

Power Airbnb's Growth Through Personalization

Seminar Group 1 Team 2



Agenda

1. Business Background

2. Business Problem

3. Analytical Problem

4. Data Preparation

5. Data Visualization

6. Modelling

7. Model Evaluation

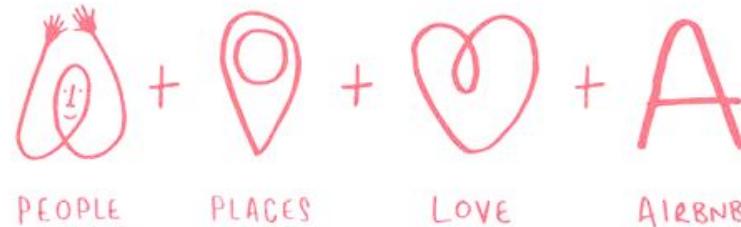
8. Recommendation

9. Feasibility Evaluation

10. Future Research

1. Business Background

1. Business Background - Airbnb



- Travel experiences platform which hosts vacation rentals and travel activities
- Two-sided intermediary platform - revenue comes from both hosts and guests
- Guests and hosts are charged **3% and 20% of booking amount** respectively

1. Business Background - Statistics



Growth rate of 46% in 2017



\$93 million in profit on \$2.6 billion in revenue in 2017



>100 data scientists



>150 million users worldwide

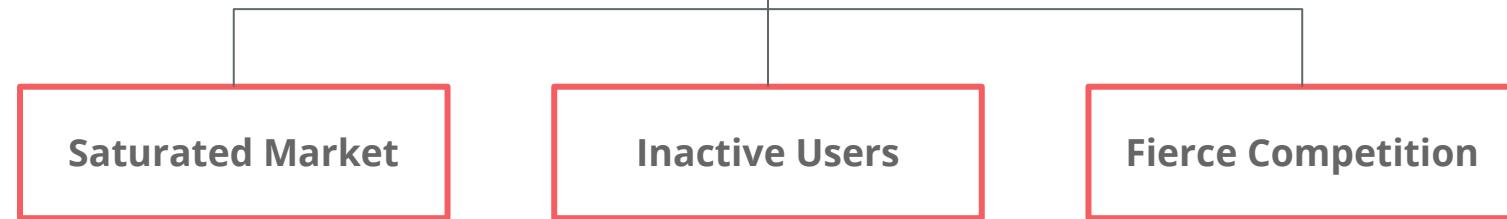


Largest market lies in the US

2. Business Problem

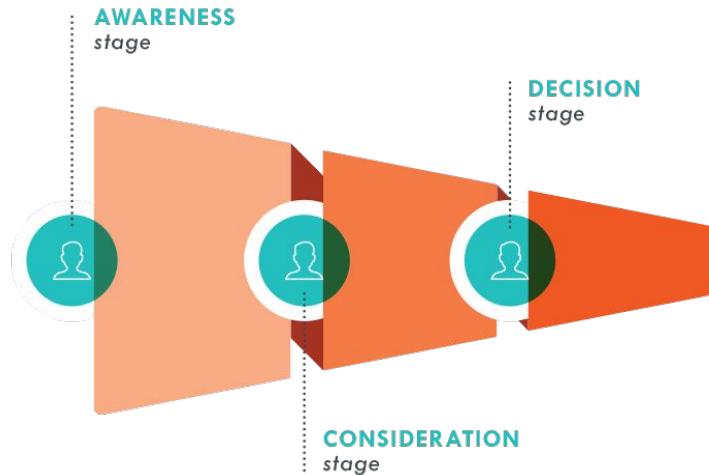
2.1 Business Problem

Stagnating Growth



2.1 Business Problem

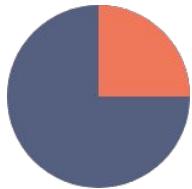
Saturated Market



- Brand awareness plateauing at 86% in biggest markets
- Current strategies overly focused on awareness
- Need to target consumers in Consideration stage

2.1 Business Problemss

Inactive Users



85% of Airbnb's user base are inactive without any bookings or listings



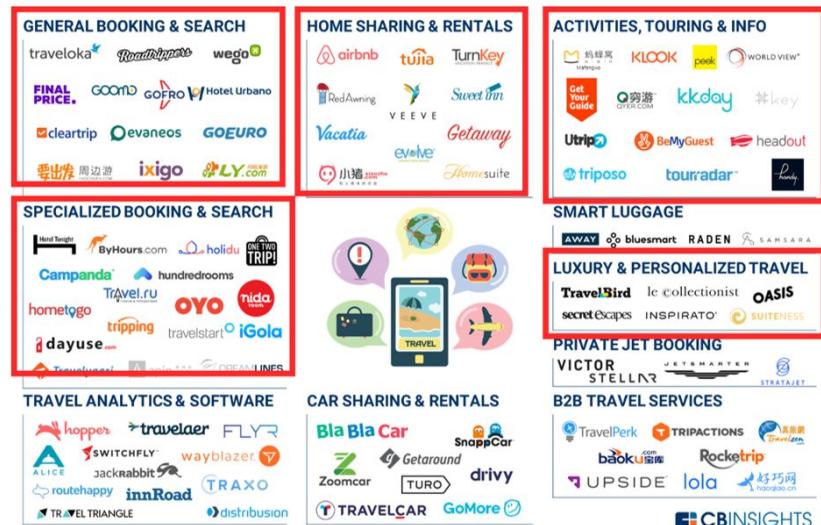
This suggests that **onboarding process** for new users needs to be more **engaging**

2.1 Business Problem

Fierce Competition

- Expansion of business to activities, luxury accommodations and hotels
- Compete directly with Online Travel Agencies giants eg. Expedia
- Competitors have aggressive personalisation strategies

AFTER EXPANSION



2.1 Business Problem

Fierce Competition

- Expansion of business to activities, luxury accommodations and hotels
- Compete directly with Online Travel Agencies giants eg. Expedia
- **Competitors have aggressive personalisation strategies**

<input type="checkbox"/>		Expedia.com.sg	Trash	Ooi Jia Xuan – Well done! We owe you THIS special on air for Singapore - You've ...
<input type="checkbox"/>		Expedia Rewards	Trash	Your monthly statement - you have 684 points worth SG\$4.89 - Your Expedia Re...
<input type="checkbox"/>		Expedia.com.sg	Trash	😊 Open ASAP! You've been selected to access THIS daily deal - You're Worth a S...
<input checked="" type="checkbox"/>		Expedia.com.sg	Trash	☺ Shocking: You've been sent a limited-time promotion - You Just Cau...
<input type="checkbox"/>		Expedia Member Pric.	Trash	;) No lie, Ooi Jia ;) We owe you FIFTY PERCENT OFF - YOU'VE BEEN CHOSEN: SP...
<input type="checkbox"/>		Expedia.com.sg	Trash	Ooi Jia Xuan – (1) new message: Reward yourself with THIS special on air from ...
<input type="checkbox"/>		Expedia.com.sg	Trash	::) Open ASAP! You've acquired 24-hour deals - You're Worth a Special Getaway ...
<input type="checkbox"/>		Expedia.com.sg	Trash	;) Message for you! We're updating you on a special event (3 days of deals ») - Yo...
<input type="checkbox"/>		Expedia Member Pric.	Trash	;) Mission accomplished, Ooi Jia ;) You've acquired FIFTY PERCENT OFF - YOU'VE...
<input type="checkbox"/>		Expedia.com.sg	Trash	Ooi Jia Xuan – Well done! We owe you THIS special on air for Singapore - You've ...
<input type="checkbox"/>		Expedia.com.sg	Trash	🎁 Congratulations, You Qualify for 10% Hotel Coupon - 3 DAYS COUPON Expedi...

