2019

Analysis of Cancellations at a Cab Portal Company

ERIN BANNON
WISCONSIN SCHOOL OF BUSINESS
OCTOBER 2019

Executive Summary

This report is an analysis of the IndoCabs cancellation problem. The data set consisted of 2,194 trips, 8% of which were cancelled. Most of the cancellations were for Travel Type 2 (city trips) so this is the focus of my analysis and is also why I focused mainly on the median trip duration of 1.39 hours rather than the average which was greatly skewed by long distance trips. Most cancellations happened with short bookings windows of 1 day or less, primarily if booked 0-6 hours before trip start time. The proportion of cancellations was highest on Monday, but all of days of the week excluding Saturday were relatively high as well. The highest proportion of bookings were from 5 am to 6am and 5 pm to 9 pm. The mobile booking channel had only 118 trips booked and of them, 15% were cancelled.

Based on my analysis, I have concluded that IndoCab's cancellation problem mainly occurs with city trips that happen during the morning and evening, especially on Mondays but also on the other days of the week except for Saturdays. This is most likely due to rush hour traffic and the high demand for transportation that comes with it as well as drivers choosing to take trips scheduled through their vendor rather than IndoCabs. I have also concluded that IndoCab's mobile booking should be terminated or improved since it had a high cancellation rate despite not many customers using this channel.

To reduce future cancellations, IndoCabs should enact a few different strategies. IndoCabs should implement surge pricing Monday through Sunday, excluding Saturday, from 5 am to 6 am and from 5 pm to 9 pm to increase supply of drivers and to slightly level out the demand for trips. To help address the conflict of vendors' schedules, consequences should be put in place for drivers who frequently cancel their IndoCabs trips in order to fulfill a trip through their vendor. If a vendor has drivers that are consistently cancelling on IndoCab trips, then IndoCabs should end their partnership with them. Another strategy that IndoCabs should test is requiring drivers to allow more time for trips than what may seem necessary. The median trip duration is 1.39 hours, thus trips from one point to another in a city may be taking longer than what drivers expect. IndoCabs should start out by requiring double the time it would normally take to complete a trip then see if this is causing too much time in between trips or not enough and go from there. Through implementation of these strategies, IndoCabs will be able to greatly reduce their cancellations.

Introduction

The data set analyzed in this report is of 2,194 IndoCabs trips that were comprised of all three trip types and varying trip lengths throughout the year 2013. The data was analyzed for patterns in booking cancellations to determine reasons for IndoCabs' cancellation problem and to make recommendations from the data findings.

Analysis

A Look at Trip Durations

Summary Statistics of Trip Duration and Booking Window

	Trip Duration (Hours)	Booking Window (Days)
Median	1.39	0.42
Average	4.36	1.86
Standard Deviation	12.59	4.72

I chose the above measures because they show how wide the range of data is and that there are large outliers that have greatly skewed the average. I chose to include the median because it is barely affected by outliers. Because of this, there is a significant contrast between the median and average. In this case, it would be most valuable to focus on the median when trying to get a good sense of a typical trip duration and how long the booking window is. I chose to include standard deviation because it reveals even more how skewed the data is by outliers. These outliers come from the long distance travel cab bookings, but there are few cancellations for this type so it is more important to look at the median because it better represents the typical travel time and booking window for the most popular travel type which is travel from one point to another in a city (Travel Type 2).

Magnitude of the Cancellation Problem at IndoCabs

Of the 2,194 bookings, 183 were cancelled, resulting in a proportion of 8% of bookings being cancelled.

