



INSTITUTO POLITECNICO NACIONAL

Escuela Superior de Cómputo
(ESCOM)

Licenciatura en Ciencias de Datos.

Nombre de la unidad de aprendizaje:

Series de Tiempo.

Grupo: 4AV1.

Nombre de la Actividad:

“Practica 1.0 Naturaleza del análisis
de las series de tiempo”.

Nombre del alumno:

Arteaga Gonzáles Edwin Yahir

Juarez Gaona Erick Rafael.

Rico Gaytan Diana Andrea.

Ruiz Merino Wendy Ivonne.

Fecha:

02/03/2025.

INTRODUCCIÓN

El presente documento tiene como objetivo analizar la afluencia en el metro a través de técnicas de series de tiempo en Python. Para ello, se han implementado diversas metodologías que permiten comprender el comportamiento de los datos y extraer información relevante para la toma de decisiones.

Objetivos del análisis

1. Carga y limpieza de datos: Se procesa la información para convertir la fecha y hora en un índice de tiempo, facilitando su análisis.
2. Visualización de la serie de tiempo: Se generan gráficos para observar patrones de comportamiento en la afluencia del metro.
3. Análisis estadístico: Se calculan medidas descriptivas como promedio, mediana, moda, desviación estándar y varianza.
4. Identificación de patrones: Se analizan granularidad, ciclos, tendencias, estacionalidad, así como los valores máximos, mínimos, picos y valles.
5. Cálculo de razones de crecimiento y reducción: Se estudia la evolución de la afluencia a lo largo del tiempo.
6. Aplicación de la metodología Box-Jenkins: Se explica cómo esta metodología puede utilizarse para modelar la serie de tiempo y realizar predicciones.

Este estudio busca ofrecer una visión detallada de la dinámica de la afluencia en el metro y cómo estos datos pueden ser utilizados para la planificación y optimización del servicio.

DESARROLLO

Para el análisis de la afluencia en el metro mediante series de tiempo, se han implementado diversas técnicas de procesamiento y visualización de datos en Python. A continuación, se describen las etapas clave del estudio:

1. Carga y limpieza de datos

Se realizó la lectura del archivo de datos, convirtiendo la fecha y la hora en un único índice de tiempo. Esta transformación facilita la manipulación de la información y la aplicación de modelos de series de tiempo. También se eliminaron columnas innecesarias para optimizar el procesamiento.

2. Visualización de la serie de tiempo

Se generó un gráfico de la serie de tiempo para observar patrones y fluctuaciones en la afluencia del metro. Esta visualización permite detectar tendencias generales, ciclos y posibles anomalías en los datos.

3. Análisis estadístico

Se calcularon estadísticas descriptivas clave:

- Promedio: Representa la afluencia media del metro en el periodo analizado.
- Mediana: Indica el valor central de la afluencia.
- Moda: Determina la afluencia más frecuente en los registros.
- Desviación estándar y varianza: Miden la dispersión de los datos con respecto a la media, ayudando a entender la variabilidad de la afluencia.

4. Identificación de patrones

Se analizaron diferentes características de la serie de tiempo:

- Granularidad: Se identificó el nivel de detalle de los datos (frecuencia horaria en este caso).
- Ciclos: Se estudió la periodicidad de la afluencia mediante la autocorrelación, observando si hay repeticiones diarias o semanales.
- Tendencias y estacionalidad: Se descompuso la serie de tiempo para detectar componentes de largo plazo y variaciones estacionales.
- Máximos, mínimos, picos y valles: Se identificaron los momentos de mayor y menor afluencia, lo que ayuda a entender los horarios críticos.

5. Razones de crecimiento y reducción

Se calcularon tasas de crecimiento y reducción de la afluencia en función del tiempo, lo que permite evaluar la evolución del número de pasajeros y anticipar cambios en la demanda del servicio.

6. Aplicación de la metodología Box-Jenkins

Se aplicaron los principios de la metodología Box-Jenkins para el modelado de la serie de

tiempo:

- Identificación: Se evaluó la estacionariedad de la serie con la prueba de Dickey-Fuller y se analizaron las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) para determinar los parámetros del modelo ARIMA.
- Estimación: Se ajustó un modelo ARIMA basado en los valores óptimos de los parámetros obtenidos en la fase anterior.
- Verificación: Se analizaron los residuos del modelo para confirmar que cumplen con los supuestos de ruido blanco, garantizando la validez del ajuste.

A través de este análisis, se logra una comprensión detallada de la dinámica de la afluencia en el metro y se establecen bases para la predicción futura de la demanda del servicio.

CONCLUSIONES

Conclusión de Juárez Gaona Erick Rafael

El estudio de la serie de tiempo ha sido clave para visualizar la evolución del transporte urbano y su comportamiento en diferentes momentos del día. La metodología Box-Jenkins para encontrar los parámetros correctos para establecer ya sea el modelo ARMIA o ARMA que aun bien como emplearlo o identificar bien estos patrones pues creo que es muy muy importante de identificar de manera correcta. Pero creo que la manera en la que se identifican las variables indispensables es muy muy importante para un correcto análisis.

Conclusión de Rico Gaytán Diana Andrea Diana

La aplicación de métodos estadísticos y de series de tiempo ha permitido comprender la variabilidad en la afluencia del metro. Gracias a este análisis, es posible anticipar momentos de alta demanda y tomar decisiones estratégicas para optimizar el servicio, evitando sobrecargas en ciertas líneas y mejorando la experiencia de los pasajeros. Sin embargo, uno de los desafíos fue la identificación de la estacionalidad en los datos debido a fluctuaciones irregulares. Para superarlo, utilizamos técnicas de descomposición de series de tiempo y análisis de autocorrelación, permitiendo una mejor detección de patrones.

Conclusión de Ruiz Merino Wendy Ivonne

El análisis de la afluencia en el metro a lo largo del tiempo ha permitido identificar patrones importantes que pueden ser aprovechados para la planificación operativa. La detección de tendencias y ciclos recurrentes demuestra que el uso del transporte público responde a horarios definidos, lo que permite mejorar la distribución de recursos en función de la demanda. Una de las dificultades encontradas fue la presencia de datos incompletos, lo que afectaba la calidad del análisis. Para resolverlo, se implementaron estrategias de imputación de datos para minimizar su impacto en los resultados.

Conclusión de Arteaga Gonzáles Edwin Yahir

La identificación de picos y valles en la afluencia del metro muestra cómo varía el número de pasajeros según el horario y el día de la semana. Este conocimiento es valioso para implementar estrategias de optimización en la infraestructura del transporte público y garantizar un servicio más eficiente para la población. Uno de los obstáculos encontrados fue la alta variabilidad en la afluencia, lo que dificultaba la detección de patrones claros. Para solucionarlo, se aplicaron métodos de suavizado y análisis de ventanas móviles, logrando una mejor visualización de la tendencia general.

CONCLUSIÓN GENERAL

El análisis de la afluencia en el metro mediante series de tiempo ha permitido comprender mejor su comportamiento y evolución a lo largo del tiempo. La combinación de visualizaciones, estadísticas descriptivas y modelos de predicción proporciona una base sólida para la toma de decisiones informadas en la planificación del transporte público. No obstante, el proceso no estuvo exento de desafíos, como la presencia de datos incompletos, la dificultad en la detección de estacionalidad y la optimización de modelos predictivos. Estos problemas fueron resueltos mediante técnicas avanzadas de análisis de datos, lo que permitió obtener resultados más precisos y útiles para mejorar la eficiencia del servicio de transporte.