2.2 Business Problem - Approach

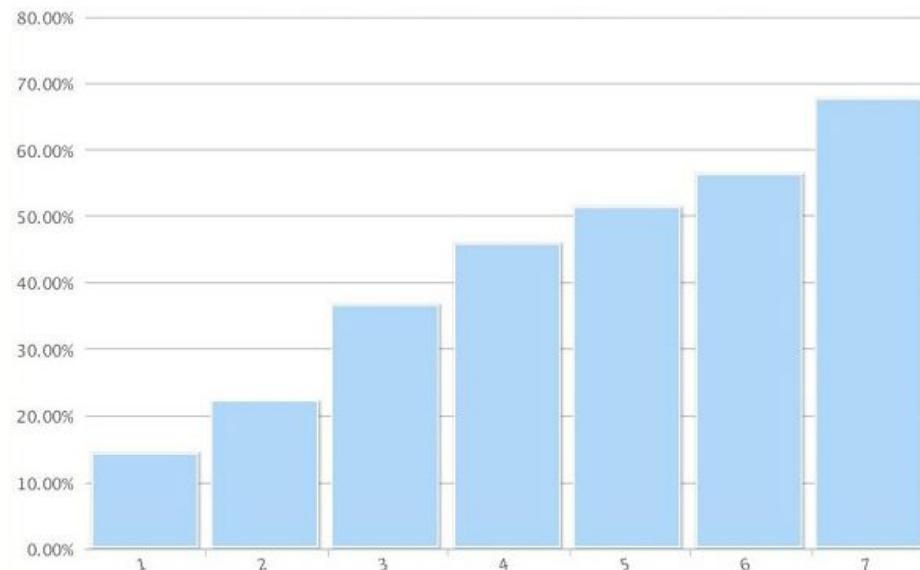
Personalisation

- Personalise **search results**
 - Already done by Airbnb's Search Ranking Algorithm
- Personalise **content**
- Personalise **promotions**

2.2 Business Problem - Approach

Targeting New Users

Likelihood of An Additional Stay



- Leads to recurring users
- Headstart among competitors

2.3 Business Problem Statement

Airbnb must improve the booking rate of new users through personalised marketing in order to sustain its growth.

**KPI 1: Overall Airbnb
Customer Booking Rate > 60%**

**KPI 2: Average Time
to First Booking <10 days**

3. Analytical Problem

3 Analytical Problem Statement

Predict the destination country for each user &
Predict booking urgency status for each user

KAPI 1: Top-5 Country
Destination Accuracy > 90%

KAPI 2: Urgency Status
Accuracy > 75%

3 How does it achieve desired business outcomes?

Recommend the right destination country to the right user at the right time



Prevent information overload and attract users' interest to increase user engagement

Accurately segment user according to booking urgency



For implementation of personalized marketing strategy, which increases booking rates

4. Data Preparation



4.1 Data Sources and Data Dictionary

Dataset of **Airbnb USA** users

- Demographics and behaviours
- Web session records

There are 12 possible outcomes of the **destination country**, including **other** and **NDF**.

United States



France



Australia



Spain



Portugal



United Kingdom



Netherlands



Canada



Germany



Italy

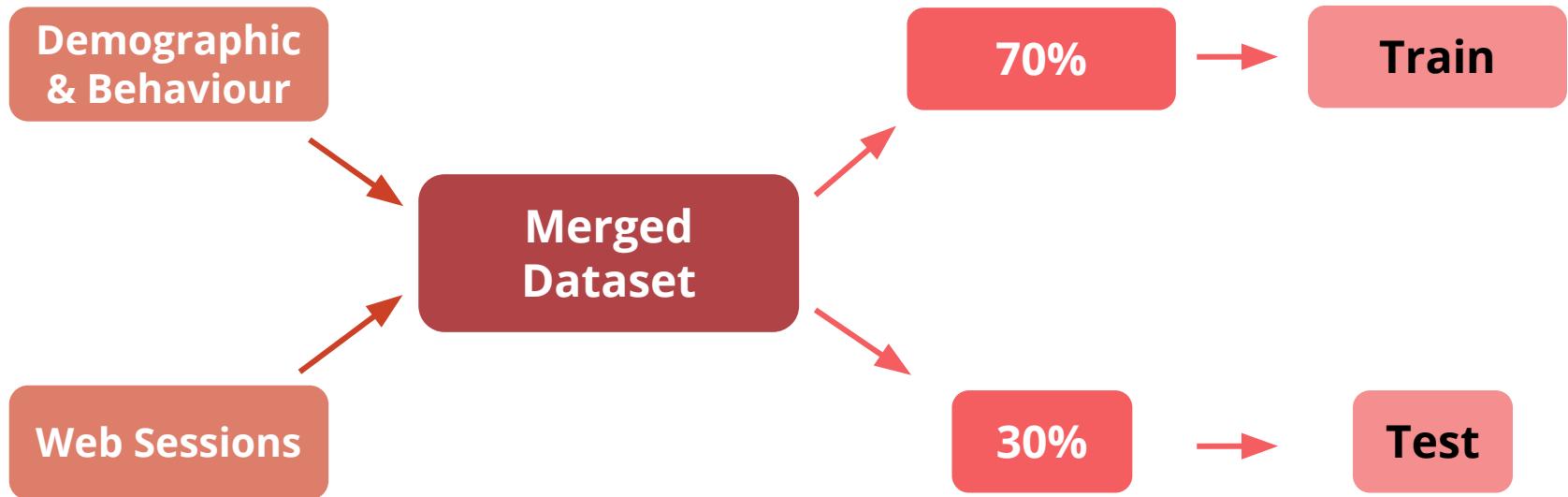


4.2 Data Cleaning



- **Outliers & NA (42.3%) in Age column**
 - Set outliers as **NA**
 - **Discretization**
 - Group NAs to '**unknown**' bucket.
- **Factorization of Month** feature
- **Counts of actions** for each user from web sessions records.
 - view search results
 - check similar listings
 - edit profile
 -

