Agenda

1. Introduction
2. Data Exploration
3. Model Development
4. Interpretation

Introduction

Data Quality Assessment

	Customer Demographic	Customer Addresses	Transactions
Accuracy	DOB: Inaccuracy Default: Inaccuracy		
Completeness	Customer IDs: Not in Sync Last Name: Null DOB: Null Job Title: Null Job Industry Category: Null Tenure: Null	Customer IDs: Not in Sync	Customer IDs: Not in Sync Online Order: Null Brand: Null Product Line: Null Product Class: Null Product Size: Null Standard Cost: Null Product First Sold Date: Null
Consistency	Gender: Inconsistency Default: Inconsistency	State: Inconsistency	
Currency	Deceased Indicator: Filter Out		
Relevancy	Default: Exclude Field	Country: Filter Out	Order Status: Exclude Cancelled
Validity	Default: Format		Product First Sold Date: Format

Introduction

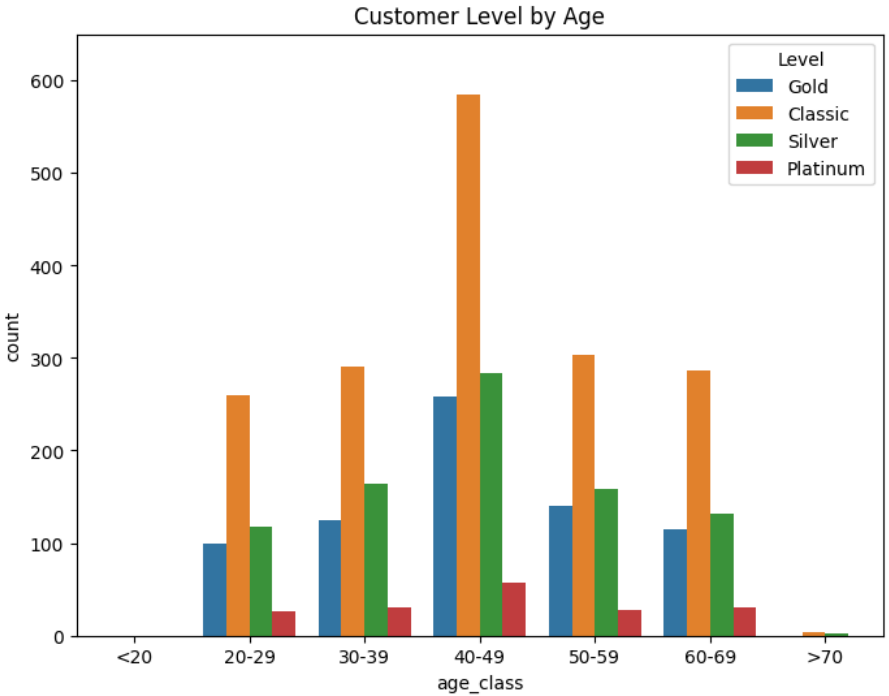
Data Cleaning

- Data containing missing values were removed from the records.
- Conflicting records were also dropped by considering the join keys between tables.
- Age, Last Purchases (Days Ago) and Profit features were added.
- Records related to the deceased indicator were excluded.

