para obtener un promedio fijado en la unidad.

$$\text{Factor de corrección} = \frac{\text{Cantidad de periodos}}{\text{Total de las mediciones}} \quad (\text{Ec.15})$$

- Por último, multiplicar el factor de corrección por cada porcentaje calculado en el paso 3.

2.1.8. Error cuadrático medio

El ECM de un parámetro que permite determinar el diferencial entre el estimador y su valor verdadero. Esto se realiza calculando el promedio de los errores y posteriormente elevarlos al cuadrado. Este cálculo emplea la siguiente fórmula. Vélez (2019).

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n} \quad (\text{Ec.16})$$

Donde:

EMC : Error cuadrático medio.

Y_t : Valor Real.

Y'_t : Estimador.

n : tamaño de muestra.

t : Periodo.

2.1.9. Modelo de Holt-Winter estacional aditivo

De acuerdo con Pérez (2017), este método aplica el alisamiento a tres parámetros, promedio, tendencia y estacionalidad. “Cada una tiene su constante de alisamiento, las que pueden ser ajustadas según la situación, pudiendo ser esto de forma individual, pues el cambio en una no altera las demás” (p.22).

Esta técnica de pronósticos se implementa para analizar series temporales que poseen tendencia y estacionalidad. La ecuación de pronóstico hecha en t para un punto a T períodos futuros viene dada por.

$$Y_{t+T} = (a_t + b_t T) + SF_{t+T} \quad (\text{Ec.16})$$

Donde:

Y_{t+T} : Es el valor estimado para un punto a T periodos futuros.

a_t : Intercepto de la línea de tendencia en t .

b_t : Pendiente de la línea de tendencia en t .

T : El numero de periodos que corresponde a partir de la serie.

SF_{t+T} : La mejor estimación del factor estacional en $t+T$.

Perez (2017).

2.2 Toma por medio de datos de ventas

La toma de decisiones por medio de datos es una gran aportación para que las empresas se empiecen a involucrar e innovar con nuevas estrategias para el negocio.

2.2.1. Ventas

¿Qué son las ventas? Las ventas a lo largo de la historia es la actividad que se practica en cualquier sector comercial, siempre existirá algo para vender. La cadena de restaurantes maneja el tipo de venta directa, la cual consiste en contacto directo entre comprador y vendedor.

La venta directa incluye un proceso en el cual, el vendedor identifica las necesidades del cliente, y crea una incitación para que exista un intercambio de bienes o servicios y de esta manera lograr satisfacer los requerimientos de ambas partes. (Stanton, Etzel y Walker, 2004)

Para la operatividad de la cadena de restaurantes es relevante conocer el comportamiento de las ventas que se realizan en el transcurso del tiempo, con el objetivo de tener un mejor panorama de la situación de la empresa y poder planificar y ejecutar acciones que contribuyan a la toma de decisiones en áreas como productividad y ventas.

2.2.2 Importancia de un pronóstico de ventas

En la actualidad las gerencias de las empresas consideran que un pronóstico de ventas forma parte de la información más importante que se maneja en las operaciones, porque esta técnica contribuye en actividades críticas como:

- Incrementar las compras de materia prima
- Contratar más operarios en la planta o ampliar las horas laborales
- Aumentar la capacidad de almacenaje
- Uso de más fletes
- Cambios en el presupuesto y planificación de ventas
- Lanzamiento de nuevos canales de ventas
- Retirar productos del mercado
- Mayor inversión en materia prima

De acuerdo con Stanton, *et al.*, (2004) un pronóstico de ventas se considera como un estimado de las ventas producidas en un periodo futuro. Estos

también son utilizados como base para decidir cuánto invertir en distintas actividades como marketing y ventas directas. “Con la base de las ventas anticipadas se planea la cantidad necesaria de capital de trabajo y uso de la planta” (Stanton, *et al.*, 2004, p.188).

2.2.3. Alcance del pronóstico de ventas

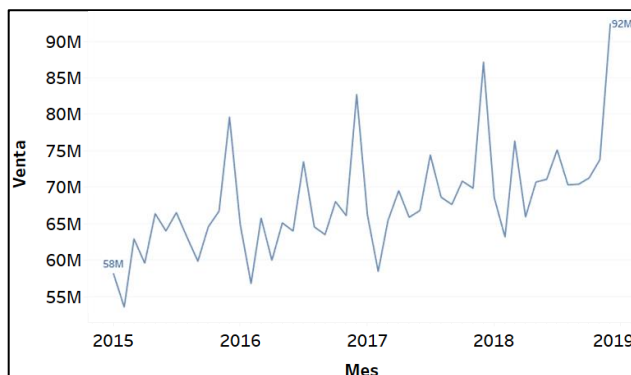
Por recomendación se considera diseñar un pronóstico de ventas para cada servicio o producto que ofrecen las organizaciones, porque de esta forma se puede tomar decisiones más estratégicas en departamentos como compras, suministros, mercadeo, producción, aprovisionamiento y contabilidad. Esta tesis aportó un monitoreo más robusto en las operaciones y ayudó que el presupuesto se cumpla de acuerdo con lo planificado.

3. PRESENTACIÓN DE RESULTADOS

3.1. Modelo de mejor ajuste a la tendencia de ventas

Para el desarrollo del primer objetivo se realizó un análisis de estacionalidad utilizando como herramienta la gráfica de correlograma, y la prueba estadística de Dickey Fuller, para determinar si existe una raíz unitaria. Por medio del correlograma se logró determinar de manera visual la correlación que existe entre las observaciones de la serie de datos, y validar si existe tendencia, aleatoriedad, estacionalidad, y ciclos. La prueba de Dickey Fuller permitió determinar la existencia o ausencia de raíces unitaria en los datos. Por último, se aplicaron métodos de pronósticos y se seleccionó el que presentó menor error. A continuación, se presentan los detalles del análisis. Al graficar la serie temporal, se logró observar que el comportamiento de los datos presenta cierto grado de estacionalidad y también una tendencia creciente. El periodo de los datos es del 2015 al 2018.

Figura 1. **Serie de tiempo, ventas de restaurantes 2015-2018**



Los datos de volumen de ventas están expresados en millones de quetzales guatemaltecos.

Fuente: elaboración propia, realizado con Tableau.

