<u>**Project Proposal**</u>

<u>**Climate Visualizer**</u>

Jacob Blomquist and William Ludwig

Basic Info:

- Title: Climate Visualizer
- Members:
    - Jacob Blomquist (U0843402)
    - William Ludwig (U0905218)
- Email Addresses:
    - Jacobrb95@gmail.com
    - williampludwig@gmail.com
- Link:
    - https://github.com/WLudwig/Climate_Visualizer

Background and Motivation:

While searching for ideas for our visualization project, we stumbled upon an extensive dataset provided by NOAA (National Oceanic and Atmospheric Administration). This dataset contains historical climate data as well as several other large data sets. We are focusing on the GHCN (Global Historical Climatology Network) which contains daily temperature, precipitation, wind, pressure, and other useful and interesting data from climate stations around the world.

We hope to make a visualization using this data to show trends in climate. This data will be able to help show climate trends over periods of time and how the climate may be changing in particular areas. This dataset has enough data that we will need to use multiple visualizations in order to display all of the data. We hope to make a visualization which will allow a user to quickly see trends from any station they choose.

Project Objectives:

With our visualization we are wanting to allow a user to visually see the climate data for the selected area/weather station. They will be able to interact with the visualizations by selecting the area/weather station they want to see data for. They will also be able to select from a drop down what specific area they want to see data for such as: temperature, rainfall, etc. All of these visualizations will be linked together and

when you interact with one view the other ones will update as well.

With our project we want to make it so that users can easily see climate data from anywhere in the USA and so that they can easily select and view whatever category they are interested in. We also want to make it so that they can see how these categories are changing and/or have changed over time.

In this project we would like to expand on the skills that we have learned throughout this course and we want to learn how to effectively manipulate/select data so that we can get and use only the data that we need. The dataset that we are pulling data from has a lot more information than what we are actually needing for our visualization. In our assignments for this class we linked together two different views, but for our project we are wanting to link together at least 3 different views. Also neither of us are very good at making things look pretty so we are looking forward to the opportunity to learn more about how to make our visualizations and webpages more visually appealing.

The benefits of our project include the ability to easily and effectively view current and past data for a certain area/weather station. We are also planning to include a comparison view where you can select multiple weather stations at once and compare them side by side. This visualization could be very beneficial to people who are wanting to see how climate has changed over the years, i.e. people who study global warming, etc.

Data:

Our data is taken from the Global Historical Climatology Network (NOAA). The link for our data is ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/. This data shows high/low temperatures, precipitation, pressure, wind, and more for every day going back to ~1900 for weather stations across the whole world. This is a ton of data so we will be filtering it down and focusing mainly on the United States and on certain categories. The full dataset is roughly 28.3 GB in size. So as we mentioned we will be filtering the data and only using a certain portion of it.

Data Processing:

We are expecting to do a lot of data cleanup. As mentioned in the Data section the full dataset is roughly 28.3 GB in size. So we are going to clean up the data and

only grab the data for the weather stations inside of the USA. We also need to process the data and put it into an easy to use/read format such as JSON or CSV. We will need to do even more clean up on the data because there are a lot of measurements in the data that we will not be using such as: soil temperature, cloudiness, direction of fastest wind, etc.

We plan on processing our data and formatting it with Python. We are thinking that Python is the best option for us to pull the data out of that huge dataset and to put it into a more friendly format that we will be able to use when we are creating our visualization.
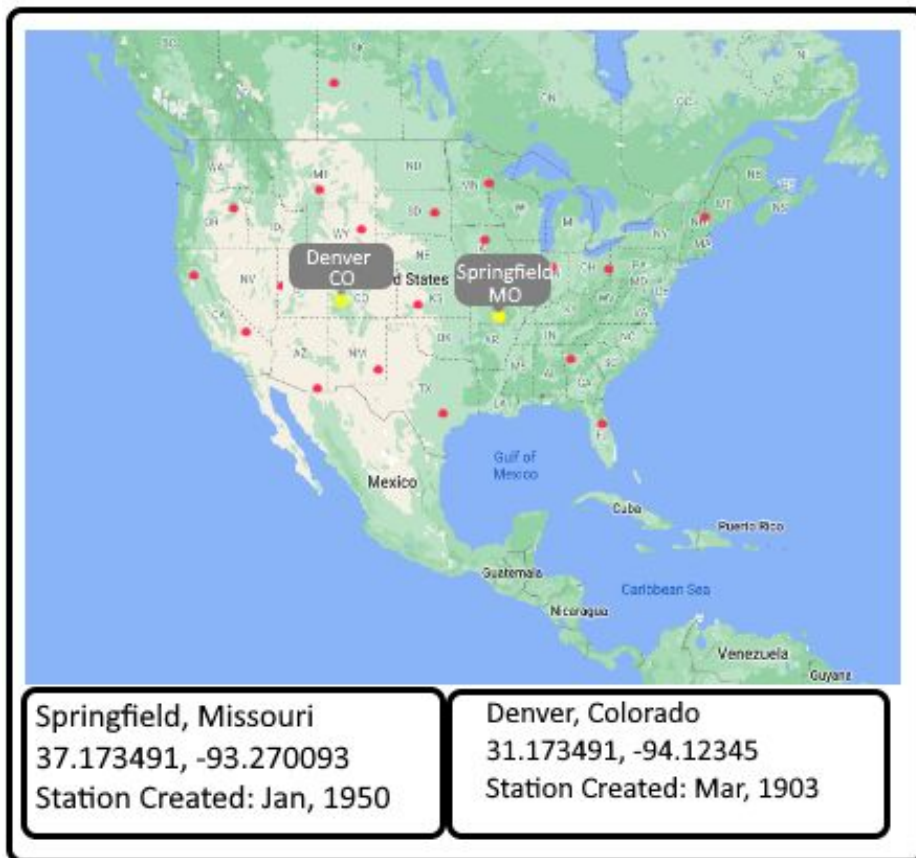
Visualization Design:

