ETL Project Proposal

Team IBM: Ivan, Bernardo, Meng-Yin

Date: 11/19/2020

ETL goal:

1. To read data from Kaggle ‘US Election 2020’ selected datasets
2. Perform data cleansing, data column selection, data column rename
3. Plan to use join to combine the dataset into a reportable data set for analysis

Proposal: US Election 2020

Data Source: Kaggle

Data sets:

1. Governors\_county.csv
2. Governors\_county\_candidate.csv
3. Governors\_state.csv

Data Extract/Transform/Load: Panda(read\_csv)/dataframe/to\_sql

Target Database: Plan to use PostGres as our target database