Jiayu Tan
Ethan Stubberfield
Tamara Micic
Elizabeth Shevchenko

When Should You Arrive at the Airport to Get to Your Flight on Time?

The following report describes the dataset, interesting insights, assumptions, the created model and other discussion relating to the pre-board screening of passengers and crew arriving at the hypothetical Borealian Airports. We come to you from *Foogle Analytics* to showcase our method for recommending passenger arrival times at Borealian Aeronautic Security Agency airports in order for them to pass through the pre-boarding screening process and make it to their flight in a timely manner. The data is provided by the Borealian Aeronautic Security Agency which is looking to predict wait times at their 4 Major Airfields: Auckland, Chebucto, Saint-François and Queenston.

The pre boarding screening (PBS) process begins with a single line from which passengers proceed to one of up to 3 servers. The number of servers varies depending on the airport busyness as well as an internal vocational policy, perhaps decided ad hoc by the airport terminal management. At the beginning of this line is a station *S1* at which boarding passes may or may not be scanned. Passengers then enter the main queue until they arrive at the end of the queue – station *S2* – where every single boarding pass is scanned and they are directed to a server entry position. Then passengers and carry-on luggage are screened by a server and they can proceed to their gate.

We were provided with 4 datasets to do our analysis: *years20262030*, *BASA_AUC_2028_912*, *dat_F_sub*, and *dat_P_sub_c*. We quickly eliminated *dat_F_sub*, as it contained only flight information and no information on passenger wait times at the airport. We then eliminated *years20262030* because it did not have nearly as many entries as the other two, and we thought that more data would provide more reliable conclusions.

Between the two datasets left, we decided on *dat_P_sub_c*. Below we show some analysis we did to compare the two. In Figure 1a below we see the passenger distribution for *BASA_AUC_2028_912* and in Figure 1b we see the passenger distribution for *dat_P_sub_c*. The passenger influx distribution is nearly identical between the two.

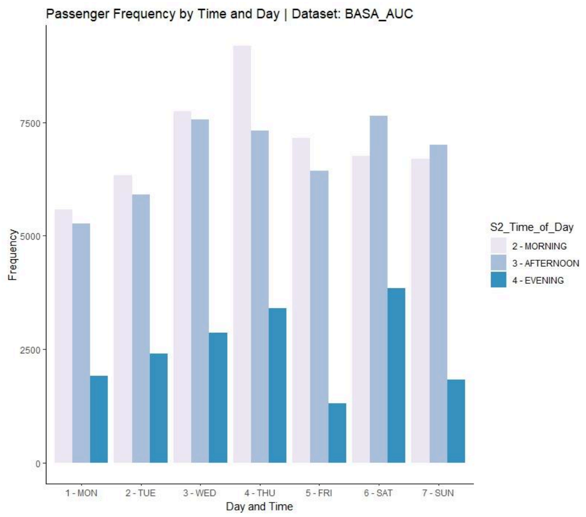


Figure 1a.

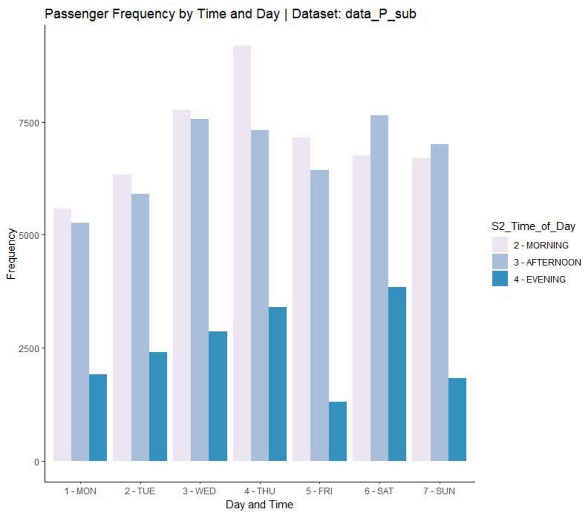


Figure 1b.

The charts below offer a more detailed visual of the differences between these datasets. We can see that there exist only fractions in differences between the two.

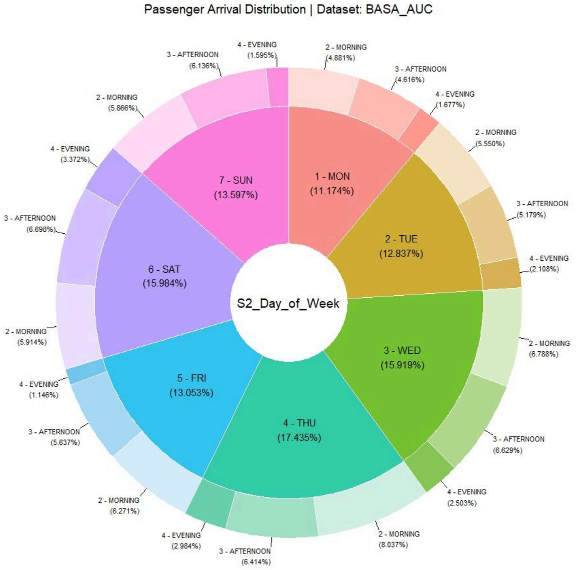


Figure 2a.

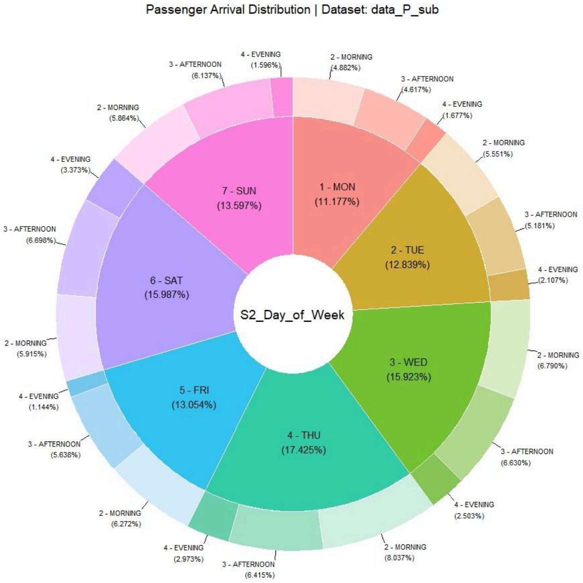


Figure 2b.

