(), , 1–15

### ARTICLE

To Queue or not to Queue: How Queueing Analysis can help Improve Airport Passenger Wait Times

Halilou Junior Tmite, Tangbufan Wei(victor), Gaurav Daven

**Abstract**

Queuing analysis focuses on the mathematical study of queuing (waiting in line) for service. It aims to analyze every component of queuing, including the arrival rates, service rates, number of severs, and the number of customers (which can be people, data packets, cars, etc.). It is widely associated with operations research or network performance since the results are used to make critical business decisions about the allocation of resources required to provide a service depending on wide variety of variables (time of day, season of year, etc.). This can help organizations build efficient and cost-effective workflow systems through addressing staffing, scheduling and customer service issues.

In this report, we use queuing analysis to predict airport passenger wait times to ensure that most passengers are serviced in a timely fashion. Our approach is to first predict the required number of servers needed to achieve the desired quality of service (QoS) level. We then use simulation with different arrival rates and number of servers to predict the proportion of passengers who would need to wait more than a certain amount of time.

**DATA SET**

20262030

**DESCRIPTION**

Passengers and flights data in four airfields from 2026 to 2030.

BASA˙AUC˙2028˙912

Passengers and flights data in airfield AUC

from Aug to Dec in 2028.

dat˙F˙sub

Flights data from Sep to Dec in 2028 in

airfield AUC.

dat˙P˙sub˙c

Passengers and flights data in airfield AUC

from Sep to Dec in 2028.

Keywords: Arrival Rate, Quality of Service (QoS) level, Service Rate, M/M/1 Queueing Model, Poisson Distribution

# INTRODUCTION

Queueing analysis can help us in evaluating performance of a queueing model (given the arrival rate of passengers and number of servers) through wait times incurred by the passengers and the service rate of the servers. These metrics can help us predict future passenger wait times and queue lengths.

This allows us to dedicate the optimal number of servers to keep the queue moving and reduce the waiting time of passengers. There are obviously budgetary problems and union considerations which may restrict the number of servers available. Another issue is that passenger arrival rates are highly variable on time (not a lot of people travel between midnight and 4:00 a.m.), seasonality (summer vacation and winter holidays travel), day of the week (Tuesdays and Wednesdays are generally less busy since people need to take more days off work to travel during those days), and other variables. This report focuses on passenger wait times for air travel, specifically for pre-board screening.

Passengers arrive at the beginning of the main queue (S1) where their boarding pass may be screened. After reaching the end of main queue, their boarding pass is screened at S2 and they proceed to one of the active servers (not busy) for further processing (Figure 1).

**Section 1: Exploratory Analysis** of the report focuses on exploratory data analysis and justification behind the selection of the dataset which is used in this project. The four datasets are compared based on the number of missing values and in which columns, the number of airfields and corresponding observations present, and the presence of any additional variables which could help in the prediction of passenger wait times. A data dictionary for the selected dataset is provided along with accompanying visualizations to help the reader better understand the dataset.

**Section 2: Queuing Analysis of Passenger Wait Times** is about deriving various variables such as average arrival rates, average number of servers, average wait time, etc. for 4-hour time clusters using the M/M/1 model assumptions. A regression model used to link the average arrival rate, average service rate, and the average number of servers required to ultimately predict the passenger the quality of service (QoS) curve (what proportion of passeneger need to wait longer than x amount of time (in minutes)).

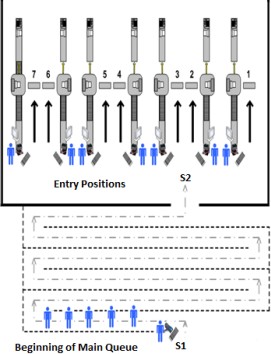


Figure 1. Pre-board screening (PBS) process for airport passengers

# Exploratory Analysis

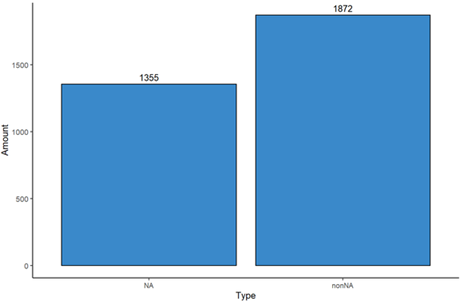
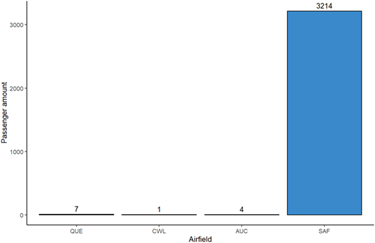
**1**

