Analysis of Cancellations at a Cab Portal Company

Teal Lucas
WISCONSIN SCHOOL OF BUSINESS
October 2024

Executive Summary

This report was created to display my analysis of the data of IndoCabs, an India based cab portal company. Although they have had to suspend operations due to the issue of drivers commonly cancelling on customers, I am confident that this issue can be solved, and IndoCabs will be able to take steps back towards success in the near future. This report provides several visual representations of the data, and recommendations for solving the issues that have caused IndoCabs to suspend operations.

Upon filtering the data I was given, my analysis of 2,853 sample bookings concluded several patterns that give explanation towards why the cancellations may be as common as they are. It was found that different types of travel IndoCabs offers has an effect on cancellations, with the highest proportion happening with Point to Point travel. I also found that average cancellations vary from day to day, and hour to hour, with the most occurring on Sundays and Mondays, in the 9^{th} and 19^{th} hours of the day. This can be fixed by incentivizing drivers to not cancel on these days at these times, with things like premium wages. Additionally, the data shows that most cancellations occur within 1 day booking windows, and the smaller the window, the more common a cancellation is. In conclusion, this report will give several representations of important findings in the data. Several varying numbers and charts are displayed, and patterns will be shown that will truly tell us what needs to be done to effectively return to successful operations.

Introduction

IndoCabs, an India based cab portal company, has had to suspend operations following difficulties in its ability to guarantee cabs for customers needing rides. A non-traditional ride service, IndoCabs does not sustain their own fleet, but they provide a platform for smaller cab providers to use their technology pricing, and logistics. Their issue has risen from inconsistencies in their vendors' schedules. These inconsistencies paired with a poor algorithm allowing customers to book a cab even though the driver may be very likely to cancel is harming the company. Due to this issue, many of their customers have had inconvenience and lost trust in the brand. Through a meticulous data analysis, I have found several numerical conclusions that provide some guidance to IndoCabs' issues. Through company report examinations, it has been hypothesized that trip durations and methods of booking may play a role in the cancellation issues. This report will tell all of my analytical findings, and draw a helpful conclusion, hopefully restoring IndoCabs' success.

Analysis

A Look at Trip Durations

	Trip Duration (Hours)	Booking Window (Days)
Measure 1 (Average)	4.33	1.97
Measure 2 (Median)	1.34	0.42
Measure 3 (Standard Deviation)	12.09	4.88

Above, 3 of the most important measures (Average, Median, Standard Deviation) are displayed. The average provides a brief view of the central tendency of our data set, displaying the overall level of our variables (Trip Duration and Booking Window). The average is highly useful for understanding the general trend of our data. The median shows the middle value of our long list of data. The benefit of the median over the average is that it still shows our central tendency without being skewed in either direction (up or down) by outliers. Lastly, our standard deviation indicates how much our data is dispersed. The standard deviation provides us with an idea of how our data is distributed around the average.

The Magnitude of the Cancellation Problem at IndoCabs

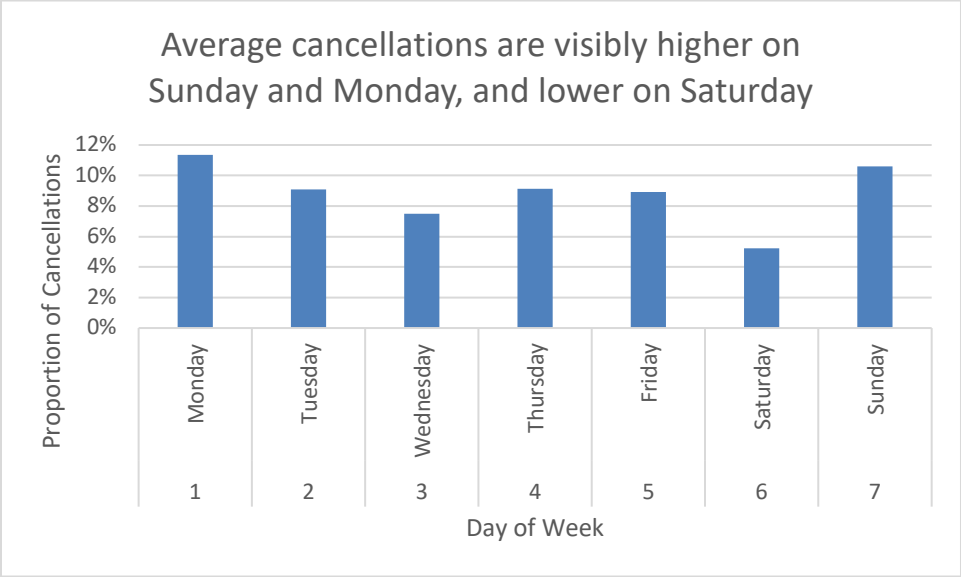
The magnitude of cancellations was easily noticeable, as 9% of our rides were cancelled. Of our 2,853 rides, 249 of them were cancelled.