Besides highlighting the apparent lack of differences between the datasets, the charts above also do a satisfactory job in describing the influx of passengers throughout the week. Keying in on the charts on the right side which represent our chosen data, we can see among the days of the week, the difference in incoming passengers is not very large. Mondays achieve the minimum number of passengers, with Thursdays reaching a maximum, with only a 6.2% difference between these extremes. The real differences lie in the times at which passengers are scanned. Passengers are scanned during mornings (6:00-11:59) and afternoons (12:00-17:59) much more often than in the evenings (18:00-23:59) within our data.

We thought that the *dat_P_sub_c* file was most applicable for creating a queuing model to analyze and predict wait times. Its 114,132 observations span the 4 month period from September 1st, 2028 to December 31st, 2028 for the Auckland (AUC) airfield. Although the very similar *BASA_AUC_2028_912* dataset also focuses on the AUC airfield, has a similar number of entries, and spans a similar time period (even including a few entries from August 2028), the *dat_P_sub_c* dataset has a marginally smaller percentage of missing *S1* values, which was useful for our analysis. It also has additional columns (the last 6 entries in Table 1) that we used for some visualizations.

Below is a summary of *dat_P_sub_c*'s variables:

Variable	Explanation	Example
Pass_ID	Unique # given to passenger when scanned at S2	5348209
valid_P_ID	Determines whether passenger ID is valid?	Every entry for this variable has a value of 1
Airfield	Airfield from which passenger departed from	AUC
S2	Datetime when passenger was scanned at S2	2028-09-01 10:07
Wait_Time	Time passenger takes to go from S1 to S2 rounded up to nearest minute(may or may not exist, not all passengers are scanned at S1)	4
C_Start	# of agents at S1 when passenger arrives (collected in 15 second intervals)	3
C0	# of agents at S2 when passenger arrives (collected in 15 second intervals)	2
C_avg	Average # of agents while passenger is in PBS queue (collected in 15 second intervals)	2.5
Sch_Departure	Datetime flight was scheduled to depart	2028-09-01 12:04
Act_Departure	Datetime flight actually departed	2028-09-01 12:04
BFO_Dest_City	Destination of flight (if Borean airfield only a TLA)	VES064, BORQUE
BFO_Destination_Country_Code	Country code for flight destination	BOR
order	Unique # assigned to passenger when scanned at S2, corresponds to order	5377821
Departure_Date	Date Flight Departed	2028-09-01
Time_of_Day	Time of Day Flight Departed	3 - AFTERNOON
Period_of_Week	Period of Week Flight Departed	1 - WEEKDAY
Day_of_Week	Day of Week Flight Departed	5 - FRI
Month	Month Flight Departed	09 - SEP
Season	Season Flight Departed	3 - Summer
Year	Year Flight Departed	2028
WT_flag	Indicator Used to Determine if Passenger did not wait (1 if Wait_Time == NA, 0 otherwise)	{0,1}
S2_Sch_Flag	Flag indicating passenger is scanned at S2 after scheduled flight departure (1 if scanned after, 0 else)	{0,1}
S2_Act_Flag	Flag indicating passenger is scanned at S2 after actual flight departure (1 if scanned after, 0 else)	{0,1}
Sch_Act_Flag	Flag indicating flight departs before scheduled. (1 if departs before scheduled, 0 else)	{0,1}
Flight_ID	ID number assigned to each respective flight	18096
Delay_in_Seconds	Act_Departure - Sch_Departure (measured in seconds)	100200, -1400

[1]

Table 1.

[1] Some variable definitions were taken from https://www.data-action-lab.com/wp-content/uploads/2021/08/IQC_ch_3_CS.pdf

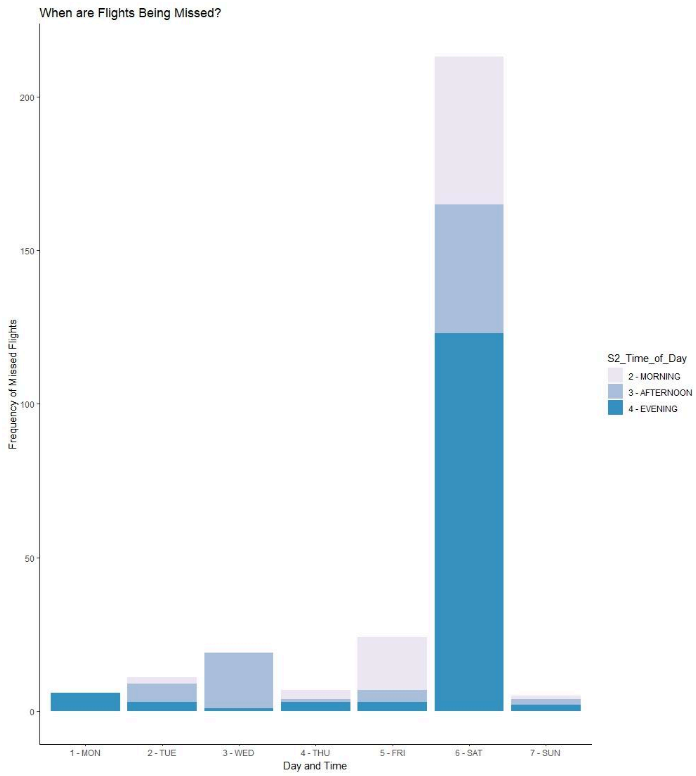


Figure 3.

One important thing we will be curious about when suggesting passenger departure times is when flights are being missed in the past. The graphic on the left highlights the frequency of missed flights departing from the Auckland airport within our collected data. These are passengers that have been scanned through the S2 checkpoint after their flight has already departed. The clear outlier within this data is the spike of missed flights on Saturdays. Why is this so? Should we suggest passengers to leave for the airport earlier on Saturdays in order to not miss their flights?