This section goes into detail about exploring the four different datasets to ultimately select the dataset used for queuing analysis. First, we compare the datasets based on if they have passenger related information (specifically wait time). Then, we compare the amount of observations available for each airfield and the amount of missing values in each dataset. Finally, we choose the dataset for the queuing analysis and provide accompanying visualizations and a data dictionary to help the reader understand it.

## Selection of Data Set

**1.1**

We first look at data set years 20262030. It is obvious that in the Figure 2(A), we only have enough data for airfield SAF (3214 lines) to do the analysis. However, in the Figure 2(B) we find out that there are 41.9 percent of the data in SAF are NA. It’s too many and we decide to leave this one. For the data set data˙f˙sub, among its 26 variables, we cannot find the variable wait time, then we leave this one as well.



(a) (b)

Figure 2. Data set years20262030 (A) Number of observations available for each airfield (B) NA observations vs Non-NA observations for the variable wait time

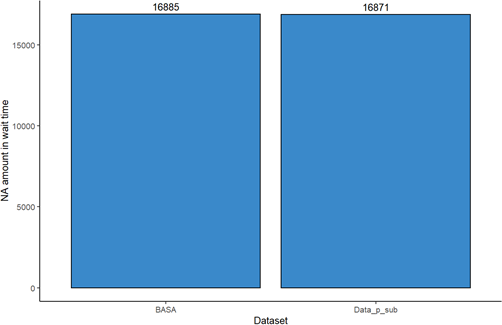


Figure 3. BASA vs Data˙p˙sub in NA amount of wait time

Finally, we compare the amount of NA in variable Wait˙time in BASA and data˙p˙sub in Figure 3, the amount in BASA is slightly more than data˙p˙sub. The c˙start and c˙avg are also blank for those lines whose Wait˙time are blank in BASA. In data˙p˙sub, we have more variables that can help our prediction, also less NA values in variables. So we finally choose it.