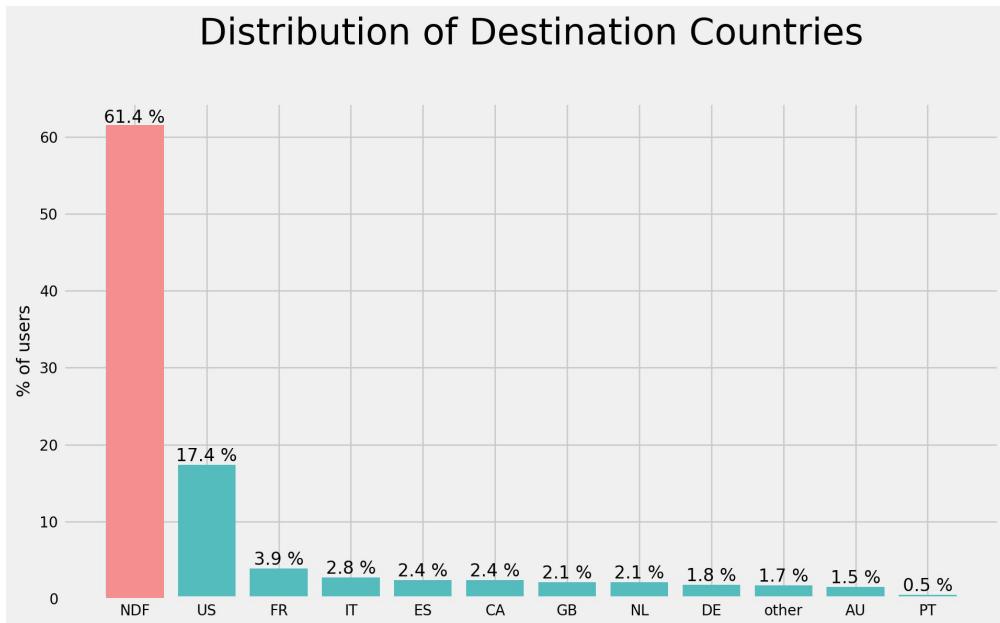
4.3 Data Splitting



5. Data Visualization

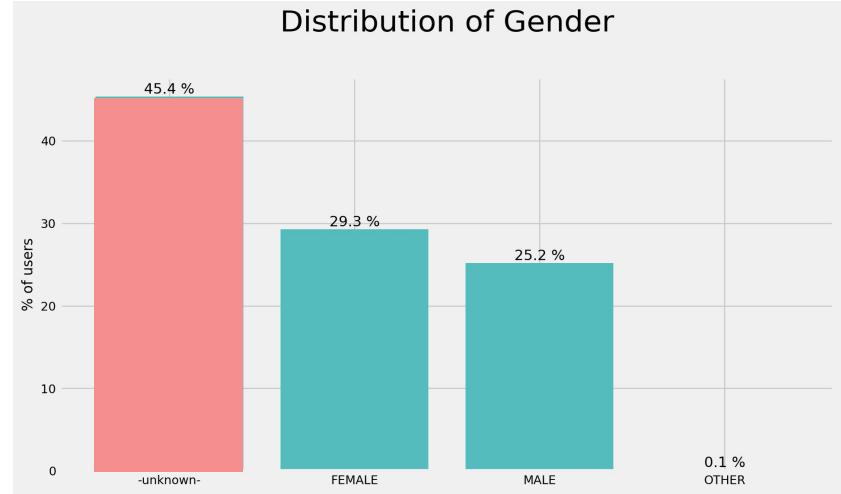
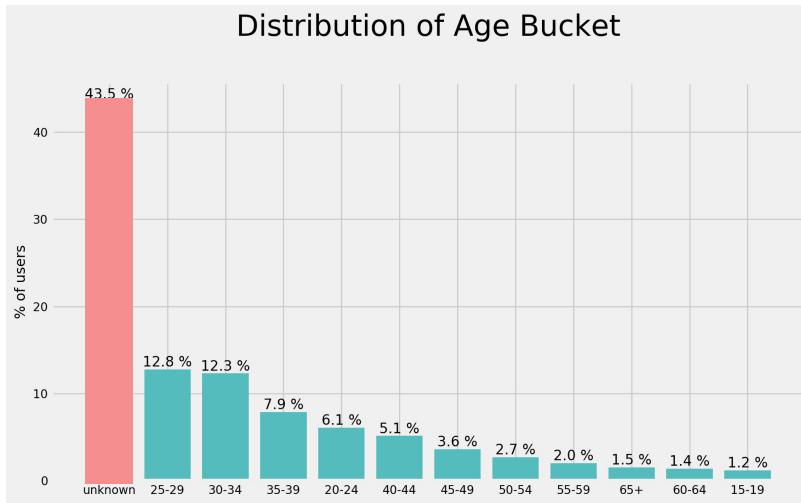


5.1 Destination Distribution



- Most users (**62%**) have never booked yet, reflecting a **low booking rate of 38%**.
- **The most popular country is US (17%)**, followed by US, France, Italy, Spain and Canada.

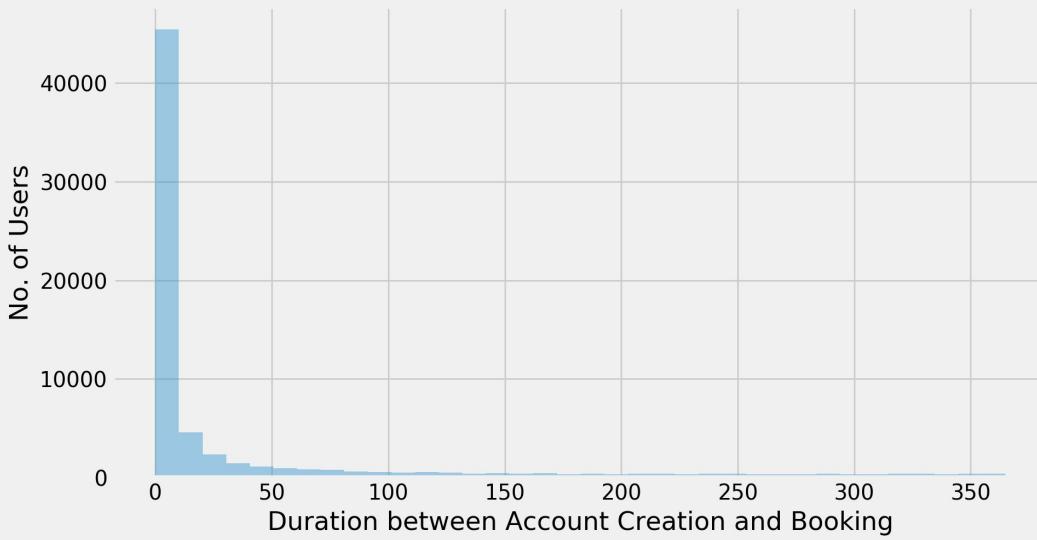
5.2 User Demographic - Age



- **Most users** are **unwilling to disclose** age and gender to Airbnb.
- The **major user group** of Airbnb are people aged in their **20s and 30s**.
 - **more acceptable** to new technology and innovative business models.
 - value the **affordability and unique experience** provided by Airbnb