We see in Figure 4 that the busiest time of day is the morning, with an average wait time in the PBS queue of about 7 minutes. The least busy time of day is the evening, with an average wait time in the PBS queue of about 4 minutes. This could potentially suggest that passengers should arrive at the airport earlier if their flight departure is in the morning or even at night.

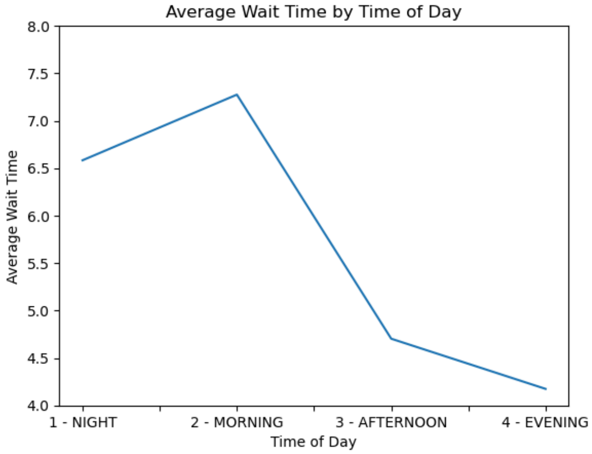


Figure 4.

We see in Figure 5 that Thursday and Monday are the most and least busy days with average wait times of about 7 minutes and 4 minutes, respectively. This graph suggests that passengers should potentially arrive earlier to the airport on Thursdays.

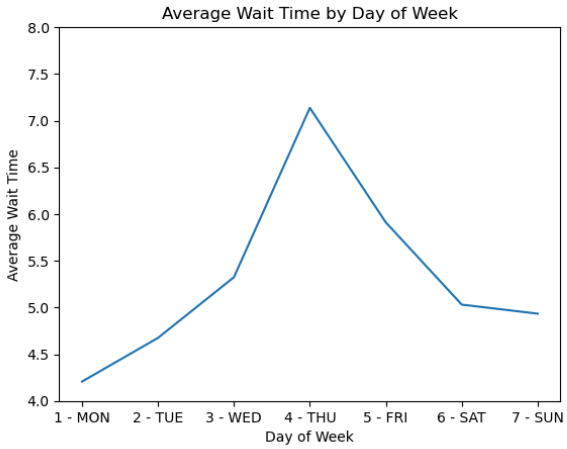


Figure 5.

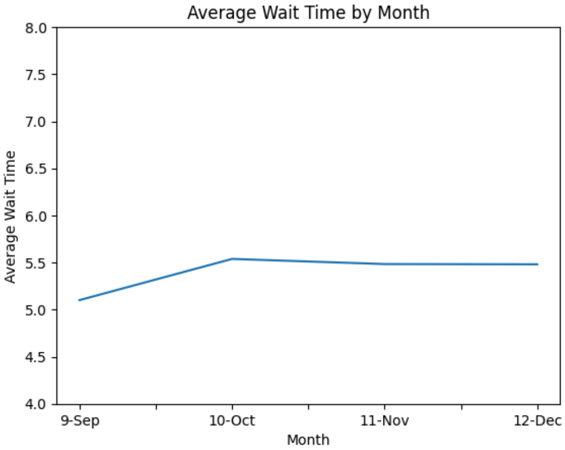


Figure 6.

In Figure 6, we see that average wait time does not vary much from month to month – September is just slightly less busy than the other months. Therefore there is no indication that passengers should adjust their arrival times to the airport depending on the month.

We also looked at the distribution of passenger wait time for each number of servers at S1 when the passenger arrives (Figure 7a) and the distribution of passenger wait time for each number of servers at S2 (Figure 7b). The observation for both is that when there are more servers, the passenger wait time is greater on average. Although, there are more instances of very large wait times when there are only 1 or 2 servers at S1 or S2 – there are instances of wait times over an hour.

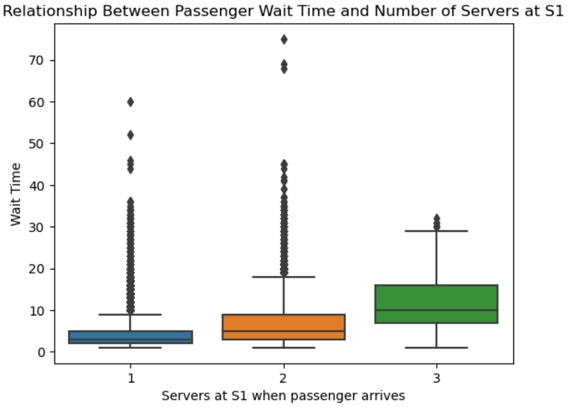


Figure 7a.

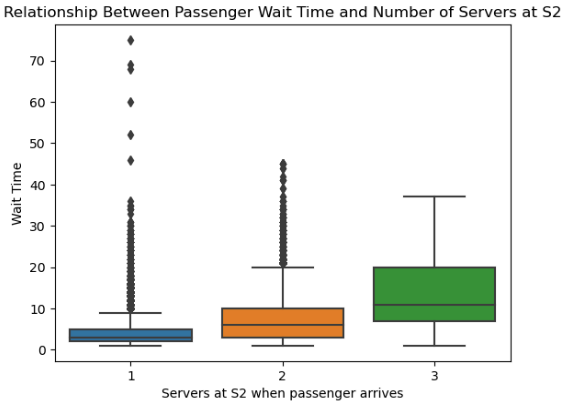


Figure 7b.

Next, we sought out to see if there are any significant correlations between variables in the dataset. For example, we wanted to see if there was a relationship between wait times and flight delays. We looked at average wait times for each time of day, and compared this to average flight delays for each time of day. The correlation we found between the two variables is 0.31. This is a relatively weak positive correlation, suggesting that there is potentially a relationship between the two variables. That is, the longer the wait times at a certain time of day, the more flights are delayed. We also did the same calculation but grouping average wait times and average delays by day of week and month, for which the correlations were 0.30

and 0.15, respectively. Although weak correlations, these correlations give the same conclusion – that the longer the average wait times are, the more flights are delayed.

