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

Rybel

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?

- Taxis are used for commuting between different geographical locations within the same city.
- The characteristics of the users that make use of this service are usually people in a mid high level socioeconomic level that use the service on a regular bases for commuting for work purposes (meeting clients, base work location, etc). The other type of users are more for entertainment purposes which means be for tourists or locals.
- The existing pain points with taxis is mainly that the limitation to moving with this service is the dependency there is on road infrastructure and the congestion that can come along with this when travelling at peak time sand according to how well the traffic management is in the city. The other paint point for taxis is that for longer distances travel, the time of travel is limited not only by the road infrastructure, highways, etc, but also by the speed limits and the maximum velocity a car can travel.
- The pain points with digital sharing service is that it is really hard to match offer and demand from drivers and riders at the same time. Usually people require this service in peak times of the day and there might be a saturation of the offer.

have over the existing state of taxis today?

What user improvements do you hypothesize a flying taxi service would

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

- The main user improvement a flying taxi service would be to reduce the total time travel from origin to destination.
- With regards to market improvement, the flying taxi service would not require such a high infrastructure building and maintenance costs, as the road infrastructural convenential taxi requires.

The non-recurring costs of building the infrastructure is huge as well as limited by the space there is in land for dense cities. However, when thinking of a flying taxi service, the fact that we can use a 3D routes means not only that the infrastructure costs is reduced to the take off and landing areas, but also that the amount of trajectories to go from point to point are much larger and therefore reducing congestion constraints and consequently reducing time travel.

Upload this dataset into Tableau Online.

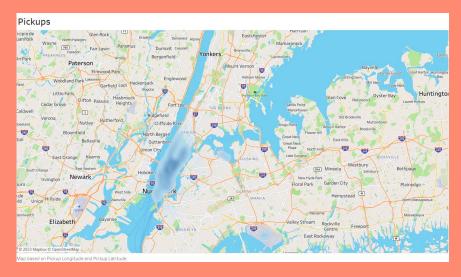
Ensure the fields are parsed correctly; field headers are included in the first row of the CSV.

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?

- There are 1 048 468 records in the table.
- Each record represents a taxi ride with the coordinates of the pick up point and drop off points, their corresponding times and the number of passengers in the taxi.
- The primary key for the table is the id column.
 This is done by making sure
 SELECT count(distinct id) FROM taxi_rides = SELECT * FROM taxi_rides
- The date range for the dataset is (1st January 2016, 1st July 2016)
- The geographical bounds for this dataset are mainly for Manhattan JFK, La Guardia and Newark Liberty airports when we look at the density map plots in the next slide. Of course there are drop offs and picks outside of these areas (e.g. Queens) but mainly the majority of the movement happen in this area. There are also some points that are clearly outliers in the data that we will remove









Here we can see what are the neighbourhoods with the highest density pick ups and drop offs

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
- distance-to-duration ratio
- price

	Mean	Median	1 std	2 stds
Duration (seconds)	962.2	662	5853.30	11706.60
Distance (miles)	3.44	2.09	4.38	8.76
Passenger Counts (#)	1.66	1	1.31	2.62
Duration to Distance Ratio (mins/mile)	78.11	4.68	15406.22	30812.44
Price (\$)	19.16	13.55	50.90	101.80

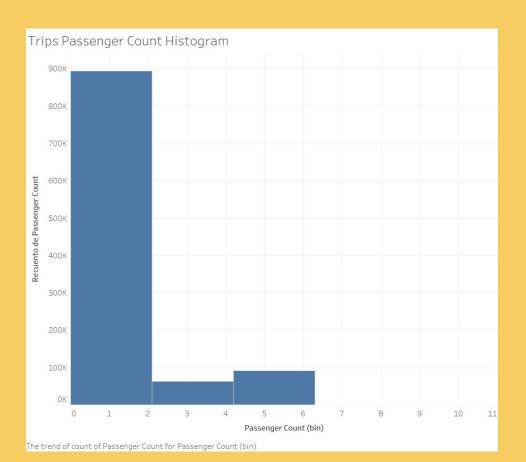
Price = 2.5 + (1.56*Distance*1.61) + (Duration/3600)*30

https://knowledge.udacity.com/guestions/645559

https://www.estimate.taxi/rates/united-states/new-york

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).



