Data Product Manager Nanodegree

Applying Data Science to Product Management
Project: Developing an MVP Launch Strategy for a Flying Taxi Service
by Tibor Zahorecz


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: Business understanding

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

Q: What are taxis used for?
A: A taxicab, also known as a taxi or a cab, is a type of vehicle for hire with a driver, used by a single passenger or small group of passengers, often for a non-shared ride. A taxicab conveys passengers between locations of their choice. This differs from public transport where the pick-up and drop-off locations are decided by the service provider, not by the customers, although demand responsive transport and share taxis provide a hybrid bus/taxi mode. Learn more on wikepedia: link

Answer Slide

Q: What are the characteristics of the users that leverage them?
A: The taxi has the following advantages over other modes of transport: private, end-to-end, the driver optimizes the journey according to the needs of the passenger, comfortable, personal.

Q:What are existing pain points with taxis?
A: Taxis move on a fixed track, we could say that their movement is affected by the traffic situation in two dimensions.

Answer Slide

Q: What are the existing pain points with digital ride-sharing services?
A: The service has the same problems as taxis, i.e. it is strongly affected by the traffic situation and resources are limited.

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

A (user improvement): Because traffic in the air is currently low (drone control), it may be the best solution for passengers who are in a hurry during peak periods or in the event of an accident.

B (market improvements): When air taxis become widely commercialized, they will definitely ease the traffic burden on city roads. They will usher in a nimble form of intracity travel, transporting people on the shortest possible route between two locations. Source: link

Section 2: Data understanding

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

Q: How many records are in the dataset? A: 1.048.468
Q: What does each record represent? A: see dataset field page
Q: What is the primary key? A: ID (a unique identifier for each trip)
Q: What date range is your dataset bound to? 2016 H1 (01-07)
Q: 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? A: see in the next slides

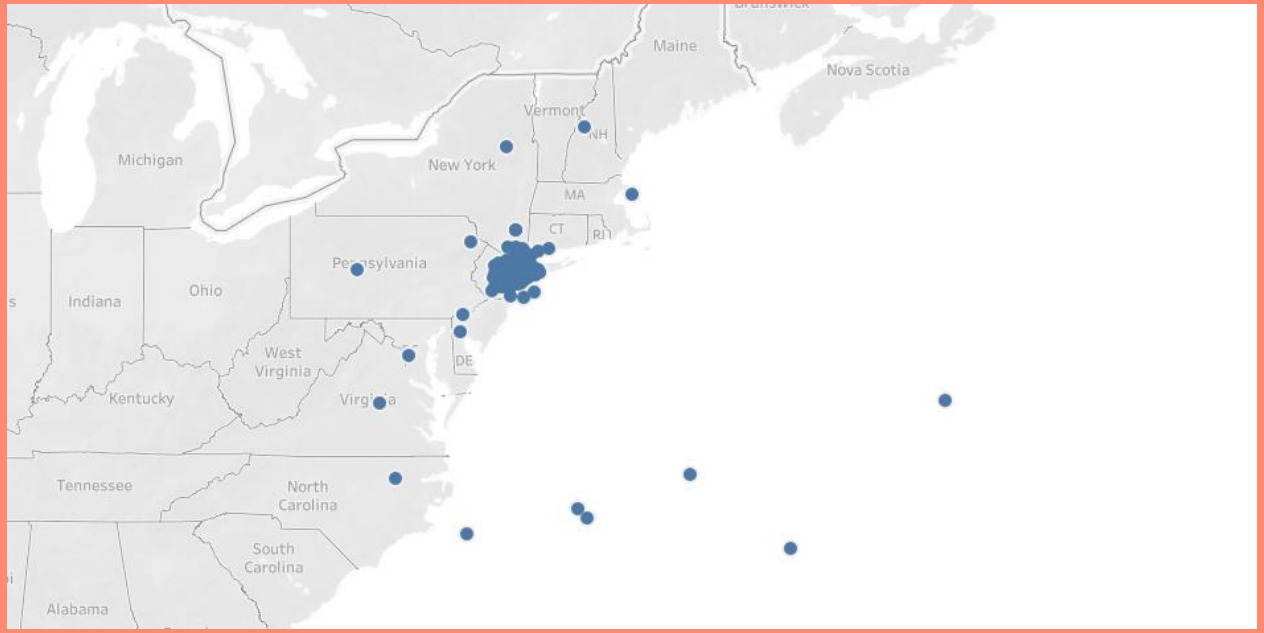
Answer Slide

Dataset fields can be find on Kaggle and with some data analysis:
source link

id - a unique identifier for each trip
vendor_id - a code indicating the provider associated with the trip record
pickup_datetime - date and time when the meter was engaged
dropoff_datetime - date and time when the meter was disengaged
passenger_count - the number of passengers in the vehicle (driver entered value)
pickup_longitude - the longitude where the meter was engaged
pickup_latitude - the latitude where the meter was engaged
dropoff_longitude - the longitude where the meter was disengaged
dropoff_latitude - the latitude where the meter was disengaged
store_and_fwd_flag - This flag indicates whether the trip record was held in vehicle memory before sending to the vendor because the vehicle did not have a connection to the server - Y=store and forward; N=not a store and forward trip
trip_duration - duration of the trip in seconds
distance is in miles

Answer Slide

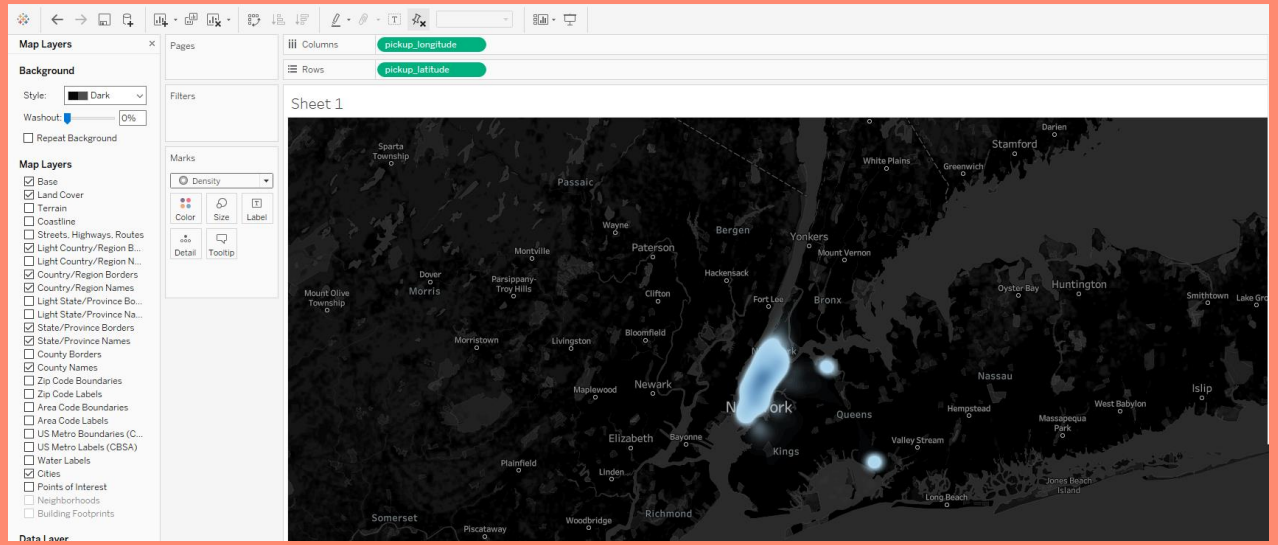
Q: What are the geographical bounds of this dataset? Is it limited to Manhattan, or is Brooklyn, Queens, Staten Island, the Bronx, and New Jersey included? Where are most of the data points centralized at? Are there outliers?



A: The data points are centralized in New York but can find outliers outside and some are even in the ocean (so data cleaning was not perfect).

Answer Slide

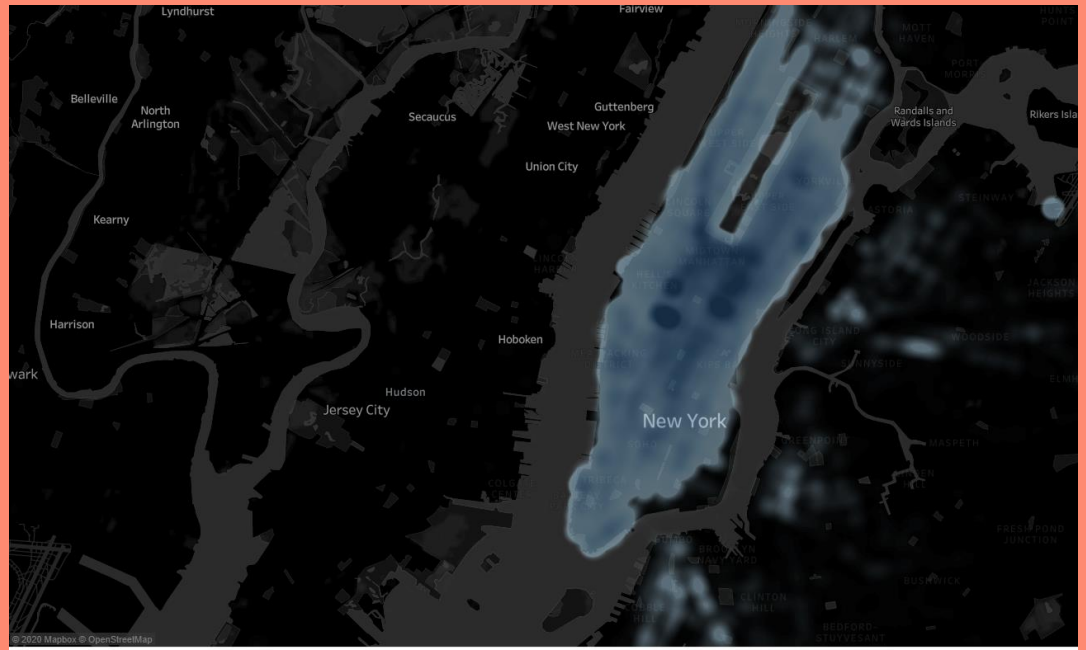
Q: What are the geographical bounds of this dataset? Is it limited to Manhattan, or is Brooklyn, Queens, Staten Island, the Bronx, and New Jersey included? Where are most of the data points centralized at? Are there outliers?



A: changing to Density marks we can see more clear where data was created.

Answer Slide

Q: 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?

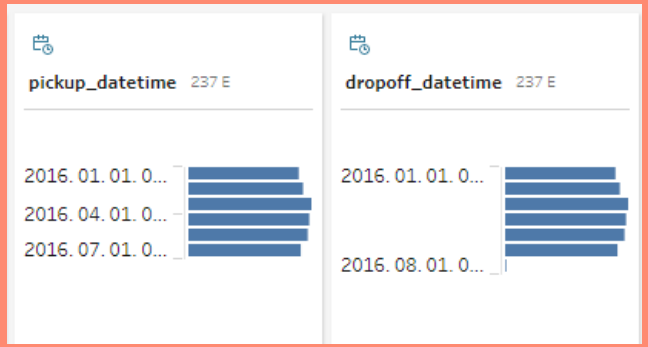


A: applying pickup_datetime: (Monday, Wedneday)
[Public link]

Section 3: Data Cleaning

Data analysis & cleaning process (Tableau Prep Builder)

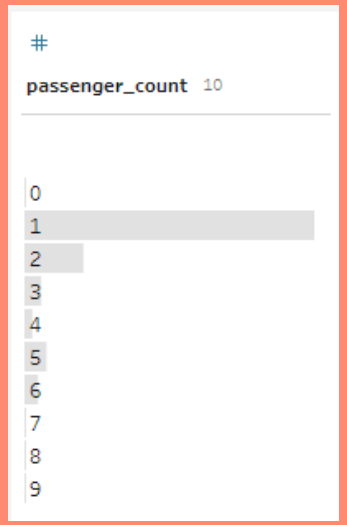
Interval: from 1st Jan 2016 to 1st Aug 2016

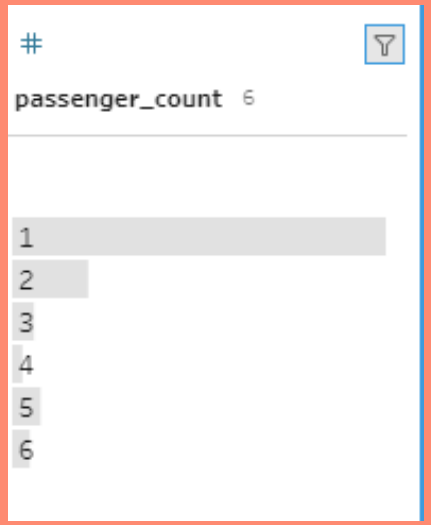


*Another way for data cleaning is Python libraries like panda and numpy. [good resource]

Data analysis process (Tableau Prep Builder)

Passanger nr: 71% 1 passanger, 14% 2 passangers, 4% 3 passangers etc.
Outlier: 0 passanger, 7-8-9 passangers ➔ modify the range according





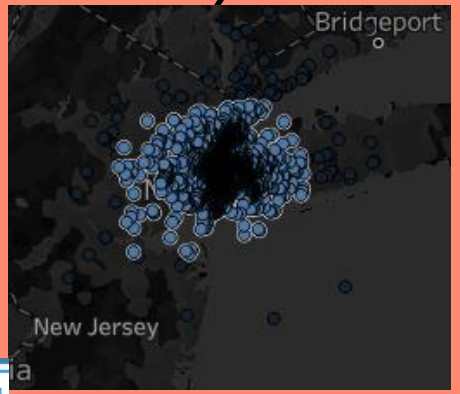
Data analysis process (Tableau Prep Builder)

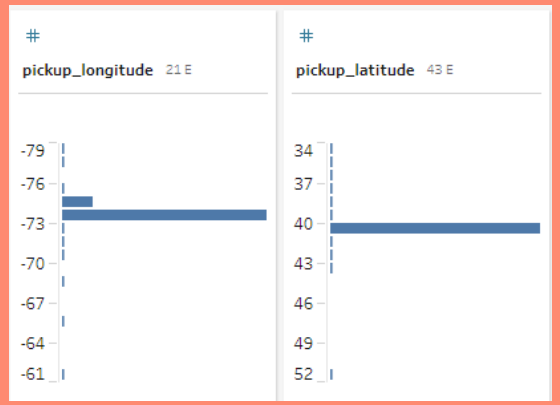
Pickup & drop off coordinates

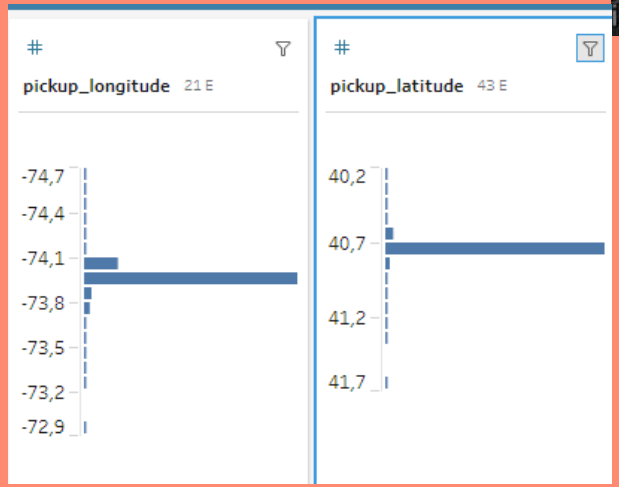
Latitude: 99% is between 40 and 41

Longitude: 99% is between - 73 and – 75.

