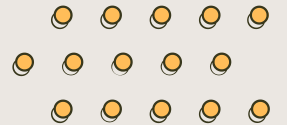


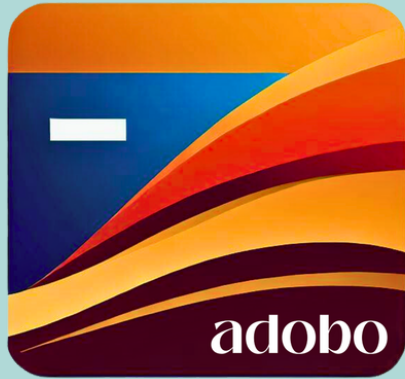
DECODE ENCODE

*Driving business growth through Clustering
and Linear Regression Modelling*



ARDEN, AUSTINE, BEA, EVIAN, JAPHET

DATASQU4D
DSFC13 GROUP 4



Adobe Advantage Cards (AAC), a credit card company, collected *comprehensive dataset of credit card transactions* which includes:

transaction details

customer demographics

amount spent in various categories



ABOUT THE CHALLENGE



DATASQU4D, one of the data science consultancy teams, was challenged to turn this **data** into actionable strategies that shall:

✓ ***Enhance customer satisfaction***

✓ ***Drive business growth***

METHODOLOGY



*Data
Preprocessing*



*Exploratory
Data Analysis*



*RFM and
K-means
Clustering*



*Cluster
Analysis*



*Linear
Regression
Modelling*



CUSTOMER DEMOGRAPHICS

Based on AAC Credit Card Transactions as of December 7, 2021



94%

88 MALE

6%

6 FEMALE

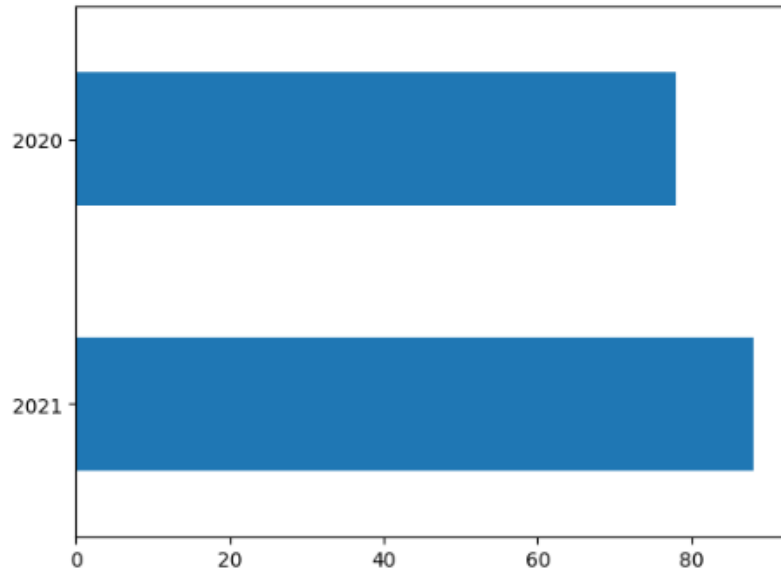
51- 95

AGE RANGE

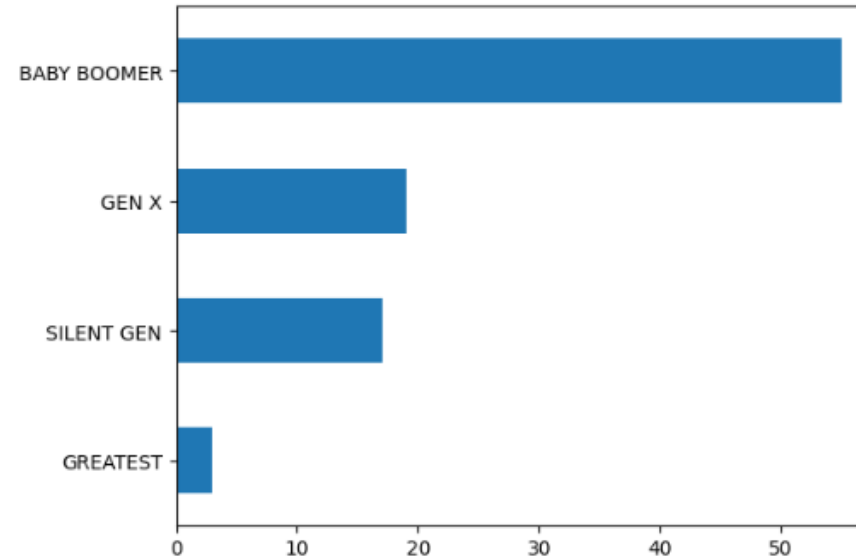
66

MEAN AGE

Active Customers per year



Generation Distribution



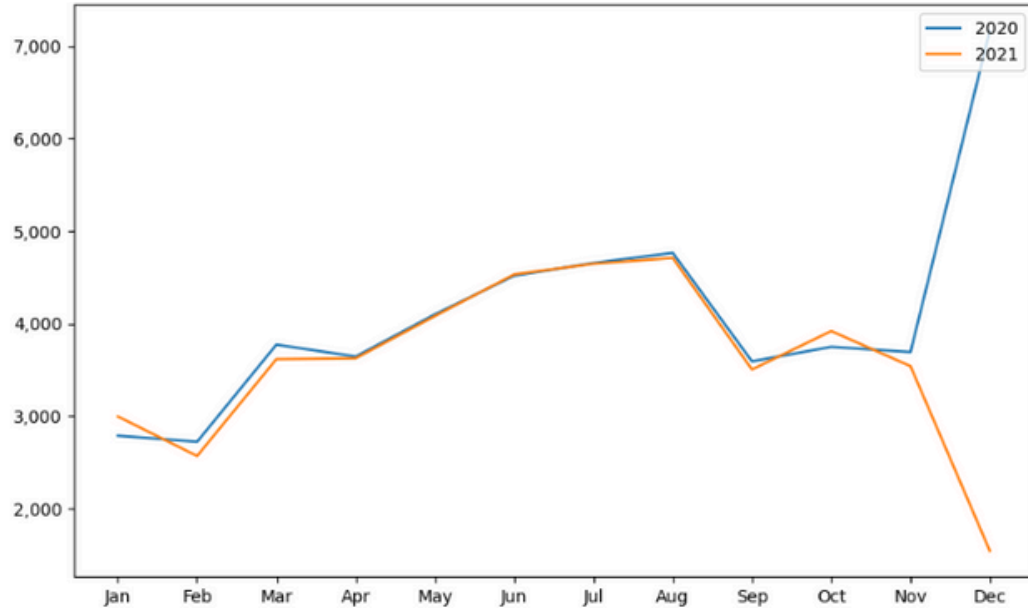


TRANSACTIONS AND SPENDING

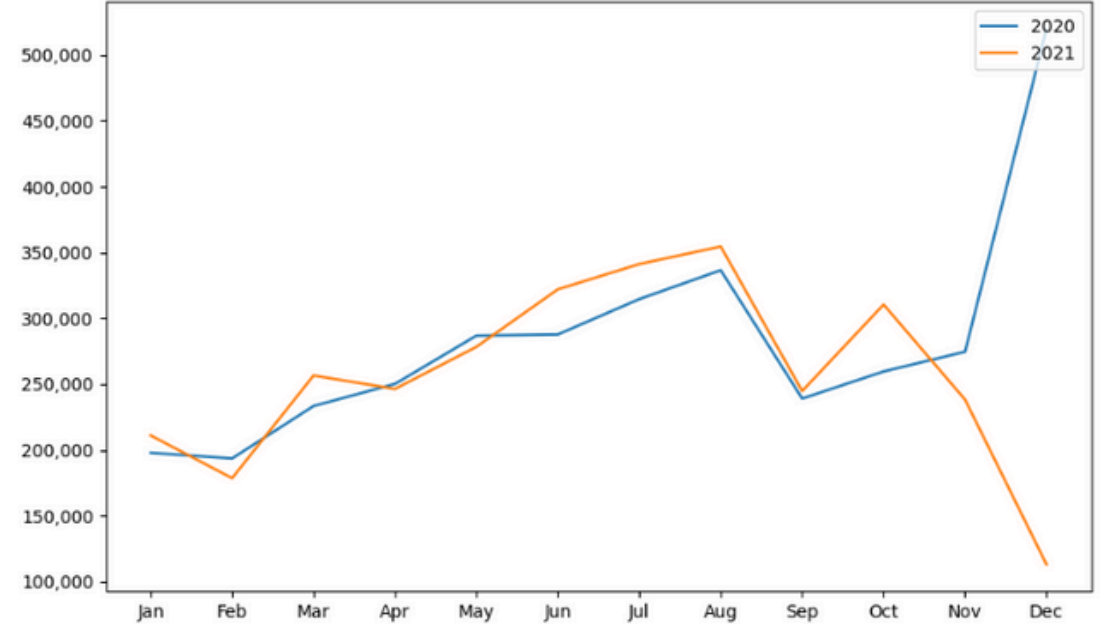
Based on AAC Credit Card Transactions as of December 7, 2021



