**Citi Bike Analyzing Project**

**Group Name: Queen**

**Group member: Chuanze Cai, Yanling Peng, Luoqi Deng, Yao Nie**

**Contents**

[Analysis one 4](#_Toc480836715)

[Analysis two 5](#_Toc480836716)

[Analysis three 7](#_Toc480836717)

[Analysis four 8](#_Toc480836718)

[Analysis Five 10](#_Toc480836719)

[Analysis Six 11](#_Toc480836720)

[Analysis Seven 12](#_Toc480836721)

[Analysis Eight 13](#_Toc480836722)

[Analysis Nine 14](#_Toc480836723)

[Analysis Ten 15](#_Toc480836724)

[Analysis Eleven 16](#_Toc480836725)

[Analysis Twelve 17](#_Toc480836726)

[Analysis Thirteen 18](#_Toc480836727)

[Analysis Fourteen 19](#_Toc480836728)

[Analysis Fifteen 20](#_Toc480836729)

[Analysis Sixteen 21](#_Toc480836730)

# Analysis one

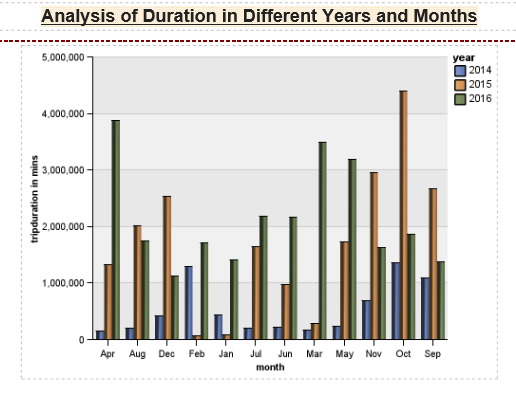
A. Analyzing duration in different years and months by Chuanze Cai

B. In this case, the problem is that the company needs to learn more about the running condition of the rental bikes throughout past three years and past thirty-six months.

C. The procedure was analyzing the whole trip duration for different months and visualize the trend and comparison.

D. From the result attached later, in the term of years, the condition is getting better and better, however, the peak is different. In the year 2014, the peak was in Feb, in 2015 the peak is in Oct and in 2016 the peak falls on Apr.

E. Based on the result, it can be conducted that the company can focus more on those months which is neither too cold nor too cold. Also, the company can keep investing on new equipment like new bikes since the trend of bikes is keep growing and profits can be made constantly.



# Analysis two

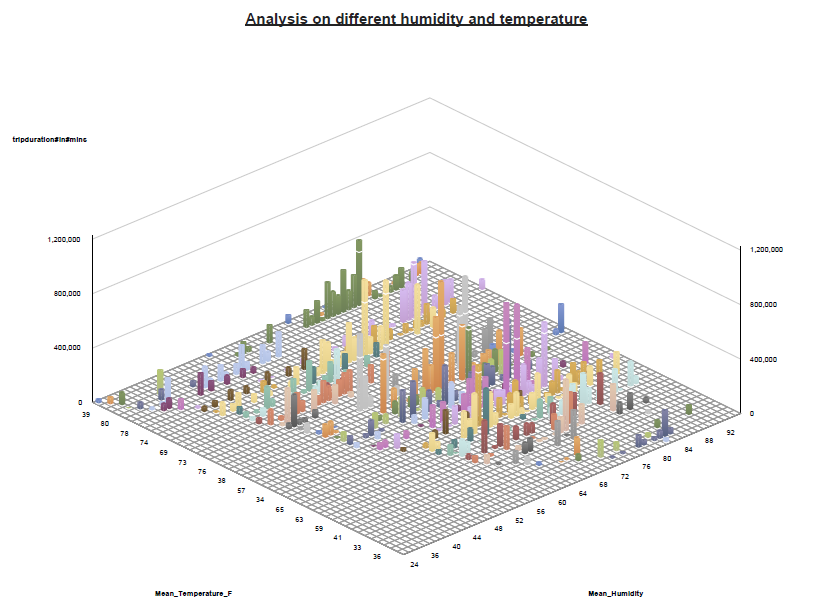
A. Analysis of duration in different seasons by Chuanze Cai

B. The company is a company that rely on outdoor activities so that weather condition can influence the revenue.

C. To visualize the situation, the best way is to generate a 3D chart and set two parameters as axis and take the height of the bar to be trip duration.

