Data Product Manager Nanodegree

Applying Data Science to Product Management

Final Project: Developing an MVP Launch Strategy for a Flying Taxi Service

Welcome to your first week at Flyber


Aybar

In this project, you will apply the skills acquired in this course to create the MVP launch strategy for the first flying car taxi service, Flyber, in one of the most congested cities in America -- New York City.

You are responsible for bringing the first flying car taxi service to market by analyzing data and building a product proposal.

You will need to use the SQL workspace provided in the Classroom, and Tableau Public, in order to successfully complete the project.

You'll present your answers, findings, and insights in the Answer Slides found in this deck. Feel free to include any additional slides, if needed.

Section 1: Data Exploration

Back to the basics of product management, identify your customer and their pain points:

- What are taxis used for?
- What are the characteristics of the users that leverage them?
- What are existing pain points with taxis?
- What are the existing pain points with digital ride-sharing services?

Answer Slide

1) I analyzed GeoSpatial data and found that maximum density of taxi rides are around Central Terminal and New York Penn Station followed by other tourist spots like Time Square,empire state building etc.

Taxis are primarily used from/to stations to/from popular tourist spots.

2) People who use the ride services travel short distances to/from station. Mostly tourists or commute to work.

3) Taxis can become expensive and are difficult to find in dense areas. Takes time to get to destination.

4) Ride sharing can be time consuming since people travelling together might have different destinations and again difficult to hail when demand is pretty high. Takes time to get to destination.

What user improvements do you hypothesize a flying taxi service would have over the existing state of taxis today?

What market improvements do you hypothesize a flying taxi service would have the existing taxi service industry & physical road infrastructure today?

Answer Slide

Taxis and Rideshare have the same problem, takes too much time to reach destination due to density of population in Manhattan.

When I analyzed the data by pick up time during the 24 hours, we found that the demand for taxis and rideshare are pretty high through out the day.

Flying cabs and rideshares can reduce the bottle neck around popular areas, help people reach destination in time while regulating the traffic on road.

Upload this dataset into Tableau Online.

Ensure the fields are parsed correctly; field headers are included in the first row of the CSV.

Let’s begin exploration!

Acquire a high-level understanding of the granularity and scope of the dataset, to inform the basis for your analyses:

- How many records are in the dataset
- What does each record represent?
- What is the primary key?
- What date range is your dataset bound to?
- What are the geographical bounds of this dataset? Is it limited to Manhattan, or is Brooklyn, Queens, Staten Island, the Bronx, and New Jersey included? Where are most of the data points centralized at? Are there outliers?

Answer Slide

Records with distance very close to 0 miles (negligible) or duration greater than 24 hours (86400 secs) have been removed.

There are over 1 million records.

Each record represents details of one ride hailed, contains passenger counts, duration of the ride, pick up and drop off coordinates and distance travelled.

The Primary Key will be ID of the travel.

The data was collected from Jan 1st,2016 to June 28,2016.

Data points are mainly centralized in Manhattan but also covers downtown Brooklyn, parts of Long Island. Bronx and New Jersey are not included. Airports such as JFK and LGA are included as well.

You notice that the dataset does not contain explicit data points out-of the-box, we'll need to enrich the dataset with relevant fields:

- You notice that ride price is not included, but figure it could be derived. Based on information about New York taxi prices gleaned from the internet, create a calculated field called `price` using the `duration`, `distance`, and `passenger count` fields.
- You hypothesize your target users will be those who take a relatively longer time getting to a destination that is relatively close, due to heavy traffic conditions and/or limitations to physical road infrastructure. To be able to analyze where this is happening, you will need to create a calculated field called `distance-to-duration ratio`.

