

# Boston Public School Bus On Time Performance Project Final Report

MSSP Team: Luying Cao, Yaqi Huang, Sijia Li, Xuechun Liu,  
Yuan Yuan, Emanuel Zanzerkia

## **Abstract**

Our analysis looked into the factors associated with the arrival delay of school buses serving Boston public schools. Overall, we found drastic reduction in delays as the semester progresses, which may be due to a combination of learning by the drivers and the students as well as adjustments made on the routes. After accounting for this learning/adjustment curve, we've identified several potential factors that are associated with the delay of individual routes, such as weather, traffic condition and the proportion of Door to Door riders. We also looked into the reduction in delays from the 8:30 to 9:30 time tiers and concluded that distance between the two schools is one of the key determining factors that help alleviate the delay. We finally built a predictive model that was able to make accurate prediction of the delays that also confirmed one of our previous findings that the proportion of Door to Door riders has positive effect on the delay time.

## **Introduction**

### **Background:**

The Boston Public School system deals with massive logistic challenge every day; busing 50,000 students around town to over 200 schools in Boston during the most congested hours of the morning. Although school buses are present in most cities throughout US, the situation in Boston is uniquely complicated by the fact that families can apply and attend any school in the district. Historically, this is due to Boston moving away from the neighborhood based system during the 1970's in an attempt at desegregate the school system.

The current school assignment method is that students in grades kindergarten through 8th grade, have a home based assignment policy while high school students can attend any school within the city. For the home based assignment policy, each family is given a list of schools based upon their home location. The family then ranks the schools in order of which they would prefer to attend. The students are then assigned to a school based upon their school rankings.

With this school assignment policy and student IEPs requiring specialized schools, the busing issue becomes extremely complex. Buses potentially have to transport students across town and even out of town for specialized schools. With such complexity, it comes as no surprise that there are some major issues that need to be resolved. Boston Public School spends more money on transporting students to school than any other school system in a major U.S. city. Approximately 10% of the school budget is allotted for transportation. This is a significant portion of the school budget. Another major issue is the on-time performance of the school buses. It is this last issue that will be the focus of this analysis and report.

When students arrive late to school not only are the students affected, the teachers and classes are affected when a large percentage of students are late. Therefore, ensuring students arrive at school on time is very important.

### **Outline of Report:**

This report will consist of four main sections. We start off describing the data that was supplied and further data that was procured from other sources. We then move onto explanations of the methods we used for our analysis of the data followed by the results from the analysis. Finally we wrap up with our final conclusions. All the supporting materials will be included in the appendix at the end.

## **Data**

### **Data Sets Provided by Boston Public Schools:**

The data sets provided by Boston Public Schools include one file of daily buses' on-time performance data from September, 2017 to January, 2018, and one file of bus routes information on January 26th, 2018.

For our analysis, we focused on the daily buses' on time performance data and looked only at Fall 2017. The data included information about schools, bus yards, bus types, bus stops, bus routes, departure and arrival times per bus. Every school has multiple routes going toward it. Drivers are assigned to a specific bus, and they may run multiple routes (time tiers) per day, but

only one route per time tier. Every route has a planned start and anchor time, arrival delays are defined by the amount of time a bus arrived after its planned anchor time.

### **Other Data That was Procured:**

During our initial exploration of the data given to us by Boston Public Schools, we realized that there may be other data that would be useful in understanding the on time performance of school buses. One of the first things we thought of acquiring was historical weather data. Using the weather underground API, we were able to get indicators on whether there was fog, rain, snow, or hail at the time of each route. Also we were able to get a measure of visibility on a 0 to 10 scale.

The next thing we tried to quantify was the amount of traffic on the routes that the buses would take. Since we did not have access to turn by turn details of each route, we had to approximate the route using its starting point and its endpoint. Furthermore, since we were not able to gain access to historical traffic patterns provided by TomTom, we decided to quantify traffic by the extra time added to a route due to traffic. Using Google's distance matrix API, we were able to obtain Google's best guess time for a bus to travel from its starting point to its endpoint. We were able to get the times for when the bus actually took the route, and a time when there would be minimal traffic (we used 3:30 am). Once we had both of these times, we were able to look and calculate the difference in time. This gave us an estimate on how much traffic adds to the travel time of each route.

Finally, we classified each route as either inbound or outbound, and as either "over bridge" or "not over bridge". To classify routes as either inbound or outbound, we first had to define an area of Boston that would be considered the center of Boston. If any route passed through this area, it would be considered an inbound route. If however this criteria was not met, then we looked at the distance from the center of this region to the starting and end points. If the distance to the starting point was greater than the distance to the endpoint, the route would be classified as inbound otherwise the route would be classified as outbound. As for classifying routes as going over a bridge, we separated the greater Boston area into three main sections,

Cambridge/Charlestown, East Boston, and Boston. If a route started in one of these areas and ended in another area, we classified that route as going over a bridge.

## **Methods of Analysis**

During the initial visual inspection, we realized certain avenues of exploration were not suitable to the final goals of the project while others warranted further exploration. We then moved on to modeling the data using multiple techniques for validation of results.

**Models** ( Specific explanation is in the appendix):

### **Random Forest**

Random forest is a tree-based learning method for classification, regression. It can be used to rank the importance of variables in a regression or classification problem in a natural way.

### **Gradient Boosting**

Besides random forest, we also used another machine learning technique called gradient boosting. It is a very similar approach to random forest but gradient boosting could give better predictions results than random forest.

### **XGBoost**

Similar to random forests and gradient boosting, XGBoost is also a tree-based model which belongs to the boosting models category. It is well known for its performance speed and prediction accuracy.

### **Multilevel linear model**

In a regression model, we create a function using the feature variables to predict the value of the response variable. The problem with simple regression is that it does not take into consideration any type of hierarchical grouping. For example if we are trying to predict student test scores, we have to take into consideration that students are grouped into different schools.

There are some variations between different schools and variations between different students. For those individual-specific difference, we would call them random effect. This method controls for group-level variation, which would give us better estimations.

## Results

We will first go over the results of the initial exploratory data analysis that led to further avenues of exploration. All other data visualizations can be found in the appendix. We will then move onto the results of the various models that were fit using the data.

To start our analysis, we wanted to understand the arrival delay response variable. For every route that was traveled, arrival delay was calculated as the difference in time between the planned anchor time and the actual arrival time. Positive arrival delay values indicate a late bus while negative values indicate a bus that has arrived early.

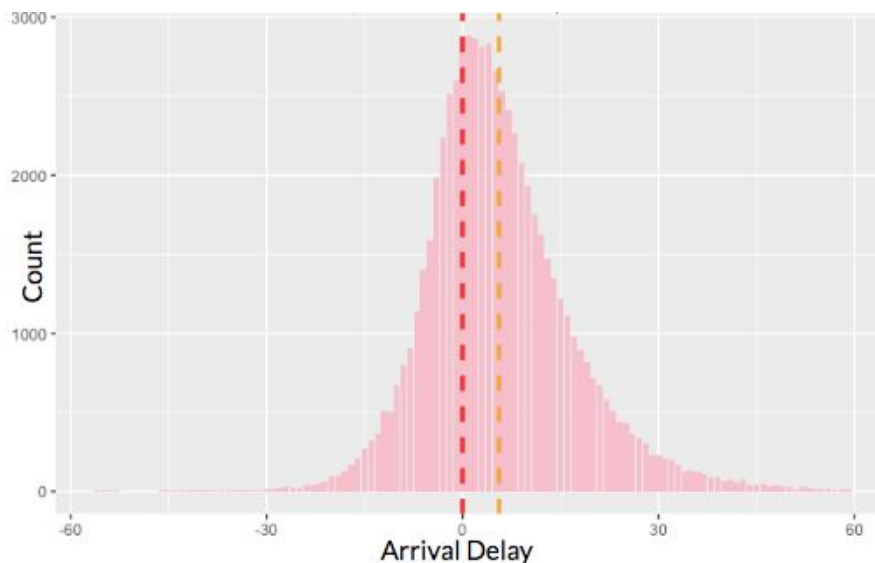


Figure 1

Figure 1 shows the overall distribution of arrival delay with the red dashed line showing an arrival delay of 0 minutes and the orange dashed line showing the mean average delay of 6 minutes. An average delay of 6 minutes doesn't seem to be too bad, but the large variance of

arrival delay is concerning. The figure shows some buses arriving at their destination almost an hour early while other buses are arriving almost an hour late. We also see that the a higher proportion of routes have positive arrival delays.

We wanted to see what happened to arrival delay as the school year progressed. The idea is that as drivers learned their routes and become more comfortable with them, the arrival delay would decrease.

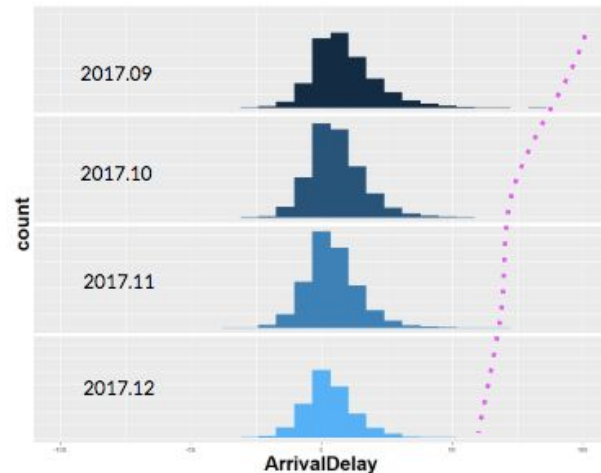


Figure 2.1

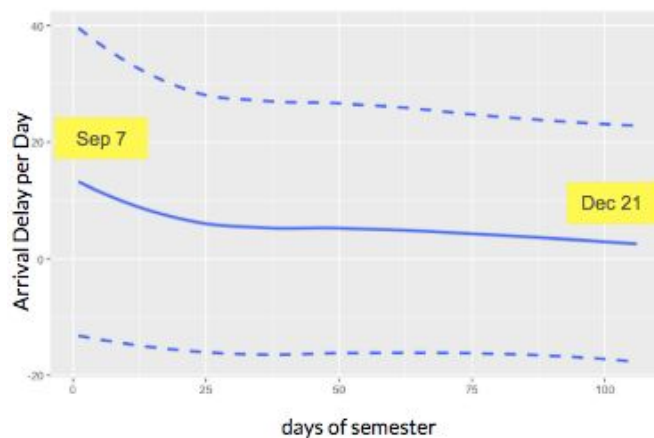


Figure 2.2

Figure 2.1 shows that the delay time decreases as the months progress. In September the proportion of buses that arrived late was the largest, but each month after that the proportion of late buses becomes smaller. The dashed curly purple line in figure 2.1 shows the improvement.

Figure 2.2 tells the same story. We used data from September 7 which is the beginning of the semester to December 21 which is the end of semester. The blue line in the center shows

that as the days of semester increases, the average delay per day decreases. The region between the 2 dashed line represents the 95% confidence interval. Even though this is a good trend, there are still large number of buses that arrive late.

We also wanted to look at arrival delay for the different school time tiers to understand what happens to the arrival delay as the day progresses.

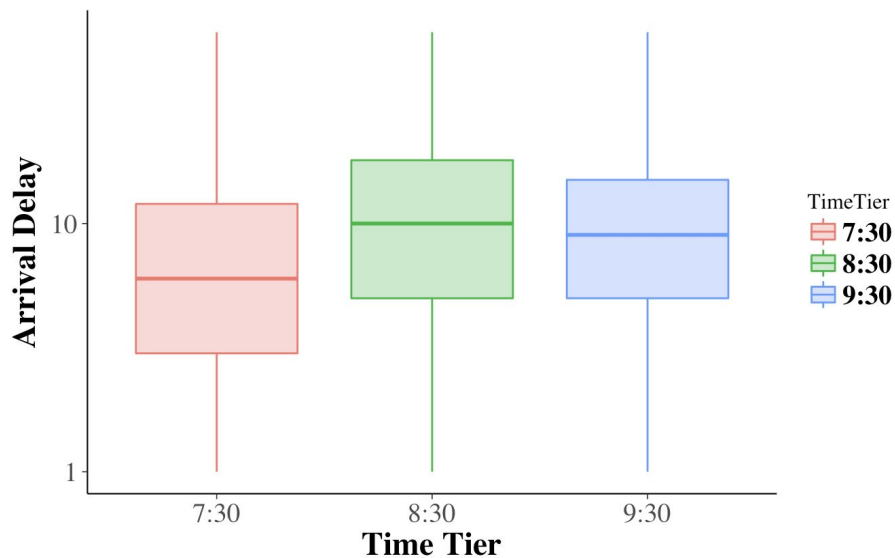


Figure 3

From figure 3, we see that buses running the 8:30 time tier generally have larger arrival delays than the 7:30 time tier and the 9:30 time tier. In general, suppose a bus is running all three time tiers in the morning. We would expect an increase in delay time along time, which means if a bus is running late for 7:30 time tier, we would expect the bus to be running even later for the 8:30 time tier and also the 9:30 time tier. However, we found that this is not the case. Examples of buses who follow this pattern are B439,MS145,HS283. The factor that could be causing this pattern is the total distance traveled from the 8:30 school to the 9:30 school.



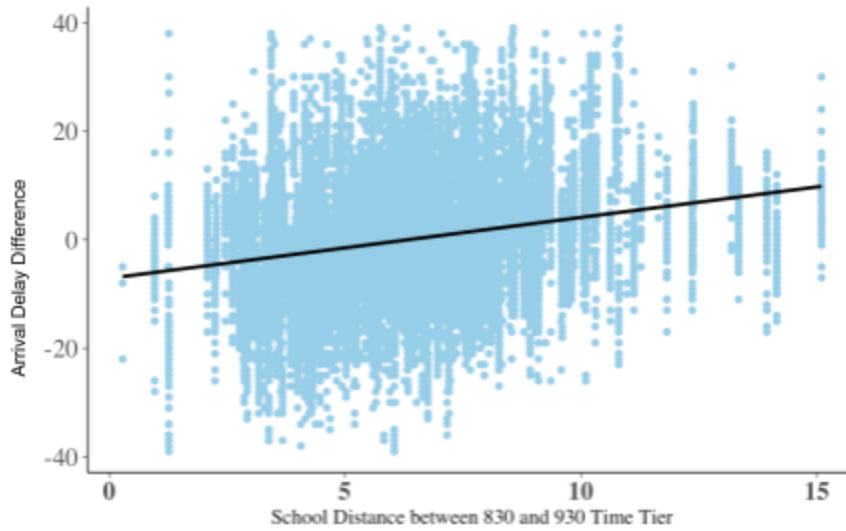


Figure 4

Figure 4 shows the relationship between distance travelled and the arrival delay difference, which was calculated by taking the difference between the arrival delay for the 9:30 time tier and the arrival delay for the 8:30 time tier on the same bus. We see in figure 4 that there is a positive trend between distance travelled and the difference in delay, which indicates that the further the distance travelled the more delayed the 9:30 time tier route will be.