	Number of Bookings	Number of Cancellations	Proportion Cancelled
Long Distance Travel	128	2	0.02
Point to Point Travel	2251	215	0.10
Hourly Rental Travel	474	32	0.07

Through the analysis of the 3 different travel types defined above, it was found that the most bookings come from Point to Point Travel. This type of travel also had the highest proportion of rides canceled. The least bookings came from Long Distance Travel, and similarly, this had the last proportion of rides canceled. This makes sense when we look at the trip duration, as for Long Distance Travel, the average duration, in hours, was 43.36. Dissimilarly, the average duration in hours of Point to Point Travel was 1.50 hours. Clearly shown, the more spontaneous travelers are, the more likely they are to cancel a trip.

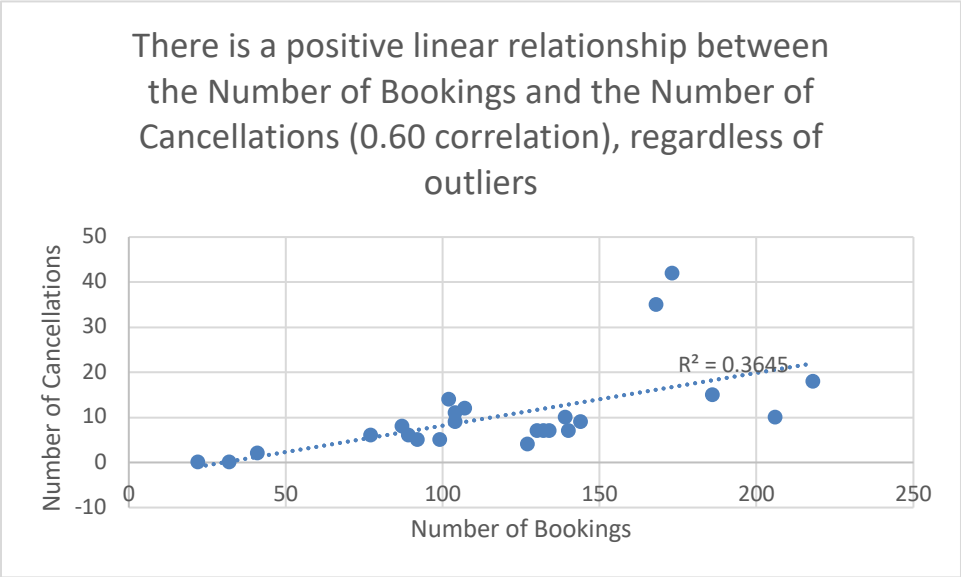
There is somewhat of a meaningful difference between cancellation rates of Online Bookings and Mobile Site Bookings. The average of cancellation rates is substantially higher from Mobile Site Booking, being 20%. For Online Bookings, this average number is only 13%. IndoCabs may want to address the higher level of cancellations from their Mobile Site, and work out a way to decrease this number. It could potentially be due to

the convenience of the Mobile Site, so some sort of incentive may make users less likely to cancel on the Mobile Site.