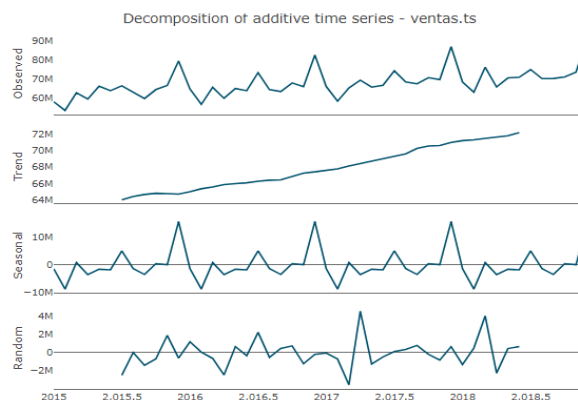
Tabla IV. Volumen de ventas de restaurantes 2015-2018

Mes	2015	2016	2017	2018
enero	58.11M	64.81M	66.15M	68.51M
febrero	53.50M	56.73M	58.38M	63.12M
marzo	62.87M	65.73M	65.37M	76.32M
abril	59.54M	59.92M	69.48M	65.88M
mayo	66.34M	65.07M	65.83M	70.67M
junio	63.96M	63.95M	66.76M	71.05M
julio	66.48M	73.47M	74.39M	75.09M
agosto	63.07M	64.51M	68.59M	70.29M
septiembre	59.80M	63.45M	67.58M	70.38M
octubre	64.52M	67.99M	70.79M	71.23M
noviembre	66.68M	66.06M	69.82M	73.74M
diciembre	79.59M	82.70M	87.14M	92.40M

Fuente: elaboración propia.

Se realizó la descomposición de la serie temporal de datos, con la finalidad de validar con más detalle cada componente, y se logró observar a nivel de gráfico, que el conjunto de datos tiene una tendencia positiva, una componente de estacionalidad que se repite en cada año y una componente aleatoria que produce cambios en los datos, ocasionados por fenómenos difíciles de explicar y que entran en el ámbito del azar

Figura 2. Descomposición de la serie de tiempo



Fuente: elaboración propia, realizado con R-Studio.

Por inspección visual se puede observar en la serie de tiempo, figura 1, que existe cierto grado de tendencia creciente con el paso del tiempo. Pero este primer análisis es muy subjetivo y se necesita de un análisis más sofisticado que permita identificar si la serie de tiempo contiene componentes de tendencia, estacionariedad o aleatoriedad. Una herramienta que permite identificar la presencia de tendencia en una serie temporal de datos es la implementación de los valores arrojados por la autocorrelación, estos permiten identificar si existe algún nivel de correlación.

Para el proceso mencionado anteriormente se empleó la autocorrelación (ecuación 4) y se realizó la gráfica, figura 3, donde se observó que existe un valor elevado de correlación en $k > 1$, lo que significa que la serie de datos muestra estacionalidad en el periodo $k=12$, y se encontró un valor relevante (cercano a 0.7). Otro aspecto que se observó fue un valor relevante distinto de cero para $k=1$, esto indicó que la serie de datos presentaba un grado importante de estacionalidad.

Figura 3. Función de autocorrelación para la serie de tiempo de ventas



Fuente: elaboración propia, realizado con R-Studio.

En la tabla V se presentaron los datos obtenidos con prueba Dickey-Fuller, la cual presentó un valor de p (0.01), menor al nivel de significancia (0.05), por lo

cual se puede determinar que se rechaza la hipótesis nula (H_0 : existe una raíz unitaria para la serie) esto indica que el conjunto de datos no posee una raíz unitaria y que está tiene un comportamiento estacionario.

Tabla V. Prueba Dickey-Fuller

Descripción	Valor
Valor observado	-4.8548
Valor crítico	3
valor-p	0.01
alfa	0.05

Fuente: elaboración propia.

De acuerdo con la tabla I, donde se muestran las técnicas predictivas según los factores estacionarios y de tendencia, el método recomendado para predecir las ventas en la cadena de restaurantes es Holt-Winter; pero cabe mencionar que en este trabajo se emplearon seis (6) técnicas distintas (móvil simple, doble, exponencial, método Holt, Regresión Lineal y método de Winters).

Promedio móvil doble: se elaboró el modelo de pronóstico utilizando el método de promedio móvil doble, con el cual se procedió a estimar el comportamiento de las ventas. Los resultados fueron los siguientes:

Tabla VI. **Resultados del pronóstico con promedio móvil doble**

2016	2017	2018	2019
63.02M	63.86M	68.79M	71.00M
72.91M	74.61M	79.10M	70.83M
73.93M	74.58M	77.74M	70.67M
67.28M	67.78M	71.87M	70.50M
65.31M	66.21M	74.30M	70.34M
54.64M	59.57M	62.37M	
57.74M	61.84M	65.92M	
63.93M	68.03M	71.69M	
69.55M	73.65M	72.16M	
70.55M	71.37M	73.72M	
67.61M	70.63M	72.40M	
68.76M	71.88M	72.20M	

Fuente: elaboración propia.

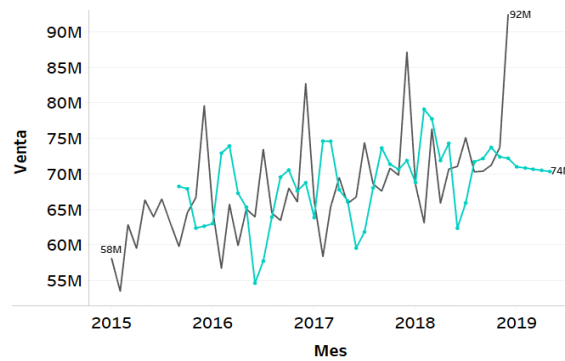
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se utilizó la ecuación (Ec.16)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 1.38 \times 10^{13}$$

Se logró observar que el valor del error cuadrático medio es demasiado elevado, por lo cual se determina que este método no es recomendable para esta serie de tiempo.

Figura 4. **Gráfica del pronóstico promedio móvil doble y venta real**



Fuente: elaboración propia, realizado con Tableau.

Promedio móvil simple: se elaboró el modelo predictivo con el método de promedio móvil simple, para estimar los niveles de ventas. Los resultados fueron los siguientes.

Tabla VII. **Resultados del pronóstico con promedio móvil simple**

2015	2016	2017	2018
	67.65M	70.05M	73.83M
	68.90M	70.72M	74.07M
	66.95M	68.32M	72.15M
	66.71M	68.15M	73.77M
58.50M	61.80M	64.85M	68.46M
60.56M	61.86M	64.77M	69.00M
63.17M	63.67M	66.86M	70.98M
64.08M	65.60M	69.12M	70.67M
64.96M	66.75M	68.89M	71.77M
63.33M	66.35M	69.33M	71.70M
63.47M	67.36M	70.34M	71.75M
63.52M	65.50M	69.20M	71.41M

Fuente: elaboración propia.

