Customer Segmentation for product marketing

Our market proposal for a successful email campaign

LULUORANGE

Our Understanding of the Business Problem

Product Campaign to increase the sales of Men's Pants in Men Category

- Analysis on Customer Data (By 2020 May-Jun)
- Channel: In Store and Online
- Customer Campaign List -> ROI

LULUORANGE

Our Proposal For A Successful Email Campaign

Decision Needed: Campaign approach to distribute promotion emails

TOP CRITERIA

+ NUMBER OF MEN'S PANTS
+ TOTAL MONEY SPENT ON MEN'S
+ TOTAL # OF PRODUCTS PURCHASED
- AVERAGE VALUE OF TRANSACTIONS
- PERCENTAGE MARKDOWN ITEMS
- TOTAL WOMEN'S SPEND

SIZE OF CUSTOMERS

3015

ESTIMATED ROI

303%+

LULUORANGE

Approach to Problem

1. Study Data Insights with Exploratory Data Analysis
2. Two-Pronged Strategy - Machine Learning & RFM Model
3. Model Design - Feature Engineering
4. Results - Target Customers & ROI

Domain Research

Engagement:
- Last visit to website.
- Email opening rate

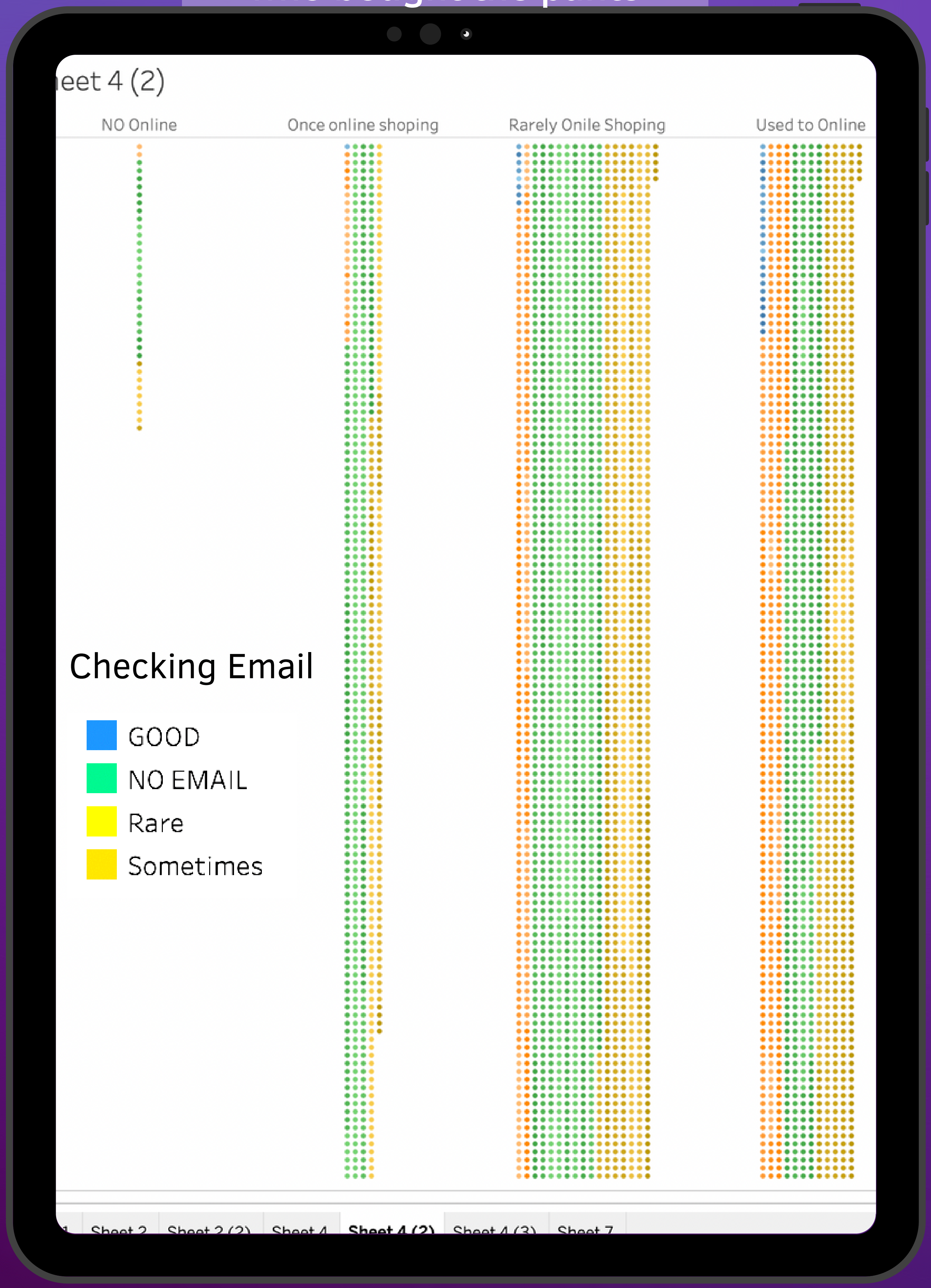
Purchase History:
- Gender-based analysis of purchase history

