

# ÍNDICE DE DOCUMENTACIÓN

## CONTENIDO

Introducción .....	2
Índice de documentos .....	2
✓ 0. Data SQL: .....	2
❖ 0. Original data: .....	2
❖ 1. BBDD Procesada: .....	2
✓ 1. Cuadernos: .....	2
❖ 0. Azure Data Study: .....	2
❖ 1. Jupyter Lab: .....	2
❖ 2. Code SQL: .....	3
✓ 2. Datos en CSV: .....	3
✓ 3. Memoria: .....	3
✓ 4. Modelo: .....	3
✓ 5. Power BI: .....	3
✓ Instrucciones.pdf: .....	3

## Introducción

En este documento se listan los archivos entregados en el trabajo final de máster.

## Índice de documentos

### ✓ 0. Data SQL:

Esta carpeta contiene dos subcarpetas:

#### ❖ 0. Original data:

Contiene el archivo original de la base de datos, denominado "TFM\_BBDD\_BK.bak".

#### ❖ 1. BBDD Procesada:

Incluye la base de datos transformada, en el archivo "TFM.bak".

### ✓ 1. Cuadernos:

Esta carpeta se subdivide en tres secciones.

#### ❖ 0. Azure Data Study:

Contiene cuatro archivos:

- Exploración y reconocimiento de datos en la tabla Store.ipynb.
- Exploración y reconocimiento de datos en la tabla Walmart\_Sales.ipynb.
- Limpieza y transformación de datos tabla Store.ipynb.
- Limpieza y transformación de datos tabla Walmart\_Sales.ipynb.
- Estos archivos ofrecen una explicación detallada del funcionamiento de los algoritmos utilizados para procesar los datos en SQL. Cabe destacar que están diseñados únicamente para fines ilustrativos y no para el procesamiento directo de datos. Para ejecutarlos, es necesario descargar e instalar [Azure Data Studio](#).

#### ❖ 1. Jupyter Lab:

Contiene todos los archivos utilizados en el entrenamiento de los modelos.

- **0 LDV-APV.ipynb:** Se encarga de la detección de valores atípicos y limpieza de datos.
- **1 EDA.ipynb:** Realiza la exploración de los datos.
- **2 TrainModel-APV.ipynb:** Incluye el entrenamiento y la comparación de distintos modelos.

- **3 MejoradoModelo-APV.ipynb:** Optimiza el mejor modelo seleccionado y valida los supuestos de la regresión lineal.
- **4 ModeloFinal.ipynb:** Guarda los modelos finales.
- **5 Generar\_Datos.ipynb:** Genera datos para un rango de fechas específico, los cuales se utilizan para realizar predicciones.

## ❖ 2. Code SQL:

Contiene los algoritmos para transformar la base de datos.

- **ClearTableStore.sql:** Limpia y transforma la tabla Store en SQL.
- **ClearTableWalmart\_Sales.sql:** Realiza el mismo proceso para la tabla Walmart\_Sales.

## ✓ 2. Datos en CSV:

Esta carpeta incluye los archivos .csv generados durante el procesamiento de datos. Es importante destacar que *"TFM\_Table\_Store.csv"* y *"TFM\_Table\_Walmart\_Sales.csv"* fueron exportados directamente desde la base de datos, mientras que los demás archivos se generaron durante las etapas de limpieza y entrenamiento de los modelos.

## ✓ 3. Memoria:

En este directorio se encuentran los archivos *"MemoriaTFM.docx"* y *"MemoriaTFM.pdf"*, que documentan detalladamente cada etapa del desarrollo del proyecto

## ✓ 4. Modelo:

Este directorio incluye el archivo *"modelos.pkl"*, que almacena internamente los modelos entrenados.

## ✓ 5. Power BI:

Este directorio incluye:

- El informe de Power BI *"Walmart\_Sales\_final.pbix"*.
- La imagen de fondo utilizada en el informe (*"Atlas.jpg"*).

## ✓ Instrucciones.pdf:

Archivo que contiene las indicaciones necesarias para reproducir el proyecto en otro entorno.