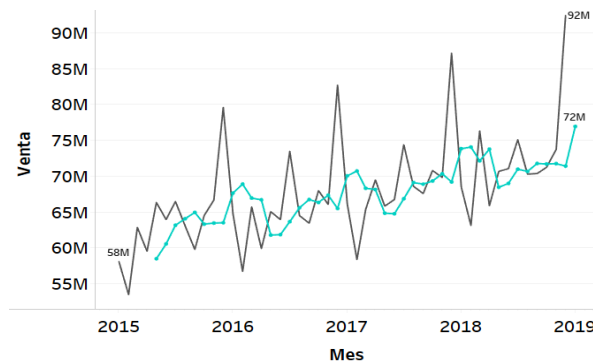
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 4.01 \times 10^{12}$$

Se logra observar que el valor del error cuadrático medio es demasiado elevado, por lo cual se determina que el método de promedio móvil simple no es recomendable utilizarlo en esta serie de tiempo.

Figura 5. Gráfica pronóstico promedio móvil simple



Fuente: elaboración propia, realizado con Tableau.

Suavizamiento exponencial doble:

Se realizó el pronóstico de ventas con el método de suavizamiento exponencial doble; esto permitió estimar la venta de los primeros 5 meses del año 2019. Los resultados fueron los siguientes.

Tabla VIII. **Resultados del método suavizamiento exponencial doble**

2015	2016	2017	2018
	84.3M	85.5M	90.5M
58.1M	71.1M	73.5M	76.0M
52.6M	49.4M	50.5M	54.6M
61.7M	59.9M	59.1M	72.4M
63.7M	61.5M	71.7M	69.7M
67.5M	63.4M	68.4M	67.2M
67.8M	66.0M	65.9M	72.5M
67.2M	75.5M	76.0M	76.7M
64.1M	69.0M	72.5M	72.3M
57.7M	60.0M	65.7M	68.5M
62.7M	66.9M	70.1M	70.7M

Fuente: elaboración propia.

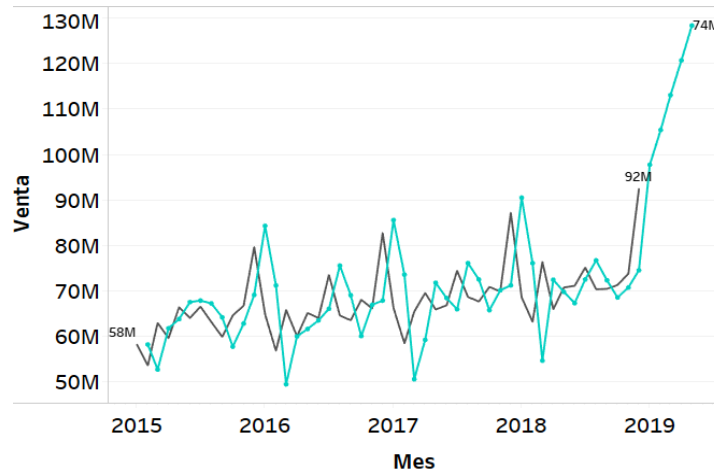
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 163,194.65$$

Se logró observar que el valor del error cuadrático medio se reduce considerablemente respecto a los dos métodos anteriores, pero continúa siendo bastante elevado; por lo que este método no es recomendable para la serie de datos que se está analizando.

Figura 6. **Gráfica de pronóstico suavizamiento exponencial doble**



Fuente: elaboración propia, realizado con Tableau.

Método de dos parámetros o Holt.

Se desarrolló la predicción de ventas con el método de dos parámetros o Holt, lo que permitió estimar la venta de los primeros 5 meses del año 2019. Los resultados fueron los siguientes.

Tabla IX. **Resultados del pronóstico con método Holt**

2015	2016	2017	2018
	64.5M	68.1M	71.6M
58.1M	64.7M	68.0M	71.4M
57.5M	63.8M	66.9M	70.6M
58.2M	64.1M	66.9M	71.5M
58.4M	63.7M	67.3M	70.9M
59.4M	64.0M	67.3M	71.1M
60.0M	64.0M	67.3M	71.2M
60.8M	65.3M	68.3M	71.9M
61.2M	65.3M	68.5M	71.8M
61.1M	65.2M	68.5M	71.8M
61.6M	65.7M	69.0M	71.9M
62.3M	65.8M	69.2M	72.3M

Fuente: elaboración propia.

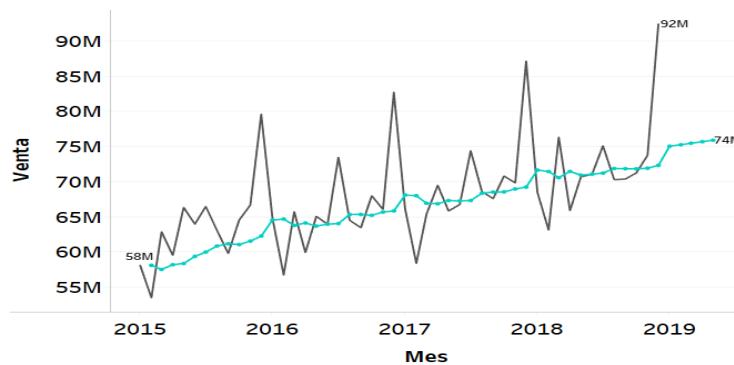
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 93,830.95$$

Se observó que el valor del error cuadrático medio se reduce respecto al método anterior, pero continúa siendo bastante elevado; por lo que este método no es recomendable para la serie de datos analizada.

Figura 7. Gráfica pronóstico método Holt



Fuente: elaboración propia, realizado con Tableau.

Método de Holt-Winter.

Se realizó la predicción de ventas con el modelo Holt-Winter, con el cual se procedió a estimar la venta de los primeros 5 meses del año 2019. Los resultados se observan en la tabla X.

Tabla X. Resultados del pronóstico con método Holt-Winter

2015	2016	2017	2018
	61.8M	66.2M	68.9M
58.1M	58.9M	59.5M	61.2M
57.9M	65.7M	68.1M	69.2M

Continuación tabla X.

58.1M	63.4M	63.1M	70.8M
58.2M	68.1M	68.9M	69.8M
58.6M	66.1M	67.4M	70.1M
59.0M	67.7M	74.6M	78.1M
59.4M	65.5M	67.4M	71.4M
59.7M	63.1M	65.9M	70.1M
59.8M	66.6M	70.3M	73.8M
60.2M	68.0M	69.5M	72.7M
60.6M	76.5M	84.1M	89.8M

Fuente: elaboración propia.