Purchase Amount and Basket Size: Amount and frequency of purchases

LULUORANGE

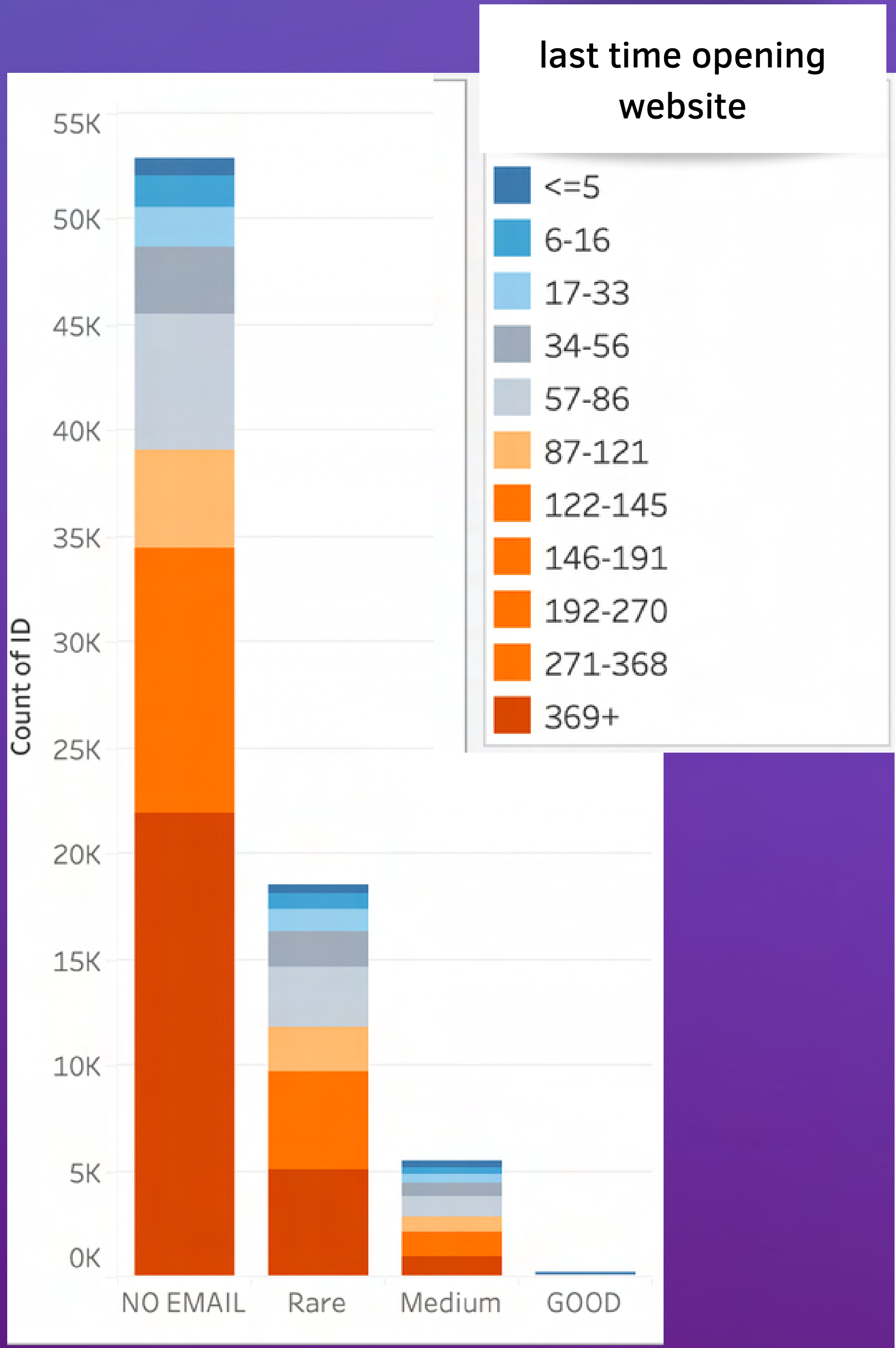
DATA INSIGHTS

Online shopping Record for who bought the pants

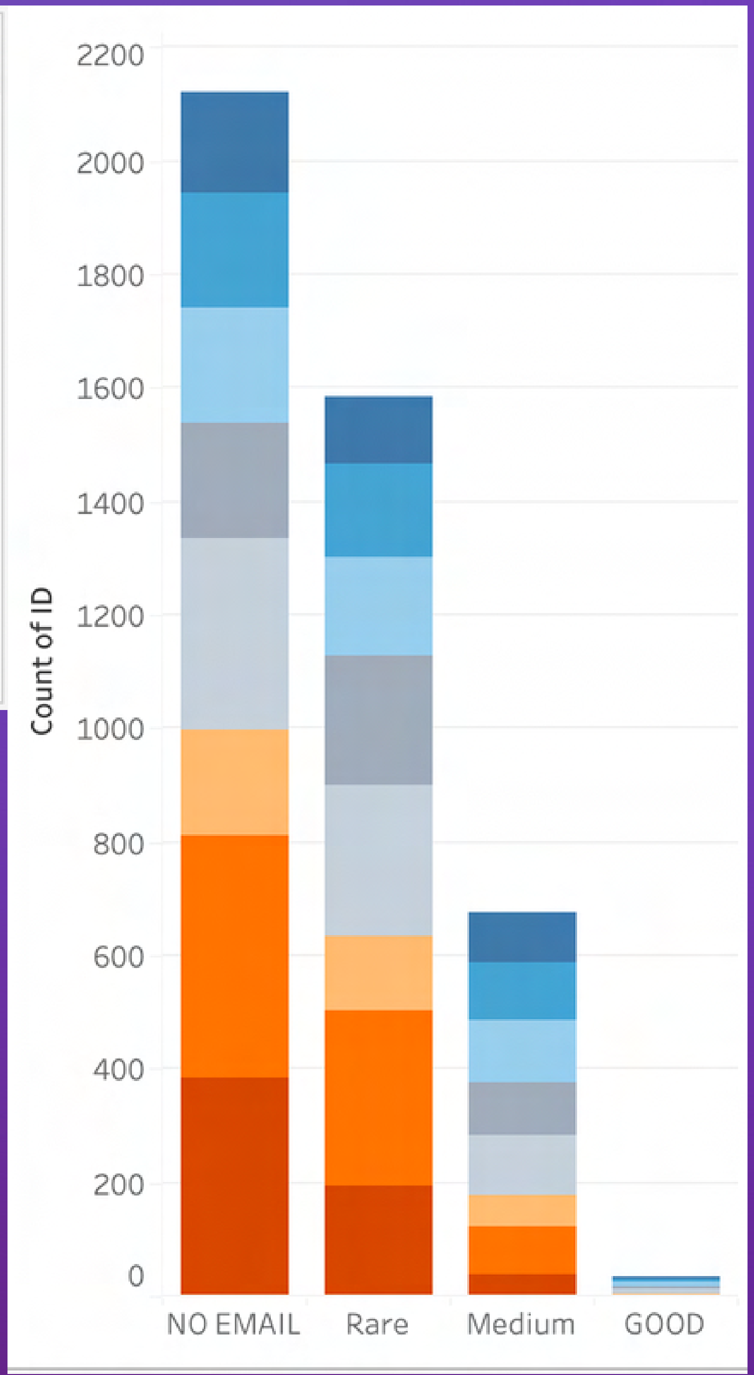


LULUORANGE

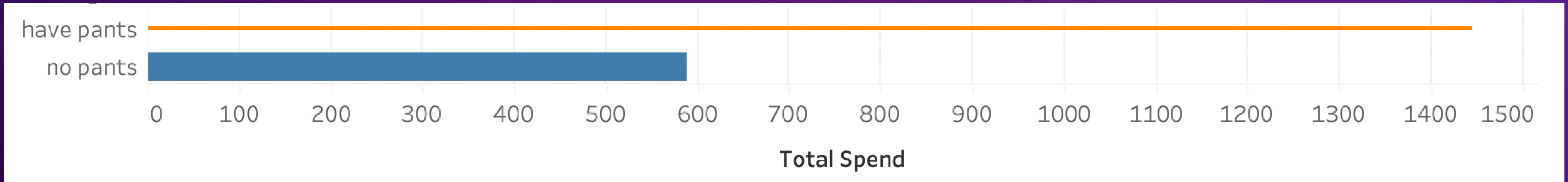
DATA INSIGHTS



costumer who did not buy the product



costumer bought the product



LULUORANGE

TWO-PRONGED APPROACH

RFM MODEL

User Profile Evaluation

- Recency score to evaluate the time gap since last purchasement
- Frequency score to evaluate how often a customer purchases a good
- Monetary score to evaluate how much a customer spend

