Push cleaned data frame to AWS

Get Cleaned Data

Use data for visualization

Load CSV data to Jupiter notebook / Google Colab from AWS database

Make Connection to Postgres and AWS using Spark

Join the tables

Load tables in Postgres

Clean data using Spark

Pull data from AWS to Jupyter notebook

Use data for Machine learning:

Linear regression

Random forest regressor

Neural Network

XGBoost

Link data to Tableau