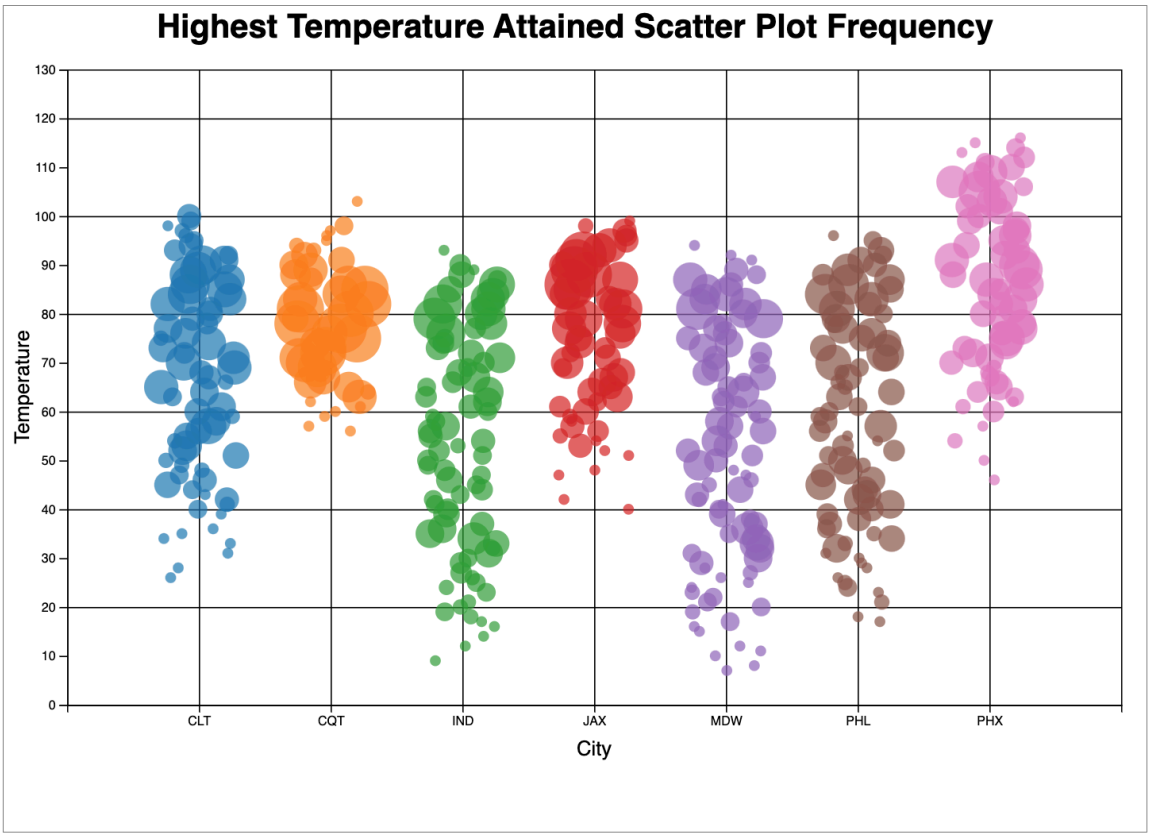
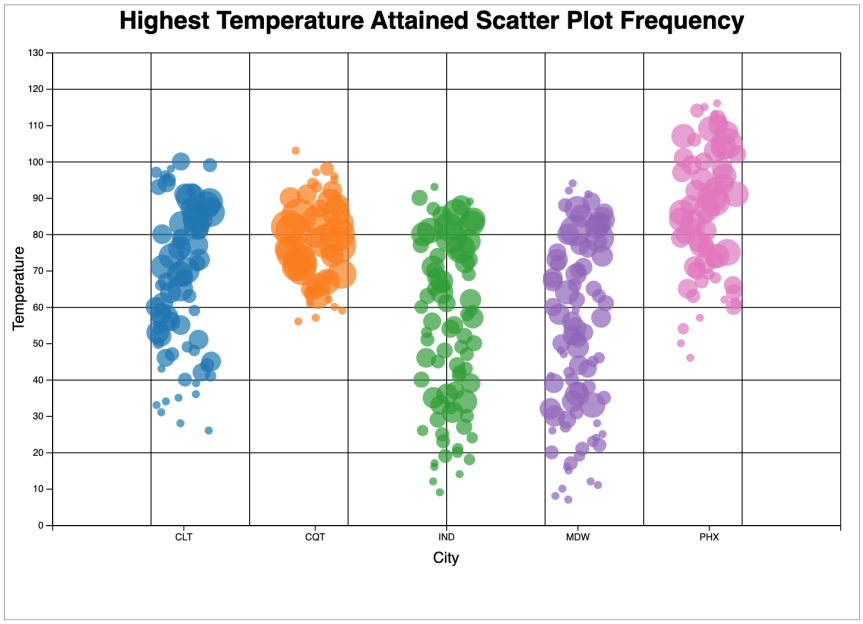
User Stories/Tasks it supports

