



Tecnológico de Monterrey

Análisis de Series de Tiempo
Reporte Final Venta de Televisores

Módulo 5: Estadística Avanzada para ciencia de datos

Inteligencia Artificial Avanzada para la Ciencia de Datos.
Grupo 501

Autora

Emilia Victoria Jácome Iñiguez
A00828347

Docente:

Blanca Rosa Ruiz Hernandez

Tecnológico de Monterrey, Campus Monterrey
miércoles, 30 de Noviembre de 2022

Contents

Resumen	2
Introducción	2
Pregunta Base	2
1. Gráfico de dispersión para verificación de la tendencia y los ciclos	3
2. Análisis de tendencia y estacionalidad	3
2.1. Descomposición de la serie en sus 3 componentes	3
3. Análisis del modelo lineal de la tendencia	4
3.1. Regresión lineal de la tendencia	4
3.2. Graficación de la recta junto con las ventas desestacionalizadas	4
4. Análisis de la pertinencia del modelo lineal.	5
4.1. Cálculo de la significancia de Beta 1	5
4.2. Cálculo de la Variabilidad explicada por el modelo	5
4.3. Análisis de los residuos	5
4.4. Aplicación de Prueba de normalidad - Shapiro-Wilk	6
5. Cálculo de errores de la predicción de la serie de tiempo.	6
5.1. CME (Error Medio)	6
5.2. EPAM (Error Medio Porcentual)	6
6. Gráfico de las ventas con predicciones vs. el tiempo	6
7. Conclusión del modelo	6
8. Cálculo del Pronóstico de Ventas de Televisores para el siguiente año	7
8.1. Gráfico del pronóstico	7
CONCLUSIÓN	8
ANEXOS	9
Liga a Repositorio de GitHub:	9

Resumen

Las ventas de televisores tienen un comportamiento fluctuante en el tiempo ya que dependen de varios factores o variables externas tales como estacionalidad, los eventos televisivos, los descuentos y promociones, etc. Por lo que son difíciles de predecir sin un estudio o análisis estadístico detrás. Se han obtenido las cifras de los ingresos por ventas trimestrales de una empresa de televisores pertenecientes a sus últimos 4 años de operación tal y cómo se muestra a continuación:

Año	1				2				3				4			
Trimestre	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Ventas (miles de \$)	4.8	4.1	6.0	6.5	5.8	5.2	6.8	7.4	6.0	5.6	7.5	7.8	6.3	5.9	8.0	8.4

A lo largo de este documento, se pretende realizar un análisis de series de tiempo para identificar una tendencia clara en los ingresos por ventas de los televisores durante los últimos 4 años con la finalidad de elaborar un pronóstico de ventas para el siguiente año y con ello poder predecir cuáles serán las ventas futuras a nivel trimestral y dar una aproximación basada en el histórico que se tiene recopilado para que la empresa pueda calcular la próxima oferta y de ese modo, asegurarse de tener suficiente abasto para cubrir la demanda de Televisores con suministros, personal y puntos de venta.

Subsecuentemente, para poder abordar esta problemática, se hizo uso de varias herramientas estadísticas vistas durante el módulo 5 de la concentración Inteligencia Artificial Avanzada para la Ciencia de Datos. Principalmente, se aplicaron varias pruebas de normalidad multivariante a los datos en conjunto y un test de normalidad a cada variable por separado. Por otro lado, se efectuó un análisis de residuos para corroborar el comportamiento de normalidad que tiene la variable. También se procedió a aplicar un análisis de series de tiempo de tipo no-estacional a través de promedios móviles donde se calcularon los errores promedio para poder evaluar la calidad de las aproximaciones y subsecuentemente poder construir el pronóstico para el próximo año.