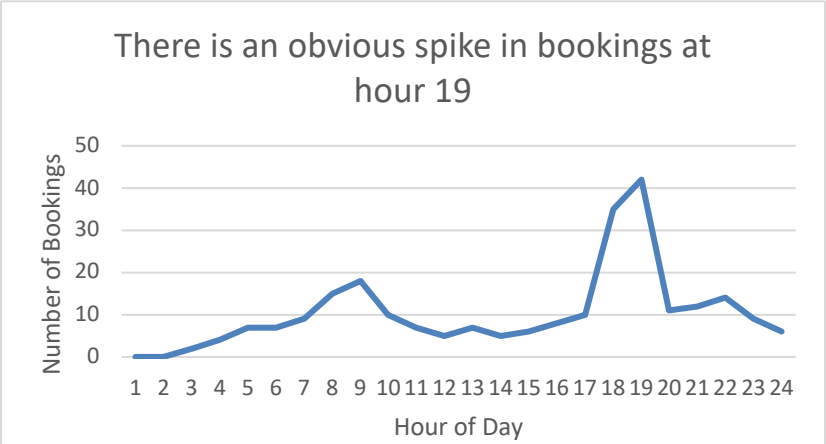


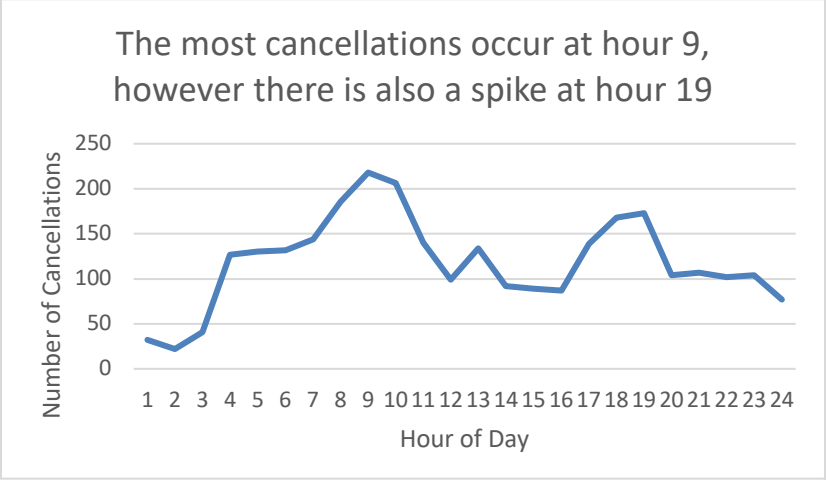
The pattern displayed in the above chart is not quite what I expected. I would've thought the average cancellations would be higher on Saturday, as that is when the most rides occur, however, the most cancellations on average occur on Sunday and Monday, which are relatively in the middle when it comes to the busiest days of the week for Bookings.

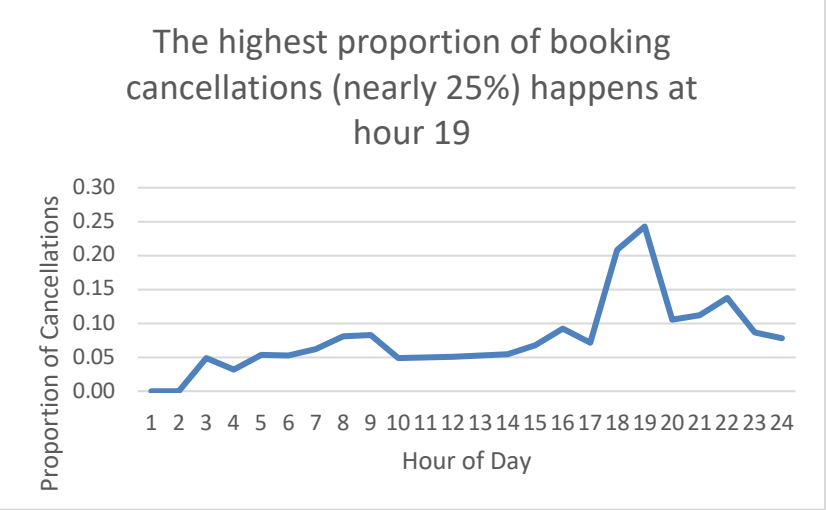
The Relationship between Booking Windows, Cancellations, and Trip Timing



