Brandon Jungen

Homework 3

Q1. California with 402 monitoring sites. I went through both the temperature and gas reading and made the state the key. Then the value was the state code, county code, and site number. Then in the reducer if it came across a new site in that state then it added it to a hash set and would not double count the site. My output had all states then I just sorted it with sort with bash and used head to get the top.

Q2. The east has a higher SO2 level at 5.4, while the west has an average of 2.2. I did this by finding states that were on the coast and then put those into the key that had the proper coast, so east or west. Then I had an average class that kept track of an average and the count, so number of times the average needed to be divided. So just calculated the average in the reducer.

Q3. 16:00 or 4:00 pm has the highest man SO2 level over all 20 years. To get this in the mapper I made sure that the year was 2000 – 2019. Then made the time the key for the reducer, and again used the average class so the reducer only needed to calculate the average for each time.

Q4. The SO2 levels over the last 40 years have gone down. The average in 1980 was 9.2, where in 2019 it was about 0.9. This was achieved by making the key from the mapper the data and then again using my average class. Then the reducer just had to calculate the average SO2 for each year. I came to this conclusion after looking at the SO2 levels sorted by year and while there were ups and downs per year, overall the SO2 did decrease.

Q5. The ten hottest states during summer months are Arizona, Texas, Nevada, Mississippi, Florida, Louisiana, Arkansas, Oklahoma, Kansas, and Georgia. This was done by the mapper making the state the key, then made sure that the month the reading was in a summer month, then sent that with the average class again to the reducer. Then the reducer just needed to calculate the average temperature and that was the average temperate for the summer months in those states. Then sorted the file on the temperature, and then took the first 10 results. Also dropped negative values and non-number values.

Q6. Using the states found in Q5 Arizona-6.2, Texas-2.7, Nevada-0.7, Mississippi-3.1, Florida-2.7, Lousisiana-3.5, Arkansas-2.6, Oklahoma-4.1, Kansas-2.8, and Georgia-3.4. I got there by making a hash set in the mapper that had the 10 states from Q5, then calculated their average SO2 levels using the average class that I made. Then the reducer just calculated the average SO2 levels for those states.