## Data Dictionary

**1.2**

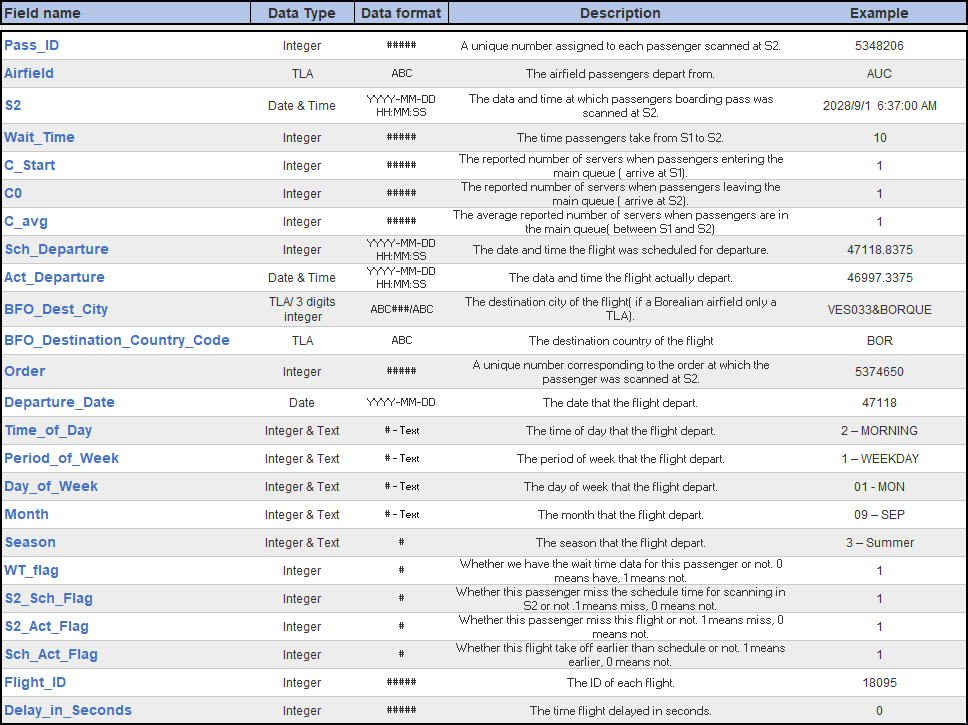


Figure 4. Data dictionary for the dat˙P˙sub˙c dataset. Missing value codes depend on the variable – they include NA.

## Data Visualization

**1.3**

The flights are punctual or not usually take a lot of concern when people buying air tickets. We are also no exception of it. The punctual here means the schedule depart time of flights are exactly the same to their actual depart time. We first turn our sight to day period. It is not hard to find in Figure (5) that there are no punctual flights in mornings and nights in all data sets. Since most of

