Declaration on Plagiarism

This form must be filled in and completed by the student(s) submitting an assignment

Name:	Palash Dange
Student Number:	19210749
Programme:	MCM
Module Code:	CA682
Assignment Title:	Data Visualisation
Submission Date:	16 Dec 2019
Module Coordinator:	Dr Suzanne Little

I declare that this material, which I now submit for assessment, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work. I understand that plagiarism, collusion, and copying are grave and serious offences in the university and accept the penalties that would be imposed should I engage in plagiarism, collusion or copying. I have read and understood the Assignment Regulations. I have identified and included the source of all facts, ideas, opinions, and viewpoints of others in the assignment references. Direct quotations from books, journal articles, internet sources, module text, or any other source whatsoever are acknowledged and the source cited are identified in the assignment references. This assignment, or any part of it, has not been previously submitted by me or any other person for assessment on this or any other course of study.

I have read and understood the referencing guidelines found at
http://www.dcu.ie/info/regulations/plagiarism.shtml,
https://www4.dcu.ie/students/az/plagiarism and/or recommended in the assignment guidelines

Name: Palash Dange   Date: 16-Dec-2019

GLOBAL TEMPERATURE ANALYSIS

Abstract

Today many scientists are working on analysing the global temperature and collecting effective conclusions on it. Earth surface temperature is rising, it may be because of natural or human activity. Why do we worry about the global temperature increase even by 1 degree, though temperature fluctuates by a few degrees daily where we live? This is because it takes a lot of heat to warm up the land and ocean altogether over the world. An increase in temperature of 1 degree overall in a century is a matter of a big concern for the earth.

Data collected over time for temperature are used to analyse and create graphs. Around the 1800s, after the industrial revolution started, numerous chemical emissions and wastage led to an increase in the level of CO_2 around the world. An increase in the level of CO_2 is the main cause of global warming and the Green House Effect. The greenhouse effect traps the warm radiations from the sun and does not let it escape from an atmosphere of earth leading to an increase in global temperature.

Various graph from the dataset portrays shows the trend of increase in global temperature overall in several countries since 1850.

1. Dataset

This dataset used to generate graphs consists of data over the years from 1850 to 2015. The dataset comprises of the values of land temperature of different countries. This dataset includes values of land and ocean temperature with its uncertainty. The data ranges from the year 1850 to 2015, this dataset has data concerning dominant countries and cities. They gathered the data before the 1940s by using mercury thermometers. After the 1940s, they moved several weather stations because of colonization and the building of airports.

They repackage this data from a fresher dataset put up by Berkeley Earth; it is a non-profit climate analysis project majorly working on land temperature. The principal focus was to pinpoint specific concerns associated with global warming and temperature record. The dataset includes numerous files that are used to generate a graph with the help of python and Tableau. The size of the dataset is around 60 MB.

Various columns in dataset is:

- Date
- LandAverageTemperatureUncertainty
- LandAverageTemperature
- Average Temperature
- Average Temperature Uncertainity
- Country.

2. Data Exploration & Data Cleaning:

- Dataset has various files in which include columns and rows with several features like Average Temperature and Uncertainty, Countries Cities & etc. This data consist a total of 5 lac rows with the help of pandas and numpy, the rows which had no temperature were removed.

- Removing the null values was necessary as it might show 0 values for several years making the graph unreadable. Using the command fillna() and dropna() null rows were successfully removed.

3. Visualisation

→ **First graph:** A time-series graph that shows global Average temperature and its uncertainty from 1850 till 2015. This data has three ongoing lines which are Max Possible Temperature, Min Possible Temperature and Average temperature in the year.

- **Colors** used are blue for max, yellow for low and brown.
- The yellow color fill of the graph was necessary to make the graph more **readable**.
- The **tooltip** of the graph includes all three at a certain year highlighting the year at the x-axis.
- For all 3 lines, the temperature is marked on the y-axis to notice the change.
- X and Y axis are properly labeled with **tick marks**.
- Toggle Spike lines are used to exactly **highlight** the values
- Graph produced with plotly can be used to zoom and get values of each year. (Refer Video for further explanation).
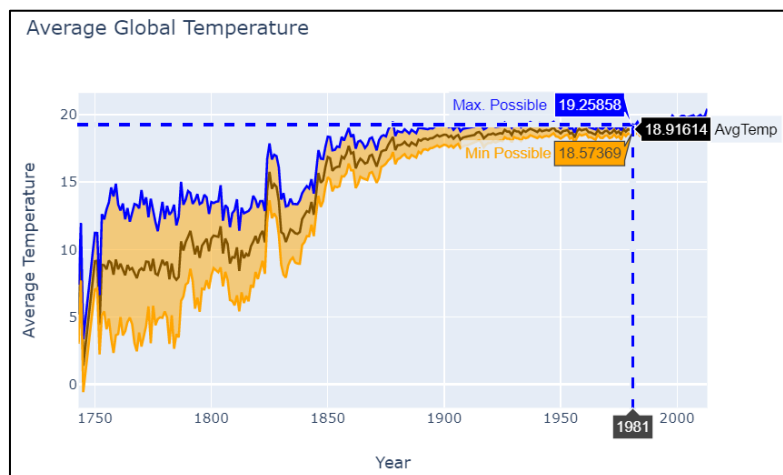