Cancellations by Travel Type

	Number of Bookings	Number of Cancellations	Percent Cancelled
Long Distance (Type 1)	87	1	1.15%
City Travel (Type 2)	1,717	165	9.6%
Hourly Rental (Type 3)	390	17	4.35%

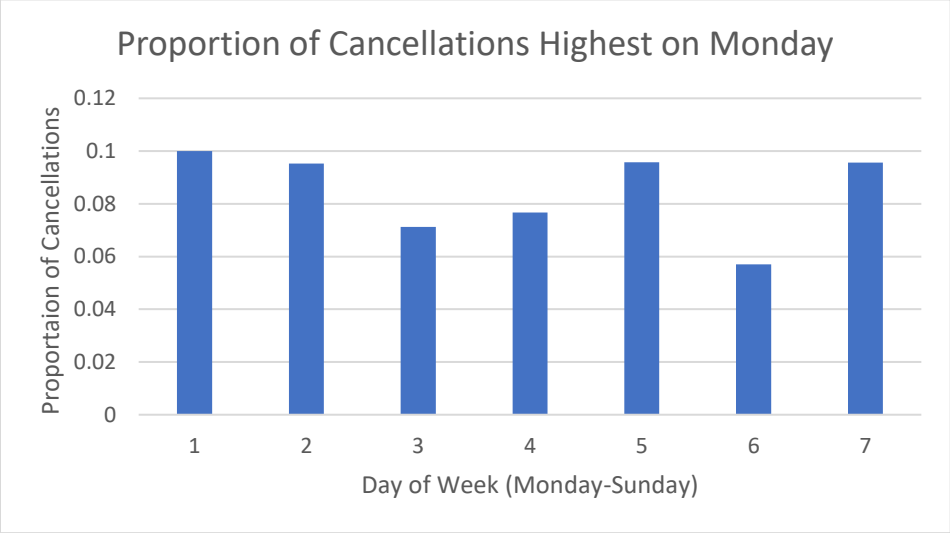
The most amount of bookings and cancellations comes from Travel Type 2 which is travel from point to point in a city. This pattern makes sense because IndoCabs stated in their overview that one of the biggest problems with the company cancelling on customers is because of their issues partnering with vendors' schedules. If cab drivers that are working for vendors are picking up customers who have not booked through IndoCabs, then they are probably taking on trips that conflict with IndoCabs booking seeing as IndoCabs has not mentioned penalizing drivers or vendors who cancel frequently. Because of this, drivers working for vendors are most likely prioritizing their vendor's customers rather than IndoCabs. The low amount of bookings and cancellations for long distance travel type makes sense because IndoCabs' service is most appealing for individuals within cities where there might be fewer options to get from place to place quickly in a city. For long distances, customers most likely have more options such as bus, train, or airplane. The only one cancellation makes sense because these drivers are most likely not affiliated with a vendor and instread are working for themselves since taxi vendors exist mainly in cities and not for long distance travel. For the same reason, I also believe this is

why hourly rental cancellations are lower because taxi vendors primarily cater towards those that are just going from one point to another, so the schedule conflicts are not as much of a problem.

