

Examen parcial 1

PUBLISHED

February 27, 2024

Inicio:

Introduzca su correo institucional completo:

fernando.marquez@iteso.mx

LUIS FERNANDO MARQUEZ BAÑUELOS

Instrucciones

- Abra el examen en el explorador (Chrome, etc.)
- Lea cuidadosamente todas las indicaciones.
- Asegúrese de que su nombre se muestra al escribir su correo institucional. En caso de que no, contacte al profesor de inmediato.
- Los puntos se otorgan a respuestas **completas**.
- La serie de tiempo que le tocó servirá para todos los ejercicios del examen.
- Al finalizar el examen, mande imprimir y guarde en formato **PDF** su examen y súbalo a CANVAS.
- Queda estrictamente prohibida toda forma de comunicación con otras personas.
- Puede subir documentos adicionales (en PDF o HTML) si así lo requiere. Asegúrese de que está bien renderizado.

Consulta de notas y uso de IA

Ej. 1

La serie que le corresponde analizar es la producción de la construcción en México.

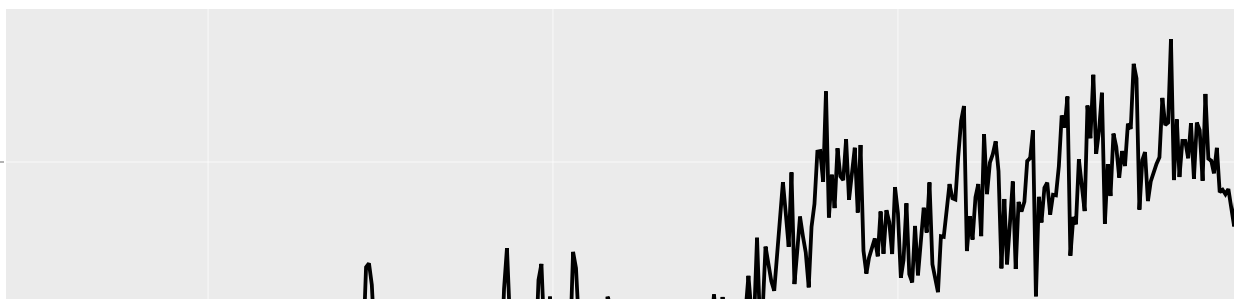
Gráfica

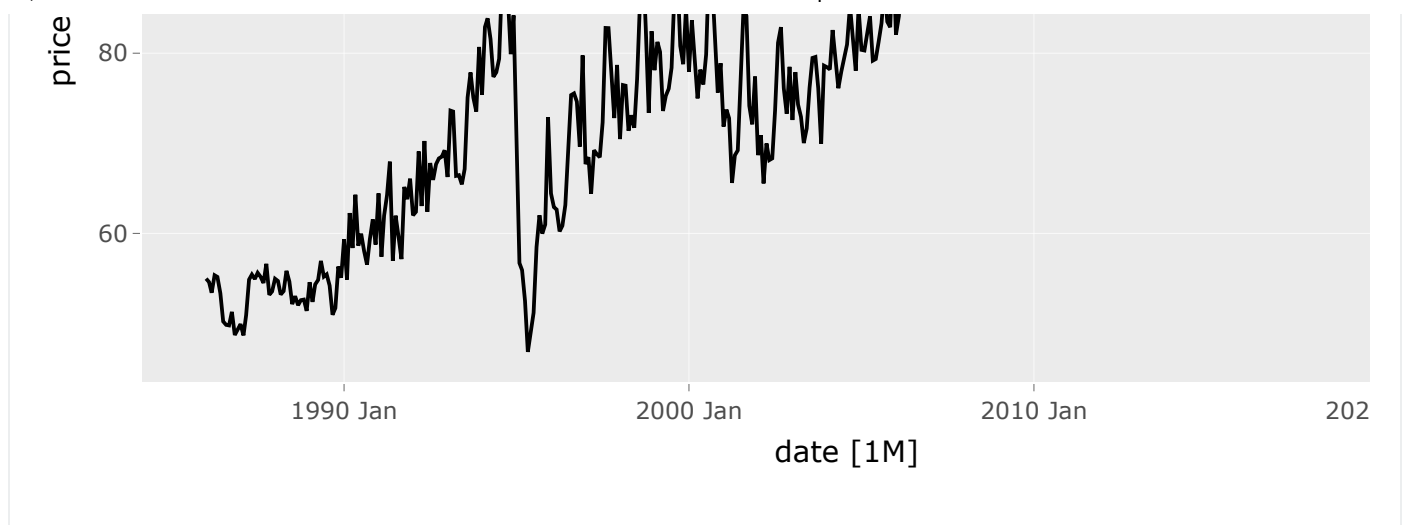
tsibble

Tabla

Código

100





a)

- Describa a la serie de la forma más detallada posible (*evite rollo innecesario*).
- Puede mencionar si detecta algún patrón.

Escriba su respuesta:

varianza variable, es decir, tiene heterocedasticidad, ya que con el tiempo en ocasiones incrementa y en otras disminuye. También tanto en 1995 como en 2020 tiene outliers de niveles de construcción muy bajos respecto al periodo anterior, esto se puede deber a la crisis de 1994 y a la pandemia en 2020. Además de tener una estacionalidad anual en los niveles de construcción.

También tanto en septiembre como diciembre suelen tener los niveles más altos de construcción en cada año.

b)

- Para efectos de modelar la serie, ¿requiere realizar algún ajuste y/o transformación? En caso de que sí, ¿cuál y por qué?

Escriba su respuesta:

La transformación que se requiere fue la transformación box-cox, si bien deja bastante similar a la serie es bueno utilizarla para ver si estabiliza la varianza de la serie, además suele ayudar al momento de hacer pronósticos aunque a simple vista no modifique mucho la serie.