D. From the visualization we can learn the distributing situation based on different humidity and temperature. And customers tend to use more between 60-65 degrees and humidity span is 60-72.

E. Based on the insights gain before, company can adjust their strategy based on the local climate condition before entering a new market or new city.



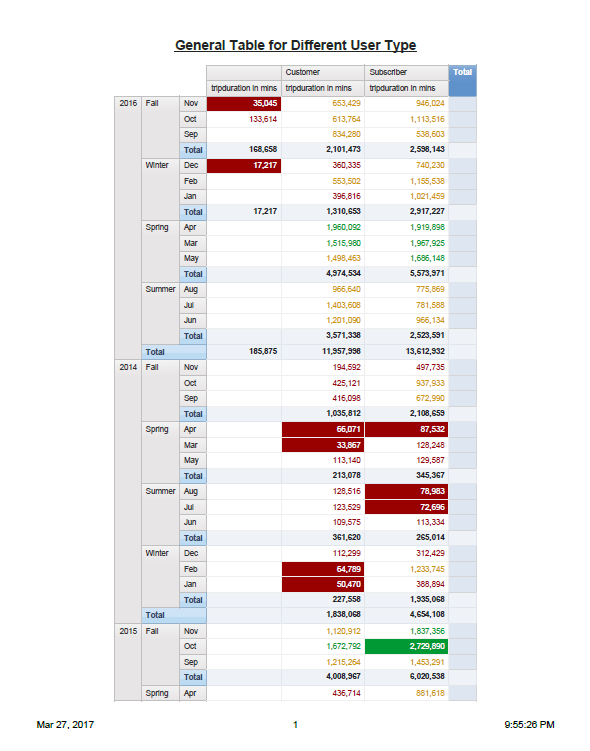
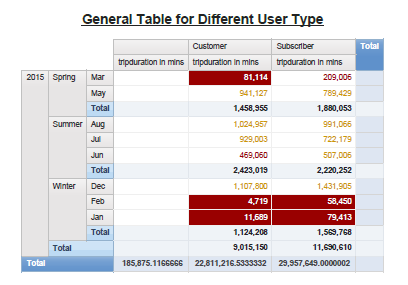
# Analysis three

A. Analyzing sales condition for different user types by Chuanze Cai

B. The company is trying to analyze the trip duration distributing condition based on different user types including customer and subscriber.

C. For this condition, the best way is to leverage the crosstab and visualize it within different months and seasons.

D. From the conclusion attached, the insight can be gain from the result. According to the conditional style we can learn that in general, subscribers perform better than customers. Also, the spring in 2016 is the best time for the company.

E. Based on the result gained, it can be learned that company can raise the price in spring and gain more profit.

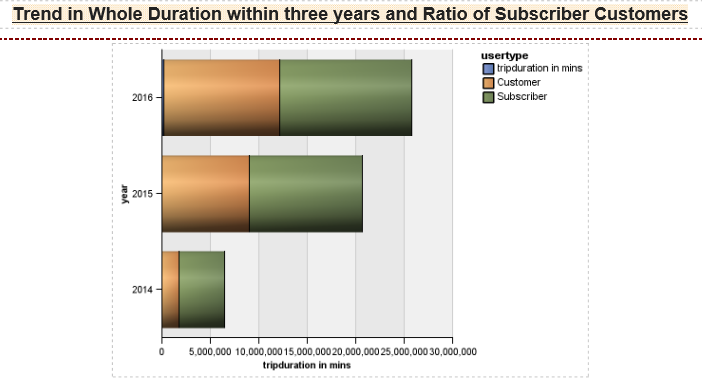
# Analysis four

A. Analyzing trend in whole duration within three years and ratio of subscriber customers.

B. Since the profit can be made from the subscribers so it is necessary for the company to learn about the ratio of subscribers to all of the customers.