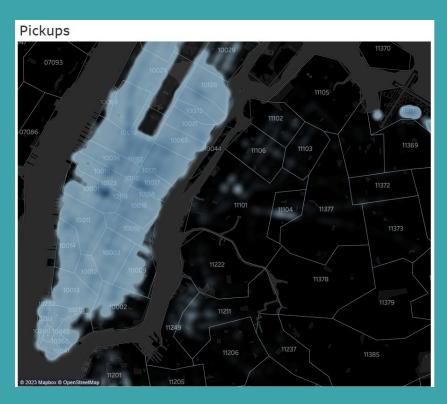
For the initial MVP launch (& most likely GA), we have a finite amount of monetary resources to build Flyber pick-up / drop-off nodes. We'll need to be strategic on where we'll place them:

- 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?

Highest density of Pick ups Neighbourhoods are in Manhattan: Midtown Manhattan, SOHO



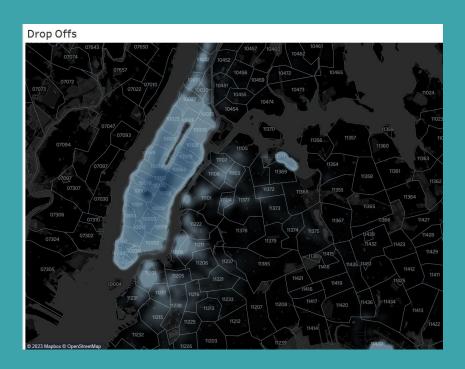
Highest density of pickups ZIP Codes are: 10001, 10178, 10018



Highest density of Drop Offs Neighbourhoods are in Manhattan: Midtown Manhattan, SOHO, airports



Highest density of drop offs ZIP Codes are: 10175, 10001, 10112



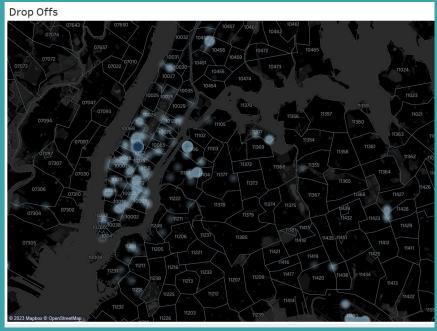
Highest density of Pick ups Neighbourhoods based on Duration to Distance ratio are: Between Midtown Manhattan & Columbus Circle and Ravenswood; Zip: 10019 & 11106





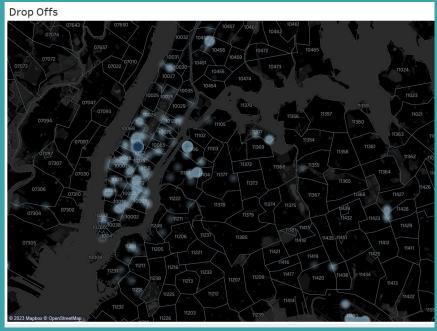
Highest density of Drop Offs Neighbourhoods based on Duration to Distance ratio are: Between Midtown Manhattan & Columbus Circle; Zip codes:10019





Highest density of Drop Offs Neighbourhoods based on Duration to Distance ratio are: Between Midtown Manhattan & Columbus Circle; Zip codes:10019



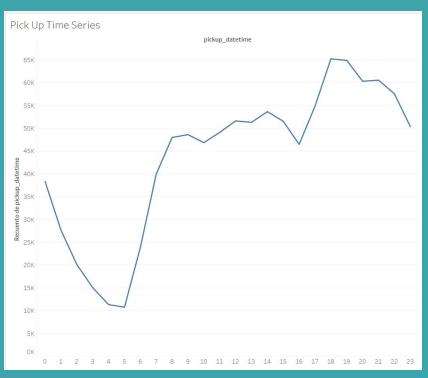


The potential neighbourhoods for both pick ups and drop off should definitely include the middle part of Manhattan since that is where we can see a higher duration to distance ratio which makes sense because there will be higher traffic in these areas.

