# Capstone Project

DATA ANALYTICS BOOTCAMP

Travel Insurance Analysis from

AirAsia Bookings Data

By Lee Boon Chek 433A

Problem Statement

- Every customer who books a flight is a potential buyer of other ancillaries viz. meal, preferred seat, extra baggage, and travel insurance.

- A customer can buy these ancillaries anytime before the flight departure (in case of meals, the current data only includes pre-booked meals and not the in-flight sale of meals).

- **Based on each customer's bookings data** (each row represents one booking), **predict whether a customer will buy Insurance** (variable name: INS_FLAG)
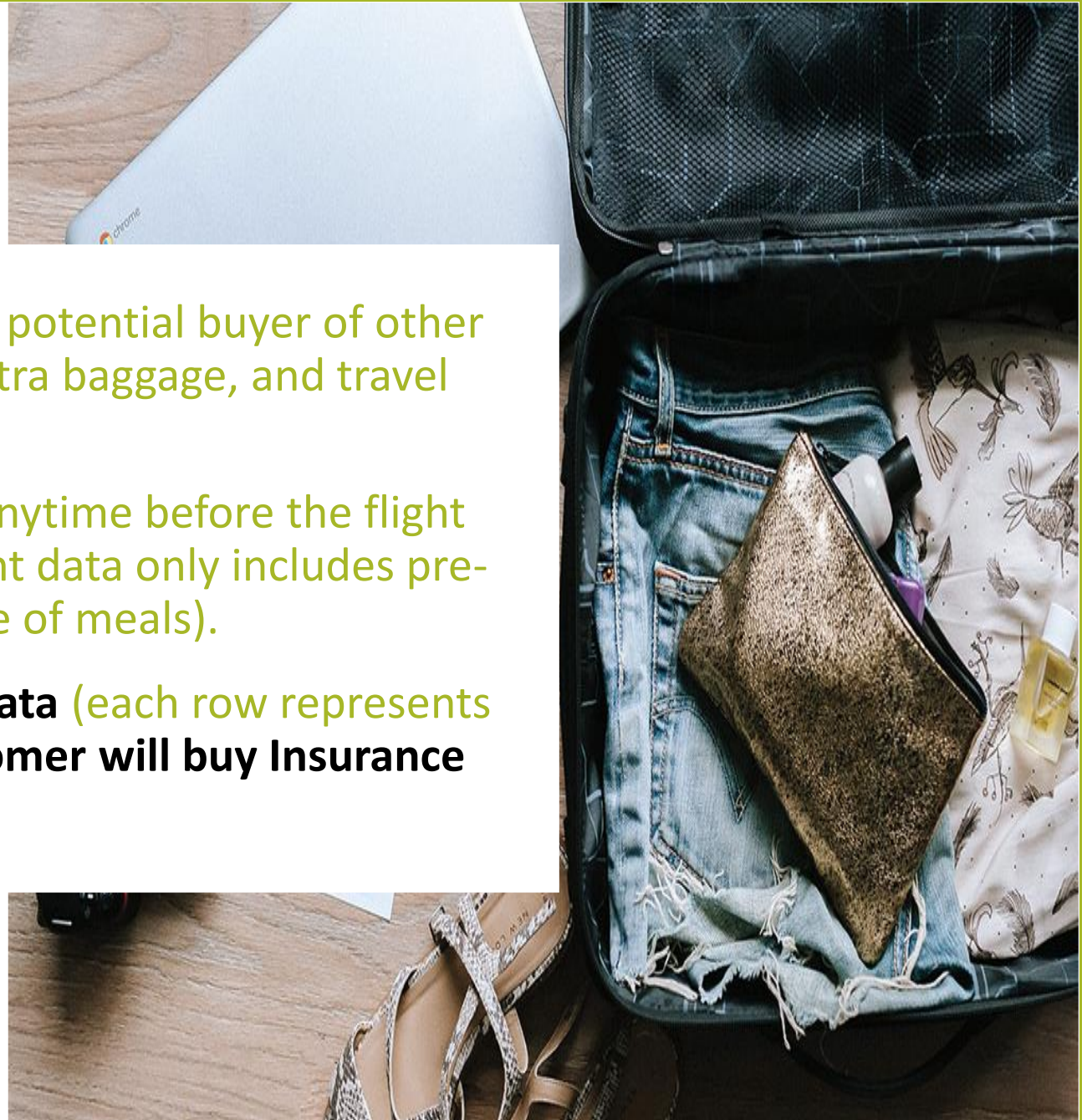
Data Acquisition

- The data used in this analysis is downloaded from below link:

- https://www.kaggle.com/datasets/mundher/airasia-passengers

- There are in total 15 columns and 50,000 bookings (rows).

| Id	Identifier	Variable Type
PAXCOUNT	Number of customer traveling	Categorical
SALESCHANNEL	Sales channel booking was made on	Categorical
TRIPTYPEDESC	Trip Type (Round Trip, One Way, Circle Trip)	Categorical
PURCHASELEAD	Number of days between travel date and booking date	Continuous
LENGTHOFSTAY	Number of days spent at destination (derived for one way trips)	Continuous
flight_hour	Hour of day of Flight departure	Continuous
flight_day	Day of week of Flight departure	Categorical
ROUTE	OriginDestination flight route (KULPEN – Kuala Lumpur to Penang)	Categorical
geoNetwork_country	Country from where booking was made	Categorical
BAGGAGE_CATEGORY	Has bought extra baggage in booking	Categorical
SEAT_CATEGORY	Has bought preferred seat in the booking	Categorical
FNB_CATEGORY	Has bought in-flight meals	Categorical
flightDuration_hour	Total duration of flight (in hours)	Continuous
INS_FLAG	Has bought insurance? (Target Variable)	Categorical

Data Cleaning and Parsing

- Assigned the type of variable to each data ID (Categorical/ Continuous).

