MetroCar Funnel Analysis

By Deirdre Hopcraft
Data Analyst
24/11/23

Introduction

This step by step analysis aims to help identify areas of improvement and optimisation through visualization. This report will focus on the information about the customer journey that provides the most insight into the steps that need to be implemented to positively impact customer retention.

Objective

Funnel Analysis is a great way to visualize data in an understandable way. Your User Funnel originally had 7 steps but rides requested/ rides accepted are essentially the same and rides completed/ rides paid are exactly the same so let's call it 5 . We lose approximately 25% at every step save the last because most riders leave reviews

Key Findings

Your Funnel only loses 5% before the review stage, I believe this is very important. Firstly I suspect that more people are likely to leave a negative review than a positive review and secondly because the reviews should state clearly what the customer's expectation is for satisfaction.

Because you do not have a review rating system I had Claude.AI read a sample of 250 reviews which resulted in 141 of 250 being negative. That is 56.4% of the sample reviews. With this being the step that your customers utilize the most I felt I needed to explore the reviews step further.

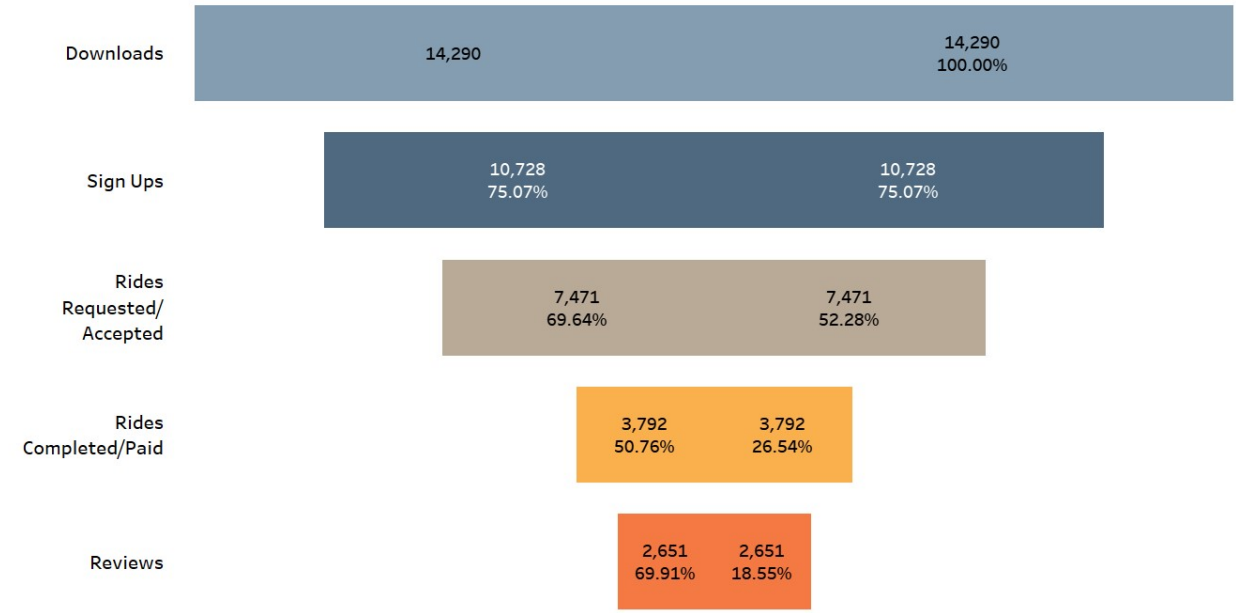
Current State of Customer Retention

I analyzed the numbers and your company retains customers for approximately 2 months averaging 30 rides over that time. I suspect that we are losing customers maybe even after the first time they leave a negative review but it would take a great deal of time to verify that due to the lack of categorization in your data. A rating system would be extremely helpful for analysis but because I felt the need to examine the crucial information in the reviews for my report I asked ChatGPT to categorize a sample of 500 reviews for me. This isn't ideal for the risk of human error and the fact that your company can't access this very important information easily. After separating the negative reviews from the positive I found the reasons within those negative reviews shocking, I think you'll agree.

Here are the top 10 negative reviews:
1."The driver was reckless and drove well above the speed limit."
2 "The driver took a completely wrong route, and the car broke down halfway."
3."The driver never showed up, and I had to find an alternative means of transportation."
4."The driver overcharged me and refused to rectify the issue."

5.“The driver was rude, and the car was in terrible condition."
6.“The driver canceled the ride last minute, leaving me stranded in an unfamiliar location."
7.“The driver took a longer route, and the app had constant glitches."
8.“The driver was late, and the vehicle had a strong unpleasant smell."
9 “The driver was disrespectful, and the vehicle was dirty and uncomfortable."
10.“The app showed incorrect arrival times, and the driver was difficult to communicate with.”

Funnel Breakdown



Within these 5 steps the drop off between Downloads and Sign Ups is acceptable and although there is too much loss between Sign Ups and Rides requested which may indicate more incentive is needed to make customers request a ride the biggest challenge and opportunity is between Rides Requested/Accepted and RIdes Completed/Paid.