C. To compare subscribing condition throughout different years, bar-chart is a good option and make easier to get the insight of it.

D. From the analysis attached, we can tell that the portion of total trip duration by subscriber is decreasing. However, from the other point of view, customers are becoming increasingly active.

E. The Company can start a new campaign and promote for new subscribers and create more revenue.

# Analysis Five

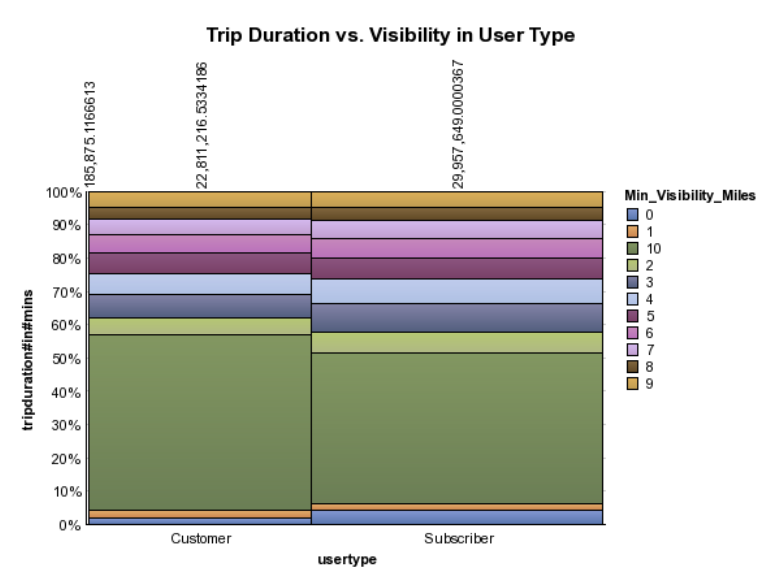
A. Title: Analyzing duration in visibility and user type by Luoqi Deng.

B. Problem statement: In this case, the problem is in what extent the visibility, customer and subscriber are more willing to use Citi bike.

C. Brief overview: The procedure was to plot the comparison graph to visualize the customer and subscriber group trip duration in different visibility.

D. Analysis: From chart below, customer and subscriber are more often to use Citi bike when visibility is in 10 miles. It can be concluded that when air is clear, people have a stronger intention to use Citi bike.

E. Managerial conclusion: Based on the result, the company should give a discount during the rain and snow season since the visibility during those times would be low.



# Analysis Six

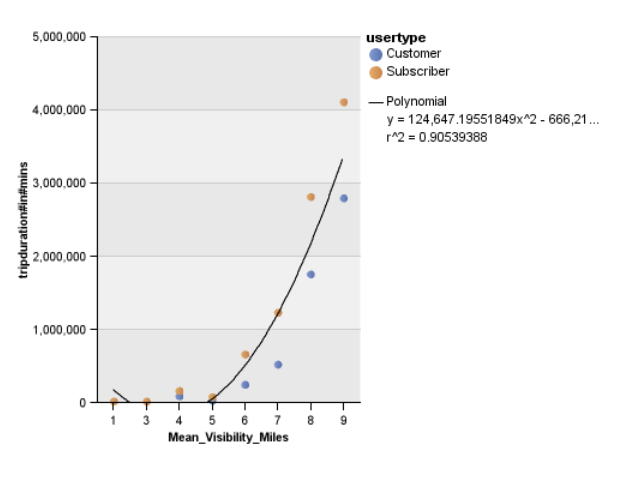
A. Title: Analyzing correlation between trip duration and visibility in user type by Luoqi Deng.

B. Problem statement: In this case, the problem is to find out the correlation between trip duration and visibility.

C. Brief overview: The procedure to plot a regression line between trip duration in different visibility points.

D. Analysis: From chart below, there is an overall positive relationship between visibility and trip duration and R-squared from the regression line reaches 0.90, suggesting the model performs very well.

E. Managerial conclusion: Based on the result, the company should try some ways to attract new customers and subscribers during the season when visibility is high.