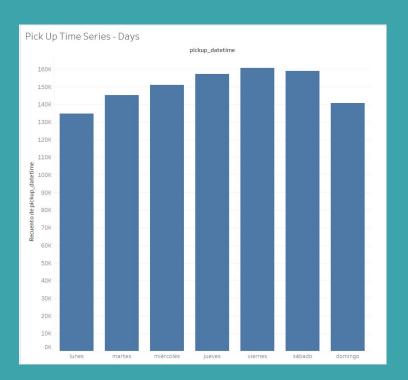
It may not make operational sense to have the service running 24/7, for now.

- 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.

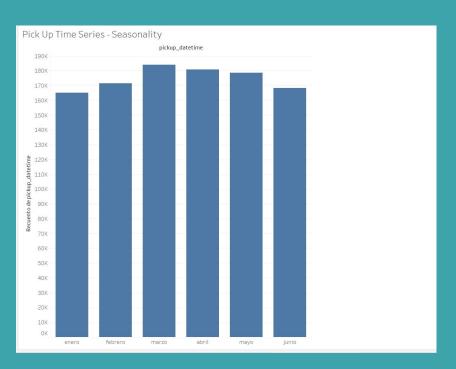
The highest peak of pick up times occurs during 18:00 and 19:00 and more broadly between 18:00 and 22:00

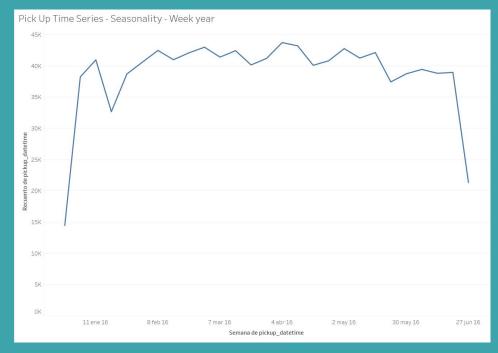


The highest peak of pick up times occurs on Friday first then Saturday followed by Thursday



The highest peak of pick up times occurs in March followed by April then May. We also see a bigger decrement in pick ups in the 3rd week of January





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 <u>this dataset</u> 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.

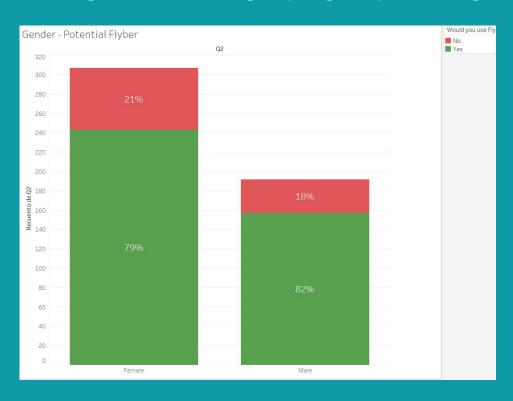
Question schema:

- 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?

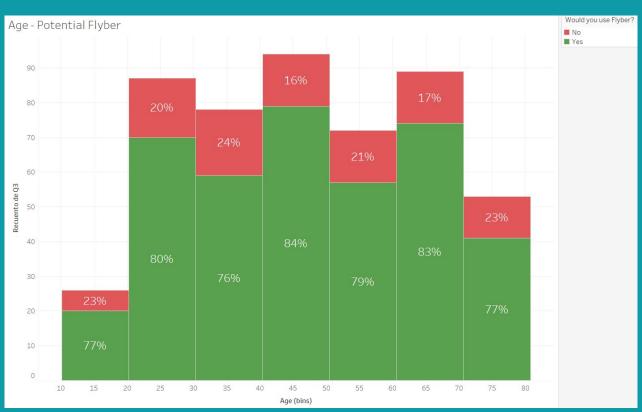
To inform our future product marketing efforts, we'll want to extract the following:

- 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?

