METROCAR

Funnel Analysis

Daniele D'Epiro

22/10/2023

Table of Content

Executive Summary

Metrocar's funnel analysis reveals a significant drop-off from ride acceptance to completion, primarily due to high demand during peak hours (8-10 AM and 4-8 PM). To address this issue, it's highly recommended increasing the number of active drivers during rush hours, promoting ride-sharing options to encourage user participation and/or implementing price surging during peak hours to attract drivers. The analysis also reveals that many users have not yet completed their first ride. Therefore, offering a discount voucher to users after they complete their first ride will certainly be beneficial.

Context

Metrocar is a ride-sharing platform that operates as a user-friendly intermediary connecting riders and drivers through its mobile application. Its customer funnel includes the following stages:

1. **App Download**: A user downloads the app from the App Store or Google Play Store.
2. **Signup**: The user creates an account in the app, including their name, email, phone number, and payment information.
3. **Request Ride**: The user opens the app and requests a ride by entering their pickup location, destination, and ride capacity (2 to 6 riders).
4. **Driver Acceptance**: A nearby driver receives the ride request and accepts the ride.
5. **Ride**: The driver arrives at the pickup location, and the user gets in the car and rides to their destination.
6. **Payment**: After the ride, the user is charged automatically through the app, and a receipt is sent to their email.
7. **Review**: The user is prompted to rate their driver and leave a review of their ride experience.

The Funnel Analysis

In any user journey, it is completely normal to anticipate a drop-off at each stage. Not all individuals who download the app will proceed to sign up, and not all signed-up users will actively utilize the service. This attrition, known as drop-off, is a typical aspect of the user acquisition and retention process. To address this, a comprehensive funnel analysis is crucial.

Here are the main reasons:

Quantifying Drop-off: The primary objective of a funnel analysis is to quantify the drop-off that occurs at each stage of the user journey. By tracking the conversion rates from one stage to the next, we can identify precisely where and how users are exiting the funnel. This data-driven approach enables us to understand the bottlenecks and identify areas where improvements are needed.

Identifying Causes: More than just quantifying the drop-off, the funnel analysis helps us uncover the underlying causes of attrition. We can identify patterns and trends that shed light on user behaviour and the pain points experienced at each stage. These insights are invaluable for designing interventions that address the root causes of drop-off.

Taking the Right Actions: Armed with an understanding of the causes of drop-off, we can develop data-informed strategies to mitigate the attrition. These strategies may involve optimizing the user experience, enhancing app features, or offering targeted incentives to users at critical stages of the funnel.

Focused Energy: Not all stages of the funnel are created equal. The funnel analysis helps us discern where it makes the most sense to focus our efforts. By concentrating on the stages with the highest drop-off or those most critical to the business objectives, we can allocate resources effectively to improve the overall user journey.

By conducting a thorough funnel analysis, we can gain valuable insights into user behaviour and systematically enhance the customer experience. This data-driven approach empowers us to make informed decisions that drive user engagement, conversion, and business growth.

Business Questions

The analysis directly addressed the following business questions:

1. What steps of the funnel should we research and improve? Are there any specific drop-off points preventing users from completing their first ride?
2. Metrocar currently supports 3 different platforms: iOS, Android, and web. To recommend where to focus our marketing budget for the upcoming year, what insights can we make based on the platform?
3. What age groups perform best at each stage of our funnel? Which age group(s) likely contain our target customers?
4. Surge pricing is the practice of increasing the price of goods or services when there is the greatest demand for them. If we want to adopt a price-surging strategy, what does the distribution of ride requests look like throughout the day?
5. What part of our funnel has the lowest conversion rate? What can we do to improve this part of the funnel?

The Dataset

For the analysis, has been used the company dataset comprises the following 5 tables:

1. ***App Downloads****:* a table with unique download key, download timestamp (exact time and date of the event) and platform (web, Android and iOS) and the age range of the customers.
2. ***Sign-Ups****:* after the sign-up, the user gets its unique ID. The timestamp and the age range is also recorded.
3. ***Rides***: here all rides are uniquely registered with the timestamp of the main stages of the service: user request, ride acceptance, pick-up, drop-off and eventually a cancellation. The table also includes user and driver IDs and the pick-up and drop-off point.

4. ***Transactions***: contains information about financial transactions based on completed rides ride ID, amount, payment status and timestamp
5. ***Reviews***: here there are information about driver reviews once rides are completed, with a unique review ID, ride, user and driver ID. Rating from 1 to 5 and the user feedback.

Query the Database

To explore and retrieve data from the database, PostgreSQL was used. When constructing the user funnel, only the first ride for each user was taken into account. The first step of the funnel, which tracks app downloads, was calculated by counting the unique download key. The second step counts all user IDs, unique values generated when a user signs up. All the other steps in the funnel refer to the count of distinct user IDs at the moment of the recorded timestamp for each event in the funnel (ride request, ride accepted, ride completed, payment, and review). Conversion rates and drop-offs between each stage were calculated using the lag() function.

funnel_step	step_name	user_count	drop_off	drop_off_rate_percentage	conversion_rate_percentage
1	app downloads	23608			
2	signups	17623	23608	25.00	75.00
3	ride requests	12406	17623	30.00	70.00
4	ride accepted	12278	12406	1.00	99.00
5	ride completed	6233	12278	49.00	51.00
6	payment completed	6233	6233	0.00	100.00
7	reviews	4348	6233	30.00	70.00

A ride funnel was also developed to gain a more in-depth understanding of the entire customer journey on a broader scale and to identify any possible reasons for the significant drop-off between ride acceptance and completion. By constructing a funnel at the ride granularity, it became evident that many rides do not get accepted, and there is still a 10% drop-off between ride acceptance and ride completion.

funnel_step	step_name	ride_count	drop_off	drop_off_rate_percentage	conversion_rate_percentage
1	ride requests	385477			
2	ride accepted	248379	385477	36.00	64.00
3	ride pickup	223652	248379	10.00	90.00
4	ride dropoff	223652	223652	0.00	100.00
5	ride payment completed	212628	223652	5.00	95.00
6	ride reviews	156211	212628	27.00	73.00