I exclude the outliers





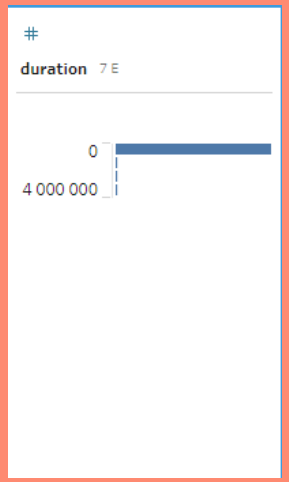


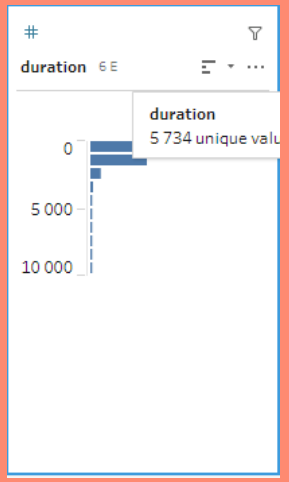
Data analysis process (Tableau Prep Builder)

Duration

I modify the range from 60sec up to 10.000sec (2,7 hour)

Remained: 6k unique value, and the below distribution



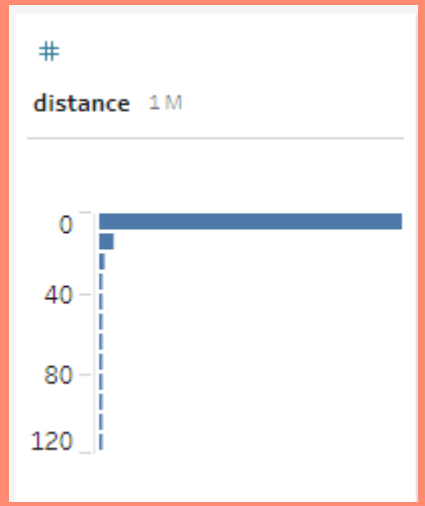


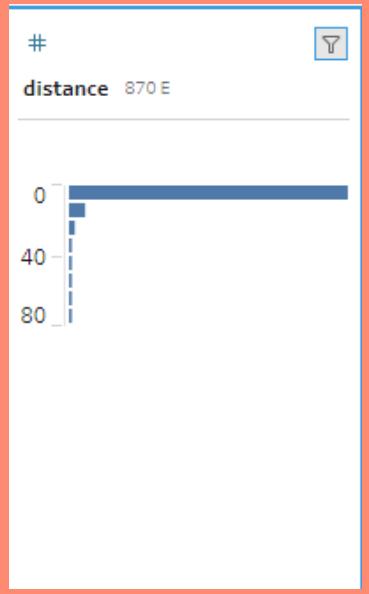
Data analysis process (Tableau Prep Builder)

Distance: 1M unique values from 0miles upto 120miles

Exclude: 0 miles and from 80miles

Distribution is still short trips: from 1-10miles





After Data analysis and cleaning process

Records: 870.558

Ourliers are exluded

Section 4: Enrich the dataset

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`.

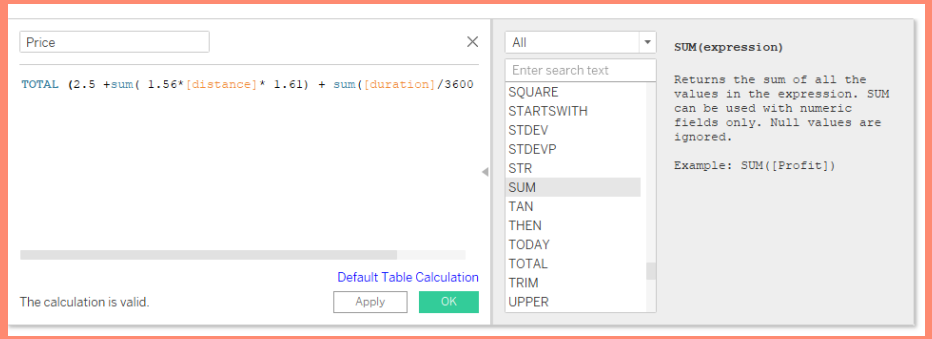
Answer Slide

Create a calculated field called `price` using the `duration`, `distance`, and `passenger count` fields.

Calculation based upon (source): Taxis in New York City have a base fare of $2.50, with an additional 50 cents added every 1/5 mile or minute. So, for a five-mile, 10-minute trip, including a 10% to 20% tip, a taxi would cost around $15 to $18. ➔ This is the simplified version what I use, in reality there are more factors as can be seen in the following study (link), more supportive sources: Quora link, Udacity forum article
price= Base Fare + (Cost per minute * duration) + (Cost per mile * distance) +(passenger_cost*no of passenger) + Booking Fee
Price = (2.5 +(1.56*distance*1.61)+(duration/3600)*30)
Distance is in miles but to convert to km, you multiply by 1.61.

Answer Slide

Create a calculated field called `price` using the `duration`, `distance`, and `passenger count` fields see below. The issue with this calculation is listed on knowledge base post so solution can be: Ride Price = [Fixed Base Fare] +([Cost Per mile] * [Distance in miles]) + ([Cost Per minute] * [Duration in hours]/60)

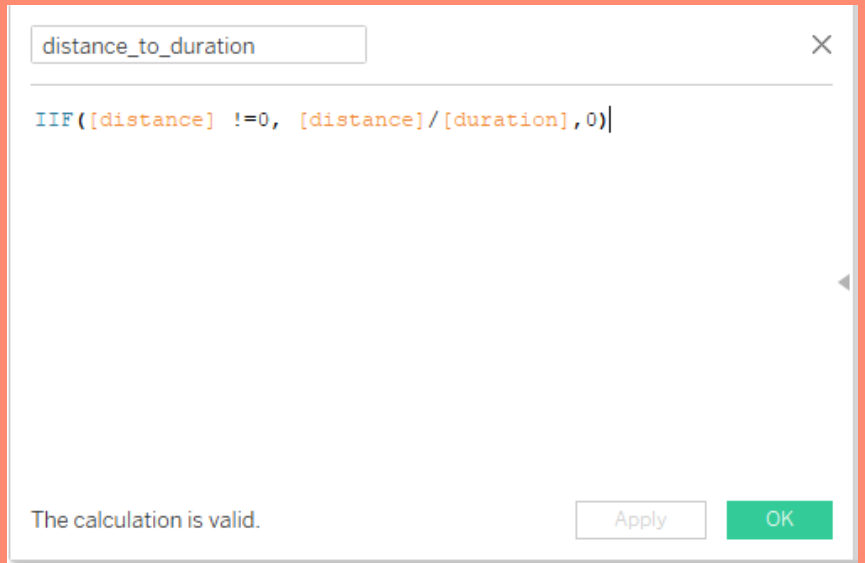


My 1st approach (see above) was wrong, so started to look for answers: Understanding how Tableau calculation types work together

Need simplification: just $2.50 + $0.50 (City Taxes) + $0.50 per 1/5 mile (distance)

Answer Slide

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**`. I used Udacity knowledge article

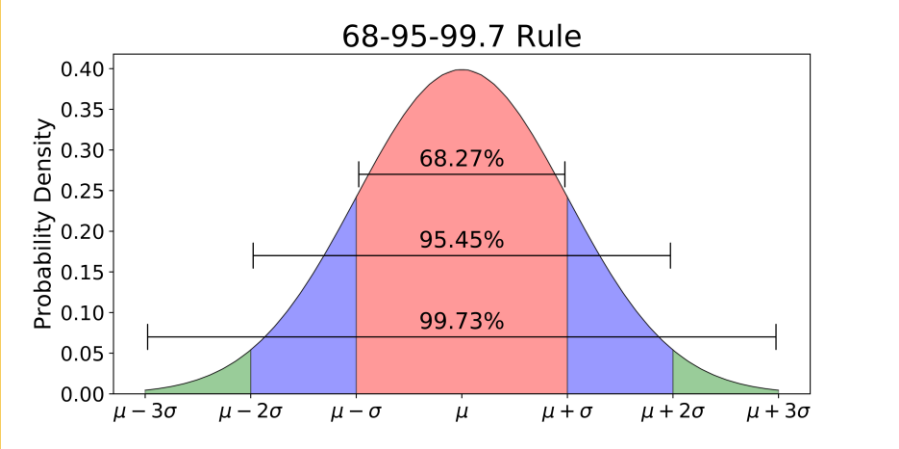


Let’s understand the scope and distribution various dimensions within the dataset. Calculate the **average**, **median**, and the **first & second standard deviation of the mean** for the following measures:

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

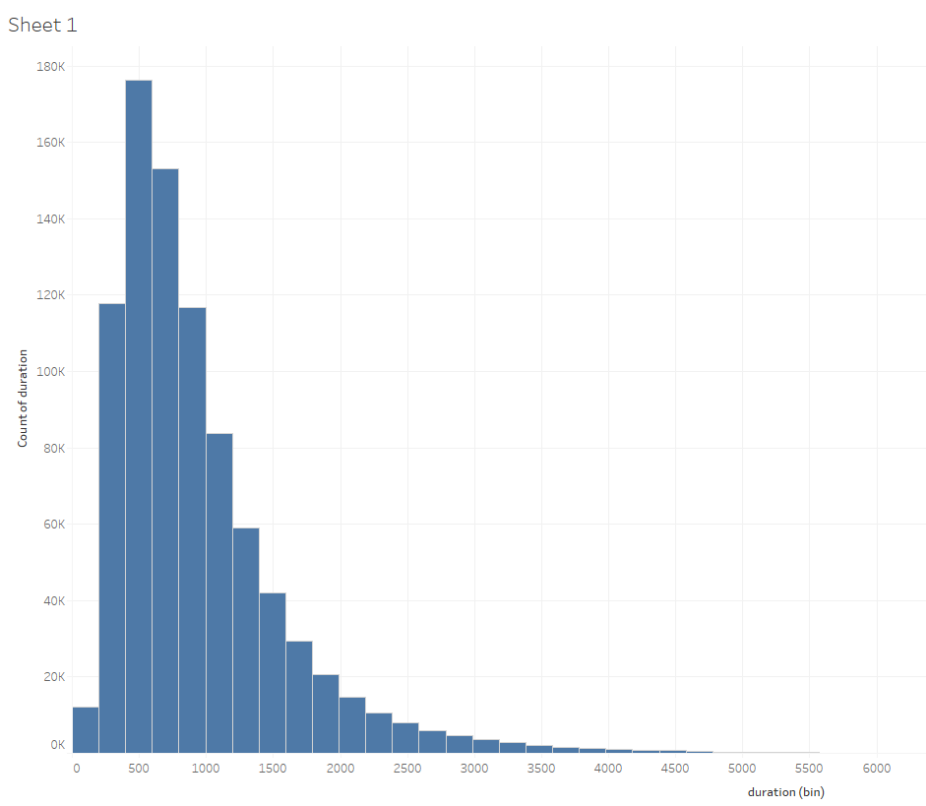
Answer Slide

Support materials: Udacity forum and towards datascience link



Answer Slide

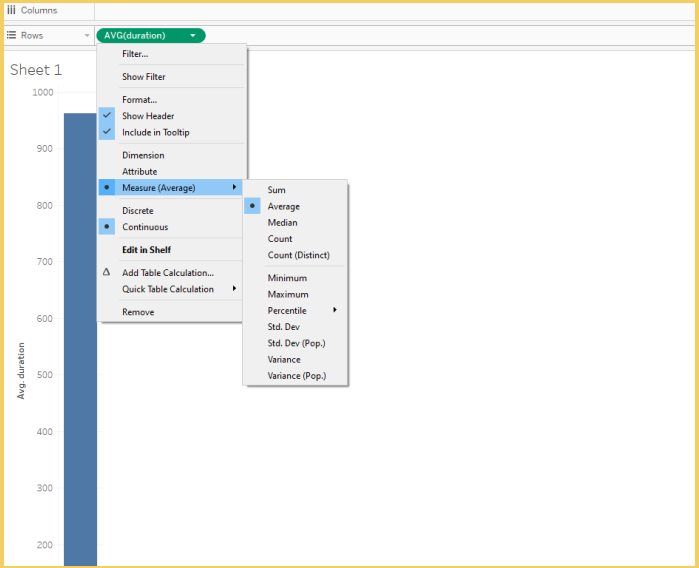
Analyzing duration (unit in seconds)



Highest peak around 399sec – 1250sec

Answer Slide

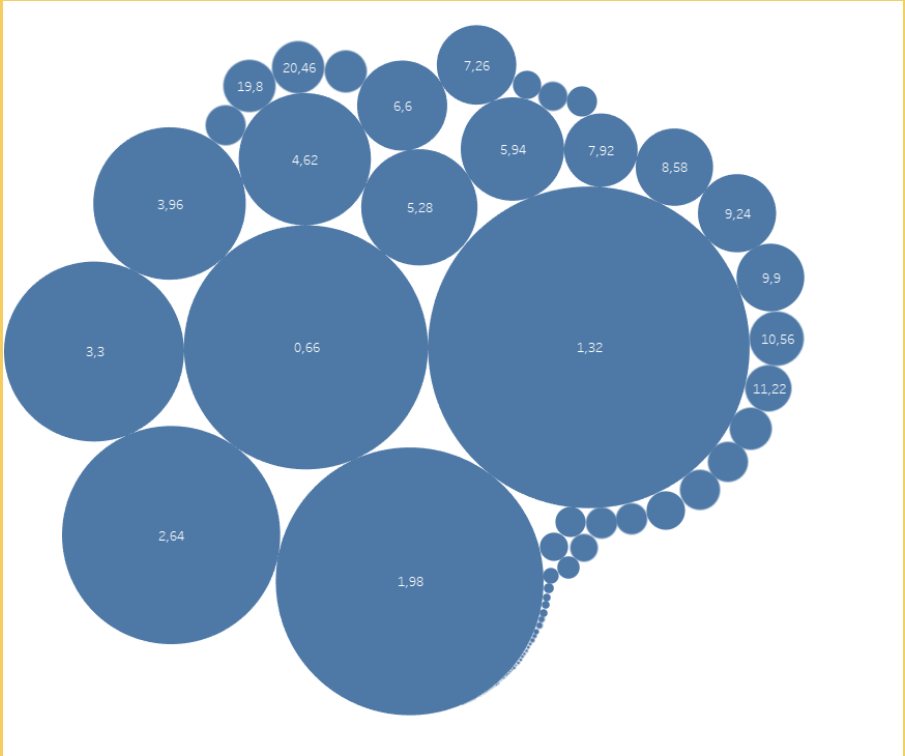
How to do: change measure to avg...