# Analysis Seven

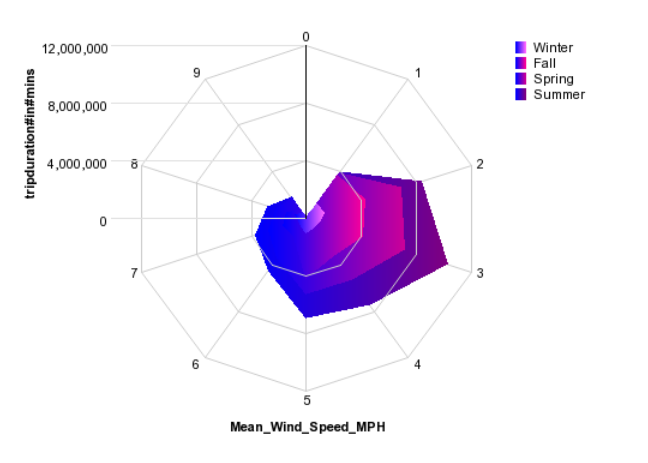
A. Title: Analyzing duration in wind speed by season by Luoqi Deng.

B. Problem statement: In this case, the problem find out how the total trip duration in different wind speed in four seasons in three years.

C. Brief overview: The procedure was to plot a radar char to show the trip duration changing by season and wind speed.

D. Analysis: From chart below, customer and subscriber are more often to use Citi bike when wind speed is between 2 miles/hour and 4 miles/hour. It can be concluded that people have a stronger intention to use Citi bike in breeze weather.

E. Managerial conclusion: Based on the result, the company should avoid entering city like Chicago that always has strong wind.



# Analysis Eight

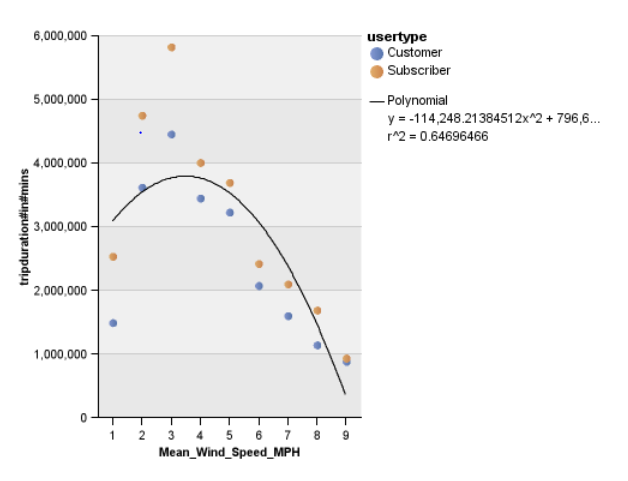
A. Title: Analyzing correlation between trip duration and wind speed by Luoqi Deng.

B. Problem statement: In this case, the problem find out the correlation between wind speed and trip duration.

C. Brief overview: The procedure was to plot a regression line between wind speed and trip duration points.

D. Analysis: From chart below, there is a positive relationship between wind speed and trip duration when wind speed is between 0 mile/hour and 3 miles/hour, and negative relationship when wind speed is between 3 miles/hour and 9 miles/hour.

E. Managerial conclusion: Based on the result, the company should take the average wind speed in a city into consideration before they open a new market.



# Analysis Nine

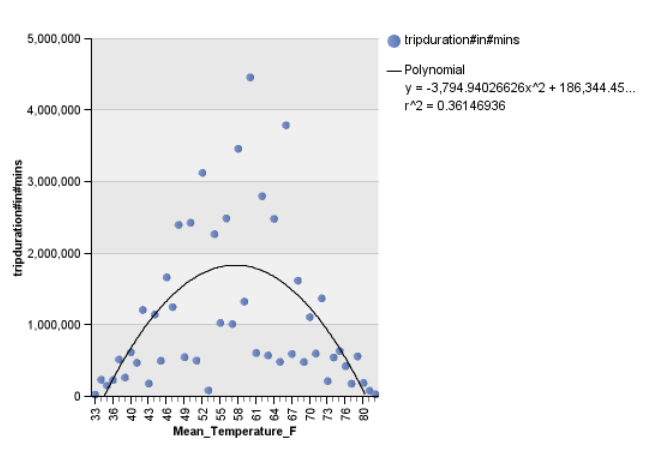
A. Title: Analyzing correlation between trip duration and temperature by Yao Nie