Creating the Dashboard

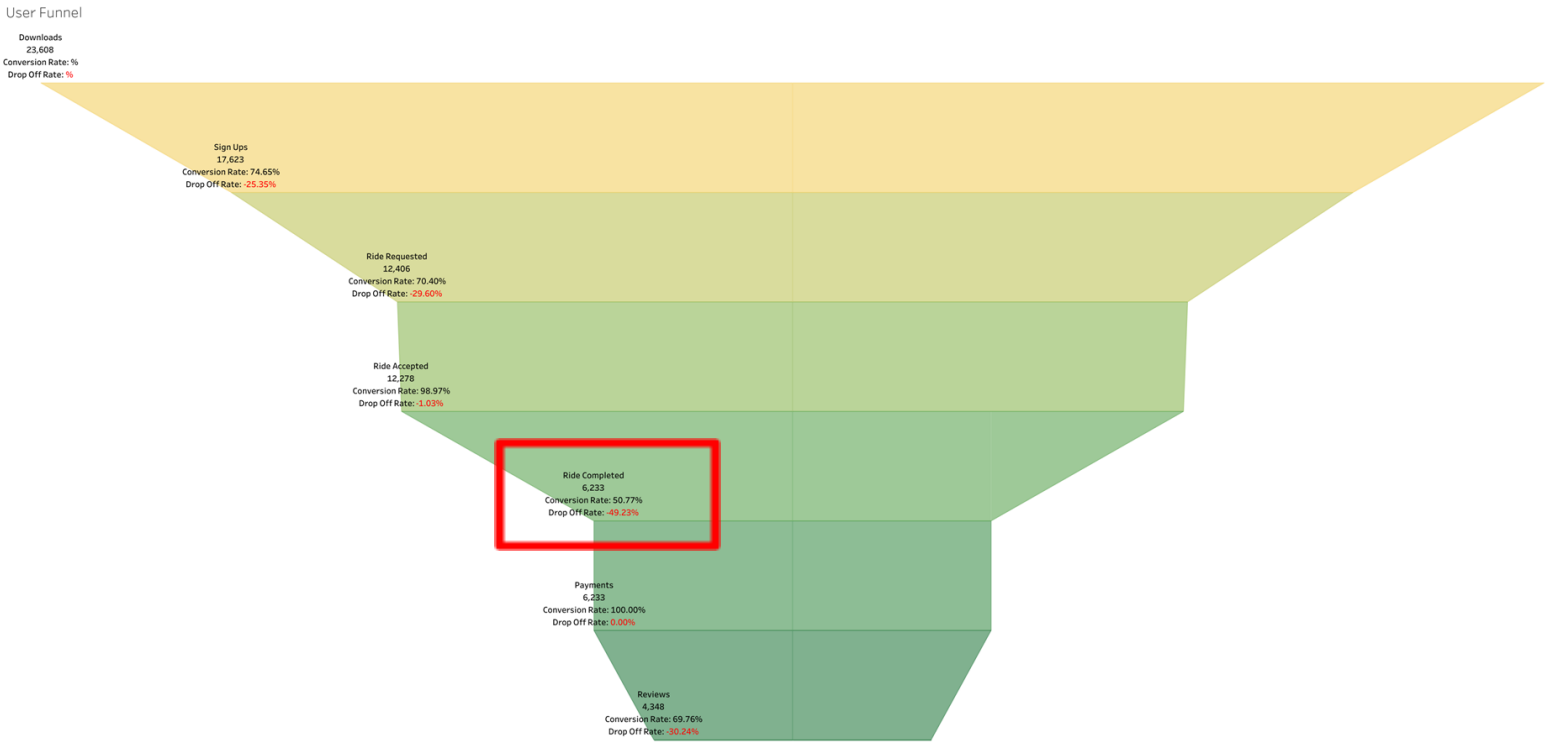
At this point, the analysis shifted to identifying all possible reasons for the drop-off. This process began by extracting all relevant data needed to plot both funnels and to enable stakeholders to clearly visualize the results and explore them further by filtering by platform, age group, and date. Tableau was used for this purpose. The relevant data were extracted, and the queries used can be consulted here. To suggest accurate data-driven solutions, I also calculated the difference in time between each stage of the funnel for each ride.

Results

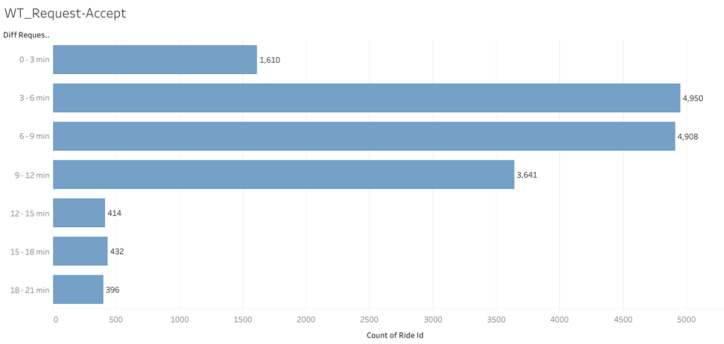
Answer to Business Questions

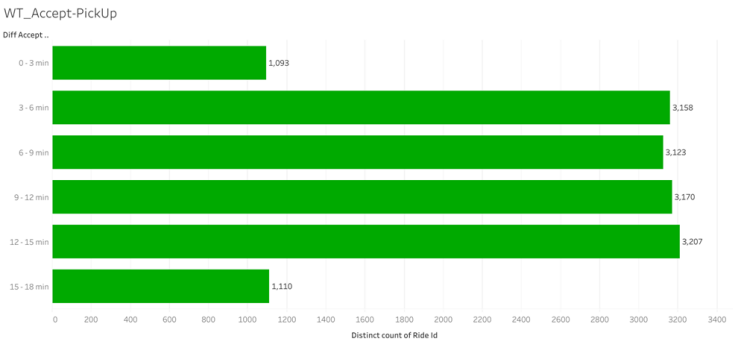
1. **What steps of the funnel should we research and improve? Are there any specific drop-off points preventing users from completing their first ride?**

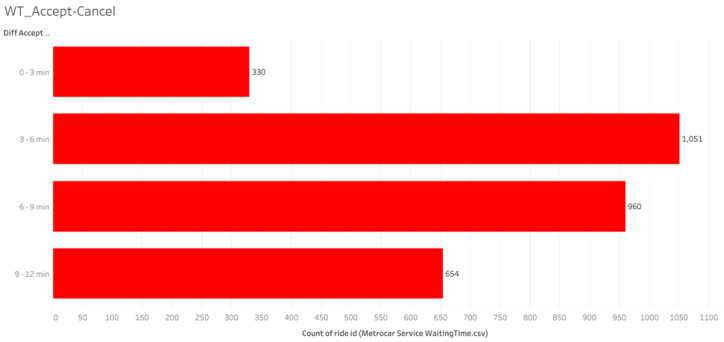
The main drop in the funnel is between the stage of Ride Accepted and Completed with a drop of almost 50%, with only 6,233 rides completed over 12,278 accepted.