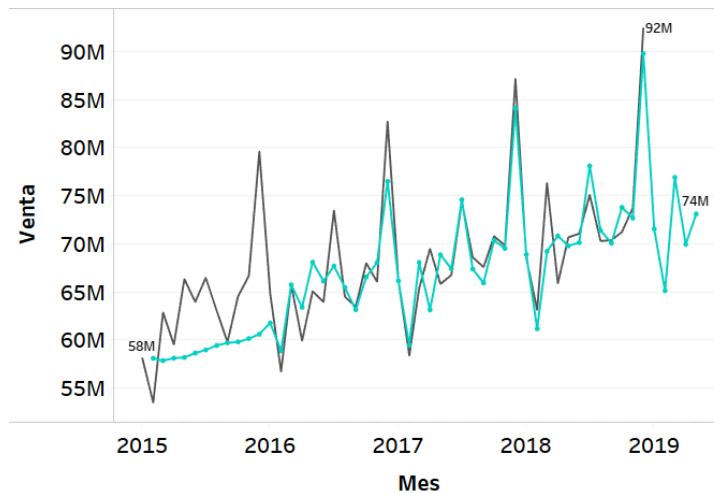
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 65,195.23$$

Se logró observar que el valor del error cuadrático medio se reduce respecto al método anterior, este modelo es adecuado, debido a que considera la tendencia y la estacionariedad en la serie de datos.

Figura 8. **Gráfica de pronóstico método Holt-Winter**



Fuente: elaboración propia, realizado con Tableau.

Regresión lineal.

Se realizó una regresión lineal con los datos históricos de ventas, en la cual se determinó una recta de ajuste al comportamiento de los datos, y se obtuvieron los siguientes resultados.

Tabla XI. **Resultados del pronóstico con una regresión lineal**

2015	2016	2017	2018
60.7M	64.4M	68.0M	71.7M
61.0M	64.7M	68.3M	72.0M
61.3M	65.0M	68.6M	72.3M
61.6M	65.3M	68.9M	72.6M
62.0M	65.6M	69.2M	72.9M
62.3M	65.9M	69.5M	73.2M
62.6M	66.2M	69.8M	73.5M
62.9M	66.5M	70.1M	73.8M
63.2M	66.8M	70.5M	74.1M
63.5M	67.1M	70.8M	74.4M
63.8M	67.4M	71.1M	74.7M
64.1M	67.7M	71.4M	75.0M

Fuente: elaboración propia.

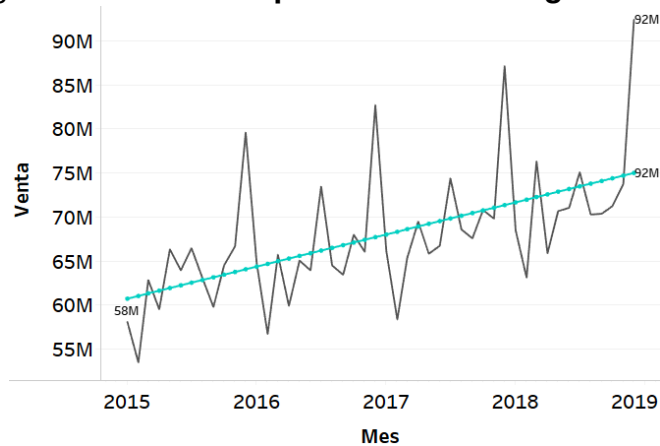
Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 86,344.57$$

El valor del error cuadrático medio se reduce respecto al método anterior, este modelo es adecuado, debido a que considera la tendencia y la estacionalidad en la serie de datos.

Figura 9. Gráfica del pronóstico con regresión lineal



Fuente: elaboración propia, realizado con Tableau.

Tabla XII. Pronósticos de las ventas en la cadena de restaurantes

Método	Promedio Móvil Simple	Promedio Móvil Doble	Suavización Exponencial Doble	Holt	Holt-Winter	Regresión Lineal
Pronóstico enero 2019	Q.77M	Q.71M	Q.98M	Q.72M	Q.69M	Q.72M
ECM	115,410	163,653	166,666	93,830	65,195	86,344

Fuente: elaboración propia.

Se determinó que el método de pronósticos que presentó el menor error en la medición es el método de Holt-Winter el cual obtuvo un error de 65,195.23. Este método fue el seleccionado para optimizar sus parámetros.

Se aplicó el método de Holt Winter Aditivo para determinar el mejor ajuste al comportamiento de ventas de la cadena de restaurantes. Se utilizó este método debido a que la magnitud de las fluctuaciones estacionales de la serie no cambia significativamente al hacerlo la tendencia.

Modelo Holt-Winter aditivo.

Se realizó este método aplicando los 3 parámetros que sugiere la teoría (α , β , γ). Los cuales se seleccionaron al azar, para el desarrollo del método; los cuales fueron optimizados con la herramienta de solve de Microsoft Excel:

Tabla XIII. Parámetros utilizados completamente al azar

Alfa	0.3
Beta	0.2
Gamma	0.7

Fuente: elaboración propia.

Dentro del análisis grafico se observó un patrón cíclico anual. Con lo cual se puede definir que el ciclo de la serie de datos es de 12 meses. Se realizó una regresión lineal para determinar la pendiente de los datos, el cual se utilizó para el desarrollo del modelo. El valor del coeficiente de determinación R^2 es menor a 0.5, esto muestra una estacionalidad significativa. Consultar tabla XI.

Tabla XIV. **Indicadores de la regresión lineal de los datos**

Estadísticas de la regresión	
Coeficiente de correlación múltiple	0.57962145
Coeficiente de determinación R ²	0.33596103
R ² ajustado	0.3215254
Error típico	6038514.78
Observaciones	48

Fuente: elaboración propia.

Dentro del análisis de la regresión lineal se validó que el porcentaje de probabilidad de la pendiente de los datos es menor que el 5 %, lo que indica que existe una tendencia significativa. Este análisis se utilizó para extraer el valor de la pendiente, el cual se empleó en el desarrollo del modelo aditivo Holt-Winter.

Tabla XV. **Coeficientes de intercepto y pendiente de la regresión lineal**

L1	Q63,704,440
T1	303,514.71

Fuente. elaboración propia.

Con los datos del ciclo de la serie de tiempo, que fue 12 meses y el promedio de los valores de este periodo, se determinaron los parámetros de longitud y para el parámetro de tendencia se utilizó el valor de la pendiente.

Tabla XVI. **Longitud del ciclo y pendiente de la serie de datos**

	Coeficiente	Error típico	Estadístico t	Probabilidad	< 95 %	> 95 %
Intercepto	60.43 M	1.77 M	34	0	56.87 M	63.99 M
Variable	303.51 K	62,915	4.8	0.002 %	176.8 K	430.1 K

Fuente: elaboración propia.

Con estos parámetros se desarrolló el modelo Holt-Winter Aditivo, y se obtuvieron los siguientes resultados.

