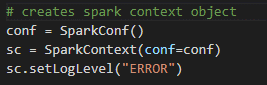
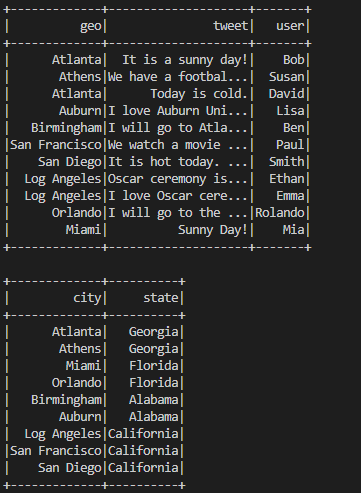
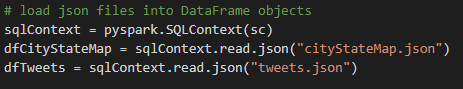
Tracy Michaels

Big Data Programming Assignment 5

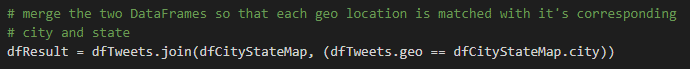
I first created a spark context object

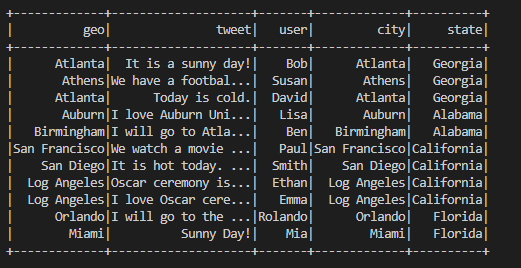


Then imported the data from provided json files and stored them in DataFrame objects

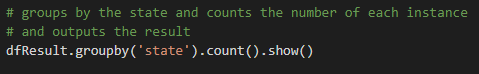


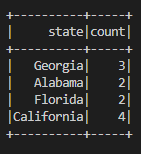
Next I merged the two dataframes into a 3rd so that matching items in the geo column of the Tweets dataframe and city column from the CityStateMap dataframe resulted in another DataFrame with where each Tweet is now associated with its appropriate state





Lastly to see the number of tweets published by state I grouped the above dataframe by state and counted each instance, then displayed the resulting table





Which is the result I was expecting and aligns with what the assignment prompt says should be the correct result