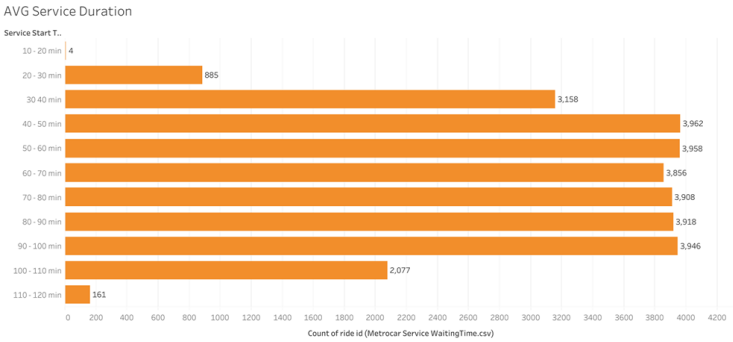


By looking at all rides (not only the first one), it's possible to see that most of the rides get accepted between 3 and 12 minutes, since the ride is accepted the majority of the user get picked up in 3 to 15 minutes. The dashboard below allow you to dig-down into the waiting time at each step of the service.

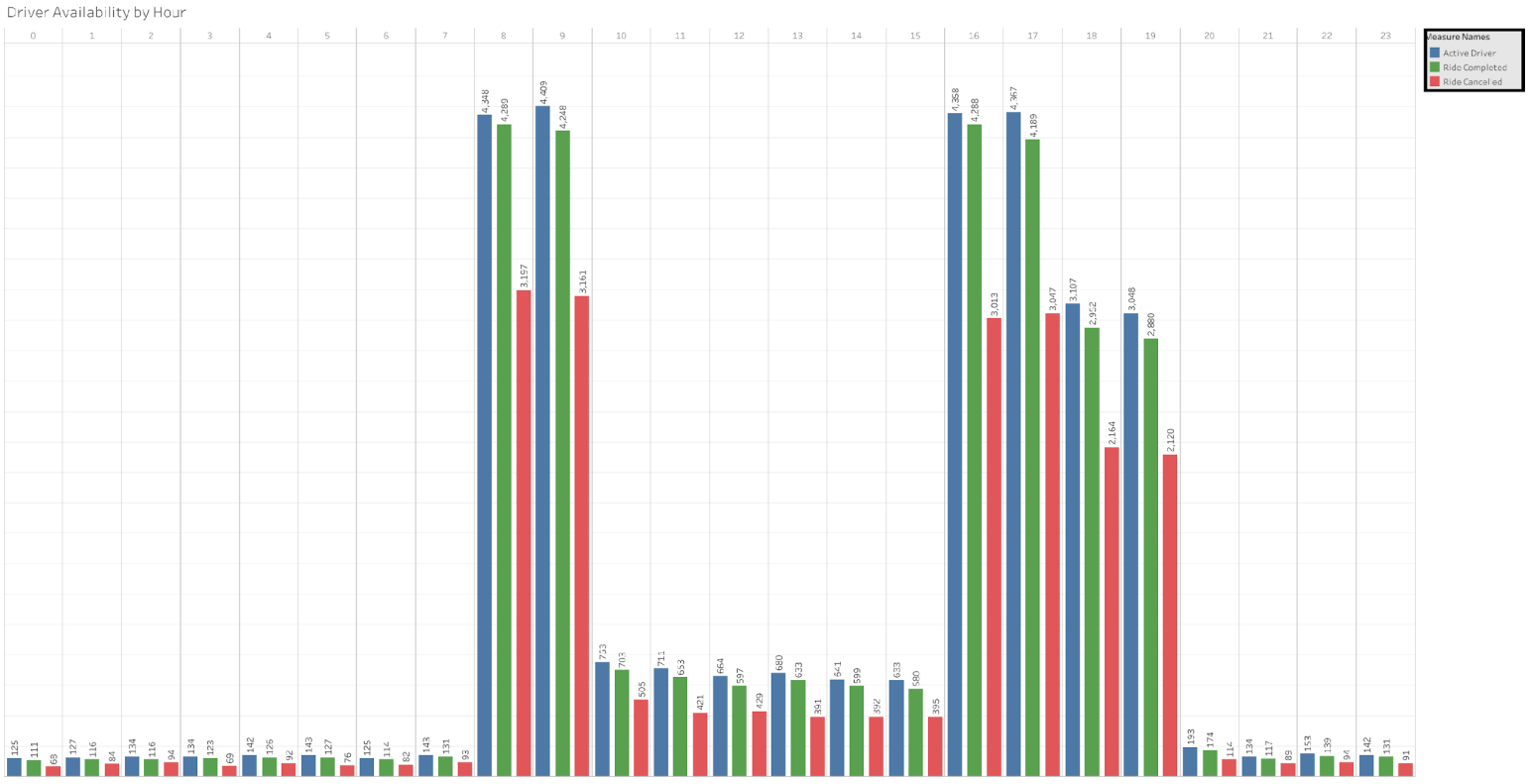




