Improving Customer Retention for Invistico Airlines

Sowmiya Muruganandam

Industry Background

- Hyper-Competitive Industry
  - Focused on how to retain existing customers and acquire new customers.
- Customer Satisfaction
  - Primary non-pecuniary indicator and precursor to determining customer loyalty.
  - Satisfaction data is more accessible than customer loyalty data.
- Customer Loyalty
  - Loyal customers tend to have higher retention and lower likelihood to switch to different companies.
  - Literature shows that loyal customers spend more money and make more frequent purchases.

Business Problem

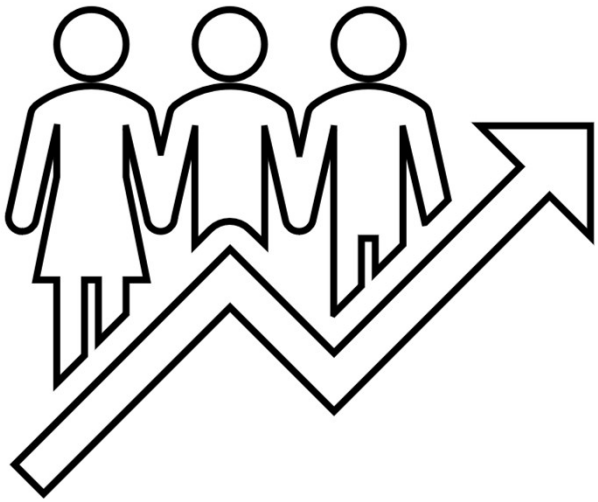
Implement marketing strategies for each "customer segmentation" to maximize customer retention by promoting customer satisfaction and loyalty.

Available Data

- Customer satisfaction survey data
  - Surveys containing feedback on flight experience and other relevant information.

Data Source

- Kaggle Data Sets
  - One year of data from an anonymous airline company.



Business Objective: Maximize Customer Retention

- Use available data to identify determinants of Customer Satisfaction and Customer Loyalty.
- Customer Satisfaction: Minimum standards that lead to loyalty.
- Customer Loyalty: Higher standards which encourage future commitment.

Customer Retention: Satisfaction + Loyalty

Our Approach to the Problem

- Identify Satisfied vs Dissatisfied customers
  - Classification modeling
- Create Loyalty vs Customer Satisfaction matrix
  - Crosstab
- Develop appropriate marketing decisions
  - Using customer segmentations in the matrix

Analytic Tools/Methods used

Programming Language: R

Descriptive Analysis

Data Overview

- The dataset consists the details of customers from the past flights and individual's feedback on various context
- Dimension: (129880, 23)
- In the model construction, we split the dataset into 70% training and 30% testing groups

	satisfaction	Gender	Customer.Type	Age	Type.of.Travel	Class	Flight.Distance	Seat.comfort
1	satisfied	Female	Loyal Customer	65	Personal Travel	Eco	265	0
2	satisfied	Male	Loyal Customer	47	Personal Travel	Business	2464	0
3	satisfied	Female	Loyal Customer	15	Personal Travel	Eco	2138	0
4	satisfied	Female	Loyal Customer	60	Personal Travel	Eco	623	0
5	satisfied	Female	Loyal Customer	70	Personal Travel	Eco	354	0
6	satisfied	Male	Loyal Customer	30	Personal Travel	Eco	1894	0

	Departure.Arrival.time.convenient	Food.and.drink	Gate.location	Inflight.wifi.service	Inflight.entertainment
1	0	0	2	2	4
2	0	0	3	0	2
3	0	0	3	2	0
4	0	0	3	3	4
5	0	0	3	4	3
6	0	0	3	2	0

	Online.support	Ease.of.Online.booking	On.board.service	Leg.room.service	Baggage.handling	Checkin.service	Cleanliness
1	2	3	3	0	3	5	3
2	2	3	4	4	4	2	3
3	2	2	3	3	4	4	4
4	3	1	1	0	1	4	1
5	4	2	2	0	2	4	2
6	2	2	5	4	5	5	4