Distinct Customer IDs (Before)	4,000
Distinct Customer IDs (After)	3,491

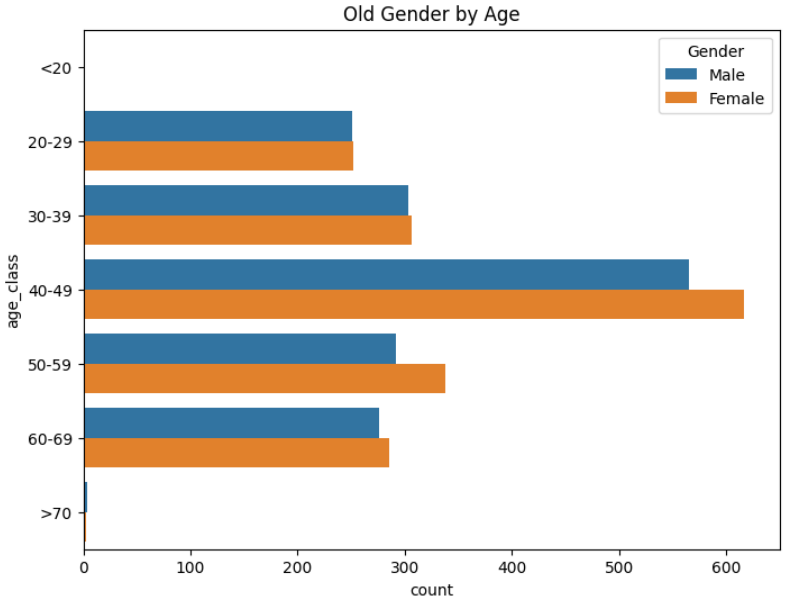
Data Exploration

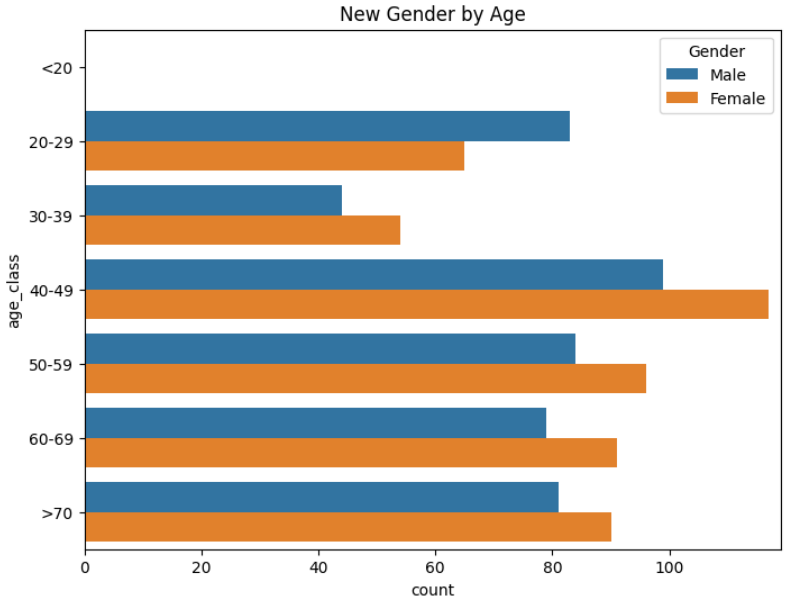
Customer Level Old Customer by Age



Data Exploration

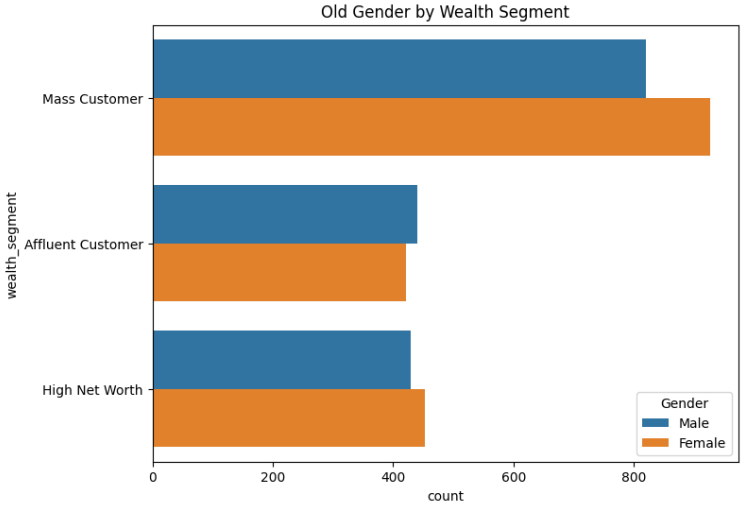
Gender by Age