Let’s understand the scope and distribution various dimensions within the dataset. Calculate the **average**, **median**, and the **first & second standard deviation of the mean** for the following measures:

- duration
- distance
- passenger counts
- duration-to-distance ratio
- price

Answer Slide

average, **median**, and the **first & second standard deviation of the mean** for the following measures:

duration in seconds : Avg=842 , median=666, 1st std dev =673 , 2nd =1346

distance in miles: Avg=3.5 , median=2.1, 1st std dev =4, 2nd =8

passenger counts: Avg=2 , median=1, 1st std dev =1, 2nd =2

duration-to-distance ratio (seconds/mile): Avg=416 , median=280, 1st std dev = 3451, 2nd = 6902

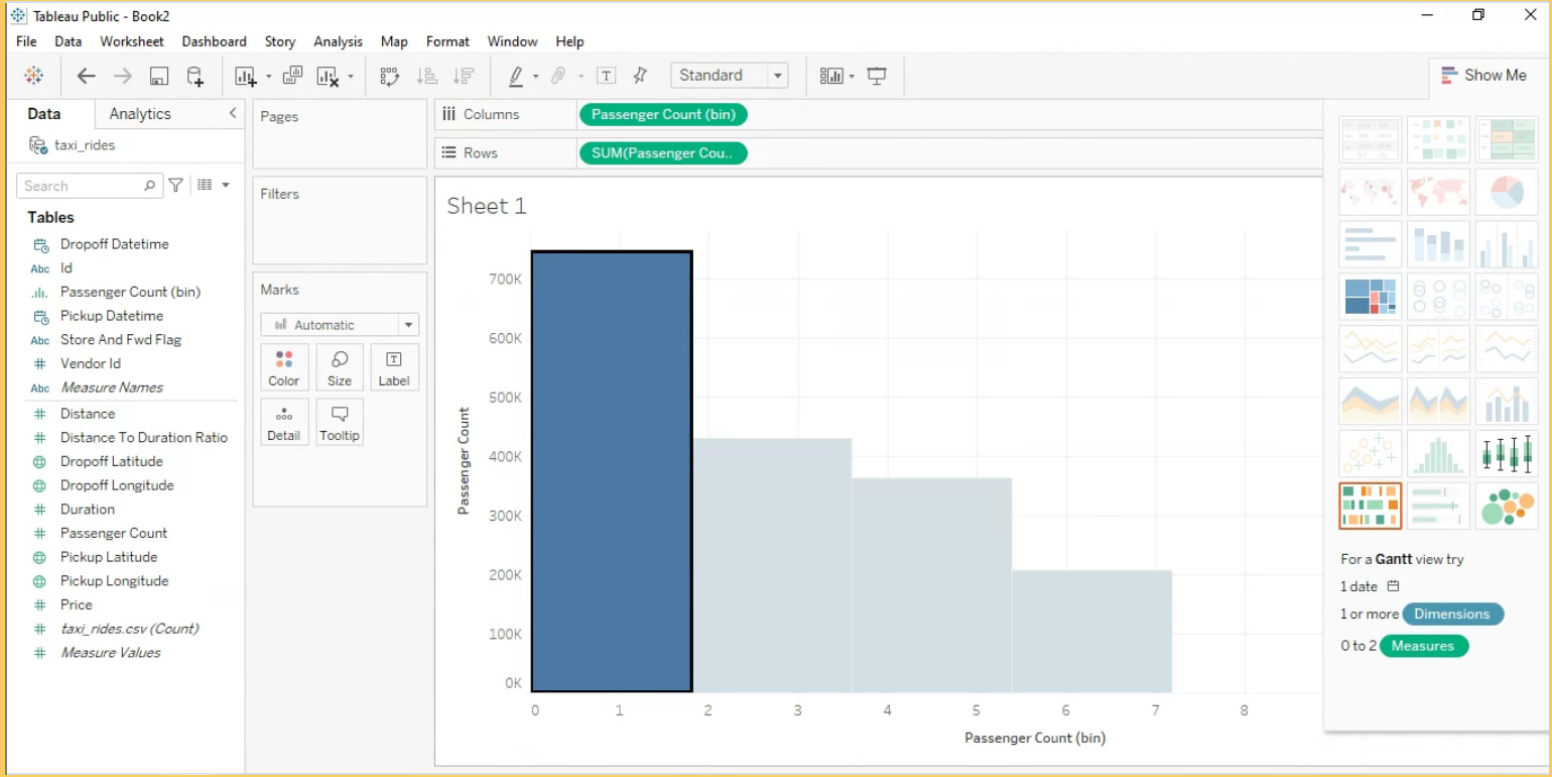
price: Avg=18.2, median=13.6, 1st std dev =15, 2nd std dev =30

Flying cars may have to have to be a lower weight for efficiency & take-off. Or you may just decide to leverage mini-copters for your initial MVP.