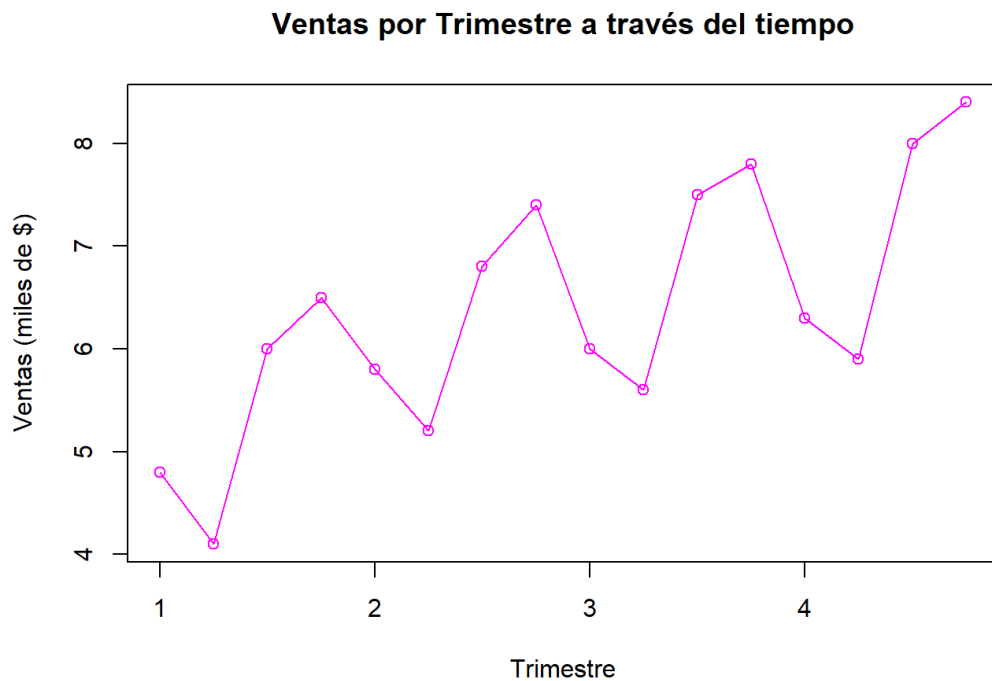
Introducción

El objetivo de este documento es analizar los pronósticos en series de tiempo no estacionarias al utilizar los datos de las ventas de televisores para familiarizarte con el análisis de tendencia de una serie de tiempo utilizando las herramientas estadísticas previamente descritas. La importancia del siguiente análisis radica en la posibilidad de las empresas en poder anticiparse ante acontecimientos futuros y de ese modo, tener mayor preparación al momento de abordar sus ventas futuras y poder diseñar estrategias que les permitan alcanzar sus objetivos de ventas anuales, trimestrales o inclusive mensuales. En efecto, los pronósticos de series de tiempo tienen un amplio panorama de aplicaciones en donde se busca aprovechar de los patrones y tendencias que se encuentran en los datos para construir un modelo futuro que permita dar visibilidad de los acontecimientos futuros y poder tomar decisiones anticipadas con respecto a los suministros, proveedores y demás factores que afectan directamente las ventas, en este caso, de televisores.

Pregunta Base

¿Cuáles serán las ventas trimestrales del próximo año para la empresa a causa de los ingresos obtenidos a través de la venta de televisores?