B. Problem statement: In this case, the problem find out the correlation between temperature and trip duration.

C. Brief overview: The procedure was to plot a regression line between temperature and trip duration points.

D. Analysis: From chart below, there is a positive relationship between temperature and trip duration when temperature is between 33 and 58, and negative relationship when temperature is between 58 and 90. R-squared of the regression line is about 36%, which is acceptable.

E. Managerial conclusion: Based on the result, the company should find ways to attract new user in Fall and Winter in order to optimize the revenue.



# Analysis Ten

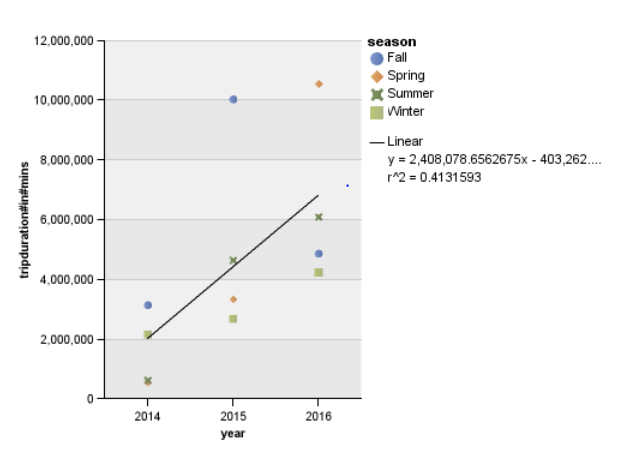
A. Title: Analyzing trip duration growth by years by Yao Nie

B. Problem statement: In this case, the problem find out the correlation between years and trip duration in total.

C. Brief overview: The procedure was to plot the scatter plot between years and trip duration, then plot the regression line on the scatter plot.

D. Analysis: From the chart below, there is a positive relationship between years and trip duration, implying that the usage in Citi bike is gradually increasing by time. R-squared in this model is 41%.

E. Managerial conclusion: As the regression suggests, the company should add equipment or station to satisfy the growing usage and to optimize the company’s profit.



# Analysis Eleven

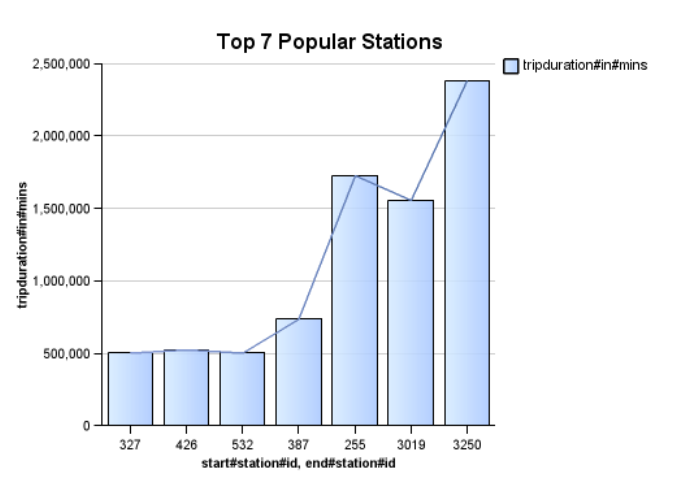
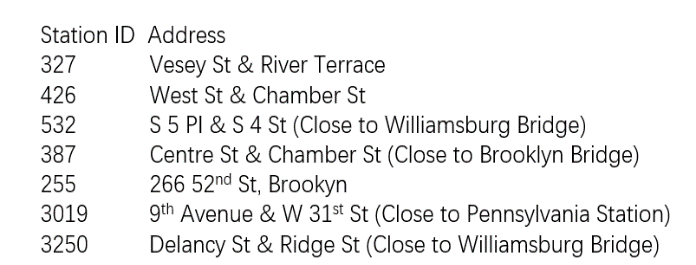
A. Title: Top 7 popular stations of Citi Bike in NYC by Yao Nie

B. Managerial problem: Managers of Citi Bike need to know about which station is popular in NYC because popular stations have larger customer flow than other stations.