Create a histogram that visualizes the number of total rides grouped by passenger counts to analyze the potential market volume of low passenger pickups (1-2 passengers).

Answer Slide

We see that demand for 1-2 Passenger count is the highest over 700K.

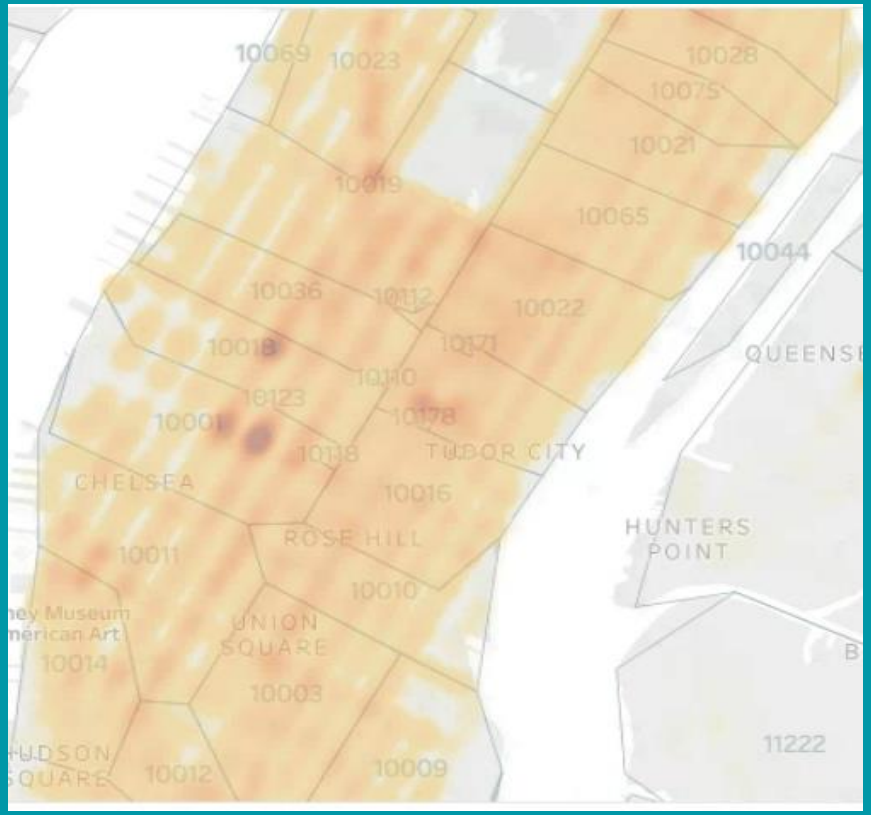


For the initial MVP launch (& most likely GA), we have a finite amount of monetary resources to build Flyber pick-up / drop-off nodes. We'll need to be strategic on where we'll place them:

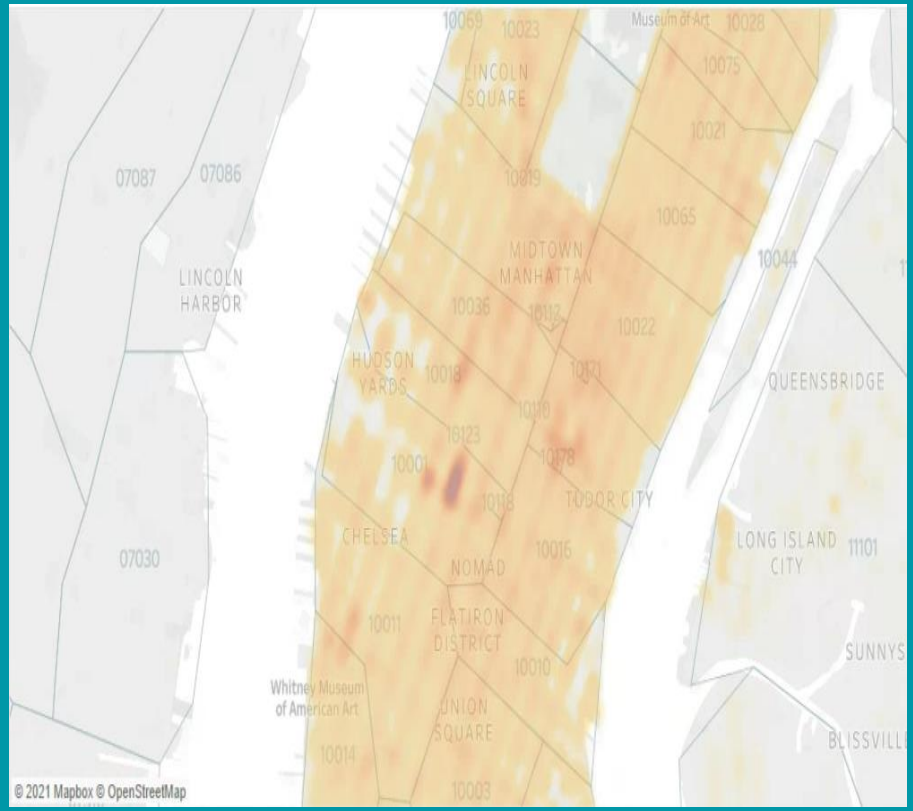
- Which neighborhoods/zip codes tends to experience a relatively higher density of pick-ups?
- Which neighborhoods/zip codes tends to experience a relatively higher density of drop-offs?
- Which neighborhoods/zip codes tends to have the highest duration-to-distance ratios, based on pick-up?
- Which neighborhoods/zip codes tends to have the highest duration-to-distance ratios, based on drop-off?
- For any of the neighborhoods identified, are there any potential areas within the neighborhood that are optimal for flying taxi pick-up / drop-off? What makes them suitable?

