

Versión 1.0.3

Realizar el análisis de predicción de filtrado en la Zona Metropolitana de Guadalajara es el siguiente:

1. Importamos el archivo fuente y realizamos la lectura.
2. Validamos que la información esté correctamente clasificada.
3. Realizamos un filtrado por tipo de consumo municipal.
4. Realizamos un filtrado por tipo de consumo municipal tal que el municipio sea igual a ZMG.
5. Validamos que la información esté correctamente clasificada.
6. Agrupamos los datos por año.
7. Validamos si tenemos registros sin datos.
8. Validamos que la información esté correctamente clasificada.
9. Almacenamos la información para su manipulación y preservación de datos.
10. Leemos la información que acabamos de almacenar.
11. Validamos que la información esté correctamente clasificada.
12. Actualizamos el formato de la fecha.
13. Validamos que la información esté correctamente clasificada.
14. Validamos si tenemos registros sin datos.
15. Realizamos un filtrado por año.
16. Validamos que la información esté correctamente clasificada.
17. Agrupamos los datos por año.
18. Calculamos la suma de los valores en cada año.
19. Almacenamos la información para su manipulación y preservación de datos.
20. Leemos la información que acabamos de almacenar.
21. Validamos que la información esté correctamente clasificada.

22. Realizamos una copia de la información.
23. Imprimimos el número de existencias.
24. Imprimimos el nombre de las existencias.
25. Verificamos si existen valores nulos en nuestros registros.
26. Graficamos y mostramos los datos de los productos ofertados con relación a los años.
27. Creamos una figura de gráfico de líneas.
28. Recorremos cada columna, agregando una serie de datos de dispersión para cada columna.
29. Mostramos la figura (gráfico).
30. Validamos que la información esté correctamente clasificada.
31. Organizamos la información en un gráfico (Mapa de calor).
32. Cargamos los datos.
33. Eliminamos columnas no deseadas (Año).
34. Calculamos la matriz de correlación de las tasas de retorno diarias.
35. Mostramos el mapa de calor.
36. Calculamos el rendimiento promedio de cada acción en el periodo de tiempo considerado.
37. Calculamos la desviación estándar de cada columna.
38. Leemos los datos del archivo de datos.
39. Validamos que la información esté correctamente clasificada.
40. Reemplazamos el nombre de "YE" por "Año" (Posible eliminación).
41. Establecemos el año como índice.
42. Agrupamos y clasificamos los datos por año.
43. Verificamos si existen valores nulos en nuestros registros.
44. Clasificamos y mostramos el de los productos según el año (por ejemplo, azul).

45. Aplicamos la prueba de Dickey-Fuller aumentada a la serie de datos de votos para determinar si una serie temporal es estacionaria o no.

46. Repetimos el proceso con cada producto.

46. Repetir el proceso para cada producto:

- Extraemos los datos de un producto X.
- Guardamos los datos en un archivo.
- Cargamos esos datos y se almacenan.
- Validamos que la información esté correctamente clasificada.
- Configuramos el modelo con las opciones deseadas.
- Ajustamos el modelo a los datos disponibles.
- Generamos fechas futuras (próximos 3 años) a partir de los datos

trabajados.

- Mostramos las fechas para verificar su correcta clasificación.
- Realizamos pronósticos de las fechas futuras generadas.
- Guardamos los resultados.
- Validamos que la información esté correctamente clasificada.

ds:

Fechas futuras para las que se realizaron.

yhat:

Estimación puntual del valor pronóstico.

yhat_lower:

Límite inferior del intervalo de confianza.

yhat_upper:

Límite superior del intervalo de confianza del pronóstico.

- Mostramos los datos de manera gráfica (los pronósticos junto con los datos

históricos).

- Generamos un gráfico con los datos del pronóstico incluyendo:
 - Tendencia anual.
 - Tendencia semanal.
 - Estacionalidad anual.
 - Estacionalidad diaria

Funciones

Usando conjunto de datos de consumo municipal de productos determinados con fin de vida y nueva introducción de productos definida, a lo largo de varios años, denotado como D, donde cada registro contiene información sobre el consumo en un año específico, el municipio al que pertenece, y otros atributos relevantes.

“Importación y Lectura de Datos”

Definimos una función `D = ImportarDatos(archivo_fuente)` que carga el conjunto de datos desde un archivo fuente especificado.

“Validación de Clasificación”

Utilizamos una función `ValidarClasificación(D)` que verifica que los datos estén correctamente clasificados.

“Filtrado por Tipo de Consumo Municipal”

Aplicamos una función `D = FiltrarPorTipoMunicipal(D, tipo)` que filtra los datos según el tipo de consumo municipal especificado.

“Ordenamiento de Datos”

Realizamos una función `D = OrdenarPorAño(D)` que ordena los datos en base al año de menor a mayor.