- Created 4 time **ranges** to the "LENGTHOFSTAY" data, **defined** the data selection as "length of stay", and used **Vlookup** to pair the time range to each booking.

- **Replaced** "1" as "YES" and "0" as "NO".

- Converted the data into **pivot table** and used pivot table to calculate p-value in chi-square test for categorical variables. The p-value was then used to eliminate the fields that are not likely to affect the purchasing of travel insurance. (Significance study)

- Calculated p-value of continuous variable using two sample t-test. All p-values were converted with log scale for better presentation.

- The combined two airport names under "ROUTE" field were **split into two columns** (Origin/ Destination), and "**Geographic Role**" was assigned in Tableu. Latitude and longitude data were retrieved from Tableu and paired to each airport in excel with Vlookup and "**IFNA**" function.
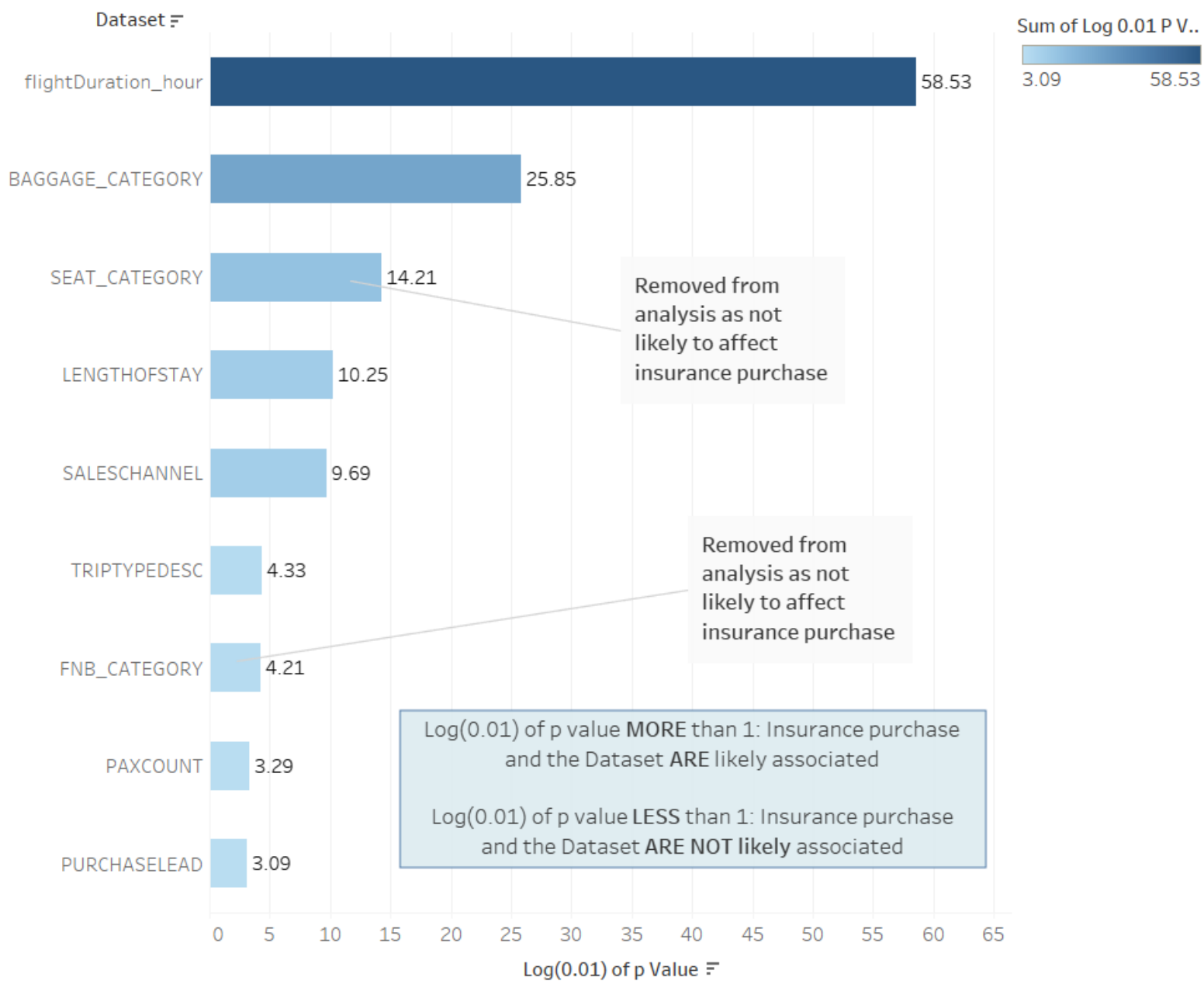
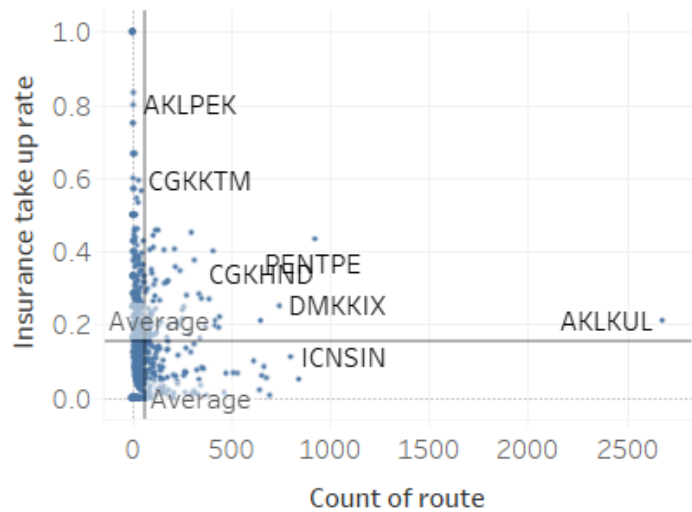# TABLEU STORYBOARD
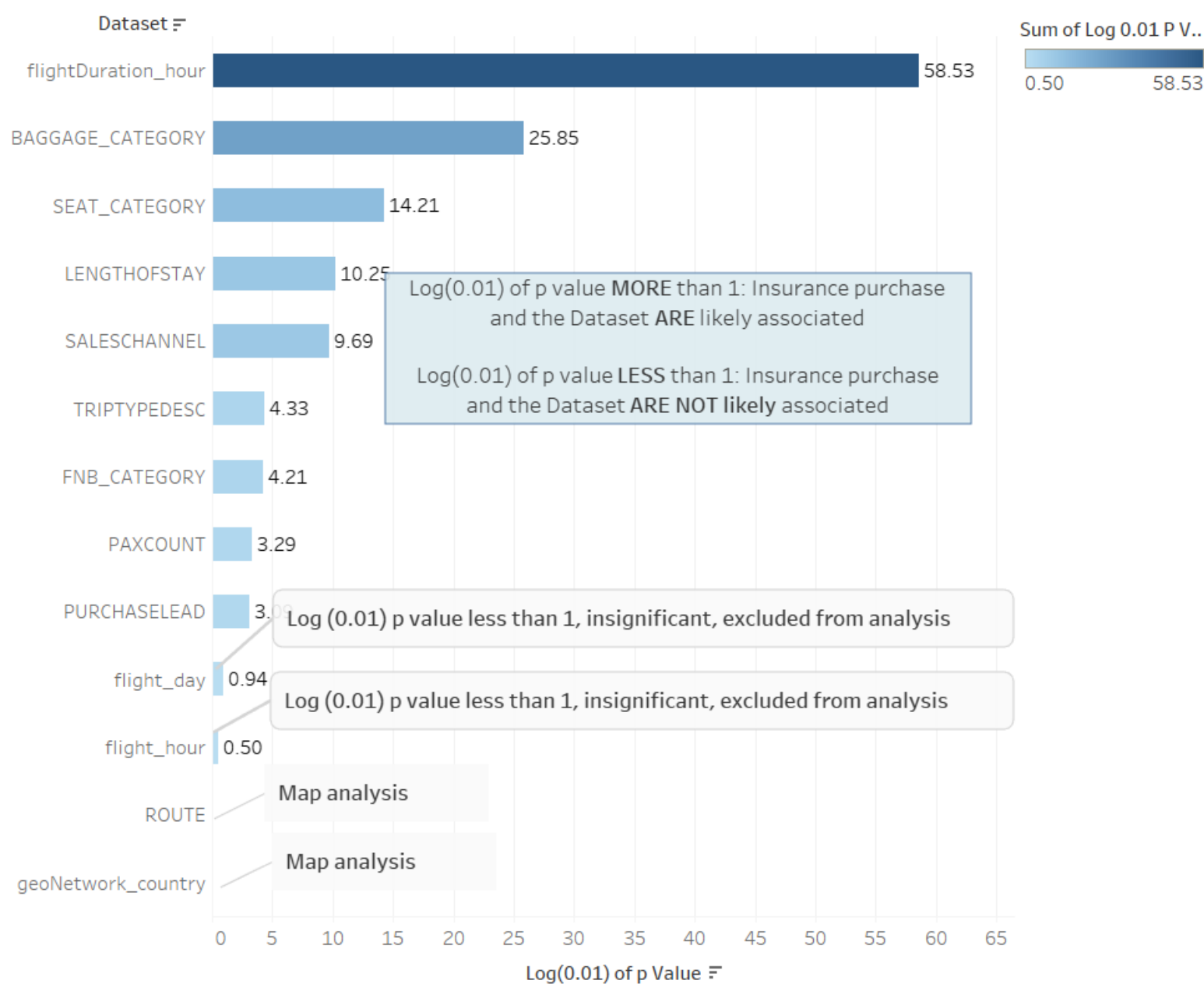
Significance Study

- In order to tell which dataset affects the purchase of travel insurance, p-values in log scale were compared.