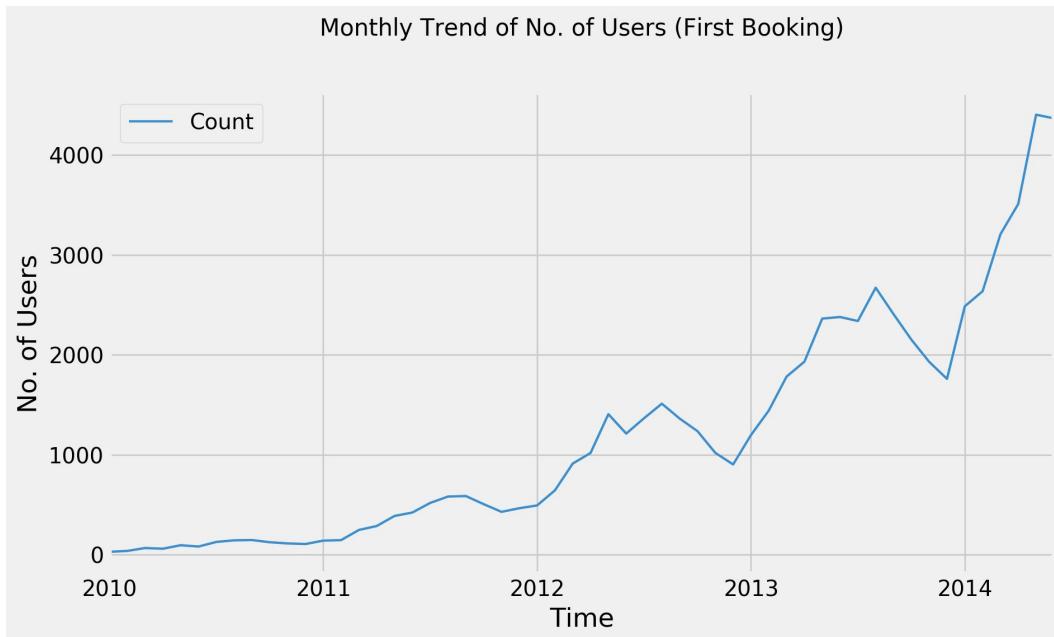
5.3 Account Creation and Booking Day

Distribution of Duration Between Account Creation and First Booking



- Most people make the booking **within 10 days** after account creation.
- It is important to **reduce the average time to 10 days** to capture more customers.

5.4 Peak Season



- The **midyear period** always experiences a surge in bookings.
 - Schools and universities **end their semesters** and families and college students head onto their **summer vacations**.

5.5 Country Destination and Month

Users are **2.1 more likely to travel to Australia** if they book in **December**, due to its location **in the southern hemisphere**, where the warmest weather falls between December and February.

The popularity of **Spain and France around June to August** is due to them locating **in the northern hemisphere**, where the weather the most ideal and travellers can enjoy long days of sunshine during that period.

The **festivals and special sights** in each country also affects their popularity:



Germany | October | Oktoberfest

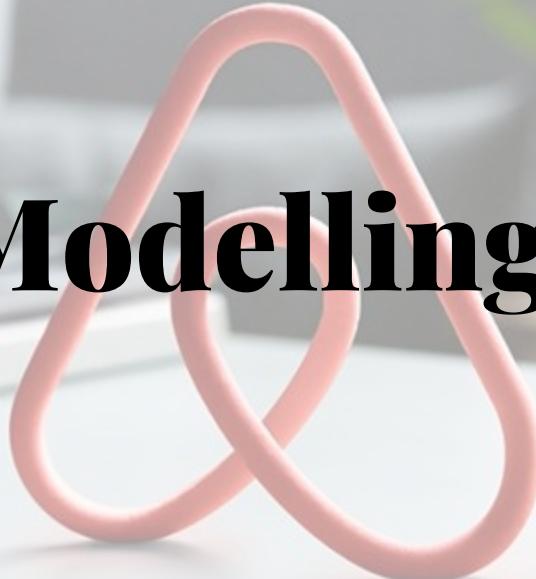


Netherlands | May | Tulip Season



Italy | April | Easter

6. Modelling



6.1 Modelling Summary

Model 1

Predict Country Destination



Recommend the right destination
country to the right user
at the right time

Model 2

Predict Status Urgency



Accurately segment user
according to booking urgency

6.1 Modelling Summary

Model	Technique	Feature Importance	Feature Effect
Logistic Regression	Feature Elimination	Coefficient	Coefficient
MARS	Captures nonlinearities & interactions between variables	Knots Value or Subsets Usage	Knots Value
Decision Tree	Pruning	Variable Importance	Decision Rules
Random Forest	Ensembling	Variable Importance	NA

6.2 Most Important Factor for Country Destination

Language

Day of Account
Creation

Age

Logistic Regression	MARS	Decision Tree	Random Forest
language	language	language	day of account creation
age	day of account creation	affiliate_channel	language
affiliate_channel	age	age	age

6.3 MARS for Country Destination

Language

- Users using Catalan are 0.994 more likely to travel to **Spain**
- Users using Italian are 0.887 more likely to travel to **Italy**
- Users using French are 0.783 more likely to travel to **France**
- Users using German are 0.452 more likely to travel to **Germany**

Day of Account Creation

- At the **end** of the year, the **later** the user creates account, the more likely that he/she travels to **Australia**.

Age

- For people aged **> 65**, they are **0.237** more likely to **travel within the home country**.

6.4 Most Important Factor for Urgency Status

action_refresh_subtotal

action_search_results

action_similar_listings

Logistic Regression	MARS	Decision Tree	Random Forest
action search results	action refresh subtotal	action refresh subtotal	action refresh subtotal
action refresh subtotal	action search results	action search results	action search results
age	action similar listings	action similar listings	action similar listings

6.5 MARS for Urgency Status

Action Refresh Subtotal

- When the user **refreshes the subtotal after changing trip characteristics** for one time, the probability that he/she is urgent to book increases by **0.521**.
- For **every refreshing** the user performs, the probability that he/she is urgent to book increases by **0.068**.

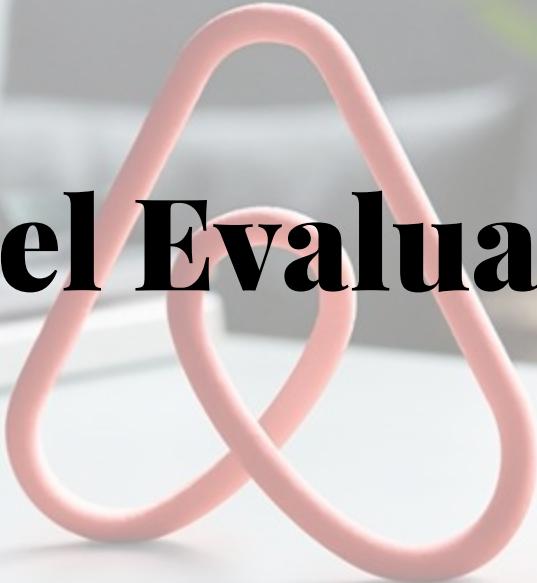
Action Search Result

- When the user **views the search result for one time**, the probability that he/she is urgent to book increases by **0.441**.
- For **every search result viewing**, the probability that he/she is urgent to book increases by **0.045**.

Action Similar Listings

- When the user **checks similar listings** for one time, the probability that he/she is urgent to book increases by **0.385**.
- For **every checking of similar listings**, the probability that he/she is urgent to book increases by **0.01**.