We also wanted to investigate schedule intensity as a factor of PBS wait time. We define schedule intensity as the number of scheduled flights per day, and the average wait times per day. The correlation is 0.11, suggesting that the more flights in a day, the more passengers waited in the PBS queue (although it is a relatively weak correlation). Also, passengers on flights with greater volumes waited more on average, with a correlation of 0.12.

The queuing model created aims to provide a full prediction on wait times at the different server availability levels of 1, 2 and 3. It is based on the AUC airfield and uses the following assumptions:

- No skipping happens while passengers are in the queue, between S1 and S2.
- No passengers leave the queue after getting into it.
- Passengers commute as individuals and thus, each arrival and process time is independent of all others. This also leads us to believe that the arrival rate for passengers is exponentially distributed.

Additionally, we would like to make the following clarifications regarding our model:

As about 14.78% of boarding passes were not scanned at S1 (NA in S1 column), we do not have access to their wait time data. As opposed to imputing the missing data or assuming a zero wait time, we have opted not to include these records in our model. Since we do not know why certain passengers were not scanned at S1, we assume that it is random. Therefore, we assume that not including these data points does not introduce bias into our model.

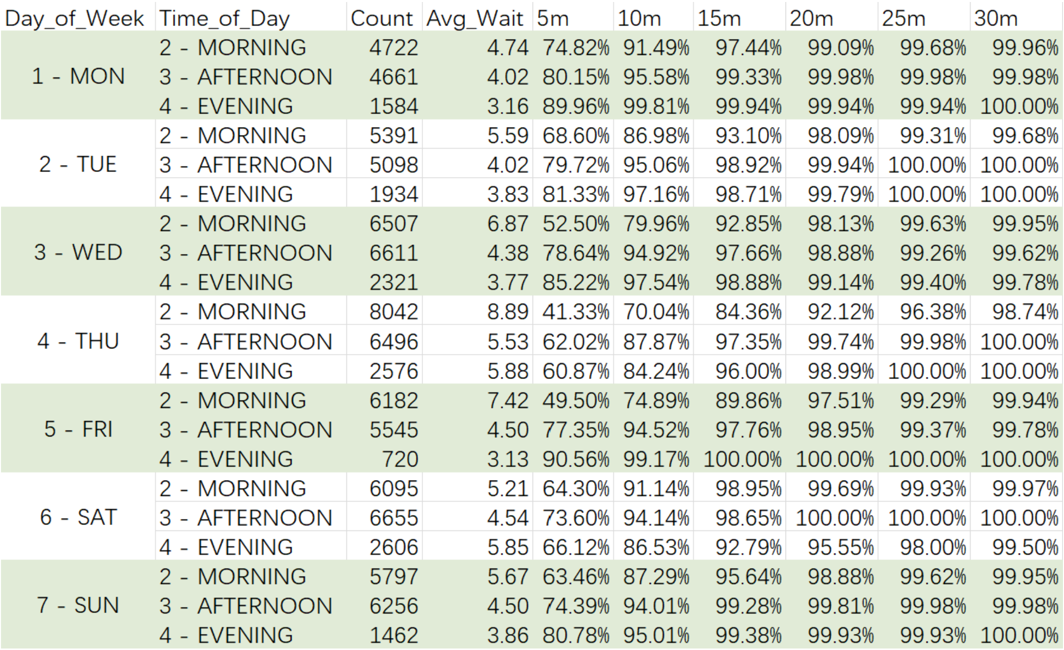
Clusters are divided by day of week, which are further divided by time of day. We experimented with 4 hour periods, however, the clusters turned out to be very small. Because of this we adjusted our clusters to be subdivided with the six hour time periods listed below.

- 6:00am - 11:59am is referred to as *Morning*
- 12:00pm - 6:00pm is referred to as *Afternoon*
- 6:00pm - 11:59pm is referred to as *Evening*
- 12:00am - 6:00am is referred to as *Night*

Our clusters are based on when a given passenger was scanned at S2. Therefore, if an individual entered the queue at a time belonging to a certain cluster and exited the queue at a time belonging to a different cluster, they would be grouped into the latter.

However, we discovered that there were no scans at S2 between the hours of 12:00am and 6:00am. We therefore assume that even though there are some departures scheduled between these hours, the S1 and S2 queue is not open between these hours.

After clustering and re-categorization, we calculate the average waiting time for each time block and their quality of service (QoS).