- flight_day and flight_hour were excluded from analysis as they are not likely to affect the purchase of travel insurance.
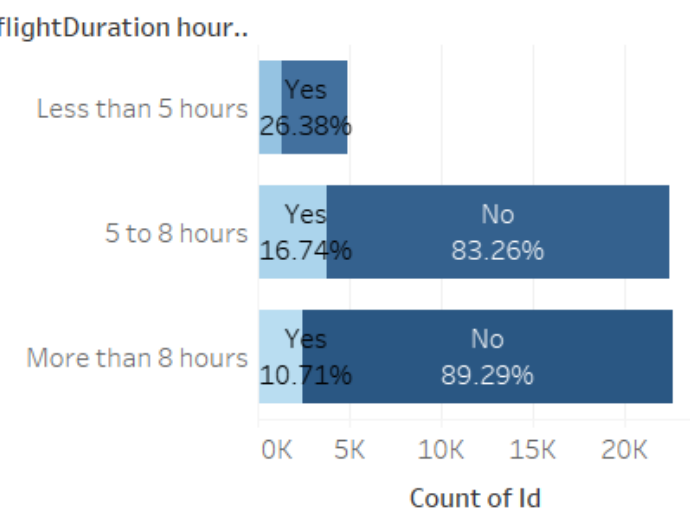
Other data exclusions

- SEAT_CATEGORY and FNB_CATEGORY are more associated to personal preference and flight hour than insurance purchase. Hence they were also excluded from this analysis

Route Count and Insurance Take Up Rate

Insurance take up rate (y-axis: 0.0 to 1.0) vs Count of route (x-axis: 0 to 2500)

AKLPEK
CGKKTM
PENTPE
CGKHND
DMKKIX
AKLKUL
Average
ICNSIN
Average

Flight Duration and Insurance Take Up Rate

flightDuration hour..

	Yes	No
Less than 5 hours	26.38%	
5 to 8 hours	16.74%	83.26%
More than 8 hours	10.71%	89.29%

Count of Id (0K, 5K, 10K, 15K, 20K)

401 flights from PEN to TPE

© 2022 Mapbox © OpenStreetMap

Route count cutoff
63

Insurance take up ra..
0.15

Ins Flag
- (All)
- NO
- ● YES

ROUTE Origin
- ☑ PEN
- PER
- PNH
- PNK
- PUS
- PVG
- REP

ROUTE Destination
- SZX
- TGG
- ☑ TPE
- TRZ
- TWU
- URT
- UTP

Ins Flag
- ■ YES

% of Total Count of Id
10.71%   89.29%

Descriptive:

- Flight duration less than 5 hours has the highest travel insurance take-up rate; and longer the flight duration, lesser the travel insurance take-up rate.
- The top flight route with high number of route and high insurance take-up rate is PENTPE.

Diagnostic:

- For longer flight duration, the travel insurance premium may be more expensive.
- High number of tourists travel in short distance due to budget and time constraint, and these group of people are more likely to purchase travel insurance. Trip from Penang to Taipei is good example.

Predictive:

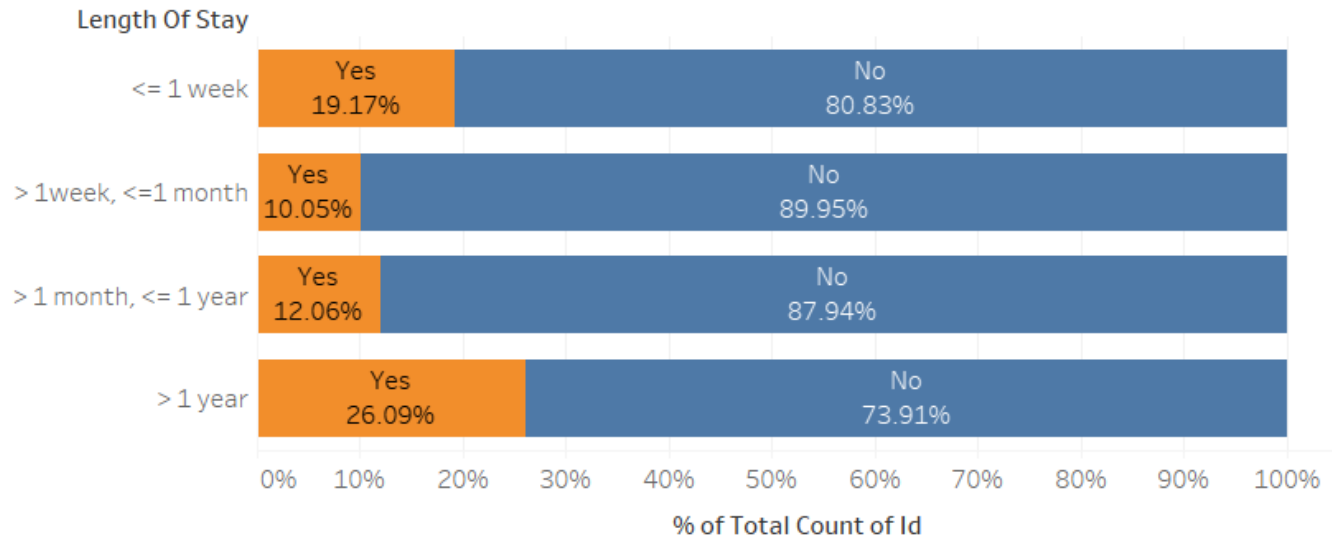- Travel insurance take-up rate would be higher for travels to tourist spot within short to medium flight duration.