Tabla XVII. Pronóstico Holt-Winter aditivo año 2019 y 2020

2016	2017	2018	2019	2020
58.42M	65.82M	69.65M	72.62M	79.91M
54.11M	59.29M	62.40M	66.47M	70.10M
63.78M	68.56M	70.33M	77.49M	81.13M
60.75M	63.90M	70.82M	71.56M	75.20M
67.85M	69.86M	71.13M	74.50M	78.14M
65.78M	68.35M	71.05M	74.70M	78.33M
68.60M	75.11M	78.33M	80.07M	83.70M
65.50M	68.56M	72.22M	74.73M	78.36M
62.53M	66.71M	70.87M	74.22M	77.86M
67.56M	71.46M	74.71M	76.31M	79.95M
70.02M	71.34M	74.09M	77.53M	81.16M
83.24M	86.56M	90.54M	95.28M	98.91M

Fuente: elaboración propia.

Luego se procedió a calcular el error cuadrático medio para este método, para la cual se empleó la ecuación (Ec.14)

$$EMC = \frac{\sum_{t=1}^n (Y_t - Y'_t)^2}{n}$$

$$EMC = 54,139$$

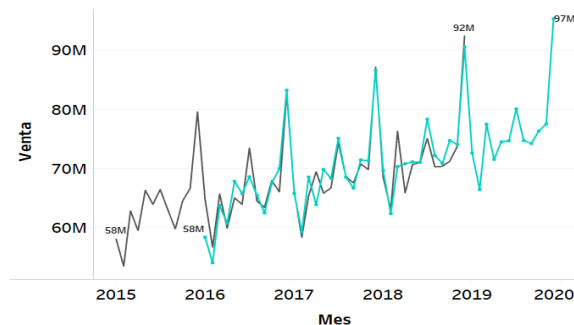
Se logra observar que el valor del error cuadrático medio se reduce considerablemente respecto al modelo Holt-Winter. Por lo que este es el modelo optimizado con el cual se realizó el pronóstico de las ventas del año 2020.

Tabla XVIII. **Pronósticos 2019 y 2021**

Mes	Pronóstico 2019	Pronóstico 2020
1	Q72,622,200.72	Q 76,264,377.27
2	Q66,465,484.44	Q 70,107,661.00
3	Q77,490,498.64	Q 81,132,675.20
4	Q71,559,246.26	Q 75,201,422.82
5	Q74,503,305.75	Q 78,145,482.30
6	Q74,695,410.57	Q 78,337,587.12
7	Q80,066,086.36	Q 83,708,262.92
8	Q74,727,672.13	Q 78,369,848.68
9	Q74,221,723.32	Q 77,863,899.88
10	Q76,309,148.81	Q 79,951,325.37
11	Q77,526,523.14	Q 81,168,699.70
12	Q95,276,085.00	Q 98,918,261.55

Fuente: elaboración propia.

Figura 10. **Gráfica de pronóstico de ventas Holt-Winter aditivo**



Fuente: elaboración propia, realizado con Tableau.

3.2. Variables categóricas en el comportamiento de ventas

Para el desarrollo del objetivo dos, el cual consistía en identificar las variables categóricas que afectan el comportamiento de las ventas, se realizó un análisis grafico en el cual se observó que efectivamente en los días festivos o de descanso, el comportamiento de las ventas cambia significativamente, lo cual provoca estacionalidad en la serie de tiempo. Pero no se identificaron valores

atípicos que provocaron que el comportamiento de los datos fluctuara de manera significativa.

El comportamiento que ocasionan estos días de descanso o festivos se logró observar en la figura 12, en la cual se analizó la gráfica del método aditivo de Holt-Winter, y se determinó que este método se adaptó a los ciclos de las ventas anuales.

Cabe mencionar que este trabajo de investigación no considera fenómenos naturales que puedan ocasionar un comportamiento inusual en las ventas, por ejemplo, huracanes, pandemias, terremotos, crímenes, entre otros.

3.3. Ciclos más periódicos de ventas

Para determinar los ciclos más periódicos de ventas en el año, el cual es el objetivo número tres de este trabajo, se utilizaron los resultados obtenidos en la tabla XV, y se determinó que los periodos donde existe mayor venta en la cadena de restaurantes son:

- Periodo de semana santa (abril)
- Fiestas de fin de año (diciembre)
- Inicio de segundo semestre del año (julio) debido al pago del incentivo bono 14.

3.4. Determinar el índice estacional óptimo por medio del análisis de series de tiempo, para proyectar las ventas a 4 años o menos

Para el cuarto y último objetivo de esta investigación, se identificaron los índices estacionales, utilizando el método de promedios móviles. Para el cálculo

se utilizó k=12, por que las ventas se encuentran a un nivel de granularidad mensual. Con esto, se perdieron k-1 valores al momento de realizar el cálculo. Se procedió a calcular el promedio móvil y el promedio móvil centrado. Por último, se calculó el índice estacional, dividiendo las ventas mensuales entre el valor del promedio móvil centrado. Los resultados fueron los siguientes.

Tabla XIX. **Índices estacionales de la serie de tiempo**

Mes	2015	2016	2017	2018
Enero		0.98096956	0.95855378	0.94925687
Febrero	0.83617939	0.85620452	0.84261783	0.87195493
Marzo	0.97621022	0.9901749	0.9393818	
Abril	0.92087573	0.90194007	0.98923956	
Mayo	1.02394533	0.97379279	0.93319958	
Junio	0.98777251	0.9509632	0.94559039	
Julio	1.02758027	1.09011419	1.04811734	
Agosto	0.97046302	0.95504644	0.9635998	
Septiembre	0.91517802	0.93639744	0.94808457	
Octubre	0.98434584	0.99837543	0.9905317	
Noviembre	1.01268866	0.96592708	0.97508079	
Diciembre	1.20658162	1.20441898	1.21382308	

Fuente: elaboración propia.

Para optimizar los índices estacionales calculados anteriormente, se calculó el promedio de cada fila de la tabla XV, de acuerdo con la teoría, la suma de estos promedios debería ser 12, debido a que anteriormente se eligió un valor de k=12; el cual se utilizó como factor de corrección. El resultado que se obtuvo fue de 11.74.

$$Factor\ de\ Corrección = \frac{12}{Total\ de\ las\ medias} = 1.01742$$

Este factor se multiplicó por cada índice mensual, y se obtuvieron valores porcentuales el cual indica que, las ventas para los meses de julio, octubre,

noviembre y diciembre están arriba del mes típico, mientras que los demás, están por debajo.

Tabla XX. **Índices estacionales corregidos**

Mes	Promedio	Factor	Índice corregido
Enero	0.9629267	1.02	98.0 %
Febrero	0.8517392	1.02	86.7 %
Marzo	0.968589	1.02	98.5 %
Abril	0.9373518	1.02	95.4 %
Mayo	0.9769792	1.02	99.4 %
Junio	0.961442	1.02	97.8 %
Julio	1.0552706	1.02	107.4 %
Agosto	0.9630364	1.02	98.0 %
Septiembre	0.93322	1.02	94.9 %
Octubre	0.9910843	1.02	100.8 %
Noviembre	0.9845655	1.02	100.2 %
Diciembre	1.2082746	1.02	122.9 %

Fuente: elaboración propia.