- User story: “As a Climate Scientist I want to evaluate high temperature frequencies in specific regions to understand climate change impact so that I can showcase detailed weather assessments”
    - Importance: This is important because being able to understand frequencies helps analyze data that people to make informed decisions on climate change laws
- User story: “As a Data analyst I want to study detailed analysis of temperature trends in multiple cities so that I can make in-depth insight and analysis between them”
    - Importance: This is important because data analysts can find trend and make reports on meaningful insights they gained to make recommendations for stakeholders decision
- User story: “As a city Administrator I want to utilize an interactive legends menu to compare temperature trends so that I can filter cities I am interested in without further distraction from others”
    - Importance: This is important because being able to focus on filtered cities helps allow more analysis on the targeted regions they want to analyze from in order to make informed planned decisions for high heat prevention measures
- User story: “As a public health official I want to identify temperature extremes in cities so that I can issue safety hazard plannings
    - Importance: This is important because similarly to the city administrator, a public health officials safeguard the community by instituting initiatives that prevent health risks from rising heats

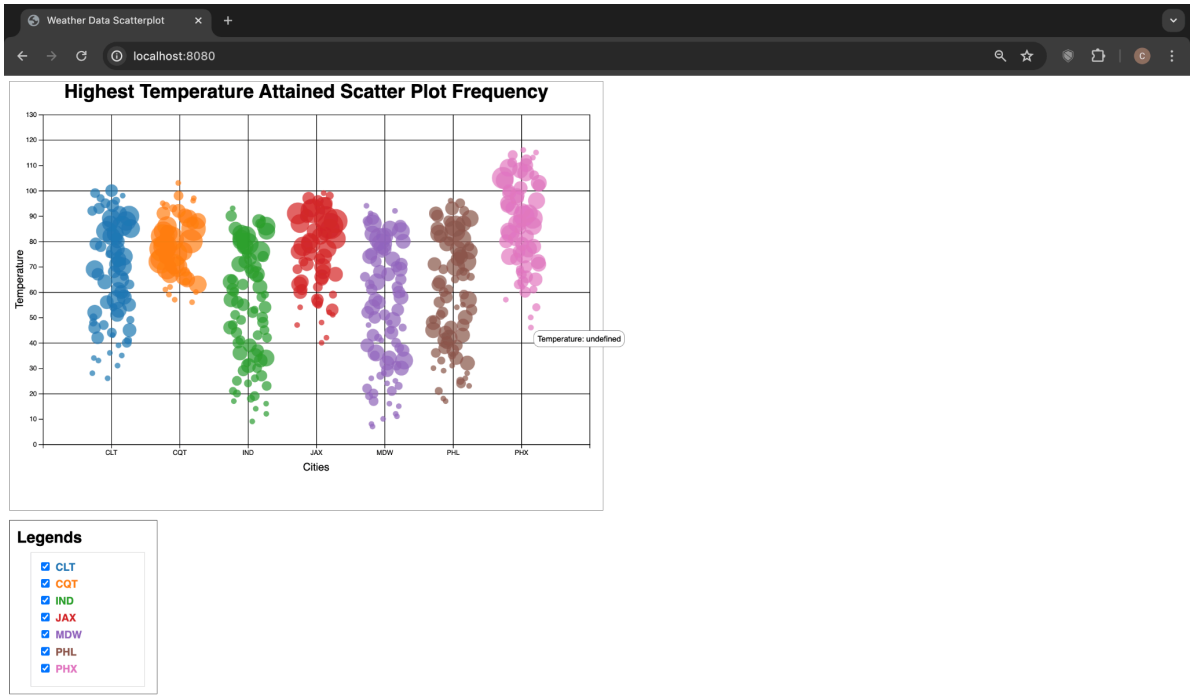
Screen shot(s) of the user interface:





Legends

- [x] CLT
- [x] CQT
- [x] IND
- [ ] JAX
- [x] MDW
- [ ] PHL
- [x] PHX



Design Overview

The visualization shows the highest temperatures attained across multiple cities in this interactive and detailed showcasing. It applies a scatter plot system that showcases data points from the feature ‘actual_max_temp’. It isn’t considered a simple scatter plot as it incorporates frequencies of the temperatures attained and is color coded into distinguishable colors that pertain to the legends menu that is interactive to select/deselect cities at the users will while also introducing practices of data transparency/sizing. With these factors in mind, the user can utilize comparing examinations and gain thorough details on patterns /trends in temperature between each city.