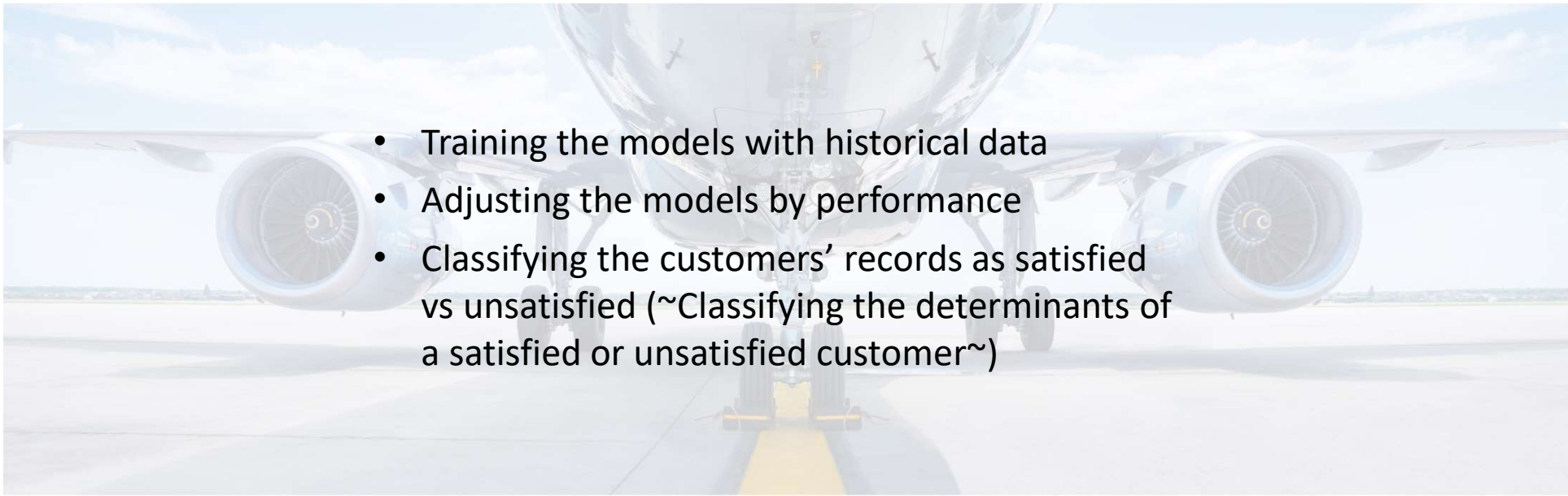
	Online.boarding	Departure.Delay.in.Minutes	Arrival.Delay.in.Minutes
1	2	0	0
2	2	310	305
3	2	0	0
4	3	0	0
5	5	0	0
6	2	0	0

> |

Classification Analysis

Models used:

- Logistics regression
- RandomForest
- Naïve-Bayes



- Training the models with historical data
- Adjusting the models by performance
- Classifying the customers' records as satisfied vs unsatisfied (~Classifying the determinants of a satisfied or unsatisfied customer~)