Cancellation Rates and Booking Channels

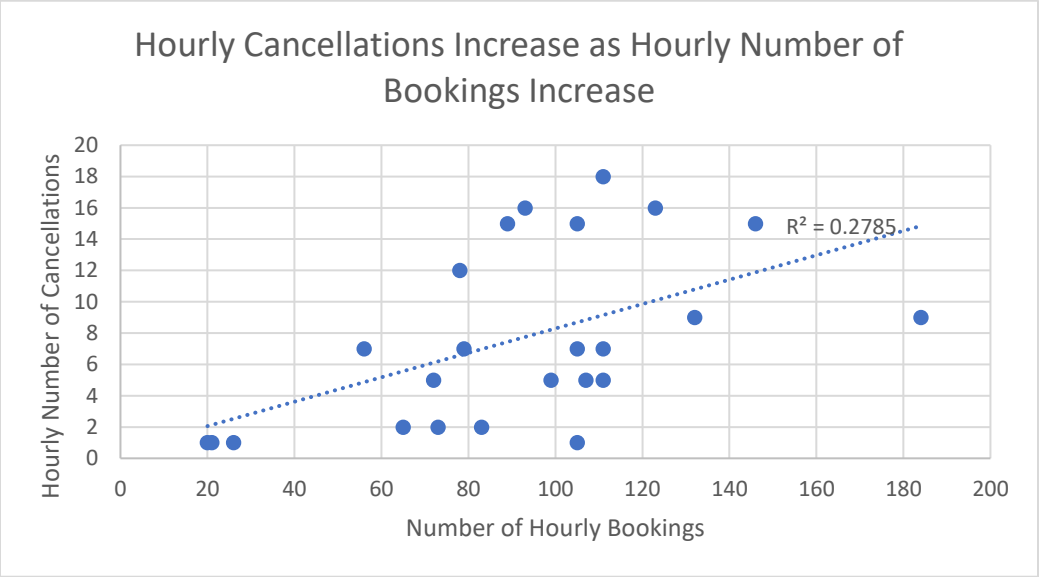
For online bookings, 844 customers booked a trip using this method and 113 of them cancelled, resulting in a 13% cancellation rate for online booking. For mobile bookings, 118 customers booked a trip using this method and 18 of them cancelled, resulting in a 15% cancellation rate for online booking. My suggestion for IndoCabs would be to eliminate mobile booking because only a small number of customers use this channel and it has a relatively large cancellation rate. I also suggest that IndoCabs keep online booking because even though there is still a large cancellation rate of 13%, a significant amount of customers, almost half of the 2,194 customers in the data set, use this channel. If IndoCabs were to take this option away, it may be too inconveniencing for many of their customers to switch to a different channel.

Cancellations by Weekday



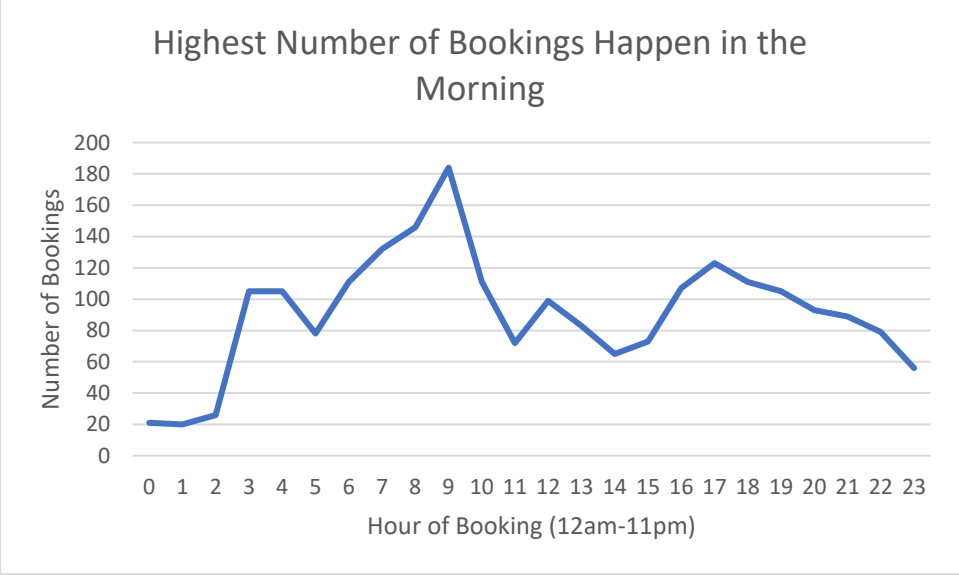
The proportions of cancellations are highest on Monday, but are also relatively high on Tuesday, Friday, and Sunday. I am not surprised that cancellations are highest on Monday because it is generally the busiest day for most people, especially in cities where the cancellations are the highest. I am surprised that Sunday is also one of the leading days of cancellations especially since Saturday's cancellation rate is the lowest of the whole week. I would have thought that Saturday's cancellation rate would have been higher than Sunday because people generally go more places on Saturday (unless there is a cultural difference among Americans and Indians). Thus, there might a problem among drivers on Sundays to fulfill trip requests.

