Team members: Sowmya Chandrasekaran, Yeshwanth Somu, Emily Robles

GitHub Repo:
https://github.com/orgs/UC-Berkeley-I-School/teams/datasci_w200_project2_emily_sowmya_yeshwanth

Dataset Ideas

Berkeley Earth Temperature Data (time series)
http://berkeleyearth.org/data/

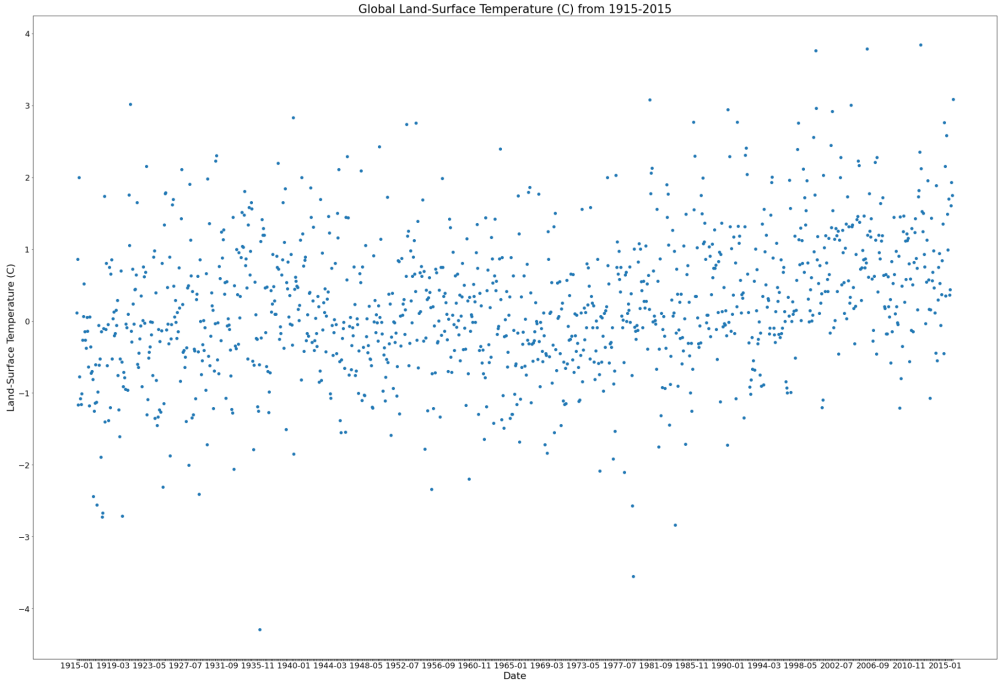
Figures

- Temps by month for each (whether comparing cities or countries)
    - Compare yearly w/different colors

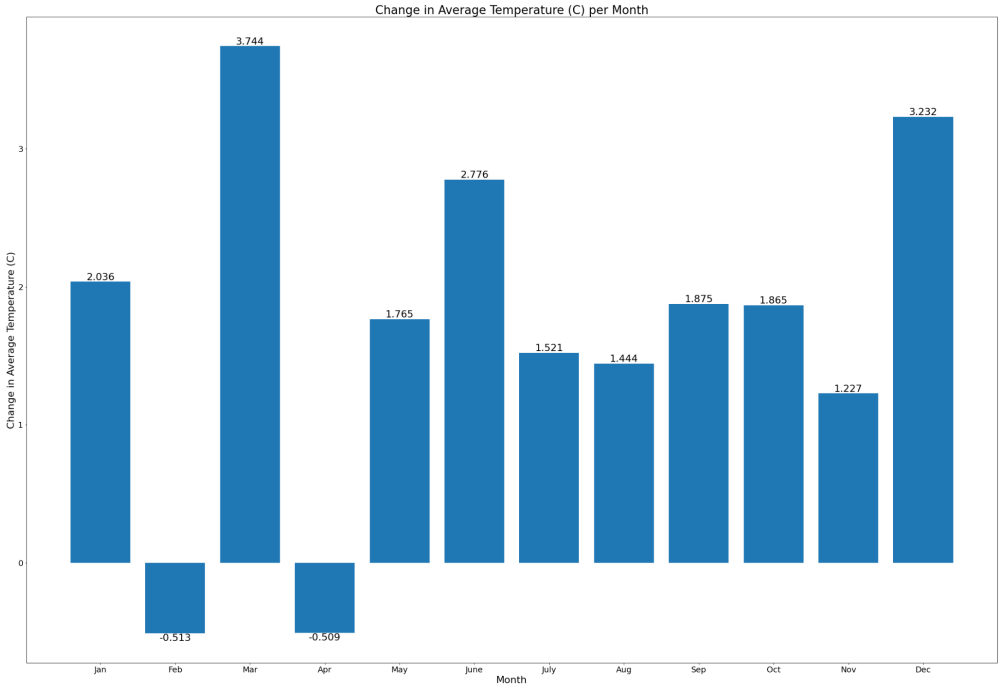
A primary dataset you intend to analyze,
Berkeley Earth Data United States monthly average temperature from 1915 - 2015.
http://berkeleyearth.lbl.gov/auto/Regional/TAVG/Text/united-states-TAVG-Trend.txt

Initial plots, figures, or tables

1. Scatter plot: Average temperature (in degrees Celsius) by month for every 10 years (1915-2015)



2. Bar graph or table: For each month, difference in average temp from 1915 start to 2015 finish



Some of the variables (column names) you intend to explore and what kind of insights you expect to glean

We will use the **Year**, **Month**, and **Monthly Anomaly** (& uncertainty) columns, in combination with the given estimated Absolute Temperatures, to derive monthly estimated temperatures for the years 1915-2015. We expect to be able to see a trend in monthly temperatures across the 100 year period, and intend to compare those with global trends from the global average dataset (supplemental dataset).

Using both of these datasets, we'd like to see how the increase or decrease in average monthly temperatures in the United States compares to the increases or decreases in temperature at the global scale.

Supplemental datasets:

Berkeley Earth Data Global land-surface average temperature dataset:
http://berkeleyearth.lbl.gov/auto/Global/Complete_TAVG_complete.txt

We will use this global dataset to compare average monthly temperatures from 1915-2015. The tables will be joined using the Year and Month columns.

Once we have joined these datasets, we plan to generate the following figures:

1. Table of average temperatures by month for both US and Global
2. Graph of difference in 2015 and 1915 US average temp and global average temp for each month
    a. Intended to answer the question "What is the average difference in monthly temperatures from 1915-2015 at both the US and global scale?"

What you plan to cover in the final report and how you plan to organize it.

We plan to first discuss our interest in global temperatures, then introduce the Berkeley Earth project and motivations behind it and the methods of bias correcting and quality control performed by the project before releasing the dataset.

We will then give an overview of our findings organized by the following main questions:

1. “How has the average monthly temperature changed in the US from 1915 - 2020?”
2. “How has the average monthly temperature changed at the global scale from 1915 - 2020?”
3. “What is the average difference in monthly temperatures from 1915-2020 at both the US and global scale?”

We’ll include our tables and figures in their respective sections. As we analyze the data, we may include smaller sub questions within each section. We will also detail our methods of data analysis that led to each finding and any translations that we performed on the data (for example, we expect to translate the Monthly Anomaly data to degrees C using the absolute temperatures given in beginning of the dataset).