Duration (sec)
Avg: 939,6 sec
Median: 761 sec
Std. Dev: 664,4 sec
2nd std dev (95th percentile): std dev* 2

Answer Slide

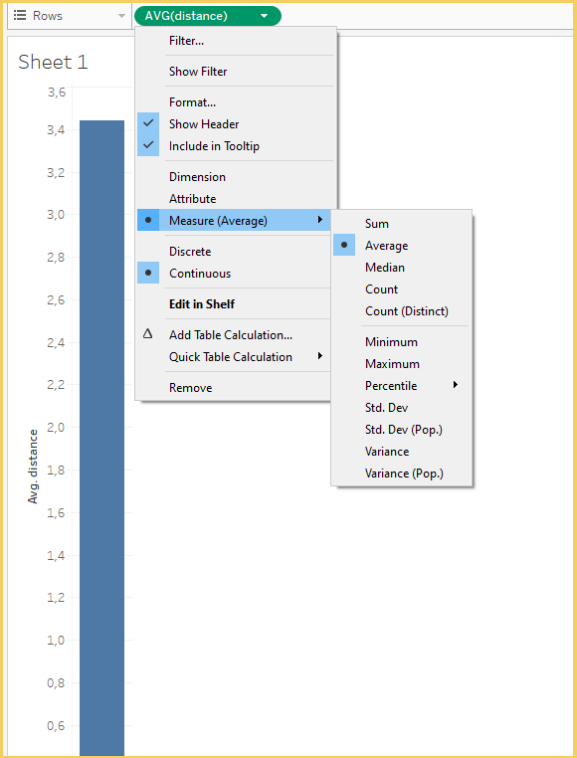
Analyzing distance (unit in miles)



The data distribution around 1.32 miles – 4.62 miles

Answer Slide

How to do: change measure to avg...



Distance
Avg: 4 miles
Median: 2,5 miles
Std. Dev: 4 miles
2nd std dev (95th percentile): 8 miles

Answer Slide

How to do: change measure to avg...

Passanger count
Avg: 1,667
Median: 1,00
Std. Dev: 1,314
2nd std dev (95th percentile): 2,63

Distance-to-duration ratio
Avg: 0.0042 (miles/sec)
Median: 0.0037 miles/sec
Std. Dev: 0.0021 miles/sec
2nd std dev (95th percentile): 0.0042

Answer Slide

How to do: change measure to avg...

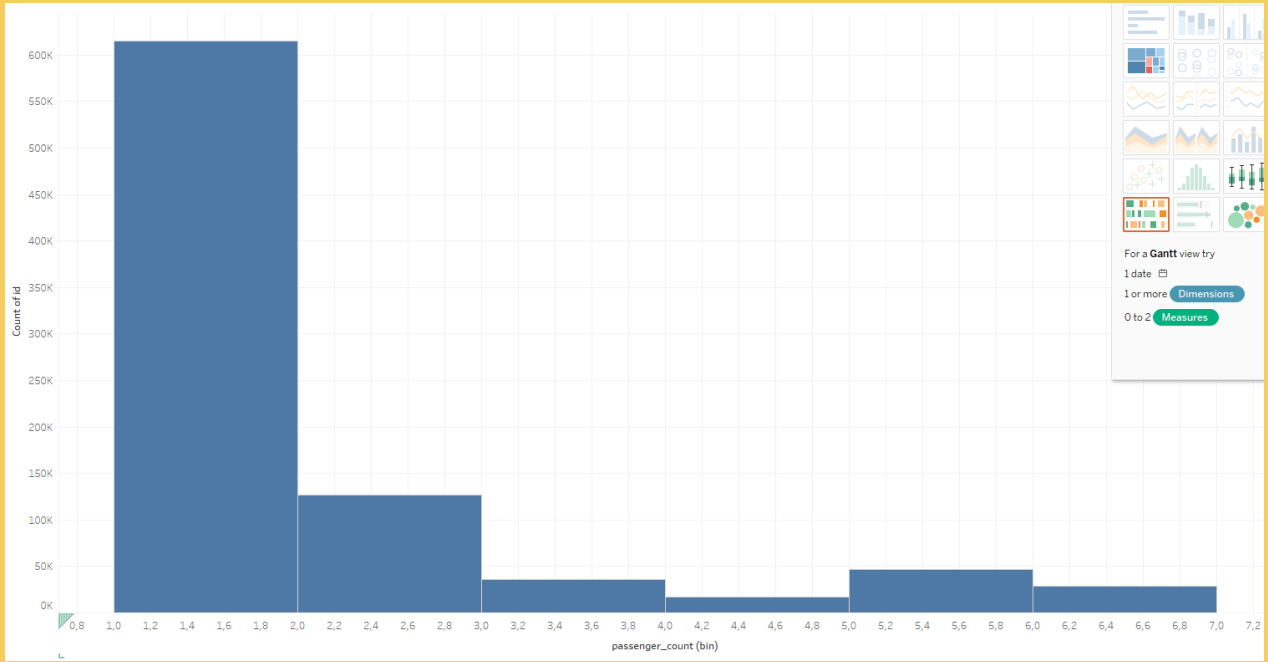
Price
Avg: 13
Median: 9,26
Std. Dev: 10.25
2nd std dev (95th percentile): 20,50

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

For 1 passanger: 615k trips, for 2 passangers: 127k trips



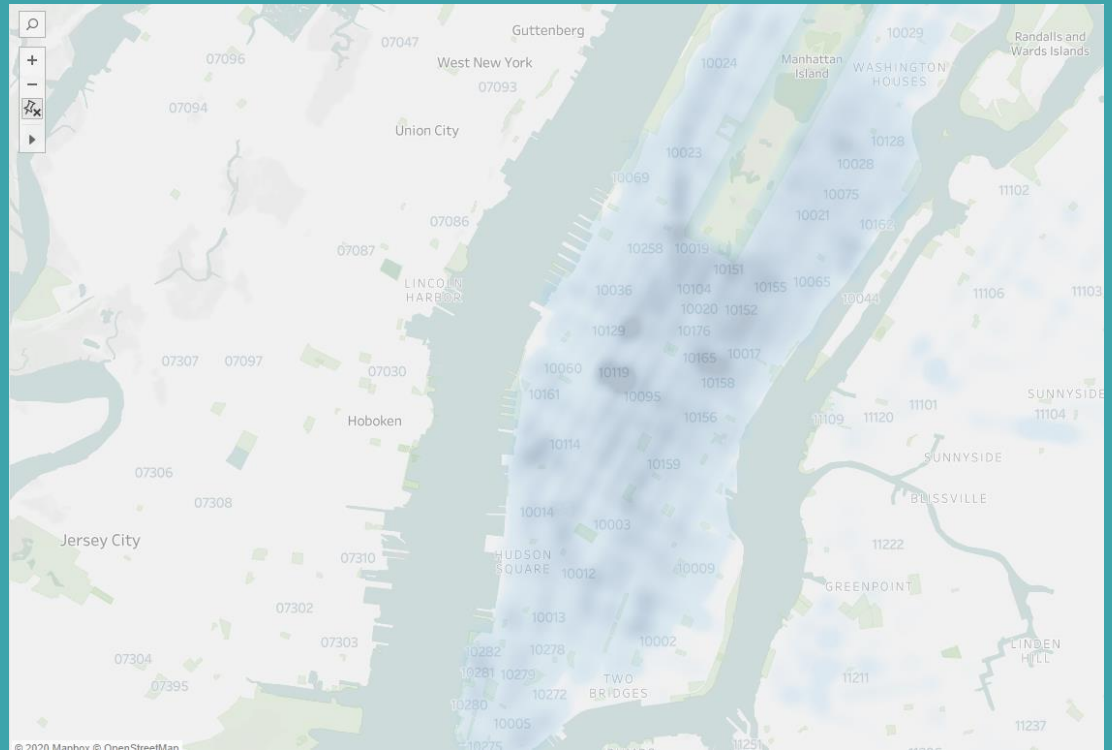
Section 5: MVP set-up

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?

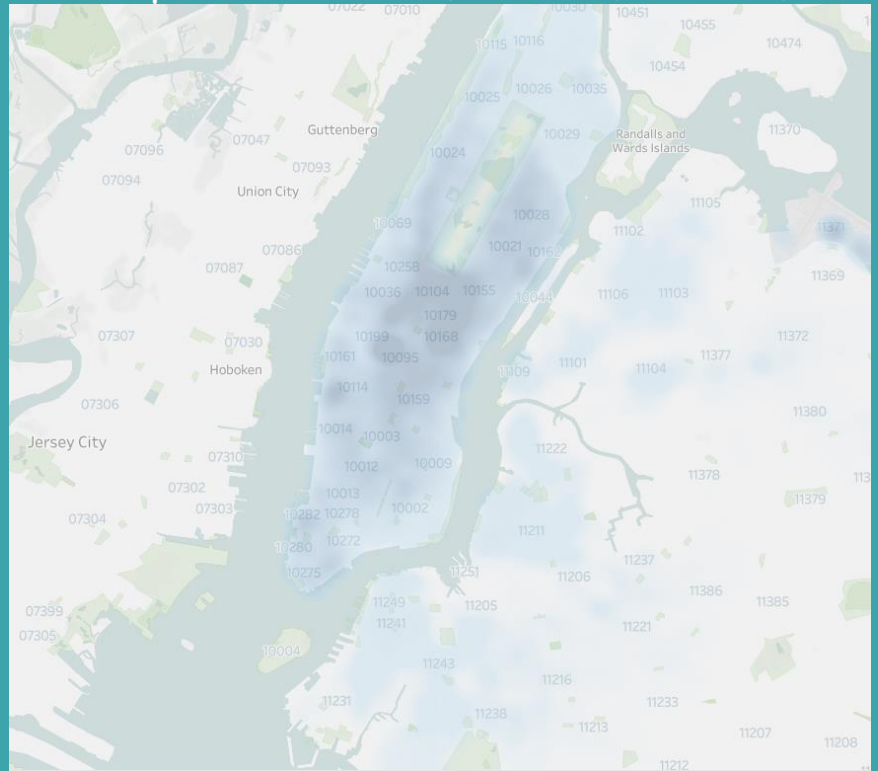
Answer Slide

- Which neighborhoods/zip codes tends to experience a relatively higher density of **pick-ups**: 10119, 10129, 10165, 10104, 10151, 10019, 10023



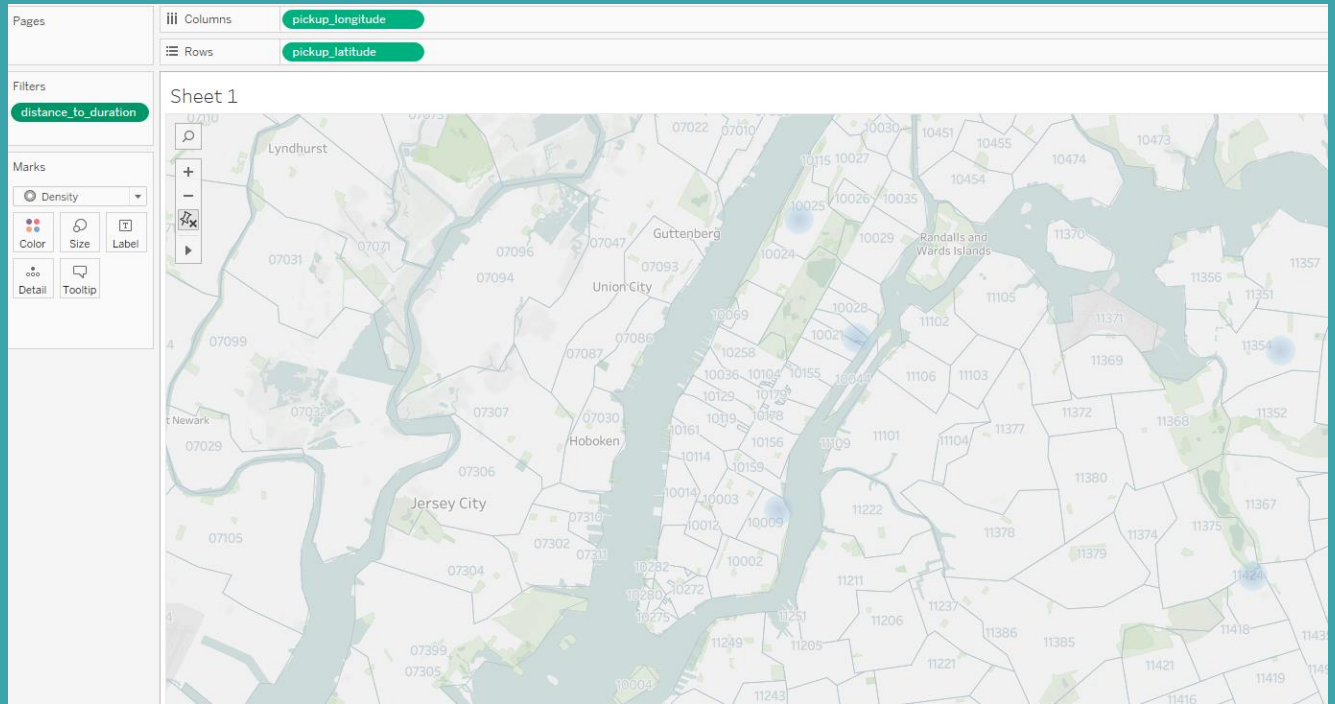
Answer Slide

- Which neighborhoods/zip codes tends to experience a relatively higher density of drop-offs: 10095, 10282, 10168, 10104, 10021, 10179, 10159



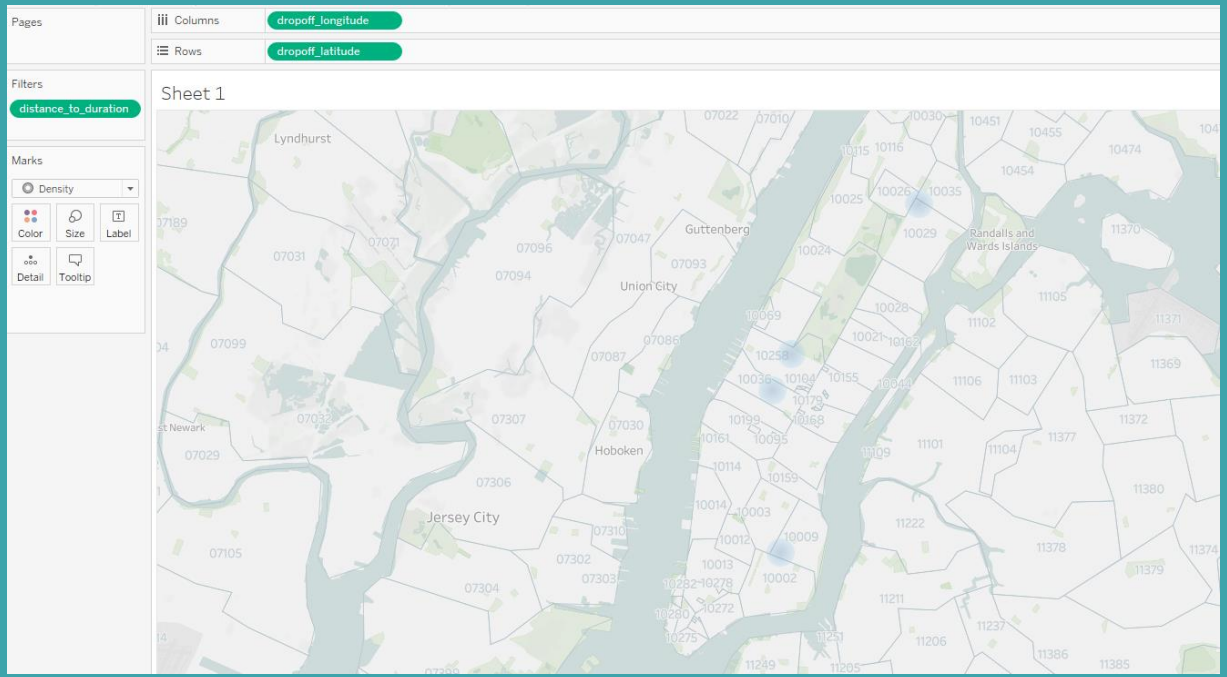
Answer Slide

- Which neighborhoods/zip codes tends to have the highest duration-to-distance ratios, based on pick-up? Applied the high range of dictance_to_duration ratio: 10009, 10021, 10025



Answer Slide

- Which neighborhoods/zip codes tends to have the highest duration-to-distance ratios, based on drop-off? Applied the high range of dictance_to_duration ratio: 10009, 10036, 10258, 10029



Answer Slide

- 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?