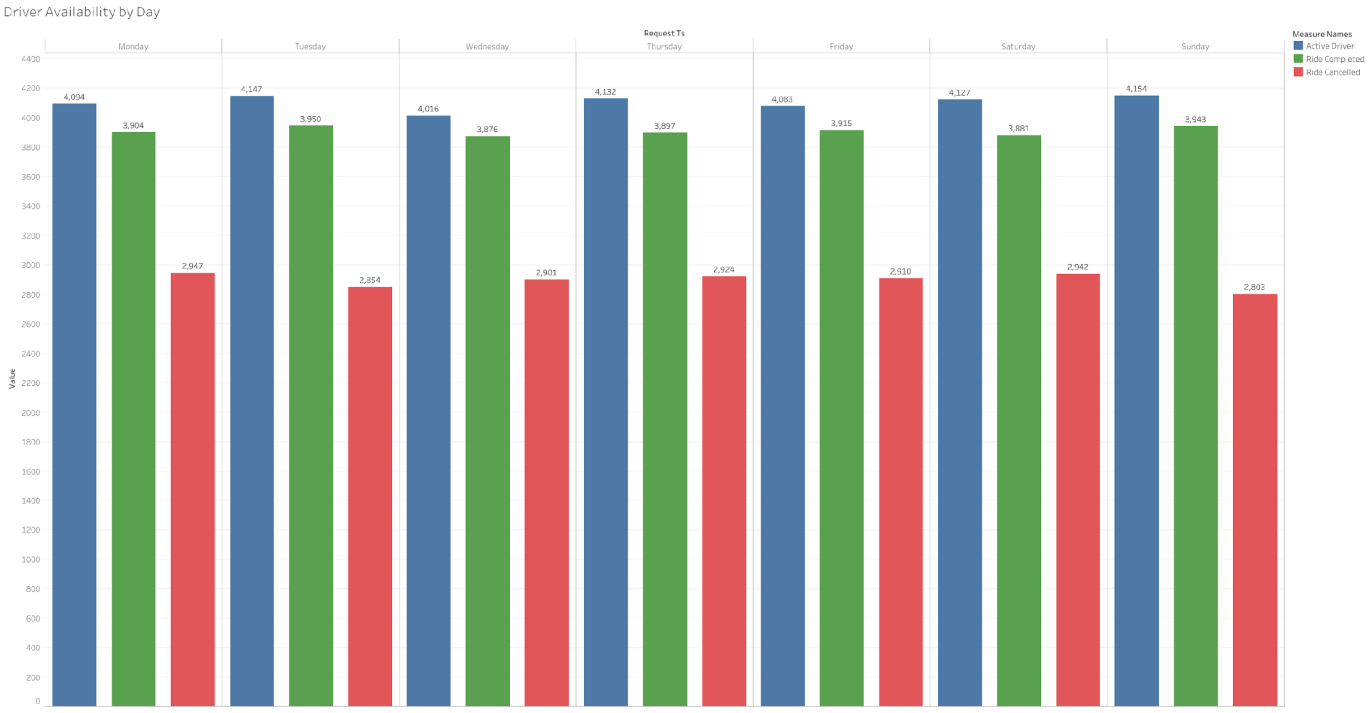




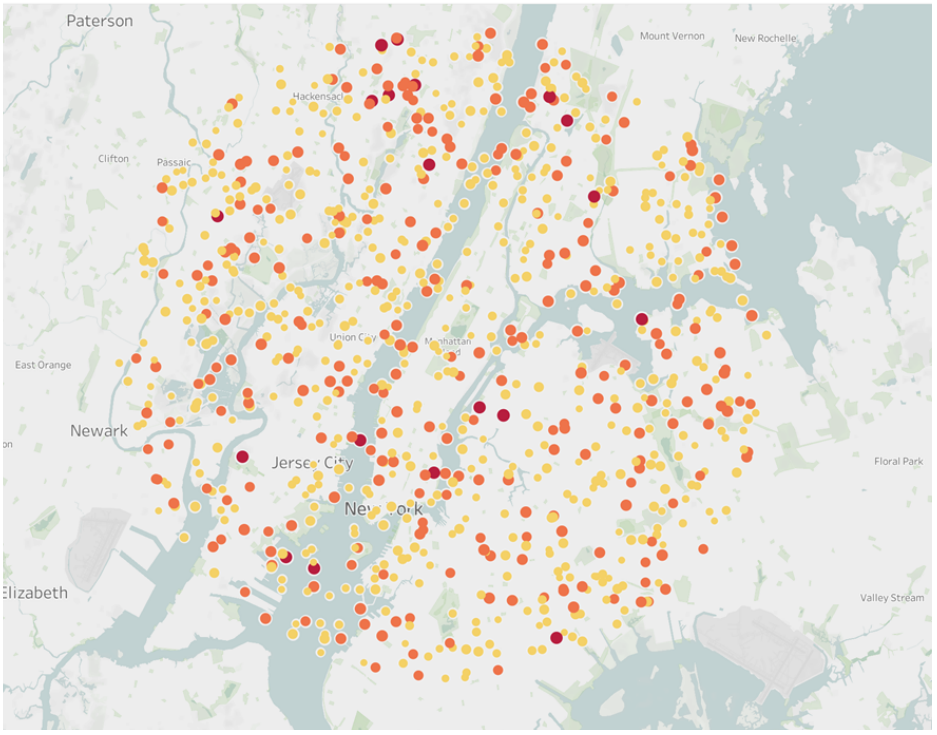
It’s clear that the waiting time has a huge impact on the cancellation rate. By looking at the 8am and 10am hours range, it’s possible to see that almost 3,200 rides get rejected each hour.



During the all week we can see a similar trend, on average, active drivers, ride completed and cancelled have a high degree of similarity.