7. Model Evaluation



7. Model Evaluation



Performance

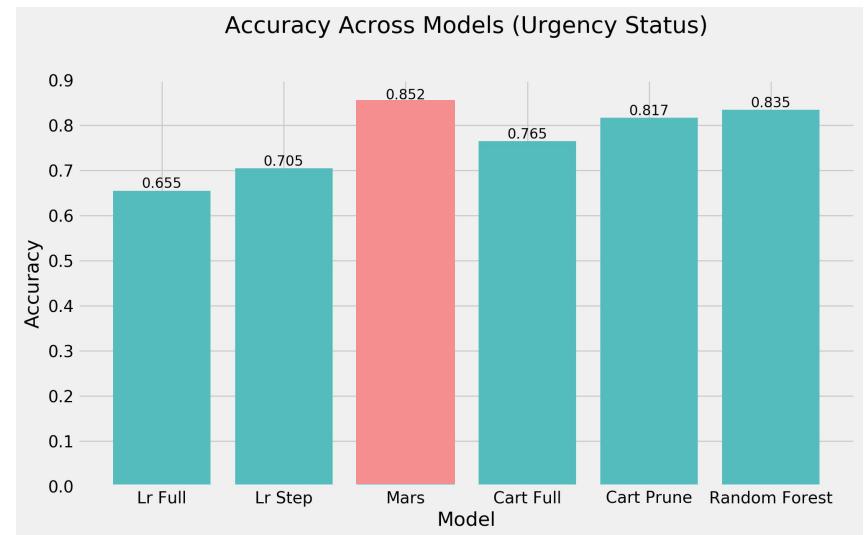
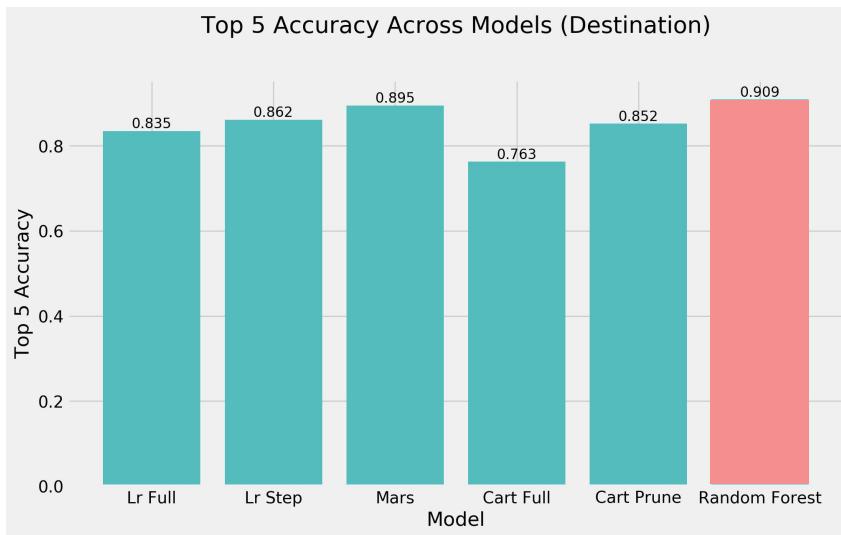


Interpretability

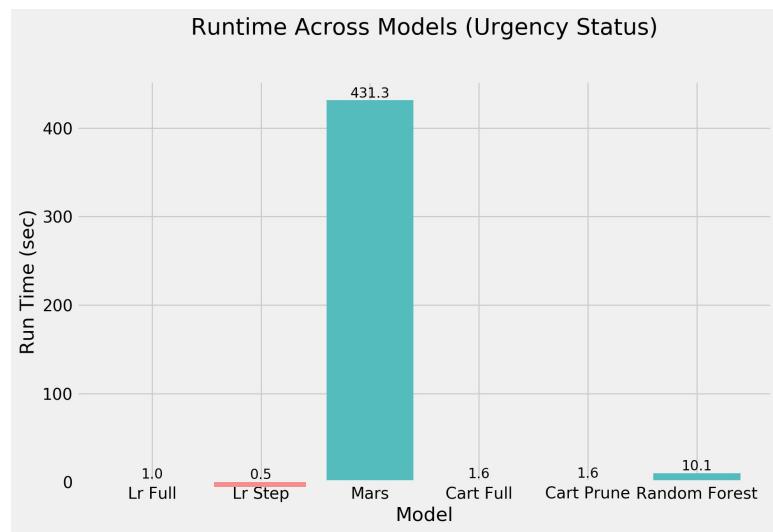
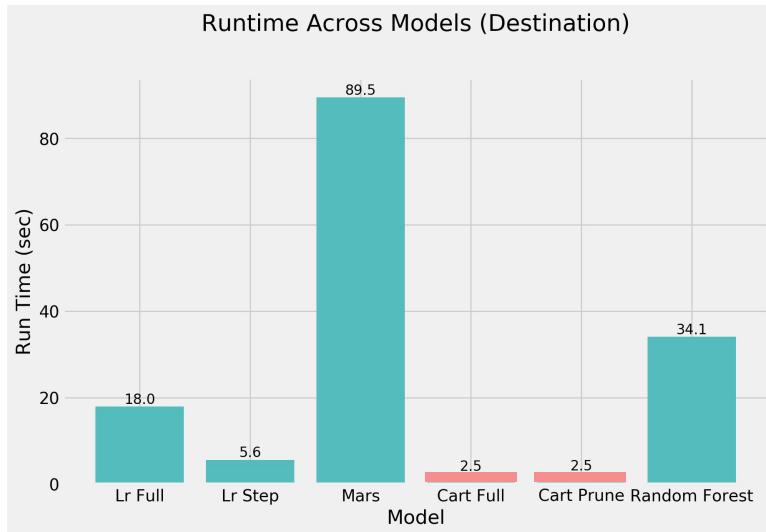


Efficiency

7. Model Evaluation - Performance



7. Model Evaluation - Efficiency



7. Model Evaluation - Interpretability

Model	Feature Importance	Feature Effect
Logistic Regression	✓	✓
MARS	✓	✓
Decision Tree	✓	✓
Random Forest	✓	X

7. Model Evaluation

The model **performance** is of the **highest priority**

Training time of models are also **not a concern** as they still within the **same scale**.

If the **interpretability** of the results (eg. for **reliability**) is emphasized by management, models with both **high level of interpretability and performance** are preferred.

