# The City Austin's Weather & Bike Share Trips Data Analysis

---

CIS 5270: BUSINESS INTELLIGENCE

Spring 2019

By Aakanksha Tasgaonkar

Submitted to: Dr. Shilpa Balan

## A. URL's of data sets:

https://www.kaggle.com/grubenm/austin-weather

https://www.kaggle.com/jboysen/austin-bike/activity#austin_bikeshare_trips.csv

The Austin, Texas has a subtropical humid climate. This climate is characterized by very long and hot summers; short, mild winters; and pleasantly warm spring and fall seasons in-between. Austin averages 34.32 inches (872 mm) of annual rainfall and it is distributed mostly evenly throughout the year, though spring and fall are the wettest seasons. Sunshine is common during all seasons, with 2,650 hours, or 60.3% of the possible total, of bright sunshine per year. Austin falls in USDA hardiness zones 8b (15 °F to 20 °F) and 9a (20 °F to 25 °F).[1] Over all, the Austin, Texas area has very pleasant weather, ideally suited for the abundance of outdoor activities the area is famous for. The uncomfortable high humidity experienced on the Texas Gulf Coast is seldom experienced here. [1]

In the Austin area, transportation is a hot topic in this city. It has one of the longest commute times in the state. According to the Texas A&M Transportation Institute, the stretch of IH-35 that runs through downtown was ranked the 3rd most congested road in Texas. [2] Bike shares are becoming a popular alternative means of transportation. The City of Austin makes data available on >649k bike trips over

2013-2017. The main goal of bike sharing programs is to provide a fun, healthy, low-cost transportation option for locals and visitors alike. Austin B-cycle is a public bike sharing program that rolled in to Austin in December 2013. Austin B-cycle nearly broken the national records for check outs during SXSW, making it one of the most successful bike-sharing programs in the country. Currently there are 76 B-cycle stations around downtown Austin. [2]

## B. Data Cleaning:

1. Heading in the dataset is not written completely for abbreviation used.

Before

| | B | C | D |
|---|---|---|---|
| 1 | TempHigh | TempAvgF | TempLowF |
| 2 | 74 | 60 | 45 |
| 3 | 56 | 48 | 39 |
| 4 | 58 | 45 | 32 |
| 5 | 61 | 46 | 31 |
| 6 | 58 | 50 | 41 |
| 7 | 57 | 48 | 39 |

After

| B | C | D |
|---|---|---|
| Temperature High | Temperature Average | Temperature Low |
| 74 | 60 | 45 |
| 56 | 48 | 39 |
| 58 | 45 | 32 |
| 61 | 46 | 31 |
| 58 | 50 | 41 |
| 57 | 48 | 39 |

The heading in the data set was not written completely. So, edited it.

2. Year column needs to be changed to proper date format.

Before

After

| subscribe | trip_id | year |
|---|---|---|
| Walk Up | 9.9E+09 | 2015 |
| Local365 | 12617682 | 2016 |
| Local365 | 9075366 | 2016 |
| 24-Hour K | 9.9E+09 | 2014 |
| Walk Up | 14468597 | 2017 |
| Local30 | 9.9E+09 | 2015 |

| subscribe | trip_id | date |
|---|---|---|
| Walk Up | 9.9E+09 | 3/19/2015 |
| Local365 | 12617682 | 10/30/2016 |
| Local365 | 9075366 | 11/3/2016 |
| 24-Hour K | 9.9E+09 | 11/23/2014 |
| Walk Up | 14468597 | 4/16/2017 |
| Local30 | 9.9E+09 | 5/29/2015 |

The year should be Proper format. So, changed it to mm/dd/yyyy.

3. Repetition of month i.e. Removed the month cell, which needs to be converted into single column as Year.

Before

| month | start_station_id | start_station_ | subscribe | trip_id | date |
|---|---|---|---|---|---|
| 3 | 2536 | Waller & 6th S | Walk Up | 9.9E+09 | 3/19/2015 |
| 10 | 2494 | 2nd & Congres | Local365 | 12617682 | 10/30/2016 |
| 3 | 2538 | Bullock Museu | Local365 | 9075366 | 11/3/2016 |
| 11 | 2497 | Capitol Statio | 24-Hour K | 9.9E+09 | 11/23/2014 |
| 4 | 2707 | Rainey St @ C | Walk Up | 14468597 | 4/16/2017 |
| 5 | 2540 | 17th & Guadal | Local30 | 9.9E+09 | 5/29/2015 |

After

| start_station_id | start_station_ | subscribe | trip_id | date |
|---|---|---|---|---|
| 2536 | Waller & 6th S | Walk Up | 9.9E+09 | 3/19/2015 |
| 2494 | 2nd & Congres | Local365 | 12617682 | 10/30/2016 |
| 2538 | Bullock Museu | Local365 | 9075366 | 11/3/2016 |
| 2497 | Capitol Statio | 24-Hour K | 9.9E+09 | 11/23/2014 |
| 2707 | Rainey St @ C | Walk Up | 14468597 | 4/16/2017 |
| 2540 | 17th & Guadal | Local30 | 9.9E+09 | 5/29/2015 |

Repetition of month so deleted the whole month column as the month is there in the date column.