Pick up and Drop Off Density –Geo Spatial Maps

- PICK up



- DROP off



Answer Slide

Higher density of pick-ups zipcodes:10001(Penn Station),10018, 10178,

Higher density of drop-offs zipcodes:10001(Penn Station),10018.

Density Highest duration-to-distance ratios based on pick-up : 10014, 10011,10036,10001,10023,10024,10028,10065,10128, 10003,11101,11104

Density based on Highest duration-to-distance ratios, based on drop-off zones : 10014, 10011,10036,10001,10023,10028,10065,10128,

Building rooftops would be potential pickup and drop off zones just like copter pads.

It may not make operational sense to have the service running 24/7, for now.

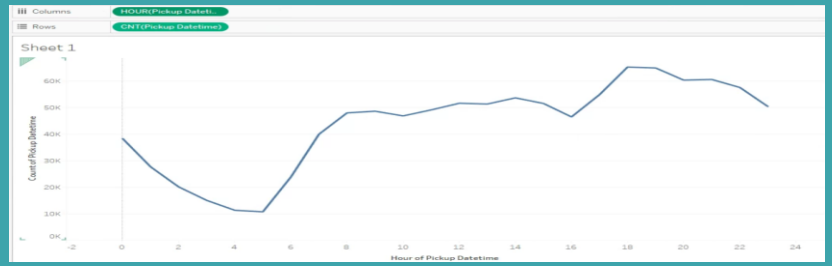
- What times throughout the day experience relatively higher volumes of ride pick-ups?

- What days throughout the week experience relatively higher volumes of ride pick-ups?

- Pinpoint any periods throughout the year that experience trend fluctuation or seasonality around ride pick-up volumes. This will help us in our post-launch analyses to determine if any spikes or dips were influenced by seasonality or through actual feature adoption/regression.

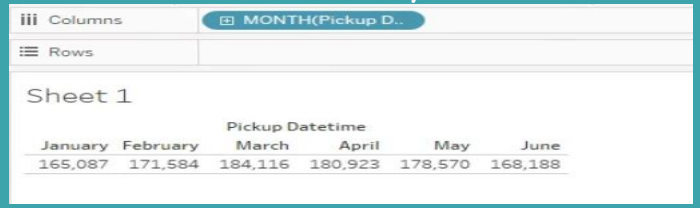
Answer Slide

The time between 5 pm to 10 pm in a day experience relatively higher volumes of ride pick-ups.



Thursdays,Fridays and Saturdays have the highest volumes of rides

This is missing Crucial Holiday season data as we have data only from Jan-June 2016. The highest volumes being in months March, April and May.



You and the user research team ran a quantitative survey on existing taxi and/or rideshare users in New York City to determine sentiment around potentially using a flying taxi service.

Dive into the survey results dataset in order to extract insights from explicit feedback.