  Based on the previous analysis I advise the 10009 for the south part and 10025 for the north part as building nodes.

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

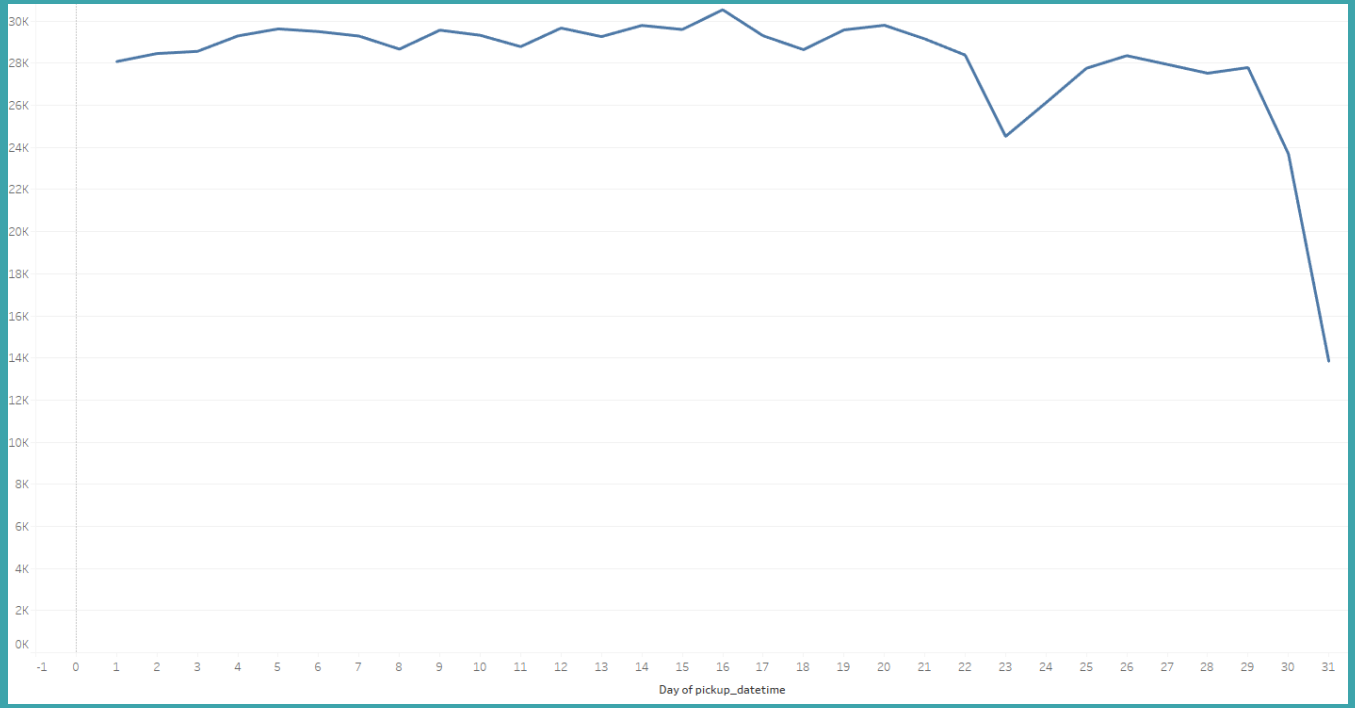
- What times throughout the day experience relatively higher volumes of ride pick-ups?
- What days throughout the week experience relatively higher volumes of ride pick-ups?
- Pinpoint any periods throughout the year that experience trend fluctuation or seasonality around ride pick-up volumes. This will 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

What times throughout the day experience relatively higher volumes of ride pick-ups? As expected in the morning and after working hours

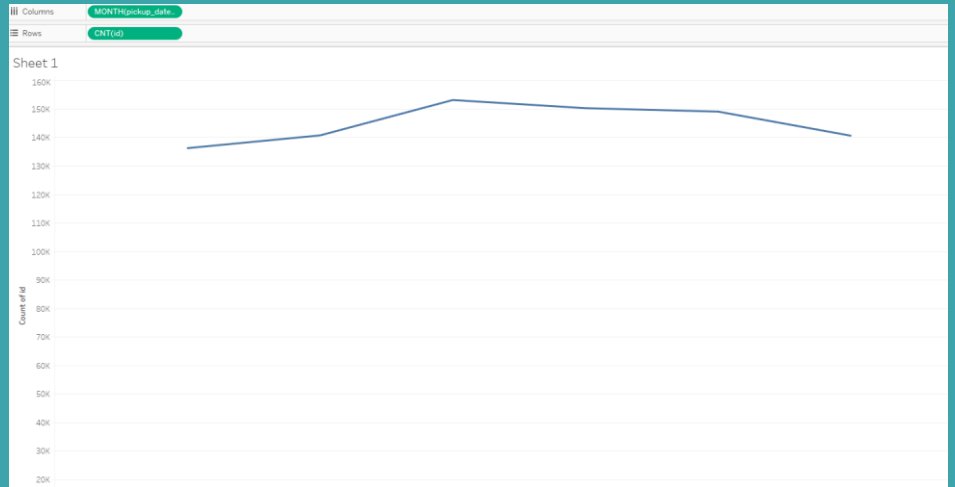
Answer Slide

What days throughout the week experience relatively higher volumes of ride pick-ups? The beg. and middle of the months are strong



Answer Slide

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.
Our dataset is limited only for a half year so quite hard to find seasonality... Spring time is stronger



Section 5: MVP set-up, validate your hypothesis by user research

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:

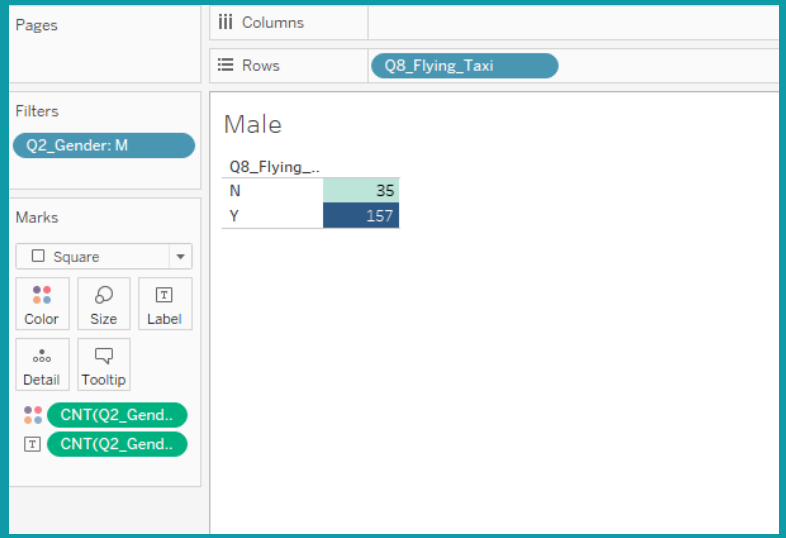
Q1 - What is your email?
Q2 - What gender do you identify as?
Q3 - What is your age?
Q4 - What is your annual income? (income bands)
Q5 - What neighborhood do you reside in?
Q6 - Do you currently use taxis? (Y/N)
Q7 - Do you currently use ridesharing services? (Y/N)
Q8 - Would you use a flying taxi service, 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?

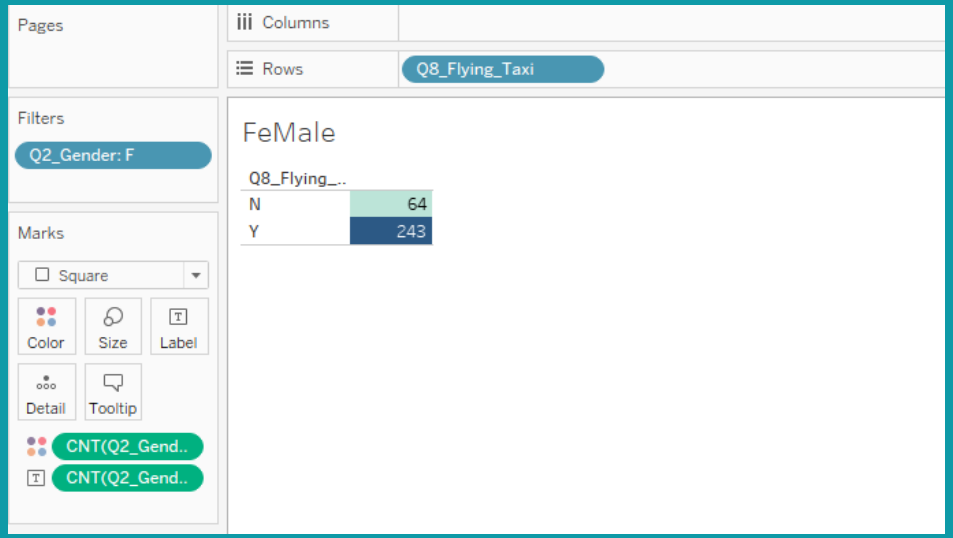
To inform our future product marketing efforts, we'll want to extract the following:

- 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

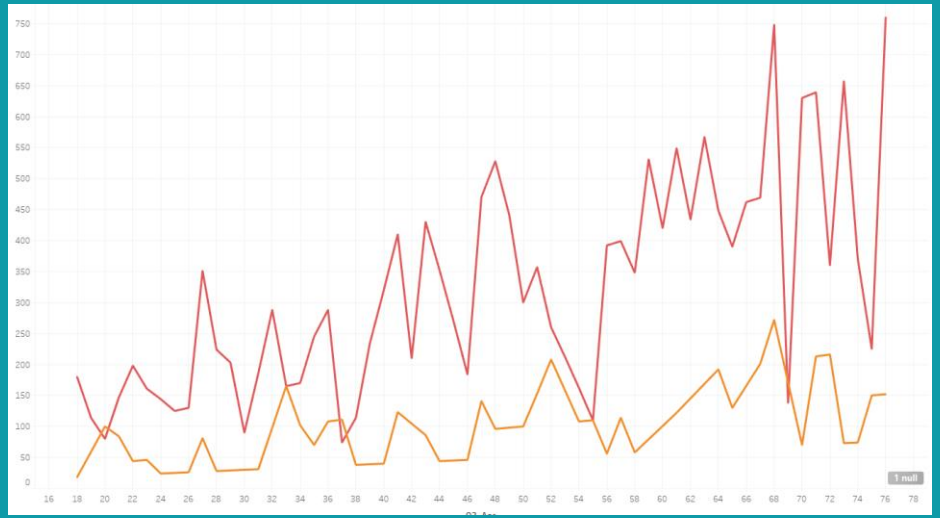
- Is there an inclination of better Flyber adoption based on gender?
  For this analysis need: Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N) vs Q2 - What gender do you identify as? Male: 157yes (82%), Female: 243 yes (79%)

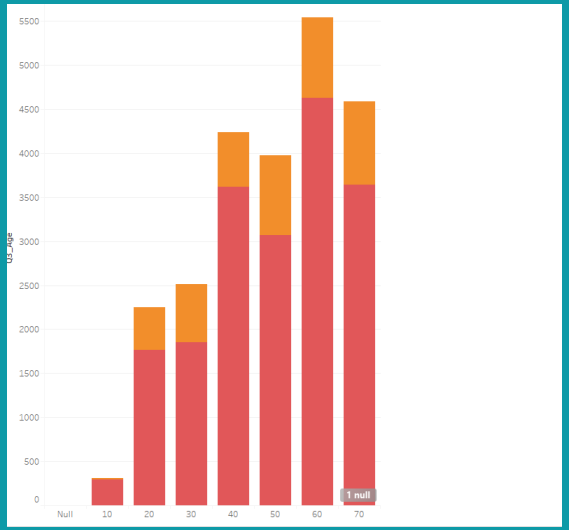




Answer Slide

- Is there an inclination of better Flyber adoption based on age? For this analysis need: Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N) vs Q3 - What is your age? Analysis: I suprised as old generation was open to the service as well (created bins).





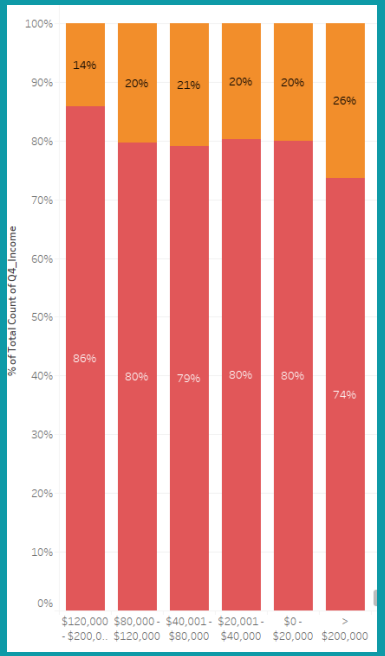
Answer Slide

- Is there an inclination of better Flyber adoption based on income level?

  For this analysis need: Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N) vs Q4 - What is your annual income? (income bands)

  * support article for percentage calculation: link and put on Tableau

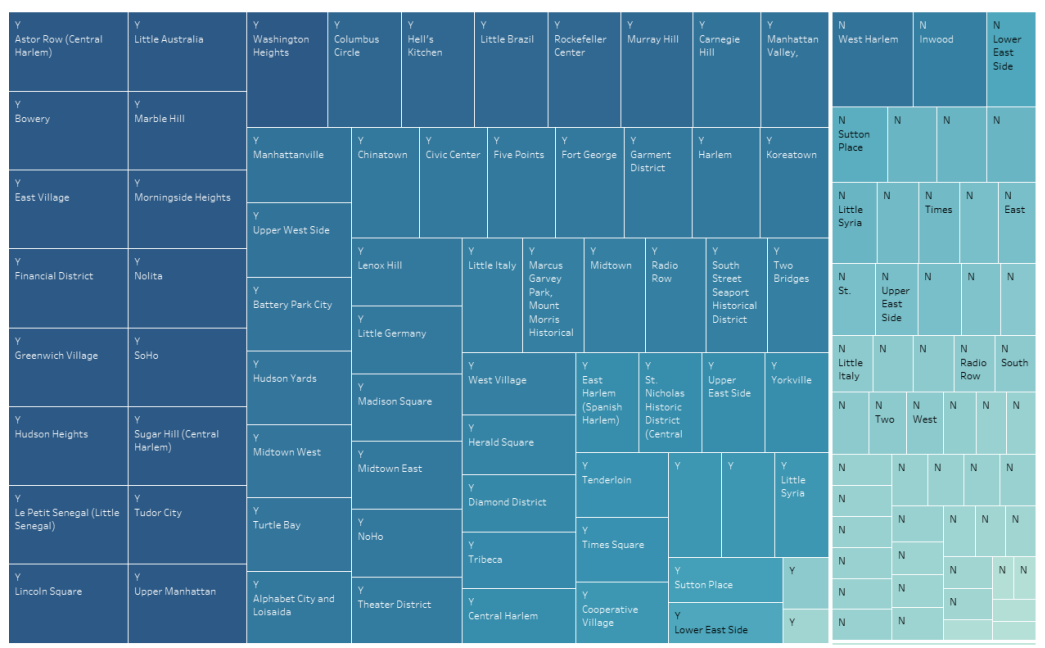
  can be seen the $120k-200k income range is more open



Answer Slide

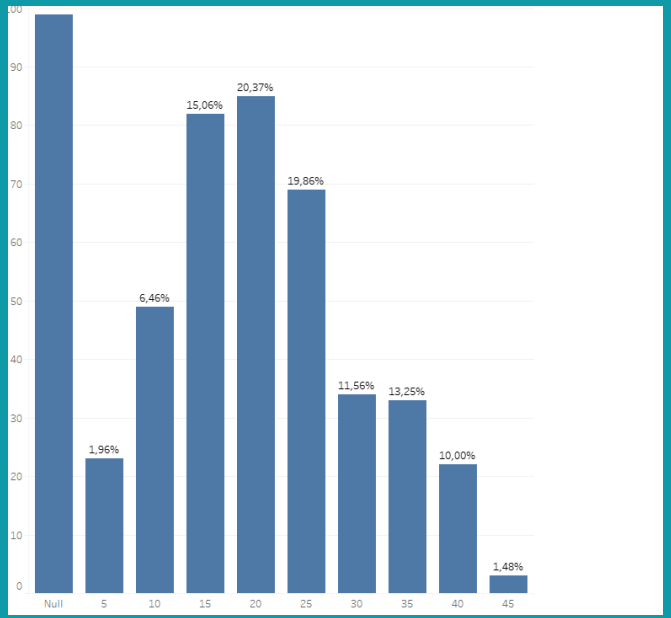
- Is there an inclination of better Flyber adoption based on neighborhood of residence?