Description of any aspect of the interface/visualization that you feel needs explanation. (in-depth analysis of the design overview)

The visualization uses the data columns date to record the day that the cities temperature was attained and the ‘actual_max_temperature’ which includes the measured temperature used that day to use as a data point for that scatter plot highest temperature.

It goes beyond a simple scatterplot since it involves the frequency of highest temperatures attained showcasing the most common high heat temperatures they attained.

As you can see from the visualization I allowed for all the cities to be all selected to fully distinguish the use of color scheme to allow an easier visual appearance for the user to see which city and its points are.

You can also see I implemented the principle of 'overplotting' with the use of jittering as data points are scattered but still going in the vertical direction thus helping to see more data points.

<u>The interactive features it represents</u>

- Interactive filter system
    - To implement an interactive filter system I initialized a legends box that has all of the cities be able to be selected/deselected.The legends menu itself is a check box menu so that the users are able to click on the boxes inside to toggle the visibility of the cities. As cities are deselected/selected they span outwards or inwards to allow adjustment principles for easy readability thus allowing the scatterplot to be clearly interpreted regardless of the amount of cities showcased. Being able to allow a filter system to look at a specific city or cities will allow for comparison and further analysis/focus on when it initializes itself as a toggle system to allow which city the user prefers to look at.
- Click based tool tip for temperature insight
    - I also implemented a tool tip to look into each data point's temperature. Though it isn't able to read the temperature(as you can see it reads the temperature undefined from the 3rd screenshot when I clicked on PHX's lowest temperature data point), I was able to implement it and make it functional using a click-based system when using the mouse to have a pop-up come up

<u>Visual Encoding</u>

- Grid lines Usage
    - It uses a grid line scale to get a clear overview of highest temperatures to lowest temperatures each city exhibits, this allows the position of each cities data points to be encoded on a scale that is representable for looking at the temperatures of each cities data point which is supported by the y-axis (temperature)
- X axis categorical data

    - The x-axis uses categorical data to represent the cities within the ‘Weather_Data” zip file to ensure related data points are correspondent to their city
- Position Encoding
    - The scatter plot implements positioning as the data points are mapped vertically on the x-axis (for cities) and y-axis (for high temperature values). This lets the user understand the range and variation in temperatures for each corresponding city thus allowing comparison amongst them.
- Color Encoding
    - It uses a color scheme to distinguish relatable points to their specific city which allows visual recognition to define an easier tracing of whose cities temperatures are belonging to them.
- Size/Opacity Encoding
    - The data points use an opacity for density reasons and sizing for the frequency of temperatures to collect the commonly obtained ones. This lets the user to both identify the more important/significant data in a quicker response time while also introducing the reduction of cluttering for through the use of jittering to allow more visible points

Is the visualization an effective representation of the data? Is it clear and useful, and does it effectively communicate different aspects of the data?

The visualization is an effective representation of the data as we are not only comparing temperatures attained from each city but also looking at the most common temperatures each city pertains to. As you can see from the data it uses a frequency scatter plot to attain commonly renowned temperatures each city has gained and also shows that each city is both different in temperature highs/lows while also having different patterns of temperatures. Also, We can see from the data that many of the cities are different in temperature levels as a few are more frequent in the higher levels of temperature while others coincide at lower temperature extremes. This is effective at the overall glance level to allow the user to see which cities' temperatures tend to be on a daily basis.

Does the visualization support different analytical questions and/or communicative objectives about the data? These objectives should be made clear in the description.pdf file you submit.