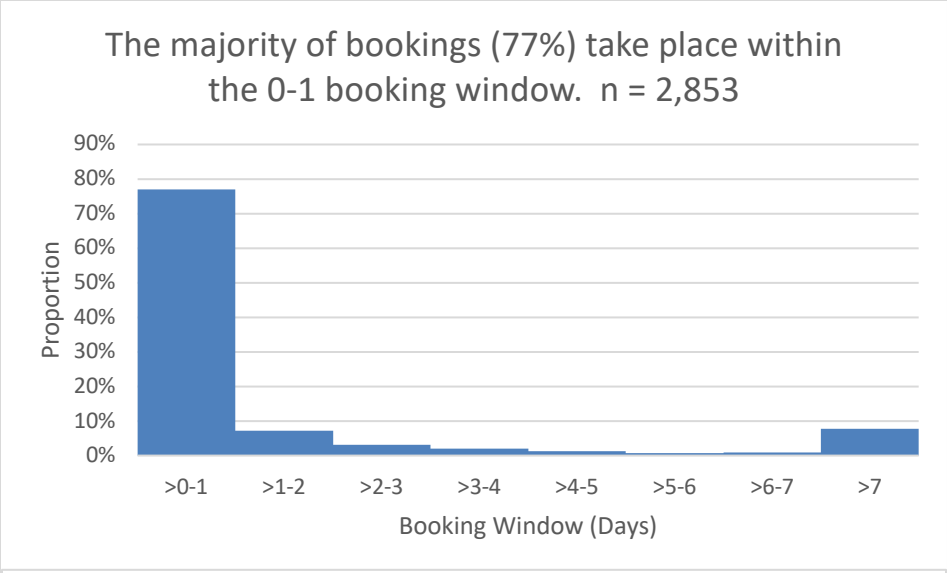
The R-squared value of 0.3645 is neither substantially high nor low. It indicates that our model does capture a fair amount of the relationship between our two variables, but there is still some variation that is not explained. With a correlation of 0.60, our variables have a moderate positive correlation.

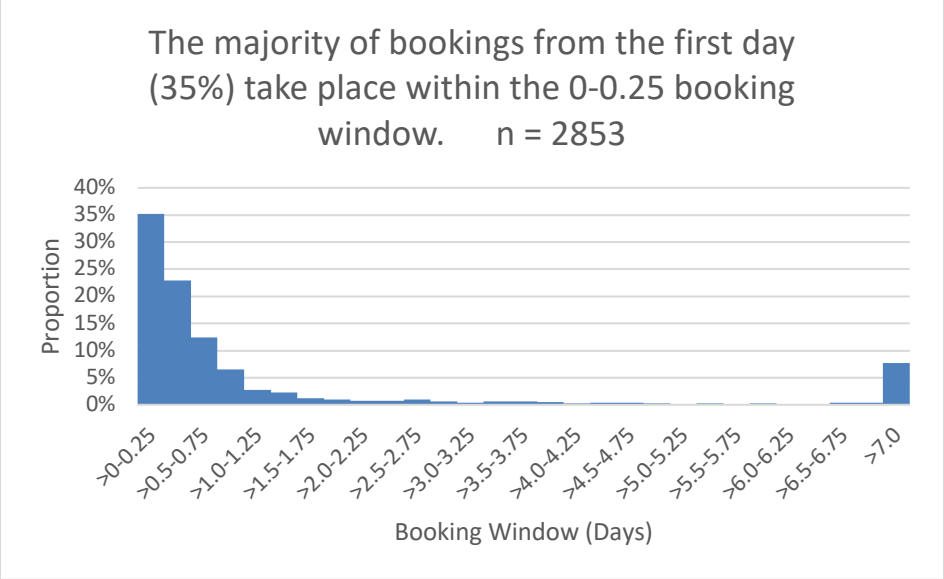






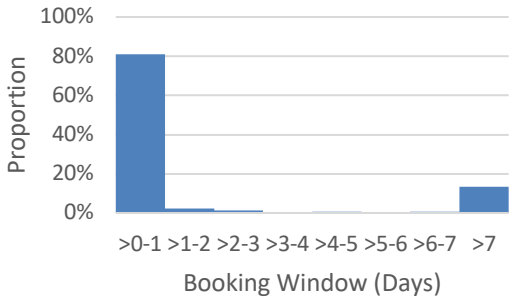
I find the chart showing the number of cancellations to be the most informative, as that is an important criterion in solving IndoCabs' issues. It should also be noted that in hour 19, the most bookings occur, and the highest proportion of cancellations occurs, however the most cancellations occur at hour 9.