There are still some buses who do not follow this pattern, which means there is an accumulation in delay performed by this bus, such as HS242, HS226, MS231. Take an example of MS231. Figure 5 shows the overall performance done by the bus MS231 for the fall semester on a daily base:

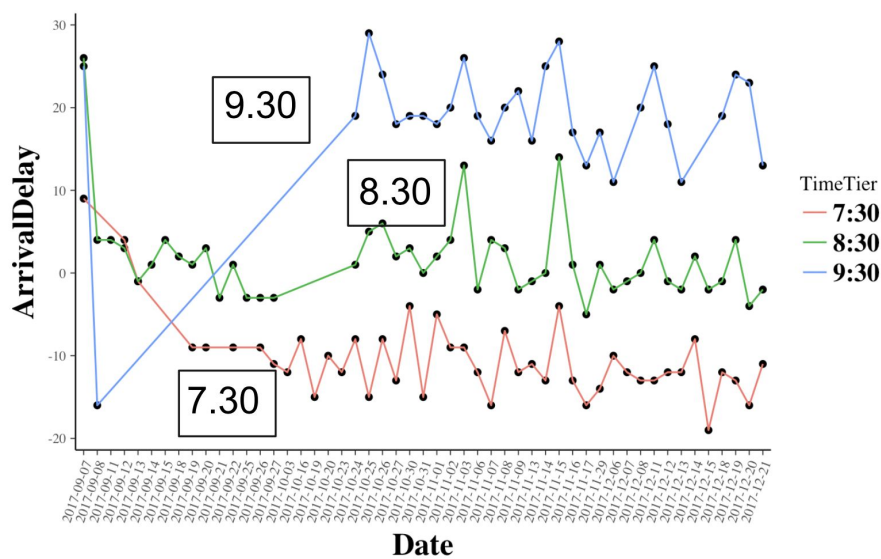


Figure 5

We can see that starting from Oct 24th, this bus started running all three time tiers every day. Overall, the 9:30 performance is worse than the other two time tiers. We can see that there is an accumulation in delay. In this case, we still consider whether distance could be a factor here. Figure 6 shows the locations of the three schools the bus travelled to and the general direction it travelled.

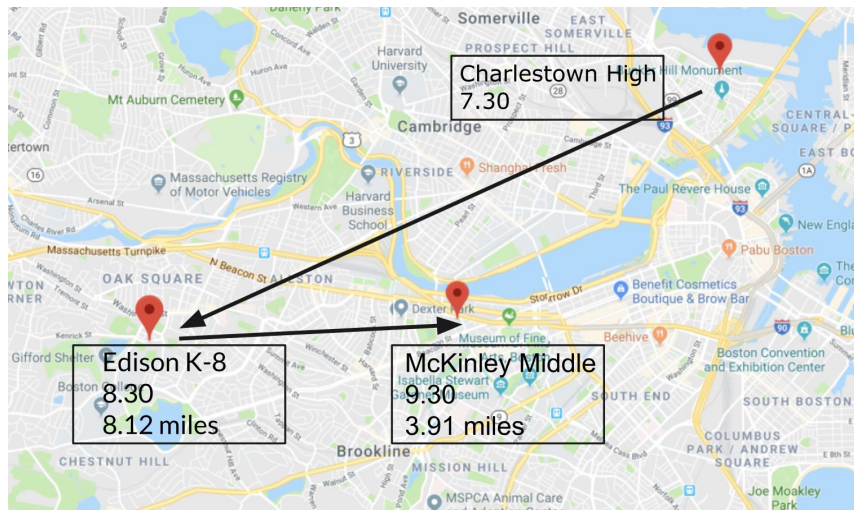


Figure 6

Bus MS231 travelled from Charlestown High School to Edison K-8 School for 8:30 time tier with the a distance travelled of 8.12 miles. The bus then travelled to McKinley Middle School for the 9:30 time tier with a distance travelled of 3.91 miles. We can see that in this case, the shorter distance travelled did not make the bus perform better at the 9:30 time tier.

## Correlation Analysis:

Correlation analysis is a way to find relationships between the factors of interest and the response variable which in this case is the arrival delay. When a factor has a positive correlation with the response variable, it means that for every increase in that factor, there is an increase in the response while a negative correlation shows the opposite reaction. Furthermore, the larger the absolute value of the correlation is, the stronger the relationship between the factor and response variable is. The following figures show the correlation of the variables of interest in this analysis with respect to the arrival delay.

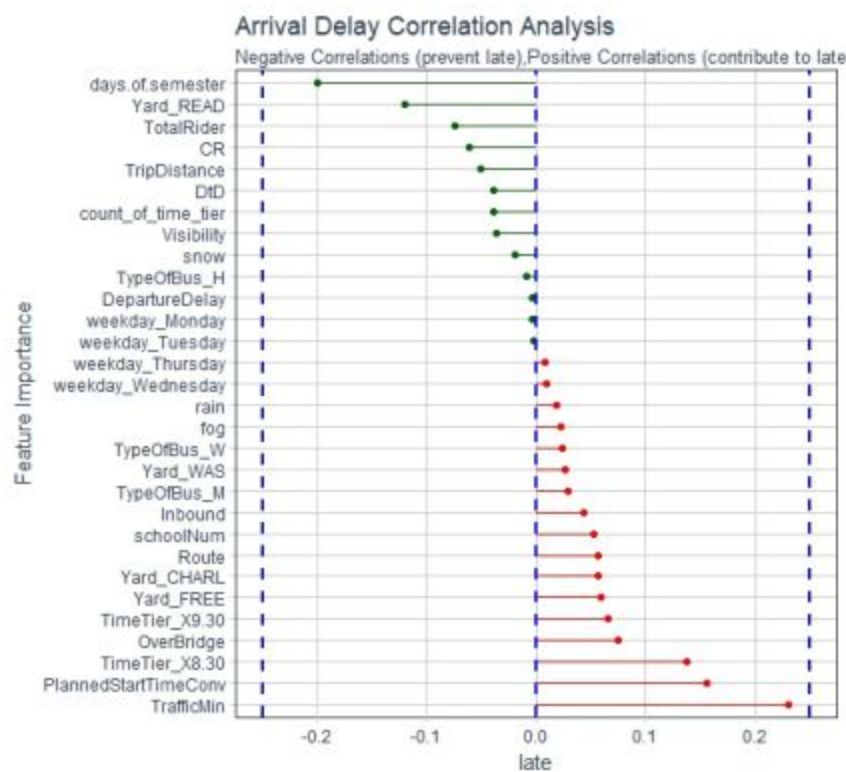


Figure 7

From figure 7, it seems as though traffic minutes, planned start time, and 8:30 time tier are the factors with the strongest positive correlation. That is they contribute the most to increasing arrival delays. On the other hand, day of semester, bus yard Readville and total riders are the top three factors with the strongest negative correlation. That is they contribute to most to decreasing arrival delays. Figure 7 shows the correlation for the whole data set, but we also wanted to look at how these correlations change over time. We accomplished this by analyzing

the correlations for the September data and for the December data. The following figures display this information.

### September

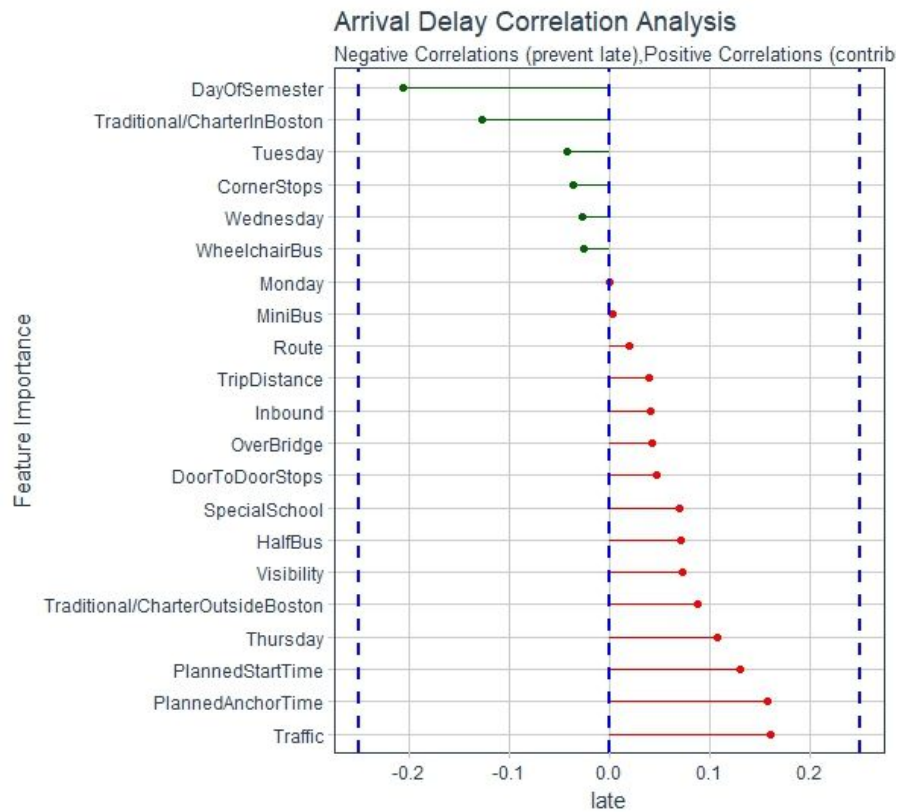


Figure 8.1

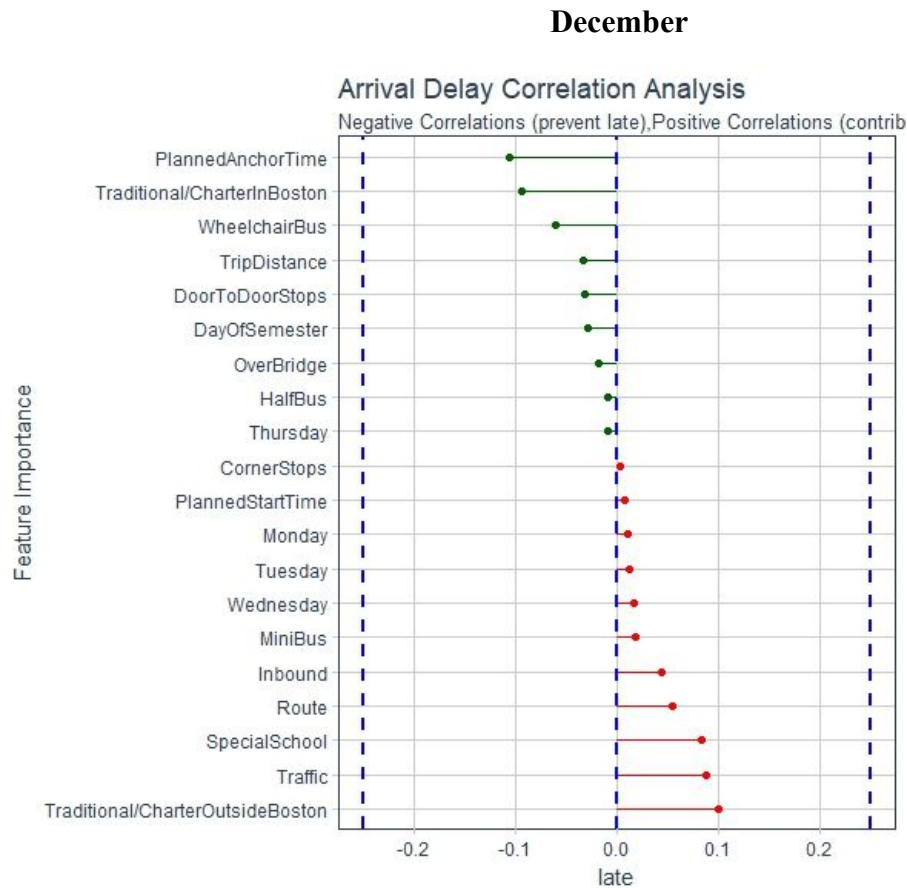


Figure 8.2

From figures 8.1 and 8.2, we see that traffic and over bridge routes are still the top positive correlated factors for September and December, but the top negative correlated factor changes from day of semester to trip distance.

Correlation analysis is useful exploratory data analysis tool since it is fairly easy to interpret and allows us to look into all factors at the same time. However, correlation analysis is limited to only giving us a rough sense of what is happening with the data and the interpretation should be taken with caution since it does not consider between factor interactions nor does it take into account the inherent structures in the data. In order to make claims about dependence of certain predictors in the presence of other factors, one needs to fit a model that can control for such dependencies.

### Random Forest Model :

As was mentioned in the Methods of Analysis section, random forest models are easy to understand and make fairly accurate predictions. For this analysis, we wanted to use the feature grading aspect of random forests so that we could identify which features were the most impactful when making the predictions. Figure 9 displays the features in order of their importance.

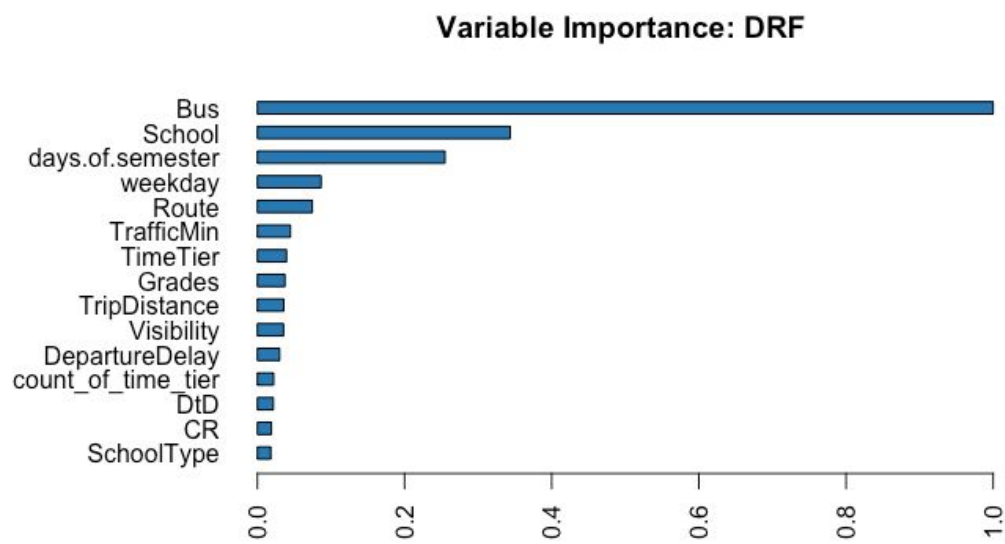


Figure 9

From the figure, we can see that some specific buses have a huge influence on the arrival delay. The days of semester also shows up as being important, which matches some of the results we displayed at the beginning of the analysis. The error rate of this model was a bit higher than desired so we moved on to a gradient boosting model.

### Gradient Boosting:

After finishing the process of gradient boosting, we were able to find the top 15 most important variables in the model. From figure 10, we could see that bus, school, and date are the most important features in the model. It shows identical results to the random forest model. However, buses and schools are not factors that we can control and we can only clearly identify the first three factor contribution. Therefore we switched to extreme gradient boosting in order to identify factors that we could control.