the ground staff work at day time, the productivity for the night and morning staff definitely is too low. There are also many other factors such like weather and visibility can influence the flights as well. The pilots will take longer time in morning and night to recognize runway of flights than day time. That’s the reason why most part of the punctual flights land in the afternoon rather than evening. Better landing conditions , ground staff and more flights causing it leads evening with 29832.

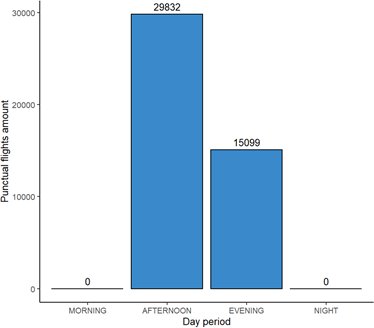
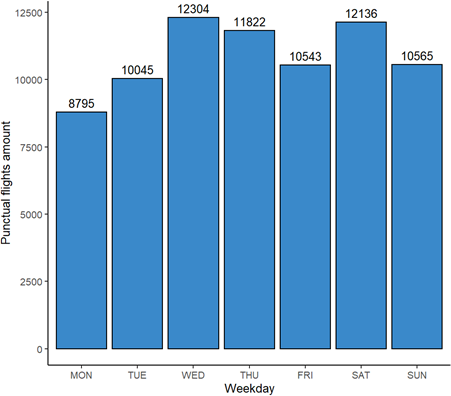
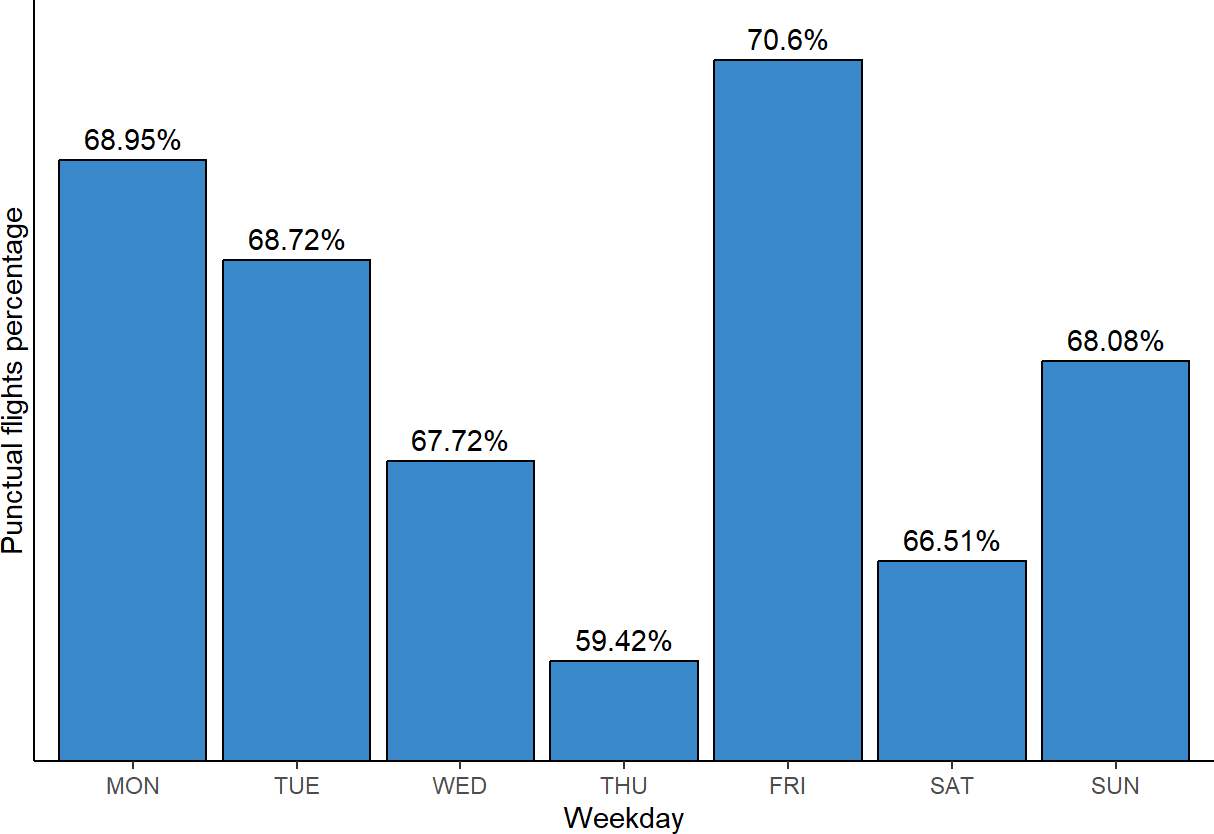