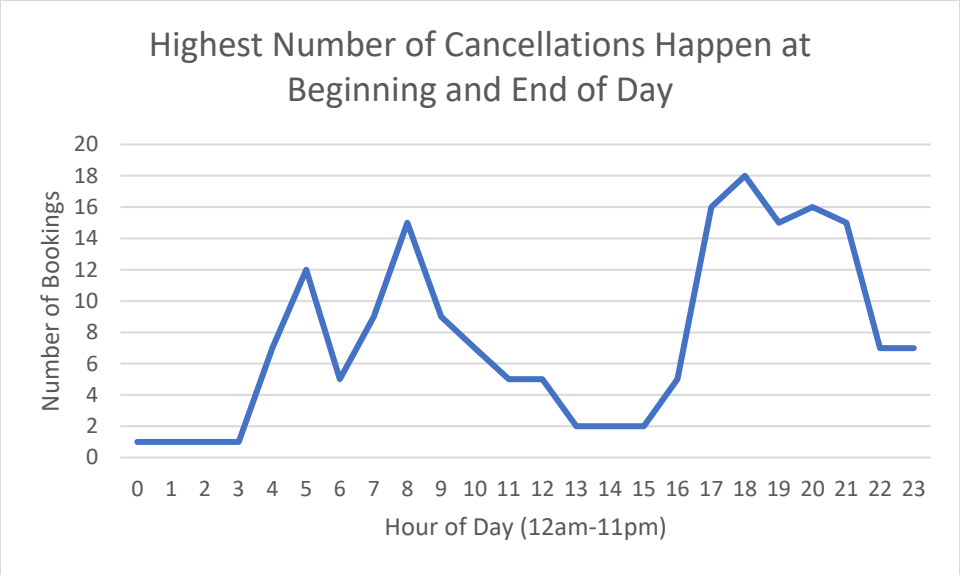
The Relationship between Booking Windows, Cancellations, and Trip Timing

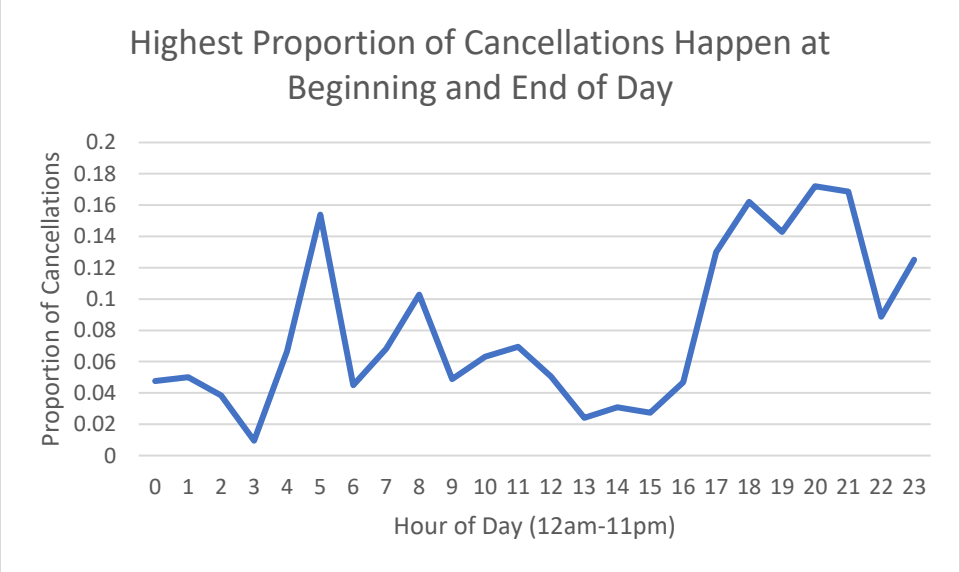


The R-squared value for the two variables, hourly number of cancellations and hourly number of bookings, is 0.2785. The correlation between the two variables is positive 0.5277. Thus, there is a moderately strong correlation of the number of bookings and cancellations at each hour. As, bookings increase, cancellations increase.