- Explore temperature trends obtained in various cities
    - Communicative objective: The common temperature each city has.
    - My scatterplot shows the distribution of each city's temperatures. This goes in hand with the clustered data points as it allows the user to immediately identify the common values and range for temperatures in each corresponding city. The sizing that introduces frequencies of commonly obtained temperatures helped support which temperature is found within each city

- Comparison (overview)
    - Communicative objective : Comparing and contrasting cities statistics for analysis.
    - With the combination of data point sizing, we can easily compare trends/patterns among the cities. Pertaining to the grand overview of the visual, rather than looking at each city separately, we can take into account all the statistics from each city at once. This question pertains to how we can compare trends among each city in a wider range in order to gain a contextual insight of the broad dispersion of data points from every city to understand the general variations in temperature and their ranges while also grasping the distribution of temperature frequencies .
- Identify patterns
    - Communicative objective : Lowest/Highest temperature between each city.
    - The scatter plot displays the range of temperatures to highlight the lowest and highest temperatures between each city. This spread in a vertical direction and the grouping/clustering of data points allow an easy visual readability to see temperature patterns across each distinct city at an overview glance.
- Interactive filtering
    - Communicative objective : Reducing/Increasing data point showcasing.
    - The interactive legend lets the user select/deselect the cities names to allow data points to be visible or disappear based on visibility purposes. This feature allows comparative analysis, allowing the user to focus on specific cities chosen and get a more in-depth comparison.

Does the visualization effectively apply the ideas we learned all semester? Does it follow good visualization design principles?

The visualization does follow good visualization design principles, specifically the principles of ‘color’, ‘get it right in black and white’ from ‘Munzers Rules, and gestalt principles. Down below are in-depth analysis on how they follow effective visualization design principles:

- Color
    - It showcases the use of color schemes for my categorical types,specifically using differentiating colors for each city, this allows the cities to be easily distinguishable from one another.
    - I also used the principle of ‘safe colors’ to ensure easy recognition for the audience in order to speed up thinking with consistent colors that are devoted to their own city thus lowering confusion to a minimum on what the data points are representable towards.
- Get it right in black and white (munzer rule)
    - In order to allow accessibility for my visualization for the visually impaired, specifically for color blind people, I used the approach of transparency and sizing to see different data points more visible as this ensures data points aren't hindered from each other.
- Similarity (Gestalt principle)
    - The visual follows the principle of similarity as it invokes how every clustered data point shares a respective color contrast system to each of their cities. This aspect for a visual characteristic helps show an effective learning outcome I learned from this semester as I use color to support this principle to improve navigation for the user as it allows them to reduce mental cognition/perception to figure out whose data points belong to what city. This is also supported in the legends as the symbols are color coated to their respective city and in such correlate to the data points ensuring that the points are related to their corresponding city.
- Proximity (Gestalt Principle)
    - I follow this principle as each data point is grouped together to its own city while also separating the cities with evenly spaced margins to distinguish temperature data from one city to another. To go more in-depth with this, the data points are also clustered to their own axis/grid line in a vertical spread which further compliments on how each cities data points are more unique to their corresponding city thus they are perceived more related.

- Overplotting (Jittering)
    - I followed the use of jittering, a process to make more data points more visible as I randomized the noise for each point to allow more traversed positioning but still instituting alignment purposes as there is still spacing for the user to easily distribute whose cities data points are. This helps the reduction of points from being hidden from one another thus letting the user see a better representation of the dataset.

Does the visualization exhibit some creativity?

For this visualization criteria on creativity, I exhibited parts of scatter plot frequency with the use of different opacities and different sizing data points for commonly obtained temperatures between each city while also exhibiting interaction and a storytelling presentation:

- Interaction
    - The interactive state within this visualization is the filtration of cities present as it allows the user to gain more analysis when they aren't obstructed from irrelevant cities when targeting specific cities they choose from thus allowing better perception and analysis to be obtained from while also implementing a click based system to check a data points specific temperature for in-depth comparison among the cities chosen
- Different opacities
    - I issued the use of data transparency to see more data points as some are hindered behind other points. This way you can see more of the data points efficiently as the combination of different sized points would allow the showcasing of the most frequent data used but allows more visibility when points aren't hindered from the visible eye
- Different sizing data points
    - For the visualization, within each city, you can see how the points differ in sizing. I represented data points this way by collecting all of the available data points to offer the most common/attained temperatures between each city.
- Presentation/Storytelling
    - For presentation purposes, the culmination of both opacity and sizing allows the user to essentially deal with the issue of the principle of 'overplotting', when there's a common degree of visual hardship when seeing the relationship between the cities and their respective points.

- We can further analyze common temperatures each city exhibits to see if they usually trend colder or hotter and in retrospect compare between cities in general to allow deeper analysis on which one