Figure 5. Punctual flight in every day period

We can see in Figure 6(A) that among all the week days, Monday and Tuesday are two out of seven days have the least punctual flights with 8795 and 10045 separately. But their punctual rates are high, with 68.95 percent and 68.72 percent separately. In our consideration, that because the total amount of flights are not too much. Staff have more time to deal with each flight. Comparing to the first two days of week, it is obvious that the punctual flights amount reach the peak in the mid-week days, with 12304 on Wednesday and 11822 on Thursday. However, in Figure 6(B) we can find that the punctual flights rates are lowest .Similarly, that is because the total amount of flights increase, staff need to deal with more flights at the same time.

The boxplot of wait times (in minutes) (Figure 7 ) shows that mean amount of wait times (on weekdays and weekends) for the vast majority of passengers is approximately 5 minutes. The maximum amount of wait time for passengers on weekdays was 69 minutes whereas the value for weekends was 75 minutes. Conversely, the minimum amount of wait time for passengers on both weekdays was weekends was 1 minute. This figure shows that the wait times for most passengers are non-significant (less than 10 minutes) with only some outliers where the wait times were really long.

(a) (b)

Figure 6. Punctual flight in every weekday (A) Amount (B) Percentage

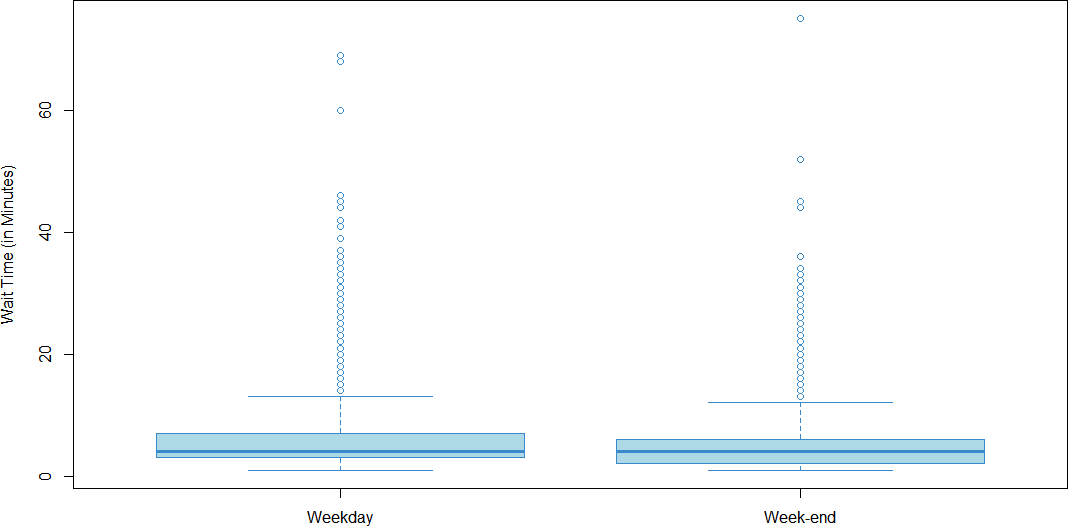


Figure 7. Boxplot of passenger wait times for weekdays and weekends

# Queueing Analysis of Passenger Wait Time

**2**

For the selected dataset, we first removed the missing values for passenger wait times. Then we filtered the dataset based on season and only selected observations for the Autumn season.

## Objectives

**2.1**

We use available data to provide estimates of passenger arrival rates *λ*, processing rates *µ*, number of servers c for each combination of checkpoint, time period, day of the week, season (cluster),

For each cluster, given *λ*, *µ*, and c, calculate the quality of service (QoS) level curve (p(x), x) (i.e. percentage p of passengers which wait less than x minutes).

For each cluster, predict the average number of servers c required to achieve a prescribed QoS level (p(x), x) given an arrival profile *λ*.

## Deriving Variable for Prediction

**2.2**

* + 1. *Arrival Rate*

The average arrival rate *λ* is the passengers arrival rate. We first get the number of observations in certain time cluster, then divide it by the total time units in this cluster.

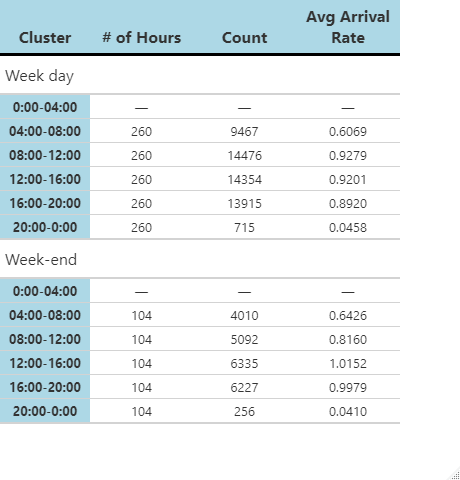


Figure 8. Average arrival rate by 4 hours time slot cluster

* + 1. *Average Number of Server: c*

The average number of server c is the sum of all servers that are available at S2 (C0) divided by the number of observations in each time cluster. The distribution here means the probability to have certain number of server when arriving at S2 in each time cluster.

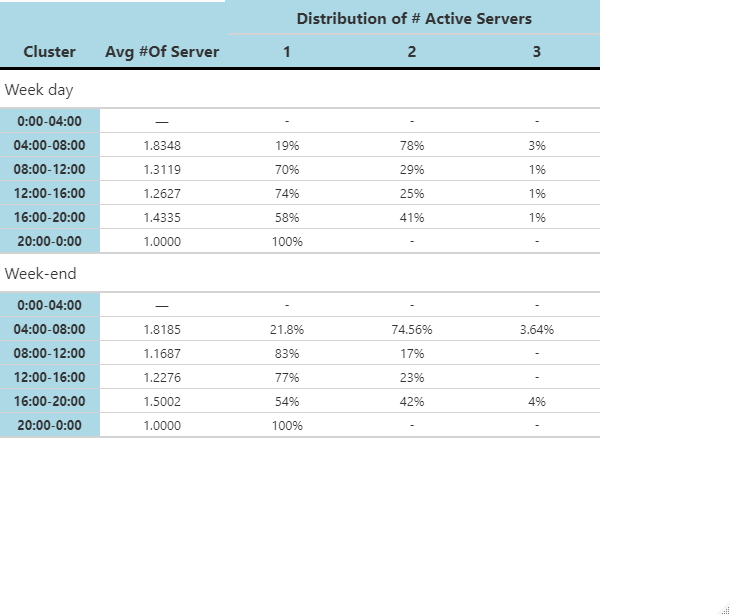


Figure 9. Average number of servers by 4 hours time slot cluster

* + 1. *Average Wait Time and Performance Level*

The average wait time Wq is sum of wait time divided by the number of observations in each time cluster. It’s performance level is a cumulative probability that the wait time for passengers is less than certain level.

