**Exploring Weather Trends Project 1**

The tools that I use: Excel

CSV files: city\_data.csv, city\_list.csv, global\_data.csv

**Steps and SQL that I used to extract the data from the data base**

1. Get the data from city list and discover what is the nearest city to me.

SELECT \*

From city\_list

The city nearest to me is San jose, california

1. Get the data from city\_data regarding city ”San jose”

SELECT \*

FROM city\_data

WHERE city = 'San Jose';

1. Get the global data.

SELECT \*

FROM global\_data;

Download all these to csv files and then open them up with Excel.

* What tools did you use for each step? (Python, SQL, Excel, etc)
  + How did you calculate the moving average?
  + What were your key considerations when deciding how to visualize the trends?
* **Line chart** with local and global temperature trends
* At least **four observations** about the similarities and/or differences in the trends
* **Below are the outline** of steps taken to prepare the data to be visualized in the chart

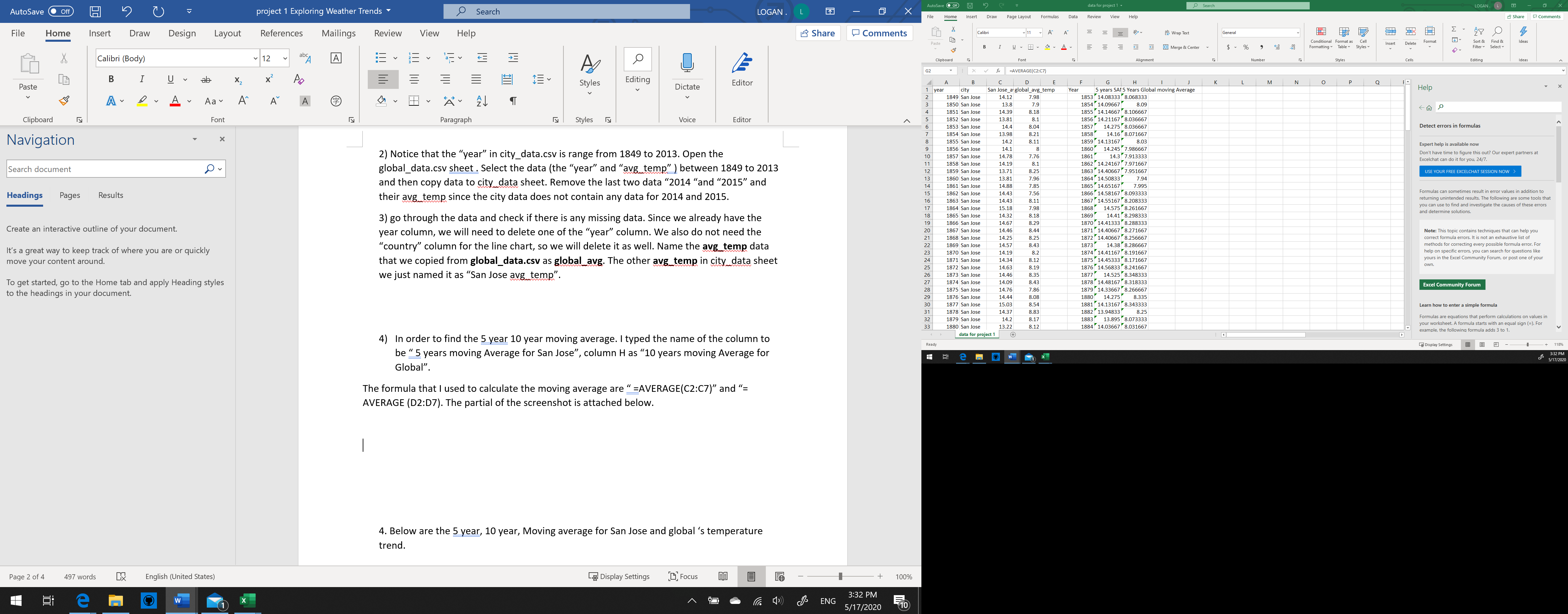
1) Open up the excel file city\_list.csv to look up the nearest city to me and find that the nearest city to me is “San Jose”. Then open up the city\_data.csv and global\_data.csv in excel.

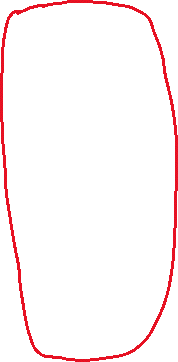
2) Notice that the “year” in city\_data.csv is range from 1849 to 2013. Open the global\_data.csv sheet . Select the data (the “year” and “avg\_temp” ) between 1849 to 2013 and then copy data to city\_data sheet. Remove the last two data “2014 “and “2015” and their avg\_temp since the city data does not contain any data for 2014 and 2015.