Random Forest



Predict Country
Destination

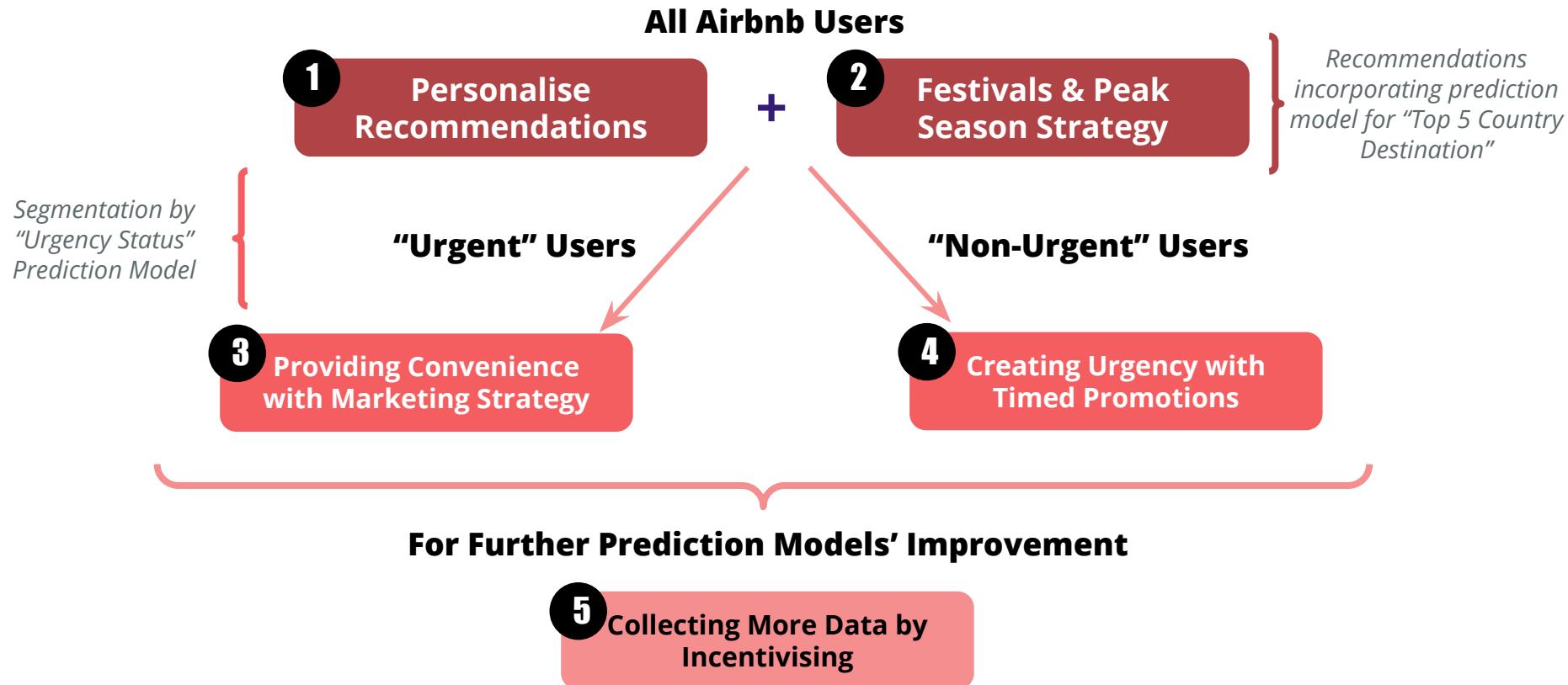
MARS



Predict Urgency
Status

8. Recommendations

8. Recommendation Categories



8.1 Personalise Recommendations



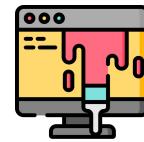
57% of U.S. travellers feel that brands should tailor their information based on personal preferences (Loo, 2017)



Removing Irrelevant Content from Website Improves Conversion Rates (Taylor, 2014)



Using the “Top 5 Country Destinations” Prediction Model:



Include predicted countries on main website's “Recommended for you” section for users



Include predicted countries on Airbnb App's “Recommended for you” section for users



Airbnb's separate blog content website could be incorporated into the website such that it will be the centerpiece of the country's homes and experiences segment

8.1 Personalise Recommendations

Website Snippet after incorporating idea

The screenshot shows a personalized travel guide for Copenhagen. At the top, a section titled "Recommended for you" displays five travel destinations with their average night rates: Copenhagen (\$167), Berlin (\$96), Prague (\$94), London (\$165), and Los Angeles (\$178). Below this, a large banner for Copenhagen features a vibrant flower field and the text "Step into Spring". To the right, a sidebar lists three travel experiences: "BOAT RIDE Hey Captain, let's sail!", "CRAFT CLASS Knitting in the nordic Hygge tradition", and "BIKE RIDE Copenhagen Discovery On Stylish Bike". To the left, another sidebar lists three local attractions: "BEER TASTING BeerWalks with Copenhagen Locals" and "ENTIRE APARTMENT - 1 BED Charming apartment in the heart of Copenhagen". Red arrows point from each of these sections to the corresponding red-highlighted areas in the main banner.

Recommended for you

- Copenhagen \$167/night average
- Berlin \$96/night average
- Prague \$94/night average
- London \$165/night average
- Los Angeles \$178/night average

Copenhagen Inspiration, Homes and Experiences

Step into Spring

Find the best Cherry Blossom viewing spots, discover colorful spring festivals, and enjoy the beauty of the blossoming season

Explore More

Blog/ Marketing Content for predicted country #1

Homes from predicted country #1

Experiences from predicted country #1

BOAT RIDE
Hey Captain, let's sail!
\$85 SGD per person
4.81 ★ (545)

CRAFT CLASS
Knitting in the nordic Hygge tradition
\$74 SGD per person
4.93 ★ (29)

BIKE RIDE
Copenhagen Discovery On Stylish Bike
\$85 SGD per person
4.93 ★ (14)

BEER TASTING
BeerWalks with Copenhagen Locals
\$55 SGD per person - 2 hours, Drinks included
4.99 ★ (246)

ENTIRE LOFT - 1 BED
★CityCentre Penthouse, PrivateTerrace, 25min...
\$226 SGD per night - Free cancellation
4.04 ★ Superhost

ENTIRE APARTMENT - 7 BEDS
Penthouse, 4-5 rooms + 5 balconies
\$382 SGD per night - Free cancellation
3.90 ★

ENTIRE APARTMENT - 1 BED
Charming apartment in the heart of Copenhagen
★CityCentre Penthouse, 4-5 rooms + 5 balconies
\$382 SGD per night - Free cancellation
2.46 ★

8.2 Festivals & Peak Season Strategy



Surge of bookings during middle of the year identified to be **peak season** (Section 5.4)



Trend where more users make bookings to **specific country destinations** during **certain months** (Section 5.5)



Using the "Top 5 Country Destinations" Prediction Model:



Peak Season: **Increase marketing and send out content and listings** on user's **predicted** countries to entice users to book



Off-Peak Season: **Provide promotions and discounts** to attract customers to book during low peak season



Festivals: Send out **marketing content** nearing to the months on the festivals according to the user's **predicted** country destinations

8.3 Recommendations for Different User Segments

"Urgent" Status Customers



Providing convenience to customers will ensure they purchase with you instead of competitors (Hyken, 2018)

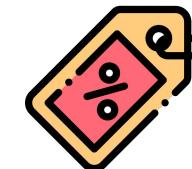


Using these customer's past actions data, streamline the most appropriate listings to email these users to provide convenience of the right room at their fingertips

"Non-Urgent" Status Customers



60% of travellers would consider an impulse trip if they receive a good hotel deal (Loo, 2017)



Injecting urgency by sending **timed promotions and discounts** for booking listings in their predicted top 5 countries and search history

8.4 Collect More Data

1

44% of users did not specify **age** and 46% did not provide **gender** (Section 5.2)

2

Demographic data are the **most important factor** for Predicting Country Destination (Section 6.2)

3

63% of millennial consumer agree that they are willing to share data with companies for **personalisation** and to receive **offers** (Altexsoft, n.d)



Examples



Airbnb can introduce an **Airbnb Rewards system** that will allow users to earn points from their bookings and redeem points for free night stay and vouchers for experiences or restaurants.