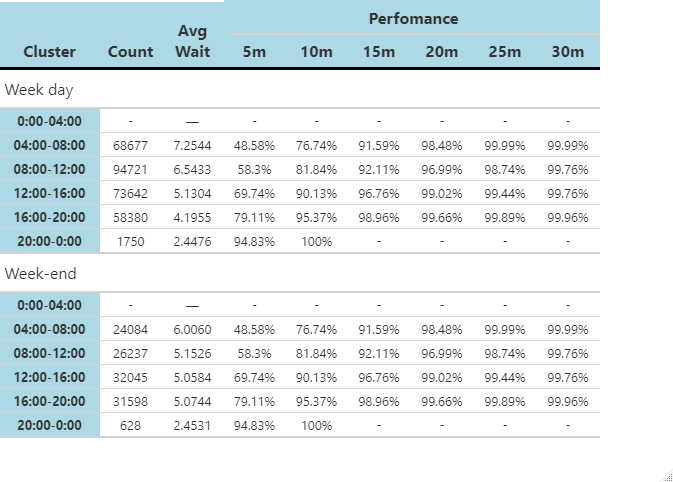
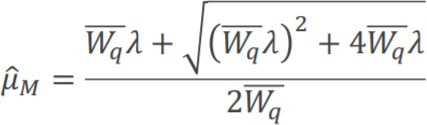


Figure 10. Average wait time by 4 hours time slot cluster

* + 1. *Estimate Server Rate*

The arrival rate *λ* is already known, and the wait time Wq can also be computed, so the estimate server rate *µ* can be recovered with this function in each cluster.



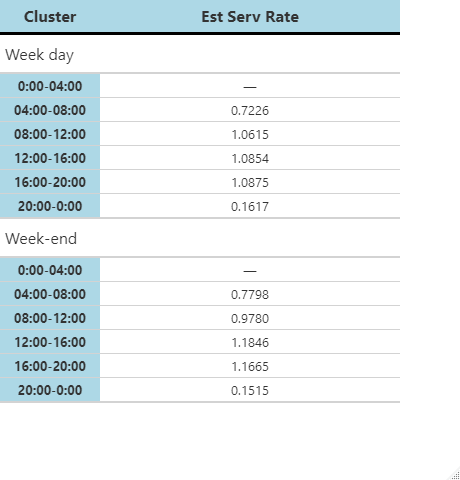


Figure 11. Estimate server rate by 4 hours time slot cluster

## Regression Analysis

**2.3**

Figure(12) is the illustration of variables that we are going to use in the regression model.

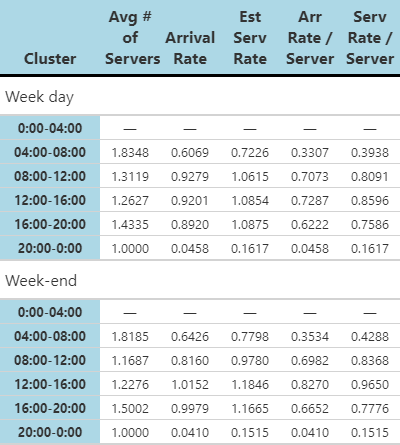


Figure 12. Regression variables

* + 1. *Predicted Arrival Rate per Line vs. Server Rate per Line: a and b*

We know the *µ* = *a · c* + *b · λ*. They we build a regression model which shown in Figure(13). So the a\*c for this model is 0.09116 and b is 1.04208. For the c here, we use the average of c0 of the whole dataset after filtering. So for this model, we get a for 0.06451, b for 1.04208.