c)

- ¿Sería útil realizar una descomposición para esta serie?, ¿por qué?
- Si su respuesta fue positiva, ¿cuál tipo de descomposición propondría para este caso?

Escriba su respuesta:

sus componentes de tendencia y estacionalidad y su comportamiento a lo largo del tiempo. Podemos confirmar si realmente tiene estacionalidad la serie, lo cual es cierto. Además de ver la tendencia y estacionalidad para saber luego que modelos pueden ser mejores para su modelado, como el snave o combinado.

La descomposición elegida fue la STL ya que ayuda a suavizar la tendencia y al usar su opción de "robust" ayuda a que los outliers no afecten la tendencia y esta serie si cuenta con outliers.

Ej. 2

De acuerdo a los datos que le tocó, cree un escenario de negocios en el cual podrían contratarle para modelarlos y crear pronósticos. Responda lo siguiente:

- ¿Cuál sería el propósito del pronóstico?
- ¿Qué utilizarían de datos de entrenamiento y cuáles de validación?

a)

Escriba su respuesta:

vemos que la producción ha aumentado (precio). Esto puede ayudar a la empresa a estimar ganancias y costos que pueda tener a futuro ya que el precio los afecta. Además con esto puede realizar un plan de gastos para tener un estimado del dinero necesario en base a la posible producción que habrá en el país.

Tenemos datos de 37 años, para entrenamiento es conveniente utilizar una mayor cantidad de datos por lo que se usarán el 80% de los datos que serían 30 años (hasta 2016) y para validación el 20% restante que son de enero 2017 hasta marzo 2023.

b)

- ¿Qué horizonte de pronóstico establecerían?

Escriba su respuesta:

En base a la cantidad de datos que es de 40 años, se puede establecer como horizonte los siguientes 6 años.

- ¿Cuál métrica de error utilizarían para estimar el error de pronóstico?, ¿por qué?

Escriba su respuesta:

Como métrica de error usaremos el MAPE, ya que es un error porcentual entre el modelo y los datos reales, este se elige porque es muy utilizado en general en empresas y negocios, estamos más familiarizados con él, y es bastante sencillo de interpretar ya que es solo una diferencia en porcentaje.