C. What was done: Draw a bar chart to show total trip duration amount in each station by using Cognos Report studio. Apply a filter to show only top 7 stations according to the of trip duration amount.

D. Analysis: The result shows the top 7 popular stations in NYC. It shows that the most popular stations are close to major public transportation areas, such as Pennsylvania Station, Brooklyn Bridge and Williamsburg Bridge.

E. Managerial conclusion: Managers can improve quality of facilities in the top 7 popular stations and other stations near to major public transportation area. Also, managers can consider to add more stations and staffs in these areas to satisfy customers’ needs.



# Analysis Twelve

A. Title: Trip duration of users in different ages by Yao Nie

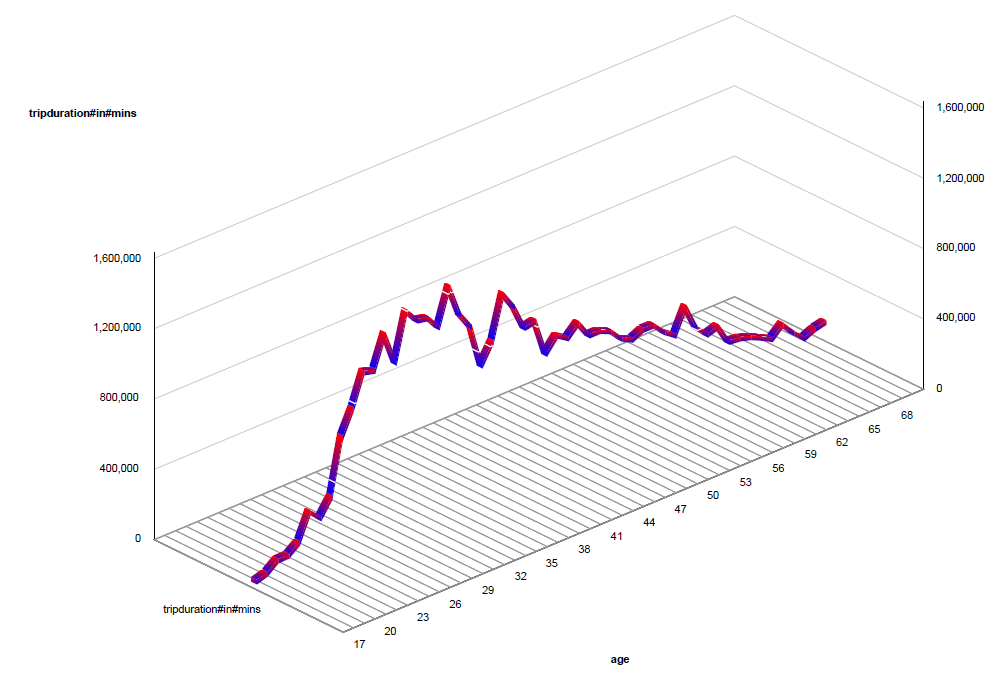
B. Managerial problem: In order to understand users deeply, correlation between users’ age and their trip duration is essential for managers.

C. What was done: Use Cognos Report studio to show trip duration of users in each age.

D. Analysis: The result shows that users in 30-50 years old are more likely to ride than other users, with a high total trip duration. Most people in 30-50 years old are office staffs and they are more likely to use Citi Bike for commute.

E. Managerial conclusion: Users from 30 to 50-year-old are recommended to be considered as the most important target customers.