Prescriptive:

- Prioritize marketing budget to target customer booking tourist destination with medium flight duration.

Length of Stay

Length Of Stay

Length Of Stay	Yes	No
<= 1 week	19.17%	80.83%
> 1week, <=1 month	10.05%	89.95%
> 1 month, <= 1 year	12.06%	87.94%
> 1 year	26.09%	73.91%

% of Total Count of Id

Baggage

Baggage ..

Baggage	Yes	No
No	11.50% Yes	88.50% No
Yes	16.67% Yes	83.33% No

% of Total Count of Id

Descriptive:

- There are more people opt for travel insurance when they are staying at the destination for less than a week or more than a year.
- People who check in baggage have higher probability of buying travel insurance

Diagnostic:

- In a short trip less than a week, more people are willing to spend on travel insurance to be free of worry during the trip.
- For trip longer than 1 week, the travel purpose is less likely for leisure purpose, there is less risk involved at the travel destination. The insurance is also more expensive compared to shorter trip.
- For long trip more than a year, long-stay travel insurance is useful for anyone who is studying abroad, working abroad temporarily.
- Longer stay also normally involves more baggage and hence insurance that covers missing baggage would be attractive to this group of people.
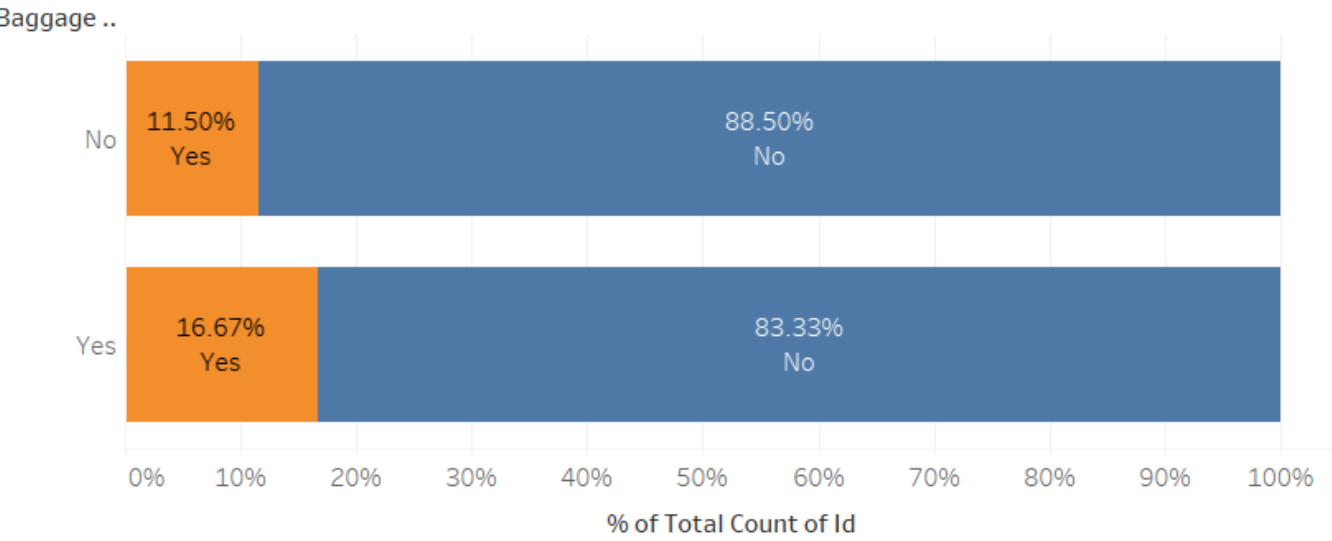
Predictive:

- People with short length of stay but more baggage has higher chance of purchasing travel insurance
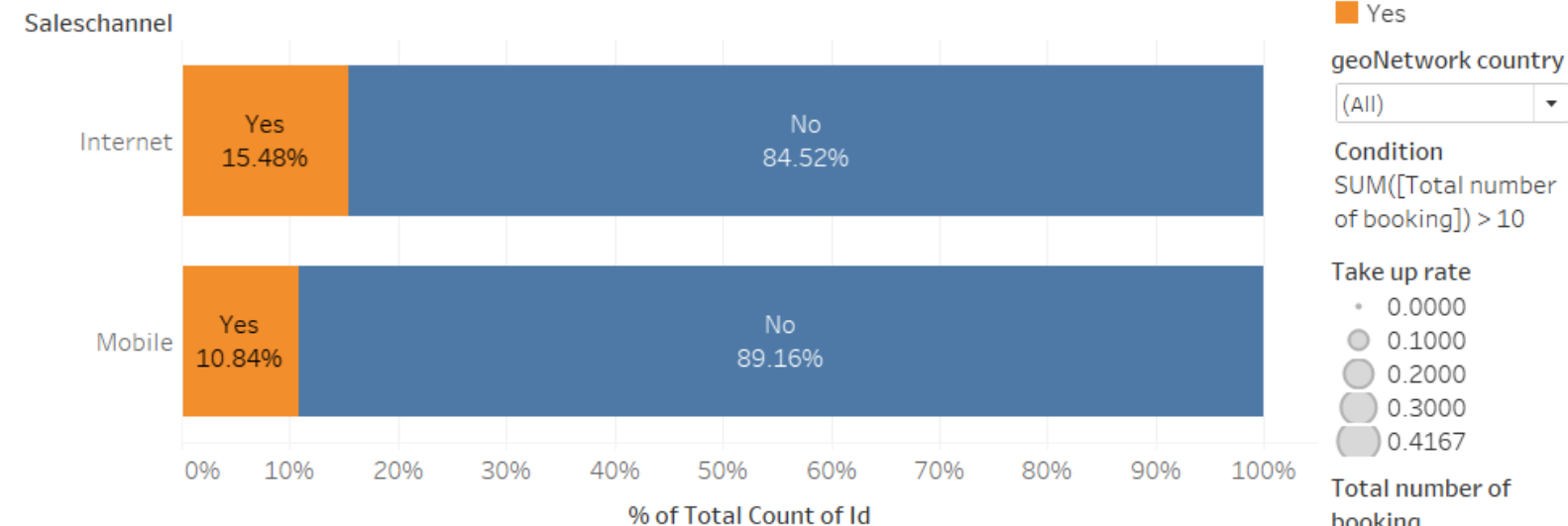
Prescriptive:

- Market a cheaper insurance package that bundles travel insurance with baggage check-in.
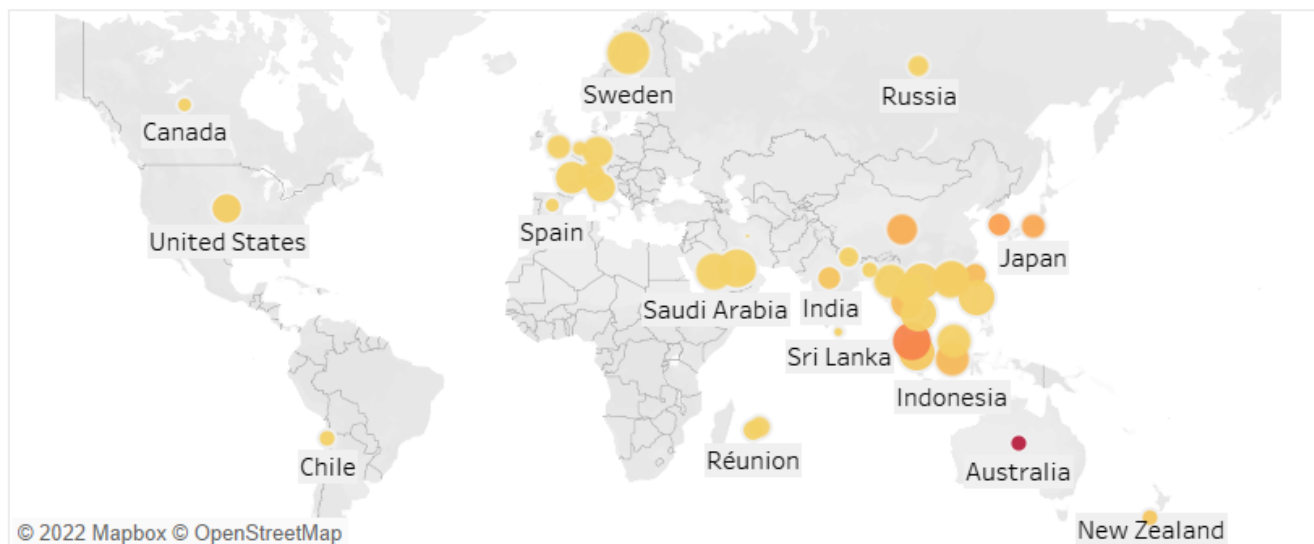
Sales Channel

Saleschannel

Internet
Yes 15.48% | No 84.52%

Mobile
Yes 10.84% | No 89.16%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
% of Total Count of Id

Ins Flag
- No
- Yes

geoNetwork country
(All)

Condition
SUM([Total number of booking]) > 10

Take up rate
- 0.0000
- 0.1000
- 0.2000
- 0.3000
- 0.4167

Total number of booking
12 — 17,872

Country from where booking was made with travel insurance

Canada, United States, Sweden, Russia, Spain, Saudi Arabia, India, Japan, Sri Lanka, Indonesia, Réunion, Chile, Australia, New Zealand

© 2022 Mapbox © OpenStreetMap

Descriptive:

- More people purchase travel insurance when they make travel booking through computer internet compared to mobile.
- Higher insurance take-up rate when bookings are made in Sweden, Qatar, Saudi Arabia, Macau, and Malaysia.
- Australia has highest number of travel booking but very low insurance take-up rate.