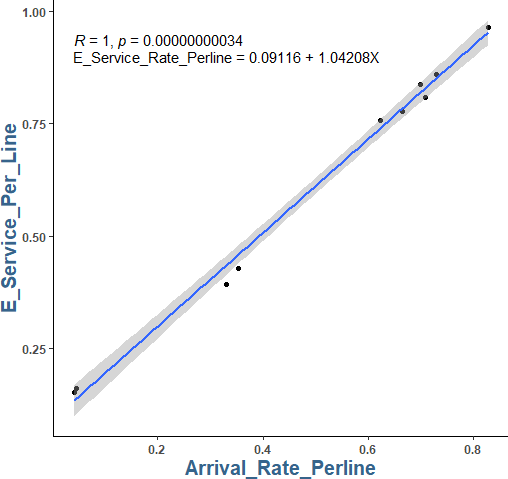


Figure 13. Linear regression

* + 1. *Predicted Mean Number of Servers: cD*

We use the a,b, *λ* we got above to find the cR.The W0 is the Lambert function. Then we have some scenarios with the x that is time unit. As we know that the cD is the cR times d. Then we compute it then get the Figure(14)



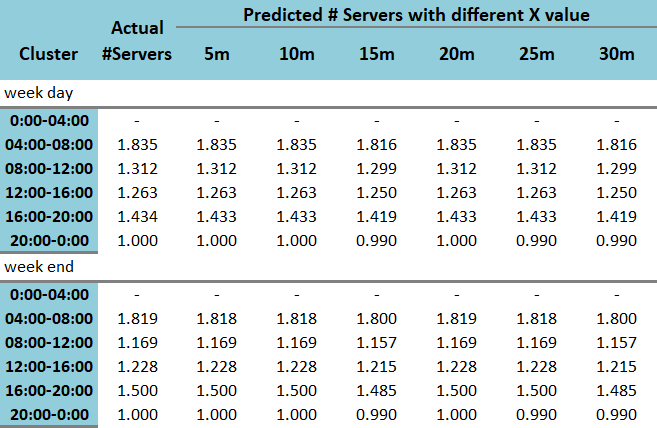


Figure 14. Predicted Number of Server

* + 1. *Predicted Number of Servers vs. Actual Number of Servers: cD*

Then we get a new linear regression model between c and cr to get the d. We use x = 15 as an example, the d in Figure(15) is the slope which is 0.0044.