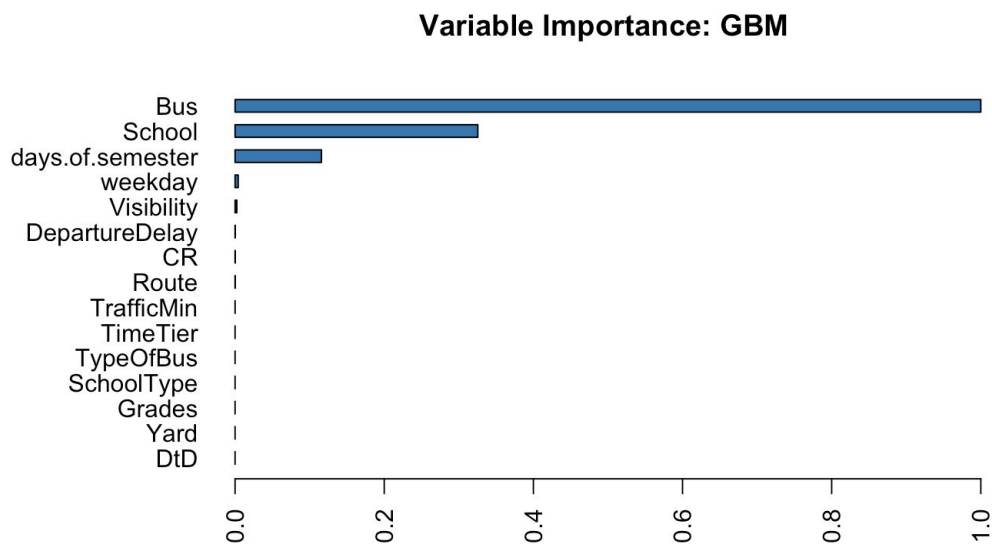


Figure 10

### XGBoost:

Slightly different from the model output for random forests and gradient boosting, xgboost was able to give variable importance on factor levels. Figure 11 illustrate how the model decomposes our response variable arrival delay using different factors. It shows that different factors have different amount of impact and different directions of impact on arrival delays. Like mentioned before, the model has good prediction accuracy. In this case, the model is able to predict arrival delays 10 minutes within its true value.

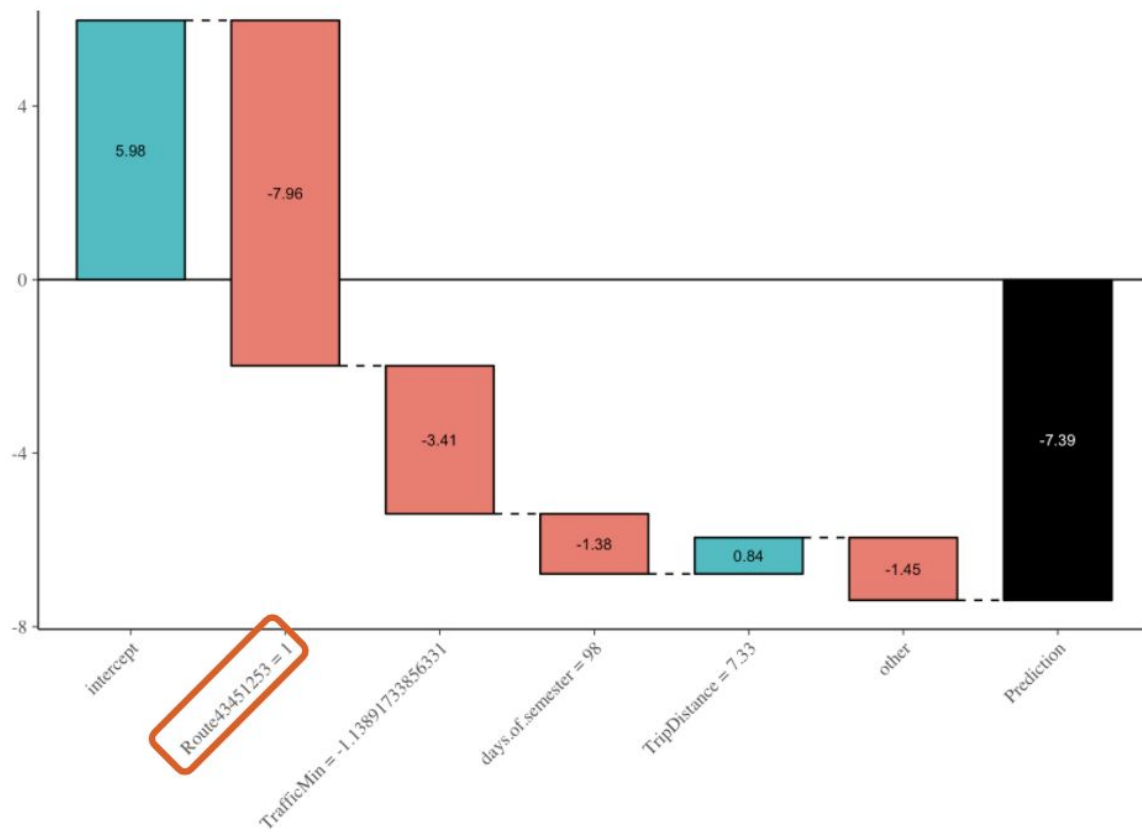


Figure 11

The factors that have relatively big impact (positive or negative impact) on arrival delays are the progress of semester, traffic, allotted travel time, trip distance and the number of total riders. Results are shown in figure 12.



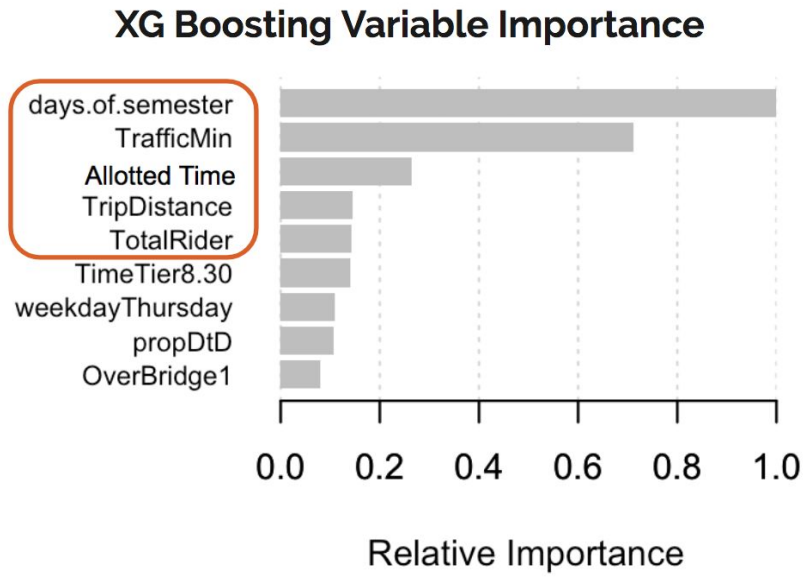


Figure 12

As the semester goes on, drivers get better at arriving on time. This might be due to them getting more familiar with their routes. Not surprisingly, traffic plays an important role. Trip distance and allotted travel time are highly correlated with shorter trips actually having a concentration of higher arrival delays compared to longer trips. This can be seen in figures 13.1 and 13.2

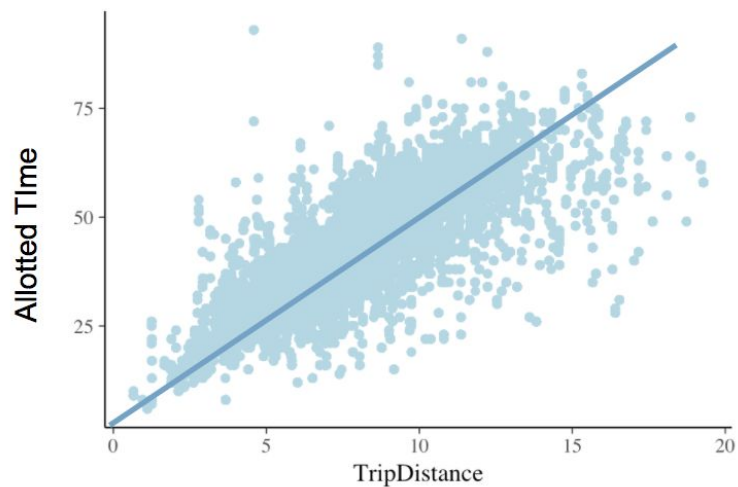


Figure 13.1

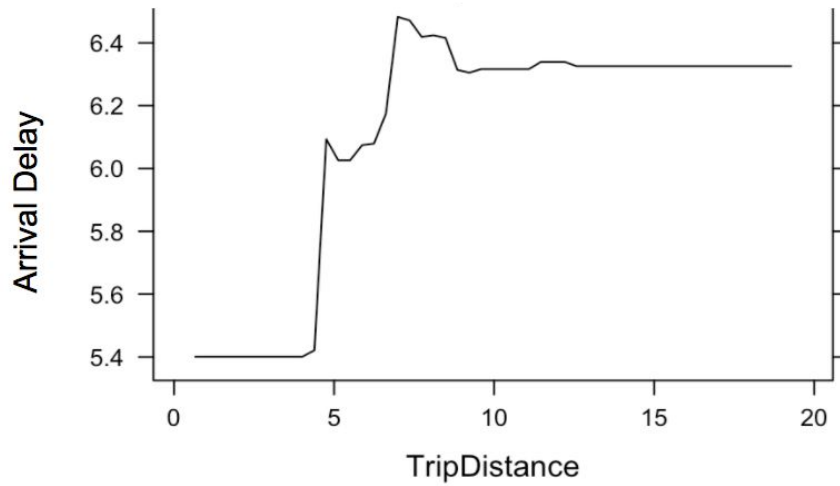


Figure 13.2

The proportion of door to door riders also affects arrival delay. Having higher proportion of door to door riders on board is more likely to delay the school bus. This is displayed in figures 14.1 and 14.2

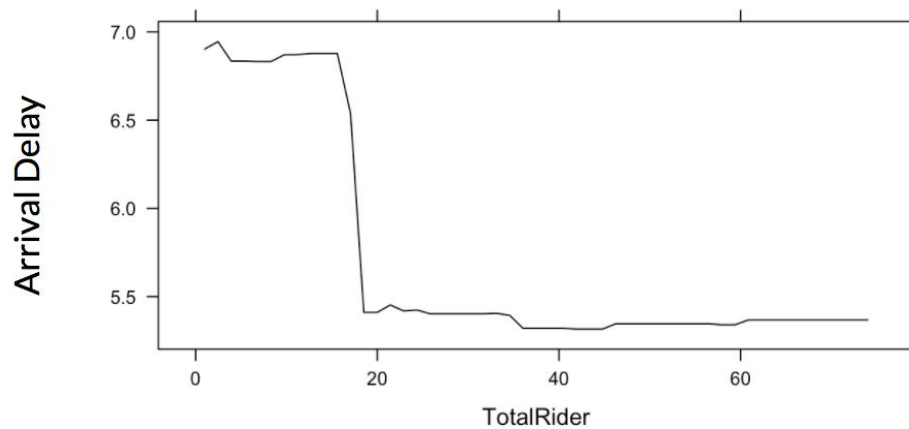


Figure 14.1

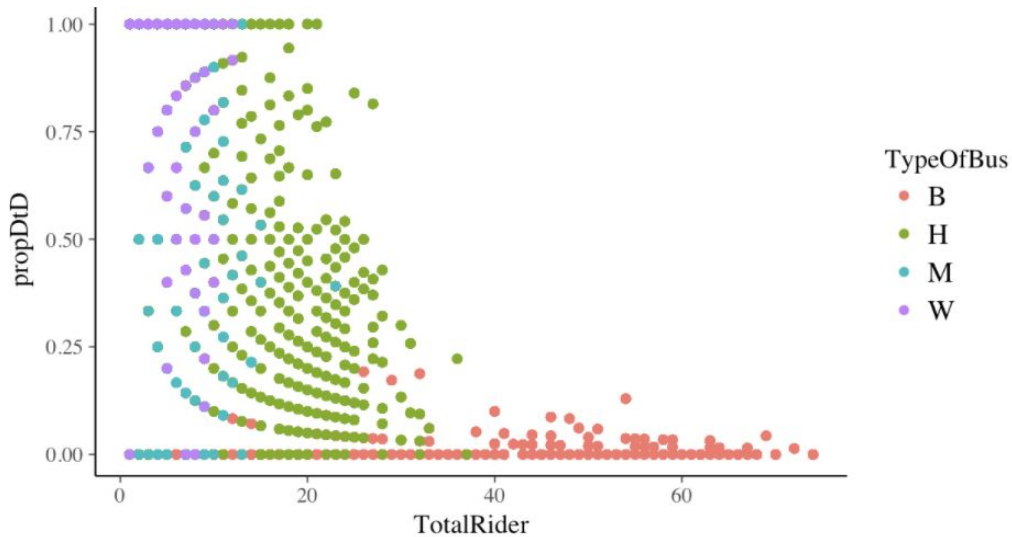


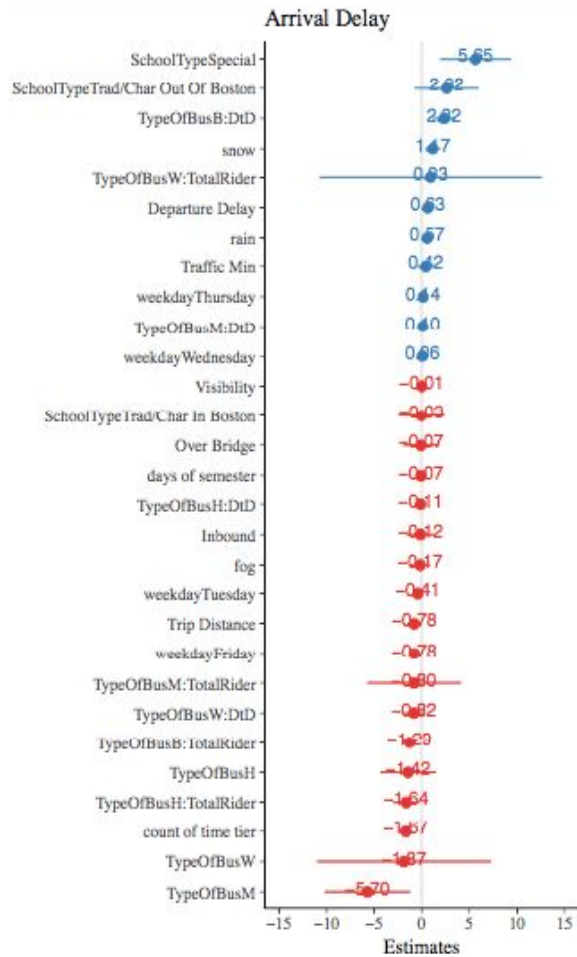
Figure 14.2

Since XGboost is a predictive model, its usage is not limited to current data. Provided with updated information in the new semester, the model would be able to predict unknown arrival delays for future routes.