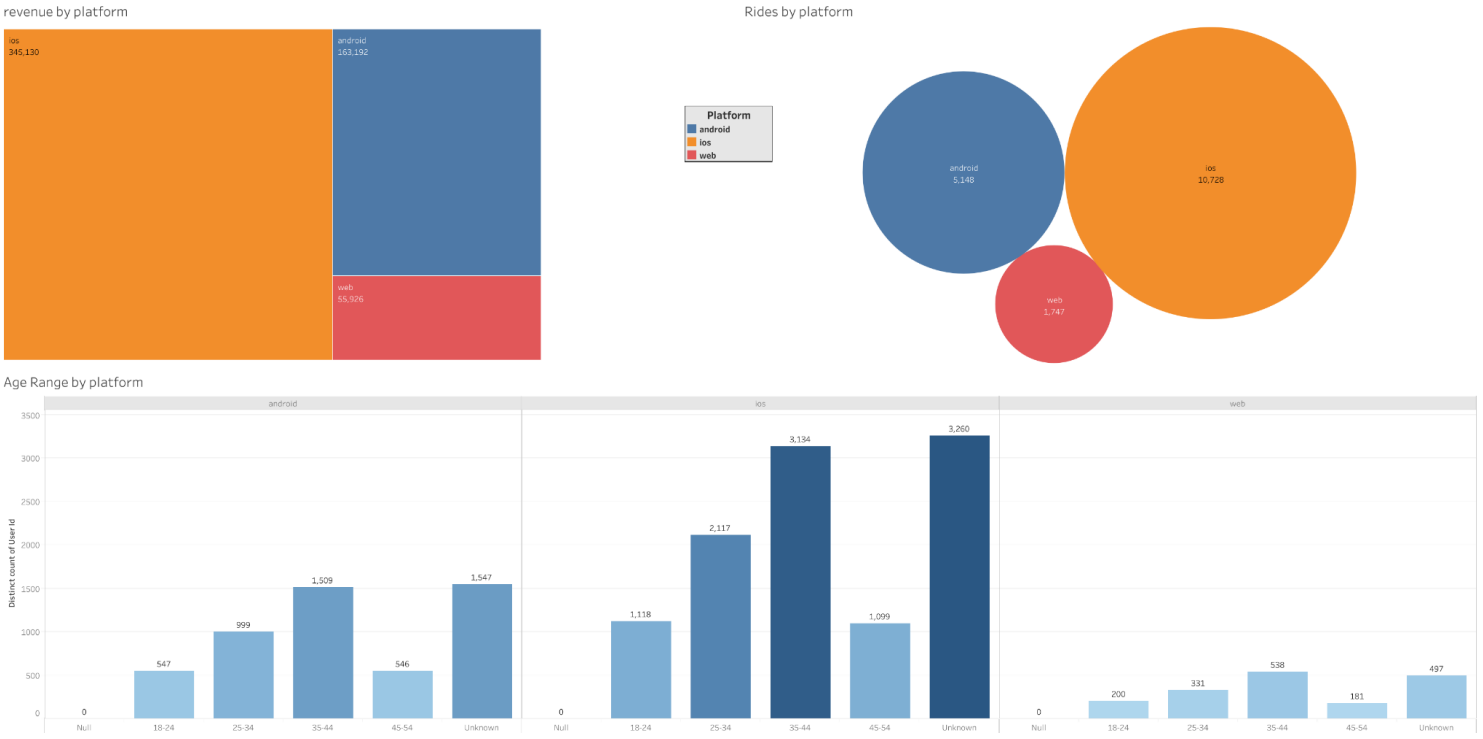
I also went further to understand better if there was any area with an outstanding number of ride cancelled. To achieve that, the coordinate recorded in the pickup location field has been divided in latitude and longitude. Doing so, I constructed the map below that does not show any area with an outstanding number of ride cancelled. In red, the location where ride cancelled are closer to 20.



The only warning is that some pickup point recorded are located in the middle of the sea and for this reason it's worth consulting the drivers to understand if this situation is happening quite often, slowing down the service. To conclude, The best solution is to propose a ride-share option where multiple users can share the same ride. This way will definitely increase the efficiency of the service by also making rides less expensive.

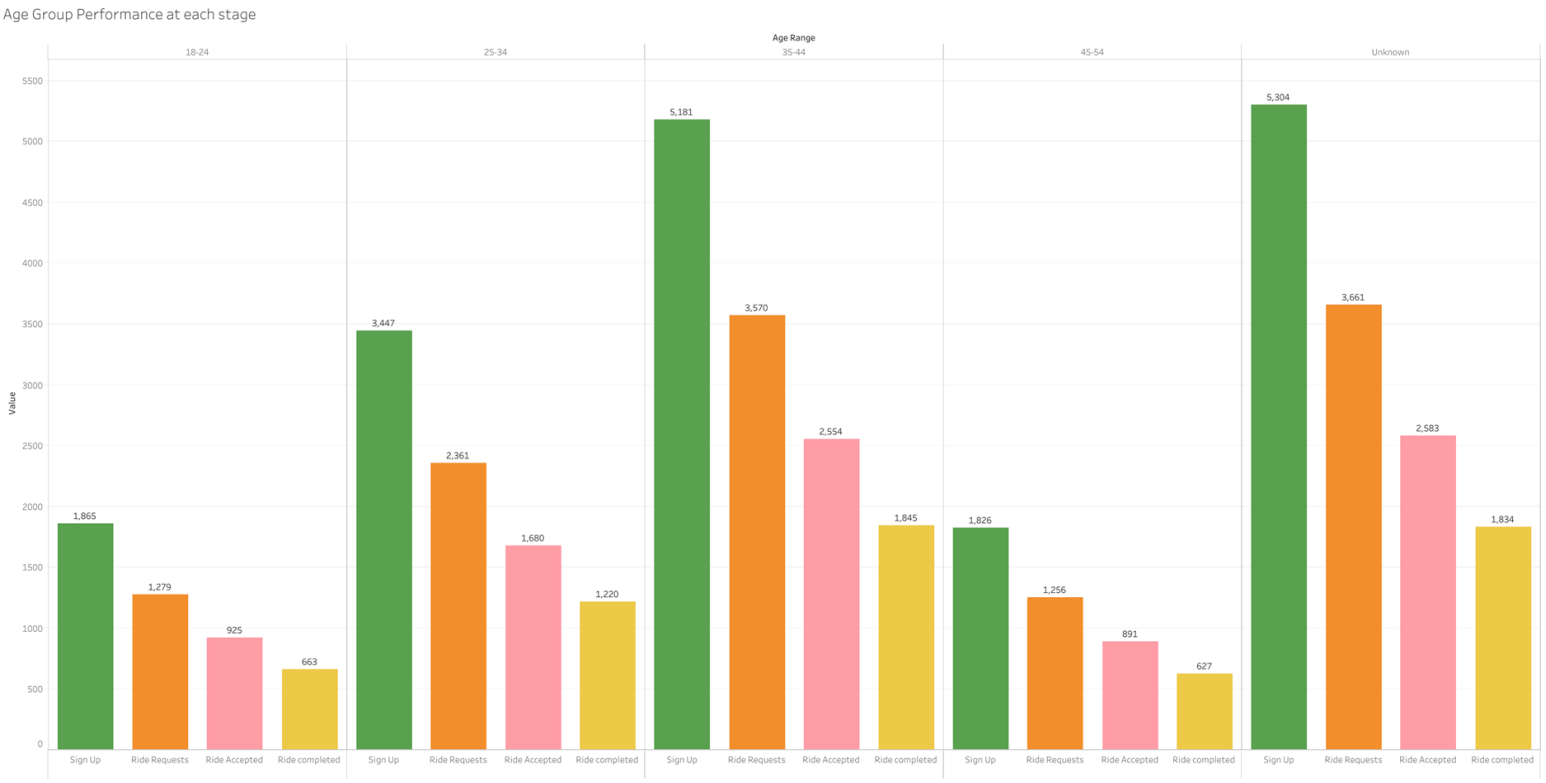
2. Metrocar currently supports 3 different platforms: IOS, Android, and web. To recommend where to focus our marketing budget for the upcoming year, what insights can we make based on the platform?