4. Removing the blank values

Before

| WindA | WindG | Precipi | Events |
|---|---|---|---|
| 4 | 31 | 0.46 | Rain , Thunderstorm |
| 6 | 25 | 0 | |
| 3 | 12 | 0 | |
| 4 | 20 | 0 | |
| 2 | 16 | T | |
| 3 | 17 | 0 | |
| 1 | 11 | T | |

After

| WindA | WindG | Precipi | Events |
|---|---|---|---|
| 4 | 31 | 0.46 | Rain , Thunderstorm |
| 1 | 18 | 0.16 | Rain |
| 3 | 19 | 0 | Fog |
| 6 | 21 | 0.1 | Rain |
| 4 | 24 | 0.01 | Rain |
| 9 | 31 | 0.06 | Rain , Snow |
| 9 | 32 | 0.02 | Rain |
| 6 | 23 | 0.05 | Rain |

Removed all the blank spaces in the excel sheet. By filtering out the blank spaces and null value.

5. Splitting the column.

Before

| start_time | subscriber_type |
|---|---|
| 3/19/2015 19:12 | Walk Up |
| 10/30/2016 2:06 | Local365 |
| 3/11/2016 16:28 | Local365 |
| 11/23/2014 15:12 | 24-Hour Kiosk (Au |
| 4/16/2017 15:39 | Walk Up |
| 5/29/2015 15:12 | Local30 |

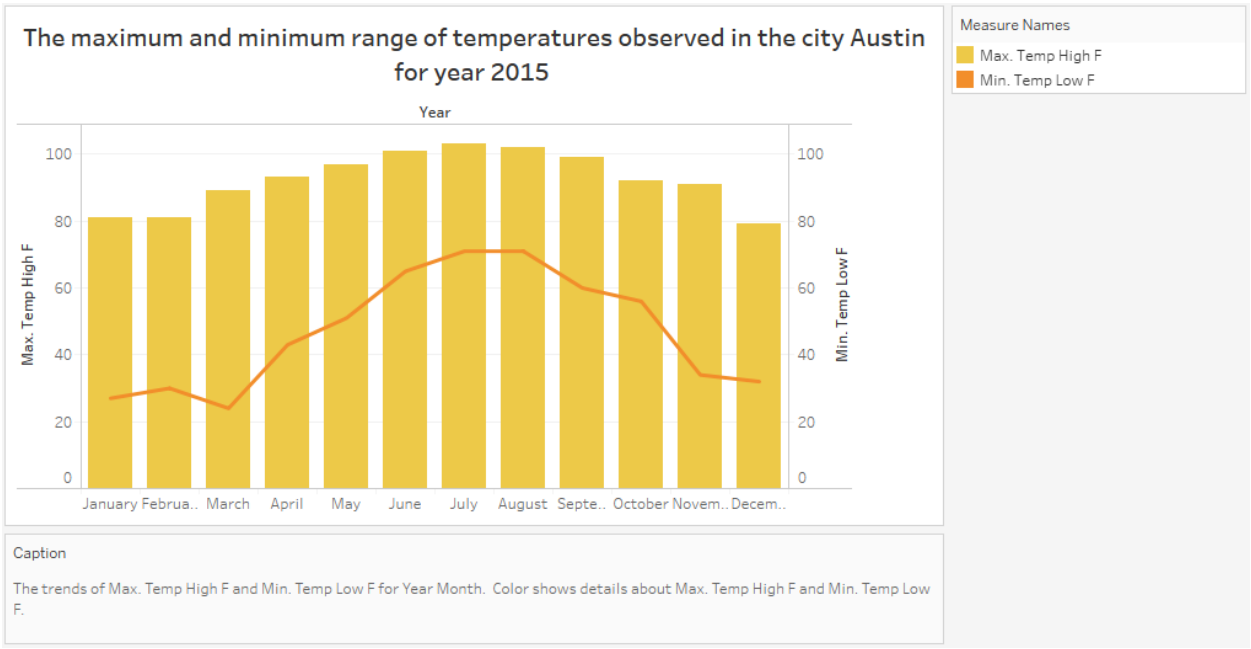
After

| start_time | | subscriber | trip_id |
|---|---|---|---|
| 7:12:00 | PM | Walk Up | 9.9E+09 |
| 2:06:04 | AM | Local365 | 12617682 |
| 4:28:27 | PM | Local365 | 9075366 |
| 3:12:00 | PM | 24-Hour K | 9.9E+09 |
| 3:39:13 | PM | Walk Up | 14468597 |
| 3:12:00 | PM | Local30 | 9.9E+09 |
| 12:24:16 | PM | Walk Up | 11114967 |
| 1:12:00 | PM | Walk Up | 9.9E+09 |

Splitting the columns and then creating two separate columns for start date as well as for the time