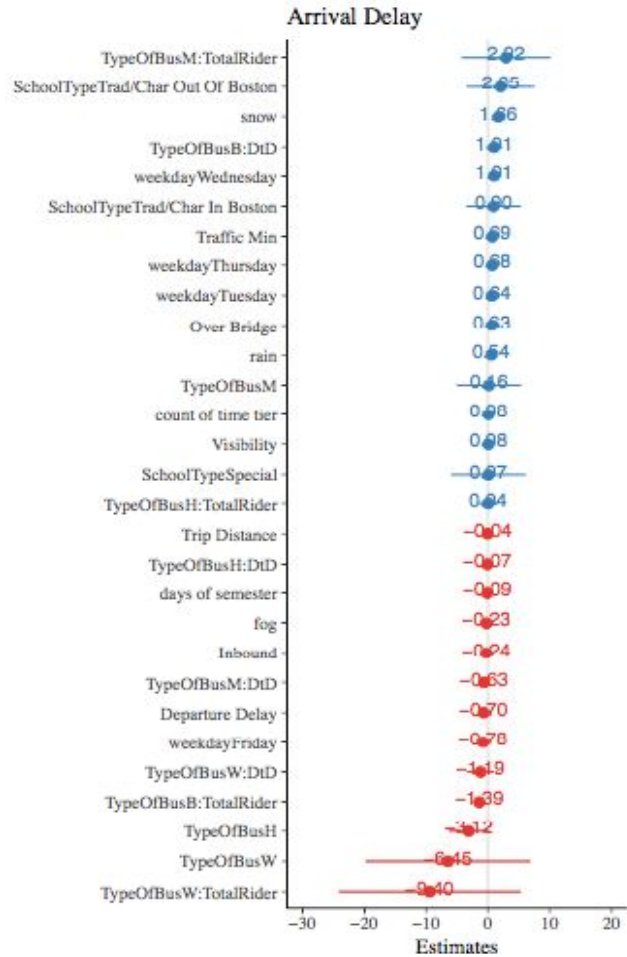
### Multilevel linear model:

We now move on to looking at the results of the the multilevel linear model. In the previous section, we presented the correlation analysis. The results from the correlation analysis should be taken with caution since it only looks at marginal correlation between the outcome and each predictors. Another important issue that correlation analysis lacks, which is particularly important for this analysis, is that it cannot consider the group-level variation. By group we mean schools, buses, and routes where routes are nested within a school but buses can serve multiple schools and routes.

Multilevel linear model is a regression model that is able to take into the inherent structure in the data giving us better quantification of uncertainty. Figure 15 shows the factors that would increase or decrease delay time at different time tiers. If a variable doesn't cross 0, it roughly means that variable is statistically significant at around alpha level of 5% . Blue coefficients presents factors that would induce delay while red coefficients are factors associated with decreases in the delay.



7:30



8:30

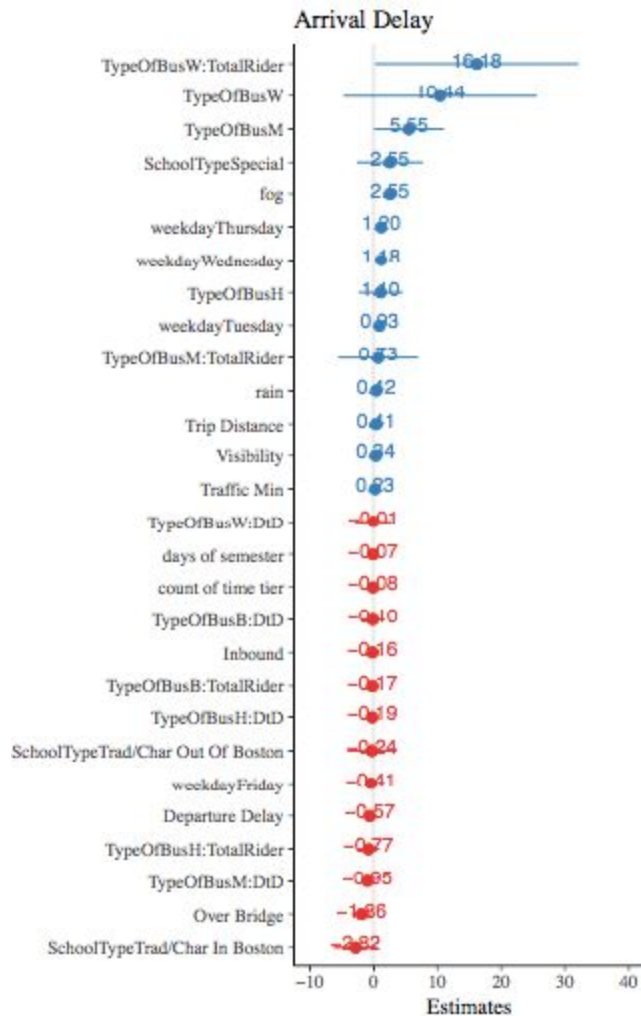


Figure 15

9:30

At 7:30, we find special schools would be about 6 minutes later than private school on average. For the fifth one, big bus combines Door to Door riders, it means that for big buses, one increase on the the number of Door to Door riders, the delay time would increase 3 approximate minutes. Snow shows the positive influence on the delay time. The day which has snow could be 2 minutes later than days which doesn't have snow. Wheelchair buses have huge variation of delay time. Mini buses seem work well at 7:30.

Compare with time tier 7:30, we found most of variables are not significant for time tier 8:30. It seems buses would usually be late on the Wednesday. The variables (big buses with more Door to Door riders) also shows up. Again, it has the positive effect. But the variation of

delay for wheelchair buses become larger. The last one, wheelchair buses with more total riders, has negative effect. This result is different from what we get for 7:30. It confirms there is variation among different time tiers.

At 9:30, we found that mini buses would be about 5 minutes later than big buses on average. For wheelchair buses, if the number of total riders increase by 10, the average delay time would increase by 17 minutes. This gives us a thought that those students may need more time.

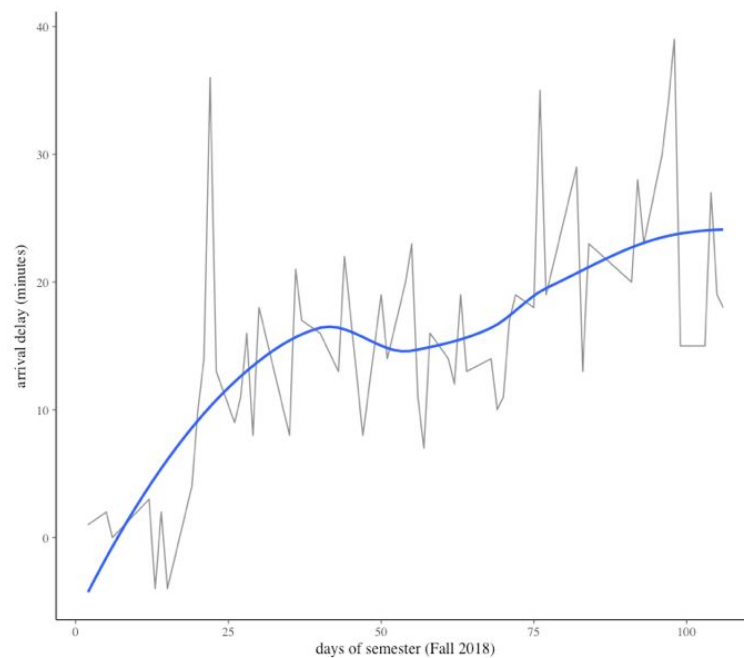


Figure 16

This model also can do analysis for individual route. So we can identify those routes which not perform well as semester progresses. Their delay time has increased from September to December. Here we provide one example (Figure 16). As we show before, for the whole data set, the delay time decrease as semester progresses. But there is huge variation among individual. (The whole list for routes consistently perform bad is in the appendix)

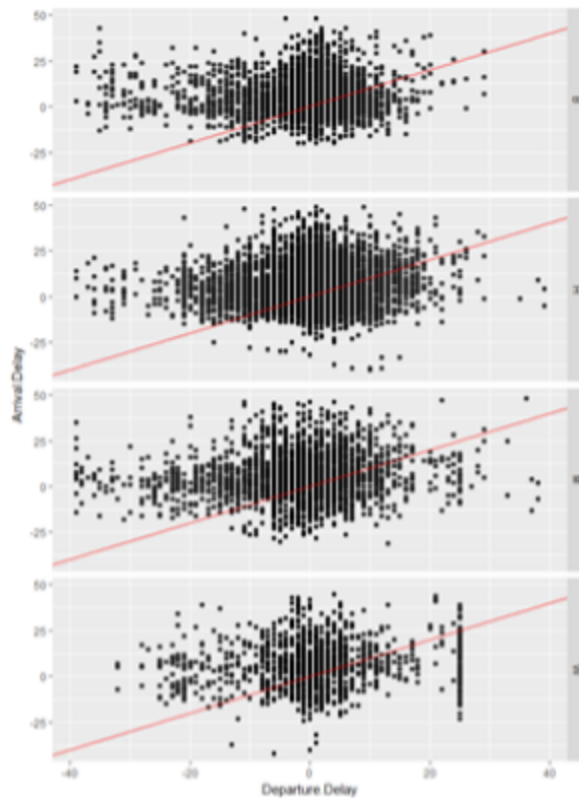
## **Conclusions**

After finishing our analysis, we found that arrival delays decreased as the semester goes on. There is no one single factor that influences arrival delay. Instead, arrival delay is associated with multiple factors including weather, traffic condition, and proportion of door-to-door riders. Therefore, the change of only one single factor would not improve the overall on-time performance of school buses. We also looked into the combined effect of factors and each of their contribution to on-time performance. Ranking the contribution of factors from our models provided us with directions for further analysis, such as traffic budgeting, allotted time optimization, etc. However, at the end of this project, our team realized that there was still room for improvement. For example, if we had a chance to study the actual routes, we might be able to compare whether some specific routes are good or not. In addition, since the data we were analyzing was only for one semester, our analysis would be more accurate and reasonable, and we would be able to make predictions and comparisons along with timeline if we could explore the data for a longer period of time.

## Appendix

This section contains some of the initial exploratory analysis that was conducted but did not find a place in the final presentation.

One of our initial ideas was that if a bus left its starting point late then it would arrive at its destination late. We expected to see that there would be a correlation between departure delay and arrival delay.

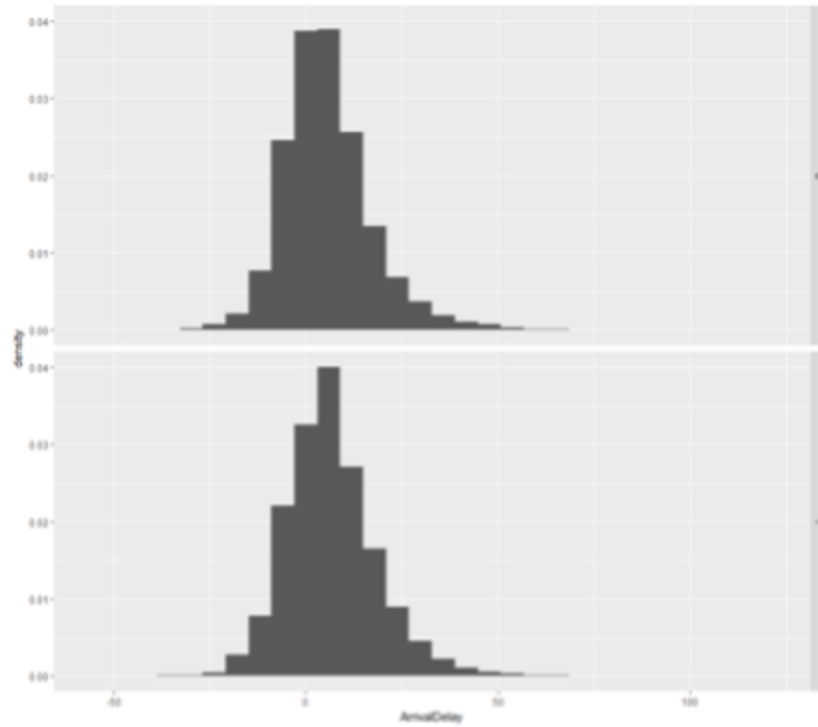


*Figure 16*

Figure 16 shows a plot of arrival delay compared to departure delay for each type of bus. The red line shows the expectation that for every minute a bus departs late, it would arrive a minute late. However, from the plot we see that there doesn't seem to be any correlation between departure delay and arrival delay.

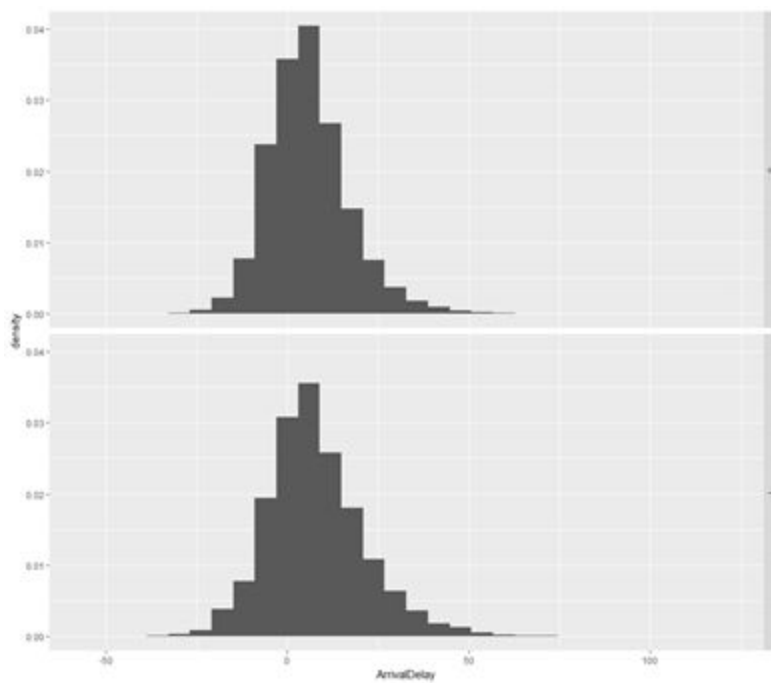
We classified routes as being inbound and outbound. We would expect that inbound routes would face larger delays compared to outbound routes.





*Figure 17*

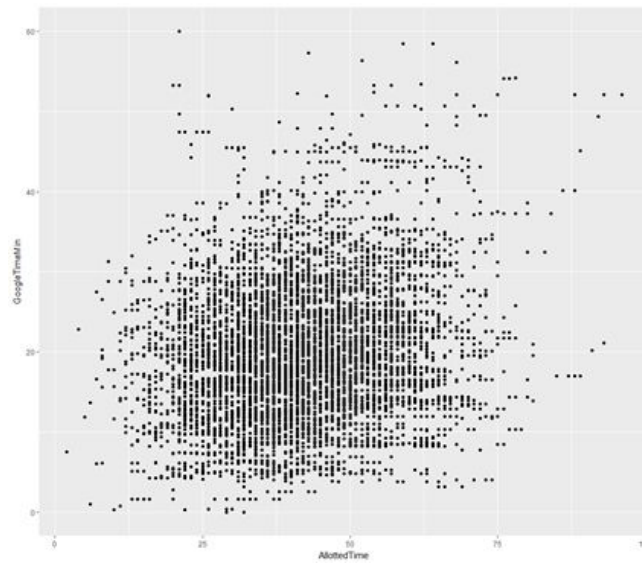
From figure 17, we see that the distribution of delay times for inbound routes and outbound routes are very similar. Similarly, we classified routes as either going over a bridge or not.