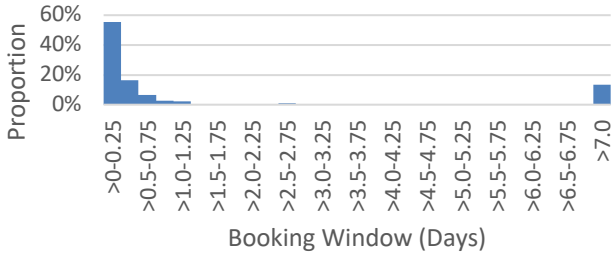


I feel the bin size of 0.25 is a much better indication of the ride cancellations, as 77% of cancellations occur within a 1 day booking window. The bin size of 0.25 portrays a more detailed dive into the proportion of cancellations in each part of each day, not just the day in its entirety. As shown above, the highest proportion of cancellations (35%) occurs in the first quarter of the first day, and this number decreases each 0.25 bin all the way to 7.

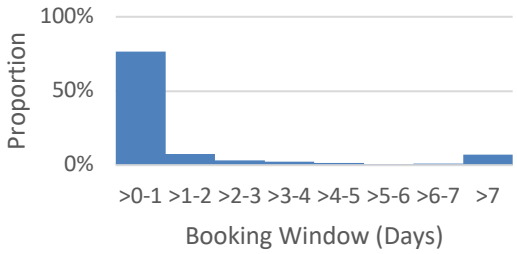
The majority of cancellations (81%) take place within the 0-1 booking window.
n=249



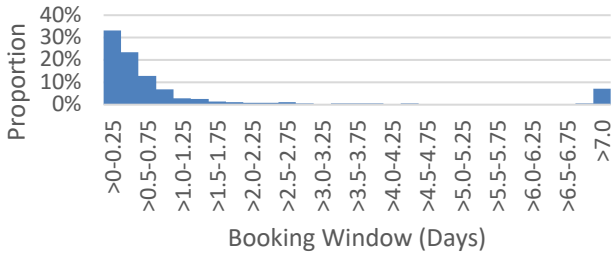
The majority of cancellations from the first day (55%) take place within the 0-0.25 booking window.
n = 249



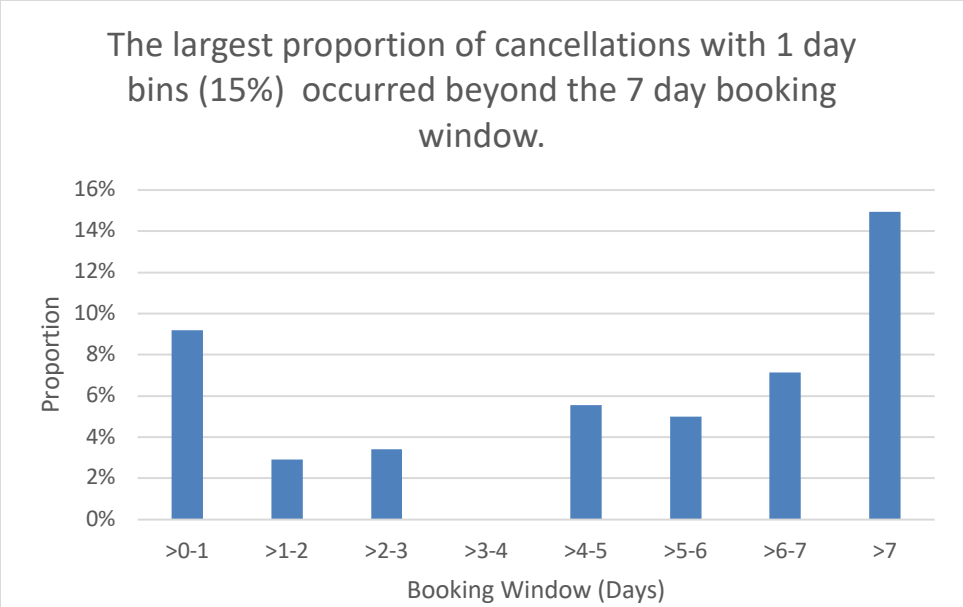
The majority of bookings that were not cancelled (77%) take place within the 0-1 booking window.
n = 2604