Using this rewards system, Airbnb can encourage customers to **provide more demographic data** like job or country they are living at currently by providing **extra rewards points**.



Simple voting boxes could be added into the website and app to get insights from fun, casual questions to further improve prediction model's accuracy.

8.4 Collect More Data

Voting boxes

What defines me best when I am on vacation?

ADVENTURE JUNKIE

INDOOR LOVER

Local's eat rule! Bring on all the unusual dishes and stories behind them

TRUE

FALSE

What defines me best when I am on vacation?

ADVENTURE JUNKIE

54%

INDOOR LOVER

46%

Local's eat rule! Bring on all the unusual dishes and stories behind them

TRUE

78%

FALSE

22%

9. Feasibility



9.1 SWOT analysis for Top 5

<ul style="list-style-type: none">• Top 5 algorithm has a matching rate of 91%• Able to implement with the Airbnb's current Agile practices• Blanket approach to all users.  Strengths	<ul style="list-style-type: none">• 45.6 million users using airbnb services (statista)• 45.6 million dataset (time/resources)• Feasibility study on airbnb database is required  Weaknesses
<ul style="list-style-type: none">• Potential advertising revenue from recommended listing• Potential huge increase in booking rates  Opportunity	<ul style="list-style-type: none">• Performance with larger dataset would change (annual evaluation)• Imbalanced data, normality of dataset might skew the result  Threats

9.2 SWOT analysis for Different User Segments

<ul style="list-style-type: none">Model has a accuracy of 85%Strategy is highly scalable (internal benchmarks)	<ul style="list-style-type: none">Effectiveness depends on model accuracyUrgent - Reliance on customer allowing airbnb to send email to themNon - Urgent - Promotional/discount strategy (financial burden)
 Strengths <ul style="list-style-type: none">Urgent - Possible implementation of search history as a predictorNon - Urgent - Potential collaboration with airbnb host for discount	 Weaknesses <ul style="list-style-type: none">Urgent - Potential area of conflict with data protection act of many countriesNon - Urgent - Conversion from non-urgent to urgent
 Opportunity	 Threats

9.3 Financial Feasibility (Cost)



Considerations	Cost
(1) Adding new codes to existing underlying code	6,000 (Giover, 2017)
(2) Redesigning the web pages, ios and android application	7,650 (estimatemyapp)
(3) Blog content material content generation	5,000 x 5 per month
(4) Promotion and discount	Revenue cut up to 20%

9.3 Financial Feasibility (Revenue)



Considerations	Revenue
<p>Recommendation engine:</p> <ul style="list-style-type: none">• Top 5• Peak season and non-peak• Different user segments	<p>Revenue growth of up to an additional 35% (Amazon)</p>

10. Limitations and Further Research

10.1 Data Collection



Information such as **searching history and user's interest** would definitely help as a predictor in our model.

Performance of our model is thus **limited** due to the lack of such information.

Limitation will be addressed by **collecting more data**

10.2 Broader Scope



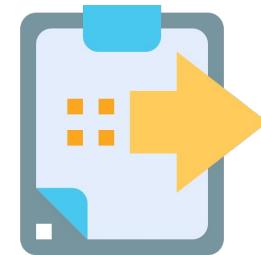
Broader Coverage of Users

Implement for all users around
the world



Narrow Down to City

Dream city for the user could
achieve more personalization



Predict Next City Destination

Target not only new users but
also booked users

10.3 Mode Rigorous Model Selection

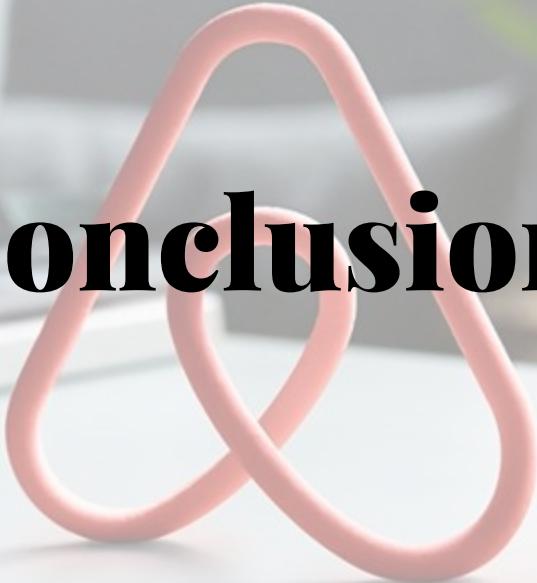


Black Box Models Difficult to Interpret

Train more rigorous and advanced models
such as **Neural Network** and **XGBoost**.