YoY Monthly Transaction Counts



YoY Monthly Transaction Amount



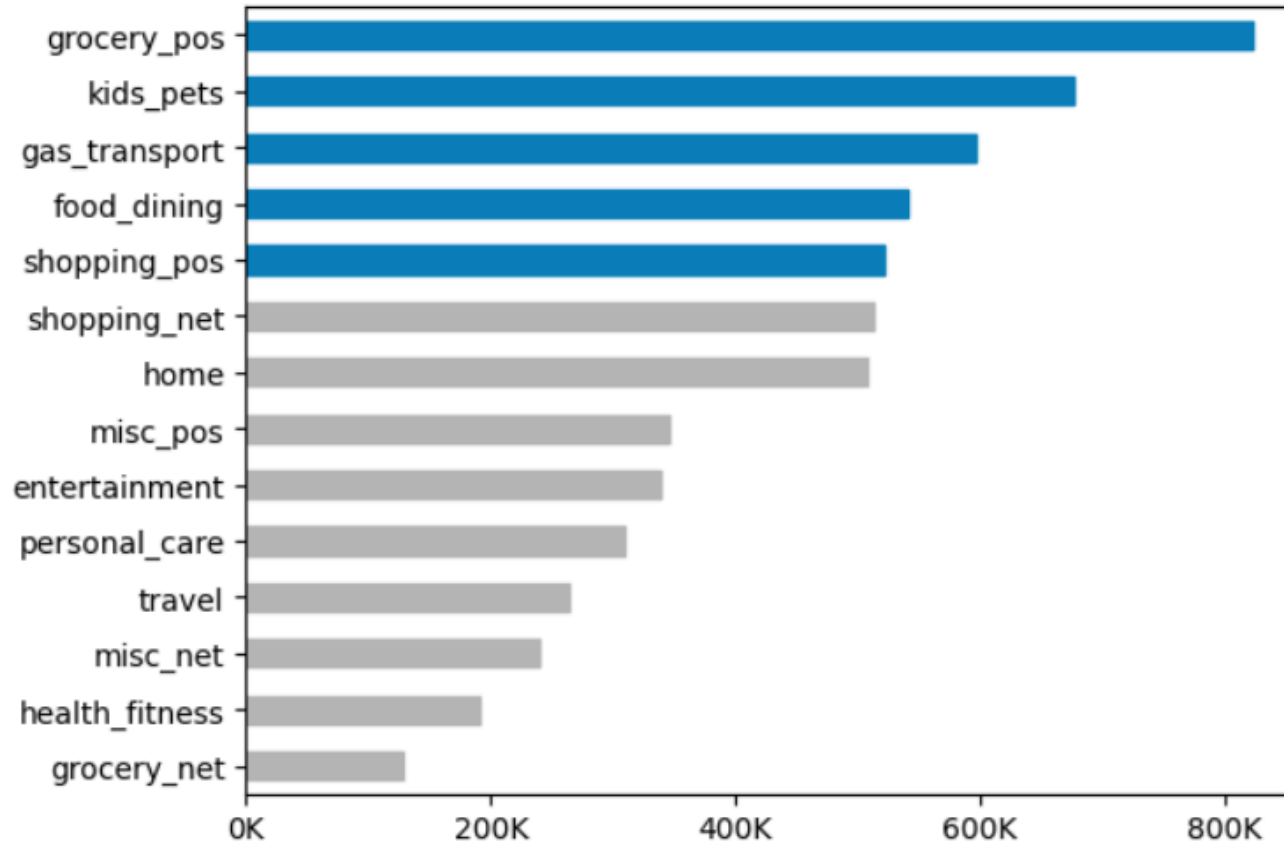


CATEGORY SPENDING

Based on AAC Credit Card Transactions as of December 7, 2021



All Time Top 5 Categories by Total Amount



AAC Credit Card Holders are **family-oriented, mobile, socially active, and financially confident . . .**

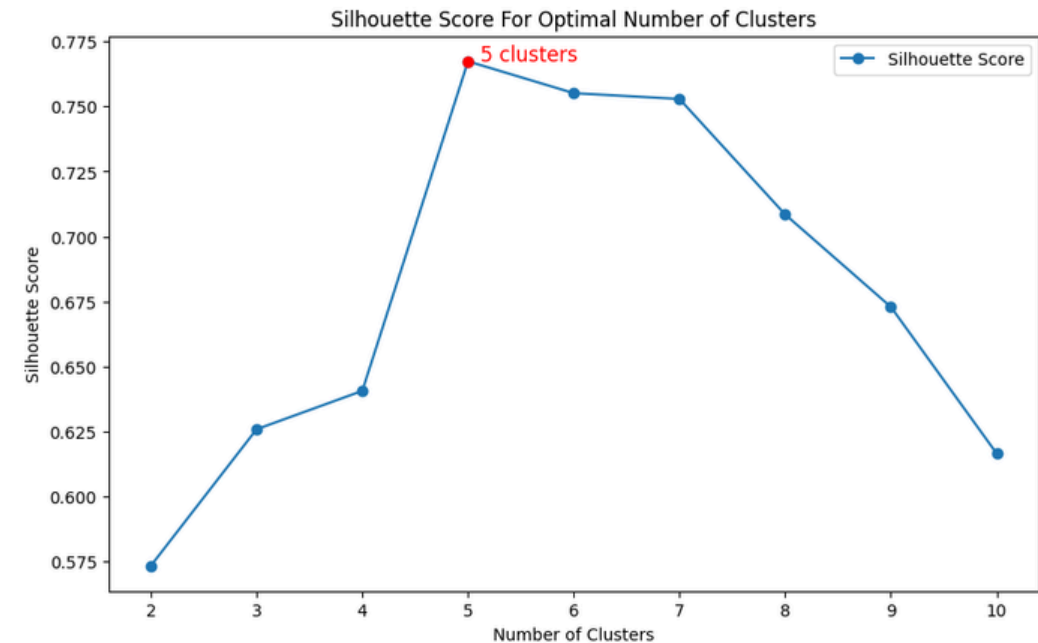
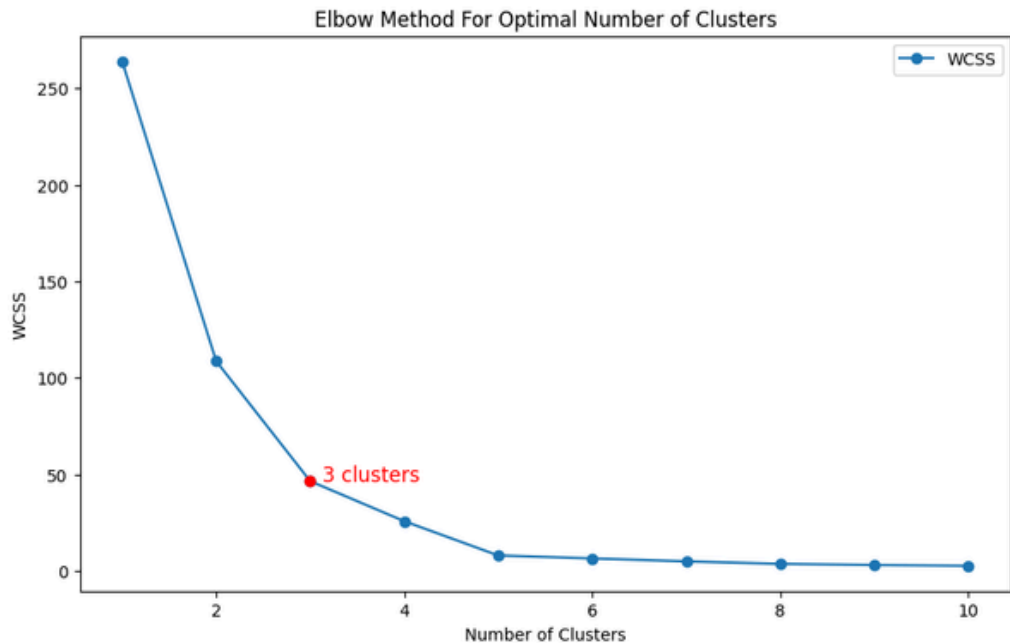