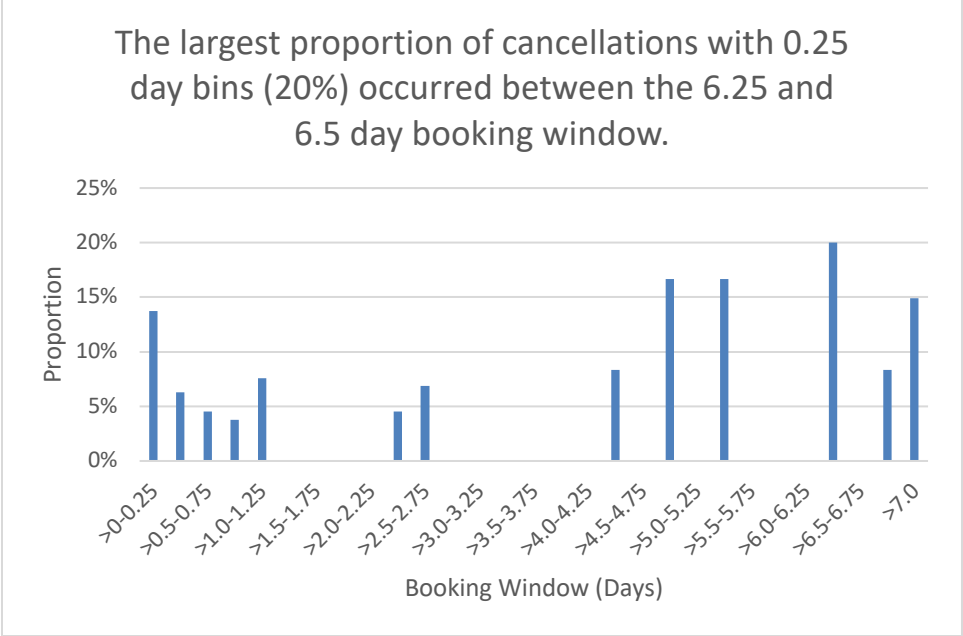


The majority of bookings that were not canelled from the first day (33%) take place within the 0-0.25 booking window.
n = 2604



Booking windows clearly impact cancellations. The majority of cancellations take place within the 0-1 day booking window. However, I feel the 0.25 day bins are much more helpful in communicating insights. The 0.25 day bins are more in depth, and allow more specific data to be displayed.



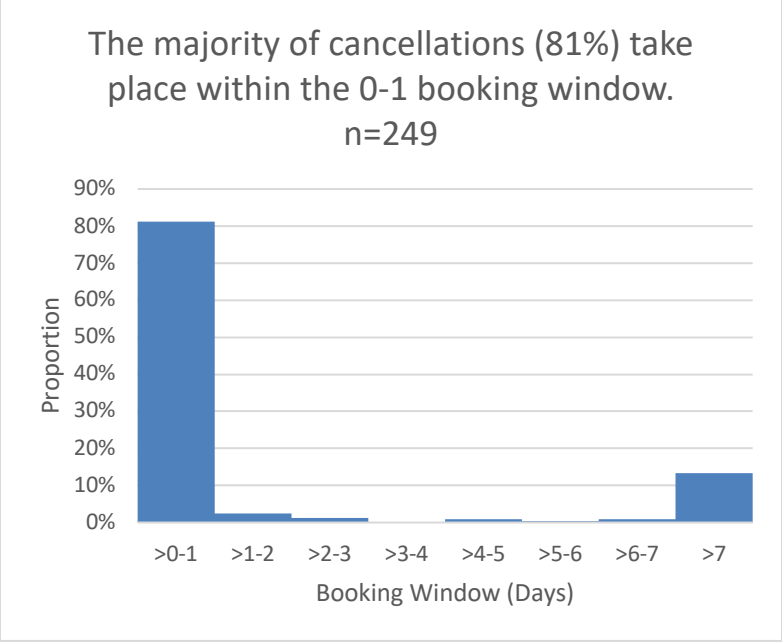


The ratio of cancelled trips to total trips with the 1 day bins is the greatest past the seventh day. The ratio of cancelled trips to total trips with the 0.25 day bins is greatest between 6.25 and 6.5 days.