3) go through the data and check if there is any missing data. Since we already have the year column, we will need to delete one of the “year” column. We also do not need the “country” column for the line chart, so we will delete it as well. Name the **avg\_temp** data that we copied from **global\_data.csv** as **global\_avg**. The other **avg\_temp** in city\_data sheet we just named it as “San Jose avg\_temp”.

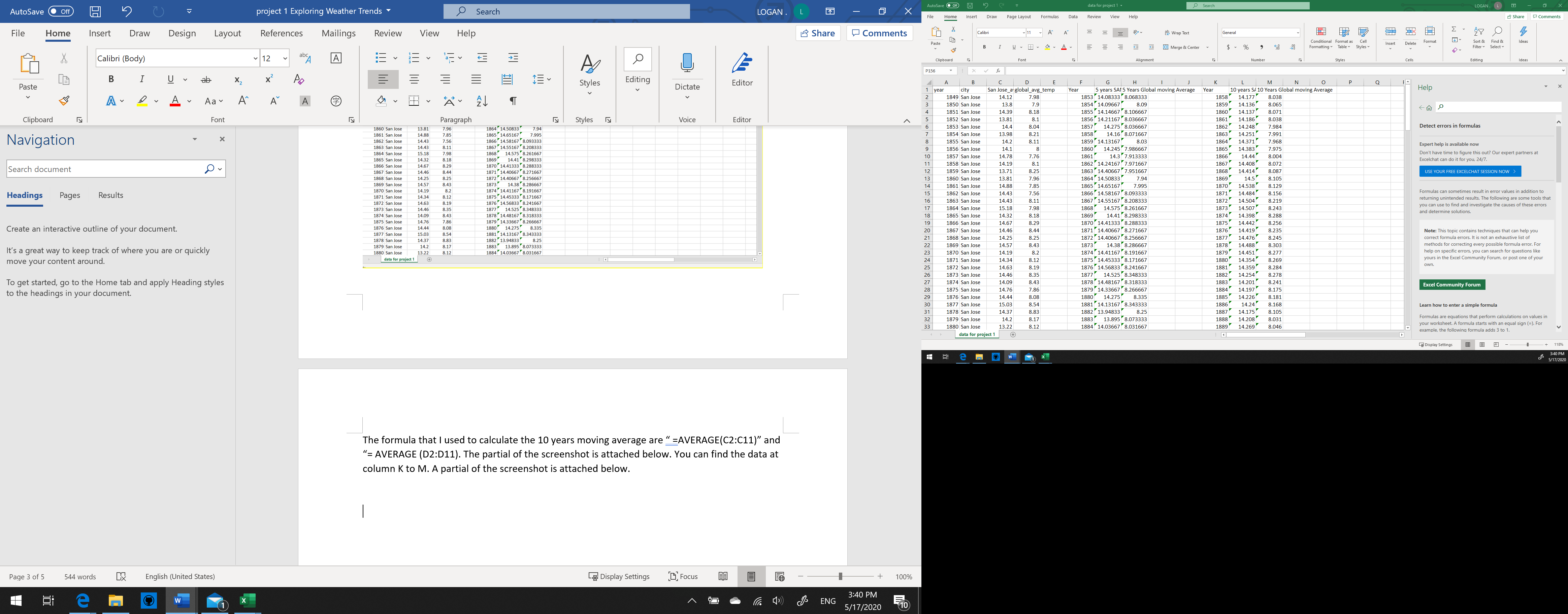
1. In order to find the 5 year 10 year moving average. I typed the name of the column to be “ 5 years moving Average for San Jose”, column H as “10 years moving Average for Global”.

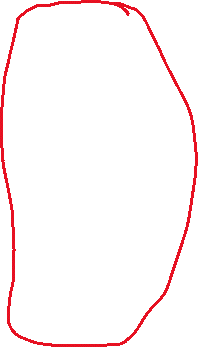
The formula that I used to calculate the 5 years moving average are “ =AVERAGE(C2:C7)” and “= AVERAGE (D2:D7). The partial of the screenshot is attached below.





The formula that I used to calculate the 10 years moving average are “ =AVERAGE(C2:C11)” and “= AVERAGE (D2:D11). The partial of the screenshot is attached below. You can find the data at column K to M. A partial of the screenshot is attached below.





5. Then based on the data, I created the following 5 years, 10 years, Moving average for San Jose and global ‘s temperature trend.

At least **four observations** about the similarities and/or differences in the trends

1. San Jose City overall has a lower temperature degree compare to the Global temperature.
2. The Temperature trend in the two sets of comparisons are steady growth.
3. The temperature trend between San Jose City and global is around 6 degrees.
4. The highest temperature trend in San Jose is in year 1898 with 14.5 degree and the global is in year 2009 9.6 degree