The dashboard below suggest allocating most of the marketing budget in the iOS platform, which for today is the one that is working the best. The web channel budget can be reduced, it is not only underperforming, but also the app makes the service experience more user-friendly. I would suggest re-using the reduction of the marketing budget for the web channel for the Android platform because around 70% of user worldwide own an Android powered device.(Statista)



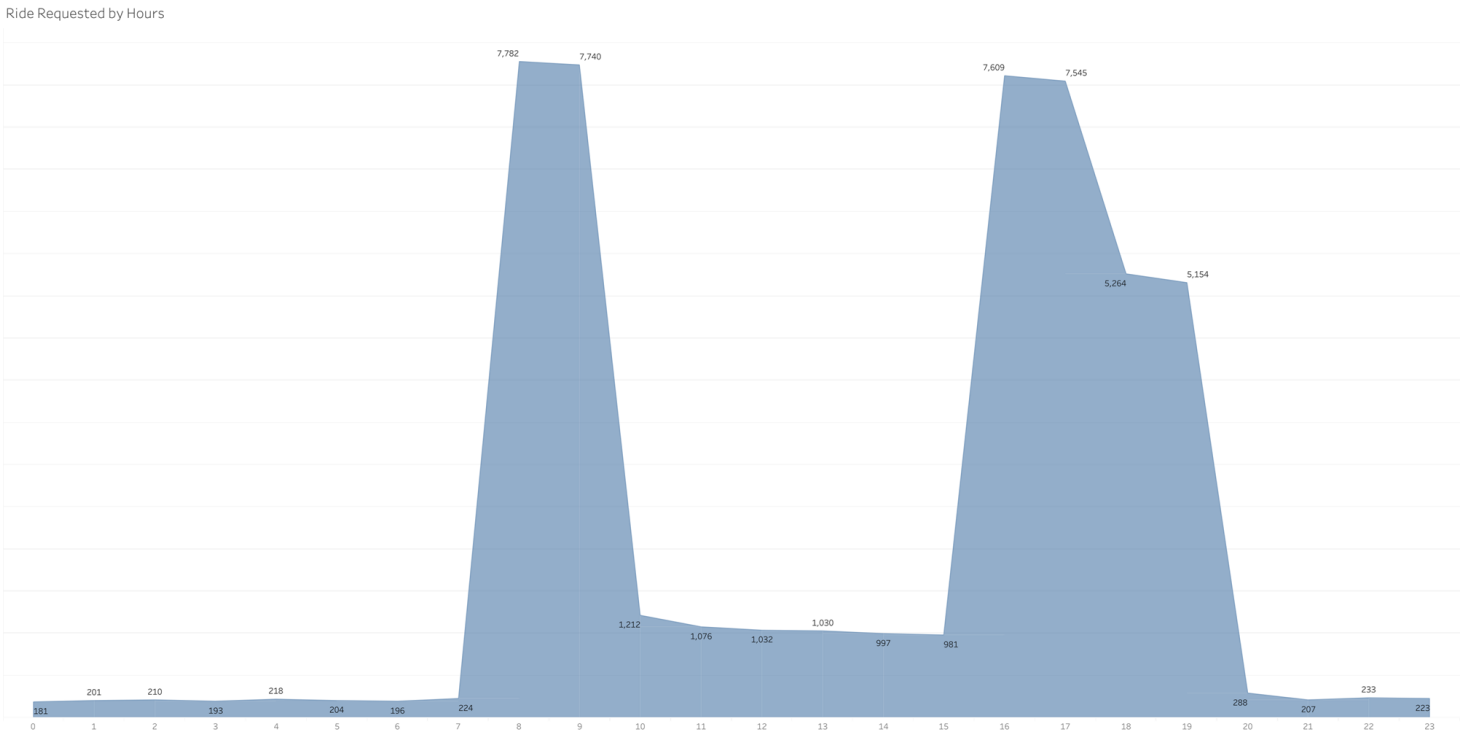
I suggest allocating the 55% of the marketing budget to the iOS platform, 40% to Android to unleash its potential and the remaining 5% to the Web.

3. What age groups perform best at each stage of our funnel? Which age group(s) likely contain our target customers?



The age between 35 and 44 years old perform the best with around 5200 users, but we need also take into account that 5300 users did not share their age, so the real group age can potentially be a different one. Obtaining a better understanding of users in the 'unknown' category could definitely lead to a better answer, however most of the market research published about ride-sharing showcase that 45% of users between 18 and 29 years old uses the service and the 36% belong to the 30-49 age group according to many market research company. With this in mind, I suggest focusing the attention on the 25-34 and 35-44 age groups. These age groups could also be more open to implement a multiple user ride-share option, to decrease the cancellation rate.

4. **Surge pricing is the practice of increasing the price of goods or services when there is the greatest demand for them. If we want to adopt a price-surging strategy, what does the distribution of ride requests look like throughout the day?**



There are two peak demand periods during the day, one from 8am to 10am and the other between 4pm and 8pm. A surge pricing technique can encourage more driver to enter the market when demand is high, helping to balance supply and demand. This can result in faster service and better availability for customers. On the other side, price surging can generate negative media coverage and public backlash, damaging a company's reputation. It may also lead to regulatory scrutiny in some markets. It can also lead to customer churn during surge periods, as some users may seek alternatives or choose to wait until prices return to normal. By looking at the graph above, it’s clear that most of the requests fell between 8-10 am when many professionals start work and between 4 and 8 pm when they came back home.