Diagnostic:

- In most countries, computer internet is still the most accessible and convenient form to purchase flight ticket and travel insurance. The benefits of travel insurance are better presented by computer platform compared to mobile platform.
- Countries with high number of booking such as Australia, Japan, and South Korea have comparatively high number of advance mobile user, posing high potential of growth.
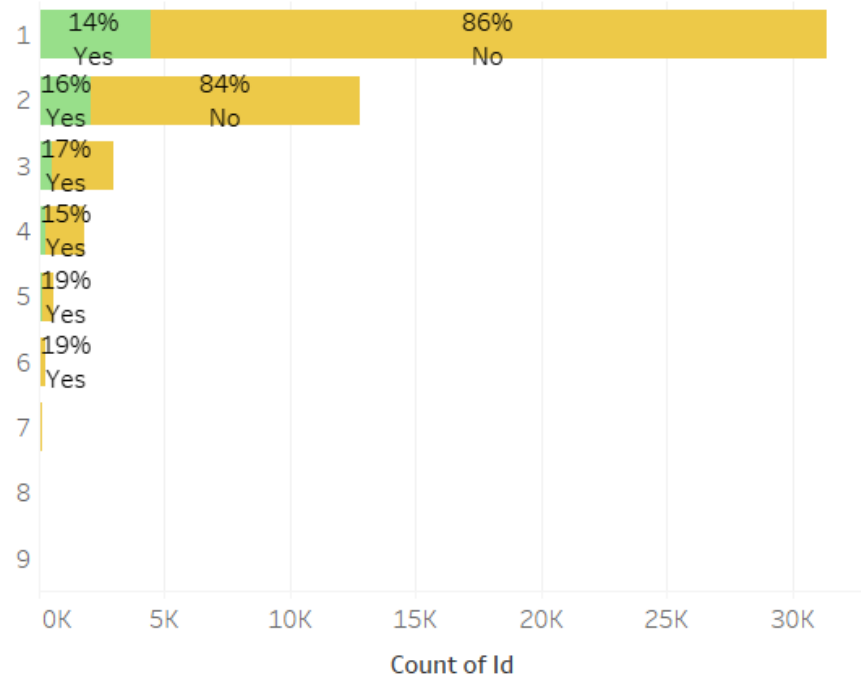
Predictive:

- The % of mobile user buying travel insurance will increase if the insurance benefits can be made simple and clear on the mobile and information navigation is user-friendly.

Prescriptive:

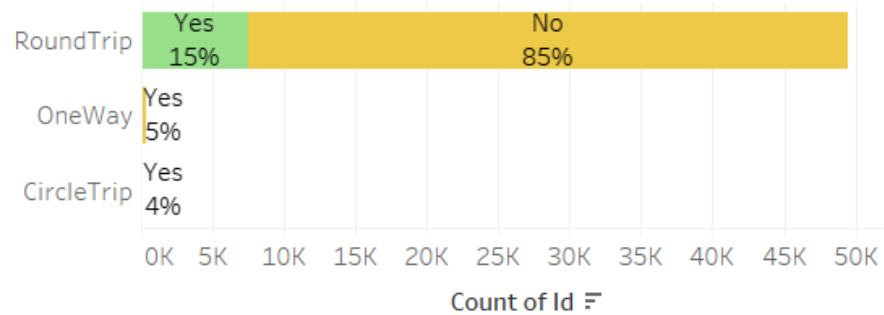- Improve mobile flight ticket and travel insurance platform.
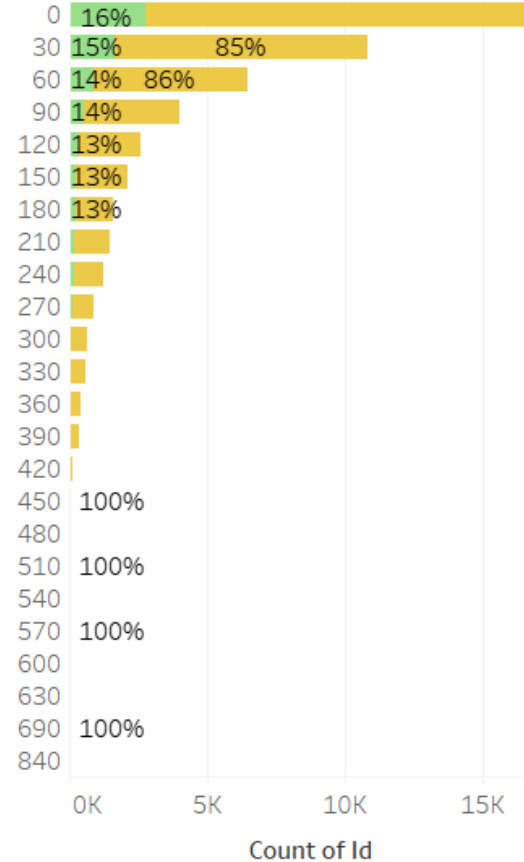
Number of customer traveling

Paxcount..

1	14% Yes — 86% No
2	16% Yes — 84% No
3	17% Yes
4	15% Yes
5	19% Yes
6	19% Yes
7	
8	
9	

Count of Id (0K, 5K, 10K, 15K, 20K, 25K, 30K)

Number of days between travel date and booking date

Purchaselead (..

0	16%
30	15% — 85%
60	14% — 86%
90	14%
120	13%
150	13%
180	13%
210	
240	
270	
300	
330	
360	
390	
420	
450	100%
480	
510	100%
540	
570	100%
600	
630	
690	100%
840	

Count of Id (0K, 5K, 10K, 15K)

Ins Flag
- No
- Yes

Trip Type

Triptypedesc

RoundTrip	Yes 15% — No 85%
OneWay	Yes 5%
CircleTrip	Yes 4%

Count of Id (0K, 5K, 10K, 15K, 20K, 25K, 30K, 35K, 40K, 45K, 50K)

Descriptive:
- Higher insurance take-up rate when more than 1 customer travelling.
- Higher insurance take-up rate for shorter period between booking date and travel date.
- Much higher insurance take-up rate for round trip compared to one-way a circle trip.