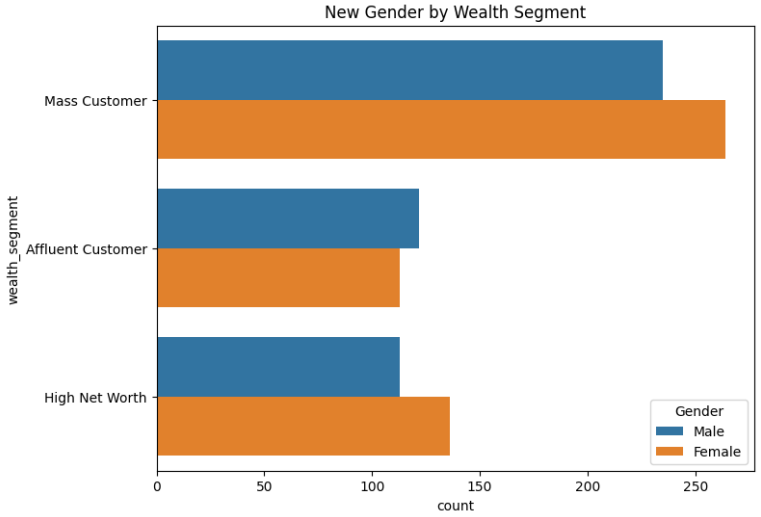




Data Exploration

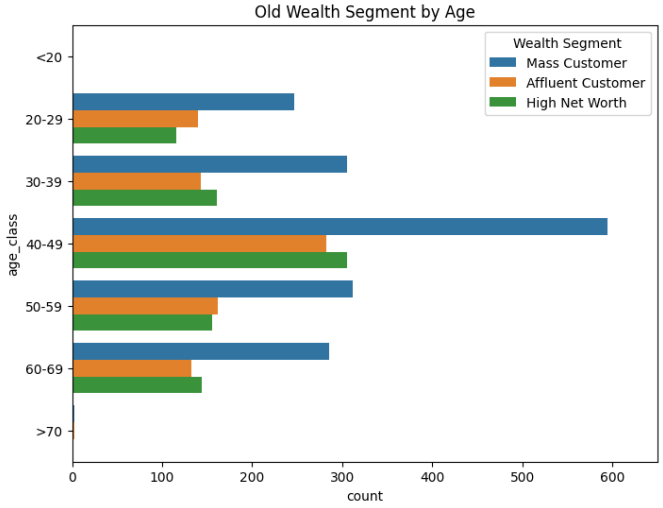
Gender by Wealth Segment

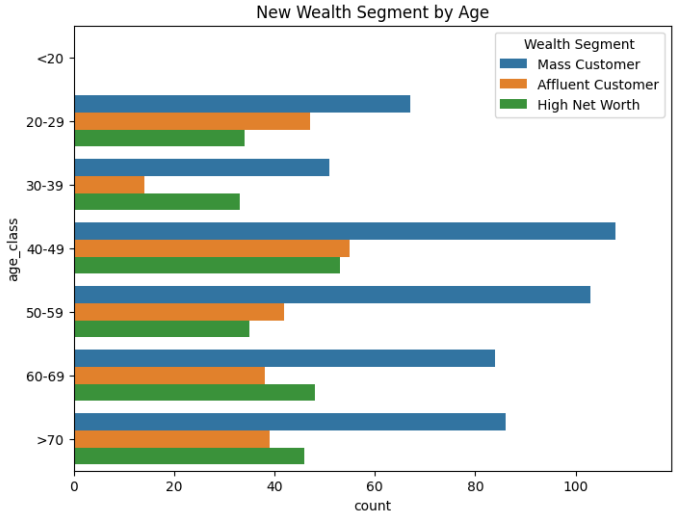




Data Exploration

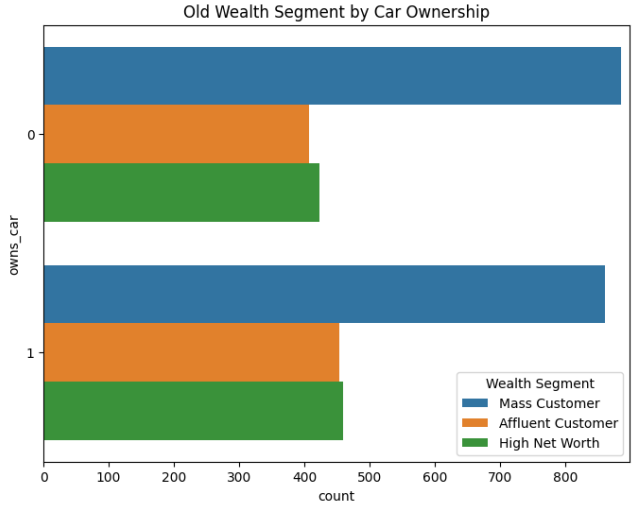
Wealth Segment by Age

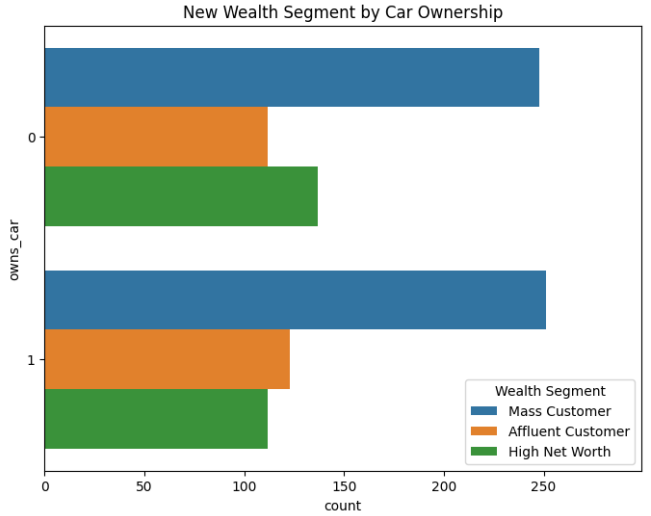




Data Exploration

