

Universidad Panamericana
Maestría en Ciencia de Datos
Econometría

Actividad Desestacionalización de Series de Tiempo

Enrique Ulises Báez Gómez Tagle
Luis Alejandro Guillén Alvarez

26 de agosto de 2025

Índice

1. Contexto	2
2. Preguntas	2
3. Link al repositorio con código fuente	9

1. Contexto

México es el mayor productor y exportador mundial de aguacate, una fruta que ha ganado popularidad global por su valor nutricional y versatilidad en la cocina. El aguacate, conocido por su riqueza en grasas saludables, vitaminas y minerales, es un componente esencial en dietas equilibradas y ha visto un aumento significativo en su demanda en mercados internacionales.

La industria del aguacate es un pilar fundamental de la economía agrícola mexicana. En las últimas décadas, el aguacate ha pasado de ser un producto regional a convertirse en una de las principales exportaciones agrícolas del país. La producción y exportación de aguacate no solo contribuyen de manera significativa al PIB agrícola, sino que también generan empleo en diversas regiones, particularmente en los estados de Michoacán y Jalisco, que son los principales productores.

A continuación en la tabla no1 se muestran los datos bimestrales de la exportación de aguacate en toneladas:

Cuadro 1: Exportación bimestral de aguacate (toneladas)

Año	Bimestre	t	Toneladas
2019	1	1	224,604.04
2019	2	2	201,125.90
2019	3	3	154,184.06
2019	4	4	140,857.77
2019	5	5	213,070.50
2019	6	6	268,644.20
2020	1	7	316,793.52
2020	2	8	283,659.09
2020	3	9	217,428.69
2020	4	10	198,535.98
2020	5	11	300,467.80
2020	6	12	378,764.26
2021	1	13	410,806.71
2021	2	14	367,955.31
2021	3	15	282,092.18
2021	4	16	257,532.45
2021	5	17	389,706.87
2021	6	18	491,240.80
2022	1	19	405,077.56
2022	2	20	362,765.23
2022	3	21	278,092.64
2022	4	22	253,986.68
2022	5	23	384,156.60
2022	6	24	484,431.36
2023	1	25	426,151.46
2023	2	26	381,653.51
2023	3	27	292,635.95
2023	4	28	267,180.48
2023	5	29	404,208.92
2023	6	30	522,430.98

2. Preguntas

1. Si establecemos como “regla de dedo” que una serie presenta poca variación si el coeficiente de variación es menor a 30 %. En el caso de las exportaciones de aguacate ¿cuánto vale el coeficiente de variación? Exprese su resultado en porcentaje con dos decimales.

Coefficiente de variación (CV) cuantifica la dispersión relativa de una variable:

$$CV = \frac{s}{\bar{x}} \times 100 \%, \quad (1)$$

Obtenemos

$$\bar{x} = 318\,674.72 \text{ toneladas}, \quad (2)$$

$$s = 100\,061.44 \text{ toneladas}. \quad (3)$$

Sustituyendo,

$$CV = \frac{100\,061.44}{318\,674.72} \times 100 \% = \mathbf{31.40 \%}. \quad (4)$$

Usando la regla de que una serie presenta poca variación si $CV < 30 \%$, el valor obtenido de 31.40 % indica que las exportaciones *sí muestran* una variación relativa considerable-

2. Considerando las estadísticas descriptivas de la serie de tiempo de las exportaciones de aguacate ¿qué puede concluir? Incluya los gráficos relevantes para su análisis.

Para la variable exportaciones de aguacate (en toneladas) se calcularon las medidas básicas de tendencia central y dispersión:

$$n = 30, \quad \text{número de observaciones}, \quad (5)$$

$$\bar{x} = 318\,674.72 \text{ toneladas}, \quad (6)$$

$$s = 100\,061.44 \text{ toneladas}, \quad (7)$$

$$\text{mín}(x) = 140\,857.77, \quad \text{máx}(x) = 522\,430.98, \quad (8)$$

$$Q_1 = 254\,873.12, \quad Q_2 = 296\,551.88, \quad Q_3 = 388\,319.30. \quad (9)$$

La media se ubica en torno a 318 mil toneladas, con una desviación estándar considerable de 100 mil, lo que indica una alta variabilidad. El rango intercuartílico nos muestra que el 50 % central de los datos se concentra entre aproximadamente 255 mil y 388 mil toneladas, mientras que tenemos valores extremos que alcanzan mínimos cercanos a 141 mil y máximos de más de 522 mil toneladas.

La Figura 1 muestra la serie bimestral; la Figura 2 presenta la distribución de las exportaciones; la Figura 3 muestra el boxplot de las exportaciones.