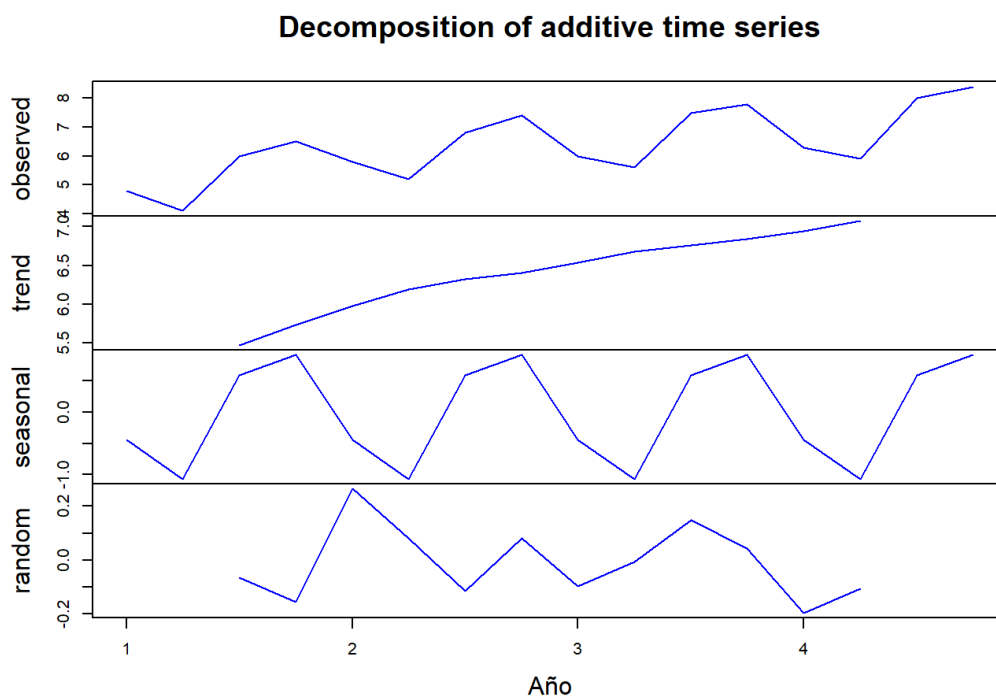
1. Gráfico de dispersión para verificación de la tendencia y los ciclos



En la gráfica de Ventas por Trimestre a través del tiempo, se puede visualizar cómo las ventas siguen un patrón bien definido que se repite cada año, o cada cuatro trimestres, pero también se puede apreciar un aumento en la magnitud de estas.

2. Análisis de tendencia y estacionalidad

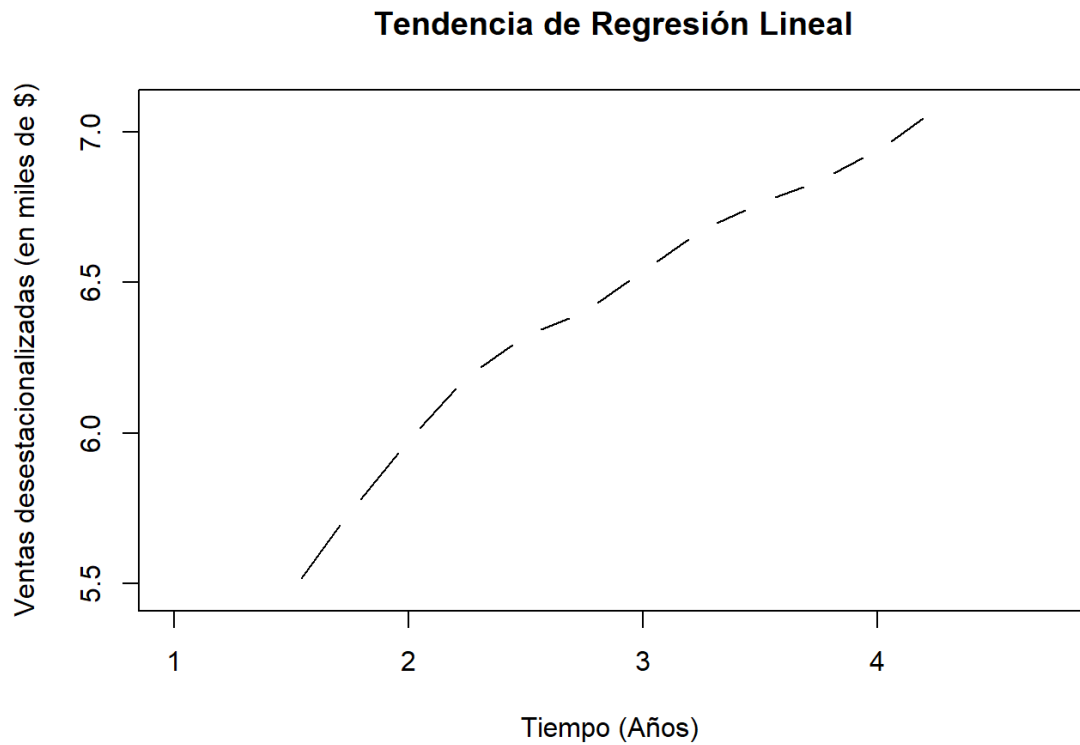
2.1. Descomposición de la serie en sus 3 componentes