Customer Touch Points

Consistently the bad experiences of your customers seem to be related to your drivers which makes me wonder what your drivers are experiencing.

In that top 5 negative reviews customers say sometimes your drivers aren’t showing up, there is no information in your data as to why but it’s possible that they are not showing up because they are quitting. A quick Google search revealed the employee churn rate for Uber which is listed as the worst is 1.8 years, Lyft as the best is around 5 years. After some calculations I saw that yours is a little under 80% of one calendar year.

Factors Influencing Retention

According to Mobilitylab.org one of the top reasons people use rideshares is to be able to go out and have some drinks while still being safe. Investopedia.com concurs and adds that riders also choose fast, convenient and cheap with a deciding factor being the added peace of mind that the drivers have been background checked. These are examples of behavior that illustrates

the fact that your business is based on repeat customers and understanding their motivation is key. Your competitors are approaching customer retention in app and out by highlighting eco friendly and personalized service.

Additional findings of Interest

- Peak hours : 8am - 10am, 4pm to 6pm
- 385,477 rides requested from 12,406 distinct users
- Positive trends in Downloads, Sign Ups and Ride Requested

Recommendations

1. use or create a review rating system for both drivers and customers that can be easily sorted within your data and escalated within your company. For customers I recommend that you make the review process very user friendly, going from leaving a number of stars and suggested reasons, to being very specific using their own words, including their contact information for follow up and link this review to their ride id so you can trace it back to the driver id if you are serious about keeping them as a customer.

2. Get driver feedback. you need to know why they are not showing up and you need to resolve this recurring issue of them not refunding customers when they feel they've been overcharged.

3. Create customer retention marketing within your app that focuses on safety (background checks, car standards) , offer personalization(saved places and routes etc.) or your company's eco footprint.

4. Create driver retention strategies.My suggestions on this are things like, partial reimbursement for service receipts or premium fuel. Discounts on car washes and detail and rewards for drivers based on the customer review rating system. You may not have to reinvent the wheel as a rideshare company but you might have to think outside the box to set yourselves apart as an employer.

Conclusion

You need to investigate and then invest in your customers and your drivers and you need to track your results so that the data will clearly answer your questions in the future. Thank you for your time.

Click here for my Loom Presentation
https://www.loom.com/share/4d6812345ad542bdb0cb20d5883abf57?sid=bbbe9e01-5603-4f27-9461-81568dd00b84

Click here to view my Tableau Dashboard
https://public.tableau.com/views/MetroCarFunnel_16990323932940/MetroCarInteractiveDashboard?:language=en-US&:display_count=n&:origin=viz_share_link

Click here to view my GoogleSheets
https://docs.google.com/spreadsheets/d/e/2PACX-1vQrRZa1E45NcbZvlvKfd4PIqkN-0Jef14lxnUzSlzpQkNKp6YbgvdmhSOTrhZTAf5L4tN1bK0Yy6GK2/pubhtml

Funnel Analysis Query

```
WITH rides_status AS (
   SELECT
    dropoff_ts, ride_id
  FROM ride_requests
  WHERE dropoff_ts IS NOT NULL
  GROUP BY ride_id
),
totals AS (
  SELECT
    COUNT(*) AS total_users_rides,
    COUNT(DISTINCT transaction_id) AS paid
  FROM transactions
RIGHT JOIN rides_status rsu
  ON transactions.ride_id = rsu.ride_id
),


funnel_stages AS(
  SELECT
  1 AS funnel_step,
  'rides' AS funnel_name,
  total_users_rides AS value
  FROM totals
UNION
```

```
SELECT
2 AS funnel_step,
'payments' AS funnel_name,
paid AS value
FROM totals)
  SELECT*,
  value::float/LAG(value)OVER(
    ORDER BY funnel_step)AS payment_rate
    FROM funnel_stages
    ORDER BY funnel_step;

WITH  user_platform AS(
    SELECT user_id, platform, age_range
    FROM app_downloads
    LEFT JOIN signups
    ON app_downloads.app_download_key = signups.session_id
    WHERE user_id IS NOT NULL),

                    app_downloads AS(
      SELECT 1 AS step,'download' AS NAME,
      platform, age_range,
      COUNT(DISTINCT app_download_key) AS users_count,
      0 AS count_rides
      FROM app_downloads
      LEFT JOIN signups ON app_downloads.app_download_key = signups.session_id
      GROUP BY platform, age_range),


    signed_up AS(
SELECT 2, 'signed_up'AS NAME, platform, signups.age_range,
COUNT(DISTINCT user_id) AS signed_up_user_count,
0 AS count_rides
FROM signups
LEFT JOIN user_platform USING(user_id)
GROUP BY platform, signups.age_range),

        rides_requested AS(
          SELECT 3, 'rides_requested'AS NAME, platform, age_range,
          COUNT(DISTINCT user_id),
          COUNT(DISTINCT ride_id)
          FROM ride_requests
          LEFT JOIN user_platform USING(user_id)
          WHERE request_ts IS NOT NULL
          GROUP BY platform, age_range),
```

```
rides_accepted AS(
  SELECT 4, 'rides_accepted' AS NAME, platform, age_range,
  COUNT(DISTINCT user_id),
  COUNT(DISTINCT ride_id)
  FROM ride_requests
  LEFT JOIN user_platform USING(user_id)
  WHERE accept_ts IS NOT NULL
  GROUP BY platform, age_range),

rides_completed AS(
  SELECT 5,'rides_completed'AS NAME, platform, age_range,
  COUNT(DISTINCT user_id),
  COUNT(DISTINCT ride_id)
  FROM ride_requests
  LEFT JOIN user_platform USING(user_id)
  WHERE pickup_ts IS NOT NULL
  GROUP BY platform, age_range),

rides_paid AS(
  SELECT 6, 'rides_paid' AS NAME, platform, age_range,
  COUNT(DISTINCT user_id),
  COUNT(DISTINCT ride_id)
  FROM transactions
  LEFT JOIN ride_requests USING(ride_id)
  LEFT JOIN user_platform USING(user_id)
  WHERE charge_status = 'Approved'
  GROUP BY platform, age_range),

reviews AS (
                      SELECT 7, 'reviews' AS NAME, platform, age_range,
                        COUNT(DISTINCT user_id) AS users_count,
  COUNT(DISTINCT review_id) AS count_reviews
                        FROM reviews
  LEFT JOIN user_platform USING(user_id)
  WHERE review_id IS NOT NULL
  GROUP BY platform, age_range)
  SELECT*
  FROM app_downloads
  UNION
  SELECT*
  FROM signed_up
  UNION
  SELECT*
```

```
FROM rides_requested
UNION
SELECT*
FROM rides_accepted
UNION
SELECT*
FROM rides_completed
UNION
SELECT*
FROM rides_paid
UNION
SELECT*
FROM reviews

ORDER BY step;
```