Upload this dataset into Tableau Online or a SQL database (the classroom contains a workspace with the data for you as well).

Ensure the fields are parsed correctly, field headers are included in the first row of the CSV.

Question schema:

Q1 - What is your email?
Q2 - What gender do you identify as?
Q3 - What is your age?
Q4 - What is your annual income? (income bands)
Q5 - What neighborhood do you reside in?
Q6 - Do you currently use taxis? (Y/N)
Q7 - Do you currently use ridesharing services? (Y/N)
Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N)
Q9 - If yes to Q8, how much would you be willing to pay per mile for such a service? (USD)
Q10 - If no to Q8, what is the reason?

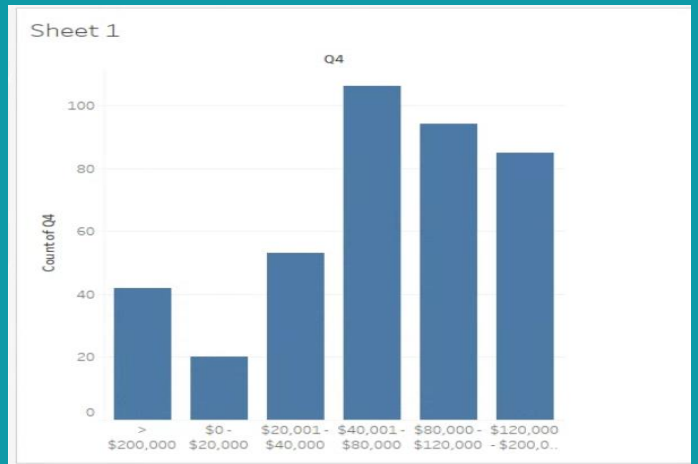
To inform our future product marketing efforts, we'll want to extract the following:

- Is there an inclination of better Flyber adoption based on gender, age, income level, or neighborhood of residence?
- What is the distribution of potential price per mile based on gender, age, income level, and neighborhood of residence?
- What is the different personas/segments of negative sentiment towards not using a flying taxi car service?

Answer Slide for Inclination to Use Flyber

The User Research group had more female participants then Male, so it is difficult to determine based on gender who would be more inclined to use Flyber. Since the histogram has the same pattern with regards to gender for total number of participants and participants who will be inclined to try Flyber we can conclude that gender does not play a role in determining inclination towards Flyber.

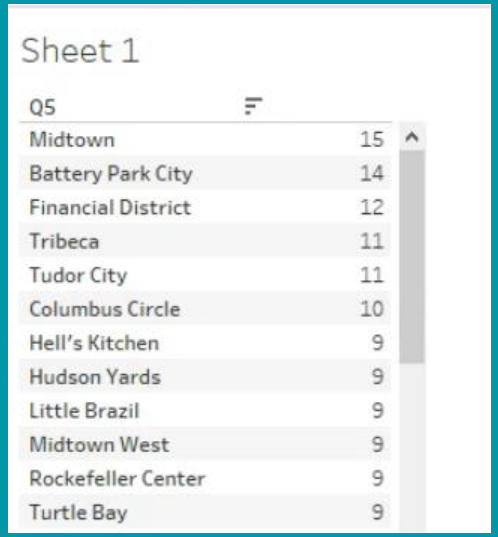
The research group comprises more of people from 40K -200K and the same group of people have inclination to use Flyber.



Answer Slide for Inclination to Use Flyber contd..

Inclination to use Flyber does not depend on age. Distribution of number of people inclined to use Flyber is similar to age distribution of participants.

Midtown,Battery Park, Financial District,Tribeca and Tudor have highest inclination
for using this service.