Flyber potential usage- Gender: Slightly higher percentage of adoption for Male

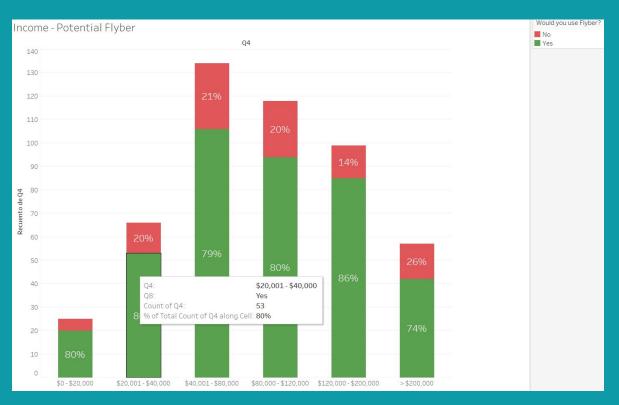


Flyber potential usage- Age: Highest potential 40 -50 followed by 60 - 70

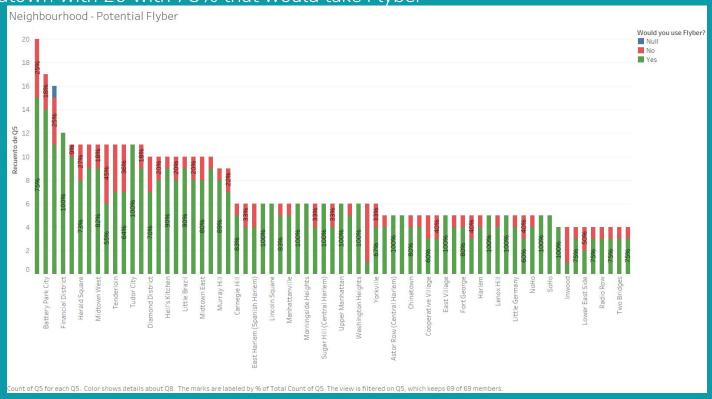


Flyber potential usage- Income: Most potential highest income in the bracket 120

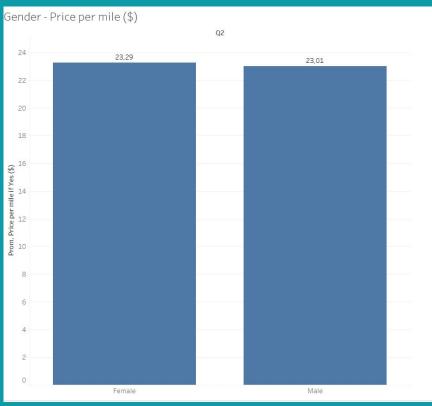
-200 K\$



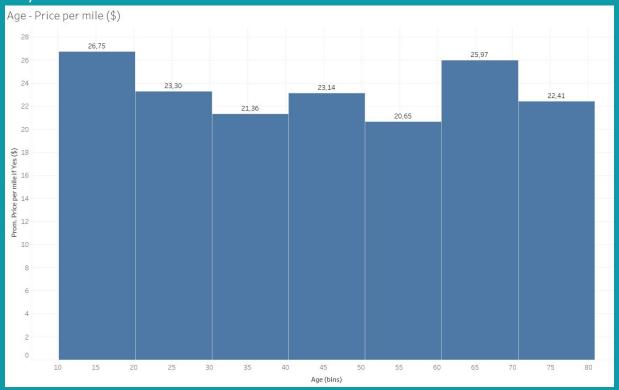
Flyber potential usage- Neighbourhood: Very heterogeneous neighbourhoods. Most common is Midtown with 20 with 75% that would take Flyber



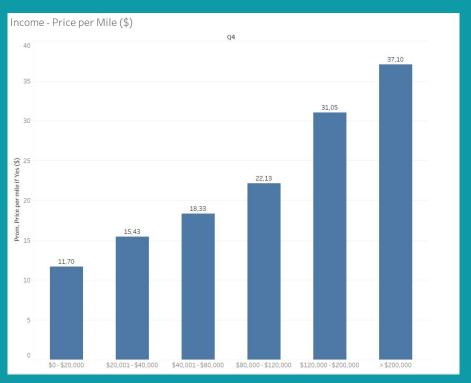
Flybe Gender - Price per Mile (\$): Very similar across genders