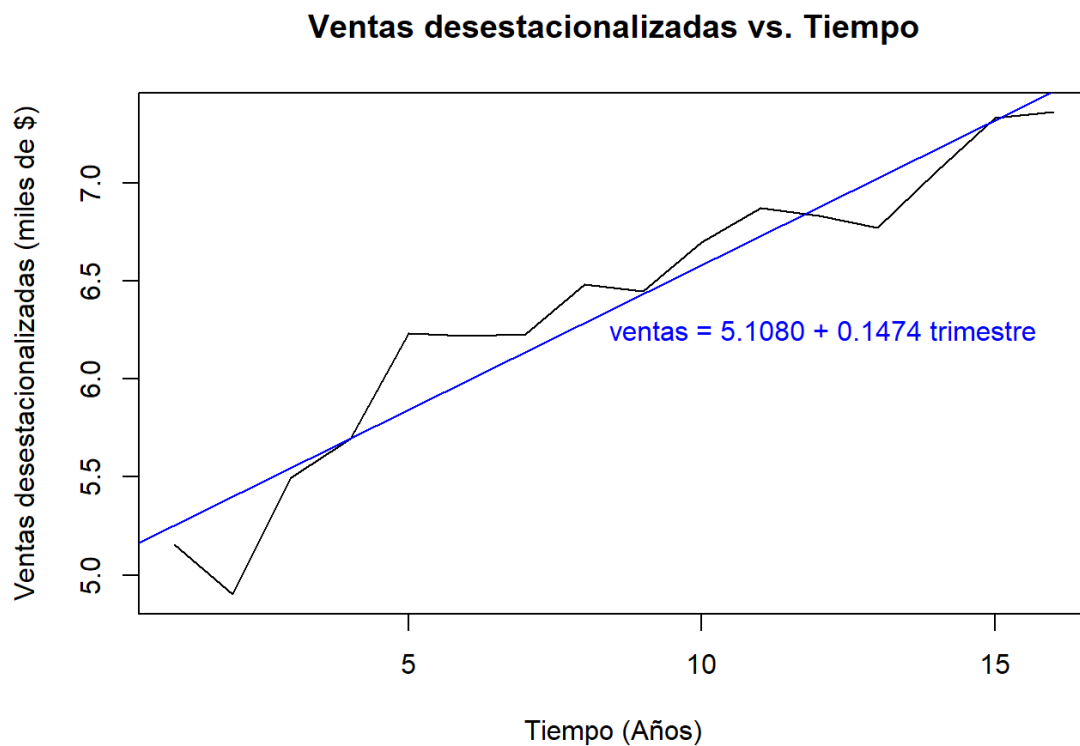
3. Análisis del modelo lineal de la tendencia

El modelo de la tendencia de ventas de televisores muestra un comportamiento de linealidad positiva.