CUSTOMER SEGMENTATION

Choosing Appropriate Number of Clusters for the Data Set



Elbow Method and Silhouette Score for Optimal Number of Cluster



- The **3-cluster model** offers a simpler segmentation of **high, medium, and low-value customers**. While the 5-cluster model provides more detailed insights, identifying 5 distinct subgroups

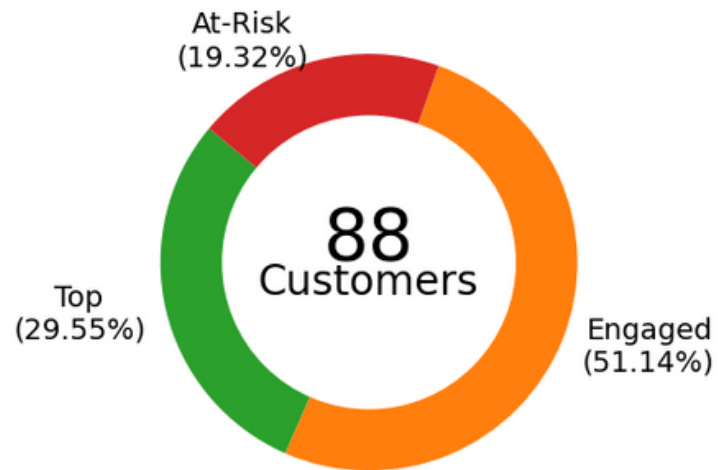


CUSTOMER SEGMENTATION

Based on AAC Credit Card Transactions as of December 7, 2021



Customer Distribution Across Clusters



K-Means Clustering Results

	Mean Recency	Mean Frequency	Mean Monetary	Number of Customers
Top	25	906	\$63,802	26
Engaged	25	435	\$30,000	45
At - Risk	192	8	\$5,130	17