*Figure 18*

Again the expectation is that buses that cross over bridges would face larger delay times. However from figure 18, we see that the distribution of delay times for over bridge routes was the same as delay times for routes that do not go over a bridge.

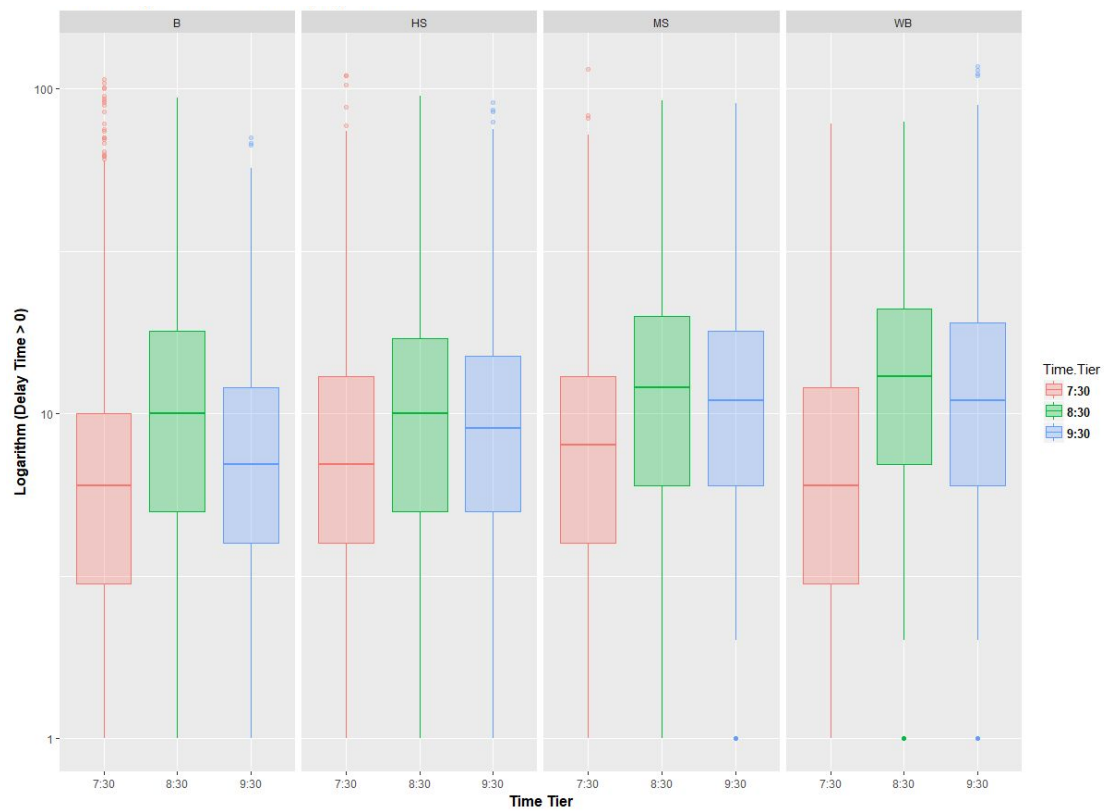
We also looked at the Google predicted time it takes for a bus to go from its starting point to the final destination and compare it to the time the bus has been allotted for its route. The idea is that if the Google predicted time is close to the allotted time, then there isn't time left to pick up students.



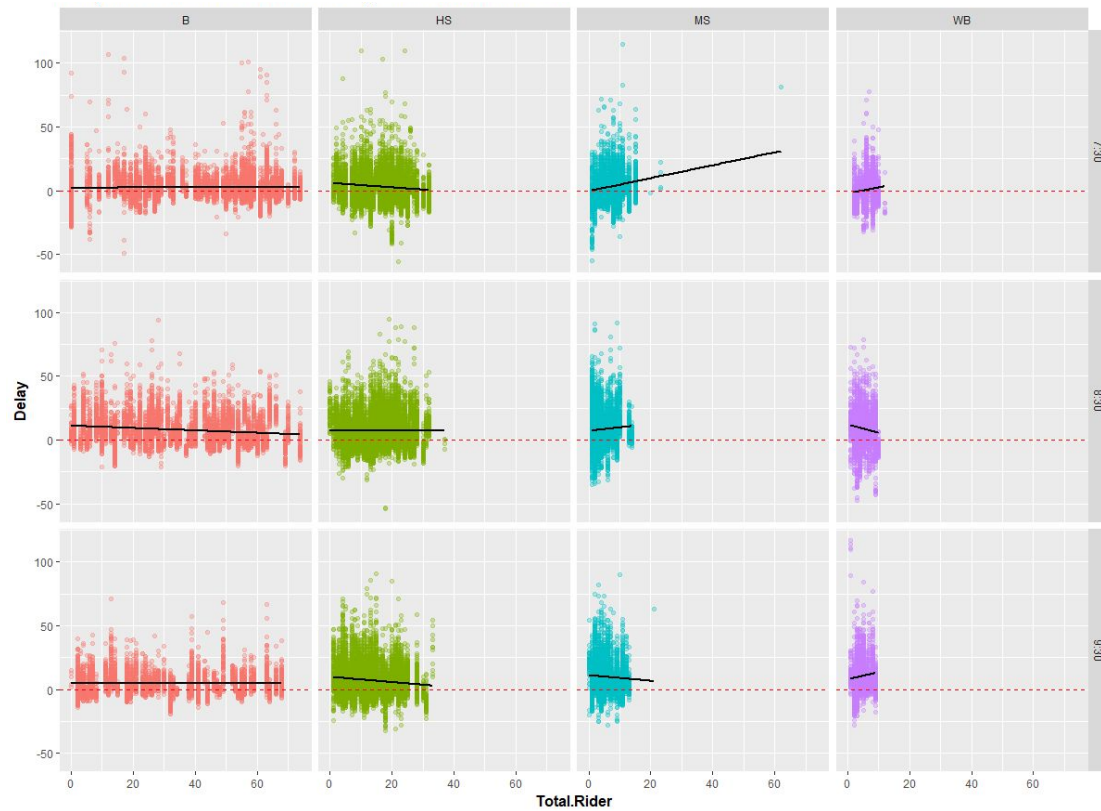
*Figure 19*

Figure 19 shows that the difference and delay time showed no correlation either.

Figure 20. Delay in different time tiers by bus type



There are no patterns shown in the arrival delay with the total number of riders mediating by the interactions between bus type and time tier.



*Figure 21*

Even though there are considerable amounts of early arrival outliers for the yard READ in the dataset, the routes from the yard READ do not have less departure delay or have more on-time departure

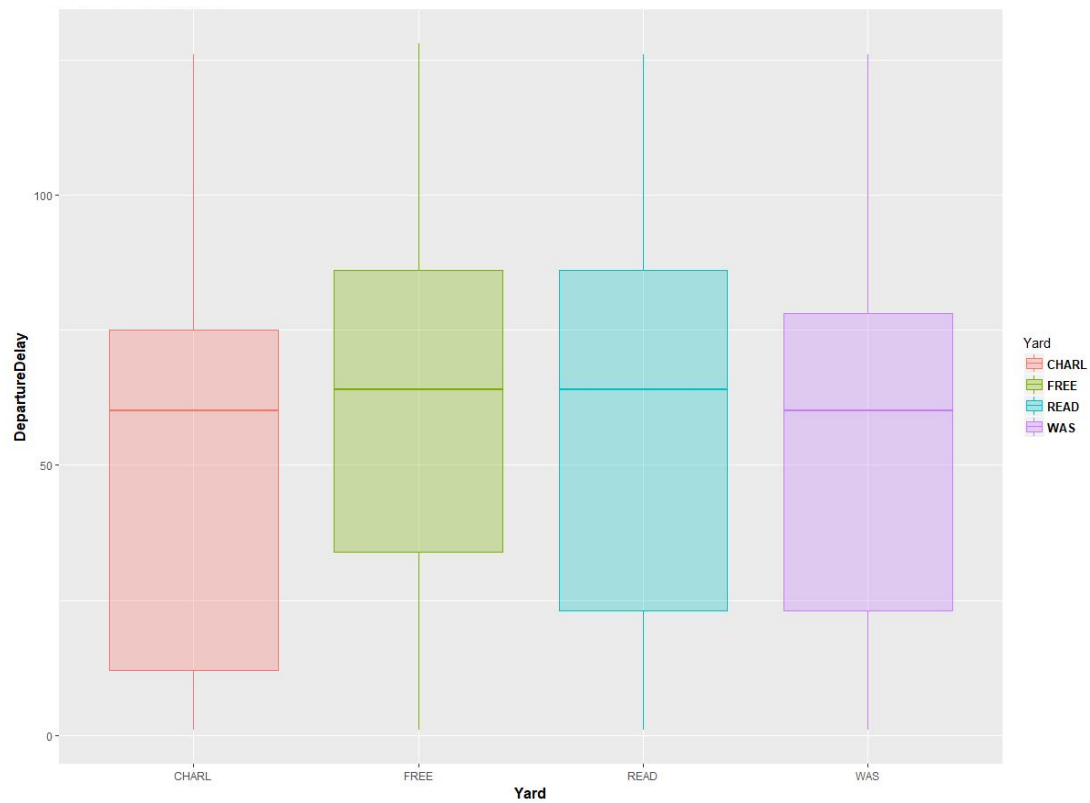


Figure 22

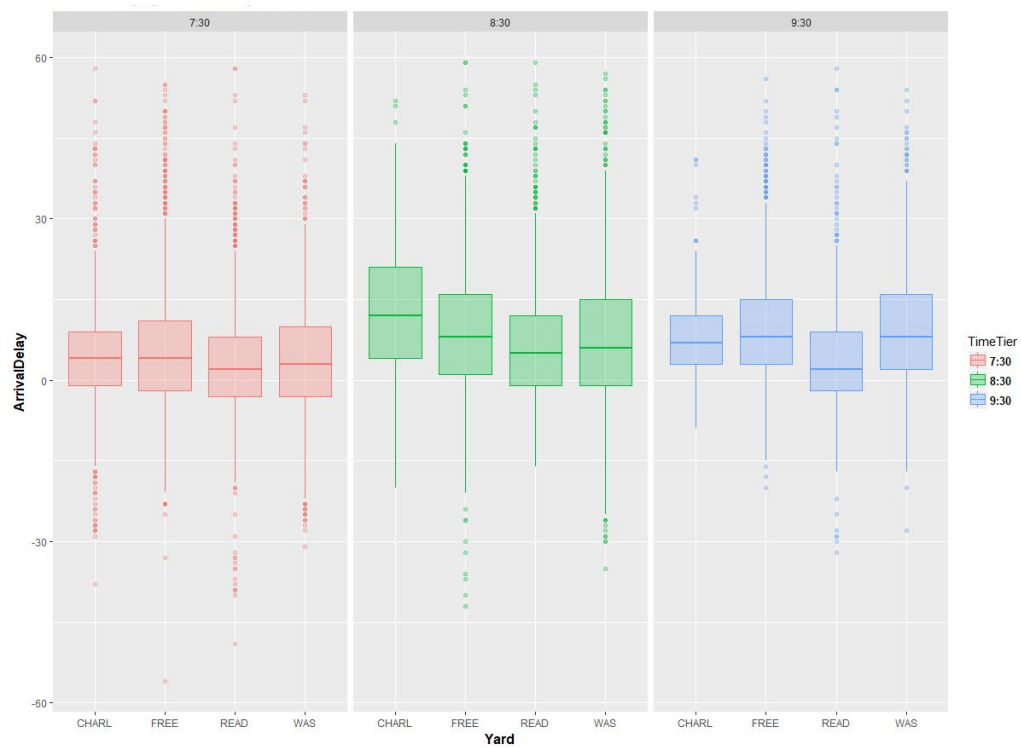


Figure 23

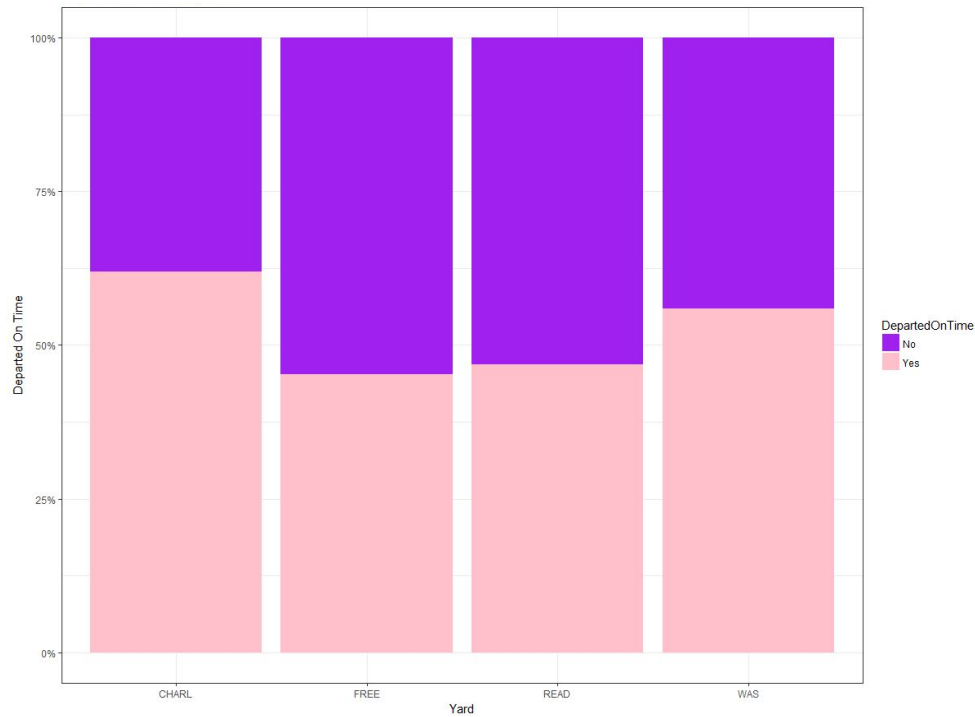


Figure 24

*Proportion of each type of bus in the dataset.*

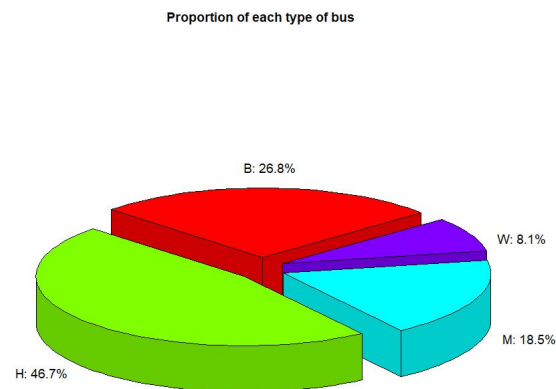


Figure 25

Since it takes time to get on the buses, we anticipated there is a positive relationship between total number of riders with the delay time. However, when we generally looked into this relationship, the range of delay time shrunk as the number of total riders increased. Although

the data showed opposite results with our expectations, we found that the routes with larger number of total riders usually have longer trip distance. We conjectured that more tolerance was considered when planning the time for a longer route.

