ETL Project Report: US Unemployment Rate for 2015

The original data is from two Kaggle dataset:

<https://www.kaggle.com/jayrav13/unemployment-by-county-us#output.csv>

<https://www.kaggle.com/muonneutrino/us-census-demographic-data>

The data represent the local Area (county-level) unemployment rate for 2015 except New York City.

Project Description:

Transforming the data using Python and Pandas per follow steps:

**E**xtract:

* + The Unemployment rate for each state and county is from csv file
  + The total, men and women population is from CSV file
  + Annual unemployment rate for each state and county is from json file. To extract the data annual rate from json I have used nested “For Loop”

**T**ransform:

* + Clean and structure the data for 2015 for each state and county

**L**oad:

* + Load the data into relational DB(Postgresql)
    - Advantages:
      * Organized
      * simplicity
      * Easy to mange
      * Generic language
      * Authorization and privilege control

Najmeh Dehghani