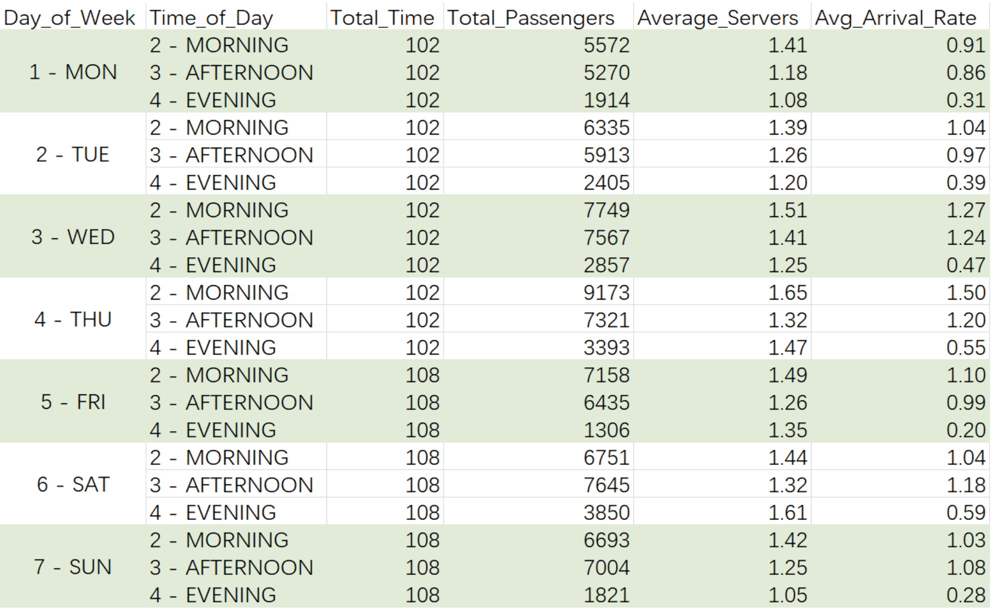


Day_of_Week	Time_of_Day	Count	Avg_Wait	5m	10m	15m	20m	25m	30m
1 - MON	2 - MORNING	4722	4.74	74.82%	91.49%	97.44%	99.09%	99.68%	99.96%
	3 - AFTERNOON	4661	4.02	80.15%	95.58%	99.33%	99.98%	99.98%	99.98%
	4 - EVENING	1584	3.16	89.96%	99.81%	99.94%	99.94%	99.94%	100.00%
2 - TUE	2 - MORNING	5391	5.59	68.60%	86.98%	93.10%	98.09%	99.31%	99.68%
	3 - AFTERNOON	5098	4.02	79.72%	95.06%	98.92%	99.94%	100.00%	100.00%
	4 - EVENING	1934	3.83	81.33%	97.16%	98.71%	99.79%	100.00%	100.00%
3 - WED	2 - MORNING	6507	6.87	52.50%	79.96%	92.85%	98.13%	99.63%	99.95%
	3 - AFTERNOON	6611	4.38	78.64%	94.92%	97.66%	98.88%	99.26%	99.62%
	4 - EVENING	2321	3.77	85.22%	97.54%	98.88%	99.14%	99.40%	99.78%
4 - THU	2 - MORNING	8042	8.89	41.33%	70.04%	84.36%	92.12%	96.38%	98.74%
	3 - AFTERNOON	6496	5.53	62.02%	87.87%	97.35%	99.74%	99.98%	100.00%
	4 - EVENING	2576	5.88	60.87%	84.24%	96.00%	98.99%	100.00%	100.00%
5 - FRI	2 - MORNING	6182	7.42	49.50%	74.89%	89.86%	97.51%	99.29%	99.94%
	3 - AFTERNOON	5545	4.50	77.35%	94.52%	97.76%	98.95%	99.37%	99.78%
	4 - EVENING	720	3.13	90.56%	99.17%	100.00%	100.00%	100.00%	100.00%
6 - SAT	2 - MORNING	6095	5.21	64.30%	91.14%	98.95%	99.69%	99.93%	99.97%
	3 - AFTERNOON	6655	4.54	73.60%	94.14%	98.65%	100.00%	100.00%	100.00%
	4 - EVENING	2606	5.85	66.12%	86.53%	92.79%	95.55%	98.00%	99.50%
7 - SUN	2 - MORNING	5797	5.67	63.46%	87.29%	95.64%	98.88%	99.62%	99.95%
	3 - AFTERNOON	6256	4.50	74.39%	94.01%	99.28%	99.81%	99.98%	99.98%
	4 - EVENING	1462	3.86	80.78%	95.01%	99.38%	99.93%	99.93%	100.00%

Table 2. Average queuing time and distribution in the original data where *Count* represents the amount of people waiting in the given cluster, *Avg_Wait* represents the average wait time for each cluster and the rest of the columns represent the likelihood a passenger will wait that amount of time or less.

Overall, Thursdays have the longest average wait times. Mornings also have longer average wait times than afternoons and evenings. This aligns with our observations made from Figure 5 and Figure 6.

From Table 2 alone, we cannot provide much information for passengers. Therefore we need to get more information such as average arrival rate and average service rate for each time period.



Day_of_Week	Time_of_Day	Total_Time	Total_Passengers	Average_Servers	Avg_Arrival_Rate
1 - MON	2 - MORNING	102	5572	1.41	0.91
	3 - AFTERNOON	102	5270	1.18	0.86
	4 - EVENING	102	1914	1.08	0.31
2 - TUE	2 - MORNING	102	6335	1.39	1.04
	3 - AFTERNOON	102	5913	1.26	0.97
	4 - EVENING	102	2405	1.20	0.39
3 - WED	2 - MORNING	102	7749	1.51	1.27
	3 - AFTERNOON	102	7567	1.41	1.24
	4 - EVENING	102	2857	1.25	0.47
4 - THU	2 - MORNING	102	9173	1.65	1.50
	3 - AFTERNOON	102	7321	1.32	1.20
	4 - EVENING	102	3393	1.47	0.55
5 - FRI	2 - MORNING	108	7158	1.49	1.10
	3 - AFTERNOON	108	6435	1.26	0.99
	4 - EVENING	108	1306	1.35	0.20
6 - SAT	2 - MORNING	108	6751	1.44	1.04
	3 - AFTERNOON	108	7645	1.32	1.18
	4 - EVENING	108	3850	1.61	0.59
7 - SUN	2 - MORNING	108	6693	1.42	1.03
	3 - AFTERNOON	108	7004	1.25	1.08
	4 - EVENING	108	1821	1.05	0.28

Table 3. Total Time represents the total days that a cluster occurs multiplied by 6 hours (the length of a cluster). Total Passengers represents the total number of people in the cluster.