“Validación de Registros sin Datos”

Utilizamos una función `ValidarRegistrosSinDatos(D)` para asegurarnos de que no haya registros vacíos.

“Almacenamiento de Datos”

Creamos una función `AlmacenarDatos(D)` que almacena los datos procesados para su manipulación y preservación.

“Filtrado por Año”

Implementamos una función `D = FiltrarPorAño(D, año)` para filtrar los datos por año específico.

“Agrupamiento de Datos por Año”

Utilizamos una función `G = AgruparPorAño(D)` que agrupa los datos por año y calcula la suma de los valores en cada año.

“Cálculo de Rendimiento Promedio”

Calculamos el rendimiento promedio de cada acción en el período de tiempo considerado mediante una función `R = CalcularRendimientoPromedio(G)`.

“Cálculo de Desviación Estándar”

Determinamos la desviación estándar de cada columna con una función `D = CalcularDesviacionEstandar(G)`.

“Gráficos”

Creamos gráficos para visualizar los datos y resultados del análisis.

“Pruebas Estadísticas”

Realizamos pruebas estadísticas como la Prueba de Dickey-Fuller Aumentada para evaluar la estacionalidad y estacionariedad de las series temporales.

“Modelado y Pronóstico”

Configuramos modelos estadísticos para realizar pronósticos sobre los datos futuros.

“Validación de Resultados”

Verificamos la calidad y precisión de los resultados obtenidos.

Este modelo matemático se basa en funciones y operaciones que representan los pasos clave del algoritmo. Sin embargo, la implementación detallada de estas funciones requeriría conocimientos específicos de programación y estadísticas, así como datos reales para trabajar. Cada función en este modelo matemático podría representarse mediante ecuaciones matemáticas y algoritmos específicos según las necesidades del análisis de datos real.

Posible notación matemática

"Filtrado por tipo de consumo municipal tal que el municipio sea igual a ZMG".

- Sea D el conjunto de datos original.
- Sea F el conjunto de datos resultante después del filtrado por tipo de consumo municipal igual a ZMG.

$$F = \{ d \in D \mid d.\text{municipio} = \text{"ZMG"} \}$$

"Calcular la suma de los valores en cada año".

- Sea V el conjunto de valores a sumar.
- Sea S el conjunto de sumas por año.

$$S = \{ \sum v_i \mid v_i \in V, \text{año}_i = \text{año}_j \}$$

"Validamos que la información esté correctamente clasificada".

- Sea C una función que valida la clasificación de datos. Si C retorna verdadero, entonces los datos están correctamente clasificados.

$C(D) = \text{Verdadero}$ si los datos están correctamente clasificados, Falso en caso contrario.

"Ordenamos los datos en base al año de menor a mayor".

- Sea D el conjunto de datos original.
- Sea O el conjunto de datos ordenados en base al año de menor a mayor.

$O = \text{Ordenar}(D, \text{año})$.

"Validamos si tenemos registros sin datos".

- Sea V un conjunto de registros que contienen datos.
- Sea S un conjunto de registros sin datos.

$V = \{d \in D \mid d.\text{contieneDatos} = \text{Verdadero}\}$
 $S = \{d \in D \mid d.\text{contieneDatos} = \text{Falso}\}$

"Realizamos un filtrado por año"

- Sea D el conjunto de datos original.
- Sea F el conjunto de datos resultante después del filtrado por año.

$F = \{d \in D \mid d.\text{año} = \text{añoEspecifico}\}$

"Calculamos la matriz de correlación de las tasas de retorno diarias."

La acción de calcular la matriz de correlación de las tasas de retorno diarias se puede expresar de la siguiente manera en notación matemática:

Supongamos que tenemos un conjunto de activos financieros, y para cada activo, tenemos una serie de tasas de retorno diarias registradas durante un período de tiempo. Denotemos estas series de tasas de retorno como R_1, R_2, \dots, R_n , donde R_i es la serie correspondiente al activo i .

Para calcular la matriz de correlación entre estas tasas de retorno diarias, podemos utilizar la fórmula de correlación de Pearson entre dos series R_i y R_j :

$$\text{Corr}(R_i, R_j) = \frac{\sum_{t=1}^T (R_{i,t} - \bar{R}_i)(R_{j,t} - \bar{R}_j)}{\sqrt{\sum_{t=1}^T (R_{i,t} - \bar{R}_i)^2} \sqrt{\sum_{t=1}^T (R_{j,t} - \bar{R}_j)^2}}$$

Donde:

$\text{Corr}(R_i, R_j)$ es la correlación entre las series R_i, R_j .

$R_{i,t}$ y $R_{j,t}$ son los retornos diarios en el periodo t para los activos i y j , respectivamente.

\bar{R}_i y \bar{R}_j son los promedios de las series R_i y R_j respectivamente.

T es el número total de observaciones (días) en las series de retorno.