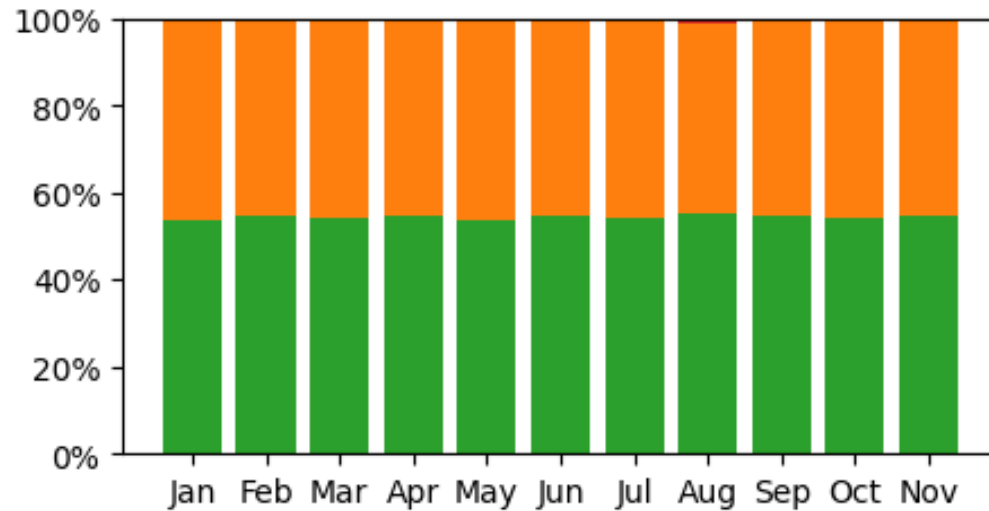


CUSTOMER SEGMENTATION

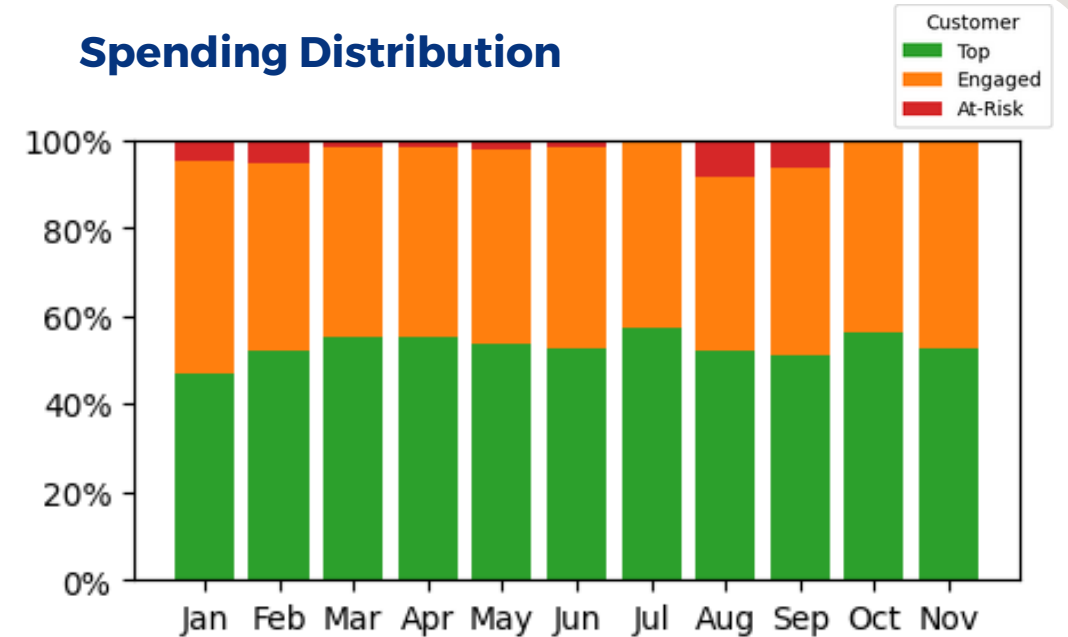
Based on AAC Credit Card Transactions as of December 7, 2021



Transaction Distribution



Spending Distribution



- **Top Customers:**
50-55% of transaction volume & expenditure
- **Engaged Customers:**
Significant contribution to volume & expenditure
- **At-Risk Customers:**
Minimal transaction activity. Only notable expenditure for 4 months

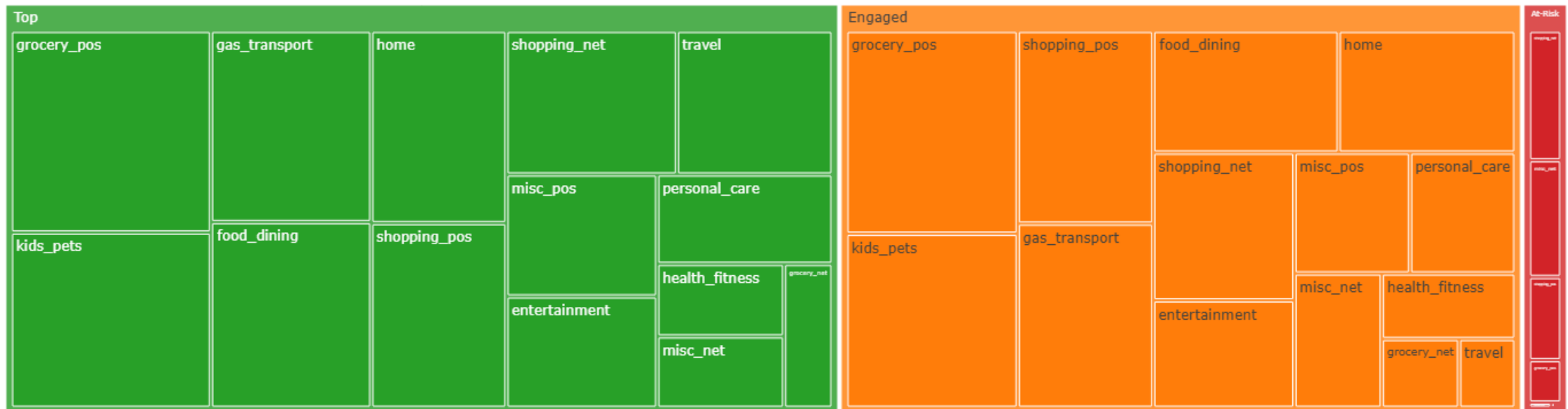


CUSTOMER SEGMENTATION

Based on AAC Credit Card Transactions as of December 7, 2021



Customer Spending per Cluster and Category



- Top and Engaged customers spend the most in **physical grocery stores** (grocery_pos).
- At-Risk customers spend the most in **online shopping** (shopping_net)

