

Exploring Weather Trends - Project

Analysis

- CRITERIA
 - MEETS SPECIFICATIONS
- 1. Student is able to extract data from a database using SQL.
 - The SQL query used to extract the data is included.
 - The query runs without error and pulls the intended data.
- 1. Student is able to manipulate data in a spreadsheet or similar tool.
 - Moving averages are calculated to be used in the line chart.
- 1. Student is able to create a clear data visualization.
 - A line chart is included in the submission.
 - The chart and its axes have titles, and there's a clear legend (if applicable).
- 1. Student is able to interpret a data visualization.
 - The student includes four observations about their provided data visualization.
 - The four observations are accurate.

Importing libraries

```
In [38]: import pandas as pd
import numpy as np
```

1. Extraction of data from a database using SQL.

From city_data

```
select * from city_data
```

Input

HISTORY

MENU

SCHEMA

city_data

city_list

global_data

1

select * from city_data

Success!

EVALUATE

Output

71311 results

Download CSV

year	city	country	avg_temp
1849	Abidjan	Côte D'Ivoire	25.58
1850	Abidjan	Côte D'Ivoire	25.52
1851	Abidjan	Côte D'Ivoire	25.67
1852	Abidjan	Côte D'Ivoire	
1853	Abidjan	Côte D'Ivoire	
1854	Abidjan	Côte D'Ivoire	

From city_list

```
select * from city_list
```

Input

HISTORY

MENU

SCHEMA

city_data

city_list

global_data

1

select * from city_list

Success!

EVALUATE

Output

345 results

Download CSV

city	country
Abidjan	Côte D'Ivoire
Abu Dhabi	United Arab Emirates
Abuja	Nigeria
Accra	Ghana
Adana	Turkey

Adelaide

Australia

From global_data

```
select * from global_data
```

Input		HISTORY ▾	MENU ▾
SCHEMA	1	select * from global_data	
city_data	▾		
city_list	▾		
global_data	▾		
		EVALUATE	
Output		266 results	Download CSV
year	avg_temp		
1750	8.72		
1751	7.98		
1752	5.78		
1753	8.39		
1754	8.47		
1755	8.36		
1756	8.85		

2. Manipulation of data in a spreadsheet or similar tool.

Using Jupyter Notebook

```
In [39]: # Read city_data
city_data = pd.read_csv('city_data.csv', parse_dates=True)
city_data.head()
```

```
Out[39]:
```

	year	city	country	avg_temp
0	1849	Abidjan	Côte D'Ivoire	25.58
1	1850	Abidjan	Côte D'Ivoire	25.52
2	1851	Abidjan	Côte D'Ivoire	25.67
3	1856	Abidjan	Côte D'Ivoire	26.28
4	1857	Abidjan	Côte D'Ivoire	25.17

```
In [40]: # Read city_list
city_list = pd.read_csv('city_list.csv')
city_list.head()
```

```
Out[40]:
```

	city	country
0	Abidjan	Côte D'Ivoire
1	Abu Dhabi	United Arab Emirates
2	Abuja	Nigeria
3	Accra	Ghana
4	Adana	Turkey

```
In [41]: # Read global_data
global_data = pd.read_csv('global_data.csv')
global_data.head()
```

```
Out[41]:
```

	year	avg_temp
0	1750	8.72
1	1751	7.98
2	1752	5.78
3	1753	8.39
4	1754	8.47

```
In [42]: # Grouping global_data based on year
yearly_grouped = city_data.groupby('year')
yearly_grouped.groups

# Year wise temperature data
global_temp = yearly_grouped.mean()
global_temp.head()
```

```
Out[42]:
```

	avg_temp
year	
1743	4.725443
1744	11.230253
1745	1.678481
1750	10.424459
1751	10.727838

```
In [43]: # Data for Patna city
patna_temp = city_data[city_data['city']=='Patna']
patna_temp.head()
```

Out[43]:

	year	city	country	avg_temp
47467	1796	Patna	India	24.99
47468	1797	Patna	India	26.49
47469	1798	Patna	India	24.27
47470	1799	Patna	India	25.25
47471	1800	Patna	India	25.20

```
In [44]: # Year wise temperature data
patna_temp = patna_temp.iloc[:,[0,3]]
patna_temp.set_index('year', drop=True, inplace=True)
```