Fig 1.1

→ *Second graph*: This graph correlates Land temperature the year is now divided into 4 seasons Winter, Summer, Spring, and Autumn. This is a scatter plot that shows the increase of temperature over the season from 1850 to 2015.

- **Spring season** ranges from month March to May.
- **Summer season** range: June to Aug

- **Autumn season** range: September to Dec
- **Winter Season** range: Jan to March

Colors used:

- Orange to identify warmth of summer
- Green to identify the spring season
- Autumns are indicated using red color
- Winter is highlighted by the blue shade.

Grid background is used with the X and Y axis which are designated properly as per the graph data. These axes are marked with specific tick marks of interval 20 years, so the aggregation has number of points for analysis.
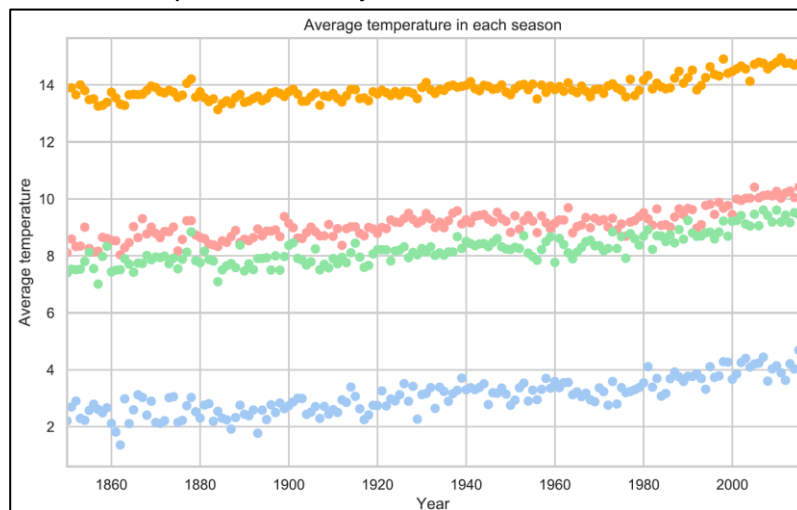

Fig. 1.2

→ *Third graph* Tableau animated graph that shows how the temperature changes over the period with colours for different countries from the period of 1850 to 2015. (Refer video for explanation)

Libraries Used: Plotly and Seaborn
Tools used: Tableau
IDE – Visual Studio Code

4. __Conclusion__

Temperature change over the years can be seen in the time series graph after the 1850s there is a drastic change in the increasing trend of temperature. Numerous factors like the industrial revolution, chemical and electrical advancement around the world led to an increase in global temperature. The use of electrical thermometers has led to a decrease in the uncertainty of temperature. This means the data recorded/predicted was approximately the same as the actual temperature, from the graph we can observe, temperature before the 1950s was not predicted accurately and had a large deviation from the actual temperature. The reason behind this is the use of mercury thermometers instead of an electrical thermometer.

Climate is affected by a lot of characteristics and not just temperature. Climate is also affected by the amount of rainfall, air pressure, and various other factors. Limitation of this dataset only to temperature and its uncertainty abide us from making exact conclusions over the global climate of the earth. As we can see from the graph there has been an increase of 10-12 degree rise of temperature as per this dataset, which has affected the climate. Repercussions of which can be seen today as well. Pollution and depletion of the ozone layer are a few major issues.

As 3D graph shows the exact visualization and is easy to understand at times. Because the data only had a temperature variable it was technically impossible to create a 3d graph. Though I was able to visualize countries with the time change on the world map with Tableau.

References

1. Dataset link: https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data
2. Time Series: https://www.statisticshowto.datasciencecentral.com/timeplot/
3. Seaborn Library: https://python-graph-gallery.com/seaborn/
4. Visualization with Seaborn: https://jakevdp.github.io/PythonDataScienceHandbook/04.14-visualization-with-seaborn.html
5. Tableau Tutorials : https://www.tutorialspoint.com/tableau/index.htm
6. Global Temperature : https://www.ncdc.noaa.gov/global-warming/last-2000-years
7. Climate Change past 2000 years : https://www.bbc.com/news/science-environment-49086783