3.1. Regresión lineal de la tendencia



3.2 Graficación de la recta junto con las ventas desestacionalizadas



4. Análisis de la pertinencia del modelo lineal.

Residuals:

Min	1Q	Median	3Q	Max
-0.5007	-0.1001	0.0037	0.1207	0.3872

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	5.10804	0.11171	45.73	< 2e-16 ***
x3	0.14738	0.01155	12.76	4.25e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.213 on 14 degrees of freedom

Multiple R-squared: 0.9208, Adjusted R-squared: 0.9151

F-statistic: 162.7 on 1 and 14 DF, p-value: 4.248e-09

Se puede ver que el p-value es menor a 0.05, por lo tanto se rechaza la hipótesis nula y se determina que Beta 1 si es significativo. Esto significa consecuentemente que el modelo se ajusta casi perfectamente a los datos reales ya que logra explicar aproximadamente un 92% (valor de R cuadrado) de la variabilidad de estos.

4.1. Cálculo de la significancia de Beta 1

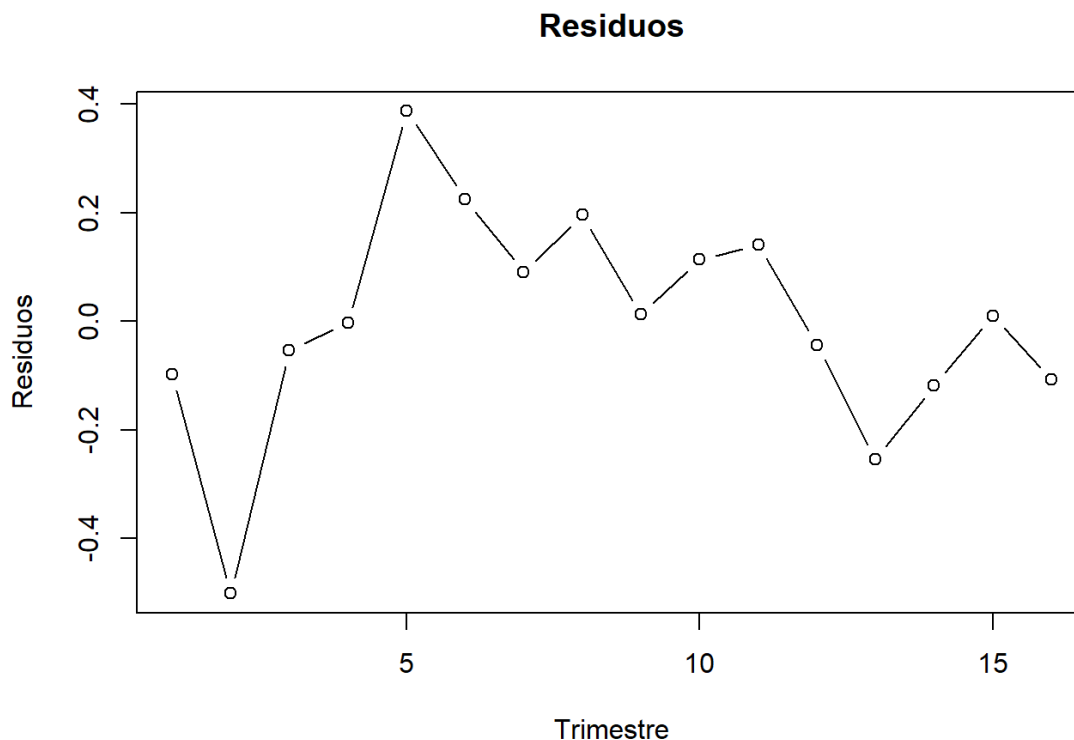
4.247717e-09

4.2. Cálculo de la Variabilidad explicada por el modelo

0.9207911

La variabilidad del modelo es de 92.08%, es decir, el modelo logra explicar 92.08% de los datos reales.

4.3. Análisis de los residuos



A partir de la gráfica superior, se puede denotar claramente que los residuos no tienen una tendencia bien definida por lo que se podrá afirmar que tienen una distribución aleatoria.