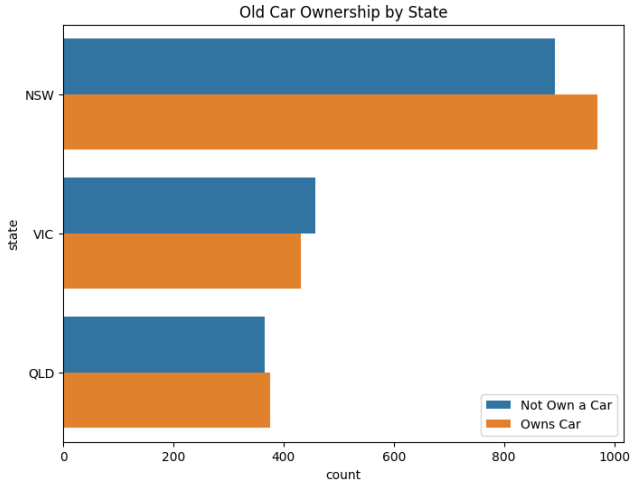
Wealth Segment by Car Ownership

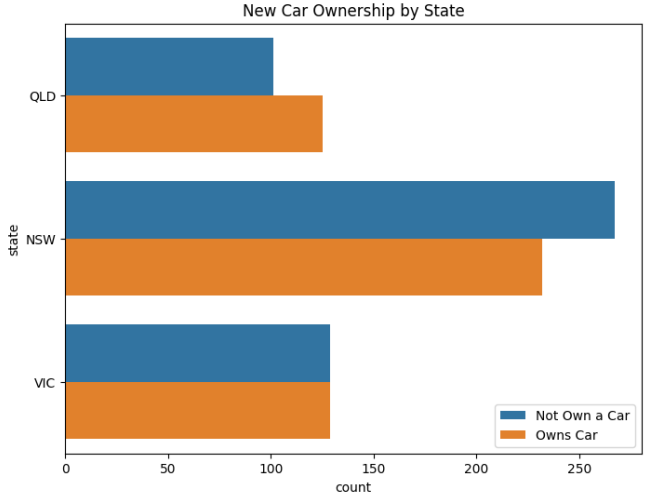




Data Exploration

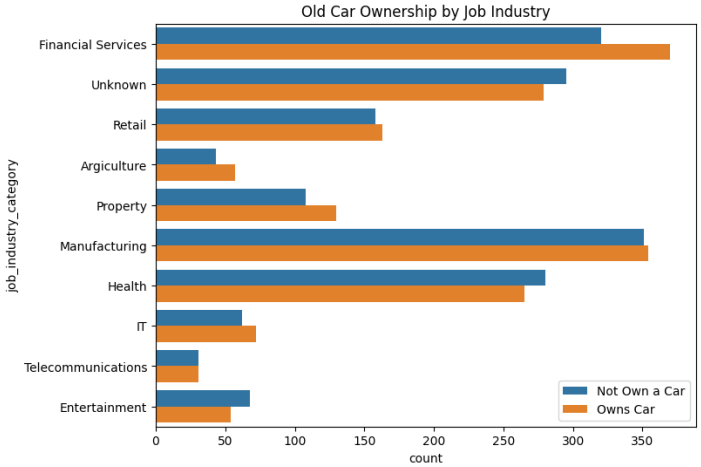
Car Ownership by State

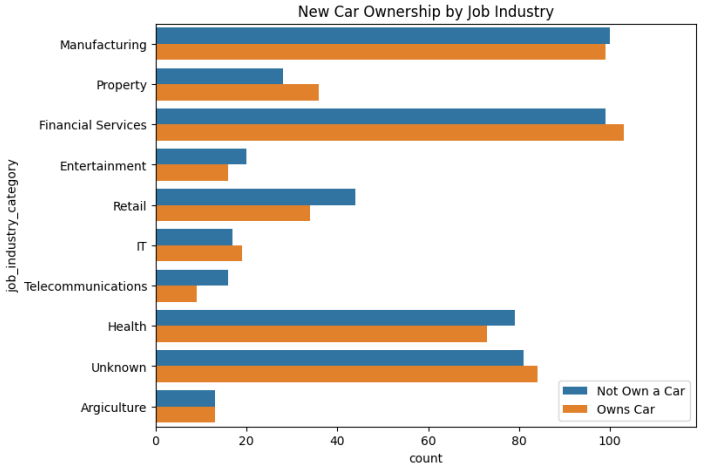




Data Exploration

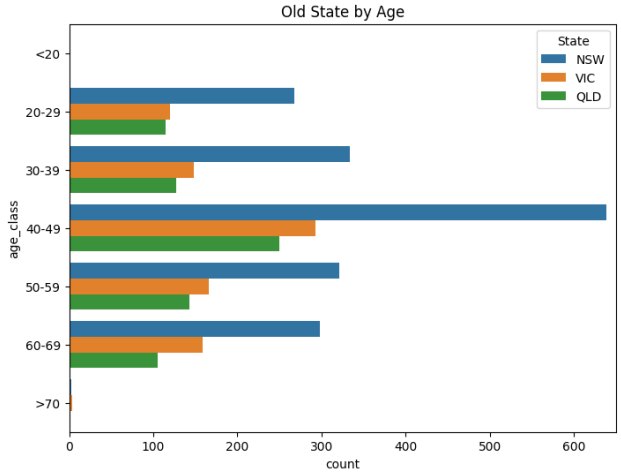
