

Exploring Wheather Trend

Table of Contents 1.Introduction 2.Data Wrangling 3.EDA 4.Conclusion

Introduction

In this project, we will analyze global and Toronto (where I live) temperature data and compare the temperature trends of Torontoto to overall global temperature trends.At the end we will be looking to answer these qustions- a.Is Toronto hotter or cooler on average compared to the global average? b.Has the difference been consistent over time? c.How do the changes Toronto temperatures over time compare to the changes in the global average?" d.What does the overall trend look like? Is the world getting hotter or cooler? Has the trend been consistent over the last few hundred years?

Tools used in the project:- SQL,Python,Pandas,Matplotlib

Data Wrangling

In [51]: *## Getting and expolring data*

First I have extracted data from database using SQL,Secondly I have observed d ata to understand how they are related to each other and afterthat I made following observations -
1. The 'year' column is the foreign key to join two tables city_data and globa l_data
2 .In global_data table there are 266 rows from year 1750 to 2013 and 2 colu mns 'year'& 'avg_temp'
AND in city data there are 271 rows from year 1743 to 2015 and 4 columns 'yea r','avg_temp','city','&'country'
3. After understanding the data, I wrote SQL query to get the required data fo r analysis-

The SQL query written to extract data-
select city.*,glob.avg_temp as avg_temp_global
from city_data as city inner join global_data as glob
ON city.year=glob.year
where country='Canada'and city='Toronto'
Order By year;

Then I have saved data as weather_data.csv on my desktop.

In [52]: *## Data cleaning*

```
In [53]: # Importing Packages
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [54]: # Loading data in pandas and displaying 5 rows
df=pd.read_csv('C:\\Users\\Sneha\\Desktop\\weather_data.csv')
df.head()
```

Out[54]:

	year	city	country	avg_temp	avg_temp_global
0	1750	Toronto	Canada	6.29	8.72
1	1751	Toronto	Canada	6.84	7.98
2	1752	Toronto	Canada	-1.10	5.78
3	1753	Toronto	Canada	5.76	8.39
4	1754	Toronto	Canada	5.94	8.47

```
In [55]: # Displaying last 5 rows
df.tail()
```

Out[55]:

	year	city	country	avg_temp	avg_temp_global
259	2009	Toronto	Canada	6.28	9.51
260	2010	Toronto	Canada	7.77	9.70
261	2011	Toronto	Canada	7.30	9.52
262	2012	Toronto	Canada	8.66	9.51
263	2013	Toronto	Canada	8.46	9.61

```
In [56]: # Rename column " avg_temp"
df.rename({'avg_temp':'avg_temp_toronto'},axis=1,inplace=True)
# set " year " as index
df.index=df['year']
# delete redundant columns
df.drop(['year','city','country'],axis=1,inplace=True)
# check the result
df.head()
```

Out[56]:

	avg_temp_toronto	avg_temp_global
year		
1750	6.29	8.72
1751	6.84	7.98
1752	-1.10	5.78
1753	5.76	8.39
1754	5.94	8.47