4.4. Aplicación de Prueba de normalidad - Shapiro-Wilk

Shapiro-Wilk normality test
 $W = 0.96379$, $p\text{-value} = 0.7307$

Dado que el valor-p es notablemente mayor a α (0.05), no es posible rechazar la hipótesis nula, por lo que se puede determinar que los residuos siguen una distribución normal.

5. Cálculo de errores de la predicción de la serie de tiempo.

5.1. CME (Error Medio)

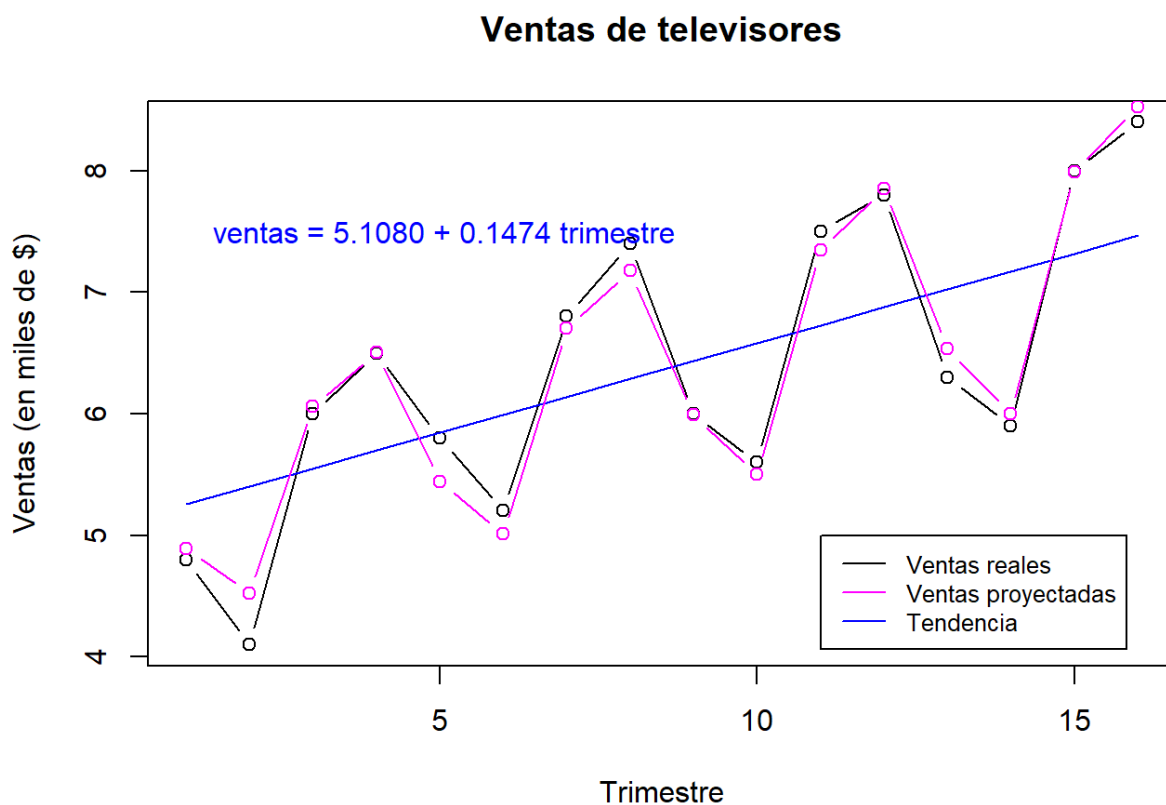
0.6957218

5.2. EPAM (Error Medio Porcentual)

0.125894

A partir del cálculo del error medio porcentual, se puede interpretar que este valor está por debajo del 20%, por ende, el nivel de error del modelo no es muy elevado.