Recommendations and Conclusion

To conclude this report, it is obvious that there are several repeating trends in the data where the issues are coming from. The booking windows provide us with a comprehensive analysis of when bookings and cancellations have taken place. We have also taken a look at what hour of the day has the most activity. With this information, I have a few recommendations for the company. A large portion of the cancellation issues come from the drivers' end. IndoCabs needs to hold their ride providers to a higher standard. Whether it is incentivizing drivers for giving more short notice rides (paying them more within the 1 day booking window), or cracking down on drivers and telling them that they only get so many ride cancellations each week, management needs to take action. It would also be beneficial to pay higher wages during busy hours, as this will boost employee morale and in-turn result in less ride cancellations. Also, it should be noted that company efforts should be directed towards perfecting Point to Point Travel. The greatest amount of capital should be used with Point to Point, as it is where most issues are occurring. Next, hourly rental. Very few, if any resources, should be put towards Long Distance Travel, as only 2% of rides are cancelled with this travel type.

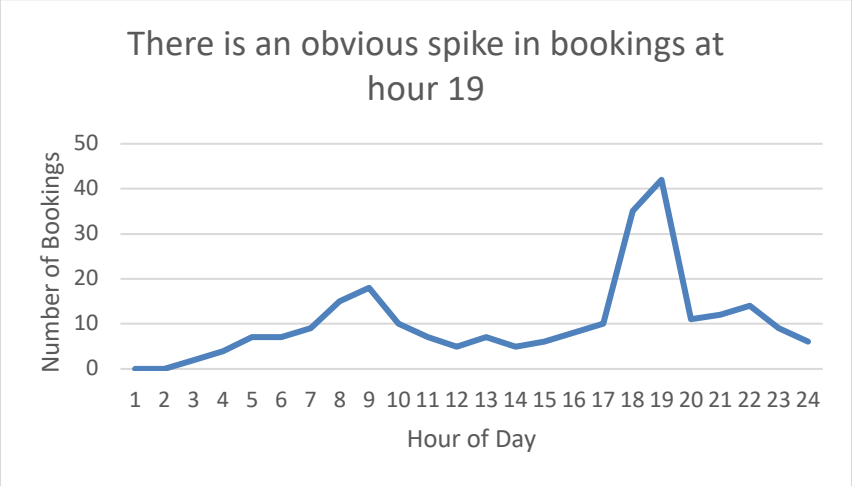
Elevator Charts



I feel the importance of this chart falls within the fact that it emphasizes where the most cancellations occur. This is one of the most important statistics we are analyzing, and it is very easy to read. The 0-1 day bin bar is clearly the highest.

	Number of Bookings	Number of Cancellations	Proportion Cancelled
Long Distance Travel	128	2	0.02
Point to Point Travel	2251	215	0.10
Hourly Rental Travel	474	32	0.07

I feel this table is important because it shows what type of travel is most common, and what proportion of those rides are cancelled. This allows the company to know where resources should be allocated, because there isn't a reason to fix something that does not need to be fixed.



This line graph is important because based on my recommendations, there should be incentives to drivers during busy hours. Hour 19 is clearly the busiest hour of the day, and this graph displays that perfectly.

Notes on Data Preparation

The data referenced in this report, although complex and dense, was prepared meticulously and in an organized way. Data that was outdated or simply irrelevant and unusable was removed. The large set of data provided by IndoCabs was sorted and filtered precisely, and this filtered data was then analyzed effectively. There were no duplicates or erroneous records in the data set because of the precise filtering. Several charts were used to display this analysis, and although many of them show similar statistics, all statistics are integral to the proper analysis. This analysis allowed me to give effective recommendations for the company moving forward.