## C. Data Visualization:

**Question 1. What is the maximum and minimum range of temperatures observed in the city Austin?**

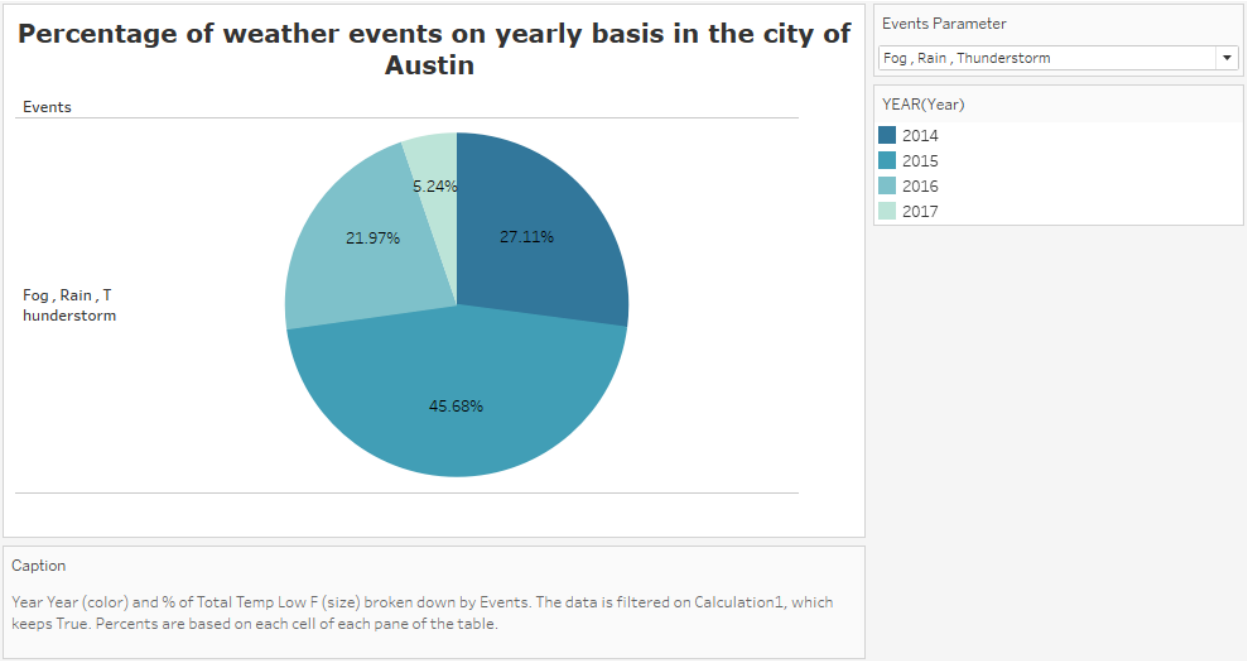



**[Tools Used: Dual Axis Chart]**

The above bar graph represents the maximum and minimum range of temperatures observed in the city of Austin on monthly basis for the year 2015. Dual Axis shows comparison between high temperature (bar graph) and low temperature (Line graph). The highest temperature is in the month of June with 103F whereas the lowest is temperature is observed in the month of March with

24F. The range for maximum temperature is 78F to 103F. The highest temperature steadily decreases after the month of July. The range for minimum temperature is 24F to 71F.