  For this analysis need: Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N) vs Q5 - What neighborhood do you reside in? Dark Blue where the highest % adaption Tableau Public



Answer Slide

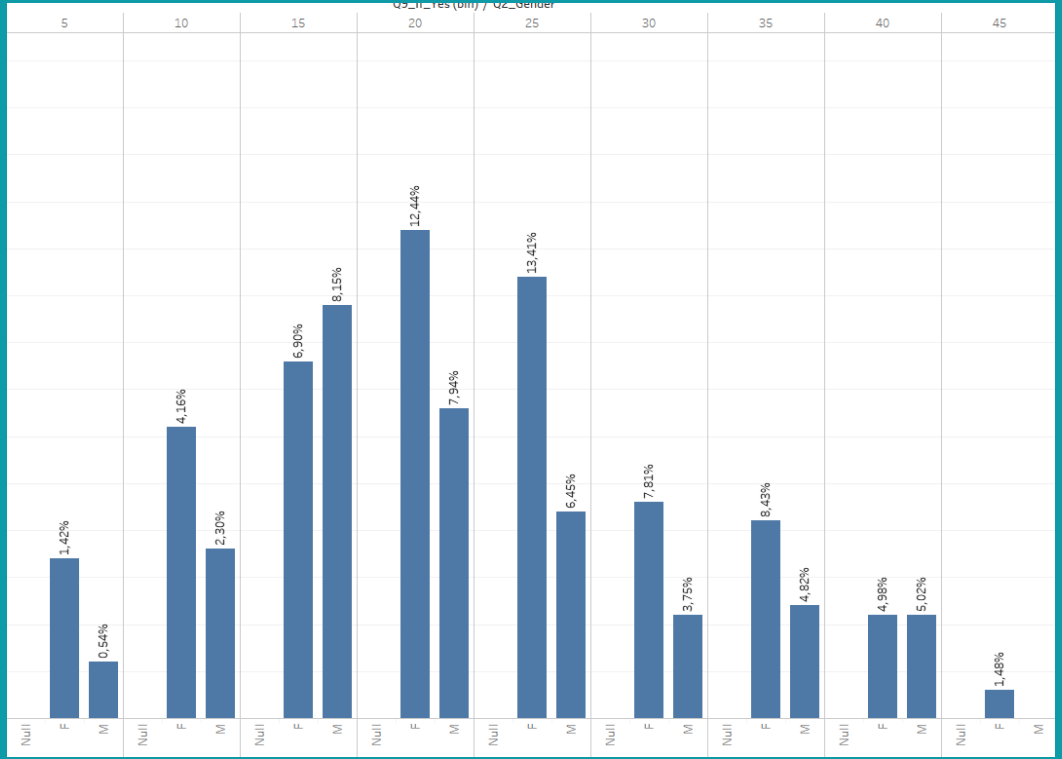
What is the distribution of potential price per mile based on gender, age, income level, and neighborhood of residence? Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N)
Q9 - If yes to Q8, how much would you be willing to pay per mile for such a service? (USD) / Q2-Q5



Price distribution (by bins): $20-$25

Answer Slide

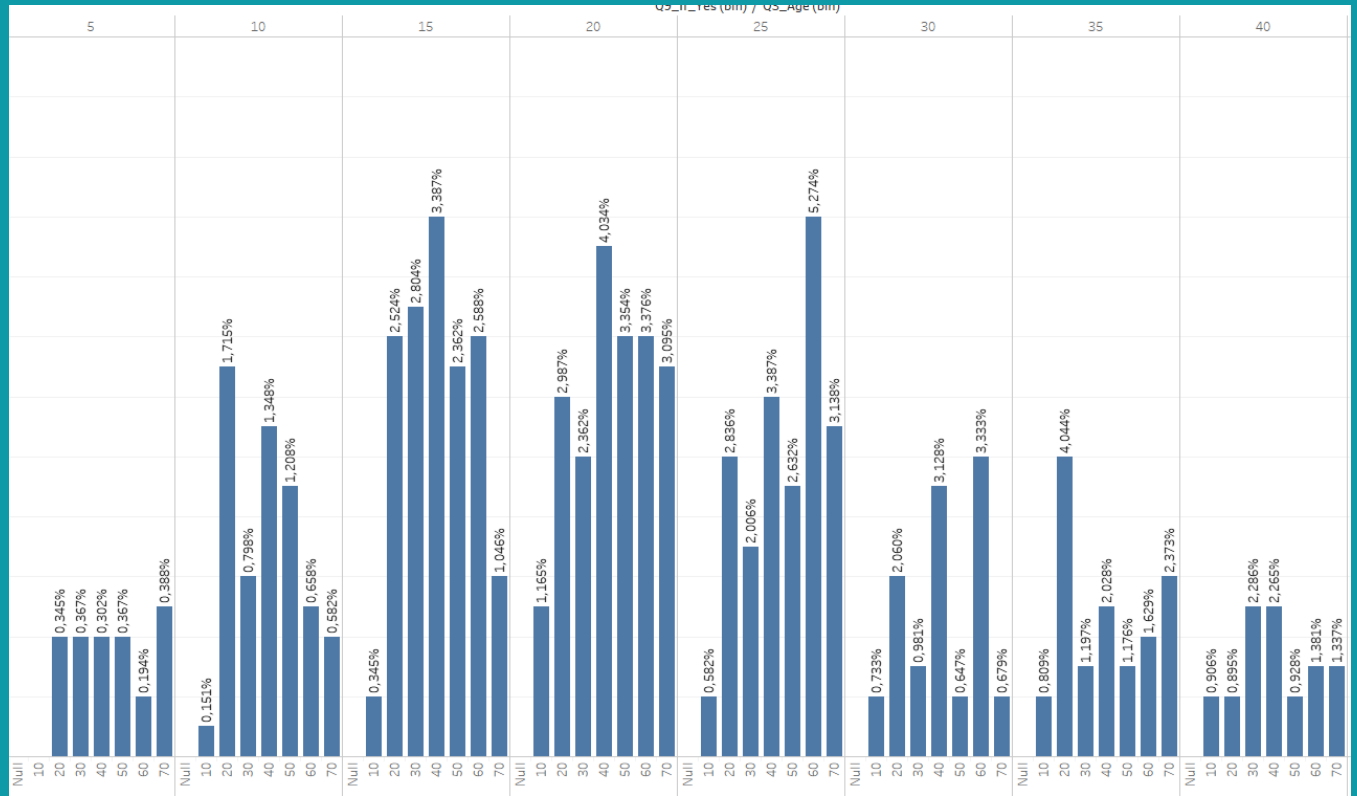
What is the distribution of potential price per mile based on gender?



Females are ready to pay more

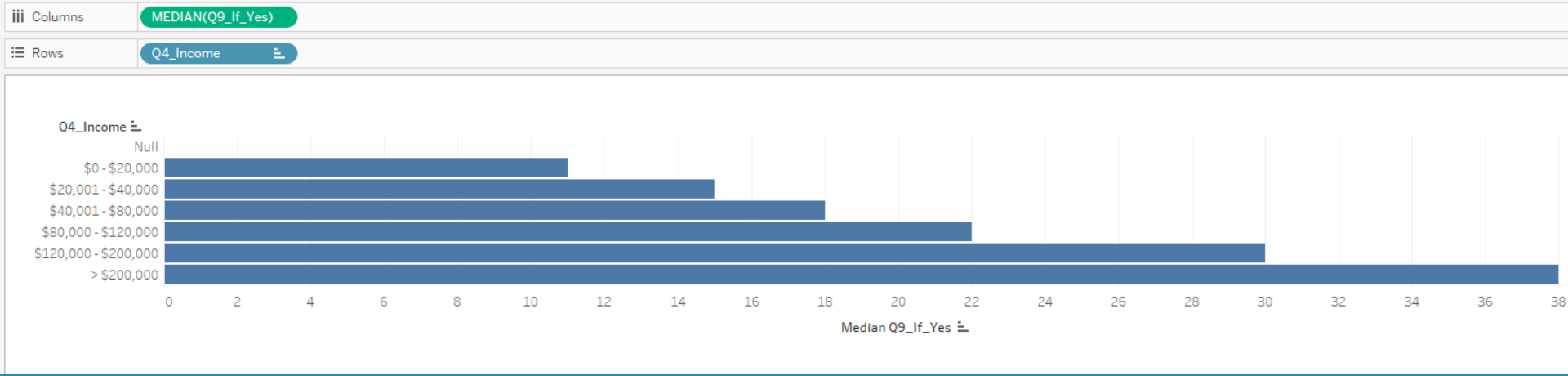
Answer Slide

What is the distribution of potential price per mile based on age?



Answer Slide

What is the distribution of potential price per mile based on income level?

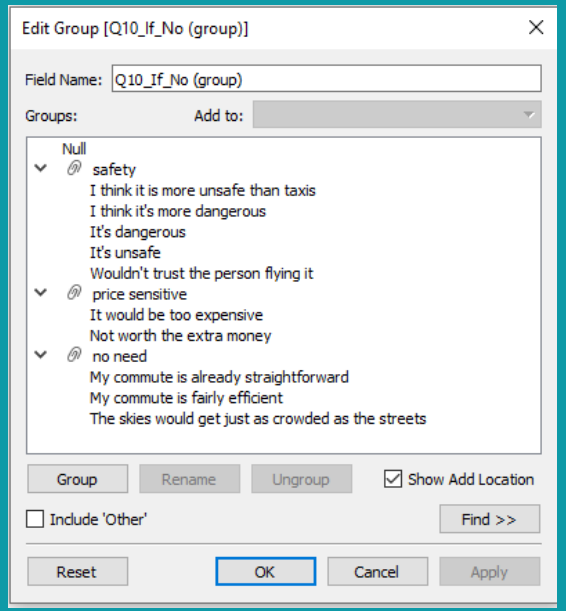


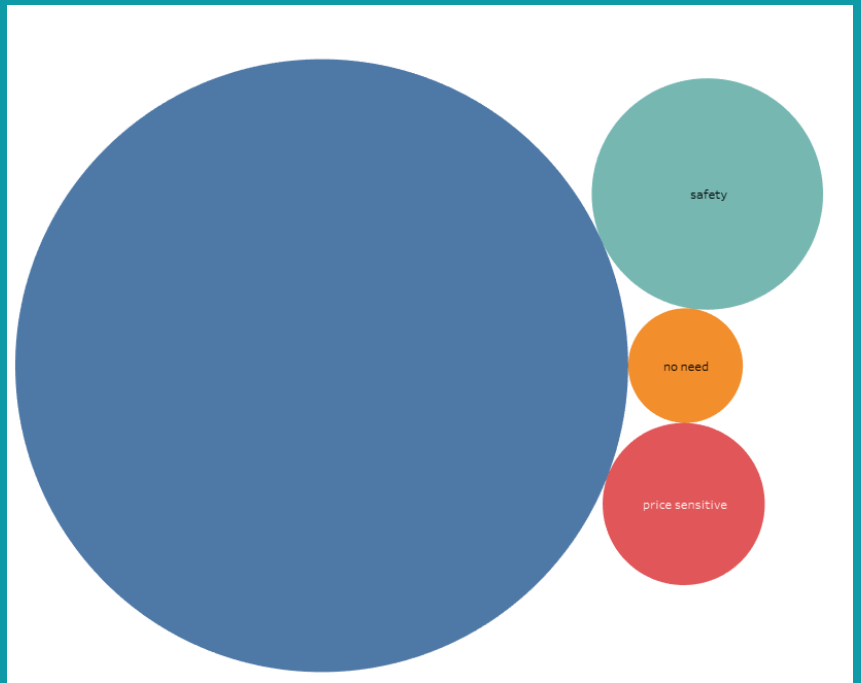
Answer Slide

What is the distribution of potential price per mile based on neighborhood of residence?

Answer Slide

What is the different personas/segments of negative sentiment towards not using a flying taxi car service?
1st I have created groups of negative sentiments, after compared with all answers





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

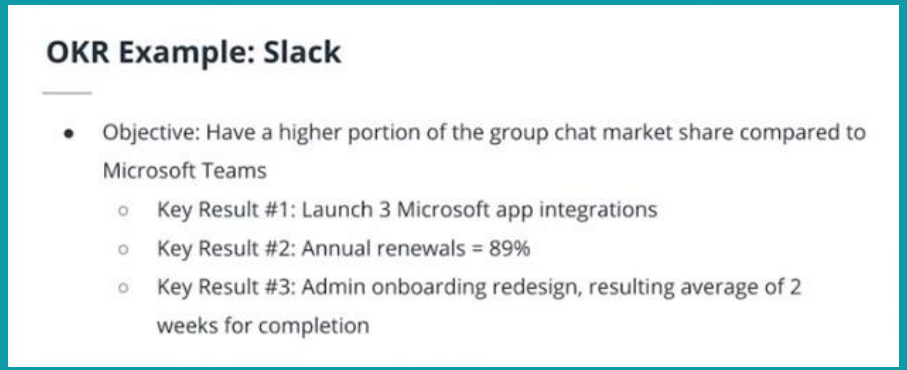
Support: Udacity lesson

Product vision - (or North Star) the ultimate goal of where you want your users to do or accomplish with your product.

Objectives and Key Results (OKR) Framework - goal setting methodology to define & track the progress towards objectives

Objectives - clearly defined goal at a high-level. Could be both quantitative or qualitative.

Key Result - metric targets that directly influence the success of meeting our objective. Could be both quantitative or qualitative, but usually the former. Multiple key results could roll up to one objective.



Answer Slide | User Engagement

Objective: Provide an ultimate flying experience; high-end & safe for flyber early user to feel safe and privileged.

Q: 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.