Car Ownership by Job Industry

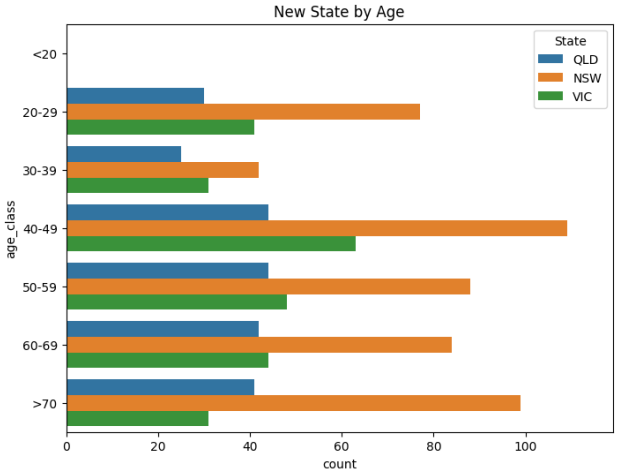




Data Exploration

State by Age





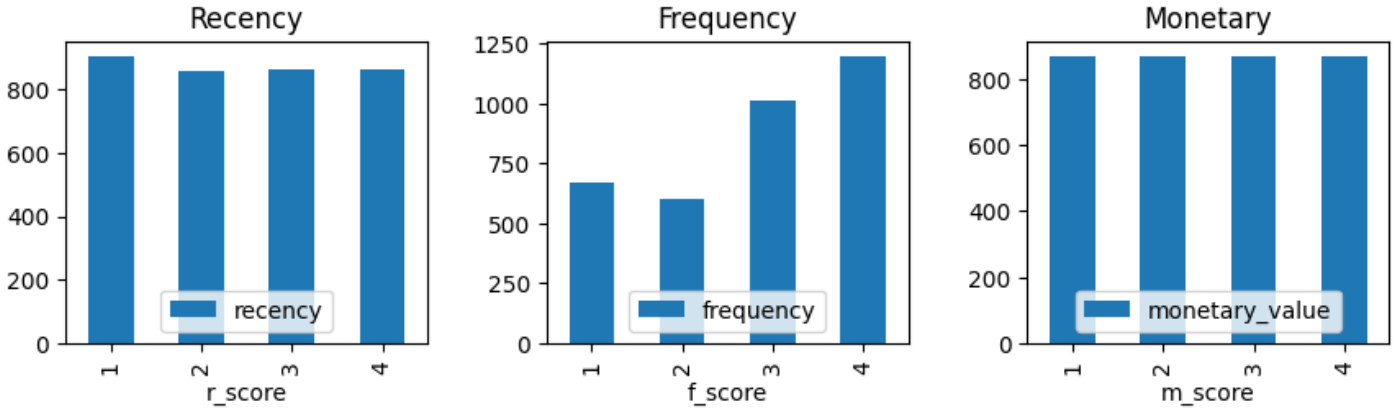
Model Development

RFM Analysis

- Recency
  - Based on the most recent transaction date of a customer, where they were divided into 4 quartiles and given a *R_Score*.
- Frequency
  - Based on the number of transactions made by a customer, where they were divided into 4 quartiles and given a *F_Score*.
- Monetary Value
  - Based on the average profit generated by each customer, where they were divided into 4 quartiles and given a *M_Score*.