Diagnostic:
- It is likely there are children or elderly family members when travelling more than 1 person.
- It is less viable to make personal arrangement if the trip is cancelled in short notice, hence making the insurance purchase a good choice.
- Round trip travel are normally travel for leisure, which is the category of travel that more people will book with travel insurance.

Predictive:
- The annual trip with family members for leisure would likely have travel booking together with travel insurance.

Prescriptive:
- Improve the flight ticket purchasing system so that when it detects family travelling to famous tourist spot, the relevant insurance benefits are highlighted to persuade the purchase.

Analysis Limitation

- Some important data are not given in the Dataset downloaded from Kaggle, such as:
  - Type of travel insurance plan purchased
  - Insured's age
  - Travel purpose
  - Flight ticket cost and travel insurance cost

- Difficulty in putting travel origin and destination as a paired data in map analysis.
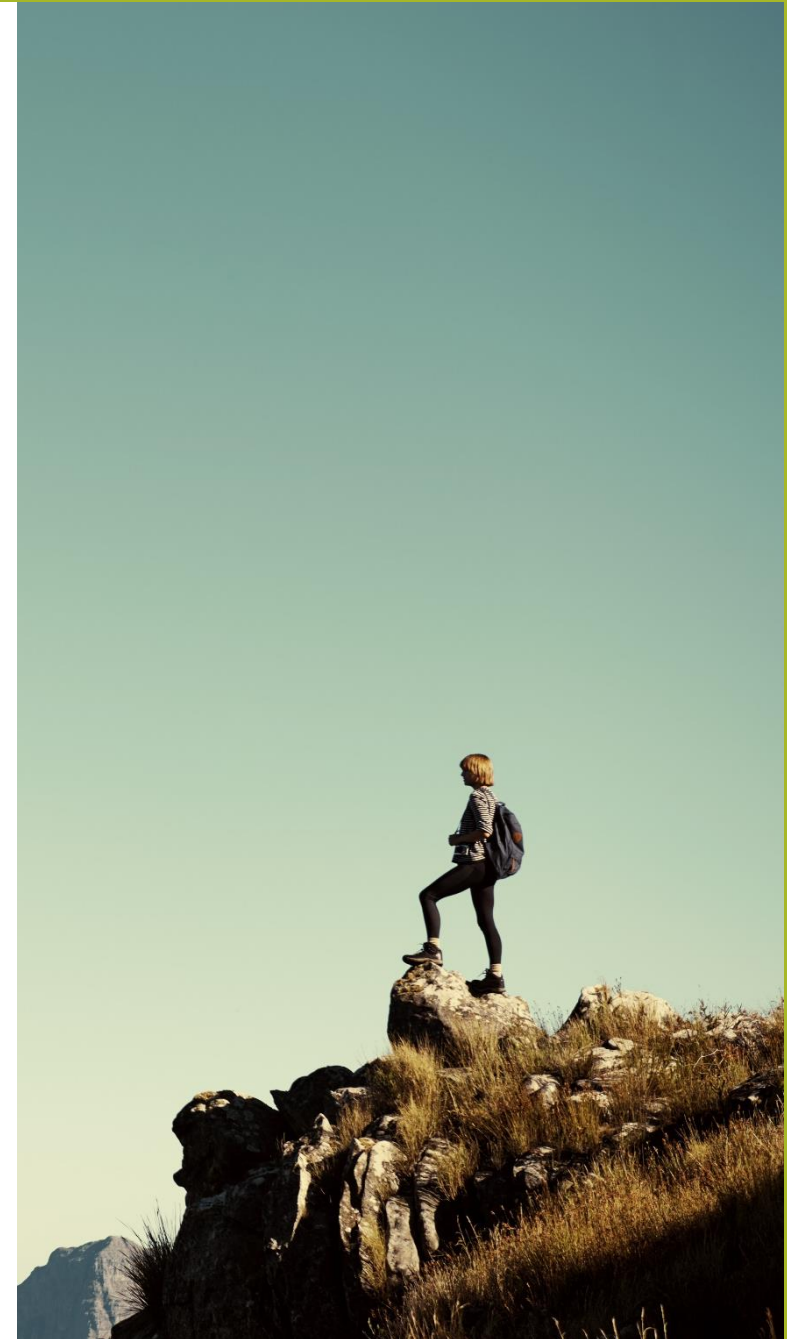
Conclusion

- Prioritize marketing budget to target customer booking tourist destination with medium flight duration.

- Market a cheaper insurance package that bundles travel insurance with baggage check-in.

- Improve mobile flight ticket and travel insurance platform to make it more user friendly when navigating information of travel insurance.

- Improve the flight ticket purchasing system so that when it detects family travelling to famous tourist spot, the relevant insurance benefits are highlighted to persuade the purchase.

Thank You

https://public.tableau.com/app/profile/lee.boon.chek/viz/CapstoneLeeBoonChek433A/Story1?publish=yes