A: The data analysis and user research analysis detailed on the previous slides showed that the MVP product (service) needs to be concentrated on some high-traffic points, user fears and prejudices need to be addressed.
I believe that **early users** for whom we need to focus on and provide the highest customer experience. They will spread the word have a key role in evangelization the service (example is Tesla go-to-market strategy).

Answer Slide | User Engagement

Objective: Provide an ultimate flying experience; high-end & safe for flyber early user to feel safe and privileged.

(KR1) number of repeated trips in proportion to users 36%
(KR2) The NPS (or Net Promoter Score) to be above 82

KPI: DAU / MAU - a ratio effective at measuring the rate of stickiness, for products that are meant to be used daily. Calculation: (#) Daily active users / (#) Monthly active users = (%) DAU/MAU Ratio
Industry Benchmarks (link): Sequoia tweeted the standard DAU/MAU ratio is 10-20% with only a handful of companies over 50%.

KPI: Average NPS of all flying [How likely is it that you would recommend [Organization X/Product Y/Service Z] to a friend or colleague?]

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

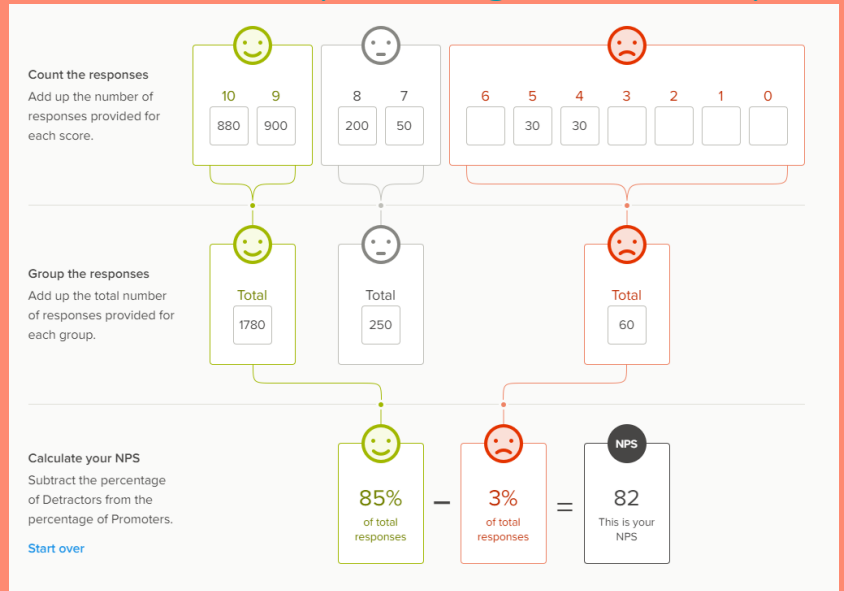
Answer Slide

KPI 1: DAU / MAU - a ratio effective at measuring the rate of stickiness, for products that are meant to be used daily.
DAU target: 2000 users in MVP phase (6 month)
MAU: 5500 users in MVP phase
ratio: 36%

* DAU/MAU ratio of 50% basically means that the average user of our service uses it 15 days from 30. DAU/MAU ratio of < 1% means that all new coming users try our service just once and do not return.

Answer Slide

KPI 2: Average NPS of all flying
target: 82 with 2000 DAU
calculation: https://delighted.com/nps-calculator



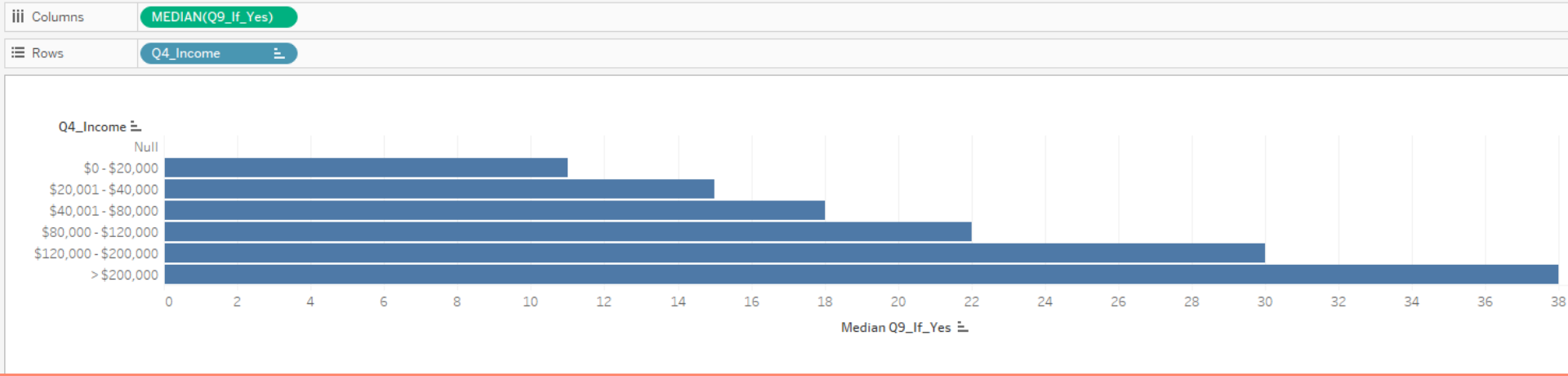
Answer Slide

KPI 3: Monthly recurring revenue (MRR) - total amount of estimated revenue that a company will receive on a monthly basis.
Calculation: MRRt = ∑ Recurring Revenuet
Taking into consideration: DAU: 2000, average distance: 4miles, my target early users price budget: 30$ (see below)
Target MRR in MVP phase: 3.6M$

* calculation: https://www.cobloom.com/blog/saas-metrics

Answer Slide

What is the distribution of potential price per mile based on income level?

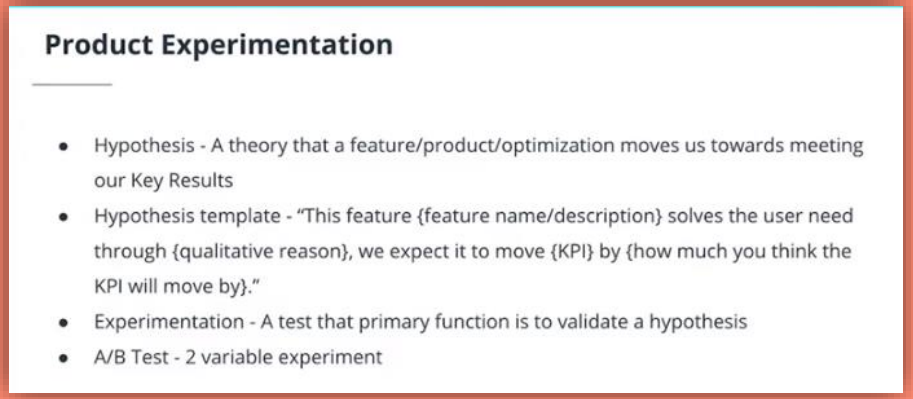


Create hypotheses around what thresholds your KPIs would need to hit in order to determine success

Answer Slide: Product Experimantation

To solve this task I used Udaxity lecture

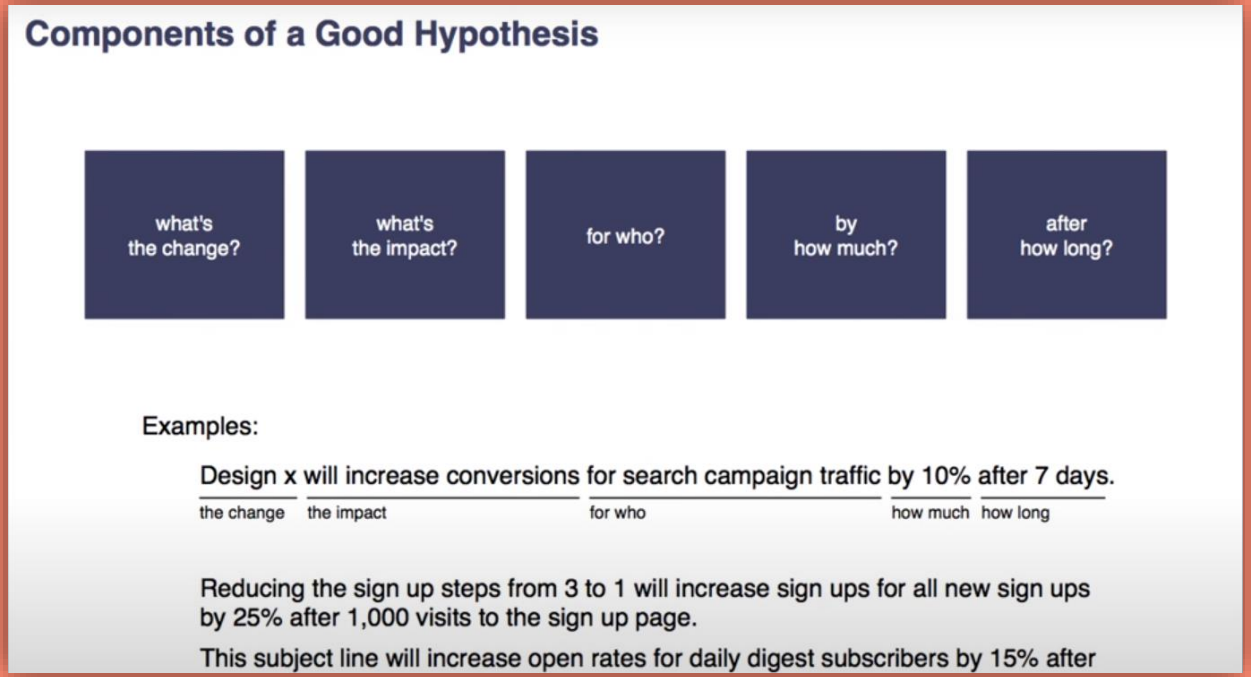
Hypothesis template: [Feature name/description] solves the user need through [qualititive reason], we expect it to move [KPI] by {how much you think the KPI will move by}.



Also used: **The Product Experimentation Playbook by Optimizly**

Answer Slide: Product Experimantation

Another approach by Product Talk: [link]



Answer Slide

Hypotheses I.: Add flying safety features like safety engine solves the user need through reducing fear from flying services we expect it to move DAU / MAU KPI by 5%.

Answer Slide

Hypotheses II.: Add luxury features like 5G connection, laptop workstation solves the user need through enjoy her time we expect it to move up NPS KPI by 5% .

Answer Slide

Hypotheses III.: Add early-booking features through mobile app solves the user need through giving control we expect it to move up MRR KPI by 12%.

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

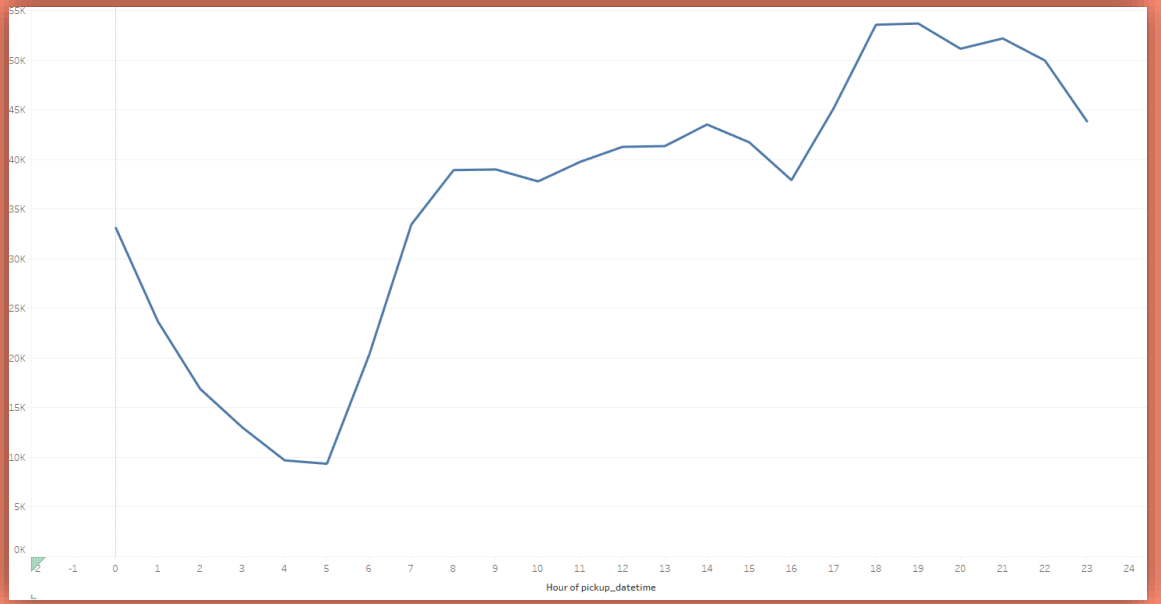
Q: What times/days of operation should the service run for?

A: Based in user research data analysis (see below)
In the morning: 8am-3pm there is a significant demand, and rush hours starts at 7pm till 9pm.

Conclusion: The Flying service should focus for this time window

Answer Slide

What times throughout the day experience relatively higher volumes of ride pick-ups? As expected in the morning and after working hours



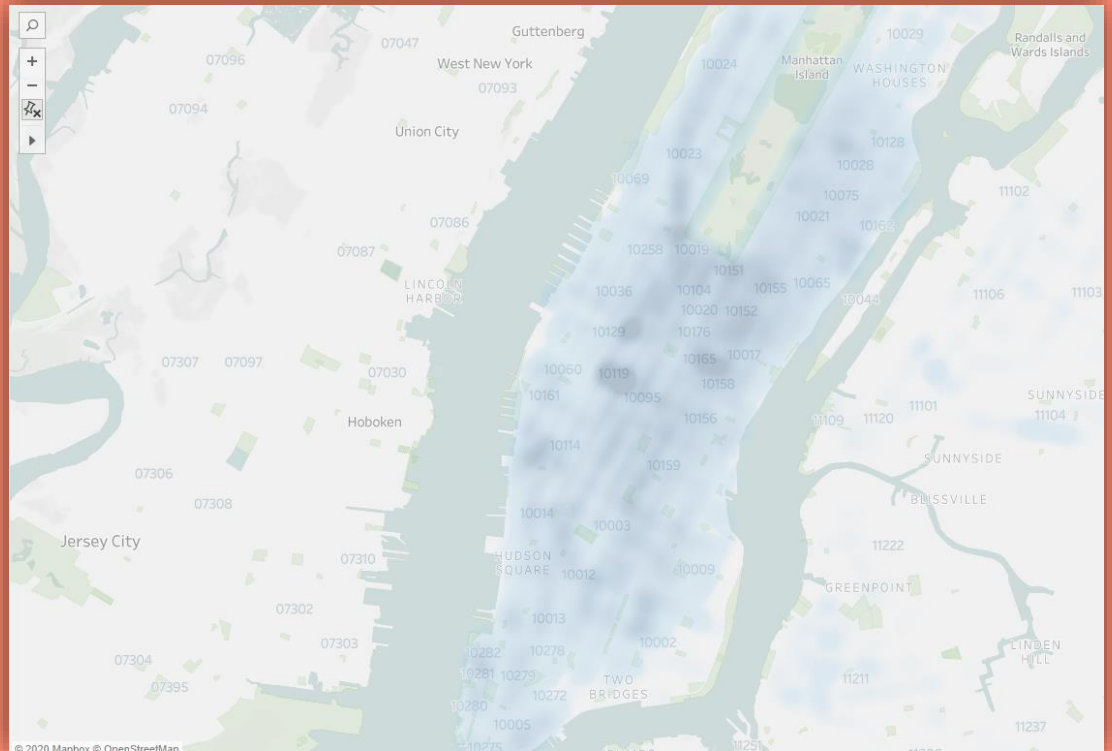
Answer Slide

Q: How many pick-up / drop-off nodes should we have?
Where should the nodes be located?

