

# Proyecto 3: Tiendas Subway

## 1.- Descripción del proyecto

Este proyecto consiste en detectar todos los negocios que están en un radio de 2km de cada tienda subway en Estados Unidos (Se proporcionará un dataset que contiene todas las tiendas en EEUU), limpiar esa información y almacenarla. Para este proyecto habrá que utilizar la API de Google Maps u otra similar.

Este proyecto se centra en buena parte en utilizar Técnicas de ETL y capacidad de autoaprendizaje/exploración

**Dificultad:** Alta

**Complejidad:** Alta

## 2.- Requerimientos

El desarrollo del proyecto requiere que seamos capaces de plantear las soluciones, herramientas y descripciones requeridas para el desarrollo normal del proyecto y su puesta en marcha.

Debes pensar en todo lo que es necesario tener para que los datos sean Actualizables, Almacenables y Utilizables, es decir;

1. Pipeline que adquiera datos, los compare con los ya almacenados y los actualice si es necesario (Añadiendo nuevos o modificando la información de los que ya están almacenados).
2. Un almacén para los datos en el flujo de datos (temporales, crudos, etc)
3. Un almacén para los datos utilizables (Estandarizados, Limpiados, Listos para utilizar, etc). Este almacén debe contener todos los campos necesarios para poder llevar a cabo los análisis que requiere la marca.
4. Una descripción detallada de los componentes, campos (con significados, formatos, relaciones, etc)
5. Un documento descriptivo del funcionamiento de todo el proceso, que servidores, herramientas, scripts se utilizan, en qué orden y qué está conectado con qué.

## 3.- Fuentes de datos

Se debe utilizar como fuentes de datos los siguientes elementos:

1. CSV con las tiendas Subway de EEUU
2. Google Maps api (u otra a elección del alumno)

## 4.- A tener en cuenta

Se debe plantear un proyecto que tenga en cuenta las necesidades que tenga una marca para planificar dónde abrir nuevas tiendas, cerrar tiendas actuales o ampliar las que tiene. Para ello es necesario conocer con detalle todo el entorno de cada tienda, para ello realiza un análisis de los negocios (tiendas de comida o no), servicios (por ejemplo, colegios, dentistas, médicos, gimnasios, etc) que puedan ser de utilidad para planificar la apertura de otros locales o el cierre de actuales.

Debes almacenar toda esta información de tal manera que se pueda utilizar para realizar análisis por lo que la tarea de análisis de la información y diseño del datawarehouse es de gran importancia