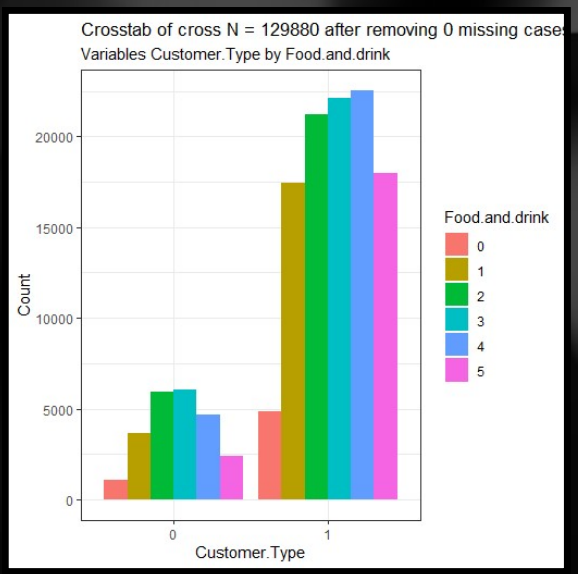
Classification Summary

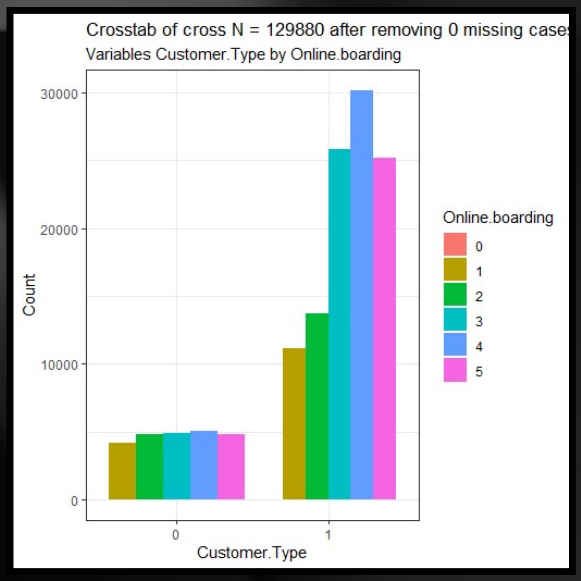
Method1 - Logistic regression							
Regression Model	Description	BIC	TP	FP	TN	FN	Accuracy
Regression 1	Full logistic regression **with all columns**	70223.32	18237	3235	14272	3220	83.4%
Regression 2&3	Logistic regression with **only survey columns**; regression 3 used in prediction (insignificant IV removed from reg2)	Regression 2=80101.86 Regression 3=80080.73	17788	3872	13635	3669	80.6%
Regression 4&5	Logistic regression **without Customer.Type (Loyalty)**; regression 5 used in prediction (insignificant Ivs removed from reg 4)	Regression 4=74940.81 Regression 5=74929.43	18068	3535	13972	3389	82.2%