5. **What part of our funnel has the lowest conversion rate? What can we do to improve this part of the funnel?**

The lowest conversion rate occurs at the “ride completed” stage, which is also the point where occurs the biggest drop-off. The reasons at this point are well known, and most of the suggestions made before are valuable to reverse the trend. It’s also worth considering offering a discount after the completion of the first ride. This can encourage more users to try the service and increase the conversion rate at this stage.

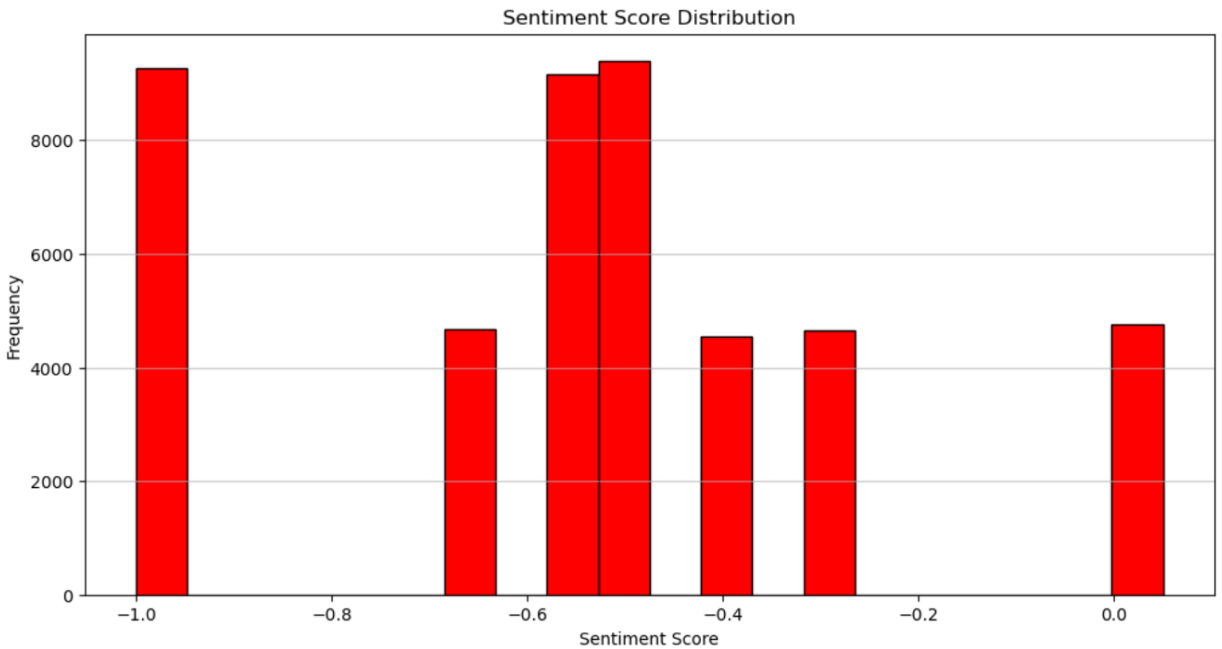
Explore the Dashboard

I would suggest that you all have a look at the dashboard available here. Here you will be able to analyse trends and patterns on your own. In particular, I have added the option to filter the result of each funnel by Age-Group, Platforms or Date.