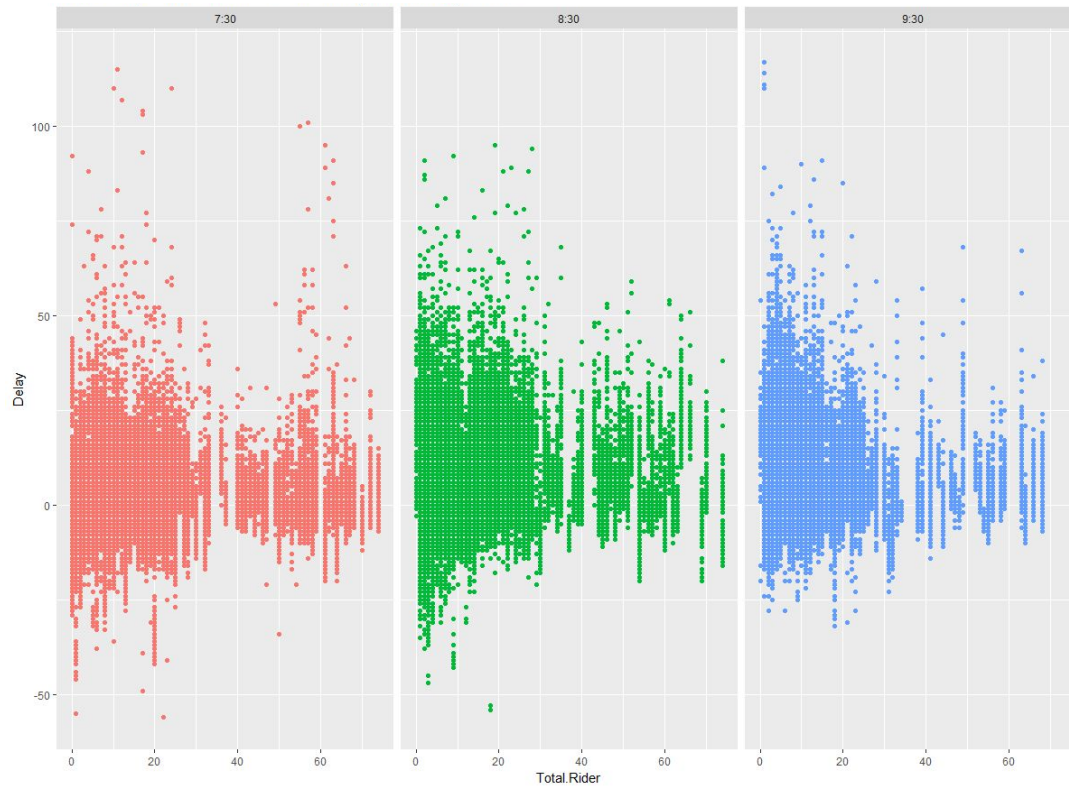


Figure 26

Figure 27. Relationship between total riders and delay time facet by Yard

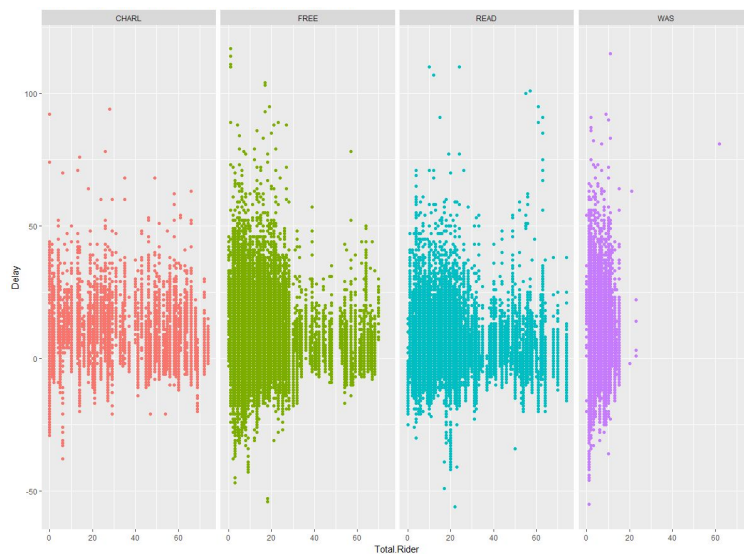


Figure 28. Delay time vs. Total Rider by Yard and Bus type

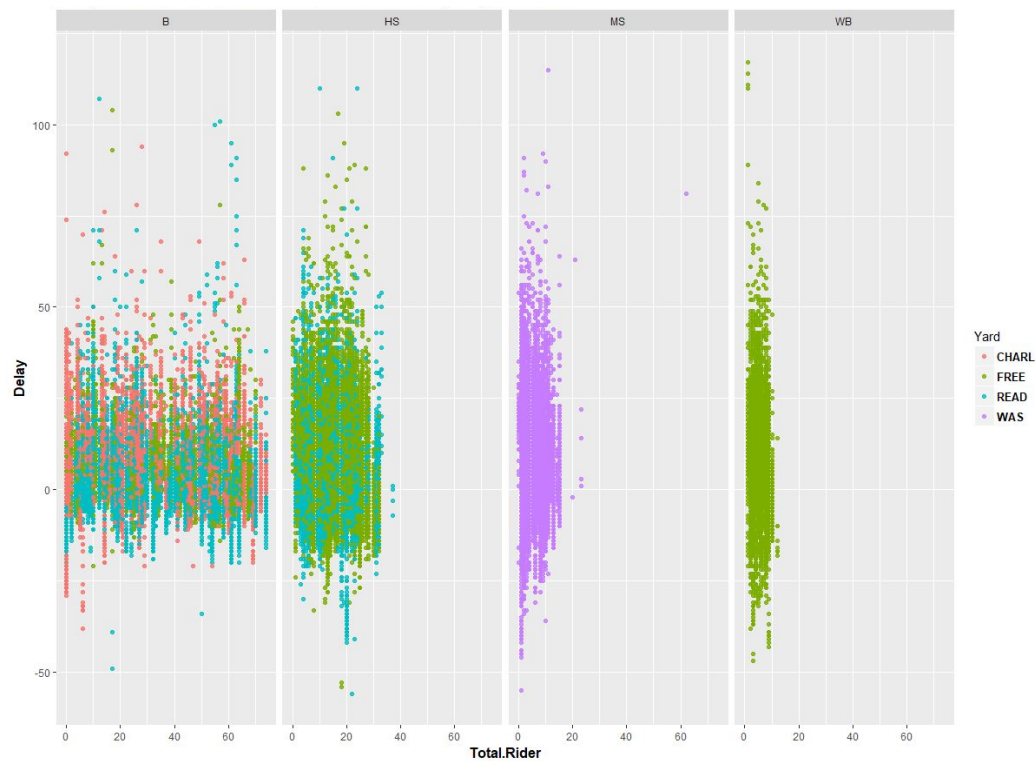
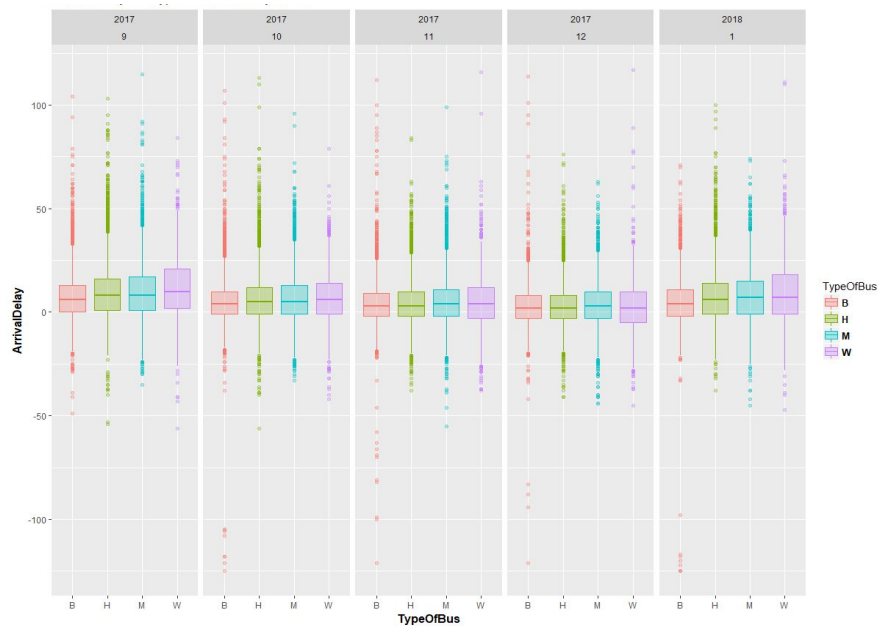


Figure 29. Delay time in months with different types of bus





Surprisingly, weather has no influences on delay time

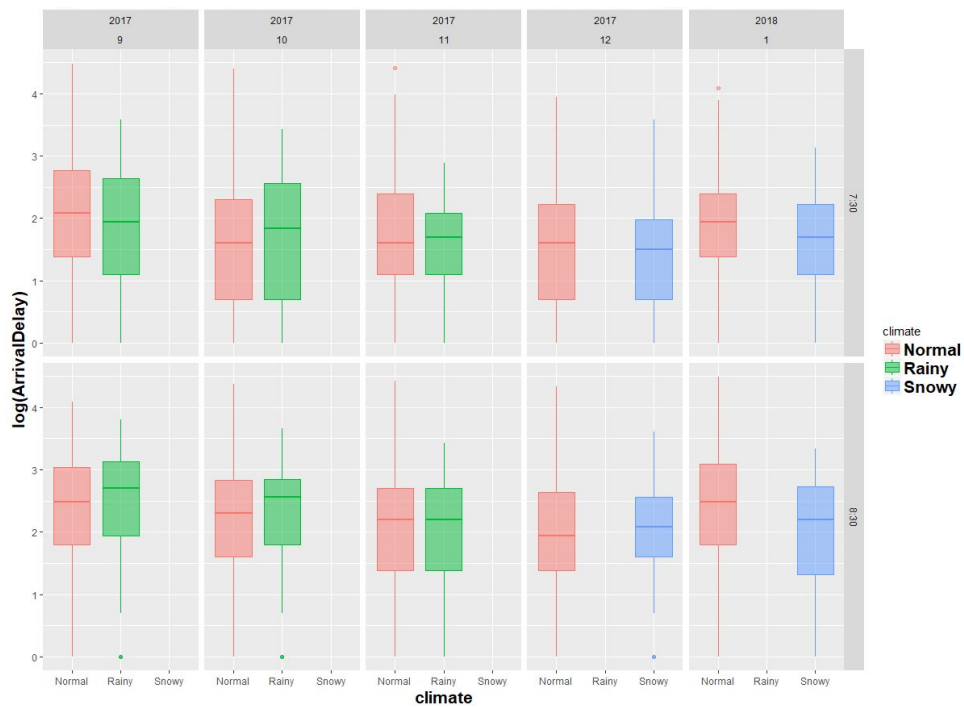


Figure 30

Neither does visibility.