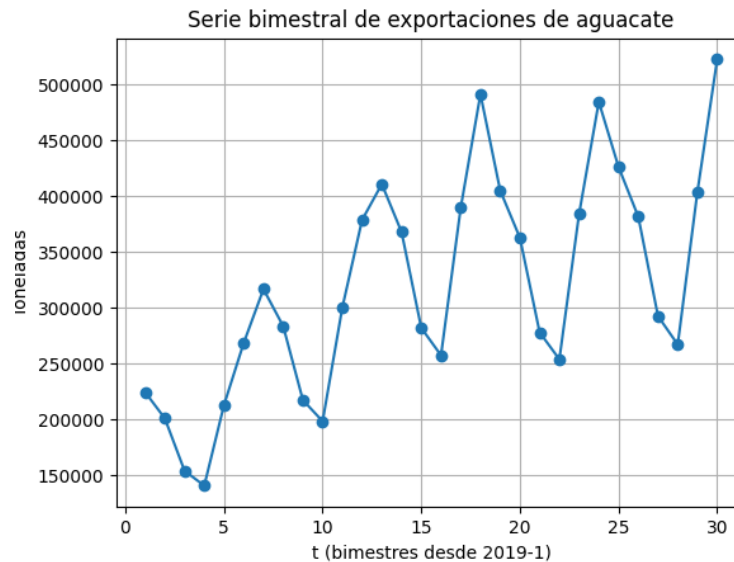


Figura 1: Serie bimestral de exportaciones de aguacate (toneladas).

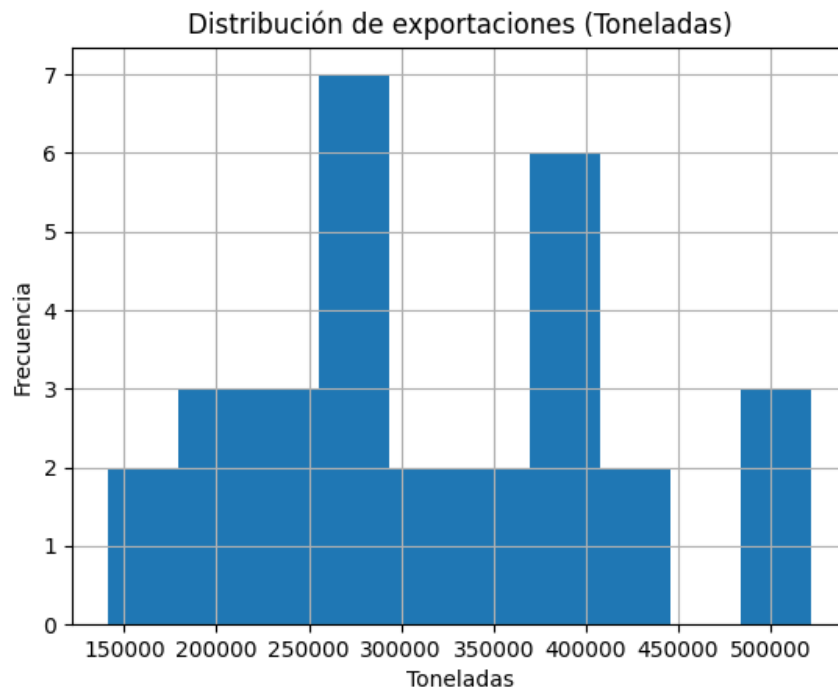


Figura 2: Histograma de exportaciones bimestrales (toneladas).