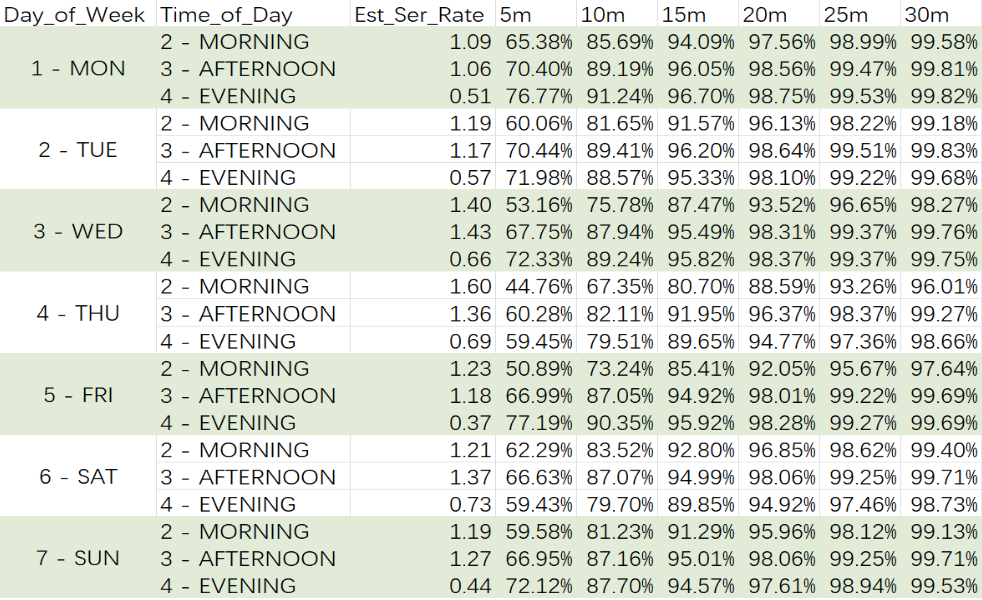
The *Avg_Arrival_Rate* unit in Table 3 is the number of people arriving at S1 per minute. *Average_Servers* is the average number of servers active at S2 while passengers move through the queue between S1 and S2. The reason we only calculate the average number of servers in S2 and exclude the number of servers in S1 is that our queuing system only occurs between S1 and S2, and therefore the performance of this queuing system depends on the number of servers in S2, not the number of servers in S1.

One key parameter we are missing is the average service rate for S2. One way to calculate the average service rate is to compute the time when S2's service starts and the time when S2's service ends, and then take the average. However, we do not have the end time of S2 in our data, so we need to estimate the average service rate of S2. We use the following formula to estimate the average service rate for each time cluster:

$$\hat{\mu} = \frac{\overline{w}\lambda + \sqrt{(\overline{w}\lambda)^2 + 4\overline{w}\lambda}}{2\overline{w}}$$

With the estimated service rate, we can finally find the estimated QoS, $\hat{p}(x) = 1 - \frac{\lambda}{\hat{\mu}} e^{-(\hat{\mu}-\lambda)x}$.

The following table shows the estimated service rate and estimate probability of waiting up to x units of time.



Day_of_Week	Time_of_Day	Est_Ser_Rate	5m	10m	15m	20m	25m	30m
1 - MON	2 - MORNING	1.09	65.38%	85.69%	94.09%	97.56%	98.99%	99.58%
	3 - AFTERNOON	1.06	70.40%	89.19%	96.05%	98.56%	99.47%	99.81%
	4 - EVENING	0.51	76.77%	91.24%	96.70%	98.75%	99.53%	99.82%
2 - TUE	2 - MORNING	1.19	60.06%	81.65%	91.57%	96.13%	98.22%	99.18%
	3 - AFTERNOON	1.17	70.44%	89.41%	96.20%	98.64%	99.51%	99.83%
	4 - EVENING	0.57	71.98%	88.57%	95.33%	98.10%	99.22%	99.68%
3 - WED	2 - MORNING	1.40	53.16%	75.78%	87.47%	93.52%	96.65%	98.27%
	3 - AFTERNOON	1.43	67.75%	87.94%	95.49%	98.31%	99.37%	99.76%
	4 - EVENING	0.66	72.33%	89.24%	95.82%	98.37%	99.37%	99.75%
4 - THU	2 - MORNING	1.60	44.76%	67.35%	80.70%	88.59%	93.26%	96.01%
	3 - AFTERNOON	1.36	60.28%	82.11%	91.95%	96.37%	98.37%	99.27%
	4 - EVENING	0.69	59.45%	79.51%	89.65%	94.77%	97.36%	98.66%
5 - FRI	2 - MORNING	1.23	50.89%	73.24%	85.41%	92.05%	95.67%	97.64%
	3 - AFTERNOON	1.18	66.99%	87.05%	94.92%	98.01%	99.22%	99.69%
	4 - EVENING	0.37	77.19%	90.35%	95.92%	98.28%	99.27%	99.69%
6 - SAT	2 - MORNING	1.21	62.29%	83.52%	92.80%	96.85%	98.62%	99.40%
	3 - AFTERNOON	1.37	66.63%	87.07%	94.99%	98.06%	99.25%	99.71%
	4 - EVENING	0.73	59.43%	79.70%	89.85%	94.92%	97.46%	98.73%
7 - SUN	2 - MORNING	1.19	59.58%	81.23%	91.29%	95.96%	98.12%	99.13%
	3 - AFTERNOON	1.27	66.95%	87.16%	95.01%	98.06%	99.25%	99.71%
	4 - EVENING	0.44	72.12%	87.70%	94.57%	97.61%	98.94%	99.53%

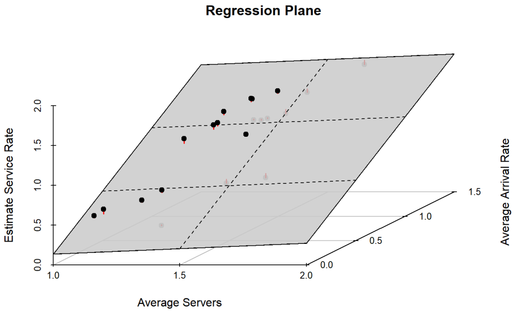
Table 4. Estimated average service rate and QoS (Generalized M/M/1)

Are our estimations trustworthy? For comparison, we go back in the other direction using a regression approach to estimate the average service rate.

