# PostgreSQL preparation

## Create the database, etl\_project.

# ETL Processes

## Unemployment Data from the Bureau of Labor Statistics

### Extract

#### Read data from <https://www.bls.gov/news.release/laus.t01.htm> to a Pandas DataFrame, unemployment\_df

### Transform

#### Use iloc to drop columns, only keeping the state and the 2018 unemployment rate (columns 0 and 9).

#### Rename columns to ‘state’ and ‘unemployment\_rate’.

#### Use astype to convert ‘unemployment\_rate’ to float datatype.

### Load

#### Create an SQLAlchemy engine object.

#### Write unemployment\_df to etl\_project database using Pandas to\_html

# Clean up the db schema

## Use ALTER TABLE ADD PRIMARY KEY statements to set primary keys on imported tables

# Create the view to show:

## State abbreviation

## Count of 2018 complaints

## Unemployment rate from 2018