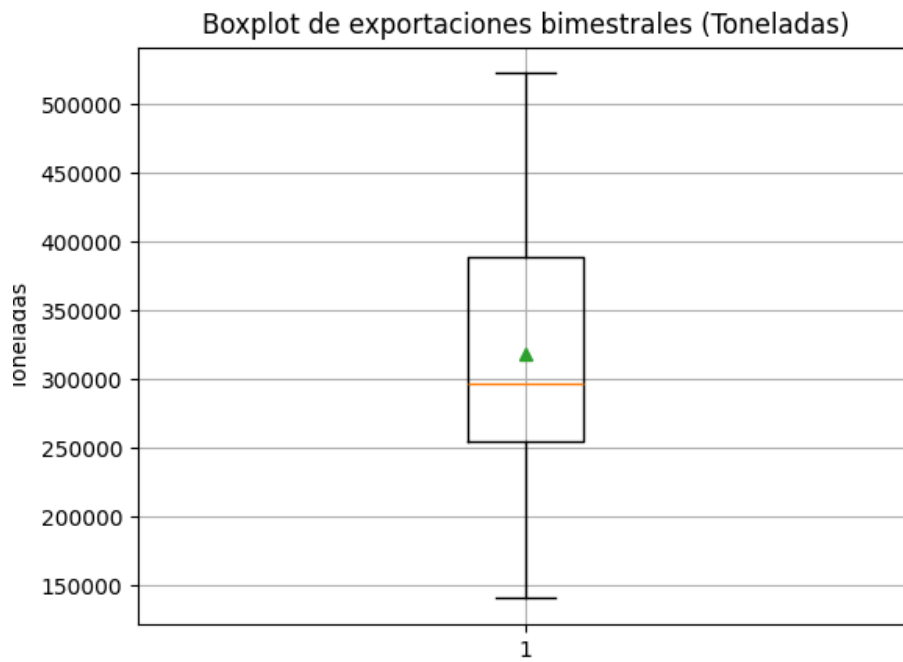


Figura 3: Boxplot de exportaciones bimestrales de aguacate (toneladas).

3. Considerando un modelo de regresión de la serie original contra el tiempo, el valor del coeficiente de pendiente muestral es en miles de toneladas es:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad (10)$$

donde y_t representa las exportaciones (en toneladas) en el bimestre t , β_0 es la ordenada al origen, β_1 la pendiente que mide el cambio esperado en toneladas por cada bimestre y ε_t el término de error. A partir de la regresión con Mínimos Cuadrados Ordinarios (MCO) se obtuvo:

$$\hat{\beta}_0 = 201,900 \text{ toneladas (aprox.)}, \quad (11)$$

$$\hat{\beta}_1 = 7,533.78 \text{ toneladas/bimestre}. \quad (12)$$

En términos de miles de toneladas:

$$\hat{\beta}_1 \approx \mathbf{7.53} \text{ mil toneladas/bimestre}. \quad (13)$$

La pendiente positiva indica que, en promedio, las exportaciones de aguacate han aumentado en alrededor de 7.5 mil toneladas por cada bimestre a lo largo del periodo 2019–2023.

4. Elabore la prueba de hipótesis correspondiente con un 0.05 de significancia para validar si las exportaciones de aguacate han venido creciendo en el tiempo.

Consideremos:

$$H_0 : \beta_1 \leq 0 \quad (\text{no existe tendencia creciente}), \quad (14)$$

$$H_1 : \beta_1 > 0 \quad (\text{existe tendencia creciente}). \quad (15)$$

A partir de la regresión lineal sobre la serie original se obtuvo:

$$t = \frac{\hat{\beta}_1}{\text{EE}(\hat{\beta}_1)} = \frac{7533.78}{1608.38} \approx 4.684. \quad (16)$$

El valor p bilateral reportado fue 6.58×10^{-5} , lo que implica un p unilateral de:

$$p \approx 3.29 \times 10^{-5}. \quad (17)$$

Como $p < 0.05$, se **rechaza la hipótesis nula**. Hay evidencia estadísticamente significativa para concluir que la pendiente es positiva. Las exportaciones de aguacate presentan una tendencia creciente en el periodo analizado.

5. Aplicando la técnica de desestacionalidad simple vista en clase, obtenga los índices estacionales e intérpretelos.

Para identificar la estacionalidad se calcularon índices multiplicativos por bimestre mediante:

$$I_j = \frac{\bar{y}_j}{\bar{y}}, \quad (18)$$

donde \bar{y}_j es el promedio de las exportaciones en el bimestre j a lo largo de los años, y \bar{y} es el promedio global de toda la serie.

Resultados obtenidos:

$$I_1 = 1.1193, \quad I_2 = 1.0024, \quad I_3 = 0.7685, \quad (19)$$

$$I_4 = 0.7017, \quad I_5 = 1.0617, \quad I_6 = 1.3465. \quad (20)$$

Interpretación.

- Valores $I_j > 1$ indican bimestres con exportaciones por encima del promedio general.

- Valores $I_j < 1$ reflejan bimestres con exportaciones por debajo del promedio.

Particularmente

- El bimestre 6 (jul-ago) destaca con un índice de 1.3465, lo que significa un 34.7 % por encima del promedio.
- El bimestre 4 (may-jun) es el más bajo, con 0.7017, es decir, un 29.8 % por debajo del promedio.
- Los bimestres 1 y 5 también muestran exportaciones superiores al promedio, mientras que bimestres 3 y 4 son consistentemente más bajos.



Figura 4: Serie bimestral de exportaciones de aguacate (toneladas) desestacionalizada.

- Obtenga la regresión lineal simple con los datos desestacionalizados y el tiempo y elabore la prueba de hipótesis correspondiente con un 0.05 de significancia para validar si la tendencia de las exportaciones de aguacate ha venido creciendo en el tiempo.