A: I adapt learning on NewYork taxi dataset analysis and select the minimum nr. of station: I advise the 10009 for the south part and 10025 for the north part as building nodes.

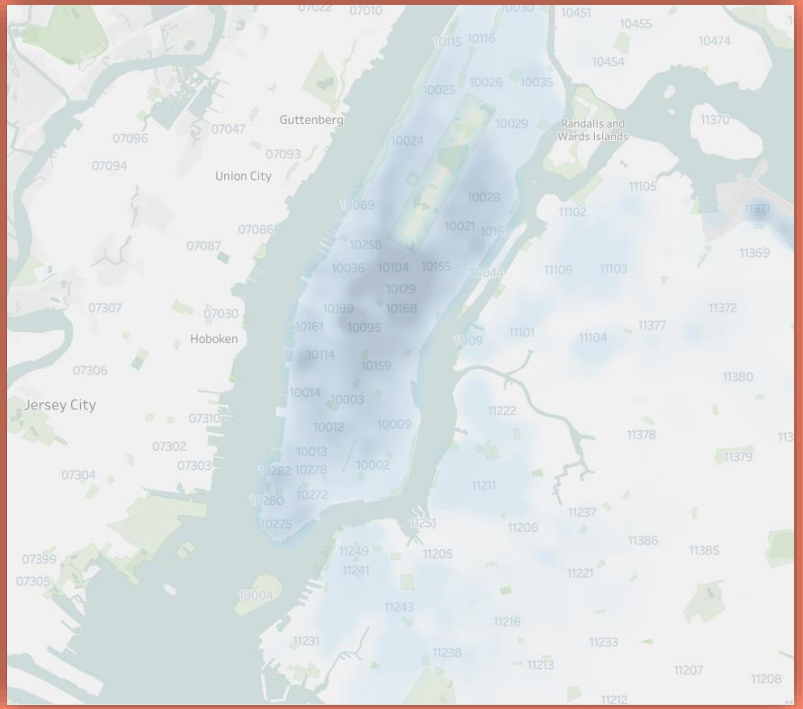
Answer Slide

Which neighborhoods/zip codes tends to experience a relatively higher density of **pick-ups**: 10119, 10129, 10165, 10104, 10151, 10019, 10023



Answer Slide

Which neighborhoods/zip codes tends to experience a relatively higher density of drop-offs: 10095, 10282, 10168, 10104, 10021, 10179, 10159



Answer Slide

Q: Should we initially use copters or homegrown hardware?

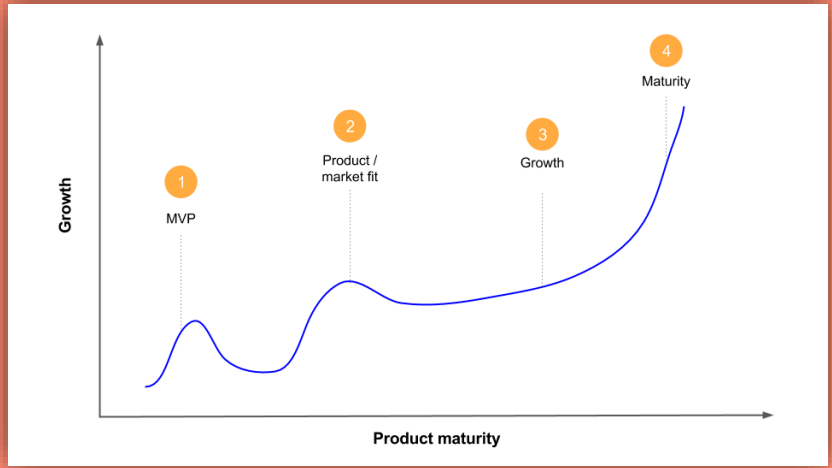
A: As our go-to-market plan is focusing on early users for whom we need to provide the highest customer experience, I advise homegrown (customized) hardware.

Answer Slide

Q: Should the pricing be fixed or dynamic? At what rates?

A: **Monetization Strategy**
Flyber service is in MVP status so monetization should be aligned to product lifecycle. We are in MVP phase.

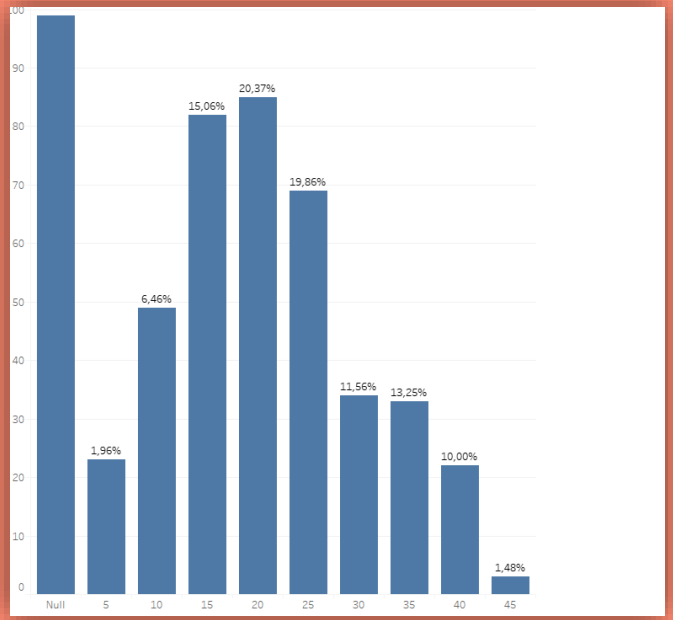
As our user engagement objective is focusing on early users with high user experience we start our service for that user segments who has high financial resources and willingness to use the service. Based on user research analysis: it is **$120k-200k income range and targeting 30$**



Source: link

Answer Slide

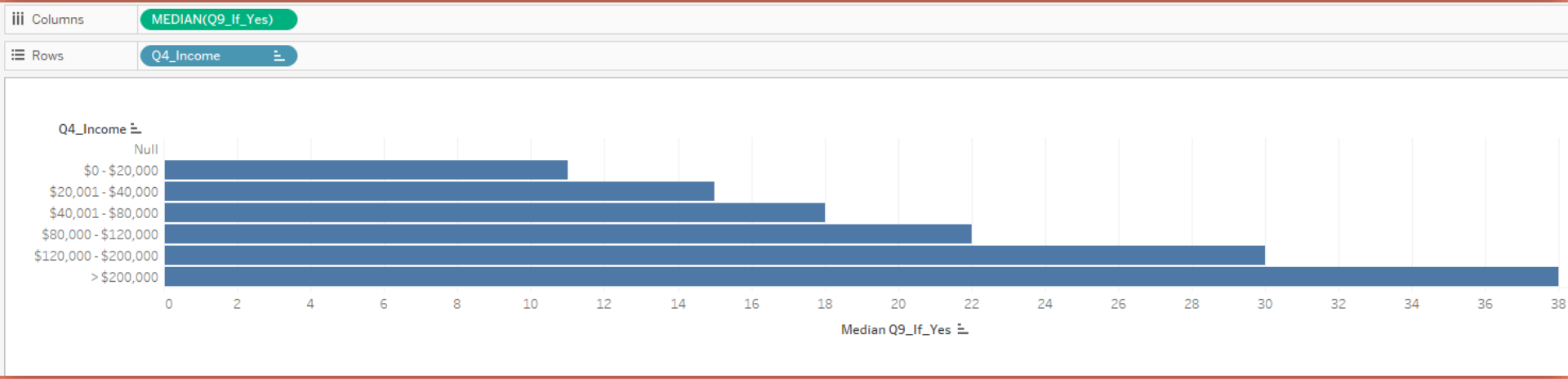
What is the distribution of potential price per mile based on gender, age, income level, and neighborhood of residence? Q8 - Would you use a flying taxi service, if such a concept existed? (Y/N)
Q9 - If yes to Q8, how much would you be willing to pay per mile for such a service? (USD) / Q2-Q5



Price distribution (by bins): $20-$25

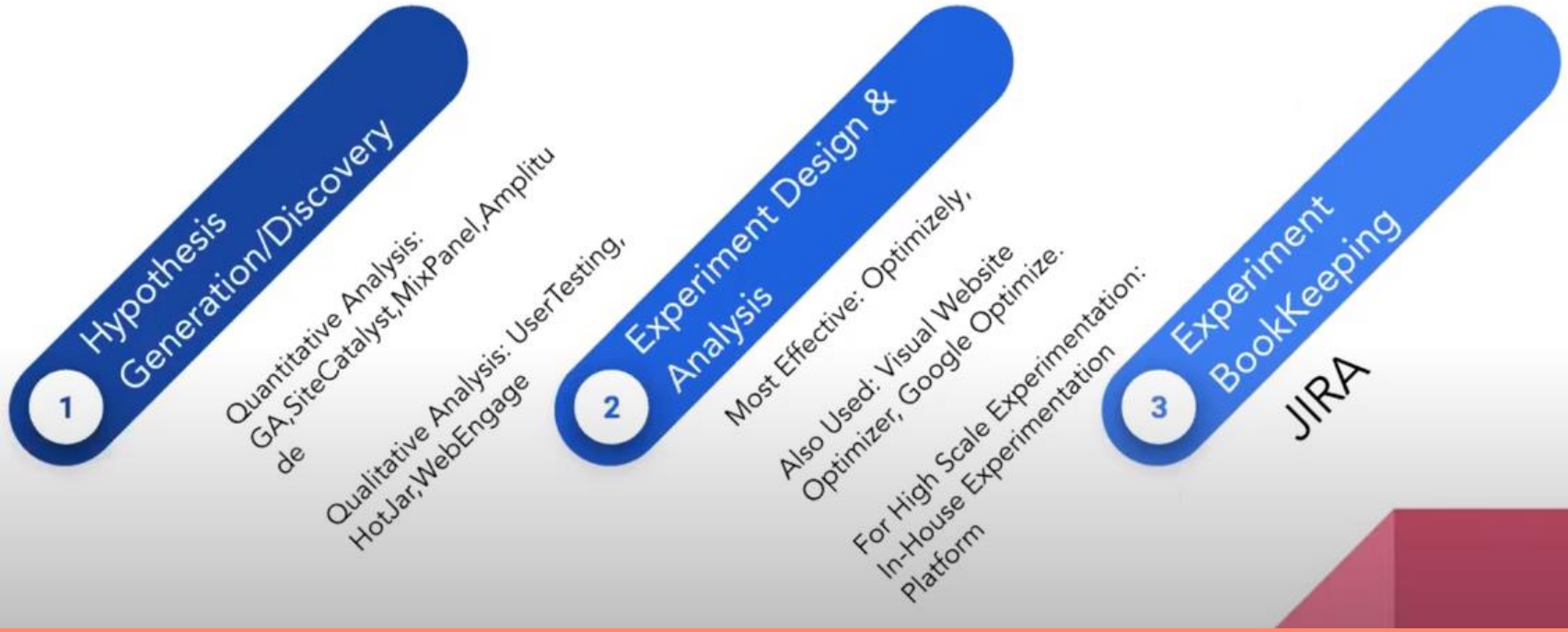
Answer Slide

What is the distribution of potential price per mile based on income level?



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

Tools Used at Different Stages of Experimentation



Source: Webinar: [Experimentation & PM Part II by Indeed Product Lead, Abhishek Rathore]

Answer Slide

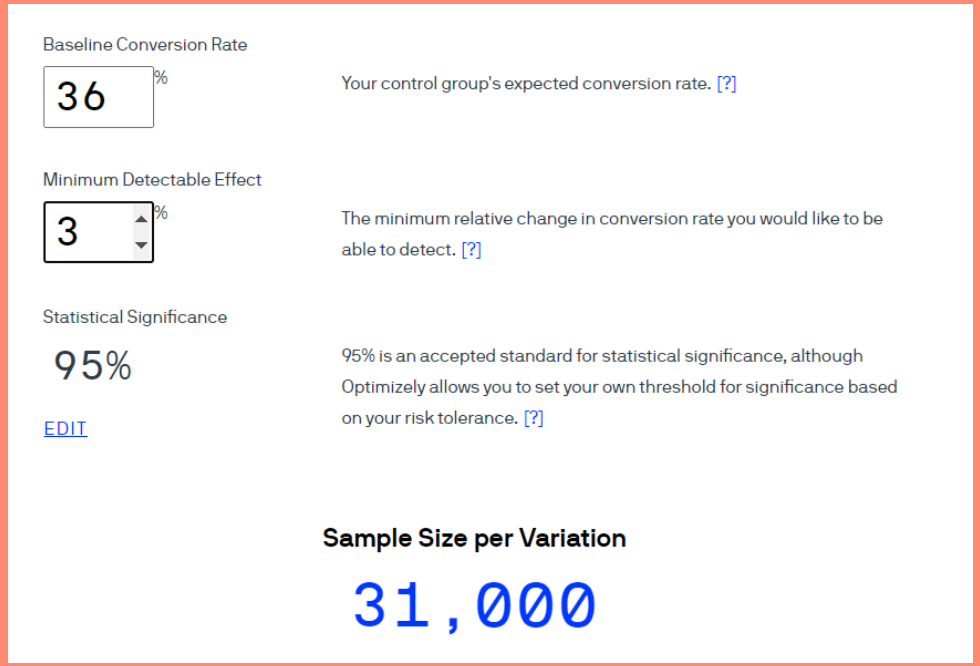
To cover this task we use Optimizly sample size calculator.
How? Link
Baseline conversion rate is the current conversion rate for the page you're testing: Start with my DA/MAU KPI: 36%
Minimum detectable effect (MDE): you need to decide how much change from the baseline (how big or small a lift) you want to detect. If you enter the baseline conversion rate and MDE into the Sample Size Calculator, the calculator will tell you what sample size you need for your original and each variation
Statistical significance: Statistical significance answers the question, "How likely is it that my experiment results will say I have a winner when I actually don't?" We usually consider 90% statistical significance. Another way to say the same thing is that we will accept a 10% false positive rate, where the result is not real (100% - 10% = 90%).

Answer Slide

Based on we establish just a few pick-up/drop-off location and testing this service is a small portion of total taxi users we should calculate with a realistic small sample size.