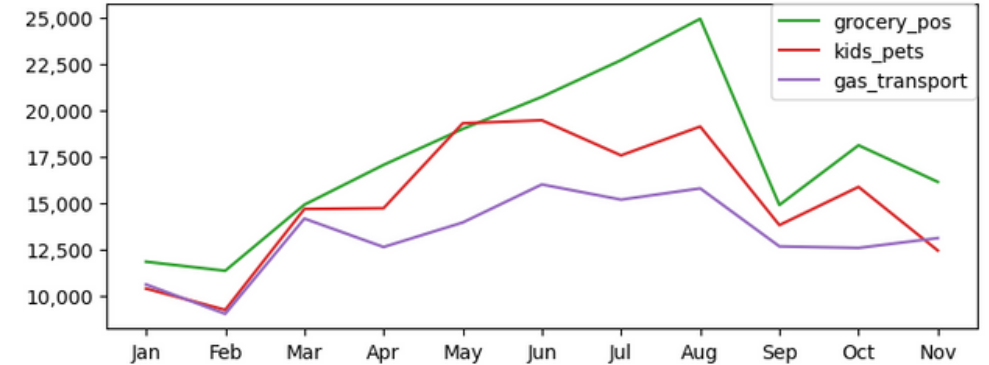
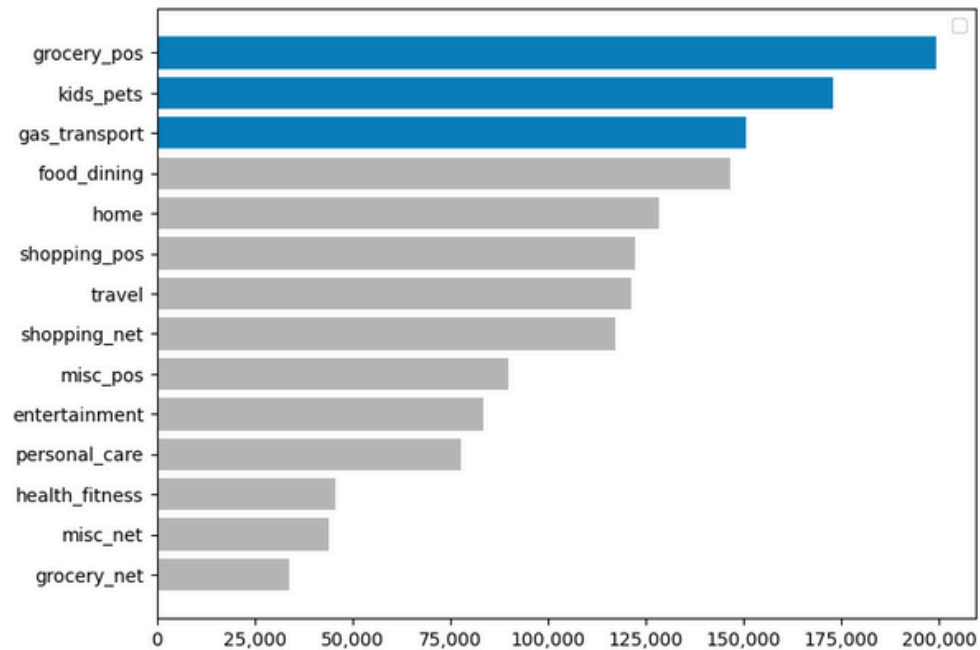


TOP CUSTOMERS

Highly active and valuable customers who purchase frequently and spend a lot



Top Customers' Top 3 Categories



- Introduce **special bonuses**, such as extra cashback or double points on key months
- Develop **loyalty programs** with exclusive offers and personalized rewards
- Explore opportunities for **cross-selling and upselling** within the top categories