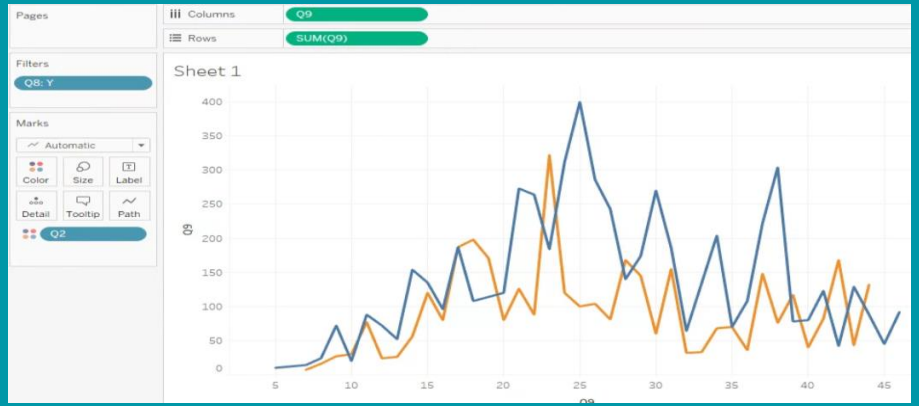


Sheet 1

Q5	
Midtown	15
Battery Park City	14
Financial District	12
Tribeca	11
Tudor City	11
Columbus Circle	10
Hell's Kitchen	9
Hudson Yards	9
Little Brazil	9
Midtown West	9
Rockefeller Center	9
Turtle Bay	9

Answer Slide for potential price per mile.

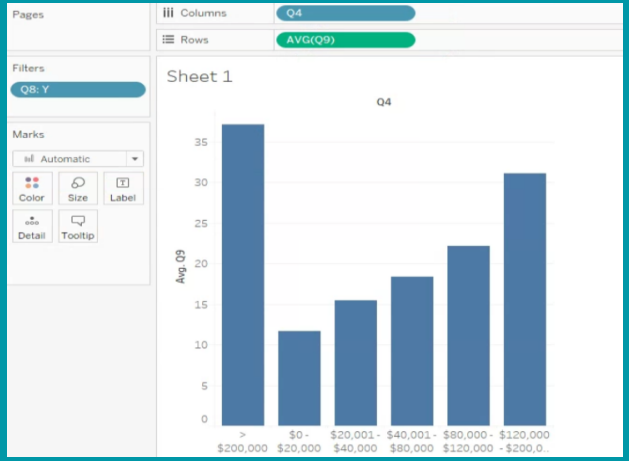
Female are inclined to pay
a max of $25/mile where as male are inclined to pay $23/mile.

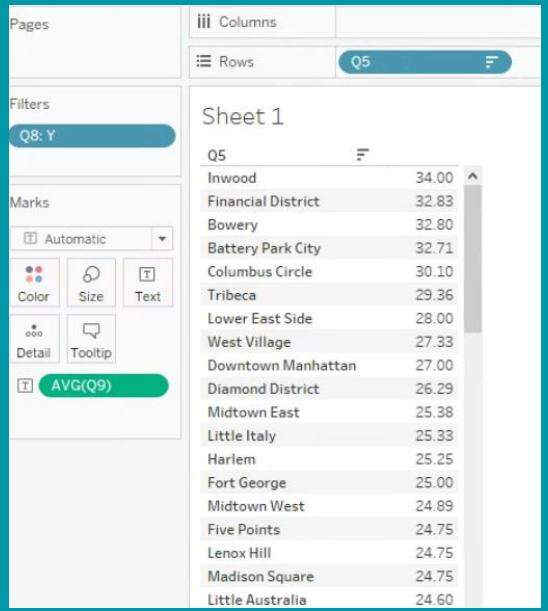


Average price/mile across various age groups seems to be roughly around $21-$25 per mile.

Answer Slide for potential price per mile contd....

High income folks are willing to pay higher

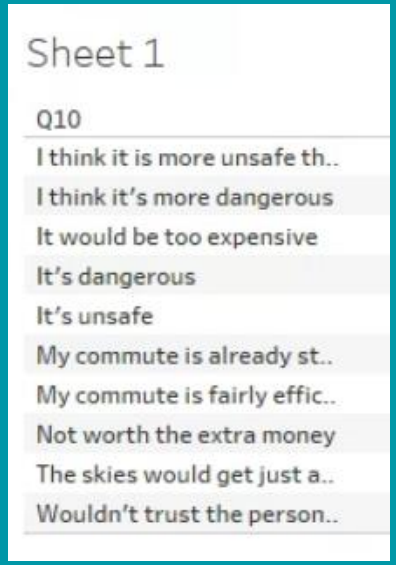




Inwood, Financial District, Bowera, Battery Park, Columbus Circle, Tribeca are inclined to pay more than $30/mile

Personas/segments of negative sentiment towards not using a flying taxi car service

Expensive, unsafe or dangerous, not trust worthy, not a necessity.