Regression Approach

$$\hat{\mu} = ac + b\lambda$$

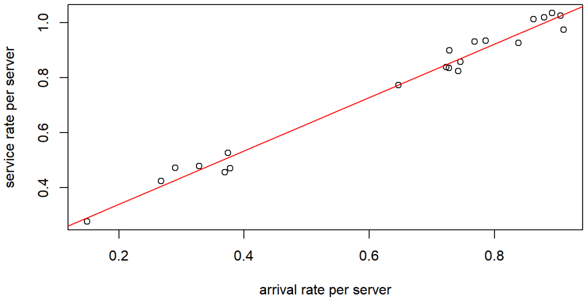
After modeling the regression using R, we get the following model and related information:



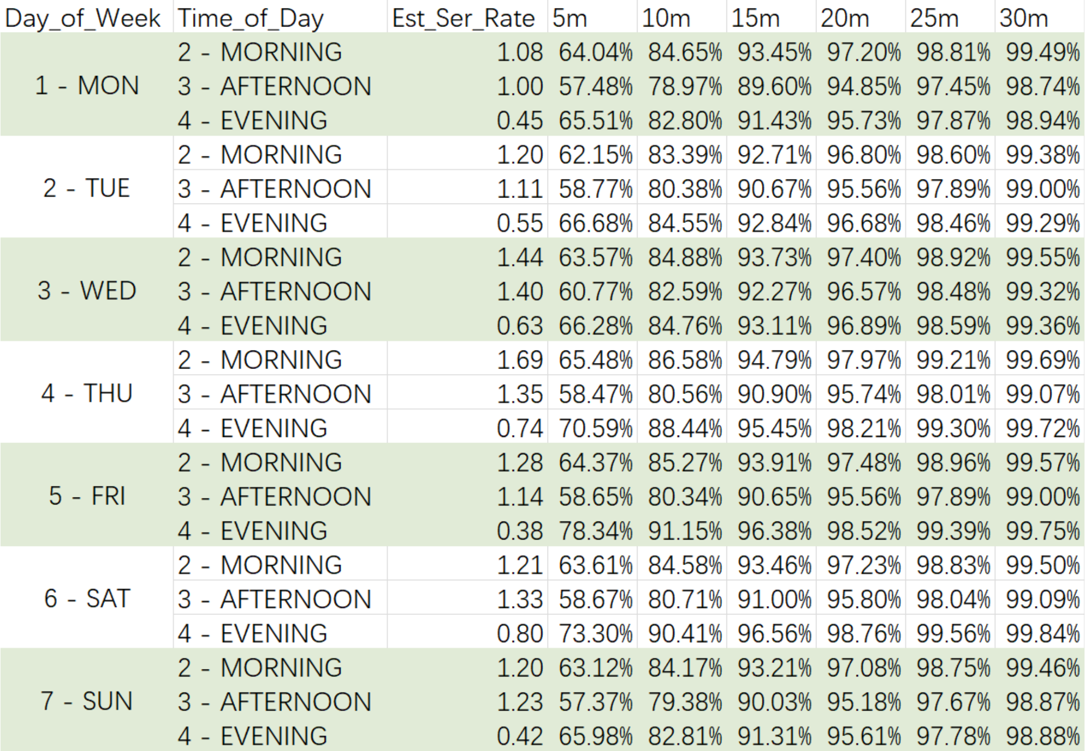
$$\hat{\mu} = 0.13600c + 0.97672\lambda$$

```
Coefficients:
                          Est Std.   Error     t value   Pr(>|t|)
df_reg$Average_Servers    0.13600    0.02094   6.495     3.19e-06
df_reg$Avg_Arrival_Rate   0.97672    0.03022   32.320    < 2e-16

Residual standard error: 0.04456 on 19 degrees of freedom
Multiple R-squared:  0.9985,    Adjusted R-squared:  0.9983
F-statistic:  6319 on 2 and 19 DF,  p-value: < 2.2e-16
```



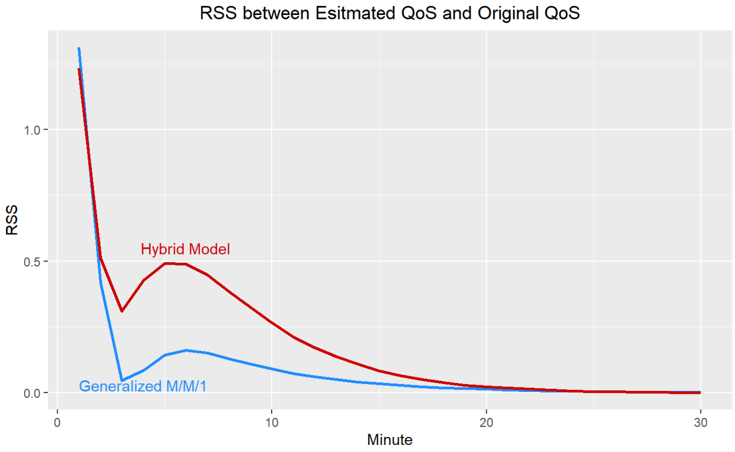
From the above information and plots, we see that the relationship between average servers, average arrival rate and estimated service rate are strong. This means that a change in any two parameters in the model will have an effect on the remaining one. For example, assuming that the number of servers in S2 remains constant, (i.e. the average number of servers also remains constant), we observe a decrease in the service rate per unit of time, which implies a decrease in arriving passengers. We then use the model to estimate the service rate and QoS in S2. We obtain the following table:



