

LA VENGANZA DE SQL

El objetivo del proyecto es generar un programa que establezca una conexión con una BDD de MySQL. Representar y analizar los datos. Y realice una predicción utilizando el mejor modelo de Machine Learning que podáis entrenar.

Para ello tenéis que utilizar las librerías:

- Pandas
- Numpy
- Pymysql
- Visualización (dos a elegir):
 - Matplotlib
 - Plotly
 - Seaborn
- Gridsearch
- Algoritmos de ML (por lo menos 2):
 - Los que deseéis

La estructura del proyecto es la misma que en otras ocasiones. De forma similar, main.ipynb llama al resto de librerías, que son las que hacen el resto del trabajo. Todas estas librerías están en /src.

- /src
 - /sql_lib
 - sql_connector.py
 - /eda_lib
 - eda.py
 - /ml_lib
 - to_ML.py

Las librerías han de tener las siguientes funcionalidades:

- `sql_connector`: establece la conexión con la BDD y ordena a las tablas que se comporten como deseemos (queries, vamos). Debe ser capaz de transformar un objeto `sql` a `pandas` y de subir los datos resultantes a la BDD.
- `eda`: limpia y representa los datos. Los datos deben guardarse en formato `.png` en la carpeta `graphs`.
- `to_ML`: crea los modelos de ML, selecciona el mejor (vuestro criterio como data scientists) y realiza la predicción sobre las tablas.

Para conectaros a la BDD tenéis que utilizar el archivo `settings.json` incluido en la tarea. Y para que veáis que no somos **tan** malos, el main (que también os adjuntamos) viene con parte de la conexión establecida... el resto lo debéis buscar vosotros/as.