Una vez que eliminamos el componente estacional, se ajusta nuevamente una regresión lineal simple:

$$y_t^* = \beta_0 + \beta_1 t + \varepsilon_t, \quad (21)$$

Con Mínimos Cuadrados Ordinarios se obtuvo:

$$\hat{\beta}_0 = 209,000 \text{ toneladas (aprox.)}, \quad (22)$$

$$\hat{\beta}_1 = 7,076.68 \text{ toneladas/bimestre}. \quad (23)$$

Prueba de hipótesis.

$$H_0 : \beta_1 \leq 0, \quad (24)$$

$$H_1 : \beta_1 > 0. \quad (25)$$

El estadístico t calculado fue:

$$t = \frac{7076.68}{666.67} \approx 10.615. \quad (26)$$

El valor- p unilateral fue de 1.27×10^{-11} .

Dado que $p < 0.05$, se **rechaza** H_0 . Hay evidencia estadísticamente significativa de que la pendiente es positiva. Aún eliminada la estacionalidad, la tendencia de las exportaciones de aguacate sigue siendo creciente, con un incremento promedio de alrededor de 7.1 mil toneladas por bimestre.

7. Proporcione un intervalo al 95 % de confianza para el valor de la tendencia de la serie de exportaciones de aguacate en el segundo bimestre del 2024.

Usando el modelo de regresión de la serie original

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \quad t = 1, \dots, 30, \quad (27)$$

estimamos la media esperada (tendencia) para $t = 32$ (bimestre 2 de 2024) mediante

$$\hat{y}_{32} = \hat{\beta}_0 + \hat{\beta}_1 32. \quad (28)$$

El intervalo de confianza al 95 % para la media condicional en $t = 32$ es

$$\hat{y}_{32} \pm t_{\alpha/2, n-2} \text{EE}(\hat{y}_{32}), \quad (29)$$

donde $\text{EE}(\hat{y}_{32})$ es el error estándar de la media pronosticada. Numéricamente se obtuvo

$$\boxed{[381\,595.30, \, 504\,368.80] \text{ toneladas}}, \quad \hat{y}_{32} = \mathbf{442\,982.05}. \quad (30)$$

El siguiente intervalo de predicción al 95 % incorpora la variabilidad del error idiosincrático y es más amplio:

$$\hat{y}_{32} \pm t_{\alpha/2, n-2} \sqrt{\text{EE}(\hat{y}_{32})^2 + \hat{\sigma}^2}. \quad (31)$$

En este caso:

$$\boxed{[275\,161.00, \, 610\,803.10] \text{ toneladas}}. \quad (32)$$

La media esperada de la tendencia para 2024-bimestre 2 es aproximadamente 443 mil toneladas; un valor observado individual podría caer, con 95 % de confianza, dentro del intervalo de predicción indicado, que es más amplio por incluir la incertidumbre del término de error.

8. Compare los resultados de los modelos de regresión lineal simple ajustados a la serie de tiempo original y a la serie desestacionalizada, ambos contra el tiempo. ¿cuál es una explicación sólida para justificar el hecho de que el coeficiente de determinación cambie?

Recordemos que

$$R^2 = 1 - \frac{\text{SSR}}{\text{SST}} = \frac{\text{SST} - \text{SSR}}{\text{SST}} = \frac{\text{SS Explicada}}{\text{SST}}, \quad (33)$$

donde $\text{SSR} = \sum_t \hat{\varepsilon}_t^2$ es la suma de residuos al cuadrado y $\text{SST} = \sum_t (y_t - \bar{y})^2$ es la suma total de cuadrados.

Resultados:

$$R^2 (\text{serie original}) = 0.4393, \quad (34)$$

$$R^2 (\text{serie desestacionalizada}) = 0.8010. \quad (35)$$

Interpretación. Al eliminar la variación estacional, la parte de variabilidad atribuible a la estacionalidad deja de contaminar el término de error. En consecuencia, el componente de *tendencia* capturado por t explica una fracción mucho mayor de la variabilidad de la serie, elevando el R^2 de 0.44 a 0.80. En otras palabras, el incremento en R^2 refleja que la estacionalidad aportaba una porción relevante de la variación total; una vez removida, la relación lineal con el tiempo resulta más estrecha.

Conclusión. El modelo sobre datos desestacionalizados presenta un ajuste sustancialmente mejor para capturar la tendencia subyacente de las exportaciones de aguacate.