**Question 2. What is the percentage of weather events on yearly basis in the city of Austin ?**

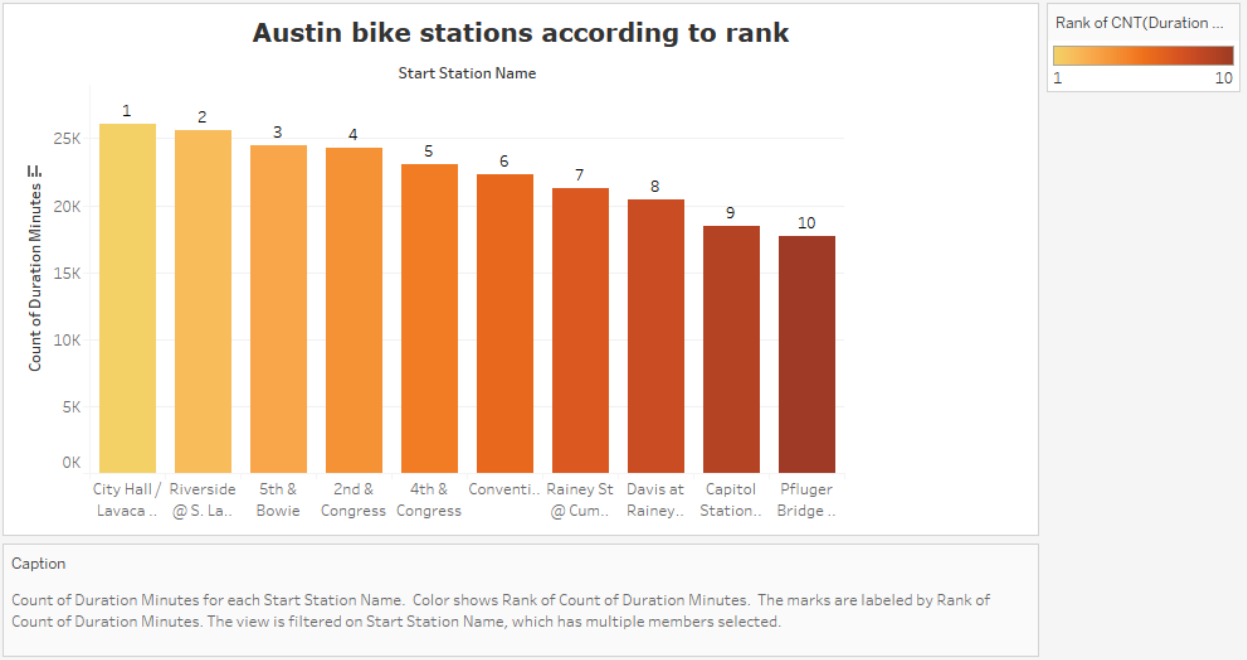



**[Tools Used: Calculated Field, Parameter]**

By analyzing the above pie chart, in the year 2015 Austin has attained the maximum fog, rain and thunderstorm. In comparison with all the years, it is observed that year 2015 itself has almost 50% of the weather event for fog, rain and thunderstorm. While, the remaining 3 years together make 50%. The minimum percentage is observed in the year 2017. There is a slight difference of

approximately 5% percentage for the year 2014 and 2016. In this pie chart there is a parameter control for all of the weather events in the city of Austin

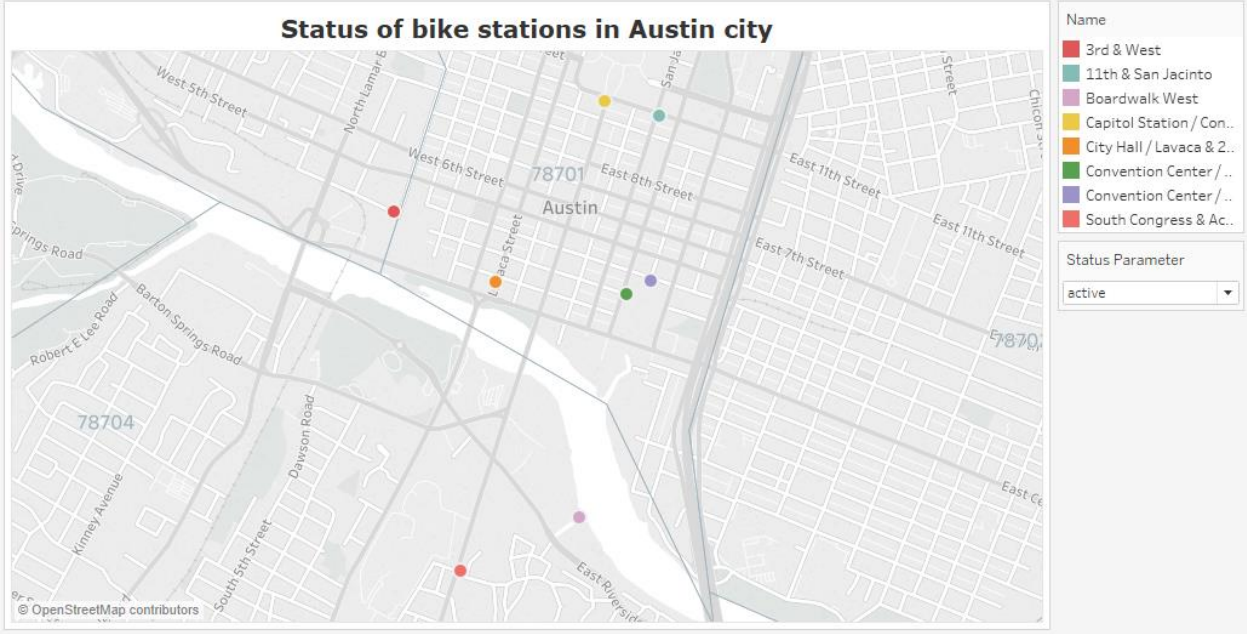
**Question 3. Top 10 bike stations in the city of Austin**




**[Tools Used: Ranks]**

The bar graph shows the rank of top 10 bike stations here all the bike stations are ranked on the basis of duration. City hall/Lavaca &2$^{nd}$ is rank the top most however Pfluger bridge is ranked the bottom. The count of duration is gradually decreasing from 26k to 17k. The station 5$^{th}$ bowie and 2$^{nd}$ congress which is in 3rd and 4$^{th}$ rank respectively has almost similar values. The bike ride duration for the top location [City hall/Lavaca &2$^{nd}$] is 35% of the lowest ranked location.