Distinct Users Query

```
SELECT COUNT(DISTINCT user_id)
FROM ride_requests;
```

Rides Request Query

```
SELECT COUNT(request_ts)
FROM ride_requests;
```

Peak hours Query

```
SELECT
   TO_CHAR(request_ts, 'YYYY-MM-DD HH24:MI') AS peak_time,
   COUNT(*) AS request_count
FROM
   ride_requests
GROUP BY
   peak_time
ORDER BY
   request_count DESC
LIMIT 50;
```

Peak hours 8am-10am, 4pm-6pm

Repeat Customers query

```
SELECT
   user_id,
```

```
    COUNT(user_id) AS occurrence_count,
    SUM(COUNT(user_id)) OVER () AS total_occurrences
FROM ride_requests
GROUP BY user_id
HAVING COUNT(user_id) > 1;
```

MetroCar has 385477 rides requested from 12406 distinct users

Metro Car Employee Churn

```
SELECT
    driver_id,
    MIN(accept_ts) AS earliest_acceptance,
    MAX(accept_ts) AS most_recent_acceptance,
    MAX(accept_ts) - MIN(accept_ts) AS employment_duration
FROM
    ride_requests
GROUP BY
    driver_id
ORDER BY
    driver_id DESC;
```

MetroCar Customer Retention Query

```
SELECT
    user_id,
    MIN(request_ts) AS earliest_request,
    MAX(request_ts) AS most_recent_request,
    MAX(request_ts) - MIN(request_ts) AS user_retention
FROM
    ride_requests
GROUP BY
    user_id
ORDER BY
    user_id DESC;
```

I queried for reviews, made a CSV file and then
I asked ChatGPT to find duplicates in the text of 500 reviews 250 at a time. I categorized them, producing a top 10 list of negative reviews.

1.”The driver was reckless and drove well above the speed limit."
2 “The driver took a completely wrong route, and the car broke down halfway.”
3.”The driver never showed up, and I had to find an alternative means of transportation.”
4.“The driver overcharged me and refused to rectify the issue.”
5.“The driver was rude, and the car was in terrible condition."

6.“The driver canceled the ride last minute, leaving me stranded in an unfamiliar location."
7.“The driver took a longer route, and the app had constant glitches."
8.“The driver was late, and the vehicle had a strong unpleasant smell."
9 “The driver was disrespectful, and the vehicle was dirty and uncomfortable."
10.“The app showed incorrect arrival times, and the driver was difficult to communicate with.”

I had Claude.AI read a sample of 250 reviews based on these keys words. "Regrettable", "Horrible", "late", "wrong", "broke", "overcharged", "limit", "canceled", "disaster", "disgusted", "disappointed", "reckless", "uncomfortable", "glitches", "rude", "disrespectful", "dirty", "unsatisfactory", "unacceptable", "nightmare", "terrible", "never", "unfamiliar", "refused", "Unsafe", "unpleasant", "recklessly", terrible", unreliable", "unprofessional". Which resulted in 141 of 250 being negative. That is 56.4% of the sample reviews.