**Trip duration of users in different age**



# Analysis Thirteen

A. Title: Correlation between trip duration and precipitation by Yanling Peng

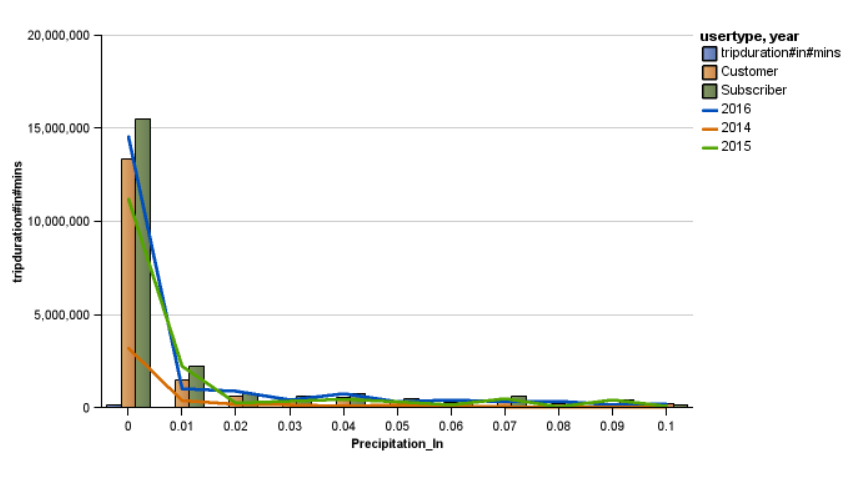
B. Managerial problem: Precipitation is an important factor of weather condition. Finding correlation between precipitation and trip duration of users is helpful for managers to make business decisions.

C. What was done: Use Cognos Report studio to show trip duration of users in each precipitation amount.

D. Analysis: There is a negative correlation between precipitation and trip duration. The result shows that users barely ride in rainy days.

E. Managerial conclusion: When managers consider expanding into new market, they should consider precipitation condition before jumping into the target market.

**Correlation between trip duration and precipitation**



# Analysis Fourteen

A. Title: Correlation between trip duration and max gust speed by Yanling Peng

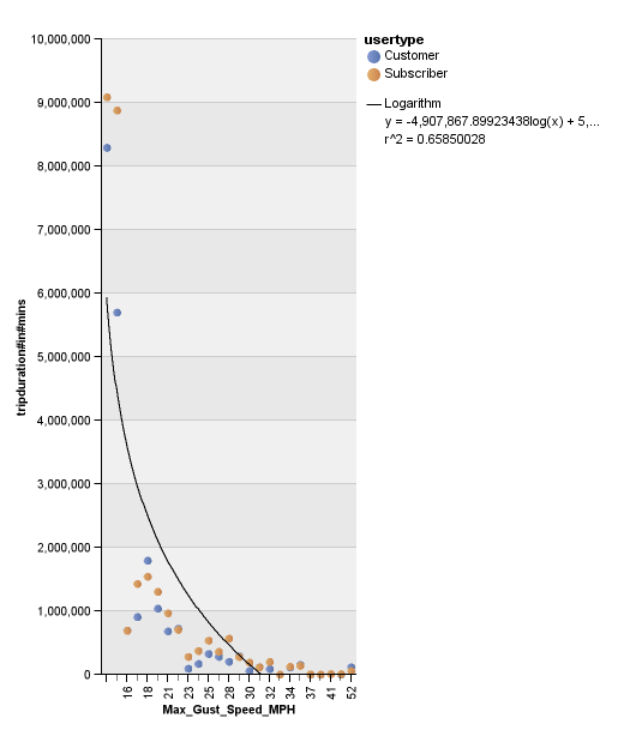
B. Managerial problem: Finding correlation between max gust speed and trip duration of users is helpful for managers to predict Citi Bike usage.

C. What was done: Use Cognos Report studio to make a regression analysis between max dust speed and trip duration.

D. Analysis: There is a negative correlation between max gust speed and trip duration. The result shows that when gust speed reaches 30 miles per hour, people’s intention to ride decreases.

E. Managerial conclusion: Managers can use this regression chart to predict Citi Bike usage according to gust speed. For example, in areas that has high gust speed, managers are recommended to decrease the price of Citi Bike.

