Our original data was in .csv format, one came from the United States Library of Congress and one from the US Census.

We first imported both csvs and then imported another csv with the State name and abbreviations, so that we could get all data with the same format.
We merged the Census data (Uninsured Americans by state) with the State name/abbreviation csv. We then used .copy() to get the abbreviations and uninsured numbers by state.
Then we changed the column names to make more sense.

We then started cleaning the data of the US legislators, by using .copy() to get the party/state data in a data frame. We then used .query to check the number of legislators to make sure we had accurate data.
We then separated the data by "party"(Democrats, Republicans and Independents). We then used .grouby(state) to get the total legislators, of each party, by state.

Then we merged the dem_df and rep_df by state, and then merged the ind_df by state. Once the merge was complete we changed the NaN to 0 by using .fillna(0). Then we merged the legislator data with the uninsured data.
We then changed the States to the index, so we could drop the territories as they had no uninsured data, and their reps do not get a vote in the house or senate. We then replaced index back to numbers.
We followed up by adjusting the Uninsured Population from 'string' to 'num', and multiplied by 1,000 as the numbers were in the thousands.
We then changed all floats to integers. And began or Postgres process.

We decided to use Postgress because this offers a cleaner and easier way to view data.
For Postgress we created the table and used python to confirm the table and insert the information.