

Jared Westerfield, Maya Sharma, Keerthana Yadavelli, and Sarah Ramsey

## Austin Weather Dataset

Our dashboard is designed to visualize the changes over time and the correlation between variables for the dataset `austin_weather.csv`. We received this dataset from Kaggle. The dataset contains statistics regarding the weather in Austin, Texas, from December 20th, 2013, to July 30th, 2017. It contains statistics about each day's temperature, dew point, humidity, sea level pressure, visibility, precipitation, wind speed, wind gust, and any other relevant weather phenomenon like thunderstorms or snow. There are also high and low values for each statistic when relevant.

## Design Process

We began our design process by working on what we wanted our visualization to answer. We selected our dataset during this meeting and planned out several questions we had regarding the data. From there, we began brainstorming what kind of graphs we wanted and how we wanted to lay out our dashboard. We spent time after this meeting individually brainstorming ideas for the second meeting.

At our second meeting, we brought our ideas together, discussed the pros and cons of our ideas, and came up with 2 plans for our dashboard. We brought sample layouts for the dashboard and decided upon a planned layout. After our meeting, we received feedback on our proposal, and we discussed making changes to our plan accordingly, notably the usage of scatterplots.

Adobe XD possible design -

<https://xd.adobe.com/view/f8f5f577-1cf9-4a18-ad01-cf30ffcfc773-ada6/screen/45329fa2-6987-4322-9de1-2f617e7797f7>

At the next meeting, we went over our first drafts for the coded dashboard, to see how it was coming along and if there were any changes we needed to implement. We've created a dashboard with several plots, as well as a timeline where we can view the change in different weather attributes over the course of our data. Additionally, we've modified our data slightly, but only adding additional attributes to the existing data objects derived from pre-existing data attributes to simplify using the data with D3.

At the following meeting, we continued with the timeline design, and added in our correlation scatterplot. We did this so we could gain the benefits of our alternate dashboard design of various graphs, along with the quieter design of our timeline. This design is our final design to which this documentation relates to.

## Data Analysis

Prior to creating the data visualizations we proposed questions that we hoped to answer with our visualizations. We have answered these questions using our visualizations below, the initial questions were as followed:

- What is the relationship between dewpoint and temperature?
- Is there a correlation between dew point and visibility?
- How has the temperature changed over the four years?
- Which months of the year have the highest amount of precipitation?
- Is there a correlation between an attribute and a weather event?
- What attributes are correlated with sea level pressure?
- Are there any other correlations that our work revealed?

What is the relationship between dewpoint and temperature?

Dewpoint and Temperature have a very strong positive correlation. This makes sense, given that dew forms typically in the mornings when air temperatures are at their lowest, since the temperature would then rise from there to reach the average temperature for the day. We can also see a similar relationship on the timeline, where both are high during the summer and low during the winter.

Is there a correlation between dew point and visibility?

Based on the possibility that lower visibility may be caused by fog, it would make sense for there to be a negative correlation between dew point and visibility, since the dew point being higher would mean fog would condense into dew. The visualization backs this up, showing a negative, albeit weak, correlation between the 2.

How has the temperature changed over the four years?

The temperature has stayed fairly consistent throughout the 4 years. It is coldest from December to March, and warmest from June to September.

Which months of the year have the highest amount of precipitation?

There seems to be no consistent patterns regarding precipitation. Large amounts of precipitation are spaced out fairly evenly throughout the timeline, with no real pattern forming. This makes sense given the dry climate of Austin.

Is there a correlation between an attribute and a weather event?

There are a few notable points where the attributes have a definite effect on the chance of a weather event occurring. When comparing Dew Point and Temperature, it is clear that weather events are more frequently occurring on days where the Dew Point is higher. While unexpected, this makes sense given that a higher dew point would be needed for rain to form.

We also noticed a similar trend with visibility in general, that days with weather events typically had lower visibility. This makes sense, given that a weather event would presumably have an effect visibility, such as fog or heavy rain.

What attributes are correlated with sea level pressure?

At first, we were perplexed as to why sea level pressure is in this data set. We did not think Austin was close enough to the Gulf of Mexico for sea level to be a factor they need to keep track of, especially considering Houston is a city in it's own right and is between the Gulf of Mexico and Austin. Additionally, the data for sea level stays fairly stagnant, varying by an inch on occasions. That being said, there were slight negative correlations visible with dew point and temperature.

Are there any other correlations that our work revealed?

Our research revealed to us some other noteworthy correlations outside of the scope of our initial questions. Wind speed and wind gust had a positive correlation. While we did not foresee this, it should be expected, as they both are measures of the strength of the wind on a given day. Additionally, humidity and dew point had a positive weak correlation. This also makes sense to be correlated, as dew point determines the temperature at which water vapor solidifies into water, and humidity is the measure of water vapor in the air, however it would make sense for a negative correlation, since a higher dew point should mean less water vapor and more liquid water, so a lower humidity. Additionally, it is surprising that these do not have a stronger correlation.