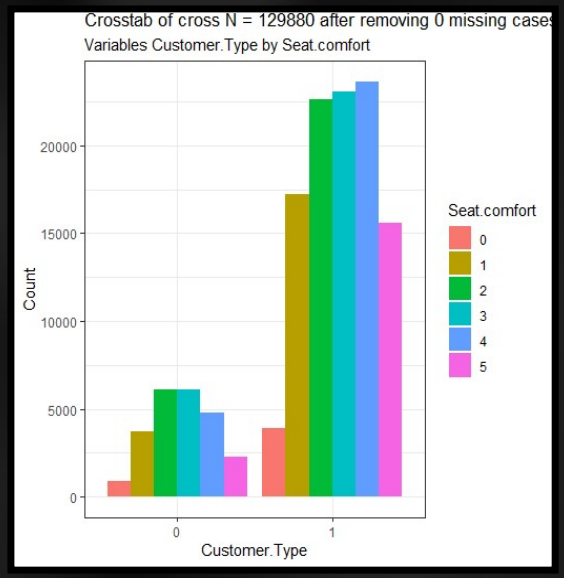
Classification Summary

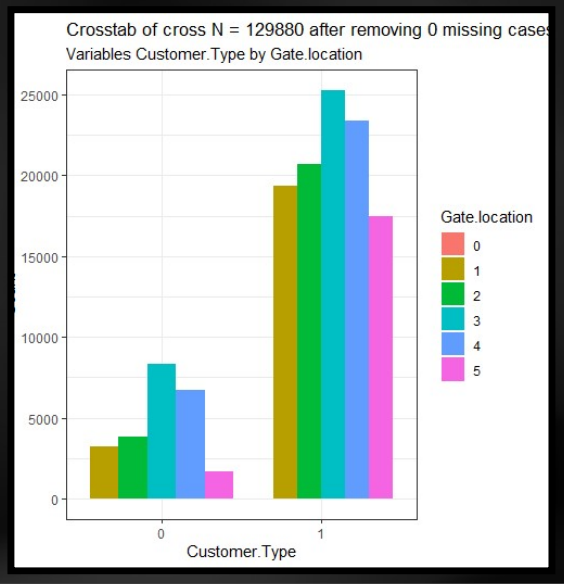
Method 2 – RandomForest						
Random Forest Model	Description	TP	FP	TN	FN	Accuracy
Model 1	ntree=50; with all columns	20332	772	16735	1125	95.130%
Model 2	ntree=200; with all columns	20326	742	16765	1131	95.190%
Method3 - Naïve-Bayes						
Model	Description	TP	FP	TN	FN	Accuracy
Model 1	with all columns	18034	4032	13475	3423	80.87%