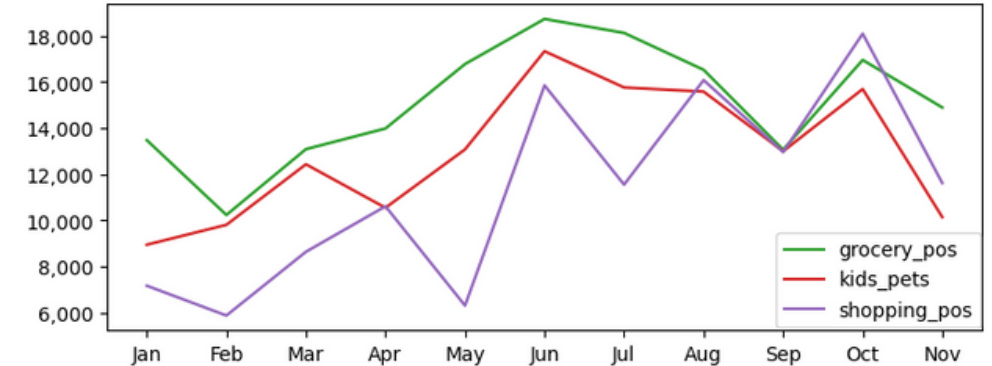
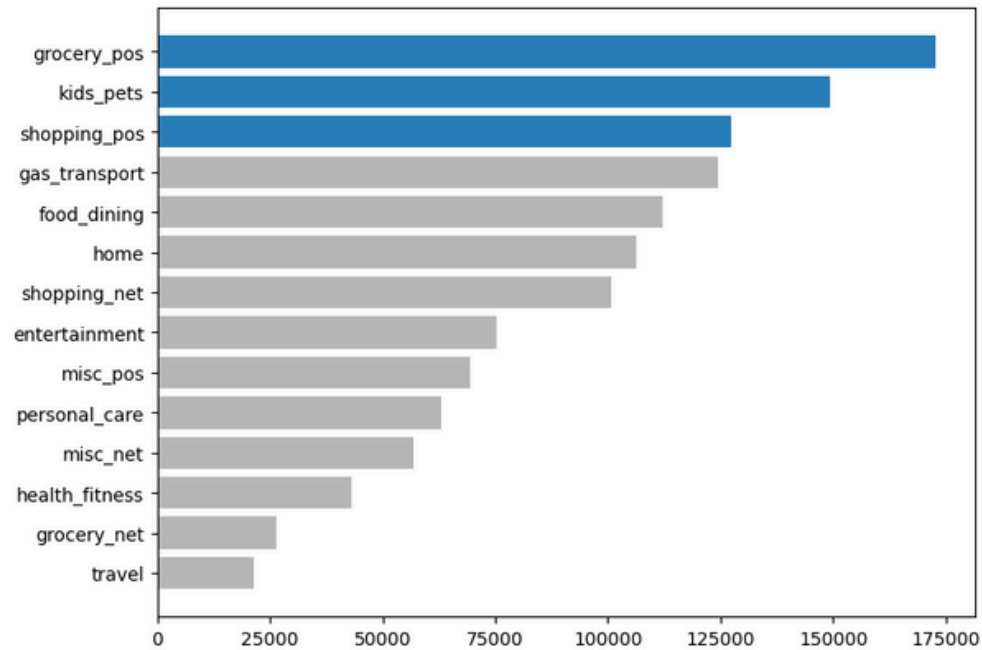


ENGAGED CUSTOMERS

Active but with a moderate level of spending compared to top customers



Engaged Customers' Top 3 Categories



- Design **targeted promotions** that encourage higher spending, such as spend-and-get offers
- Offer **seasonal incentives** during key months to boost spending.

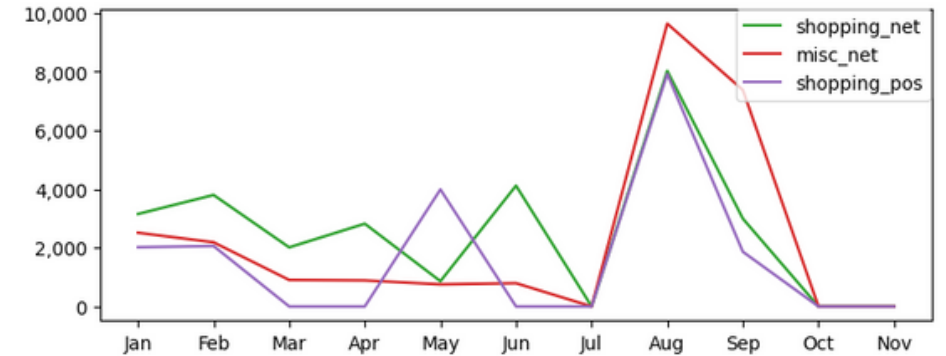
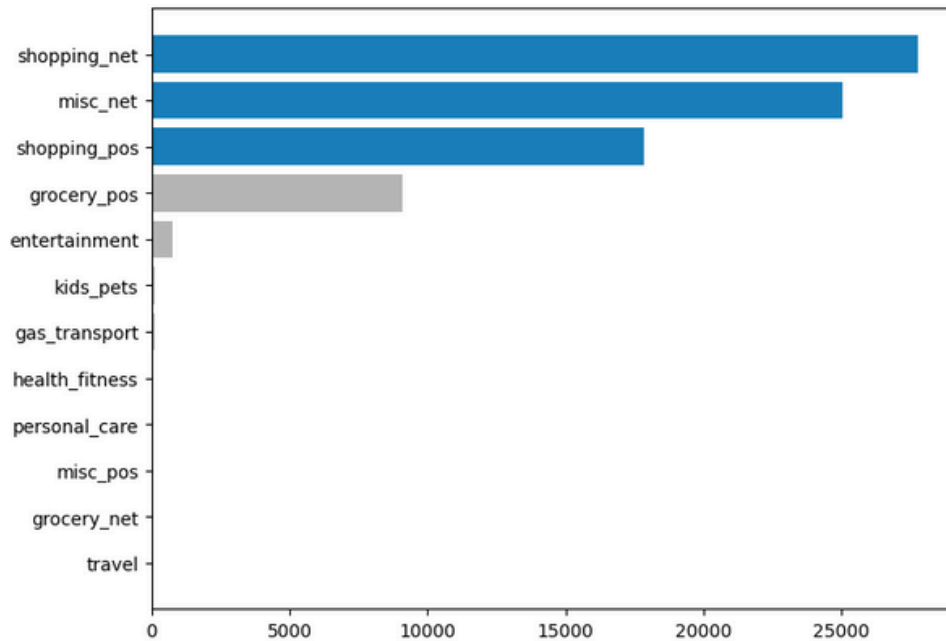