## Question 4. The status of the bike stations in Austin city



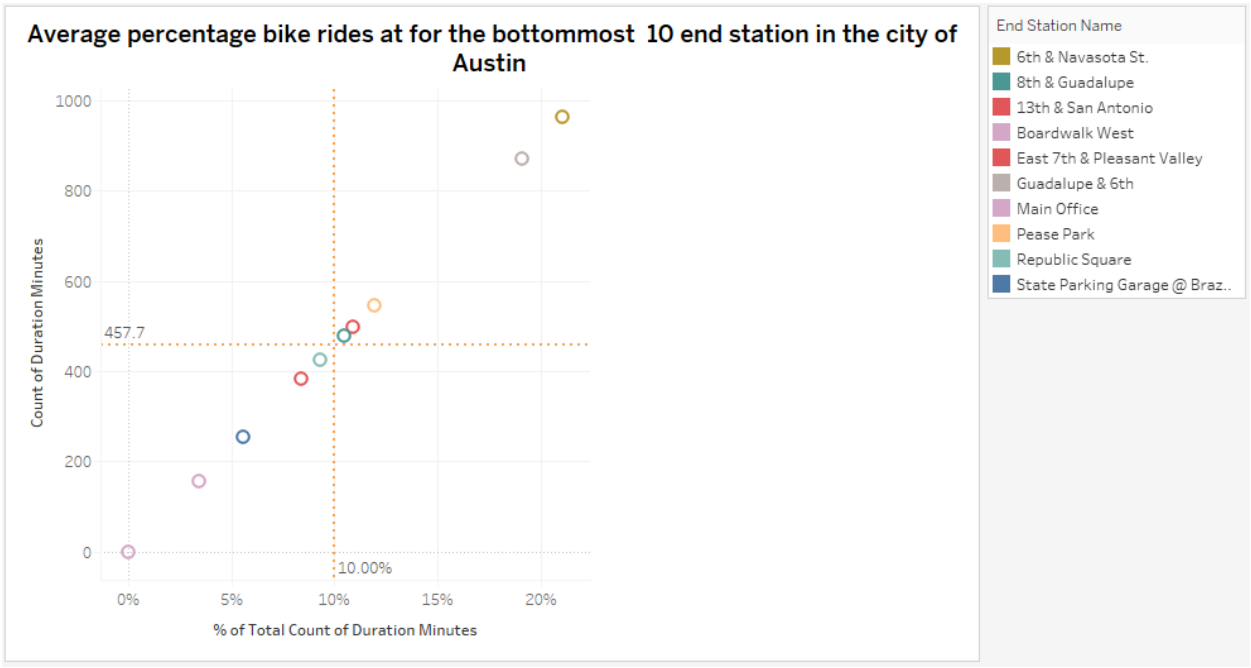

Caption

Map based on Longitude and Latitude. Color shows details about Name. The data is filtered on Status, Calculation1 and Location. The Status filter has multiple members selected. The Calculation1 filter keeps True. The Location filter has multiple members selected.

**[Tools Used: Geographical Maps, Parameter]**

The above geographical map illustrates the bike stations in Austin city. Used Parameter control for the above map. Created calculated field equation the equation is '[Status] = [Status Parameter]'. The bike station has 4 different status like Active stations, closed station, etc.

