**Proyecto: Análisis de Datos de Vuelos con PySpark**

En este ejercicio, usarás el conjunto de datos de vuelos de la Administración Federal de Aviación de los Estados Unidos (FAA) para realizar un análisis básico de los vuelos.

Asegúrate de tener instalado PySpark.

Obtén los Datos

Descarga un conjunto de datos de vuelos de la FAA. Aquí está un enlace a los datos: http://stat-computing.org/dataexpo/2009/the-data.html

Para este ejercicio, puedes descargar el conjunto de datos de 2008.

Lee los datos con PySpark.

Explorar los Datos

Explora los datos para familiarizarte con ellos. Aquí hay algunas cosas que podrías hacer:

Mostrar las primeras filas del objeto con los datos.

Ver la cantidad de filas y columnas del objeto con los datos.

Mostrar el esquema del objeto con los datos. Qué tipo de objeto es?

Limpieza de Datos

Asegúrate de que los datos estén limpios. Puedes necesitar hacer cosas como:

Eliminar filas con datos faltantes.

Convertir columnas a los tipos de datos correctos.

Análisis de Datos

Realiza un análisis básico de los datos.

¿Cuál es el vuelo más largo? Que unidad de tiempo es más apropiada? De qué depende? Seleccione una.

¿Cuál es la aerolínea con más vuelos?

¿Cuál es el destino más popular?

¿Qué día de la semana tiene más vuelos?