Hooray! End of Section 1.

You will complete Section 2 at the end of this course.

Please submit this file for review for Section 1.

Section 2: Proposal Synthesis

Identify a product objective for Flyber's launch. Your product objective will guide your KPIs, so identify what Flyber should optimize for. Your objective should be centered around one the following focus areas:

- User Acquisition
- User Engagement
- User Retention
- Profitability

Explain your reasoning. Include both why you feel your focus area is more relevant than the others for Flyber at this time of the product development cycle.

Answer Slide: Product Objective

At this stage of Product Cycle our best Objective would be User Acquisition.

We have done our due diligence in identifying our Target Users, understanding their pain points and analyzing how this product could resolve their current predicament.

Based on our User Survey, our Target Users are folks earning an Annual Income of 60K and above and are primarily from locations Battery Park, Financial District, Tribeca, Tudor , Midtown and Inwood.

Midtown,Battery Park, Financial District,Tribeca and Tudor have highest inclination for using this service and folks from Inwood, Financial District, Bowera, Battery Park,Columbus Circle, Tribeca are inclined to pay more than $30/mile.

Objective 1 : User Acquisition

Key Result 1: Increase Daily Active Users .

Key Result 2: Maintain Customer Acquisition Cost .

Key Result 3: Increase Total number of Rides taken on a Weekly Basis

Formulate 3-5 Key Performance Indicators (KPIs), to measure if the product is heading towards the right direction based on your objective

Key Result 1 : Increase number of Weekly New Active Users by at least 500.

- Increase in number of Trial/Discount signups by 100.
- Increase in number of new registered users by 300-400 through App
- Increase in number of Users as Guests i.e. tourists and non-regular users by 100

Key Result 2: Maintain Customer Acquisition Cost less than $20 per User

- Cost of advertising in Social Media, Search Engines and other modes of advertising must be less than $5 per new User
- Cost of discounts provided during Trial Signup must be less than $15 per new User

Key Result 3: Increase or Maintain at least number of rides to 5K every week.

- number of rides by New registered Users by 2K a week
- number of rides by users signed as guests 1K-2K a week.
- number of rides by repeat customers 2K-3K per week.

Create hypotheses around what thresholds your KPIs would need to hit in order to determine success

Answer Slide:

We are assuming here that by end of first year we would like to capture 15% of Taxi Market Share.

On an average there is an overall demand for 180K taxi rides per month, hence the number of rides on an average 36K per week.

Since we are aiming at grabbing 15% of market in first year, that would be 5.4K rides per week for Flyber.

As the product manager, you make decisions based on the insights you extract, we'll need to know the feature set we'll include in the MVP to measure viability, while keeping operational expenditure under control:

- What times/days of operation should the service run for?
- How many pick-up / drop-off nodes should we have?
- Where should the nodes be located?
- Should we initially use copters or homegrown hardware?
- Should the pricing be fixed or dynamic? At what rates?

Answer Slide

The maximum demand for Taxis are between 6 pm to 11 pm and on Wedensdays,Thursdays,Fridays and Saturdays.

The zipcodes 10014, 10011,10036,10001,10023,10024,10028, 10065, 10128, 10003,11101,11104 are ideal as pick up and drop-off nodes.

- They have the highest Duration to Distance ratios making them most obvious choice
- These zip codes also had the highest inclination of using Flyber during our initial User Research and are willing to pay higher cost /mile for the ride.

For the MVP, we need to test the concept of “Flying Rides” rather then spending resources on building flying cars in early stage.

Pricing should be dynamic based on demands and congestion on ground traffic. Also passenger counts should be factored in for Flyber.

Determine the MVP sample size & time period allotted estimated to come to a conclusion on your hypotheses.

Answer Slide

We need to test the “Flyber” concept for 6 months using just a mini copters that can carry Max 2 passengers. Cost of a used minicopter can range from 50K to 200K.

Sample size would be 3000 passengers .

We can start with 1 minicopters per Node. Operate them on Thursdays,Fridays and Saturdays between 6 pm to 11 pm.

Since maximum demands are around New york Penn station, 8th Avenue in Garment District, Herald Square and Grand Central Terminal, we can assign 2 mini copters to these nodes for starters.