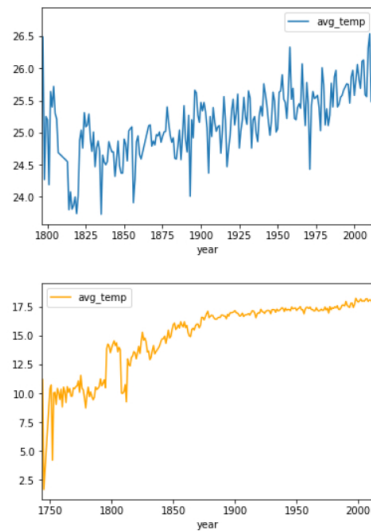
```
In [45]: patna_temp.head()
```

```
Out[45]:
```

	avg_temp
year	
1796	24.99
1797	26.49
1798	24.27
1799	25.25
1800	25.20

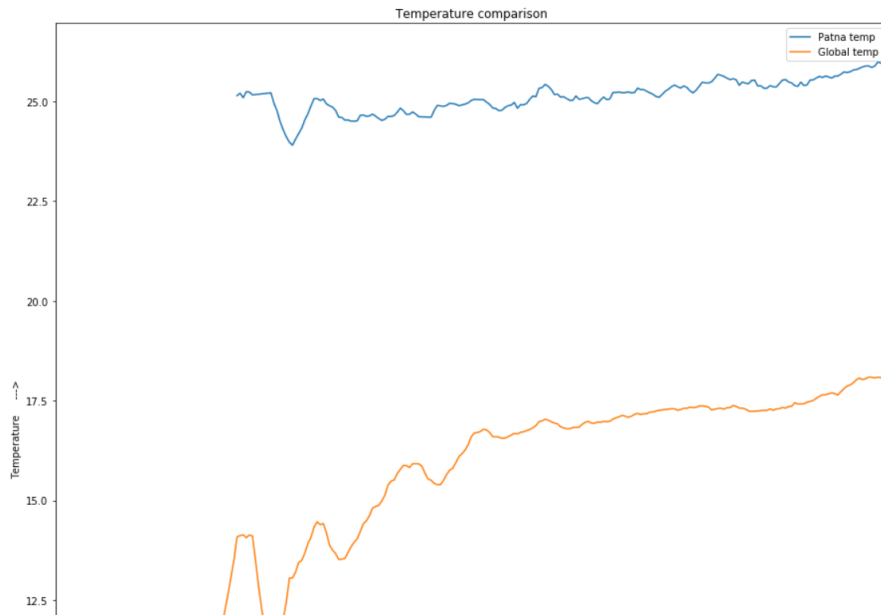
Plotting individual graphs for Patna city and global temperatures

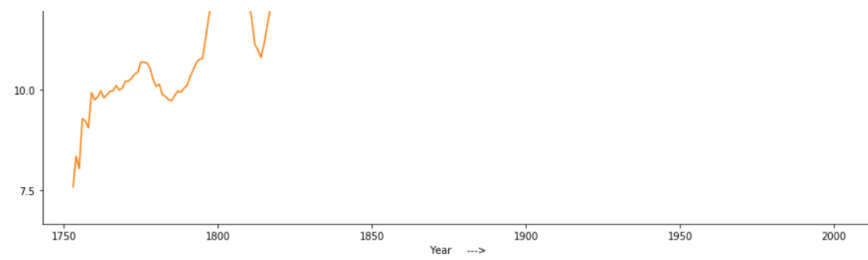
```
In [46]: patna_temp.plot(kind='line', label='Patna');
global_temp.plot(kind='line', color='orange');
```



3. Clear data visualization.

```
In [47]: # Plotting both plots on same graph
ax = patna_temp['avg_temp'].rolling(7).mean().plot(figsize=(15,15));
global_temp['avg_temp'].rolling(7).mean().plot(x='Year', legend=True);
# Legend(['a', 'b'])
ax.legend(['Patna temp', 'Global temp']);
ax.set_xlabel('Year --->');
ax.set_ylabel('Temperature --->');
ax.set_title('Temperature comparison');
```





4. Interpretation of above data visualization.

- Initial global temperature was very less
- Rate of increase of temperature is more in the global region
- The targeted region has very less significant difference in temperature considered over time
- Global temperature has changed a lot in the last 200 years
- The city taken is hotter than the average global temperature