- Displaying of Data:
    - Data will be displayed in line charts, individual weather stations will be on an interactive map. Data displayed could include temperature (max,min,average) over time, precipitation, snowfall, snow depth, soil temperature, pressure, wind speed, and much more.
- General Ideas:
    - User selects a weather station and can browse the station's data.
    - User can compare two weather stations.
    - User can select a range of dates to finely tune their data view.
    - User can select more than two stations to get a general comparison between each of them.
        - This can be accomplished by using a brush.

- Prototype design #1:
  - Selecting one weather station and showing the associated data (we did not draw all the graphs/categories that will be shown. Just enough to give an idea of what it will look like)

Springfield, Missouri
37.173491, -93.270093
Station Created: Jan, 1950
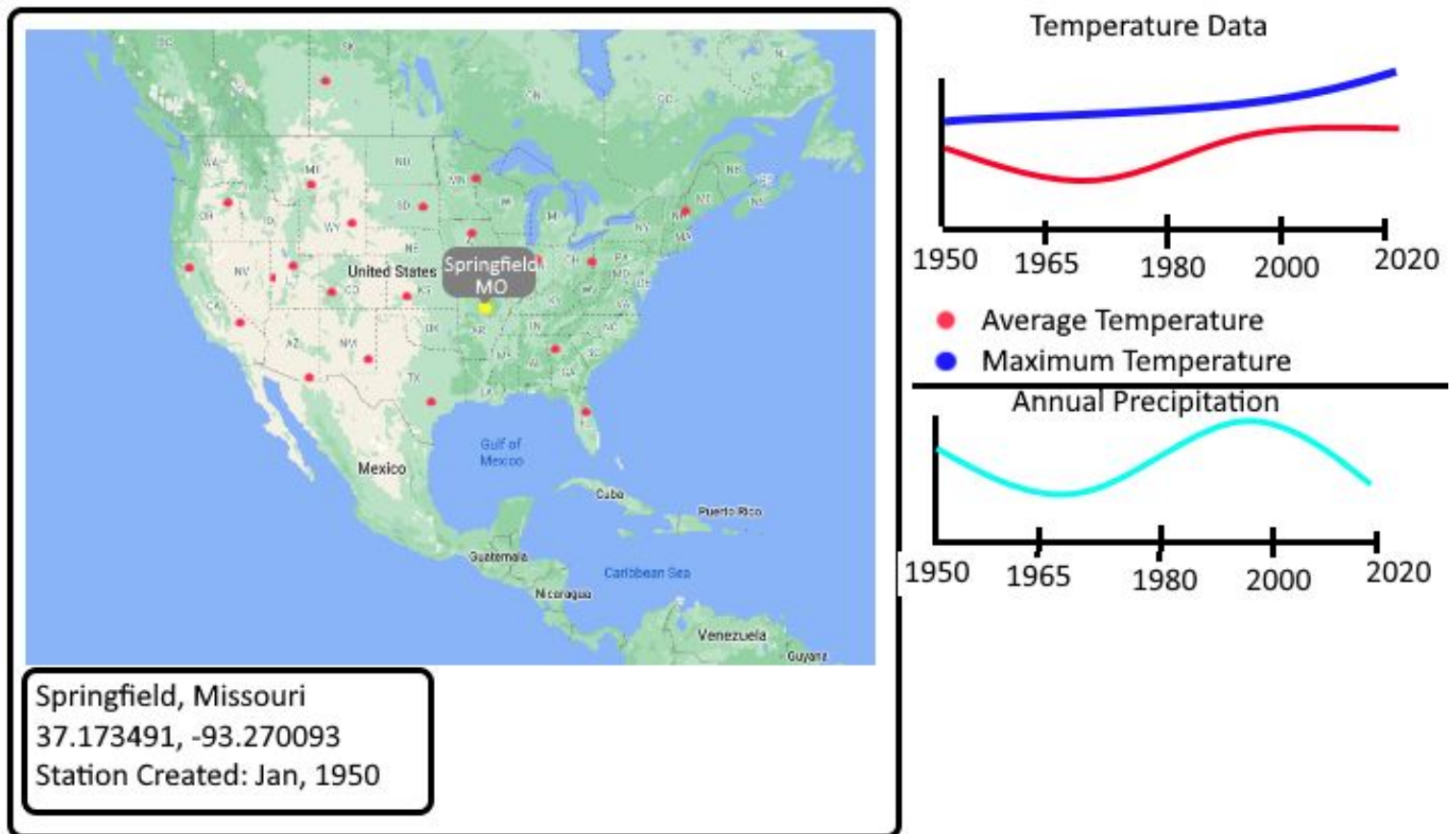
- Prototype design #2:
  - Selecting two weather stations and being able to compare their data side by side (we did not draw all the graphs/categories that will be shown. Just enough to give an idea of what it will look like).