MACHINE LEARNING MODEL

Probability Calculation

- Logistic Regression Model
  - Easy to implement
  - Efficient on binary classification problem
  - Result can be interpreted to probability

LULUORANGE

RFM MODEL DESIGN

Step 1

Calculate Recency, Frequency, Monetary score

Step 2

Fulfill operation strategy matrix

Classification	Recency	Frequency	Monetary	Operation Strategy
Critical Valuable	↑	↑	↑	Keep as-is
Critical Developable	↑	↓	↑	Improve Frq
Critical Persistent	↓	↑	↑	General Recall
Critical Preserved	↓	↓	↑	Critical Recall
General Valuable	↑	↑	↓	Boost Consuption
General Developable	↑	↓	↓	Mining Needs
General Persistent	↓	↑	↓	Loss Recall
General Preserved	↓	↓	↓	No Action (54% in dataset)

LULUORANGE

ML Model Design

DATA CLEANING

- FILLING MISSING DATA
- ONE-HOT ENCODING

FEATURE SELECTION

- XGBOOST
- L1 REGULARIZER

DATA SPLIT

- TRAIN: 80%
- TEST: 20%

MODEL BUILDING

- LOGISTIC REGRESSION
- AUC : 0.8

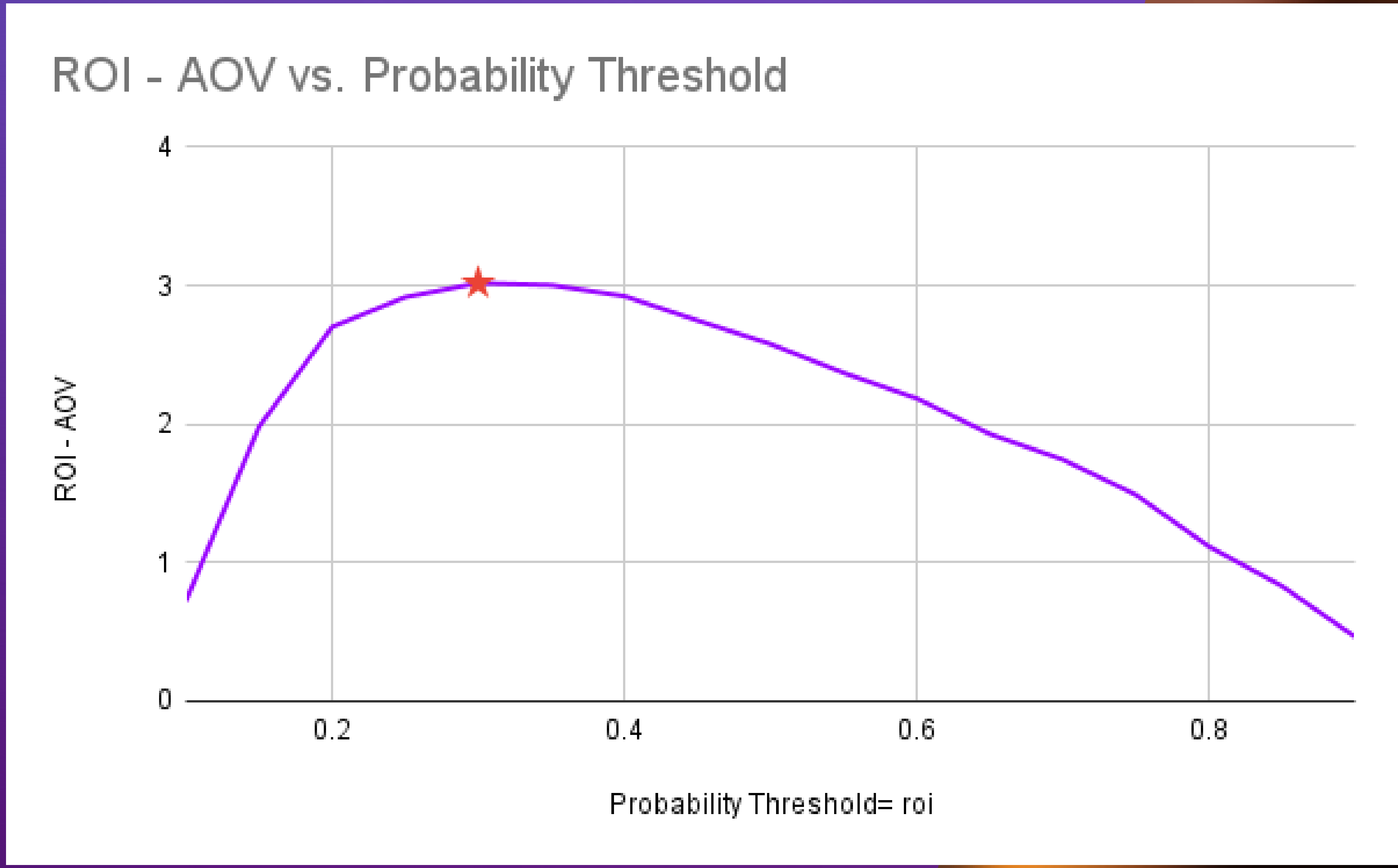
PREDICTION

- PREDICT ON TEST DATASET TO GET PROBABILITY

LULUORANGE

Final Results and Calculations of ROI

Probability Threshold	0.3
# of Customers Identified	603
Projected # of Customers in Customer Base	3,015
Estimated Cost	52,889
Projected Potential Benefits - by AOV	212,511
Projected Potential Benefits - by Pant Unit Price	213,073
ROI - AOV	3.02
ROI - Pant Unit Price	3.03



LULUORANGE

PREDICTIONS FOR EFFECTIVENESS OF EMAIL CAMPAIGN

RFM SIDE FILTER

46% of customers were likely to be excellent targets for an email campaign

SIZE OF CUSTOMERS

Based on different probability threshold calculations, we believe that the range of the target customers of this campaign is:

3015

ESTIMATED ROI

- Design Cost: $25,000
- Email Cost: $27,888.75
- Total Cost: $52,888.75
- Design Amortization Cost: $8.29/person
- Potential Benefits: $213,073.58
- ROI: **303%+**

LULUORANGE

Thank You


Chaucer Qiu


Jeff Wang


Parnian Taghipour