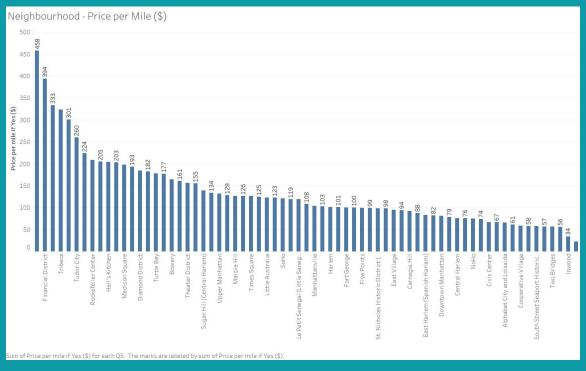
Flyber Age - Price per Mile (\$) : Surprisingly highest is the range from 10-20 years followed by 60 - 70.



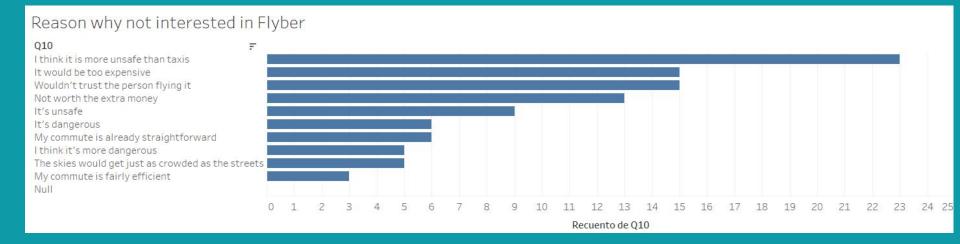
Flyber Income- Price per Mile (\$): As expected the higher the income the higher the price per mile rate the people are willing to way for Flyber if it existed.



Flyber Neighbourhood - price per mile (\$): This is higher to analyze since the maximum number of people per neighbourhood is 20 for Midtown, but in any case the neighbourhood willing to pay the highest price per mile is Battery Park City



Why are the potential clients susceptible to Flyber. Again, not a lot data for this, so would need further investigation, but here are the preliminary results from this survey. Based on this the main reason behind the people saying no, is that they believe it would be unsafe and they imagine that it would be too expensive. Other than that there is no clear mapping of a segment that rejects the Flyber be age, gender, income or neighbourhood.



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.

My project objective is focus around <u>User Acquisition</u>. Flyber is a new service that is trying to penetrate in the short distance mobility market. Knowing the history of this market, it makes sense to focus first in acquiring users even more than profitability because perhaps the net present value will be negative for the first years until there is a good bulk of users using the service.

Since the concept of Flying Taxis is new, to set up the most realistic scenario, it makes sense to start focusing on user acquisition since this will be the key and most challenging area to tackle. Later once the product is more robust we can focus more on User Engagement & User retention

User Acquisition:

- Customer Experience Let's say that the competitor of Flyber will be Uber amongst others car riding platforms. It is vital that the service shall offer is a door to door seamless experience to the customers.
- User Experience There should be a very simple and intuitive way of using the application since again the competitors have very high standards with offering a good user experience in the platform
- **Brand Awareness** This is one of the trickiest objectives whilst penetrating a new product and moreover with a service that is futuristic and might seem scary for some people. Health & Safety will therefore be key when trying to create the branding for Flyber
- Market Share This is a challenging market competing with ground transportation offering the differential of faster commuting. It will be relevant to set objectives of market share at the beginning.

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

- <u>Trial bookings:</u> N° of first time users who book their first flight with Flyber to experience it.
- **User referrals:** N° of referrals per existing users
- <u>User registration Conversion Rate:</u> N° of users that sign up for Flyber / N° of Users that download the Flyber app (or enter website)
- App downloads: N° of App downloads
- Market Share: Market share % and compare it to competitors.