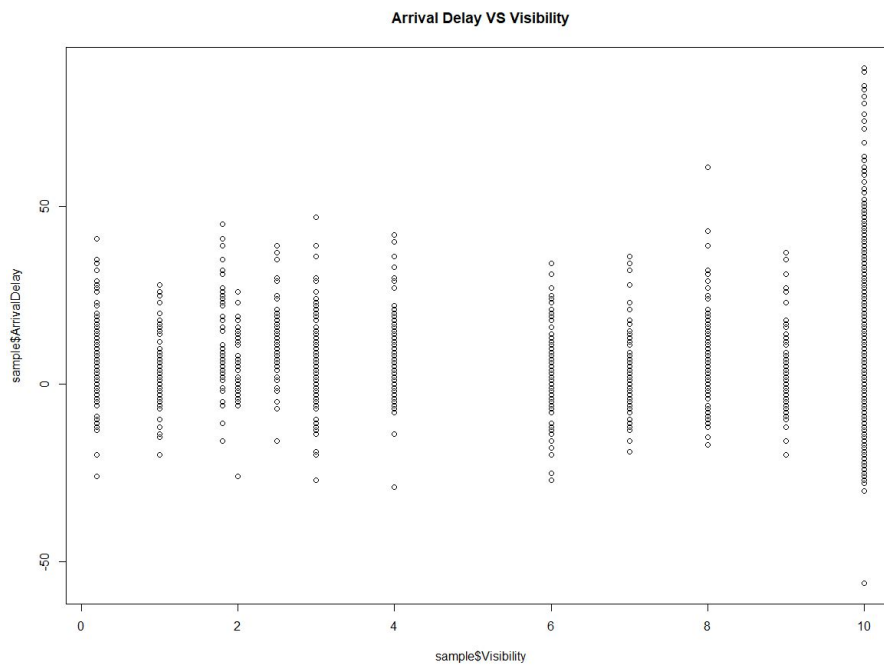


Figure 31

Figure 32. Delay Time VS Weather by Month

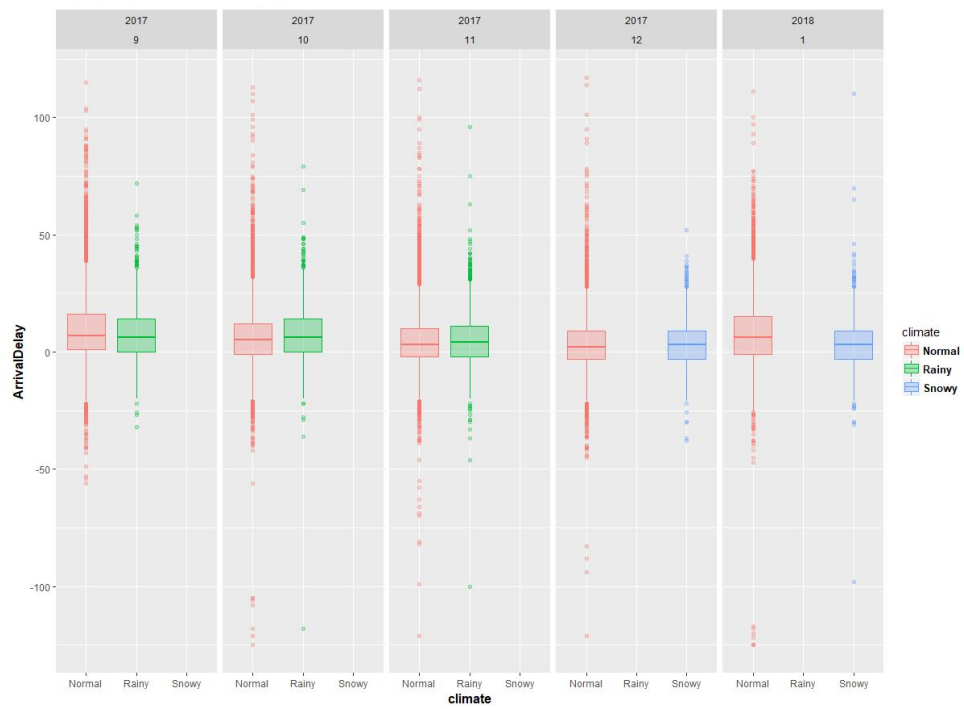


Figure 33. Delay Time VS Weather by Time Tier and Month

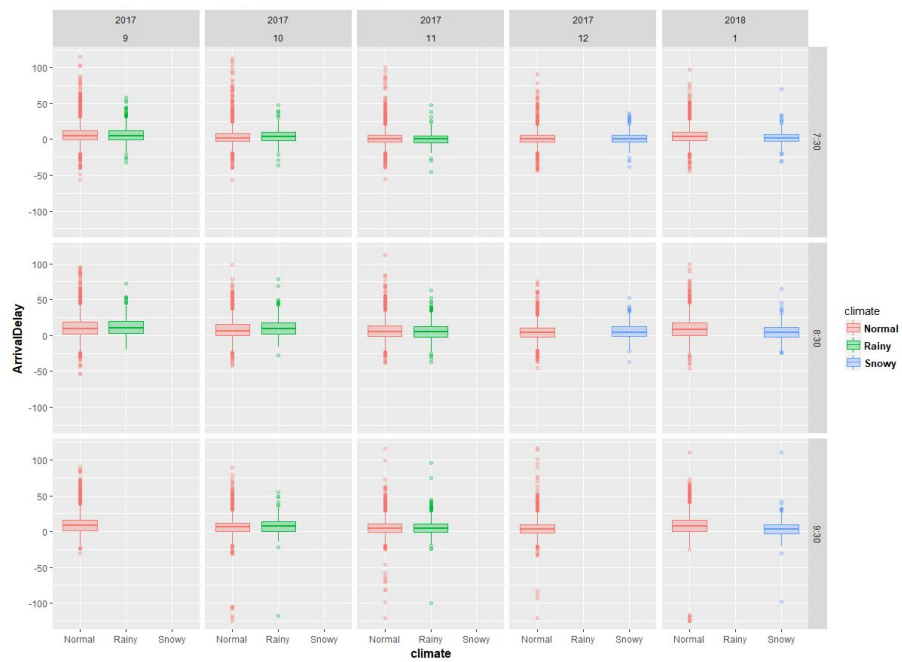


Figure 34. Delay Time VS Visibility by Month

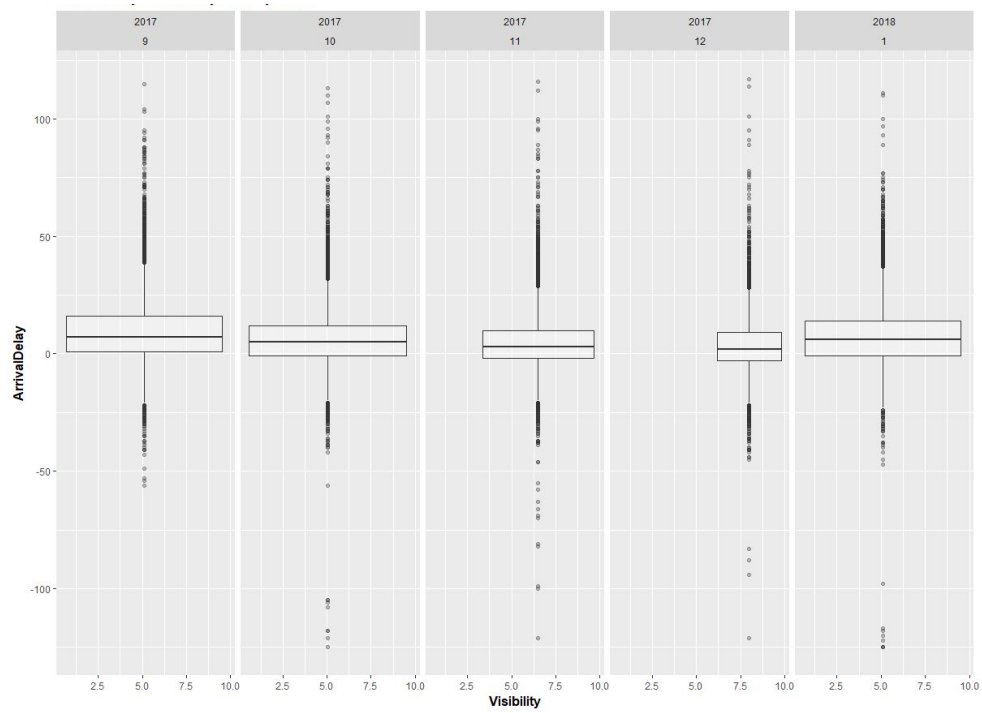


Figure 35 .Delay Time VS Visibility by Time Tier

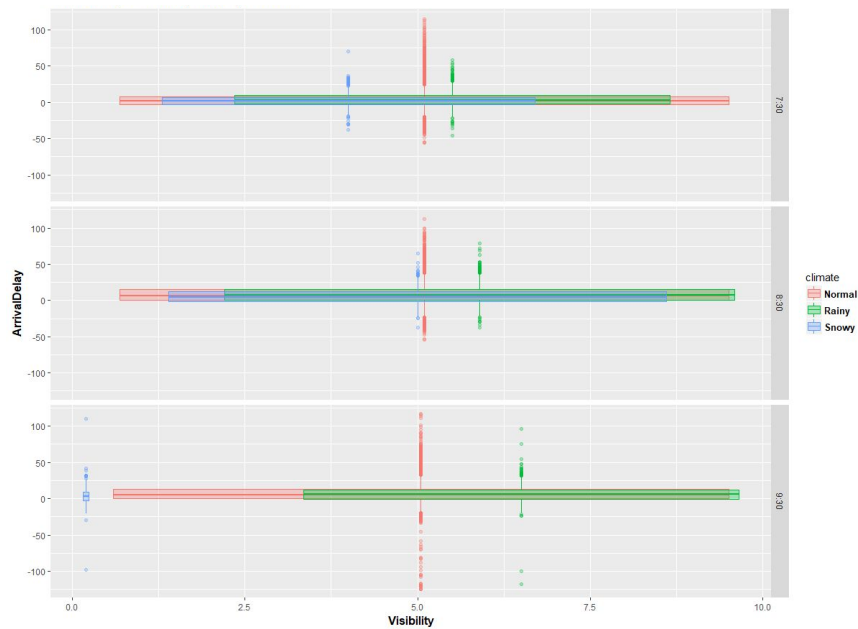


Figure 36. Delay Time VS Visibility by Time Tier and Month

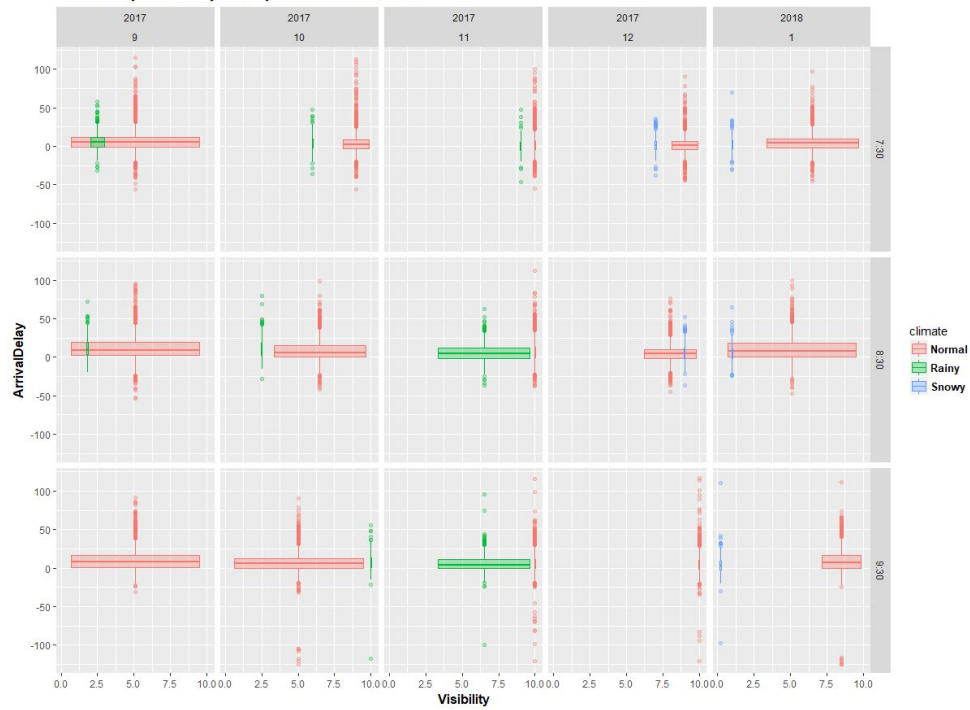


Figure 37. Delay Time VS Visibility by Type of Bus and Month

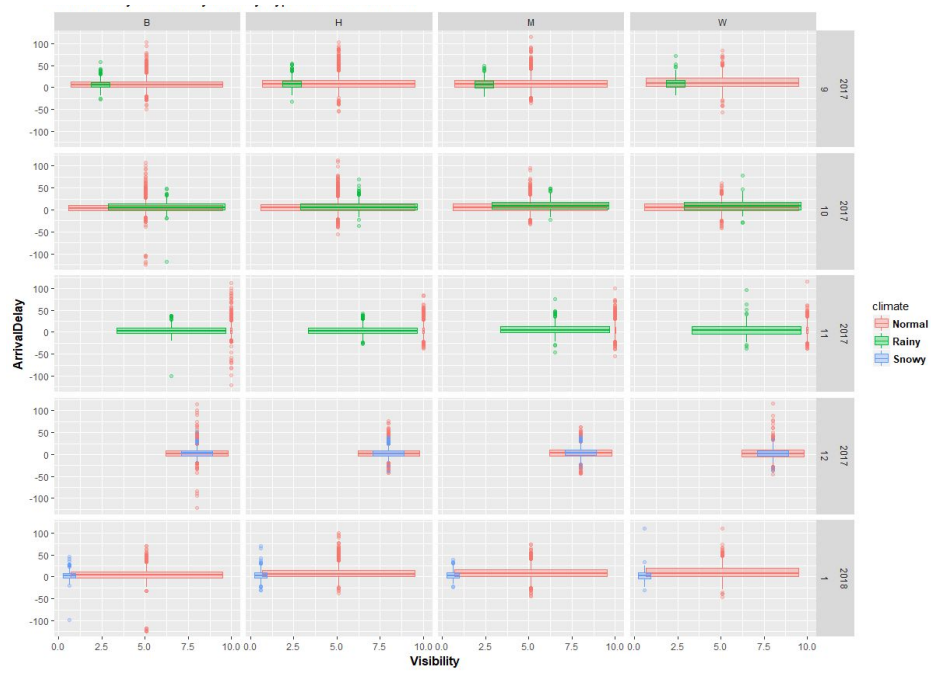
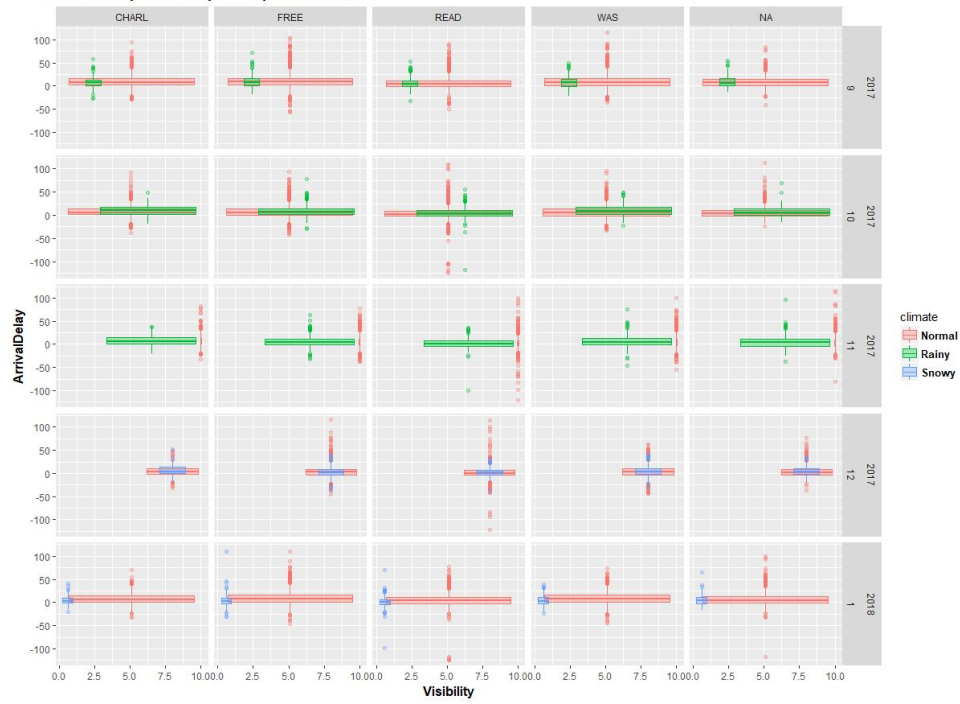


Figure 38. Delay Time VS Visibility by Yard and Month



*Figure 39. The Delay of A Rainy Day And the Day A Week Earlier And the Day A Week Later*