Denver CO

Springfield MO

Temp. Average

Precipitation

Springfield, Missouri
37.173491, -93.270093
Station Created: Jan, 1950

Denver, Colorado
31.173491, -94.12345
Station Created: Mar, 1903

Clear Selection

- Prototype design #3:
  - Selecting a larger number of weather stations through the use of a brush. The user can then select which category they want to display for the selected weather stations. When hovering over a certain line or weather station the corresponding visualization marks will be highlighted.
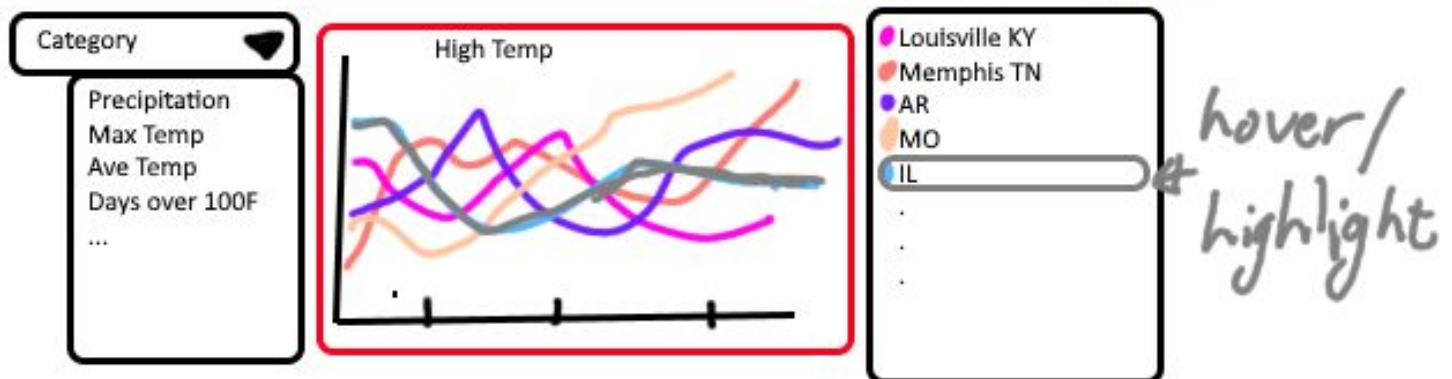
- Final Design Idea:
  - User selects a group of stations using a brush on the map. The selected stations will show up in the group compare box. A summary comparison can be selected. If one particular station is selected in the right panel, more detailed information will appear below.
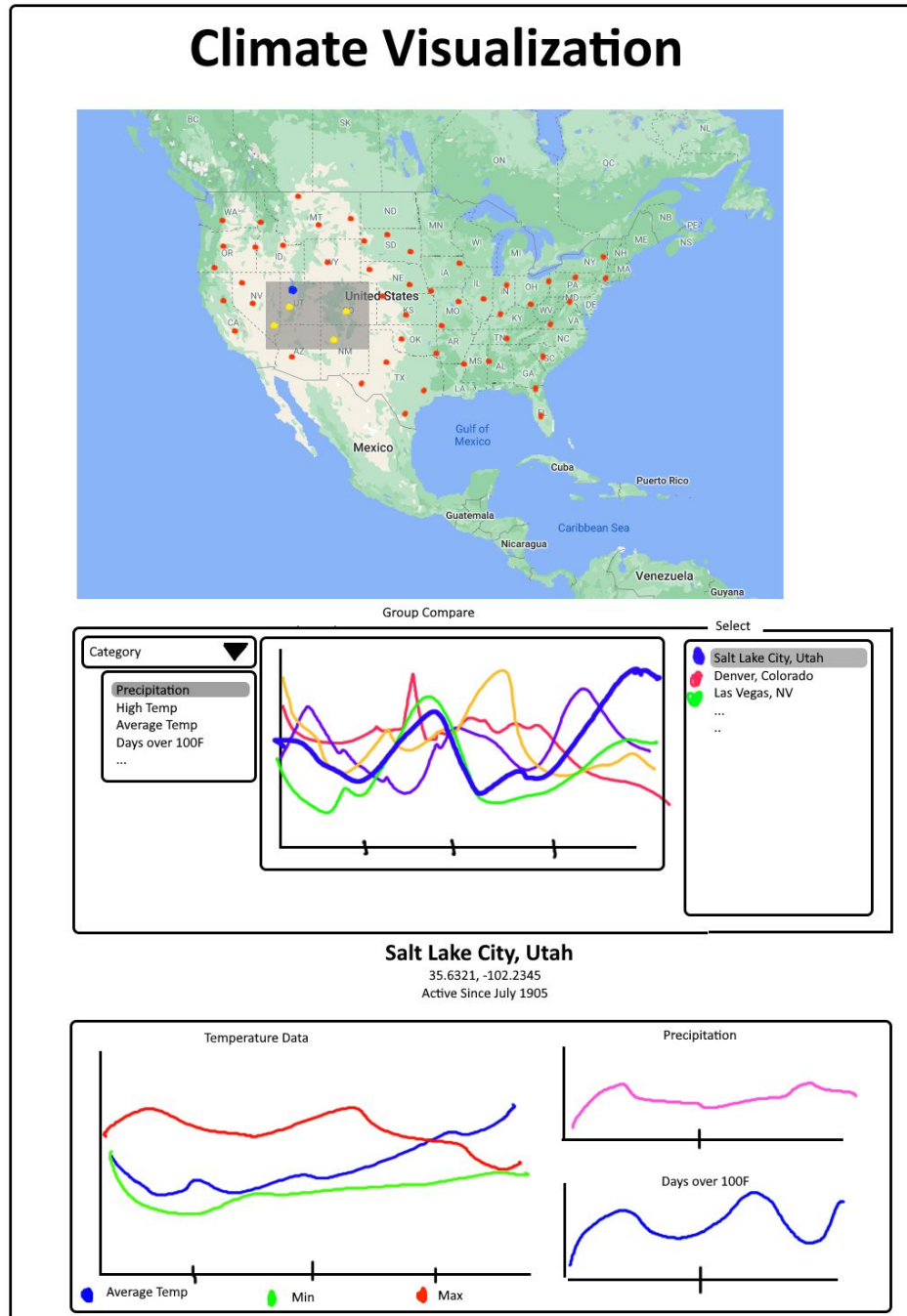
Must-Have Features:

- A Map Interface with the weather stations marked
- The ability to select a weather station and to see the climate data for it
- The ability to select data between an inputed start and end date
- The ability to brush a large group of weather stations and to be able to display which station has the min or max of a certain selected category.

Optional Features: List the features which you consider to be nice to have, but not critical.

- A search option for a certain name, or whatever weather station has the maximum or minimum value for a certain category in a certain year or range of years.
- Select some (2 - 3) stations to have a more in-depth comparison.
- Storytelling Features
- Add an option where they can select a start date and end date and data for in between those two dates will be displayed.

Project Schedule: Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.

- Week of Nov 2:
  - Data wrangling / preprocessing (Jacob)
    - Convert to CSV/JSON files.
    - Change from daily to monthly
    - Remove unneeded data.
  - Start implementing map (William)
- Week of Nov 9:
  - Implement a map with lat/lon data for the stations.(Both)
    - Selection / Hovering
    - Show details of station upon selection. (Name, lat, lon,etc)
      - Basic graphs of some data.

- - ■ Prepare for the milestone on NOV 15.
  - Week of Nov 16:
    - Multi-selection and comparison of multiple stations.(Both)
      - Different views / data type comparisons.
  - Week of Nov 23:
    - Detailed selection and comparison of a smaller subset of stations.(Both)
      - More detailed
      - Implement interactivity (interaction between views)
      - Storytelling
  - Week of Nov 30:
    - Test(Both)
    - Finalize and turn in(Both)