6. Gráfico de las ventas con predicciones vs. el tiempo



7. Conclusión del modelo

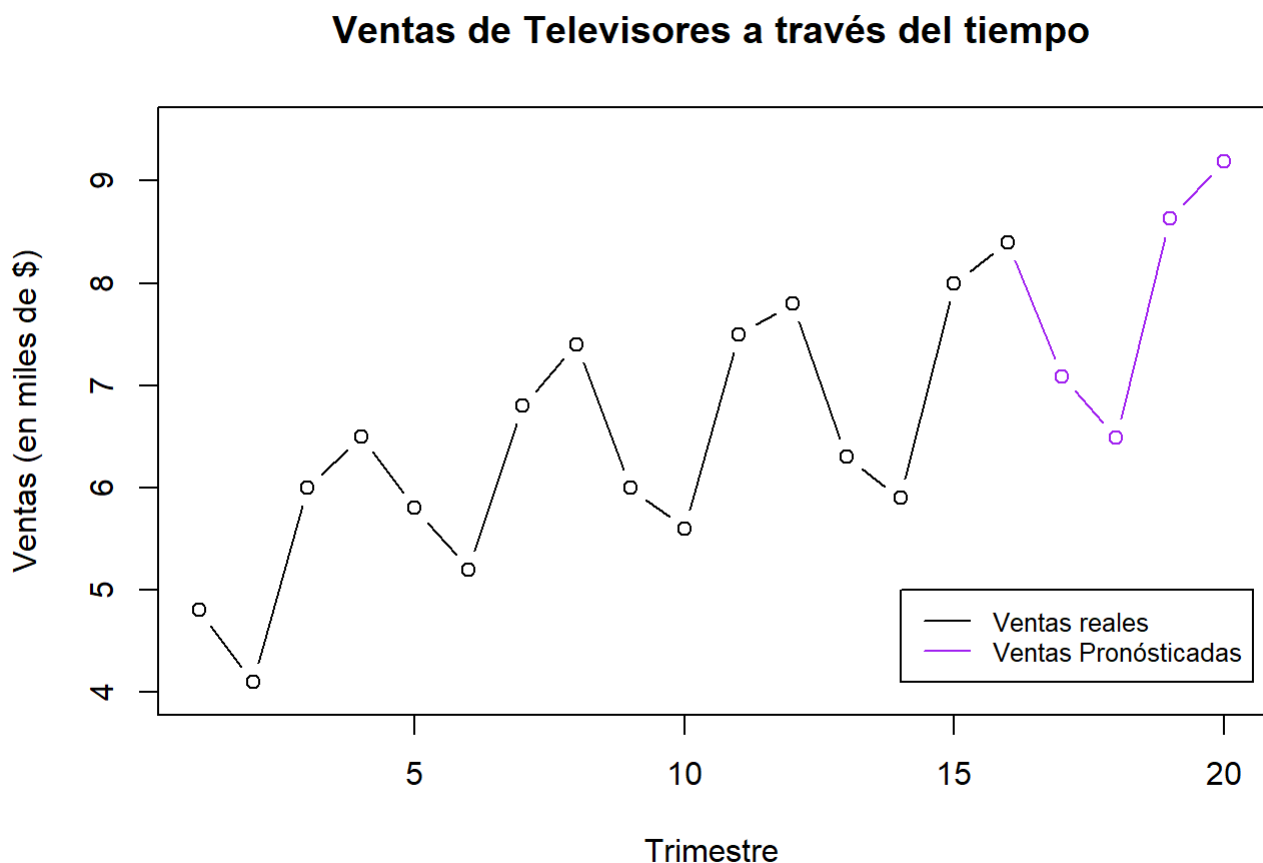
El modelado propuesto se ajusta casi perfectamente a los datos reales dado que logra explicar más del 90% de la variabilidad de los datos reales. Adicionalmente, se ha demostrado que el valor-p del test de Shapiro es superior al valor de confianza por ende, se procedió a rechazar la hipótesis nula, comprobando normalidad en los datos, lo que implica subsecuentemente que $B1 = 0$. De modo que la pendiente del modelo de regresión lineal es significativa. Por otra parte, se comprobó que los residuos carecen de una tendencia definida, es decir, tienen un comportamiento aleatorio, por ende se trata de una distribución normal en los valores. En lo que respecta a los

errores, $CME = 0.70$ y $EPAM = 0.13$, estos representan un margen de error insignificativo para considerar una alta imprecisión en el modelo.