Number of Bookings and Hour of Day

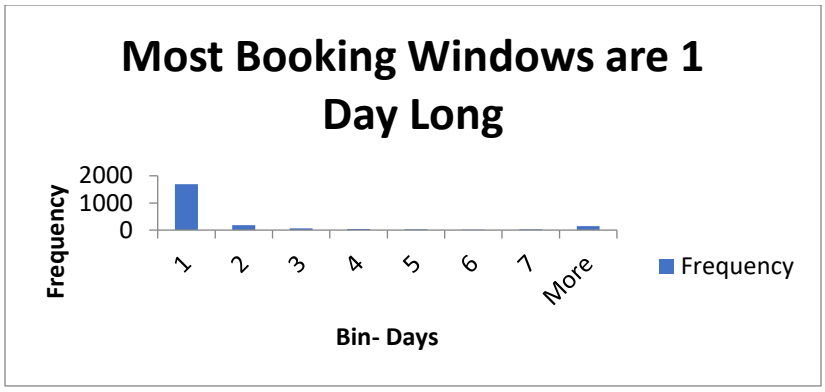


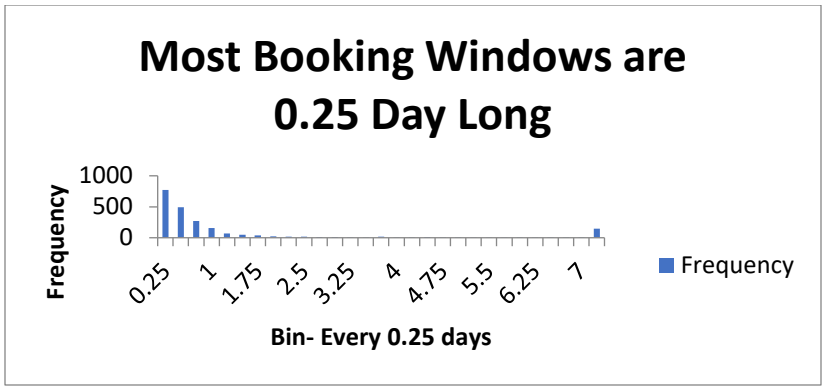




The graph that is most informative of the above three is the Proportion of Cancellations Per Hour of Day graph because this is considering all three variables; hour of day, number of bookings at each hour, and number of cancellations at each hour. This graph identifies that at 5 am and then from 5 pm to 9 pm there is a significant amount of cancellations at these times. It is also important to look at the Number of Bookings Per Hour of Day graph because this shows that the most popular hours for booking a cab is not the same as the hours with the most cancellations. Therefore, the problem of cancellations per hour does not really lie in the hours with the most amount of bookings but lies in two certain time frames: 5 am to 6 am and 5 pm to 9 pm.

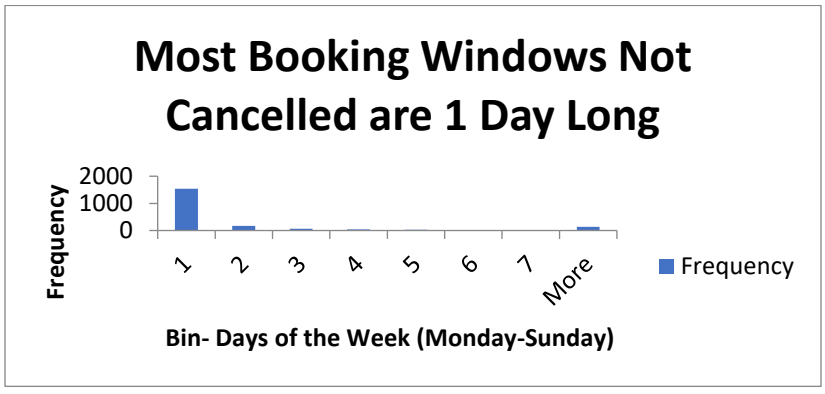
Most Frequent Booking Window is One Day or Less