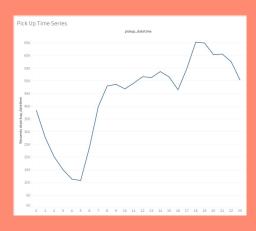
Create hypotheses around what thresholds your KPIs would need to hit in order to determine success

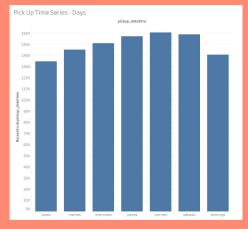
- <u>Trial bookings:</u> This is a very hard metric to achieve for a completely new product that is futuristic entering the market. Positive reviews and word-of-mouth recommendations can significantly boost trial bookings. There will be several factors affecting this price being one of course, but let's say that at the beginning we match the price to the cost of a taxi even if it leads to a loss just to acquire users. In that case defining a target of 500 trial bookings per month during the first three months could be a good target
- <u>User referrals:</u> It takes a while until users can recommend Flyber to other potential users. Let say after the 6 months we target 3 referrals per user in average.
- <u>User registration Conversion Rate:</u> This will vary also depending on marketing strategies we have with vouchers and incentives for users to get on their first ride. It may also depend how strict we are with the details needed for registration. The idea for a good user experience is that this is as lean as possible and to hit a good 30% of conversion for this metric
- <u>App downloads:</u> There will be a lot of interest in finding more about Flyber when it is launched and we expect many people to download the app. The target for this KPI is of 50000 downloads in the first year.
- <u>Market Share:</u> This will very much depend. We will consider the market share vs competitors for the specific routes that will be designed to begin with, since the idea is to compete for those specific routes. On average for all routes after 1 year of operations we want to achieve 5% of the market share.

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?

The peak times of the day are from 18:00 to 22:00 and on Friday and Saturday's. However this might vary according to the routes that we offer. Further investigation in later rounds once the routes are being matured, could be conducted. However, it is very important to always have in mind the times and the dimensioning of the whole system of systems based on this factor. Let say that the aim is to achieve rides that are more than 30 mins long, but less than 1 hour long. During peak times we shall account that we need the infrastructure and vehicles in place taking all these factors into account.





In terms of Nodes based on the highest distance to duration ratio which are potentially the areas where the differential value proposition of Flyber would stand out the most, therefore the nodes would be the following:

- Midtown Manhattan
- Laguardia Airport
- JFK airport
- Ravenswood
- Financial District

These 5 nodes making it interesting to offer commuting between these points and are far away enough capturing passengers around



these nodes and at the same time are where the distance to duration ratios

In terms of hardware of whether to use copters or homegrown hardware, we will be selecting homegrown products. From a product point of view, imagining having flying taxis in cities will be something quite impacting for the people in the city especially if the objective is to make this accessible for a person with an average income since this would mean a lot of flying taxis in the city. This means that noise & safety will be key to being able to make this method of transportation feasible. From an initial investigation it seems that current copters have exactly those two drawbacks, noise & safety so initially it makes sense to push for homegrown hardware using similar aeronautical experience from conventional helicopters. Also having an e-VTOL will allow to reduce on **CO2 emissions** which is another very important topic whilst entering the current transportation market. The e-VTOL option could imply some sort of hybrid vehicle which gives an redundant engine in case of failure of the main engine aswell so this could be a good approach.

In terms of pricing this is definitely a key topic. It could be an interesting strategy to start off initially to acquire users to fix a price that is similar to what a taxi would charge. Later on, it would be good to use a dynamic pricing strategy similar to Uber since this gives the possibility to change the prices based on offer and demand increasing the chances to use at the maximum capacity the service and at the same time making the biggest profit for Flyber.