AT-RISK CUSTOMER

No recent purchases, buy infrequently, and spend less. They might be at risk of churning



At-risk Customers' Top 3 Categories



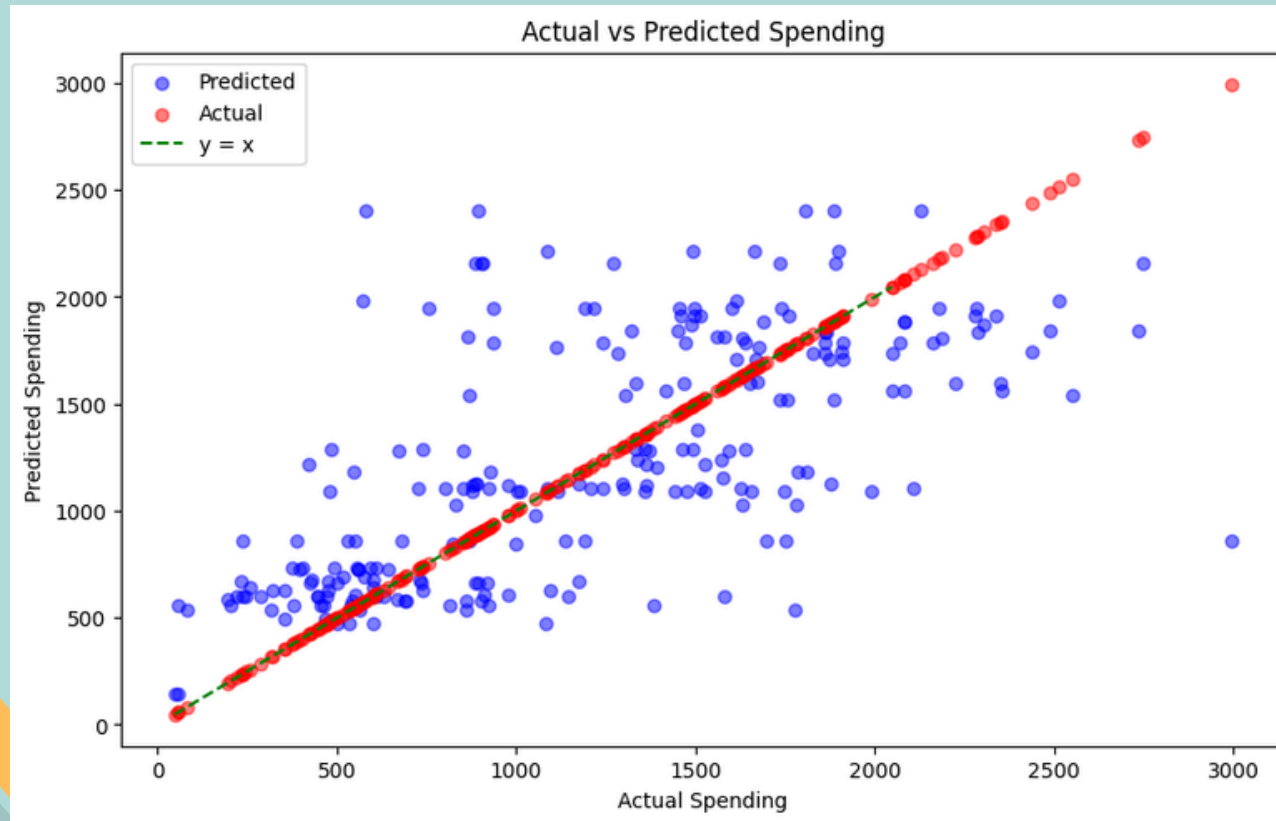
- Develop **targeted communication campaigns and strategies** to keep members informed about program updates, rewards, and special promotions.
- Create offers specifically designed to entice at-risk customers back into using their AAC credit cards. Offers include **attractive discounts and waived fees.**

Supervised Learning: Linear Regression

grocery_pos

kids_pets

gas_transport



Data date ranged: Jan 2020 to Nov 2021
Training Data: Jan 2020 to Dec 2020
Predicted Data: Jan 2021 to Nov 2021

Note: The limitation of this model is that it was trained using only one year of data.

Metrics to help evaluate the model's performance:

- **R-squared:** 0.47
- Root Mean Square Error: 513.82
- Mean Absolute Error: 382.51
- Mean Absolute Percentage Error: 41.57

STREAMLIT DEPLOYMENT

>

DECODE - ENCODE : Driving business growth through Clustering and Linear Regression Modelling

About The Challenge



Adobe Advantage Cards (AAC) - a credit card company collected

DATASQU4D

Arden Sarmiento

Austine Wong

Bea Reyes

Evian Aguilar

Japhet Pamonag



Thank you!