De acuerdo con los pasos desarrollados en este capítulo, se considera que el objetivo general de este trabajo ha sido alcanzado, debido a que se ha logrado establecer un modelo estadístico, utilizando herramientas de series de tiempo, con el cual se ha predicho las ventas que se obtendrán para el año 2019 y 2020.

$$Y_{t+t} = (1^{-10}_t + 0.129_t * 303,514.71) + 0.5876_{t+T}$$

4. DISCUSIÓN DE RESULTADOS

4.1 Modelo lineal de mejor ajuste a la tendencia de ventas

Para el cálculo del modelo lineal de mejor ajuste se utilizó la base de datos de ventas que comprenden del periodo 2015 -2018. Las cuales se registraron en los restaurantes que operaron en la ciudad de Guatemala. Precizando con el método de mejor ajuste a la tendencia de ventas, este se obtuvo desarrollando seis métodos de pronósticos, móvil simple, móvil doble, suavizamiento exponencial, regresión lineal, el método de Holt, y por el ultimo el método Holt-Winter.

Gonzales (2018), en su estudio para la predicción de la demanda de ventas de bicicletas, probó seis métodos de pronósticos, los cuales fueron, móvil simple, doble, suavización exponencial, doble, Holt y Holt-Winter.

En primer lugar, en el cálculo de la autocorrelación para la serie temporal de ventas en la cadena de restaurante, se logró identificar distintos patrones la serie temporal de datos y se determinó que la serie presenta una correlación alta, siendo casi 0.7, esto indicó que la serie tiene estacionariedad. En otras palabras, las ventas tienen comportamientos cíclicos dependiendo de la época de año.

Para comprobar si el conjunto de datos es estadísticamente estacional, se utilizó la prueba Dickey-Fuller, con la que se obtuvo que el estadístico de prueba fue más pequeño que el valor de significancia, por lo que se rechaza la hipótesis nula, que indica que la serie no posee un comportamiento no estacionario.

Con la descomposición de la serie temporal de datos que se mostró en la figura 3, se identificó que la serie temporal presentó una tendencia positiva, es decir que las ventas tienen un comportamiento creciente a lo largo de un tiempo determinado.

Roque (2016), realizó una descomposición de serie de temporal, para el desarrollo de un modelo de predicción, en la cual logró identificar la estacionalidad de la serie. Con la tendencia y estacionalidad identificadas en la serie temporal, se determinó que el mejor método para pronosticar las ventas es el de Holt-Winter, el cual presentó un valor de 65,195 de error cuadrático medio, como se detalla en la tabla I. Pero cabe mencionar que se utilizaron los otros 5 métodos para validar, por medio del error cuadrático medio, que efectivamente el método de Holt-Winter es el que presenta menor error. El estudio realizado por Gonzales (2018), demostró que el modelo de predictivo óptimo de ventas de bicicletas, fue el de Holt-Winter, el cual mostro un error cuadrático medio de 2,393.31.

Se identificó que el error cuadrático medio calculado por este autor es menor, debido a que el pronóstico que realizó corresponde solo a un tipo de articulo en específico; en este estudio se tomaron todos los menús fueron vendidos por los restaurantes. Es por ese motivo que existe una variación significativa entre los errores de los métodos Holt-Winter. El método de promedio móvil doble presentó un error de 1.38×10^{13} , siendo este un valor muy alejado de los datos de la serie temporal de ventas en los restaurantes. Gráficamente este método no se ajustó a las variaciones de la serie temporal. Este autor obtuvo un valor de error cuadrático medio de 11,524 para el método de media móvil doble. Siendo este su sexto mejor método de pronóstico. En este trabajo el promedio móvil doble fue el quinto mejor método.

El segundo método que se desarrolló fue el móvil simple, el cual presentó un error cuadrático medio de 4.01×10^{12} , comparado con resultado del método anterior, este es 29 % menor. Al trazar la gráfica de este método, por medio de un análisis visual se logró observar que la curva empezó a ajustarse a la tendencia, pero no a las fluctuaciones de la serie temporal.

Pérez, Cifuentes, Vásquez y Ocampo (2013) en su estudio para un modelo de gestión de inventarios, obtuvieron un valor de 2974.82 de error cuadrático medio. El cual fue calculado con una muestra de $n=12$, y determinaron que no era el método óptimo. “Estos primeros dos métodos se aplican cuando las series de tiempo tienen componente estacional, es decir, cuando los valores de la serie oscilan alrededor de la media” (Anderson, Sweeney y Williams, 2008, p. 324).

El tercer método utilizado fue el método de suavizamiento exponencial doble, el cual obtuvo un error cuadrático medio 163,194.65 este valor es considerablemente menor respecto a los dos métodos anteriores. En la gráfica de este método se observó que la curva se ajusta a las fluctuaciones de la serie temporal, pero de una manera desfazada a la derecha.

Menacho (2013) en su estudio sobre la comparación de series de tiempo y redes neuronales mencionó que el método de suavizamiento exponencial doble con dos constantes de suavización proporcionó un valor de error en las series de tiempo analizadas, mayor al que proporcionó el modelo de redes neuronales. El cuarto método implementado fue el de Holt o de dos parámetros, este método presentó un error cuadrático medio de 93,830.95, respecto al método anterior este se redujo en un 43 %, y gráficamente se observó que la curva se ajusta a la tendencia, pero no a las fluctuaciones de la serie temporal.

Gonzales (2018) identificó un valor cuadrático medio de 4,910 con el método Holt; el cual fue su segundo método con menor error. Al igual que en este estudio, el método Holt, es el segundo mejor con menor error.

El quinto método utilizado fue una regresión lineal simple, este método mostró un valor cuadrático medio de 86,344.57, respecto al método anterior, este es 7.79 % menor. En la gráfica de este método, la tendencia se ajusta al conjunto de la serie de datos, pero no así las fluctuaciones del conjunto de datos de la serie temporal. El sexto y último método fue el de Holt-Winter, este método presentó un valor para el error cuadrático medio de 65,195.23, siendo un 24 % menor al valor del método anterior. En la gráfica de este método se observó que la curva se ajusta bastante a la tendencia y estacionalidad de la serie temporal.

Este fue el mejor método por presentar el menor error. Esto indica que este método es el más cercano al conjunto de datos de las ventas que se realizaron del 2015 al 2018 en la cadena de restaurantes. Gonzales (2018) en su estudio, menciona que, “dentro de los métodos analizados, el mejor es Holt-Winter, debido a que presentó un error cuadrático medio bajo, respecto a los demás. Esto se debe a que este método captura la estacionalidad de la serie y realiza una extensión del método de suavización exponencial” (p. 84).