Cross Tab Analysis

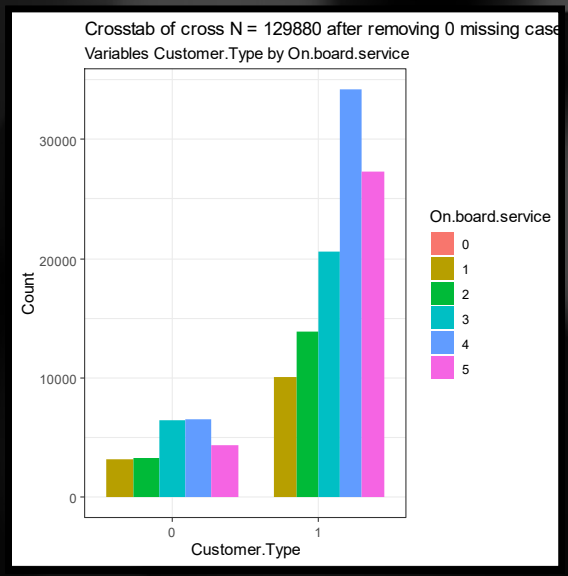


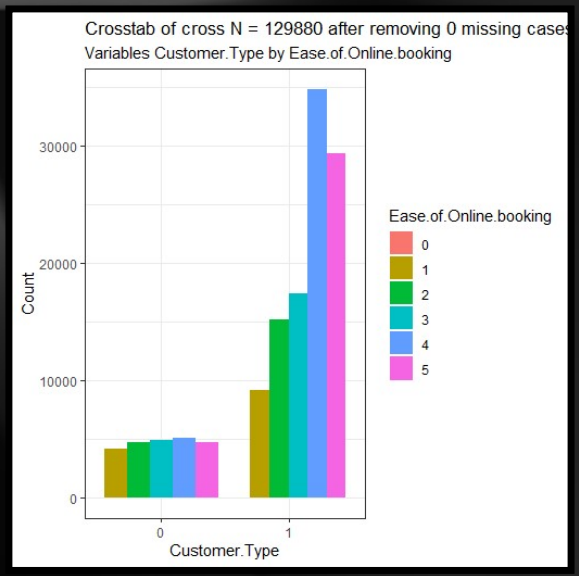


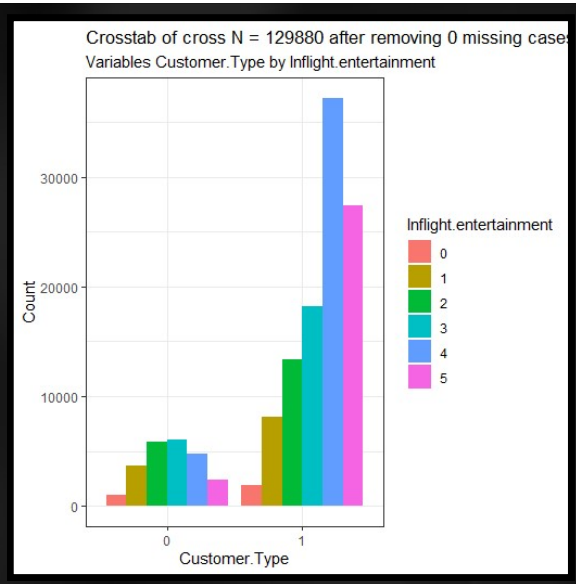


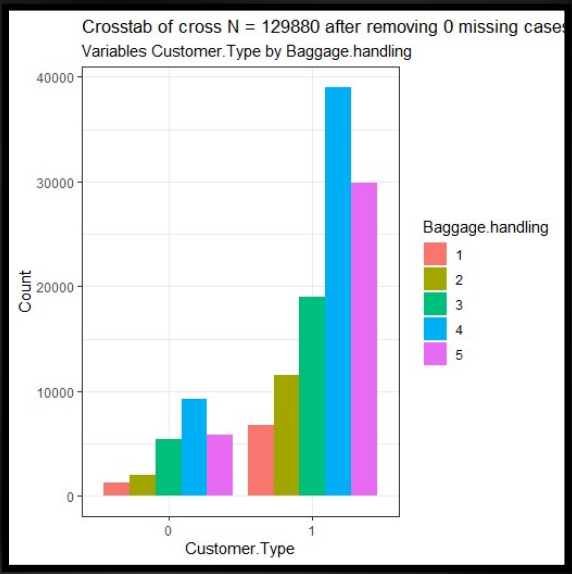


Cross Tab Analysis









Identified Pain Points from Crosstab

Loyal-Satisfied Customers (65387)

Pain Points
Baggage handling
On-board service
Gate location

Disloyal -Satisfied Customers (5700)

Pain Points
Seat comfort
Online boarding
Checkin service

Loyal-Dissatisfied Customers (40713)

Pain Points
Ease of Online booking
Baggage handling
Seat comfort

Disloyal-Dissatisfied Customers (18080)

Pain Points
Inflight entertainment
Food and drink
Seat comfort

Marketing Decisions to make

Loyal-Satisfied Customers (65387)

1. Cross Sell/Up sell Inflight services
2. Discounts/Promo Offers
3. Customization
4. Partnership Programs

Disloyal -Satisfied Customers (5700)