Rovenna Chu


Smitha Kolan


Zhi Zheng

LULUORANGE

Q&A

LULUORANGE

Appendix

LULUORANGE

Final Results and Calculations of ROI

Probability Threshold	0.25	0.3	0.35	0.45	0.55	0.65	0.7	0.75
# of Customers Identified	769	603	486	339	241	182	158	137
Projected # of Customers in Customer Base	3,845	3,015	2,430	1,695	1,205	910	790	685
Estimated Cost	60,566	52,889	47,478	40,679	36,146	33,418	32,308	31,336
Projected Potential Benefits - by AOV	237,234	212,511	190,072	152,501	121,832	97,892	88,649	78,055
Projected Potential Benefits - by Pant Unit Price	241,971	213,073	188922	151,480	120,295	97,529	87,278	77,571
ROI - AOV	2.92	3.02	3.00	2.75	2.37	1.93	1.74	1.49
ROI - Pant Unit Price	3.00	3.03	2.98	2.72	2.33	1.93	1.70	1.48

LULUORANGE

Contribution of Each Team Member

Team Member	Contributions
Chaucer Qiu	Built RFM model with Smitha as a main factor of customer evaluation. Kick-off and worked with the team, coordinated if required. Set objectives for the team & use case.
Smitha Kolan	Fine-tuned statistical RFM model with added features based on correlation factor. Set clear objectives for the team & use case. Designed and organized the flow of the presentation.
Jeff Wang	Built the ML model from start to end, including cleaning, feature transformation, feature selection, model selection and optimization, model evaluation, probability generation.
Rovenna Chu	Prepared initial machine learning models with explanatory ML, Shaped Problem Statement and Summarizing Proposals, Prepared ROI Calculations as Final Results
Zhi Zheng	Ranked the importance of machine learning features and assisted in the feature selection; Calculated the final result based on Rovenna's work
Parnian Taghipour	Working on visualizing data and EDA, Doing domain research on the problem, Introduced the function that needed to be optimized for calculating the final number of selected costumer based on the models results

LULUORANGE