Since the target audience is those people with an income of more than 40k\$, then based on the data of how much they will be willing to pay per mile is 25.14\$ based on the user data. Lets use a formula similar to the one of UberBlack but with the above price per mile which is what the users said in average they would be willing to pay in average for our target market of potential users earning income above 40k\$ \rightarrow

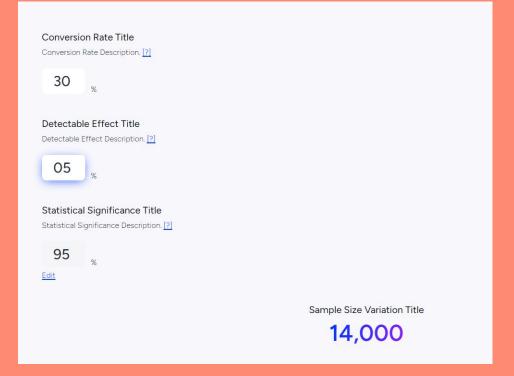
Price = 7 + 0.8 per minute + 25.14 per mile

This would mean that from mid Manhattan to JFK (15 miles) for instance lets say at an average speed of 75 mph the ride would take around 12minutes plus the take off and landing part so lets say 15 minutes. With the above formula the price would be 394\$ which if we had a 4 passenger capacity would be approximately 100\$ fare per pax if the aircraft were full.

https://www.investopedia.com/articles/personal-finance/021015/uber-versus-yellow-cabs-new-york-city.asp



If we had 30 e-VTOL aircrafts in the 6 months. The experiment we want to design implies having 14000 trips, which means around 2350 trips per month which leads to every aircraft performing 78 trips per month paper aircraft which means an average of 2.6 trips per day per vehicle



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

- 1. **appDownload:** This event will be triggered when the Flyber app is downloaded from any of the stores Android or iOS. Specific properties would be: user_id, app_download_time, mobile_operating_system
- 2. **rideCompleted:** This event will be triggered when the ride of a passenger is completed. Specific properties would be: ride_id, user_id, total_number_of_rides, ride_booked_timestamp, ride_start_timestamp, ride_finish_timestamp, ride_price
- 3. **rideCancelled:** This event will be triggered if the user books a ride but decides to cancel it. Specific properties would be: ride_id, user_id, total_number_of_rides, ride_booked_timestamp, ride_cancelled_timestamp, cancellation_charge
- 4. **userFeedback:** This event will be triggered once the rideCompleted or rideCancelled events are completed to get the feedback of the experience of the users. Specific properties will include: ride_id, user_id, thumbs_up_down_experience, would_repeat, comments

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

Thumbs up or down for Flyber

On a scale of 1 to 5, 1 being least satisfied and 5 being very satisfied:

- How likely are you to use Flyber again?
- Did you feel scared during your ride?
- Did you feel excitement during your ride?
- Were you satisfied with the quality/price of your ride?
- Would you recommend Flyber to a friend/family?
- Did you save time overall when using Flyber?

Additional comments

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

- Identify the target population. Why did you select that target population? What are their pain points?
- By looking at the Average salary bracket of the users, we find that its close to 100K\$ which is already way above the average and we need to take this into consideration. Amongst the data that we have we see that the target population of above 40K\$ is the one more interested in Flyber in general and that there is no gender difference between at this level of study.
- When looking at the price per mile that the target population is willing to pay for Flyber, it ranges all the way from 18 \$ per mile for the lower income bracket all the way up to 27\$ per mile for the highest bracket.
- We have 400 people that are willing to try Flyber versus the 99 that are not which means there is a big potential for the service
- From those 99 that are not interested it is mainly because they think that it is unsafe followed by they believe it would be too expensive. This means that if the price is competitive and the flying taxi transportation method is shown to be safer than ground transportation, there is a huge potential to win those people who are susceptible at first.
- Based on the user survey conducted, the vast majority currently use taxis, 406 vs 93 that don't. This corresponds to 81.4% that use taxis and are therefore the percentage of potential users of our service.
- Furthermore, 295 currently use ride sharing platforms versus 204 that don't. This corresponds to 59% that use ride sharing platforms which are definitely going to correspond more to the early adopters of our service since they are familiar with using internet apps for commuting which is what Flyber would be.

Create a product proposal containing claim, evidence, estimated impact, and risks

IN NYC there is definitely an issue of traffic congestion on ground transportation despite their being a subway in the city. This also leads to traffic accidents.