9. Grafique los pronósticos obtenidos con ambas técnicas y comente los resultados.

Modelo sobre serie original. Ajusta una tendencia lineal sin estacionalidad explícita; por ello los pronósticos son aproximadamente lineales y crecen suavemente (Figura 5). Se estima la tendencia con la serie sin estacionalidad y luego se *reestacionaliza* multiplicando por los índices bimestrales; por eso los pronósticos reflejan picos y valles estacionales (Figura 6).

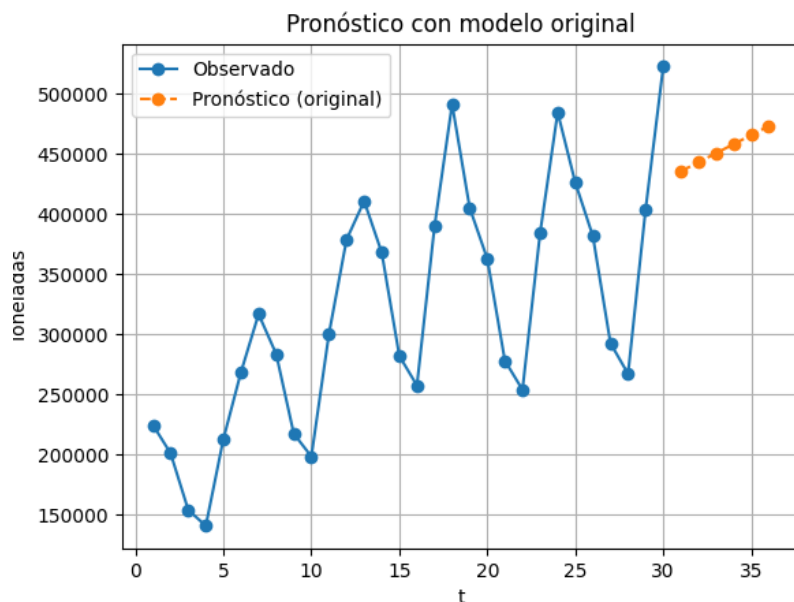


Figura 5: Pronóstico con modelo original (sin estacionalidad explícita).

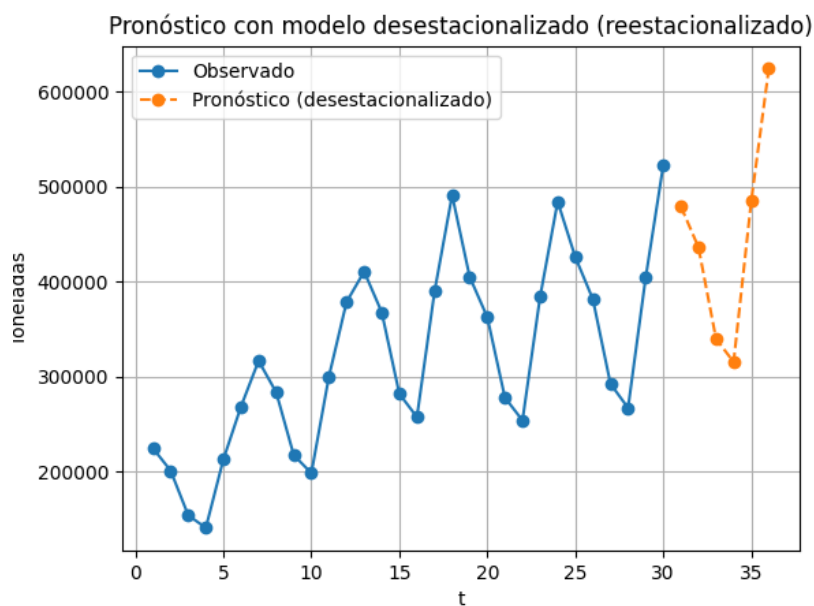


Figura 6: Pronóstico con modelo desestacionalizado (reestacionalizado).

Observaciones:

- El **modelo original** proyecta valores entre $\sim 435\text{K}$ y $\sim 473\text{K}$ toneladas para 2024-bim1 al 6, con crecimiento suave y bandas relativamente estrechas, coherente con una tendencia lineal.
- El **modelo desestacionalizado** reproduce el patrón estacional: mínimos en bimestres 3–4 ($\sim 340\text{K}$ y $\sim 315\text{K}$) y máximo en bimestre 6 ($\sim 624\text{K}$), en línea con los índices estacionales ($I_6 > 1$ e $I_{3,4} < 1$). Sus bandas son más amplias en los picos al reintroducir la estacionalidad.

3. Link al repositorio con código fuente

<https://github.com/enriquegomeztagle/MCD-Econometria/tree/main/HWs/TimeSeriesDeseasonalization>