Data analyst project 1 explanation

1. Extract the data(SQL)

- Extract the data from USA, Boston

```
SELECT *
FROM city_data
WHERE city='Boston';
```

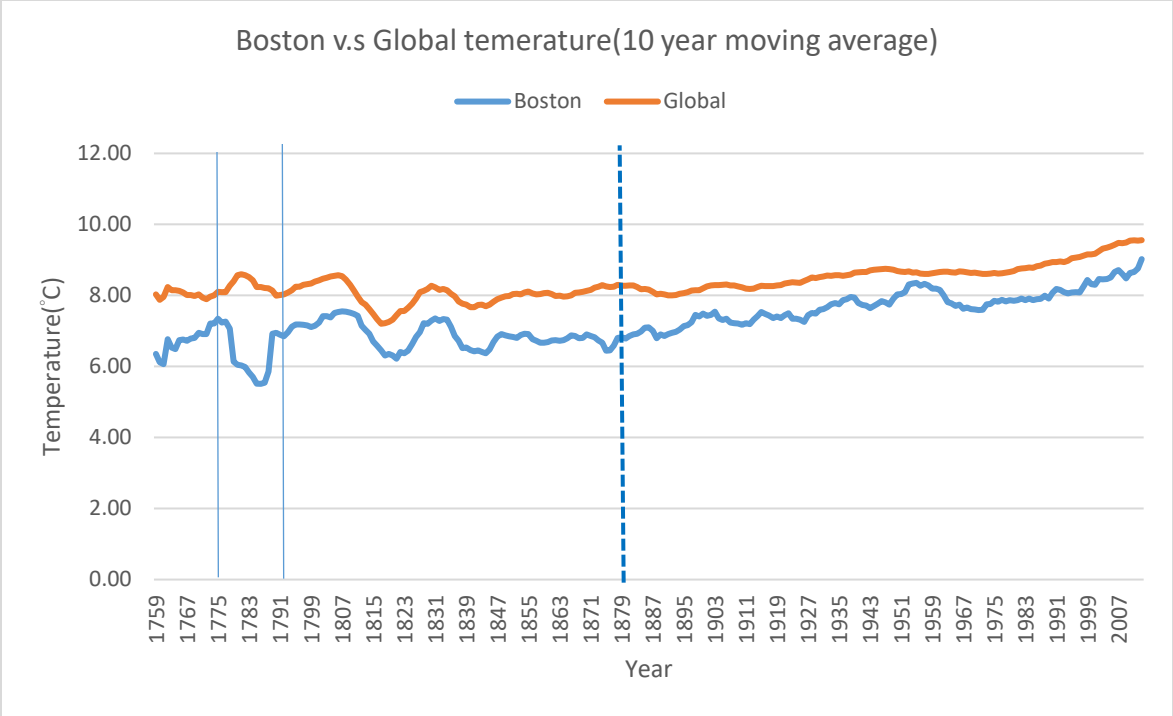
- Extract the data from global

```
SELECT *
FROM global_data;
```

2. An outline of steps(EXCEL)

- Export 2 CSV and created excel workbooks
- Combine global temperature and year data in Boston data excel book
- Compare which measures I could use, like average, 5 year moving average and 10 year moving average. Finally, I decided to use 10 year moving average since it can smooth the line chart and it would be easier to analysis.
- Boston data started in 1743 and global data started in 1750.Therefore, it was not reasonable to count past years in Boston data before 1750. On the other hand, I collected data starting in 1759 because of the attribute of moving average.
- Consider the way of visualization and I determine to use line chart because the data is related to time. And then I add axis labels and legends.

3. Line chart (EXCEL)



Observations(similarities):

- From 1759 to 2009, we could see the average temperature increased almost 2 degrees, whether from global or Boston data. Global data increased from 8 to 10 and Boston data went up from 6 to 8 respectively.
- If we draw a vertical line in around 1879 and divide two parts. It is interesting to see the trend before 1879 and after 1879. Before 1879, the temperature trend changed more radically than after 1879.

Observations(differences):

- Please look at the data between 1775 and 1791. Global temperature increase but Boston show the inverse. There might some situations happened between these years. In such almost 3 centuries, it is the only obvious difference trend between both.
- Normally, Boston is colder than most of the area in the rest of the world. It was down by 2 degrees in 18th century but it gradually got closer with global data in the 20th century.