Se implementó el método Holt-Winter Aditivo debido a que la magnitud de las fluctuaciones estacionales de la serie no cambia significativamente al hacerlo la tendencia. Para esto se optimizaron los parámetros nivel, tendencia y estacionalidad, como lo sugiere Pérez (2017) en su estudio, para identificar un mejor modelo de pronósticos. El error cuadrático medio de este método optimizado fue de 54,139, siendo 17 % menor respecto al método de Holt-Winter.

Con este método se pronosticó el valor de las ventas que obtendrán los restaurantes para los años 2019 y 2020. Los cuales se observaron en la tabla XIV del capítulo anterior.

Este método fue el que menor error cuadrático medio presentó convirtiéndolo en la mejor opción para implementarlo en la cadena de restaurantes debido a que se adapta bastante al comportamiento de la serie de tiempo. Su implementación es sencilla ya que no requiere de un analista experto ni software sofisticado.

De acuerdo con Vélez (2019), el error cuadrático medio permite calcular la diferencia entre el estimador y su valor real, este criterio se aplicó a los 6 métodos de predicción. El método Holt-Winter presentó el menor error, y fue seleccionado para optimizar sus parámetros.

4.2 Variables categóricas que afectan el comportamiento de ventas

Al graficar el conjunto de datos del histórico de ventas para la cadena de restaurantes, se logró observar que el comportamiento no presentó valores atípicos o faltantes, es decir que durante el periodo 2015-2018 no existieron acontecimientos que provocaran que la serie temporal sufriera un comportamiento fuera de lo normal o que existirán valores nulos que rompieran la tendencia y estacionalidad de las ventas.

Gonzales (2015) menciona que “los valores atípicos son provocados por factores externos que afectan la distribución normal de los datos” (p.27).

Los acontecimientos naturales como terremotos, lluvias, huracanes, pandemias no fueron considerados dentro de estas variables categóricas.

4.3 Ciclos con más periodicidad de ventas en el año

Los periodos con más volúmenes de venta en un año normal para la cadena de restaurantes se determinaron a través de los resultados que se obtuvieron en la tabla XIV y gráficamente de la figura 12. Los cuales son: semana santa, la cual puede ser en abril, las fiestas de fin de año en diciembre, e inicios del segundo semestre del año, mes de julio, debido al pago del incentivo bono 14.

Enríquez (2009), mencionó que, “con un análisis técnico visual, fue posible identificar periodos donde las ventas incrementaban y en donde mantenían una estabilidad para el sector automotriz” (p. 76).

Esto permite que la cadena de restaurantes se prepare con un dato de ventas más puntual para atender a la demanda que acontecerá en estos periodos mencionados del año mencionados.

4.4 Índices estacionales de la serie de tiempo

De acuerdo con Levin y Rubin (2004) “los índices estacionales fueron calculados por medio de la descomposición de medias móviles, en la cual se procedió a calcular el promedio móvil de orden k , siendo este 12, porque la serie de tiempo se encuentra en periodos estacionales mensuales” (p. 291).

Se identificó por medio de este análisis que los índices estacionales de julio, octubre, noviembre y diciembre se encuentran por arriba de la unidad, por lo que la cantidad de ventas promedio en esos meses se encuentra por encima de la tendencia estimada.

De acuerdo con los resultados observados en la tabla XVI, se determinó que el índice estacional más alto se encuentra en diciembre, y el índice estacional más bajo se encuentra en febrero. En el mes de julio el volumen de ventas promedio es 22.9 % más alto que la tendencia estimada, mientras que en el mes de febrero el volumen de ventas es 13.3 % ($1 - 0.867$) más baja que la tendencia estimada.

Esto le da un panorama a la cadena de restaurantes para generar nuevas estrategias en estos meses donde las ventas adquieren valores más altos y bajos respecto al promedio.

Se sugiere que cada año se prueben los métodos mencionados en este estudio, para identificar con qué método se ajusta más a la tendencia del año anterior.

CONCLUSIONES

1. El modelo de pronóstico estadístico con mejor ajuste para el volumen de venta mensual para los años 2019 y 2020, fue Holt-Winter aditivo, el cual, por su tendencia y estacionalidad, indica que el valor de las ventas aumentará en un 5.1 % y 9.8 % respectivamente, tomando como año de inicio de referencia 2018. Este pronóstico presentó el menor error cuadrático medio, con el cual fue posible generar pronósticos mensuales de ventas.
2. Según el análisis gráfico de la serie temporal, probablemente no existen variables categóricas que produjeran valores atípicos o faltantes en la serie de datos, el comportamiento presentó una tendencia incremental con estacionalidad en los periodos donde las ventas son más significativas.
3. La variación periódica inferior o igual a un año de las ventas, corresponde a marzo, abril, julio y diciembre, es decir que en estos meses las ventas son más significativas y que son más predecibles. Periodo de semana santa (abril), inicio de segundo semestre del año (julio) debido al pago del incentivo bono 14 y fiestas de fin de año (diciembre).
4. Según el factor de corrección, los meses que están por arriba de las ventas promedio son julio, octubre, noviembre y diciembre. El incremento porcentual más alto corresponde a julio, el cual presentó un valor 22.9 %. Los meses de enero, febrero marzo, abril, mayo, junio, agosto y septiembre se encuentran debajo de las ventas promedio siendo febrero el que presentó el valor más bajo con un 13.3 %.

RECOMENDACIONES

1. Para analizar y modelar un conjunto de datos de una serie de tiempo utilizando el modelo aditivo de Holt-Winter, se deben considerar aspectos como: histórico de ventas de al menos 4 años, valores atípicos, estacionalidad y tendencia.
2. Probar los métodos utilizados en este trabajo, con el fin de identificar cuál presenta el menor error cuadrático medio.
3. Utilizar los datos pronosticados para el diseño del plan estratégico de ventas, el cual permitirá saber cuántos productos necesitará producir la cadena de restaurantes, cuánto material de consumo requerirá, cuánto es el costo de almacenamiento, y el personal adecuado para cada una de las funciones correspondientes.
4. La metodología Box-Jenkins (ARIMA) es una herramienta más sofisticada y puede considerarse para análisis posteriores que se consideren complejos y precisos.