Flyber would be a great value proposition since by adding the third dimension in the sky, this means of transportation implies very quick commuting between nodes. It is true that the non recurring costs needed to initially set up the whole infrastructure and systems of systems, e.g. aircraft, and the vertiports for pick up and drop off of passengers in the chosen nodes, would be high but this type of investment pays off in long term specially for a problem that is known to be there a long time because the estimation of population in the cities is known to increase in the next decades (Table 1 next slide)

Travel Time



References:

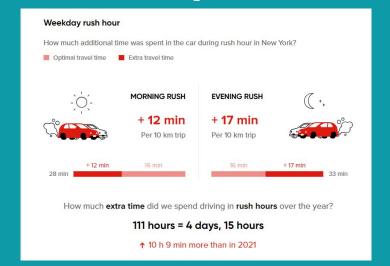
 $https://www.nyc.gov/assets/planning/download/pdf/data-maps/nyc-population/projections_report.pdf$

https://www.tomtom.com/traffic-index/new-york-traffic/

Population Density NYC

Projected Total New York City Population by Borough, 2000–2030												
		2010	2020	2030	CHANGE							
	2000				2000-2010		2010-2020		2020-2030		2000-2030	
					Number	Percent	Number	Percent	Number	Percent	Number	Percent
NYC	8,008,278	8,402,213	8,692,564	9,119,811	393,935	4.9	290,351	3.5	427,247	4.9	1,111,533	13.9
Bronx	1,332,650	1,401,194	1,420,277	1,457,039	68,544	5.1	19,083	1.4	36,762	2.6	124,389	9.3
Brooklyn	2,465,326	2,566,836	2,628,211	2,718,967	101,510	4.1	61,375	2.4	90,756	3.5	253,641	10.3
Manhattan	1,537,195	1,662,701	1,729,530	1,826,547	125,506	8.2	66,829	4.0	97,017	5.6	289,352	18.8
Queens	2,229,379	2,279,674	2,396,949	2,565,352	50,295	2.3	117,275	5.1	168,403	7.0	335,973	15.1
Staten Island	443,728	491,808	517,597	551,906	48,080	10.8	25,789	5.2	34,309	6.6	108,178	24.4

Traffic congestion NYC



Risks:

- **Regulatory compliance** Air Traffic Management will be complicated to solve. Starting this project directly that involves a lot of new technologies in a densely populated like NYC is risky since there are a lot of other flying vehicles in the air space, tall buildings and a high density of people leading to higher risks in case of a failure
- Pioneering eVTOL aircraft high challenge to design quiet helicopters that are highly safe and reliable and have two reliable sources of propulsion one the electric module for emergency landing and the other the normal diesel engine. Consider hybrid electric propulsion to reduce CO2 concentration in the cities since key metric for future transportation industry. Technology demonstrators will be needed for compliance with flying new technologies, this will be a challenge
- **Pilots upscaling** Need to train pilots specialized for flying taxis. Safety and mortality rate will be key to gaining acceptance from the society on a new transportation method. Make sure to upscale pilot training to be able to fly as many aircrafts as the business will require.

Cross functional stakeholders:

- **Aerospace engineers -** Developing new aircraft or adapting existing helicopters to satisfy the needs of the market which are less noisy and safer conventional helicopters that produce less CO2.
- Legal & Compliance Aircraft regulation (FAA), technology demonstrators flying (FAA),
 Air Traffic Management (FAA)
- **Software Engineering** Developing Seamless user friendly Flyber and website apps
- Civil Engineers Build Vertiports (Nodes) for eVTOL landing and take off.
- **User Experience** Build seamless experiences for customers on Flyber app and at the node airports, connect app with other apps for seamless door to door experience
- Logistics & Operations teams
- Marketing & Business Development
- Cybersecurity keep all software in the aircraft specially but also databases safe for cyberattacks damaging image of Flyber
- **Data teams** Storing and exploiting data to improve product features
- **eVTOL Pilots** Making sure there are enough well trained and happy pilots that can dly the aircraft safely