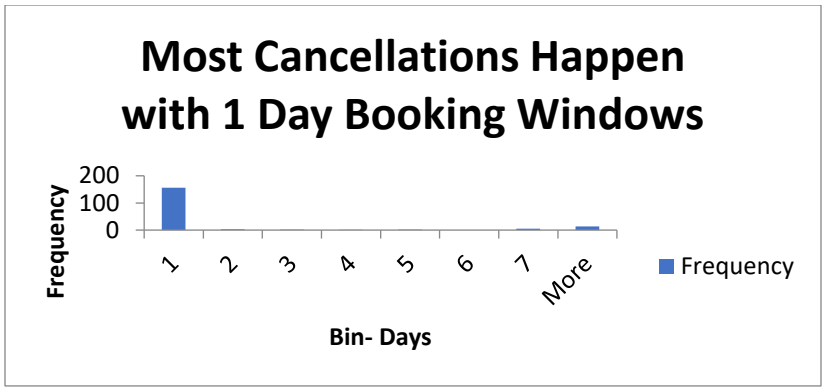


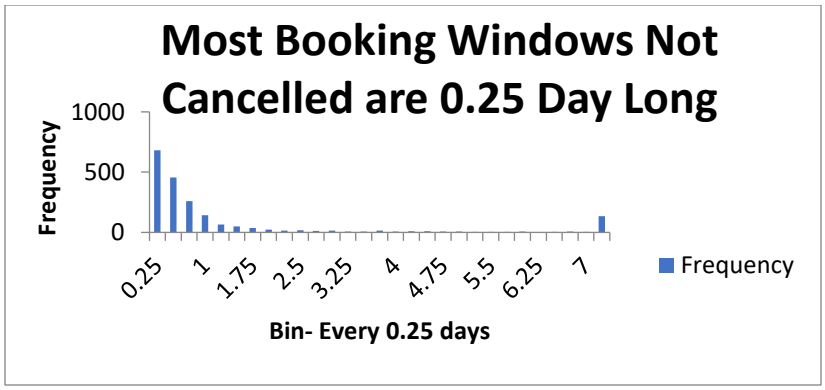


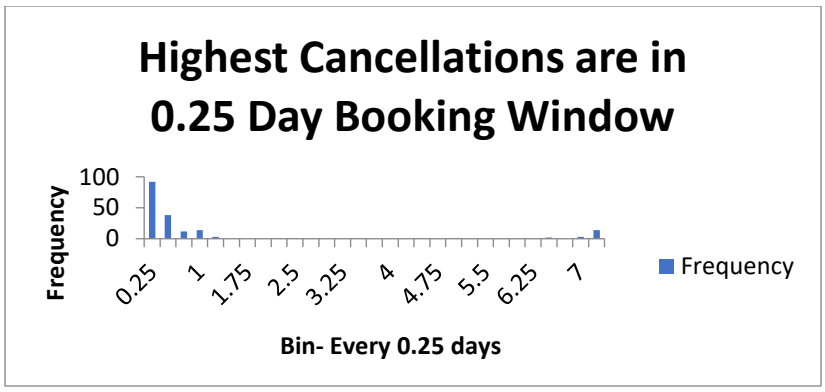
The more informative of the two histograms is the one with bins for every 0.25 days or 6 hours because most of the trip bookings happen within 1 day before the trip is set to start so it is more valuable to see that most trips are being booked by customers 0-6 hours before the start of their trip.

Booking Windows that are Less Than a Day Have Most Frequent Cancellations









The graph that is most important to look at is the histogram “Booking Windows with Cancellations” that shows cancellations by every 6 hours or 0.25 days of the week. This histogram reveals that if customers booked their trip 0-6 hours before their trip was set to start, they were more likely to have their trip cancelled. Those that booked their trip closer to a day in advance were significantly less likely to be cancelled. Customers that booked their trip more than a day in advance, were hardly cancelled on but the total amount of customers booking this early in advance is also significantly less.