**Question 5: Average percentage of bike rides for the bottommost 10 end station in the city of Austin**

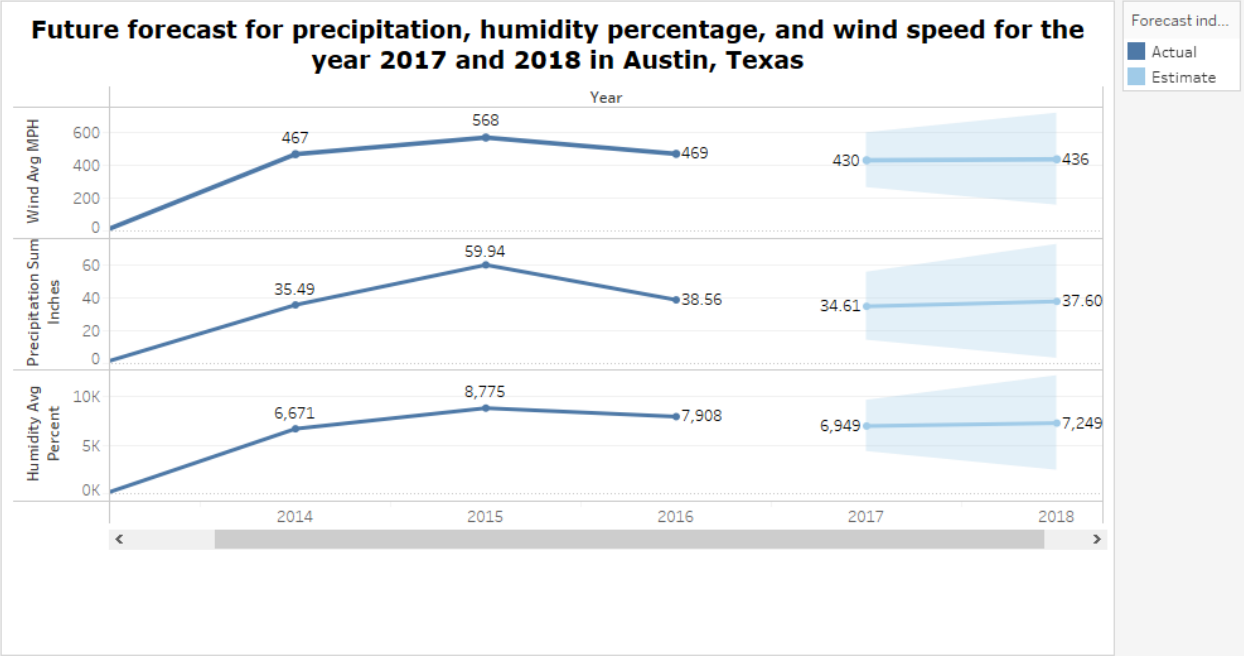



**[Tools Used: Scatter Plot, Reference Line]**

Above scatter plot displays the average percentage of bike rides for the last 10 end stations in the city of Austin. The highest percentage of bike rides are seen for 6$^{th}$ & Navasota St. end station and the lowest is seen for Main Office end station. Average percentage of bike rides are represented by a dotted reference line which is about 10% and having total count of time covered by the bikes in the Austin city

is approximately 458 minutes. End stations 8th & Guadalupe and East 7th & pleasant Valley have very close percentages of bike rides i.e. 11 % and 12 % respectively.

**Question 6: Future forecast for precipitation, humidity percentage, and wind speed for the year 2017 and 2018 in Austin, Texas?**




Caption

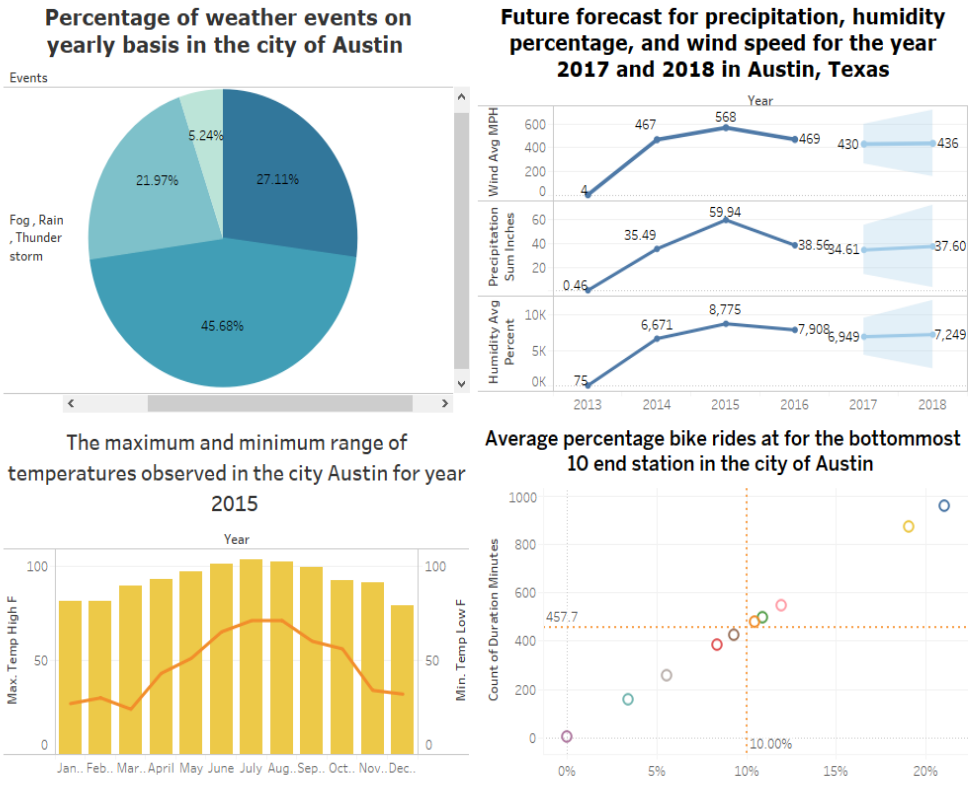
The trends of sum of Wind Avg MPH (actual & forecast) , sum of Precipitation Sum Inches (actual & forecast) and sum of Humidity Avg Percent (actual & forecast) for Year Year. Color shows details about Forecast indicator.

**[Tools Used: Forecast trend Lines]**

This diagram illustrates future prediction of the weather conditions based on the precipitation, humidity percentage, and wind speed for the year 2017 and 2018 in Austin, Texas. As per the forecast, similar trends are observed for all the weather

conditions - precipitation, humidity percentage, and wind speed. It can be seen that from the year 2014 to year 2015 there is a slight increase in their values while for the year 2016 there is a slight decrease in their values. In comparison for the years 2014 & 2016, wind speed has just 2 miles per hour difference (from 467 to 469). Also, when compared with precipitation there is only 3 inches of difference observed. Forecast trend lines feature help to get approximate figures for the weather conditions in Austin for the year 2017 and 2018 respectively.