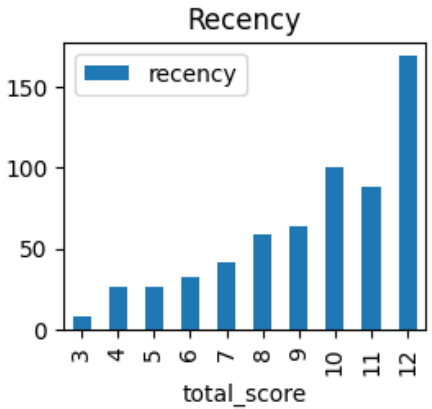
Model Development

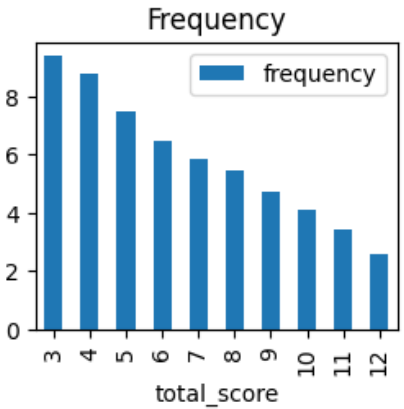
RFM Analysis

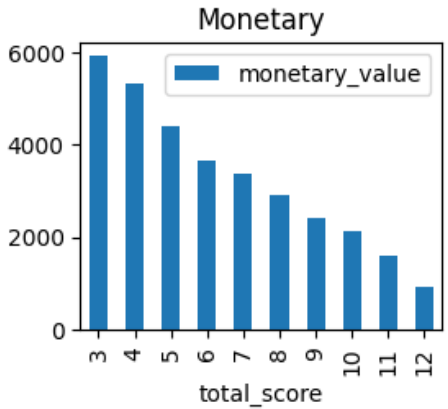


Model Development

RFM Analysis

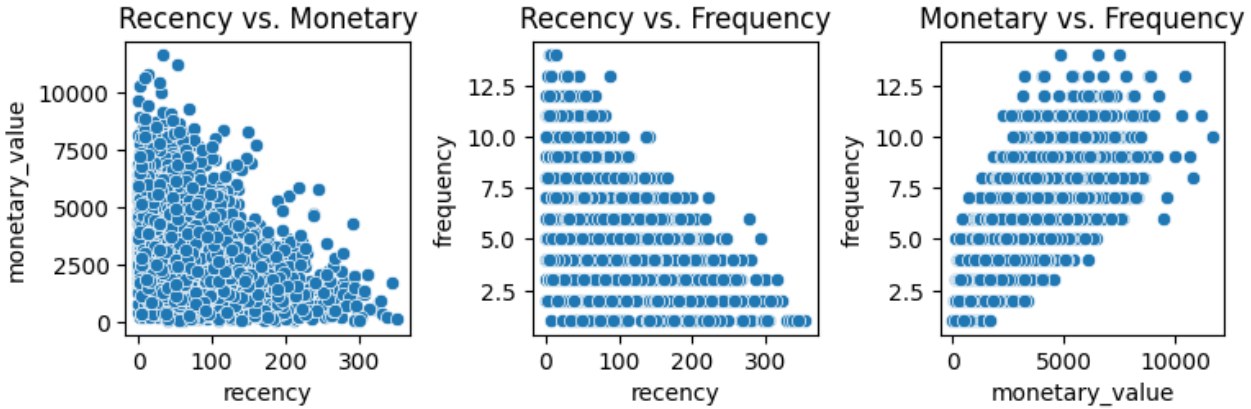






Model Development

RFM Analysis

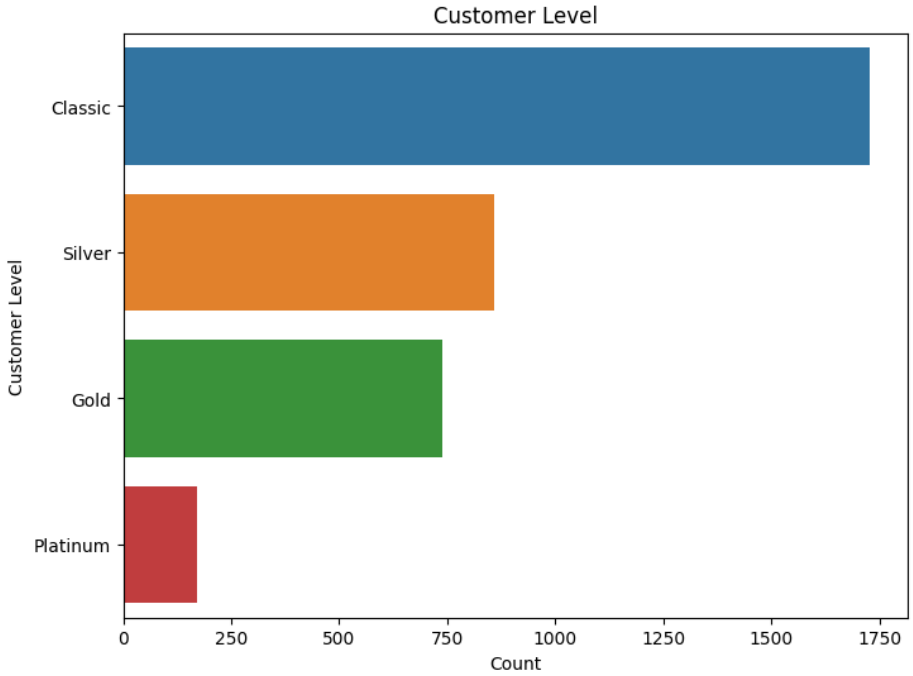


Model Development

RFM Analysis

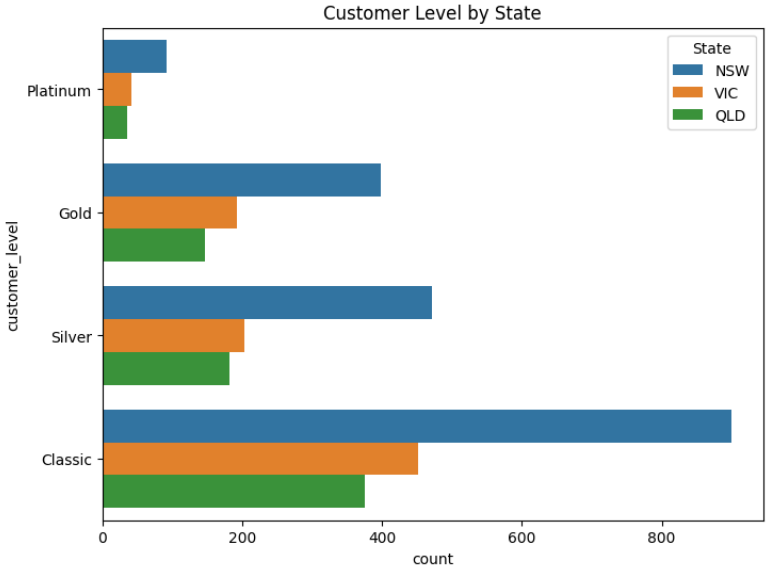
4 customer tiers were identified based on the RFM Segmentation:

1. Platinum Class: Who have made recent purchases, are frequent, and contribute the most profits.
2. Gold Class
3. Silver Class