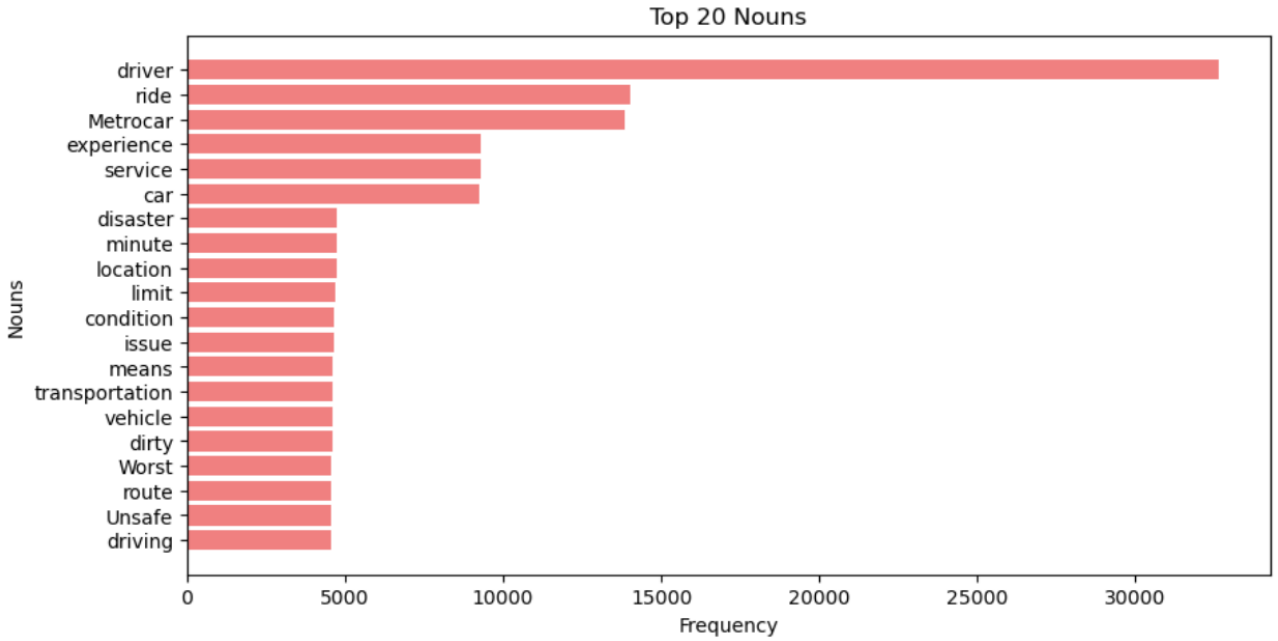
Extra

Exploring User Feedback

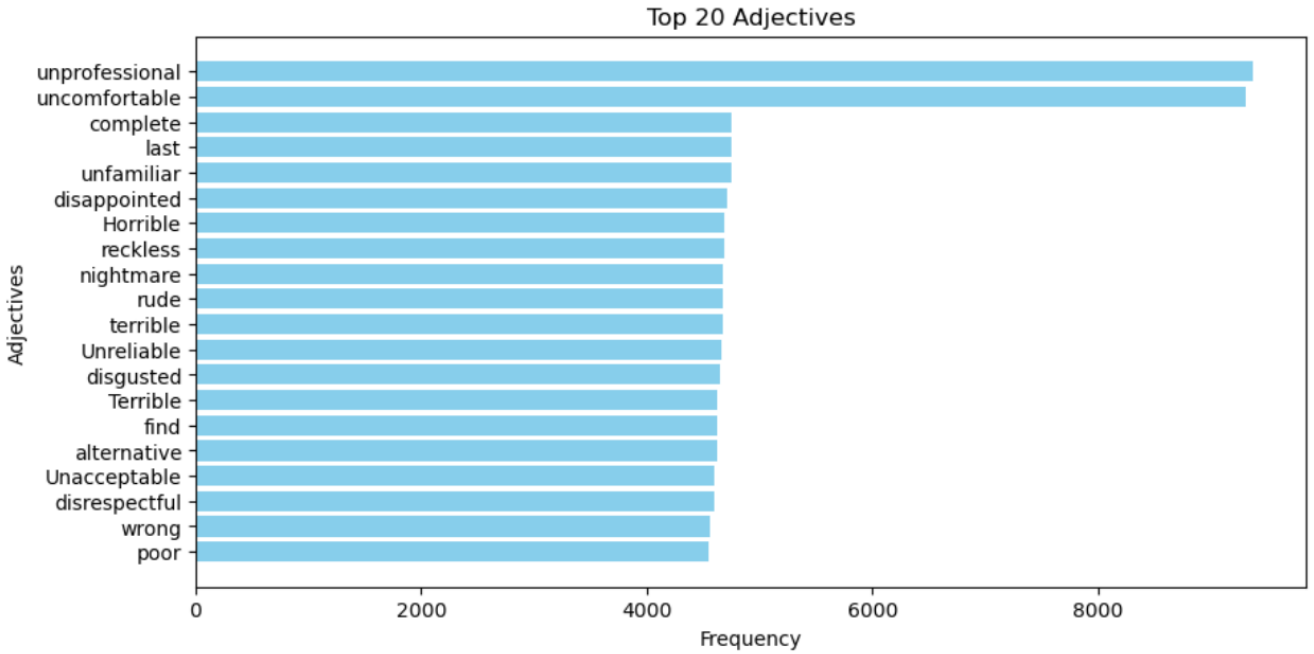
I have also performed a sentiment analysis to understand the reasons why many users have left bad reviews. For this purpose, I used the Python programming language and the following libraries: pandas, nltk, and textblob. More details about the process can be found here. First, I want to mention that due to the large number of records in the dataset, my current resources couldn't handle the analysis, which resulted in a system crash. As a result, I restricted the analysis to all the reviews with a rating of 1. I discovered that most comments are highly negative, with only a few of them scoring slightly above 0 on a scale from -1 to 1.



I also aimed to identify any nouns or adjectives in the reviews that could be related to long waiting times. The graph below shows that almost every user mentioned the noun "driver" in their review, with more than 30,000 records. This could potentially be a serious concern, as we are looking at reviews with a score of 1. Other concerning nouns mentioned include "disaster," mentioned almost 5,000 times, as well as "limit," "unsafe," and "driving." These words, when paired together, certainly don't benefit the reputation of the company.



Furthermore, the most frequently used adjectives do not inspire confidence. In fact, the majority of the adjectives in the top 20 are associated with a negative sentiment. When paired with the nouns I highlighted above, they could pose a serious threat to Metrocar's reputation.



However, before drawing any conclusions, it's essential to conduct a more in-depth sentiment analysis to see how often these adjectives and nouns are used together. With ride and driver IDs in the records, it may be relatively easy to identify who's responsible and take steps to improve the service. This could also be a suggestion for a future project to enhance Metrocar.

Additionally, it's worth noting that some words commonly found in these reviews, such as "minute," "last," and "complete," can be indicative of issues related to waiting times. Users may be referencing the time it takes for a ride to arrive or the total duration of their journey. Analysing the context in which these words appear can provide further insights into the waiting time concerns expressed by users. With this additional information, a more comprehensive sentiment analysis can be conducted to examine how these words are used in conjunction with other nouns and adjectives. This deeper analysis can help identify specific areas where improvements in waiting times are needed and allow for more targeted efforts to enhance Metrocar's service.

Conclusions

To improve the conversion rate of the currently underperforming step, I recommend considering the implementation of surge pricing and/or adding a multiple-user ride-share option. Surge pricing can increase the number of drivers available during peak hours, reducing user waiting times and the overall service duration. I also recommend introducing a ride-sharing option, allowing drivers to pick up more than one user during the same journey. This feature is offered by other competitors in the market and is on average 20% cheaper, with only an 8-minute increase in the total service time, marketed as an eco-friendly choice.(Uber) This approach aligns with the fact that the majority of Metrocar users are professionals who request rides simultaneously. Both options should be discussed with app developers to assess development costs and feasibility.

In terms of marketing budget allocation, I strongly recommend prioritizing the iOS platform. Allocating 55% of the marketing resources to iOS would be advisable. Simultaneously, I recommend reducing the allocation for the web platform to just 5%. The cost savings from this reduction can then be strategically reinvested to strengthen Metrocar's presence on the Android platform, with a 40% allocation. This strategy aligns with the widespread use of Android-powered devices and offers substantial growth potential in this market segment. It's essential to emphasize that Metrocar's primary target audience falls within the 35 to 44 age group, followed closely by the 25 to 34 age group.

Appendix

1. SQL Queries
2. Tableau Workbook
3. Sentiment Analysis Notebook
4. Statista Research on phone OS
5. Uber Ride-share Option