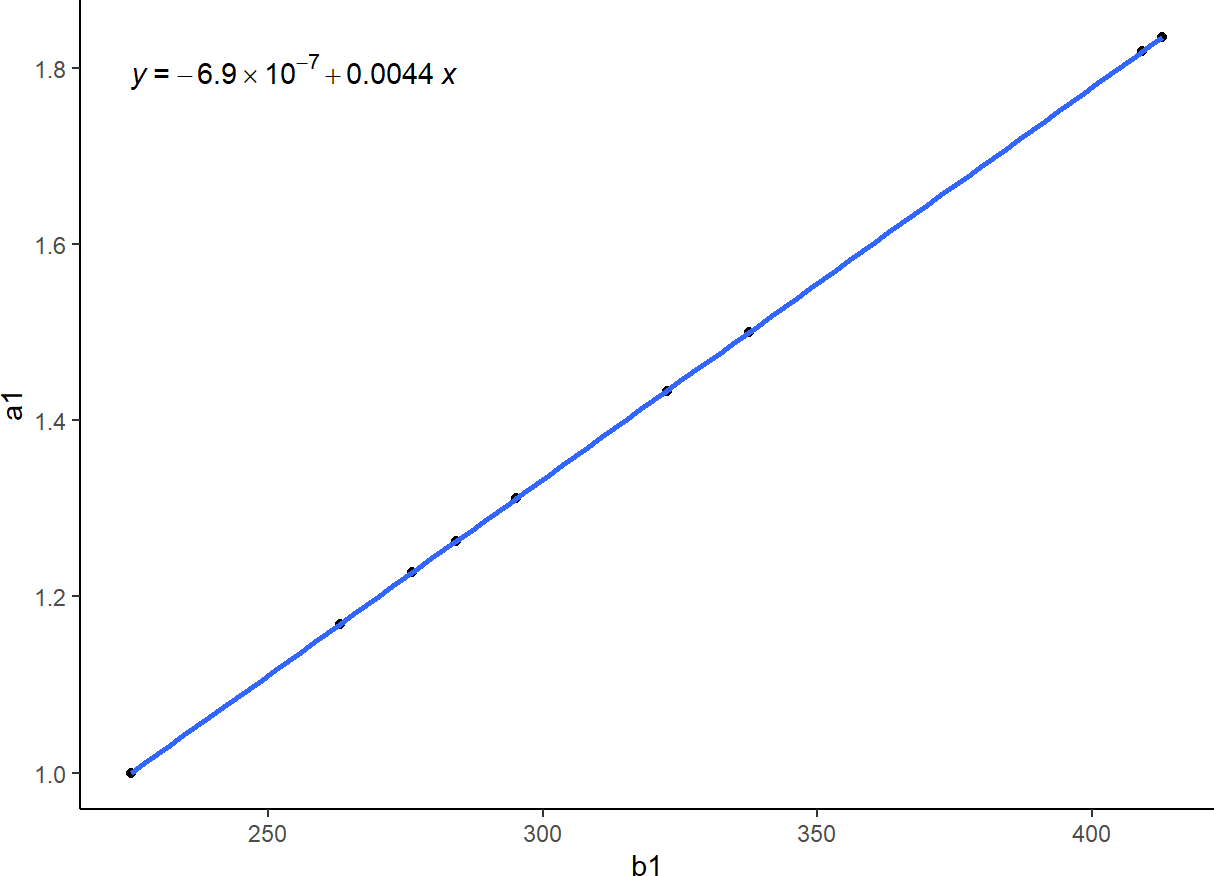


Figure 15. Linear regression to predict d

For the performance of those predicted number of servers, we put cD( our predicted values of servers) back the p equation below and get the cumulative proportion for each time cluster.



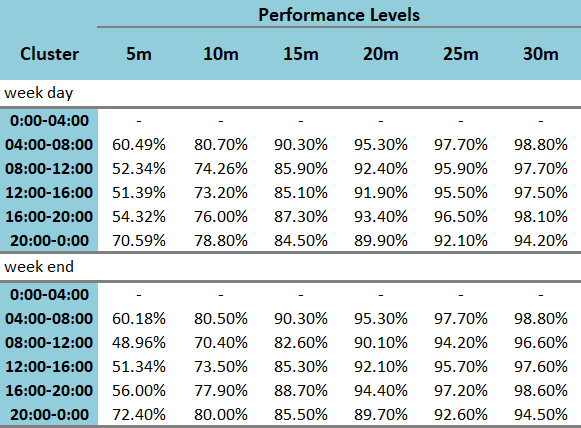


Figure 16. Performance level of predicted number of server

# CONCLUSION

In conclusion, queuing analysis can help organizations such as an airport decide on how to allocate resources efficiently to ensure budgetary requirements are met and the quality of service levels for passengers are maintained (among other important things such as union considerations, etc.). In this report, we used different variables such as arrival rates, service rates, average wait times, etc. to calculate: the quality of service (QoS) level and the average number of servers required to reach a prescribed QoS level, given an arrival profile. In the future, our recommendation to airports would be to implement automated screening lanes which can help enhance security efficiency while decreasing passenger wait times. This can also help reduce budgetary constraints for the airport.

# WORK COMPLETED BY EACH MEMBER

Halilou, Victor, and Gaurav worked together on the M/M/1 Regression model. Victor completed the data dictionary and some of the exploratory visualizations. Gaurav completed the Abstract, Introduction, Conclusion and some of the visualizations. Halilou helped create the overall format of the report and verify our results for different variables.

# REFERENCES

### List of References

Queuing Theory Definition, Elements, and Example. (2022, June 4). Investopedia. h[ttps://www.in](http://www.investopedia.com/terms/q/queuing-theory.asp)vestop[edia.com/te](http://www.investopedia.com/terms/q/queuing-theory.asp)r[ms/q/queuing-theory.asp](http://www.investopedia.com/terms/q/queuing-theory.asp)

Wikipedia contributors. (2022, November 19). Queueing theory. Wikipedia. https://en.wikipedia.org/wiki/Queueing˙theory