In [57]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 264 entries, 1750 to 2013
Data columns (total 2 columns):
avg_temp_toronto    264 non-null float64
avg_temp_global     264 non-null float64
dtypes: float64(2)
memory usage: 6.2 KB
```

In [58]: *# check null value*
`df.isnull().sum()`

Out[58]: avg_temp_toronto 0
 avg_temp_global 0
 dtype: int64

Exploratory Data analysis

In [59]: *# calculating 7 year moving average for average temp of toronto and global data and make it new column*
`df['moving_avg_toronto']=df['avg_temp_toronto'].rolling(7).mean()
df['moving_avg_global']=df['avg_temp_global'].rolling(7).mean()
df.head(11)`

Out[59]:

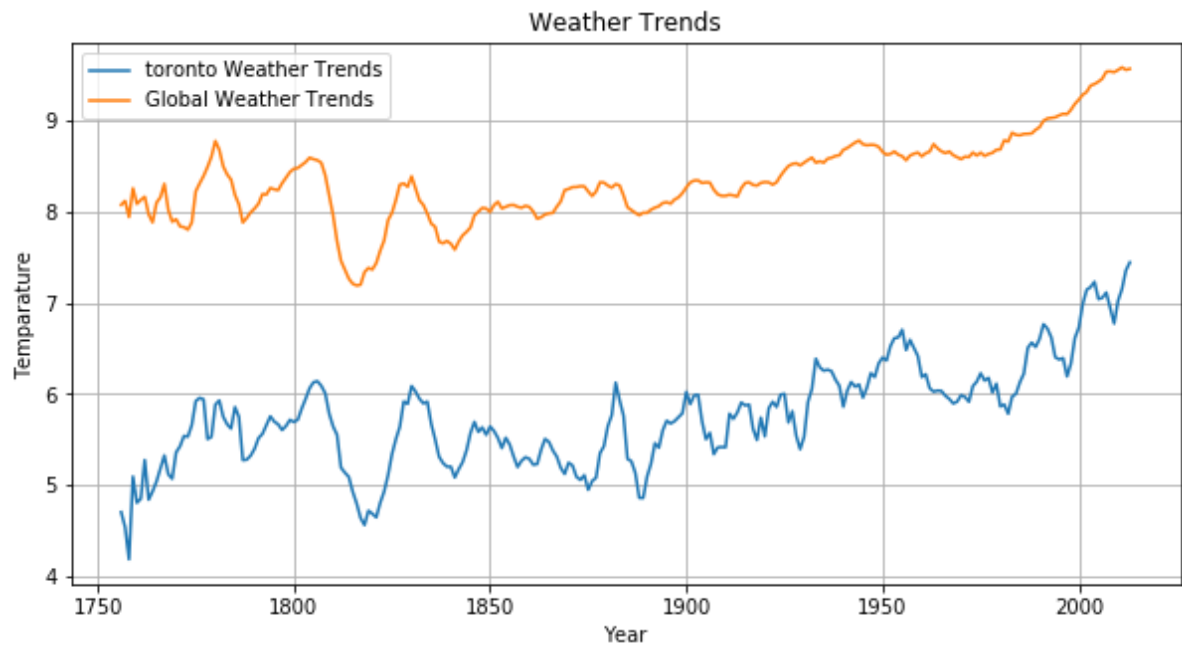
	avg_temp_toronto	avg_temp_global	moving_avg_toronto	moving_avg_global
year				
1750	6.29	8.72	NaN	NaN
1751	6.84	7.98	NaN	NaN
1752	-1.10	5.78	NaN	NaN
1753	5.76	8.39	NaN	NaN
1754	5.94	8.47	NaN	NaN
1755	2.81	8.36	NaN	NaN
1756	6.37	8.85	4.701429	8.078571
1757	5.13	9.02	4.535714	8.121429
1758	4.37	6.74	4.182857	7.944286
1759	5.27	7.99	5.092857	8.260000
1760	3.74	7.19	4.804286	8.088571

In [60]: *# Data Visualization*

```
In [61]: fig,ax = plt.subplots(figsize=(10,5))
ax.plot(df['moving_avg_toronto'], label='toronto Weather Trends')
ax.plot(df['moving_avg_global'], label='Global Weather Trends')

ax.legend()
ax.grid(True)
ax.set_xlabel('Year')
ax.set_ylabel('Temperature')
ax.set_title('Weather Trends')
```

Out[61]: Text(0.5,1,'Weather Trends')



Conclusion

1. As part of this project ,I have analyzed the weather trend using temperature data of Toronto and global from year 1750 to 2013 and concluded that Toronto is cooler than the World.The average temperature of global is more than average temperature of Toronto.
2. Toronto has higher fluctuation in average temperature compare to global average temperature. 3.However difference in both average temperature has been consistent over the time and the temperature for both Toronto and global are increasing over years particularly in recent years. 4.Furthermore, based on historical trends, we can predict that the future trend will be continually increasing.