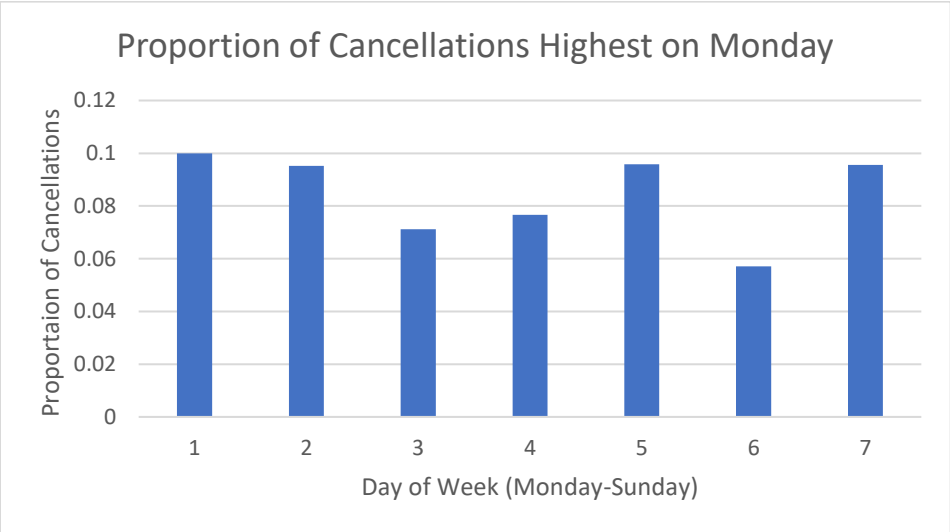
Recommendations and Conclusion

In conclusion, there are numerous patterns in the data that highlight where most of the cancellations are for IndoCabs. The median trip duration is 1.39 hours. The most cancellations occur for travel type 2, point to point travel in a city. Most bookings and cancellations occur with

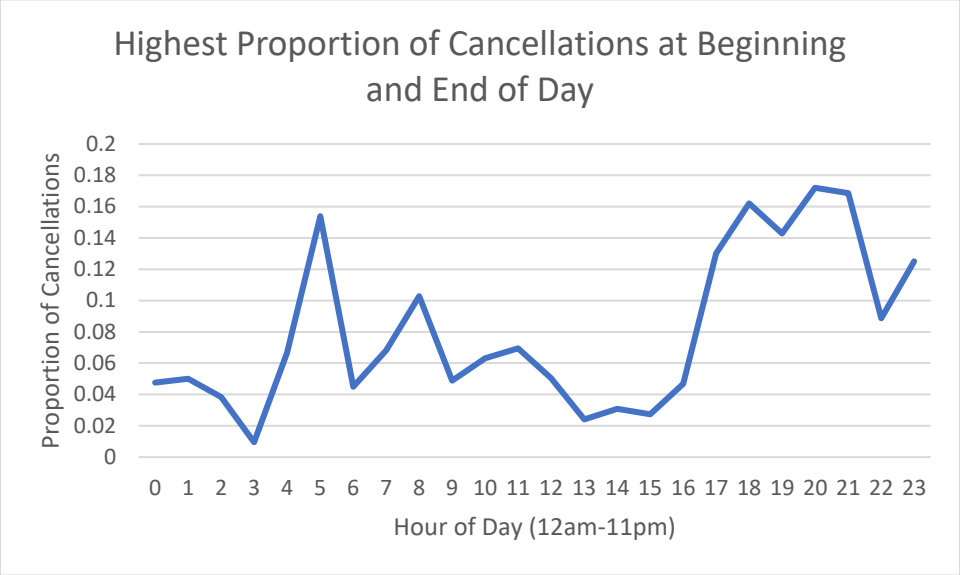
short booking windows, less than one day and especially between 0-6 hours before the trip start time. The highest proportion of cancellations happen on Mondays, but every other day except for Saturday are relatively high as well. The highest proportions of cancellations are at 5 am to 6 am and 5 pm to 9 pm. The highest cancellation rate for booking channels is mobile booking at 15% of the 118 people who used it.