Baseline: DAU/MAU rate: 36%
I want to detect 3% change in coversion rate for my expirements

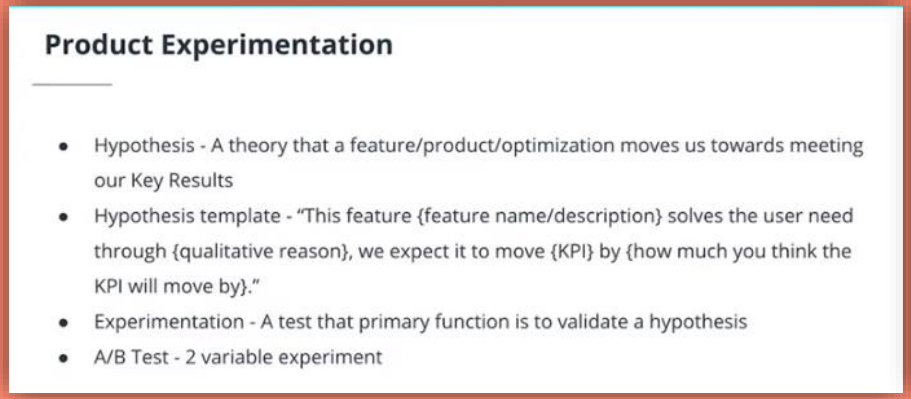
As my DAU is 2000 users, I need to run experiments 16 days

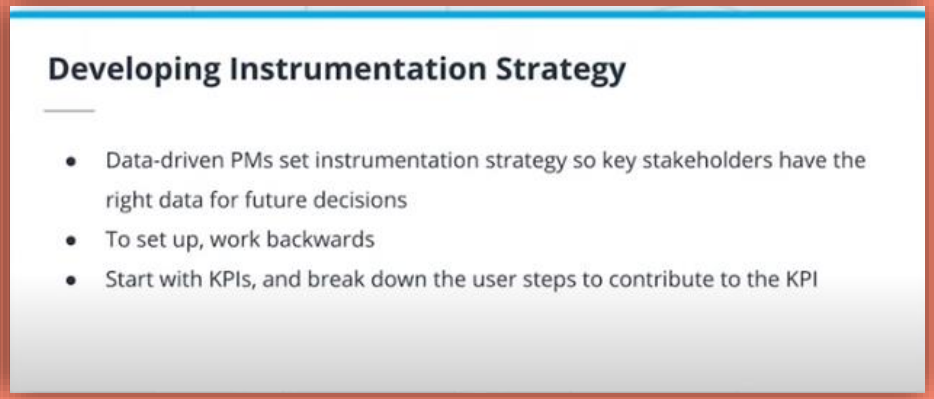


Create an instrumentation plan for the events you need collected and logged, in order to be able to physically measure your KPIs.

Answer Slide

To solve this task I used Udaxity lecture





Answer Slide Theory

Events are data points that are produced whenever an action happens.

Instrumentation is the act of ensuring that events are produced, saved, cleaned, and stored in a database for access.

Events are necessary to calculate KPIs. They are also leveraged for analytics use-cases all across the board, from training machine learning algorithms, to data science deep-dives, and even running triage on product outages.

You should capture every step of the user journey, logging each step that they take to meet a goal (like finish onboarding), so that you could analyze where if there are any leaky points in the funnel that are heavily preventing the user from meeting their goal.

It's best practice to **name your events** in camelCase nomenclature (first word uncapitalized, proceeding words capitalized), like userOnboardingCompleted or userSignedIn.

For standardization, name the event starting with the subject, then the predicate. For example: in userSignedIn, the subject is the user. What did they do? They signed in to the app.

Event properties are best named with all lower cases, and an underscore in between. For example: artist_id and ingest_timestamp**.**

Answer Slide

KPI 1: DAU / MAU - a ratio effective at measuring the rate of stickiness, for products that are meant to be used daily

Event I.- userOnboard
Definition - triggers whenever the driver hits the “User is picked up” button on the Flyber driver app.
Properties - user_count, rider_id, driver_id, pickup_timestamp, pickup_latitude, pickup_longitude

Event II.- userExperience
Definition - triggers whenever the user activte flyber’s experience on the flight like wifi, cold drinks, music.
Properties – user_experience_count, rider_id, driver_id, wifi_use, music_on, colddrinks_on

Answer Slide

KPI 2: Average NPS of all flying %

Event l.- userReferral
Definition - triggers whenever the user use her referral code on the Flyber driver app.
Properties - user_id, newuser_id, timestamp

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

Answer Slide

Can you explain what happened during the flight and how it impacted you?

What other options did you explore before choose this service?

If you could wave a magic wand, what would you have liked to have happened?

Would you recommend our service to your friends? If so, why, if not why?

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

Answer Slide

Used information: Udacity lesson ask WHY!

Pain Points

- High onboarding costs - Customer Experience Managers
- High cost of onboarding per agent, and a high volume of agents
- Job dissatisfaction from primarily negative feedback from customers

Target Users

- Primary: Customer experience agents
- Secondary: Customer experience managers

User Impact

- Increase customer experience agent satisfaction/wellness ->
- Increase customer experience agent retention ->
- Decrease total onboarding cost ->
- Increase in product upsells & retention

Market Impact

- 7 out of our 8 direct competitors in the market don't have a story for improving agent wellness
- Agents evangelize Expilliary with a positive experience as they move laterally or up in other companies

Answer Slide

Used information: Udacity lesson ask WHY!

Business Impact

- Align's with Expilliary's vision of helping customers be loved by their customers
- Aligns with year's Objective & Key Result of increasing upsell revenue by 41%

Risks & Assumptions

- Customer experience managers view this as hampering productivity, prohibiting the enablement of the feature in their instance
- Repetition fatigue from lack of diversity in notification prompts

Solutions

- Automated positive feedback
  - Example: Notification to agents if CSAT rates are above 85%
- Agent Kudos
  - Agents sending other agents Kudos (statement of congratulations)

Launch & Rollout Strategy

- EAP (Early Access Program) to generate early adopters / success stories, to evangelize to customer experience managers (mainstream market customer segment)
- Experiment on multiple positive feedback notifications across different experimental variants

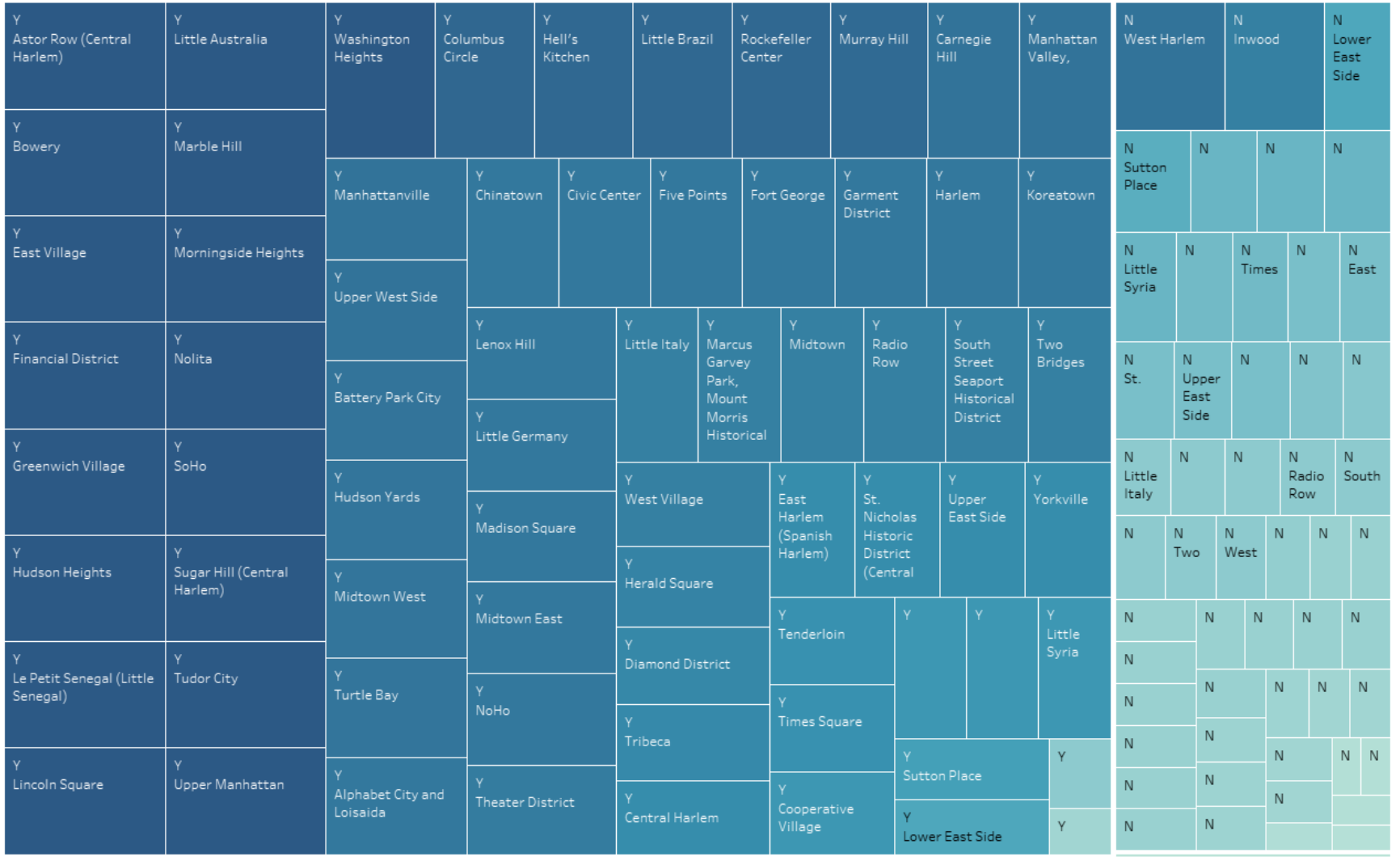
Answer Slide

Identify the target population. Why did you select that target population? What are their pain points?

Target population: High income ($120k-200k) taxi users who are under time pressure and living in dark blue zone (see next slide)

Pain points: business people who want to get from point A to B quickly and conveniently.

* traffic jam in morning and rush hours
* fear of accident
* poor taxi environment (can not work, or learn)
* poor service


Y Astor Row (Central Harlem)
Y Little Australia
Y Washington Heights
Y Columbus Circle
Y Hell's Kitchen
Y Little Brazil
Y Rockefeller Center
Y Murray Hill
Y Carnegie Hill
Y Manhattan Valley,
N West Harlem
N Inwood
N Lower East Side
Y Bowery
Y Marble Hill
Y Manhattanville
Y Chinatown
Y Civic Center
Y Five Points
Y Fort George
Y Garment District
Y Harlem
Y Koreatown
N Sutton Place
N
N
N
Y East Village
Y Morningside Heights
Y Upper West Side
N Little Syria
N
N Times
N
N East
Y Financial District
Y Nolita
Y Lenox Hill
Y Little Italy
Y Marcus Garvey Park, Mount Morris Historical
Y Midtown
Y Radio Row
Y South Street Seaport Historical District
Y Two Bridges
N St.
N Upper East Side
N
N
N
Y Greenwich Village
Y SoHo
Y Battery Park City
Y Little Germany
N Little Italy
N
N
N Radio Row
N South
Y Hudson Heights
Y Sugar Hill (Central Harlem)
Y Hudson Yards
Y Madison Square
Y West Village
Y East Harlem (Spanish Harlem)
Y St. Nicholas Historic District (Central
Y Upper East Side
Y Yorkville
N
N Two
N West
N
N
N
Y Le Petit Senegal (Little Senegal)
Y Tudor City
Y Midtown West
Y Midtown East
Y Herald Square
Y Tenderloin
Y
Y
Y Little Syria
N
N
N
N
N
N
N
Y Lincoln Square
Y Upper Manhattan
Y Turtle Bay
Y NoHo
Y Diamond District
Y Times Square
Y Sutton Place
Y
N
N
N
N
N
N
N
Y Alphabet City and Loisaida
Y Theater District
Y Tribeca
Y Cooperative Village
Y Lower East Side
Y
N
N
N
N
N
Y Central Harlem

Answer Slide

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

User impact: Increase target user satisfaction/wellbeing
Increase target user engagement, retention
Providing target user high level of safety and environment to feel to be differentiated

Market impact: Uplifting to traditional taxi service
Change of taxi service
Less traffic jam

Answer Slide

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

Solution: a custom-made hardware, designed for safety and high-end flying experience

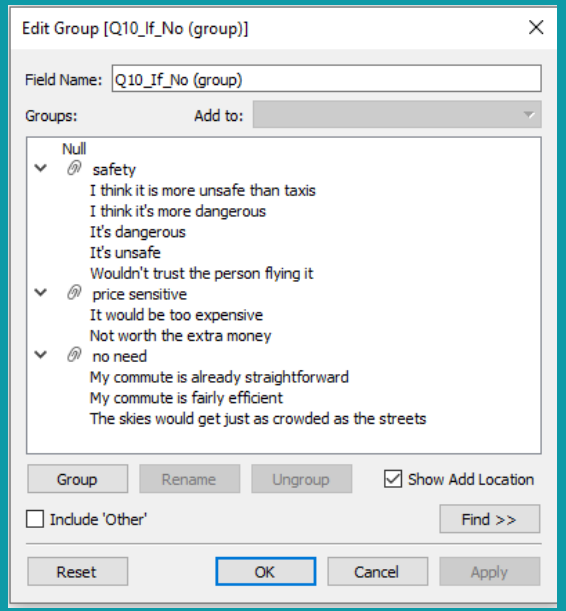
Risks: hardware engineering limitation, flying certification, human emotions (fear): see next slides

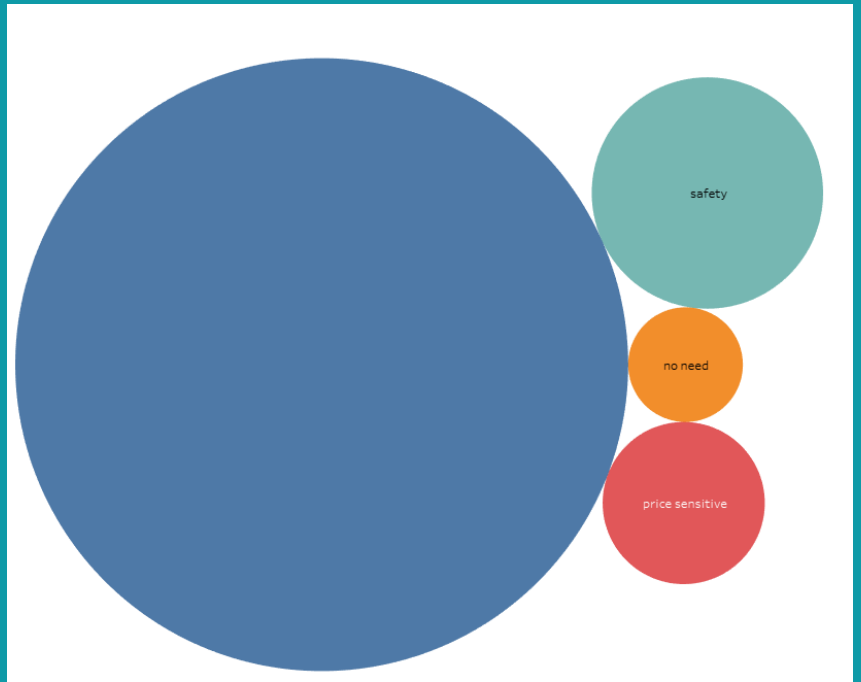
Launch & Roll-out strategy:

EAP program for early users
Sucess strories to evangalize
Experiments on early users

Answer Slide

Added slide: human fears | need to address





Answer Slide

State cross-functional stakeholder teams that will need to be involved

Answer:
Legal department: standards, EULA, Terms of Use
Marketing: positioning, value proposition, go-to-market
Engineering: custom hardware, software
Q&A: Testing
Product management:
Accounting: revenue, gross profit, other product metrics
Customer service

Answer Slide

Claim and evidence vs user pain points

Answer:
1., traffic jam in morning and rush hours_eviddence: https://www.tomtom.com/en_gb/traffic-index/new-york-traffic/ :

2., fear of accident_evidence: user research results (see above)
3., poor taxi customer experience (can not work, or learn)_evidence: Taxi industry losing war to Uber because of customer service, not technology link