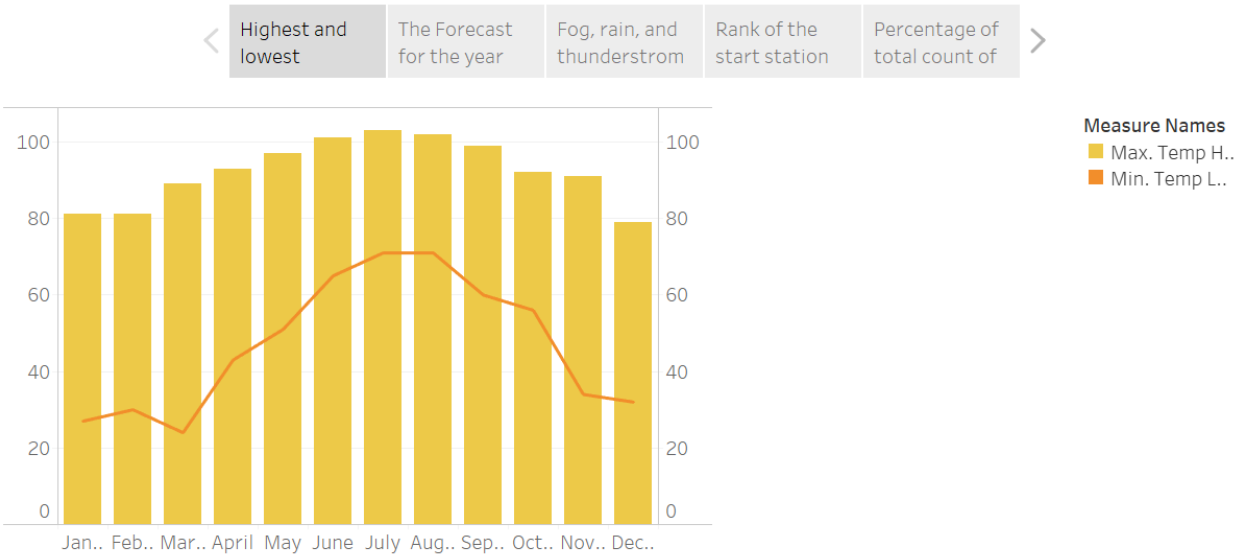
**D) Dashboard**



## E) Storytelling

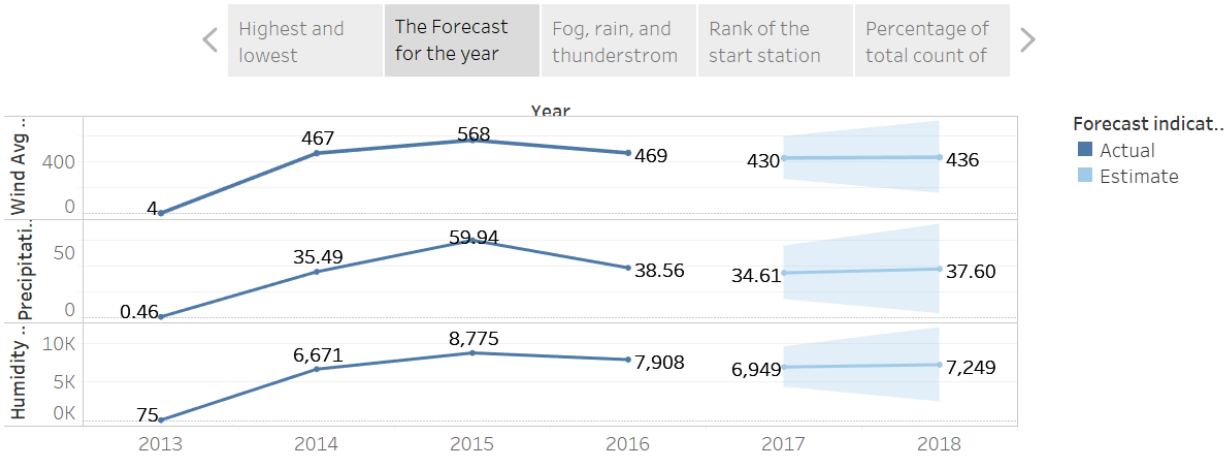
The story is to predict the weather conditions observed in the city of Austin, based on the data such as Temperature, Dew point temperature, Humidity percentage, Sea level pressure, Visibility, Wind speed, Precipitation from the year 2013 to 2017. This data will provide valuable insights on the weather situations like rain, fog, and thunderstorm. In Austin, Bike shares are becoming a popular alternative means of transportation. The City of Austin makes data available on >649k bike trips over 2013-2017. This data includes information on bike trip start location, stop location, duration, type of bike share user. [3] Bike station location data is also provided. This data will help to forecast the bike share demand as well as to understand the most popular stations in the city of Austin, Texas. It has been observed how tableau software can compare both the datasets and answer questions by dashboards. This dashboard introduces data visualization conceptual bases and propose a visual analytic and visualization platform in Austin weather surveillance for federal, state, and local public to make more appropriate decisions. The reason for measuring the overall quality of data is to improve it and perform benchmark analysis to get a clear understanding of the quality of the dataset. Tableau will enable the retrieval, visualization, and exploration of a uniformlydefined selected fields of analysis like

the highest temperature and the maximum duration of a bike trip for the Austin weather & Bike trip.