Ej. 3

Siguiendo con su respuesta del Ej. 2, responda a continuación:

a)

- ¿Cuál(es) modelo(s) estimaría para la serie que se le asignó?, ¿por qué?

Escriba su respuesta:

Para estimar la serie usaremos snaive y combinación de snaive con drift, tanto para la serie original y para la transformación Box-Cox, porque como vimos en la descomposición de esta tiene tanto tendencia como estacionalidad por lo que estos modelos se ajustan mejor a estas características.

La tendencia al cambiar con el tiempo no es muy útil el drift, al tener tendencia usar el modelo de mean no sería adecuado y como no son aleatorios los datos el método de naive tampoco ayudaría.

b)

- Construya aquí la tabla de modelo(s) (**mable**) como la pondría en su flujo de pronóstico:

Escriba su respuesta:

```
)  
const_fit
```

El código representa la tabla de modelos que utilizamos para esta serie de tiempo.

Snaive y snaive_bc(serie transformada) es el modelo de seasonal naive, este es útil para nuestra serie de tiempo ya que como vimos en la descomposición tiene estacionalidad por lo que puede dar pronósticos buenos.

Combi y combi_bc(series transformada) combina tanto a snaive como a drift que es la tendencia, son un útiles ya que además de apoyarse en la estacionalidad de la serie usa la tendencia que esta tiene para realizar sus pronósticos, lo cual es de utilidad ya que nuestra serie de tiempo cuenta con una tendencia a la alza, más aun el final ya que viene de recuperación de la pandemia.

c)

- Efectúe el diagnóstico de residuos y coloque aquí sus conclusiones al respecto. *Recuerde subir un documento de Quarto en HTML o PDF detallando esto.*

Escriba su respuesta:

Para el modelo combinado tanto en la serie original como en la transformada, los residuos muestran un poco de correlación, la cual es menor a la mostrada en el modelo snaive, y sus datos no son distribución normal pero se asemeja a la forma. Esto indica que pudieran ser un poco mejores los modelos combinados.

Para los 4 modelos el p-value es prácticamente 0.

En cuanto a su MAPE (error porcentual medio absoluto), el modelo combinado en la transformación y serie original es mejor teniendo un error ligeramente mayor al 3% en ambos casos.

d)

- Una vez que haya encontrado buenos modelos, produzca los pronósticos que mencionó en el ejercicio 2.

Escriba su respuesta:

```
autoplot(const_mx)
```

Decidí utilizar el modelo combinado con la serie transformada a box-cox. Esto se debe a que en los datos de training tuvo el error porcentual más pequeño. Si bien en los datos de test tuvo un error mayor esto se debe a que toma la tendencia y como hubo una pandemia los datos reales bajaron bastante. Sin embargo, al momento de pronosticar a futuro, si tomará en cuenta los datos más actuales del bajón por la pandemia y de como se ha ido elevando la producción de la construcción a partir de 2020 la serie recupero su tendencia a la alza, por lo que usar el modelo combinado de box-cox obtendremos un buen pronóstico a futuro ya que seremos capaces de captar la tendencia a la alza actual.

Ej. 4

- Escriba debajo sus conclusiones del análisis que realizó.

Escriba su respuesta:

Tras haber realizado todo el análisis de la serie, se puede concluir que cuenta con estacionalidad y tendencia la producción de la construcción en México. Esto es importante ya que nos permite saber que puede tener ciertos cambios dependiendo del mes del año y que hay una tendencia a futuro que es a la alza.

Debido a esto vemos que es probable que utilizar un modelo de seasonal naive o combiando son las mejores opciones. Después de hacer las pruebas con el set de entrenamiento y test de los datos vemos que en el entrenamiento el mejor fue el modelo combinado box-cox pero en el test fue el snaive, sin embargo, esto se debe a los outliers provocados por la pandemia.

Como desde la pandemia a la actualidad se retomo una tendencia a la alza y en los datos de entrenamiento donde habia esta tendencia el modelo combinado de box-cox se adaptó muy bien a los datos fue el elegido para realizar el pronóstico a futuro donde vemos que la producción de la producción crecerá a futuro.