



PROYECTO FINAL 2023

GOOGLE *YELP*

Datalogia

- Datalogia data consulting-Equipo
- Dante Alejandro Gomez – Data Analytics
- Hector Fernando Calisaya- Data Engineer
- Nicolas Fernando Tablon- Data Engineer
- Daniela Florencia Gutierrez- Data Science

Producto en desarrollo

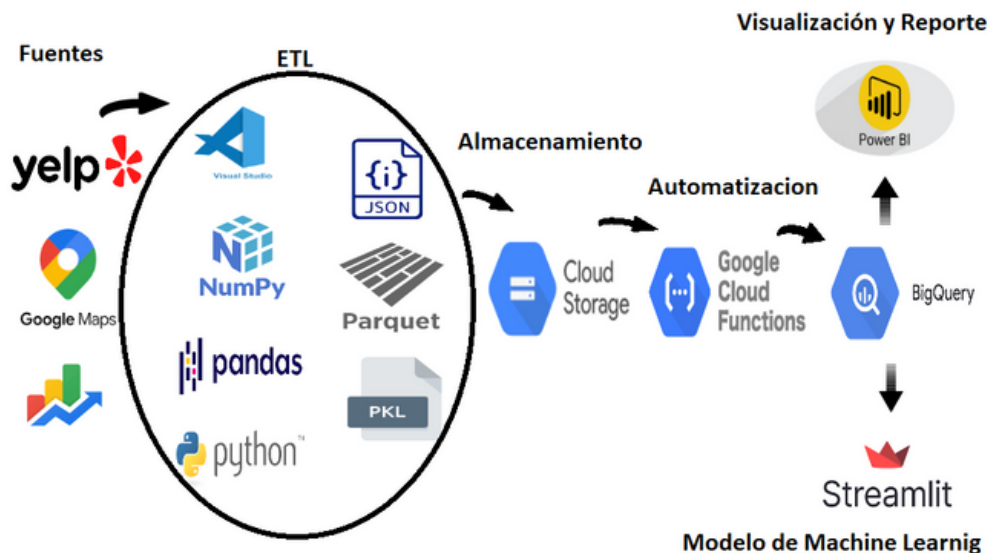
INFORMACION UTILIZADA:

Para el desarrollo del producto se empleo las siguientes carpetas y archivos:

- metadata_sitios
- review-estados /review-New_York
- business.pkl-Yelp
- review.json -Yelp
- tip.json -Yelp

MODELO DE DATA ESCOGIDO : DATA WAREHOUSE

Es una solución centralizada para almacenar, gestionar y analizar grandes volúmenes de datos.



EXTRACCION Y TRANSFORMACION DE DATOS

ALMACENAMIENTO DE DATOS

AUTOMATIZACION DE DATOS

MODELO DE MACHINE LEARNING

NLTK (Natural Language Toolkit) y Random Forest se utilizaran como modelos de machine learning usando los datos de almacenados en Google Cloud Storages.

DER(DIAGRAMA ENTIDAD - RELACION)