To limit cancellations in the future, IndoCabs should enact a few different strategies. IndoCabs should implement surge pricing Monday through Sunday, excluding Saturday, from 5 am to 6 am and from 5 pm to 9 pm. This will make it more lucrative for drivers to work during this time, so the supply for drivers will increase and some customers may chose to wait a few minutes for surge pricing to go down, so demand will level out and not be so steep during small periods of time. A higher price for trips may also help the problem of vendors' schedules conflicting with IndoCab bookings. If drivers are receiving a higher rate, this may exceed what they are being paid per trip through the local vendor that they work for, so they may choose to take the IndoCabs trip rather than cancel it. Another way to address the conflict of vendors' schedules is to put consequences in place for drivers who cancel IndoCab trips in order to fulfill a trip through their vendor. If a vendor has drivers that are consistently cancelling on IndoCab trips, then IndoCabs should end their partnership with them. This will encourage vendors to work more closely with IndoCabs to find a schedule that will work well for both parties. Another strategy that IndoCabs should test is requiring drivers to allow more time for trips than what may seem necessary. The median trip duration is 1.39 hours, thus trips from one point to another in a city may be taking longer than what drivers expect, especially since most cancellations are happening during rush hour when most people in a city are traveling. IndoCabs should start out by requiring double the time it would normally take to complete a trip then see if this is causing too much time in between trips or not enough and go from there. Lastly, IndoCabs should terminate their mobile booking channel or at the very least figure out how to improve this channel because only 118 customers of the 2,194 customers sampled use mobile booking and of those customers, 15% cancelled their trip. Through implementation of these strategies, IndoCabs will be able to greatly reduce their cancellations.

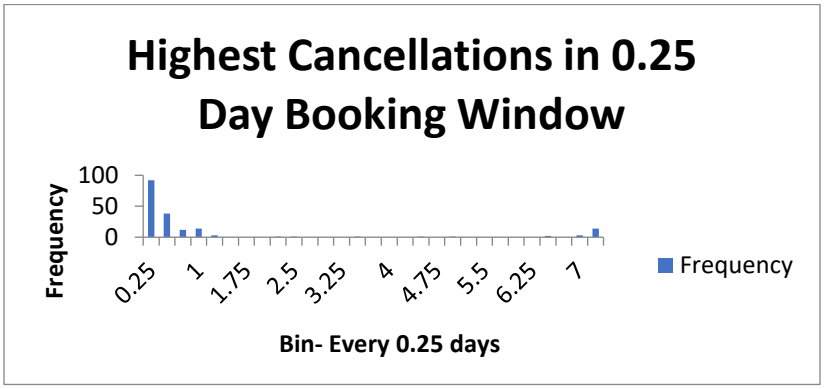
Elevator Charts



This graph gives a good sense for which days have the highest proportion of cancellations which was an important factoring in my recommendation for why and when to implement surge pricing.



This graph shows that 5 am to 6 am and 5 pm to 6 pm have the highest proportions of cancellations. These time periods were also an important factor in determining my recommendation of implementing surge pricing during the times with the highest proportions of cancellations.



This histogram highlights that a significant number of customers' trips are being cancelled when they are scheduled shortly before they are set to start. This data helped me to determine that there is a most likely a scheduling problem with vendors and that better scheduling and/or consequences for vendors need to be implemented.

Notes on Data Preparation

To prepare the data, I first removed any duplicate entries and then I removed any erroneous entries that had happened before January 1, 2013. Then I used formulas, pivot tables, histograms, and charts to find information and arrange data in a meaningful way that helped me

determine where the cancellations were and how to possibly prevent them in the future. There are few concerns that I have with the data. There were 184 erroneous entries in my sample that were caused by a trip end date not being recorded. Therefore, some of the conclusions I have made may be slightly off due to not having trip durations for these entries. Another concern that I have with data is that some of the Travel Type 1 trips (long distance trips) were not all recorded properly as I have one trip that was recorded as having spanned 10 days and a few more entries that spanned 4 or more days which seems unusually long even for this travel type. Thus, these entries may have also negatively skewed some of my data.