1. Use mark up/down pricing strategies

2. Email Campaigns to keep them coming back

3. Explore new features/services

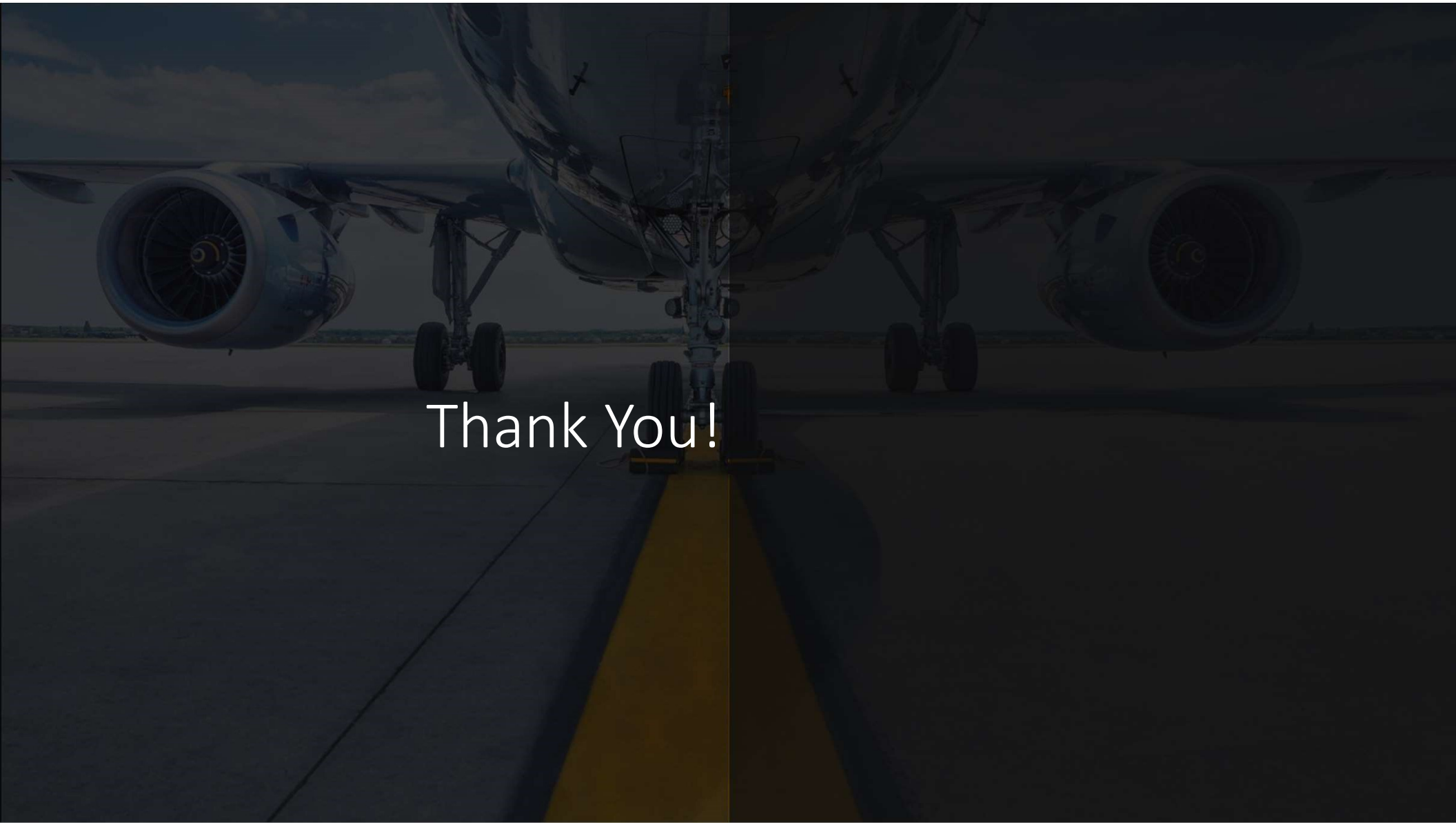
Loyal-Dissatisfied Customers (40713)

1. Focus on customer pain points
2. “We Fixed it” Email Campaigns
3. Reward Program/ Upgrades
4. Better Customer service

Disloyal-Dissatisfied Customers (18080)

Root Cause Analysis

1. Scope for New feature/product development
2. Choose to let go if the ROI is lower than the customer turnaround scenario


Thank You!

Reference

- Jana, S. (2020, March 19). Airlines customer satisfaction. Retrieved February 20, 2021, from https://www.kaggle.com/sjleshrac/airlines-customer-satisfaction
- Eunil Park, et. al. Determinants of customer satisfaction with airline services: An analysis of customer feedback big data, from https://www.sciencedirect.com/science/article/abs/pii/S0969698919304369
- Kevin Tegar B., Anggun Lestari, Sekar Widyastuti Pratiwi (2017). AN ANALYSIS OF AIRLINES CUSTOMER SATISFACTION BY IMPROVING CUSTOMER SERVICE PERFORMANCE. Retrieved February 21, 2021, from https://www.atlantis-press.com/proceedings/grost-17/25889417
- Khadka Kabu, Maharjan Soniya (2017 November). Customer Satisfaction and Customer Loyalty. Retrieved February 21, 2021 from https://core.ac.uk/download/pdf/161421179.pdf.
- Crosstab - a CROSS-TABULATION function for use with survey data. (n.d.). Retrieved February 22, 2021, from http://rstudio-pubs-static.s3.amazonaws.com/6975_c4943349b6174f448104a5513fed59a9.html