project-group2-proposal

October 29, 2021

1 CSCI 2000U - Scientific Data Analysis

1.1 Final Project Proposal

Group/Individual Names:

Adam Callahan, Trevor Lee, Angus Murchison

1.2 The dataset

Temperature Change

Source: https://www.kaggle.com/sevgisarac/temperature-change

The FAOSTAT Temperature Change data set is a collection of the average change of temperature of 190 countries and 37 territories, in Celsius, recorded monthly from 1961-2019, giving the annual, seasonal and monthly temperature outliers with respect to the 1951-1980 recordings.

- Area Code - The numerical code of area column, type of area code is an integer
- Area - Countries and Territories (In 2019: 190 countries and 37 other territorial entities.), type of area is an object
- Months Code - The numerical code of months column, type of months code is an integer
- Months - Months, Seasons, Meteorological year, type of months is an object
- Element Code - The numerical code of element column, type of element code is an integer
- Element - 'Temperature change', 'Standard Deviation', type of element is an object
- Unit - Celsius degrees °C, type of unit is an object
- #Y1961 - Mean Surface Temperature change in the year 1961
- #Y1962 - Mean Surface Temperature change in the year 1962
- ...
- #Y2019 - Mean Surface Temperature change in the year 2019

Motivation

Climate change is a very critical and ever-looming issue which many of us remain powerless over and can do nothing about but watch. It is because of this existential threatening nature, that makes it a very rewarding and reasonable data set to be interested in and analyze. We hope our findings from our results can give us some insight into what geographical and societal problem properties of countries may be in any way correlated to climate change and prove the threat this issue poses by showing what the future may look like if nothing is done about it. More specifically, with this data, we plan to derive how climate fluctuates from period to period depending on the geographical location and how much the climate has changed as the global population and pollution has increased. Some other datasets we were interested in were Chess games datasets, popular music

datasets, and solar power generation datasets however this temperature change dataset stuck out to us for previously mentioned reasons.

Code

Import data

```
[1]: #importing used libraries
     import csv
     import re
     from functools import reduce
     import numpy as np

     # this aux function reads the CSV file and returns the data in a Python␣
      ↪dictionary
     def get_data_csv():
         collection = []
         with open('Environment_Temperature_change_E_All_Data_NOFLAG.csv', 'r') as f:
             for line in csv.DictReader(f):
                 collection.append(line)
             return collection

     # the data
     data = get_data_csv()
```

Getting to know your data

How many records are there?

```
[2]: #len()/2 because there is a line for both temperature and standard deviation␣
      ↪for every row
     print("There are", len(data)/2, "records")
```

There are 4828.0 records

How many unique values are there?

```
[3]: #use sets to hold only unique values, then get length
     areas = set()
     for i in range(len(data)):
         areas.add(data[i]['Area'])

     print("There are", len(areas), "unique area values")
     #areas
```

There are 284 unique area values

What is the date range?

The data ranges from all months/seasons/full years for 58 years from 1961 to 2019

How many months/seasons were recorded per year?

2

```
[4]: month_codes = set()
     for i in range(len(data)):
         month_codes.add(data[i]['Months'])


     print("There are", len(month_codes), "recorded time periods per year (some␣
       ↪overlap)")
```

There are 17 recorded time periods per year (some overlap)

Questions

1. How much of a difference, if any, has climate change affected the yearly temperature fluctu-
   ations of first-world countries compared to third-world countries? Using Canada, the United
   Kingdom, and Japan for first-world countries and Afghanistan, Madagascar, and Bolivia for
   the third-world countries.
   - This can be achieved by analyzing each country's average standard temperature de-
     viation for the time periods of 1961-1966 and 2014-2019. From here we can perform
     the following analysis for each time period then compare the results to see how climate
     change has affected yearly temperature fluctuations and if overall country well-being
     plays any role in it:
     – We can check the 2 groups independently to see if there is any correlation between
       the standard temperature deviation between the countries.
     – We can compare the 2 groups together by comparing each group's average standard
       temperature deviation.
2. How much of a difference, if any, has climate change affected the yearly temperature fluc-
   tuations of northern countries compared to countries along the equator? Using Colombia,
   Kenya, and Indonesia.
   - This can be achieved by analyzing each country's average standard temperature de-
     viation for the time periods of 1961-1966 and 2014-2019. From here we can perform
     the following analysis for each time period then compare the results to see how climate
     change has affected yearly temperature fluctuations and if positive latitude plays any
     role in it:
     – We can check the 2 groups independently to see if there is any correlation between
       the standard temperature deviation between the countries.
     – We can compare the 2 groups together by comparing each group's average standard
       temperature deviation.
3. How have the average temperature fluctuations of Canada's seasons changed over the years?
   What does this say about our future?
   - This can be achieved by creating a graph of Canada's standard temperature deviation
     vs year for each of the seasons. From here we can do a number of things:
     – We can identify if any of the seasons have a correlation with their standard temper-
       ature deviation over time.
     – We can identify a trend line to predict what temperature fluctuation we can expect
       in the near future.
     – We can compare the different season's graphs to see how climate change affects the
       different seasons.
4. How have yearly temperature fluctuations changed over time for today's most polluted coun-
   tries? Using Bangladesh, Pakistan, and India

- This can be achieved by creating a graph of yearly standard temperature deviation vs year for each of the countries. From here we can do a number of things:
  - We can identify a trend line to predict what temperature fluctuations we can expect in the near future for each country.
  - We can compare the different country's graphs to see how climate change affects polluted countries.

5. How have yearly temperature fluctuations changed over time for today's most populated countries? Using China, India, and the United States.
   - This can be achieved by creating a graph of yearly standard temperature deviation vs year for each of the countries. From here we can do a number of things:
     - We can identify a trend line to predict what temperature fluctuations we can expect in the near future for each country.
     - We can compare the different country's graphs to see how climate change affects countries with high populations.

Potential

There is potential to help analyze factors and trends that could be contributing to climate change like location, population, development, etc.