4. Classic Class: Who have not made recent purchases, are not frequent, and do not significantly contribute to profits.



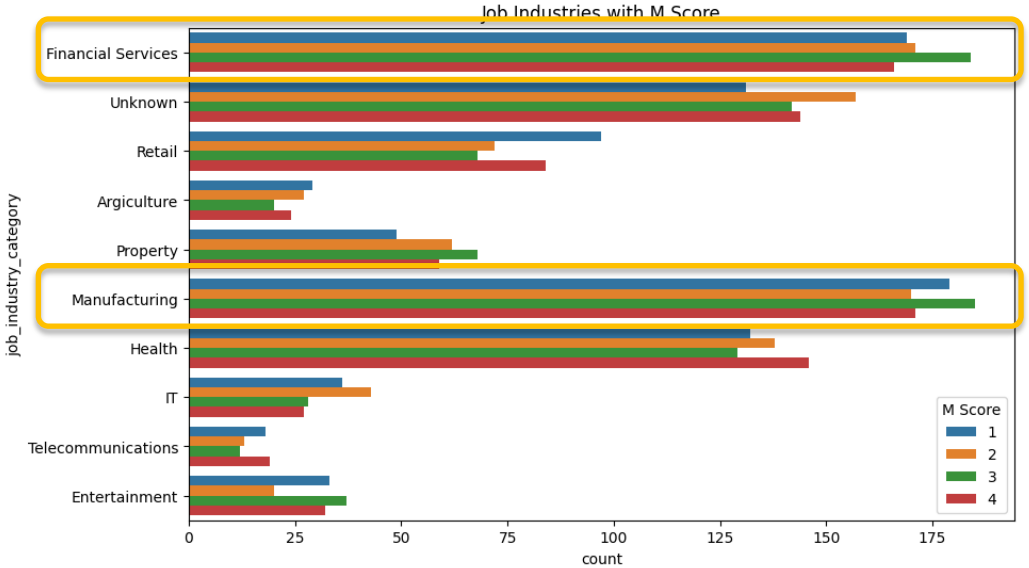
Interpretation

Customer Tiers by State



Interpretation

Profitable Job Industries



Interpretation

Customer Segments

Segment	RFM Score
Champions	3
Loyalists	4
Potential Loyalists	5
New Customers	6
Promising	7
Need Attention	8
About to Sleep	9
High Risk	10
Hibernating	11
Lost Customers	12

Interpretation

Targeting Methodology

- Customers with **high RFM scores** can be **selected and focused on**.
- These customers have made purchases recently, frequently, and generate significant profits.

Thank You