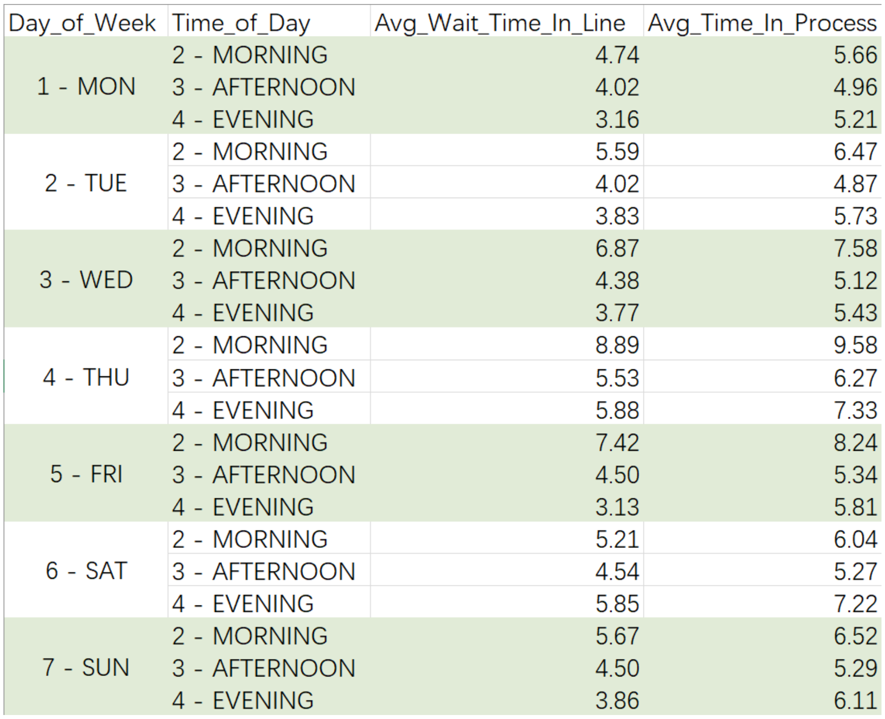
Day_of_Week	Time_of_Day	Est_Ser_Rate	5m	10m	15m	20m	25m	30m
1 - MON	2 - MORNING	1.08	64.04%	84.65%	93.45%	97.20%	98.81%	99.49%
	3 - AFTERNOON	1.00	57.48%	78.97%	89.60%	94.85%	97.45%	98.74%
	4 - EVENING	0.45	65.51%	82.80%	91.43%	95.73%	97.87%	98.94%
2 - TUE	2 - MORNING	1.20	62.15%	83.39%	92.71%	96.80%	98.60%	99.38%
	3 - AFTERNOON	1.11	58.77%	80.38%	90.67%	95.56%	97.89%	99.00%
	4 - EVENING	0.55	66.68%	84.55%	92.84%	96.68%	98.46%	99.29%
3 - WED	2 - MORNING	1.44	63.57%	84.88%	93.73%	97.40%	98.92%	99.55%
	3 - AFTERNOON	1.40	60.77%	82.59%	92.27%	96.57%	98.48%	99.32%
	4 - EVENING	0.63	66.28%	84.76%	93.11%	96.89%	98.59%	99.36%
4 - THU	2 - MORNING	1.69	65.48%	86.58%	94.79%	97.97%	99.21%	99.69%
	3 - AFTERNOON	1.35	58.47%	80.56%	90.90%	95.74%	98.01%	99.07%
	4 - EVENING	0.74	70.59%	88.44%	95.45%	98.21%	99.30%	99.72%
5 - FRI	2 - MORNING	1.28	64.37%	85.27%	93.91%	97.48%	98.96%	99.57%
	3 - AFTERNOON	1.14	58.65%	80.34%	90.65%	95.56%	97.89%	99.00%
	4 - EVENING	0.38	78.34%	91.15%	96.38%	98.52%	99.39%	99.75%
6 - SAT	2 - MORNING	1.21	63.61%	84.58%	93.46%	97.23%	98.83%	99.50%
	3 - AFTERNOON	1.33	58.67%	80.71%	91.00%	95.80%	98.04%	99.09%
	4 - EVENING	0.80	73.30%	90.41%	96.56%	98.76%	99.56%	99.84%
7 - SUN	2 - MORNING	1.20	63.12%	84.17%	93.21%	97.08%	98.75%	99.46%
	3 - AFTERNOON	1.23	57.37%	79.38%	90.03%	95.18%	97.67%	98.87%
	4 - EVENING	0.42	65.98%	82.81%	91.31%	95.61%	97.78%	98.88%

Table 5. Estimated average service rate and QoS (Regression + Generalized M/M/1)

It is easy to see that the RSS of the hybrid model is larger than that of the generalized M/M/1 for the 30-minute time period, so using the estimated service rate from the generalized M/M/1 minimizes the bias of the calculation in the next step.



Once we have the estimated average service rate as well as the average arrival rate for each time period, we can provide relevant information to passengers who are about to enter the PBS queue, such as the possible waiting time, the number of people currently pending at the PBS queue, and other information.



Day_of_Week	Time_of_Day	Avg_Wait_Time_In_Line	Avg_Time_In_Process
1 - MON	2 - MORNING	4.74	5.66
	3 - AFTERNOON	4.02	4.96
	4 - EVENING	3.16	5.21
2 - TUE	2 - MORNING	5.59	6.47
	3 - AFTERNOON	4.02	4.87
	4 - EVENING	3.83	5.73
3 - WED	2 - MORNING	6.87	7.58
	3 - AFTERNOON	4.38	5.12
	4 - EVENING	3.77	5.43
4 - THU	2 - MORNING	8.89	9.58
	3 - AFTERNOON	5.53	6.27
	4 - EVENING	5.88	7.33
5 - FRI	2 - MORNING	7.42	8.24
	3 - AFTERNOON	4.50	5.34
	4 - EVENING	3.13	5.81
6 - SAT	2 - MORNING	5.21	6.04
	3 - AFTERNOON	4.54	5.27
	4 - EVENING	5.85	7.22
7 - SUN	2 - MORNING	5.67	6.52
	3 - AFTERNOON	4.50	5.29
	4 - EVENING	3.86	6.11

Table 6.

The average queue and processing time as shown in table 6 could offer many benefits to passengers, services suggesting arrival times, as well as the airport themselves. A practical use may involve displaying these times as estimates to when passengers may expect to pass through gates.

In Conclusion,

The tables above given by our queueing model indicate a clear distinction in estimated wait times between times of day. Later periods in the day tend to have smaller wait times, which aligns with our passenger frequency charts. Therefore, we heavily encourage passengers to consider the time of day to determine at what point they arrive at the airport. Particularly, they should arrive earlier if their flight is in the morning.

Although estimated wait times generally fall in line with each other throughout the week, there are days we can highlight as being more likely to expect longer or shorter wait times. Based on our analysis, Thursdays tend to lead to longer expected wait times while earlier in the week, on Mondays and Tuesdays, there is a higher likelihood that passengers will be able to pass through checks in a shorter time frame. Based on the model, we determined that we can disregard the trend seen in Figure 3, as Saturdays do not seem to lead to higher wait times.

As *Foogle analytics*, we offer the estimates and probabilities given by our model to allow passengers to determine acceptable times to arrive at the airport. We hope relevant companies can use these numbers as a baseline for their suggestions to those who will be departing from the Auckland Airport.