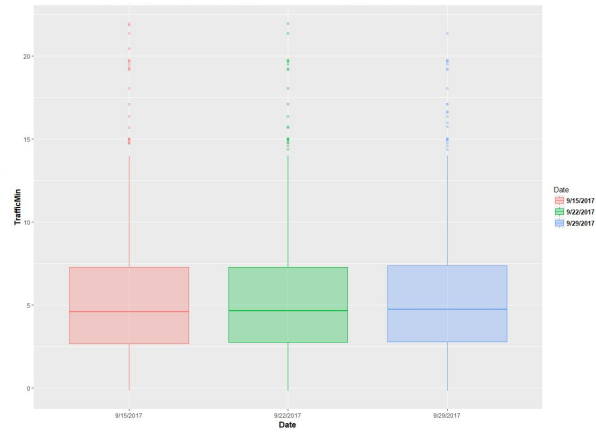
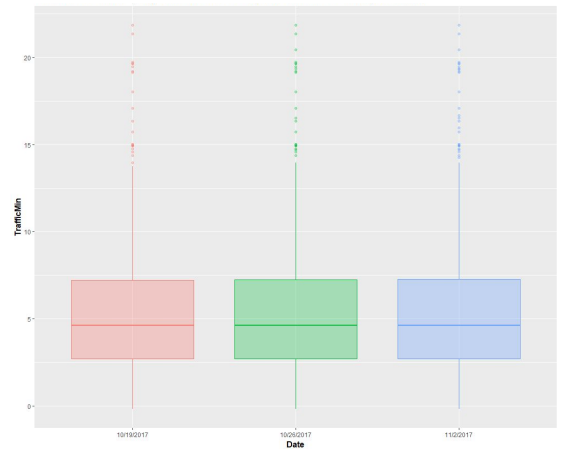
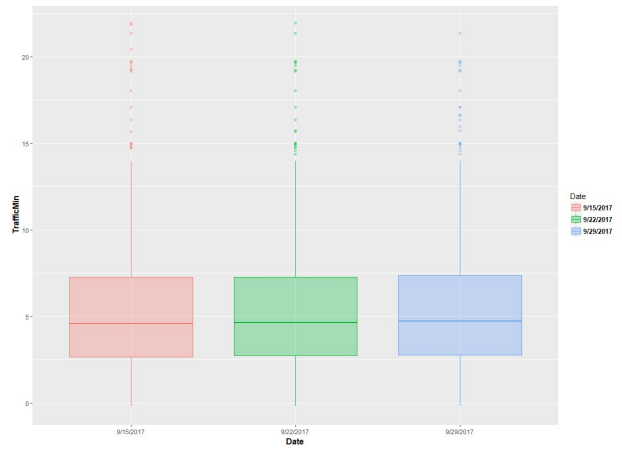
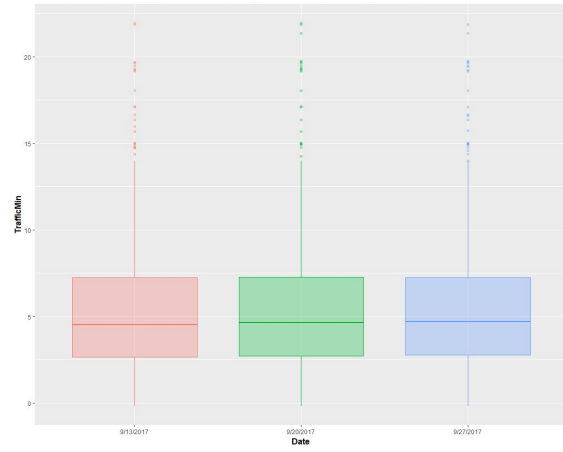
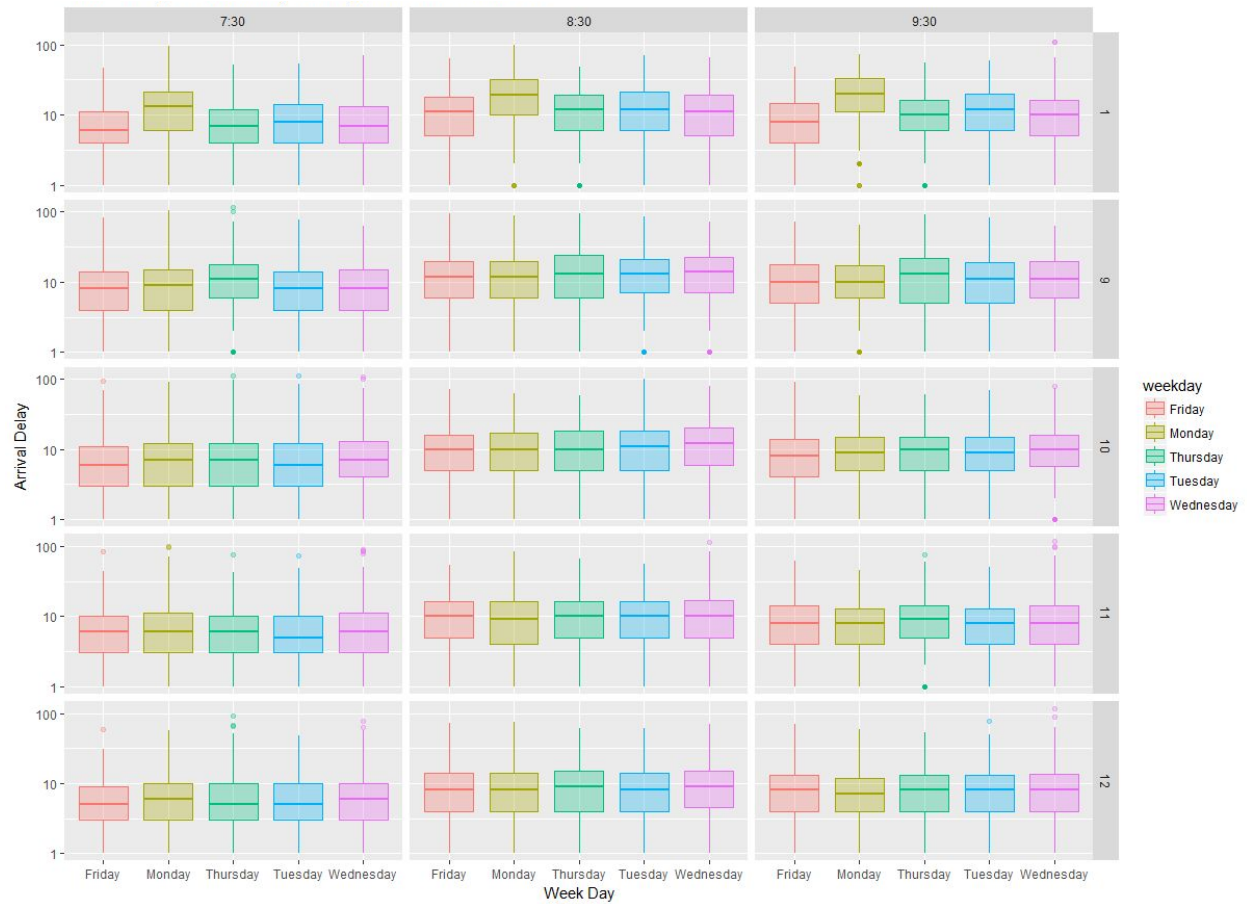
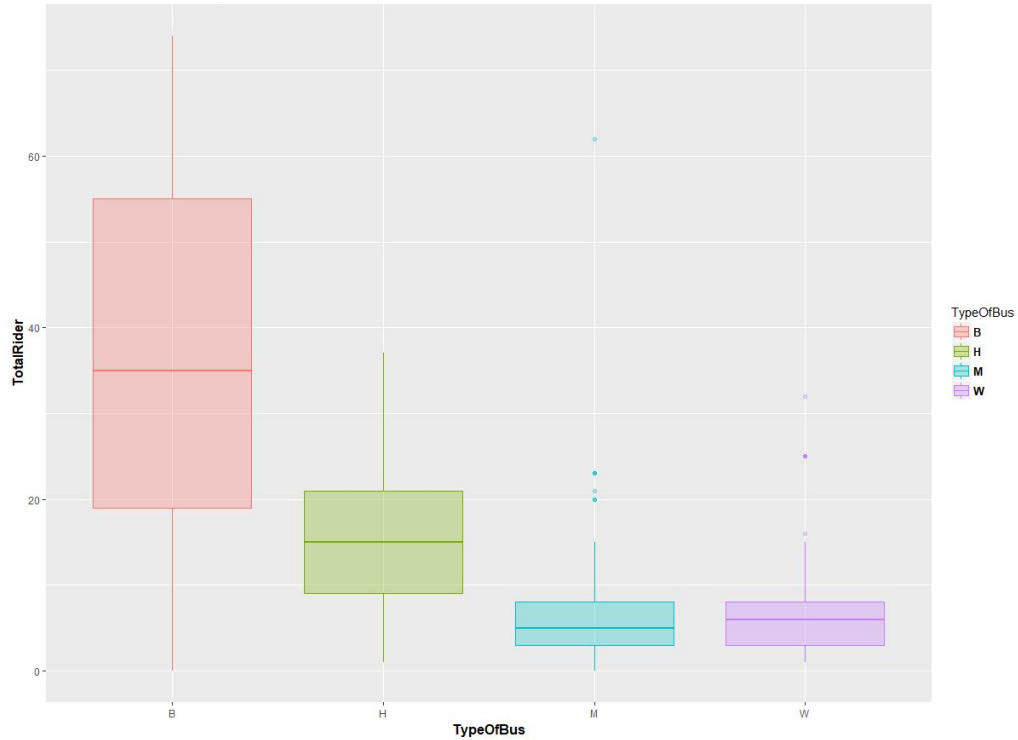


Figure 40. Delay Time VS Week Days by Time Tier and Month



Different types of bus have their own range of total rider numbers. The Big Bus have the most total rider number followed by the H bus. The minibus and the wheelchair bus have similar total rider number, which mostly under 20.



*Figure 41*

The 7:30 time tier wheelchair buses are nearly all Door-to-Door buses while the big buses nearly have no Door-to-Door stops throughout three time tiers. The variation of the Door-to-Door proportion for H bus is relatively large in the 7:30 time tier and the variation decreases on the 8:30 time tier and 9:30 time tier. On the contrary, the variation of the Door-to-Door proportion for M bus is much smaller than the variation for the H bus but expands to be the largest on the 8:30 time tier and 9:30 time tier.



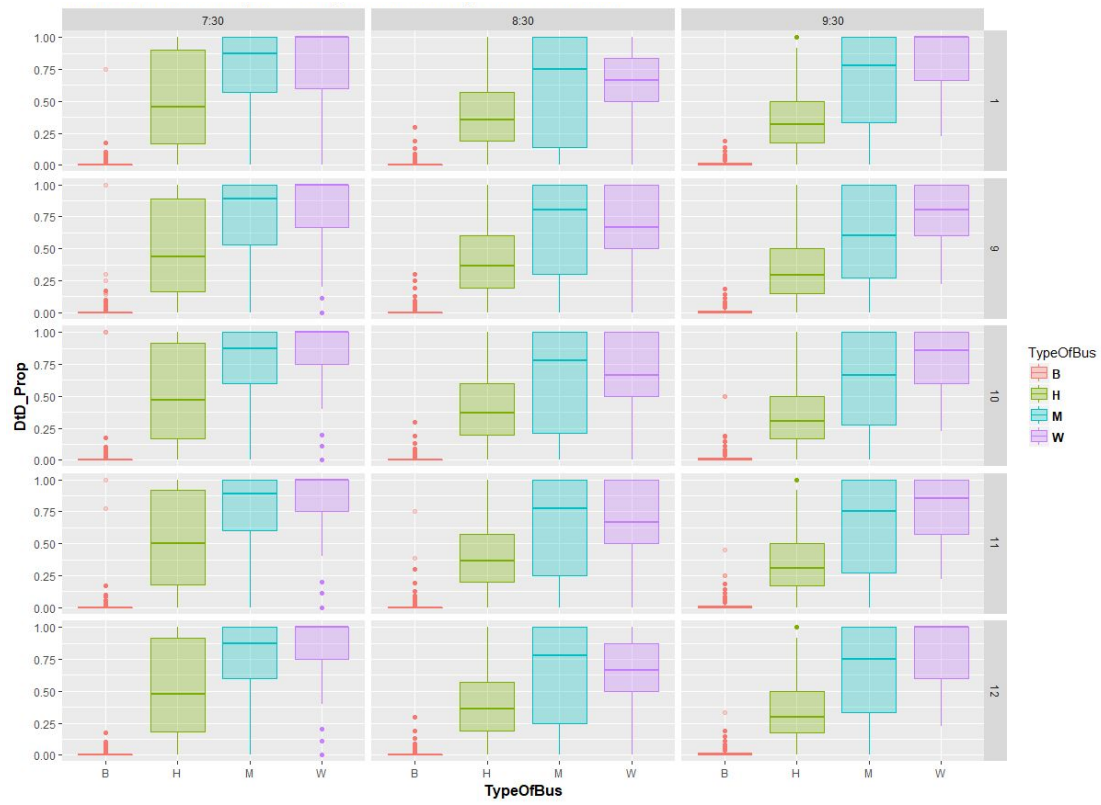


Figure 42

**Top 20 routes that are consistently performing bad:**

Route	School		Route	School
1722552	CNS Pathways		2402551	Cotting
3432551	Germaine Lawrence		3852552	Kennedy Day
4582453	Lighthouse		4722251	Longview Farm
5012251	May Ctr Randolph		10201103	Boston Latin Academy
10401551	Brighton High		10501552	Charlestown High
11402359	Henderson 2-12		11712552	Lyon 9-12
12101352	Madison Park High		12501106	West Roxbury Academy
12501112	Urban Science Academy		12501115	Urban Science Academy
12531553	West Roxbury Academy		12913561	McKinley Elementary
12941553	McKinley SEA		14601452	TechBoston Academy

The full list is included in the excel workbook “late route”

**20 Buses with shorter distance travelled but more delayed for 8:30 to 9:30 time tier:**

B308	HS215	MS239	WB909
B309	HS218	MS241	WB910
B311	HS219	MS242	WB914
B314	HS221	MS253	WB925
B315	HS220	MS255	WB935

The full list is included in the excel workbook “shorter.delayed”

## **Historical Background**

In the summer of 1974, federal District Court Judge W. Arthur Garrity ruled that the Boston Public Schools must bus students from racially segregated areas such as Roxbury and South Boston in an attempt to desegregate the school system. At the time, some schools in Roxbury were at least 90% African American while South Boston was almost completely caucasian. The ruling resulted in nearly 17,000 students being bused to different schools to increase the racial diversity of the schools. This race based busing caused a lot of turmoil in the communities, but it lasted for over a decade. In 1987, the court ordered busing was ended since the schools were deemed to be sufficiently integrated. In 1989 a controlled choice system was enacted which allowed students and parents to choose which school to attend as long as the schools were still sufficiently integrated. The families were given a list of schools from which to choose which school they wanted to attend. The final assignment of the student was based on several factors, race being among them. Finally in 1999 race was removed as a factor upon which school assignment was based.

## **The Explanation of Random Forest:**

In order to understand what a Random Forest model is, we must first understand what decision trees are. Decision trees are a way to do classification and regression. For example, let us assume we have a data set, the outcome of which is whether a person is fit. It has two levels, fit or unfit. This data set also includes information like age, eating habits and exercise habits. We call these features. The decision tree splits the dataset into smaller datasets based upon the features until the datasets are small enough so that the data points they contain are all similar. The results of a decision tree are easy to understand, but they do not have good prediction accuracy when there are too many features. But random forest could get over this disadvantage and the result would be more reliable.

## **The explanation of Gradient Boosting:**

The final product of a gradient boosting would be a single strong prediction model resulting from some weak prediction models. In most cases, when people are studying about the models, what they are referring to are decision trees.

Gradient boosting generalizes the strong model through optimizing a loss function. A loss function is a way to assess the performance of the models predictions by quantifying the error in the prediction. After a loss function is chosen and defined, the objective of the model is to minimize it. Unlike other techniques, the trees that gradient boosting uses are sequential, which means that each tree is grown using information from previous trees. The fact that those trees are sequential solves the problem of overfitting since gradient boosting trees do not extend from the original data. Also, residuals are being examined of their patterns and a decision tree is being fitted to the residuals from the model. Once no such patterns can be found in the residuals, the model might be successfully optimized. In the process of gradient boosting, the size of each of the trees may be small, with only a few terminal nodes.

## **References:**

1. Goldberg, C. (2018). Busing's Day Ends: Boston Drops Race In Pupil Placement. [online] Nytimes.com. Available at: <https://www.nytimes.com/1999/07/15/us/busing-s-day-ends-boston-drops-race-in-pupil-placement.html>.
2. Google Distance Matrix API
3. Bostonpublicschools.org. (2018). Data and Reports / 2017 School Performance Data. [online] Available at: <https://www.bostonpublicschools.org/Page/5343>.
4. Weather Underground API