Build a very basic APP/Web page for making reservations.
At the end of this timeframe, we should have achieved the following :

- captured 15% of taxi rides market.
- generate enough revenue to cover the operational cost, customer acquisition cost and make a profit margin of 6%-8% (lower margin)

Create an instrumentation plan for the events you need collected and logged, in order to be able to physically measure your KPIs.

Answer Slide

New_Registered_Users
Users_Signed_as_Guests
Repeat_Users_Hailing_Rides
New_Users_Hailing_Rides
Guests_Hailing_Rides
Payment_Per_Ride
Pickup_Zones
Dropoff_zones

Create a qualitative feedback survey questions for users after their ride, to further understand and optimize the product for future iterations.

Answer Slide

Q1 - What gender do you identify as?
Q2 - What is your age?
Q3 - What is your annual income? (income bands)
Q4 – From where do you hail the ride ?
Q5 – To where did you get dropped of?
Q6 – Did you have any safety concerns during the Flyber ride? (Y/N)
Q7 – If Yes Q6, what was the concern?
Q8- Was the ride affordable ?(Y/N)
Q9- If No to Q11, how much are you willing to pay?
Q10 – Will you recommend our services to others? (Y/N)
Q11- Will you use our services again? (Y/N)
Q12- If No to Q11, what is the reason?

Summarize everything you have learned into your final proposal

- Identify the target population. Why did you select that target population? What are their pain points?
- Create a product proposal containing claim, evidence, estimated impact, and risks
- Claims should be backed by quantitative evidence, impact should assess market needs/benefits
- Risks involve any known unknowns that we'll still need to monitor post-launch
- State cross-functional stakeholder teams that will need to be involved

FLYBER Product Proposal

The Problem :

New York city being the one among most populous and congested city in the East Coast, makes it difficult for daily and recreational commuters to reach their destinations in time.

The roads are congested making it difficult to commute via taxis and ride shares. On an average 180K taxi rides are being hailed monthly. The pickups and drop offs are concentrated to few areas especially Penn Station, Grand Central Terminal.Harold Square and Garment District.

Commuters spend a lot of time waiting in taxi to reach their destination that could be just few miles away. Calculated Duration to Distance ratios have pretty high values especially during Peak hours (6 pm to 11 pm) and during weekends.

OUR SOLUTION: Flyber

Flyber will “fly” customers to their destination via our flying cars. This will provide some breather to congestion on road.

Initial User Research has been conducted to check inclination to use Flyber. We found that out of total 500 users 400 of the Users are inclined to use.

On an average $25/mile is the price Users are inclined to pay. Users under higher income band 80K and above are willing to pay as high as $30/mile and above.

We analyzed that 100 users were not inclined to use Flyber mainly due to safety concerns or cost of the ride.
If these two issues are addressed we could persuade more to register with Flyber services.

MVP Concept

Our analysis showed that majority of the rides almost 700K had only 1-2 passenger per ride. Hence our MVP concept could be a group of mini copters. An used minicopter costs around 50K-150K.

We are currently planning to operate Flyber during peak demand mainly between 6 pm to 11 pm and on Thursdays, Fridays and Saturdays as indicated by our initial analysis.

The launch nodes planned to be deployed to zipcodes 10014, 10011, 10036, 10001,10023,10024,10028, 10065, 10128, 10003,11101,11104, as these were identified as having highest duration to distance ratios based on our initial analysis

Our User Research showed that residents from some of these zipcodes also are the ones that are willing to pay more for the flyber services.
The pricing could be set to $27/mile, based on our User Research data.

Impact and Risk

Flyber services could significantly improve the quality of transport in congested cities like New York. We are looking to have a market share of 13%-15% by then end of 1st year.

This concept needs significant investment towards buying mini copters atleast 12 of them. If demand increases we might need to invest in more.

This concept also involves the need for trained pilot to fly the copter which could increase operating cost .

Need to watch out for other flying bodies in air.

Need to consider bad weather conditions when there is a possibility that we might not be able to operate Flyber.

Cross Functional Stack holders

Need to collaborate with Aviation supplies and technicians to purchase and maintain mini copters.

Need to collaborate with Air Traffic Control board to regulate arial traffic on a daily basis.

Need to hire or train the pilots to fly copters.

SW Engineering team to build the basic website for registration and collect feedback from users.