**Correlation between trip duration and max gust speed**



# Analysis Fifteen

A. Title: Correlation between trip duration and humidity by Yanling Peng

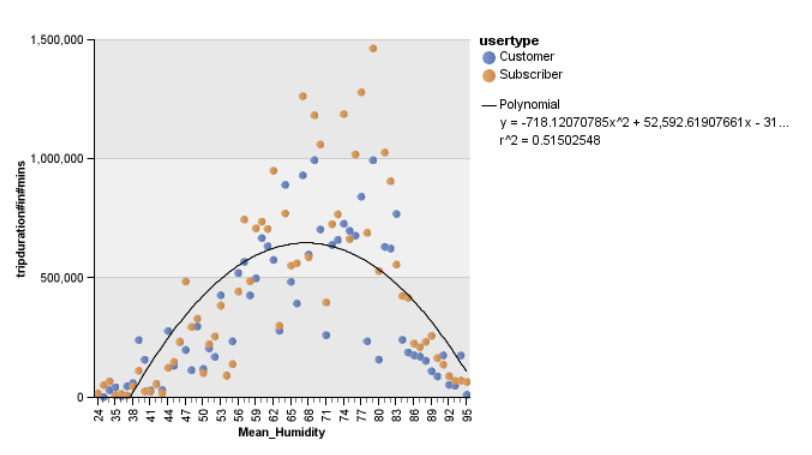
B. Managerial problem: Finding correlation between humidity and trip duration of users is helpful for managers to predict Citi Bike usage.

C. What was done: Use Cognos Report studio to make a regression analysis between mean humidity and trip duration.

D. Analysis: There is a positive correlation between mean humidity and total trip duration before mean humidity reaches 68. And once the mean humidity reaches 68, the correlation becomes negative.

E. Managerial conclusion: The result shows that when humidity reaches around 68, the trip duration of users is the highest. Managers can use this regression chart to regression chart to apply appropriate strategies toward different humidity conditions.

**Correlation between trip duration and humidity**



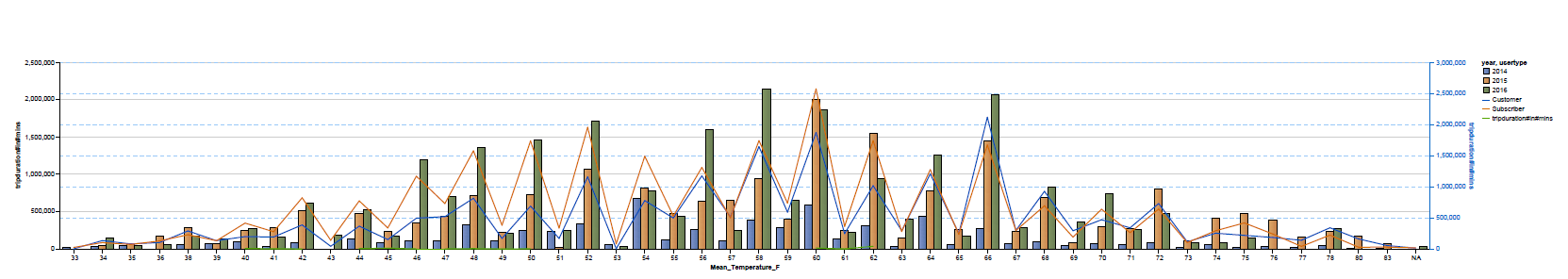
# Analysis Sixteen

A. Title: Correlation between trip duration and temperature by Yanling Peng

B. Managerial problem: Finding correlation between temperature and trip duration of users is helpful for managers to predict Citi Bike usage.

C. What was done: Use Cognos report studio to create a bar chart to display the average trip duration in different temperatures.

D. Analysis: The result indicates a gradually increase in Citi Bike customers. Also, when temperature reaches 58 degree to 70 degree, Citi Bike customers tend to grow.

E. Managerial conclusion: The result shows that Citi Bike customers have gradually increased from 2014 to 2016. In addition, the most amenable temperature for people to ride is from 58 degree to 70. The company is recommended to apply appropriate strategies towards different temperature conditions, such as apply promotions in harsh days and invest more resources in amenable weather days.

**Trip duration in different temperature**