8. Cálculo del Pronóstico de Ventas de Televisores para el siguiente año

Trimestre	1	2	3	4
Ventas Pronósticadas (en miles de \$)	7.085872	6.491284	8.632585	9.195263

8.1. Gráfico del pronóstico



En el gráfico superior se puede ver el pronóstico de ventas por ingresos de televisores de la empresa en cuestión. Es posible apreciar claramente un patrón de estacionalidad que se repite cada 4 trimestres (anualmente) donde los ingresos por ventas disminuyen los dos primeros trimestres del año y aumentan en los dos últimos, siendo el tercer trimestre dónde se produce el mayor incremento en ventas y el cuarto dónde se alcanza el mayor nivel de ventas. Por ende, en las ventas pronósticadas (sección de color morado), se puede ver que en efecto, se vuelve a repetir el mismo patrón de estacionalidad que en los periodos pasados pero con una magnitud mayor de acuerdo a la tendencia de crecimiento positivo que se calculó en los incisos pasados.

CONCLUSIÓN

En síntesis, las ventas trimestrales pronosticadas para el próximo año de la empresa en cuestión obtenidas a partir del análisis de series de tiempo no-estacionales son:

Trimestre	1	2	3	4
Ventas Pronósticadas (en miles de \$)	7.085872	6.491284	8.632585	9.195263

Las cuales oscilan en un rango de \$6.5 - \$9.20 miles de dólares con un error porcentual del 12% el cuál refleja una alta confiabilidad en los resultados, en comparación al histórico de ventas que se ha recopilado. Para un análisis más exhaustivo, se sugiere a futuro considerar más variables o factores cuantitativos que afecten las ventas de televisores para dicha empresa con el propósito de examinar el efecto que tiene cada uno de estos factores en la generación de ingresos por ventas de televisores. Para ello, se sugiere emplear un análisis de series de tiempo en conjunto con un análisis de ANOVA, y de ser necesario o en caso de tener varios factores afectando las ventas, también se podría considerar aplicar análisis de componentes principales para simplificar el modelo.

Por otra parte, se recomienda ampliamente tener mayor cantidad de datos históricos de más trimestres atrás para minimizar el error promedio y que los resultados finales del pronóstico sean más cercanos a la realidad. En adición, también se podría recomendar realizar un análisis de ventas con mayor granularidad, por ejemplo, con una estacionalidad semanal o mensual para tener patrones más precisos de acuerdo a ciclos de tiempo más acotados. Esto traerá más complejidad a la construcción del modelo, pero también mayor precisión a la obtención de resultados pronosticados. De esta manera, considero que se podría abordar de mejor manera la necesidad de la empresa en calcular sus ventas proyectadas y tener mayor especificidad en sus proyecciones de suministros para una determinada semana del año donde aparentemente sus ventas tienen mayor apogeo, y de esa manera no perder clientes por falta de productos disponibles a la venta y cubrir dichas ventas de la mejor manera posible a través de la contratación de personal provisional de apoyo durante ese periodo. Por otro lado, estos resultados también le permitirán a la empresa a diseñar estrategias de marketing y ventas para incrementar sus ingresos en aquellos periodos donde las ventas son relativamente bajas y se requiere implementar algún tipo de incentivo para animar la compra de parte de sus clientes.

ANEXOS

Liga a Repositorio de GitHub:

<https://github.com/emilyvic/Artificial-Intelligence/tree/main/Estadistical%20Models/TV%20Sales>