Accuracy of the model is priority

11. Conclusion



11. Conclusion



Data is a
extremely
valuable resource
for airbnb

User Travel
Destination

Urgency status
prediction models

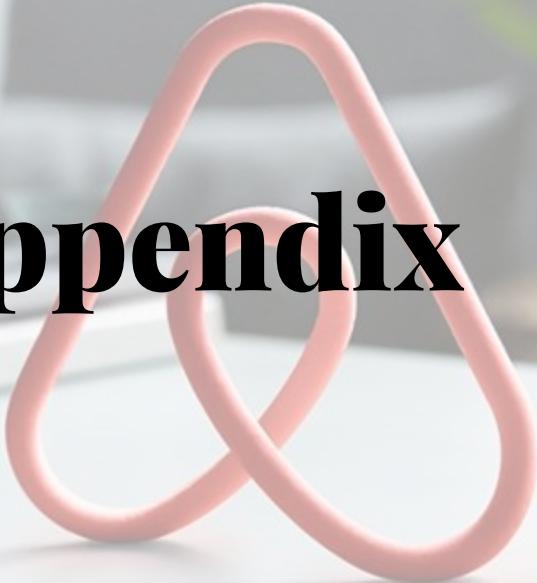
Personalization

Airbnb Growth

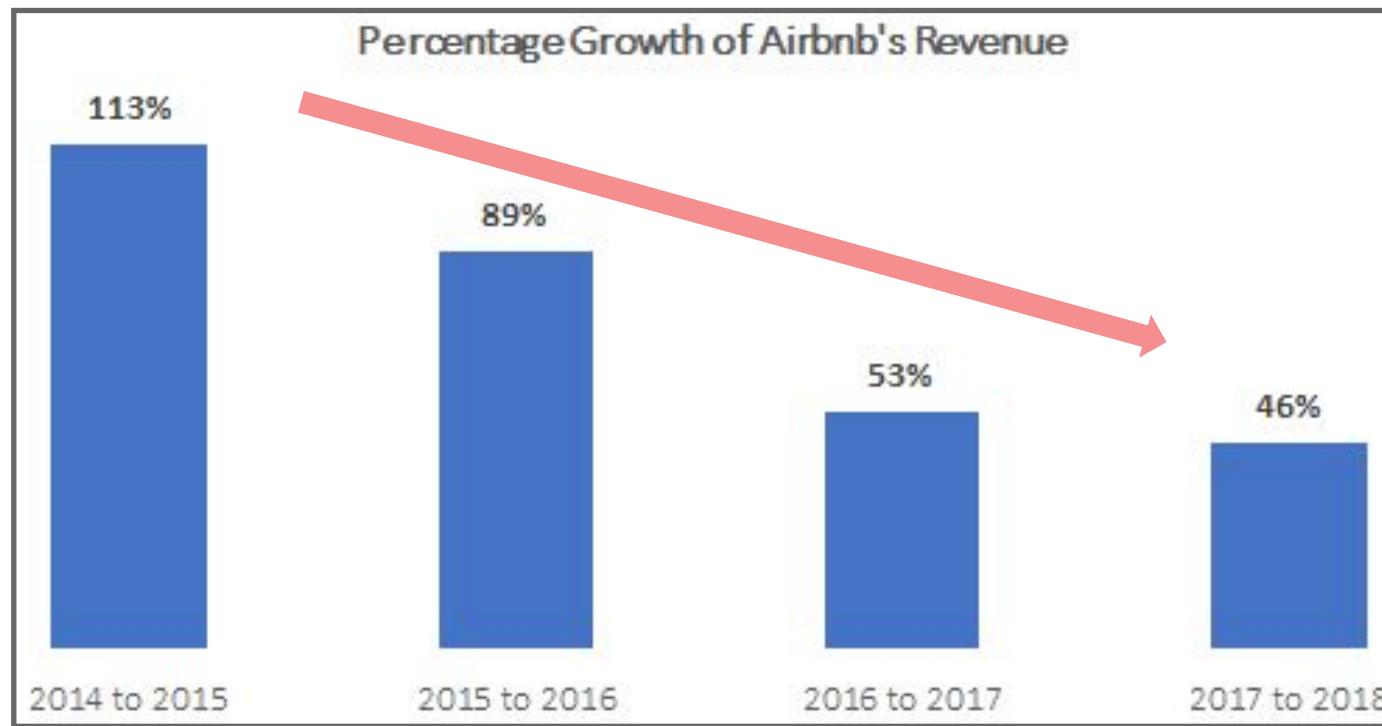
12. Live Dashboard



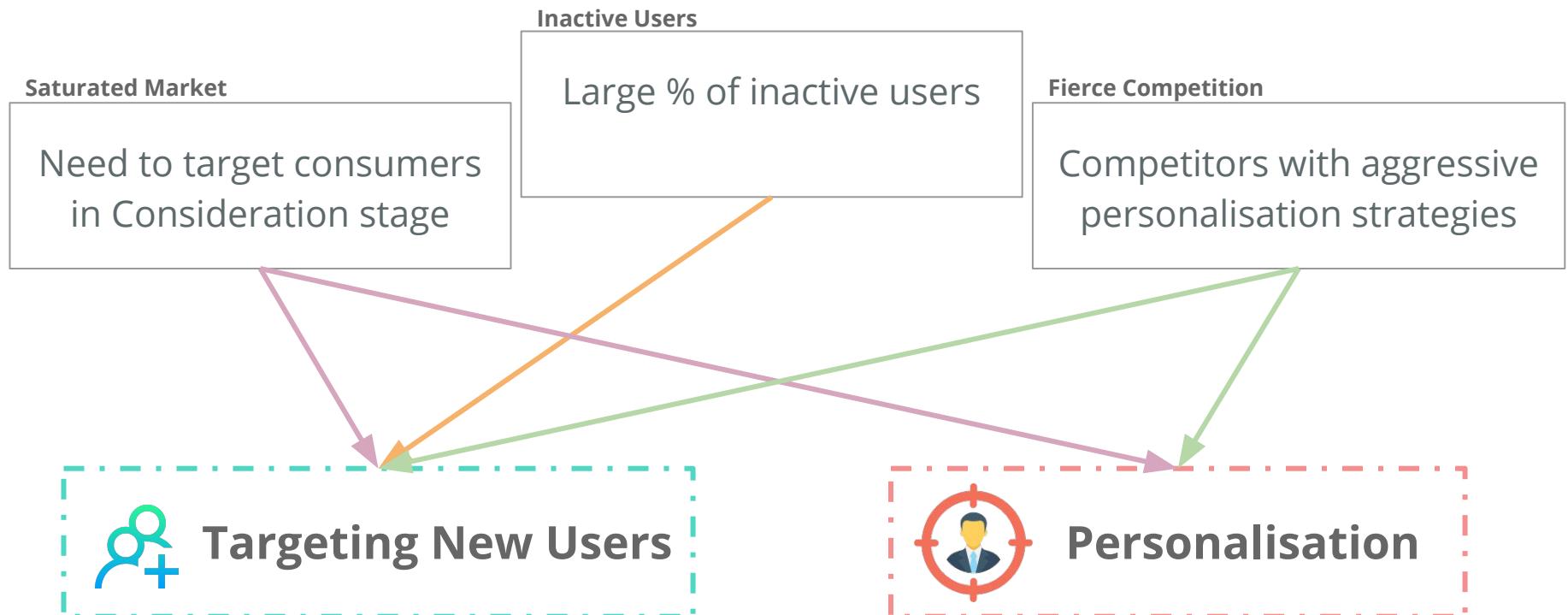
Appendix



Appendix 1: Stagnant Growth of Airbnb



Appendix 2: Business Problem & Approach



Appendix 3: Current Home Page of Airbnb

Explore Airbnb



Homes



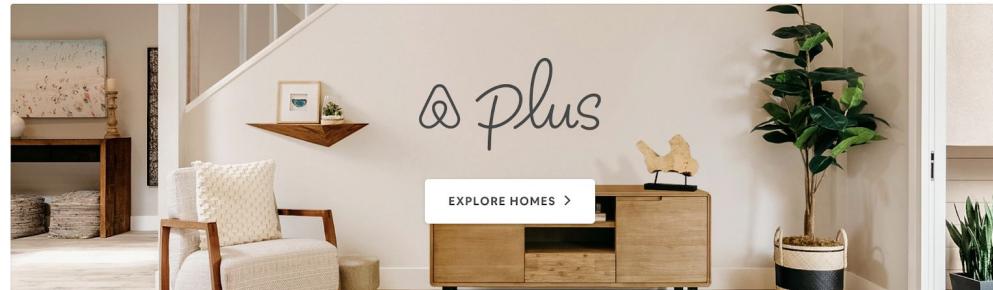
Experiences



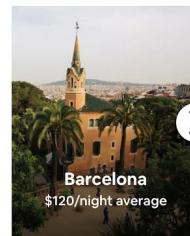
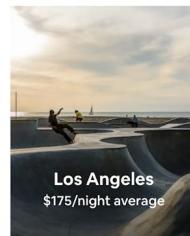
Restaurants

Introducing Airbnb Plus

A new selection of homes verified for quality & comfort



Recommended for you



[Terms, Privacy, Currency & More](#)