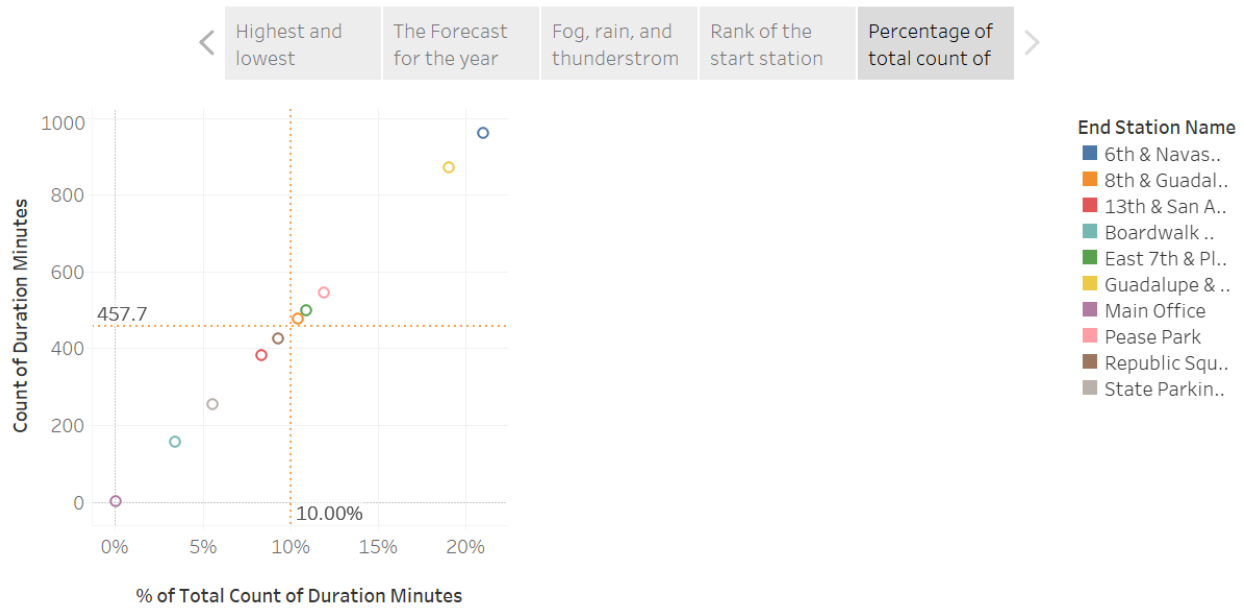



The Austin city's weather which can result in more numbers of bike share ride. If the temperature is pleasant then there will be large numbers of bike rides. According to figures from the in the year 2015, 24 F is the lowest for the month of march and 103 F is the highest for the month of July but the lowest temperature measured during that time was 23 degrees Fahrenheit (-5 Celsius) on January 23, 2015.[2] As is typical for Memorial Day, people in Houston spent their 2015 holiday enjoying barbecues and time on nearby beaches. That downtime, however, became a distant memory for many when they woke up the next day.

Within a nine-hour span from the night of May 25, 2015, to the morning of May 26, as much as 11 inches of rain fell on parts of the region. [4]




The future prediction of the weather conditions based on the precipitation, humidity percentage, and wind speed for the year 2017 and 2018 in Austin, Texas. As per the forecast, similar trends are observed for all the weather conditions - precipitation, humidity percentage, and wind speed. It can be seen that from the year 2014 to year 2015 there is a slight increase in their values while for the year 2016 there is a slight decrease in their values. In comparison for the years 2014 & 2016, wind speed has just 2 miles per hour difference (from 467 to 469). Also, when compared with precipitation there is only 3 inches of difference observed. Forecast trend lines feature help to get approximate figures for the weather conditions in Austin for the year 2017 and 2018 respectively.




Above scatter plot displays the average percentage of bike rides for the last 10 end stations in the city of Austin. The highest percentage of bike rides are seen for 6th & Navasota St. end station and the lowest is seen for Main Office end station. Average percentage of bike rides are represented by a dotted reference line which is about 10% and having total count of time covered by the bikes in the Austin city is approximately 458 minutes. End stations 8th & Guadalupe and East 7th & pleasant Valley have very close percentages of bike rides i.e. 11 % and 12 % respectively. They probably closed these kiosks because they weren't as popular as the other ones, causing less revenue? I think this locations are still a great opportunity. New events and festivals can be planned in these areas so bars and restaurants can get

benefit from this. It would be good to analyze each location and see what was the reason they closed these kiosks.

To conclude, The Austin city's weather which can result in more numbers of bike share ride. If the temperature is pleasant then there will be large numbers of bike rides. The main goal of bike sharing programs is to provide a fun, healthy, low-cost transportation option for locals and visitors alike. [3]

## Reference


**1. https://www.tripsavvy.com/austin-average-monthly-temperatures-255108**

**2.https://www.currentresults.com/Yearly-Weather/USA/TX/Austin/extreme-annual-austin-low-temperature.php**

**3.https://medium.com/datadriveninvestor/innovation-in-urban-mobility-analyzing-austin-b-cycle-sharing-program-data-56ce0cd92131**

**4.https://www.click2houston.com/weather/remembering-houstons-2015-memorial-day-flood**