Project 1 :

Explore Wheater Trends

Mathieu Rella

Summary

Data Extraction

Finding Nearest City

```
Select * from city_list
Where country = 'France';
```

This query result to show only one city, the french capital,Paris, city where I live in

Extract Paris average temperature per year in

```
Select year, avg_temp from city_data
Where city = 'Paris';
```

With Where condition = Paris we only need year and average temperature

Extract Global average temperature per year

```
Select * from global_data;
```

Here all the data from global_data will be useful

- Year
- Average Temperature

Data Processing

python

1. Import library

```
import pandas as pd
import matplotlib.pyplot as plt
```

For learning purpose I decide to use python, and two library : panda & matplotlib

2. Read CSV

```
#CSV Import

Global_Wheather = pd.read_csv('data/Global-Results.csv')
Paris_Wheather = pd.read_csv('data/Paris-Results.csv')
```

I import Global & Paris CSV file using read_csv from Panda using 2 variables.

3. Rolling Average

```
#Rolling Average

    #Paris Rolling Average
pr = pd.DataFrame(Paris_Wheather)
pr['pr10y-ma'] = pr.iloc[:,1].rolling(window=10).mean()

    #Global Rolling Average
gl = pd.DataFrame(Global_Wheather)
gl['gl10y-ma'] = gl.iloc[:,1].rolling(window=10).mean()
```

Create a rolling average column on 10 years on 2 new dataframe (pr:paris; gl:global), using .rolling function on the avg_temp column with .iloc

4. Dataframe merge

```
# Join both Dataframe
merged_inner = pd.merge(left=gl, right=pr, left_on='year', right_on='year')

merged_inner.rename(columns={'avg_temp_x':'Global_avg_temp',
                             'gl10y-ma':'Global-10y-ma',
                             'avg_temp_y':'Paris_avg_temp',
                             'pr10y-ma':'Paris-10y-ma'},
                             inplace=True)


merged_inner.shape
tp = merged_inner

del tp['Global_avg_temp']
del tp['Paris_avg_temp']
```

Merge pr & gl into merged_inner, I then renamed the columns for clarity, and dropped useless one (here : paris and global avg_temp) in a new variable tp

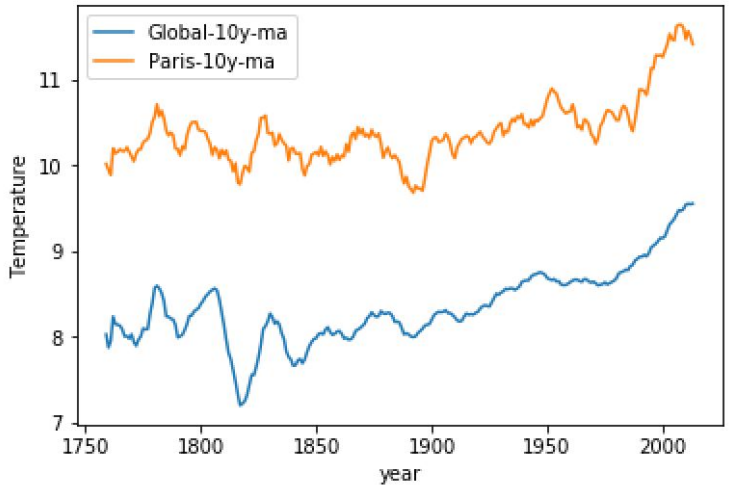
Graph Line

5. Graph Line

I made the average rolling Graph Line Using matplotlib library

```
# Graph Line

plt.figure(figsize=[15,10])
tp.plot(x='year')
plt.legend(loc=2)
plt.xlabel('year')
plt.ylabel('Temperature')
```



Observation

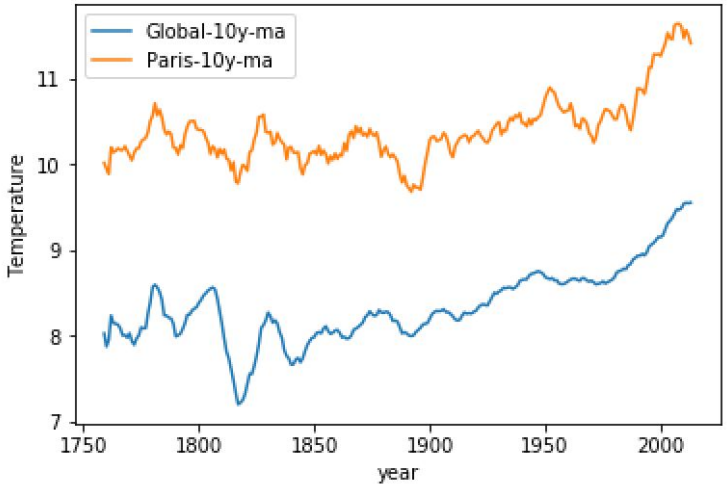
1. Paris is warmer than the global average temperature, and those of about 2 degrees Celsius.

2. Since 1750, This difference has been consistent

3. Paris and global average temperature follow the same trend with similar fluctuations.

Ex : Drop in 1816 due “to severe climate abnormalities know as the year without summer.” wikipedia

4. The world is getting hotter and hotter since 1750, the global average temperature increased by 1.5 degrees, this bullish trend is consistent



Thanks for your attention

Source code : Here