REFERENCIAS

1. Alvarado Guzmán, J. (2016) *Modelos de pronósticos basados en valores singulares de la matriz Hankel para la captura anchovetas y sardinas en el área norte de Chile*. (Tesis de maestría). Pontificia Universidad Católica de Valparaíso, Chile. Recuperado de http://opac.pucv.cl/pucv_txt/txt-6500/UCD6589_01.pdf
2. Anderson, D. R.; Sweeney, D. J y Williams, T. A. (2008). *Estadística para administración y economía*. México, D.F: Cengage Learning Editores, S.A. Recuperado de <https://www.upg.mx/wp-content/uploads/2015/10/LIBRO-13-Estadistica-para-administracion-y-economia.pdf>
3. Arango Marin, J. A.; Giraldo García, J. A., y Castrillón Gómez, O. D. (diciembre 2013). Gestión de compras e inventarios a partir de pronósticos Holt-Winters y diferenciación de nivel de servicio por clasificación ABC. *Scientia et Technica Año XVIII*, 18(4), 743-747. Recuperado de <https://www.redalyc.org/pdf/849/84929984023.pdf>
4. Arellano, M. (2001). 5campus: Introducción al Análisis Clásico de Series de Tiempo. Chile: Valparaíso. Recuperado de <http://www.5campus.com/>
5. Botero, S. y Cano, J. A. (agosto 2008). Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia. *Cuadernos de economía*, 27(48), 173-207. Recuperado de

https://www.researchgate.net/publication/28236638_Analisis_de_series_de_tiempo_para_la_prediccion_de_los_precios_de_la_energia_en_la_bolsa_de_Colombia

6. Cerreño, R. I. (2012). *Evaluación y mejora de los pronósticos de venta en una empresa comercializadora de maquinaria pesada utilizando la metodología lean six sigma*. (tesis de maestría). Universidad Nacional de Ingeniería, Perú. Recuperado de https://www.lareferencia.info/vufind/Record/PE_3502c9b7414d289ea4f8e81a7069eb8a.
7. Frausto Enríquez, J. H. (2009). *Pronósticos de ventas para la administración estratégica de los recursos en una empresa del sector automotriz*. (Tesis de maestría). Instituto tecnológico y de estudios superiores de Monterrey, México. Recuperado de <https://repositorio.tec.mx/handle/11285/569469>.
8. Frías Miranda E. V. (2006). *Aplicación de las series de tiempo al pronóstico de la demanda en la empresa de manufactura moderna*. (Tesis de pregrado). Instituto Politécnico Nacional. México. Recuperado de <https://www.repositoriodigital.ipn.mx/handle/123456789/5864>
9. Gil Vera, V. D. (febrero 2016). Pronóstico de la demanda mensual de electricidad con series de tiempo. *Revista EIA*, 13(26), 111-120. Recuperado de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-12372016000200009.

10. González Ramírez, F. (2018). *Construcción de un modelo para pronosticar la demanda de bicicletas de la empresa especializada Colombia, usando series de tiempo*. (Tesis de Maestría). Universidad Nacional de Colombia, Colombia. Recuperado de <https://repositorio.unal.edu.co/bitstream/handle/unal/69716/1053803034.2019.pdf?sequence=1&isAllowed=y>.
11. González Vargas, C. A. (2015) *Pronóstico de demanda de refacciones de automóvil, mediante series de tiempo, redes neuronales artificiales y modelos híbridos*. (Tesis de maestría). Universidad Autónoma de México, México. Recuperado de <http://www.ptolomeo.unam.mx:8080/xmlui/handle/132.248.52.100/7988>.
12. Hanke, J. E., y Wichern, D. W. (2006). *Pronósticos en los negocios*. México: Person Educacion. Recuperado de https://www.academia.edu/43174176/Pron%C3%B3sticos_en_los_negocios_9na_Edici%C3%B3n_John_E_Hanke_FREELIBROS_ORG.
13. Madrigal Espinoza, S. D. (2011). *Pronósticos de series temporales con estacionalidad*. (Tesis de doctorado). Universidad Autónoma de Nuevo León, México. Recuperado de <http://eprints.uanl.mx/2546/1/1080090514.pdf>.
14. Mason, R. D.; Lind, D. A.; Marchal, W. G.; Hernández Lozano, M. C. y Paniagua Bocanegra, F. (2002). *Estadística para Administración y Economía*. Bogotá: UNAPEC. Recuperado de

<https://catalogobiblioteca.unapec.edu.do/cgi-bin/koha/opac-detail.pl?biblionumber=7965>

15. Melo Mayta, E. (2016) *Modelo de predicción mensual de mortalidad general intrahospitalaria en el hospital regional Manuel Núñez Butrón-Puno, 2008-2016-I*. (Tesis de pregrado). Universidad Nacional del Altiplano –Puno, Perú. Recuperado de <https://repositorioslatinoamericanos.uchile.cl/handle/2250/3274457>
16. Pérez Molina, L. R, (2017). *Procedimiento para el pronóstico de ventas de la empresa de Materiales de construcción de las tunas*. (Tesis de maestría). Universidad de las Tunas, República de Cuba. Recuperado de <http://roa.ult.edu.cu/handle/123456789/4252>.
17. Richard, I.L, David, S.R, (2004). *Estadística para administración y economía*. Ciudad de México, México: Person Educacion. Recuperado de <https://profefily.com/wp-content/uploads/2017/12/Estad%C3%ADstica-para-administraci%C3%B3n-y-economia-Richard-I.-Levin.pdf>.
18. Roque Montalvo, I. L. (2016) *Análisis comparativo de técnicas de minería de datos para la predicción de ventas* (tesis de pregrado). Universidad Señor de Sipán, Perú. Recuperado de https://repositorio.uss.edu.pe/bitstream/handle/20.500.12802/2695/Tesis_Roque_Montalvo_Irene.pdf?sequence=1&isAllowed=y.
19. Sánchez, P. A. (abril 2008). Cambios estructurales en series de tiempo: una revisión al estado del arte. *Revista Ingenierías Universidad de*

Medellín, 7(12), 115-140. Recuperado de <http://www.scielo.org.co/pdf/rrium/v7n12/v7n12a07.pdf>.

20. Stanton W.; Etzel M. y Walker B. (2004). *Fundamentos de Marketing*. (13ª Edición). Ciudad de México, México: Mc Graw-Hill-Interamericana. Recuperado de <https://mercadeo1marthasandino.files.wordpress.com/2015/02/fundamentos-de-marketing-stanton-14edi.pdf>.
21. Tusell, F. (2011). *Análisis de Regresión. Introducción Teórica y Práctica basada en R*. Bilbao, España: Person Educación. Recuperado de <https://www.uv.es/lejarza/mcaf/materialR/nreg1.pdf>.
22. Vélez Montoya, D. (2019) *Una adaptación del procedimiento bootstrap en la estimación del error cuadrático medio en áreas pequeñas con aplicación a datos colombianos*. (Tesis de maestría). Universidad Nacional de Colombia, Colombia. Recuperado de <https://repositorio.unal.edu.co/bitstream/handle/unal/75926/1039453675.2019.pdf?sequence=1&isAllowed=y>.
23. Verdoy, P. J.; Beltrán, M. J. y Peris, M. J. (2015). *Problemas resueltos de estadística aplicada a las ciencias sociales*. Castellón de la Plana, España: Publicacions de la Universitat Jaume I. Recuperado de <https://core.ac.uk/download/61451006.pdf>.