Sprocket Central Pty Ltd

Data Analytics Approach

Sudipto Ghosh
Data Analytics Consulting Virtual Intern

Agenda

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

Introduction

Data Quality Assessment

	Customer Demographic	Customer Addresses	Transaction Data
Accuracy	DOB: Inaccuracy Job Industry Category: Misspelling		
Completeness	DOB: Blanks Job Title: Blanks Job Industry Category: Blanks Tenure: Blanks Customer IDs: Not in Sync	Customer IDs: Not in Sync	Standard Cost: Blanks Brand: Blanks Product Line: Blanks Product Class: Blanks Product Size: Blanks Product First Sold Date: Blanks Customer IDs: Not in Sync
Consistency	Gender: Inconsistency	States: Inconsistency	
Currency	Deceased Customers: Filter Out		
Relevancy	Default: Exclude Field		Order Status: Exclude Cancelled
Validity			Product First Sold Date: Format

Introduction

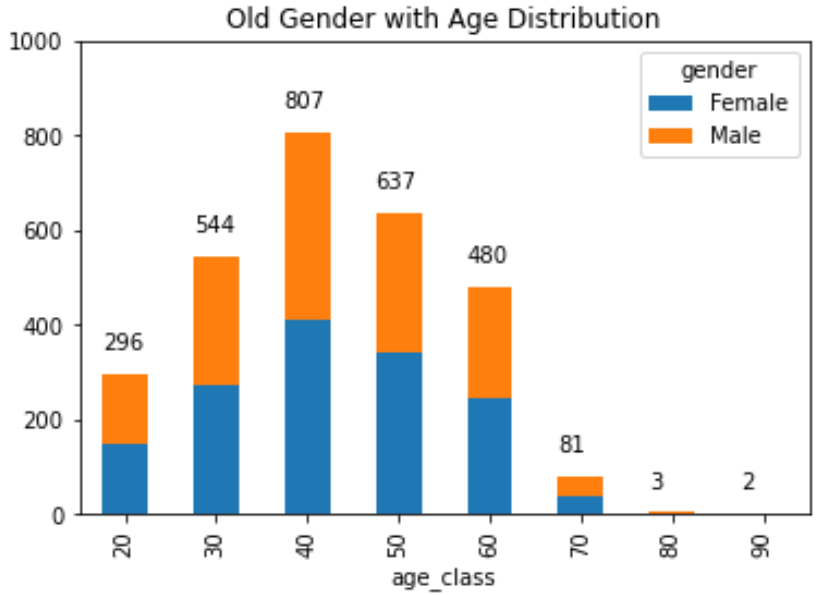
Data Cleaning

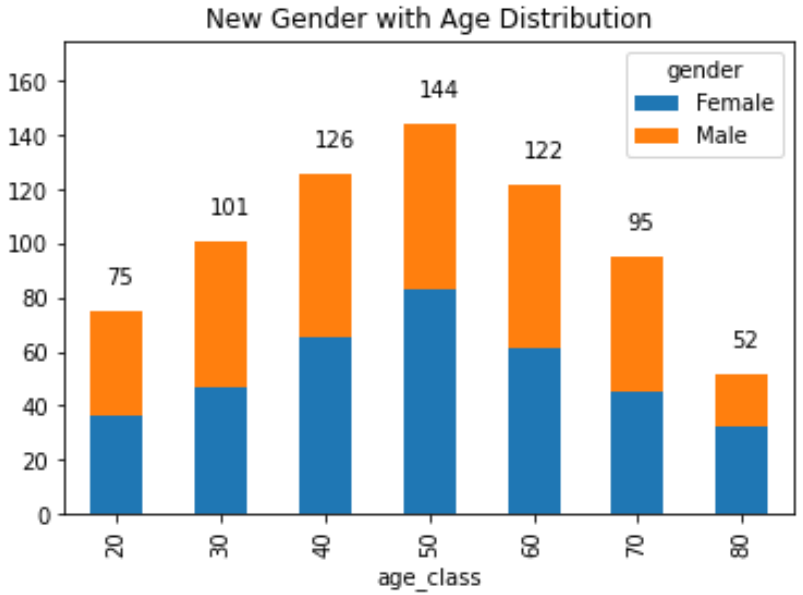
- Records with Missing Fields were Dropped.
- Join Keys between Tables were considered and conflicting Records were Dropped.
- Age, Last Purchase (Days Ago) and Profit Fields were Added.
- Records pertaining to Deceased Customers were Dropped.
- Transactions more than a year old were Dropped.

Distinct Customer IDs As Received	4,000
Distinct Customer IDs after Data Cleaning	3,492

Data Exploration

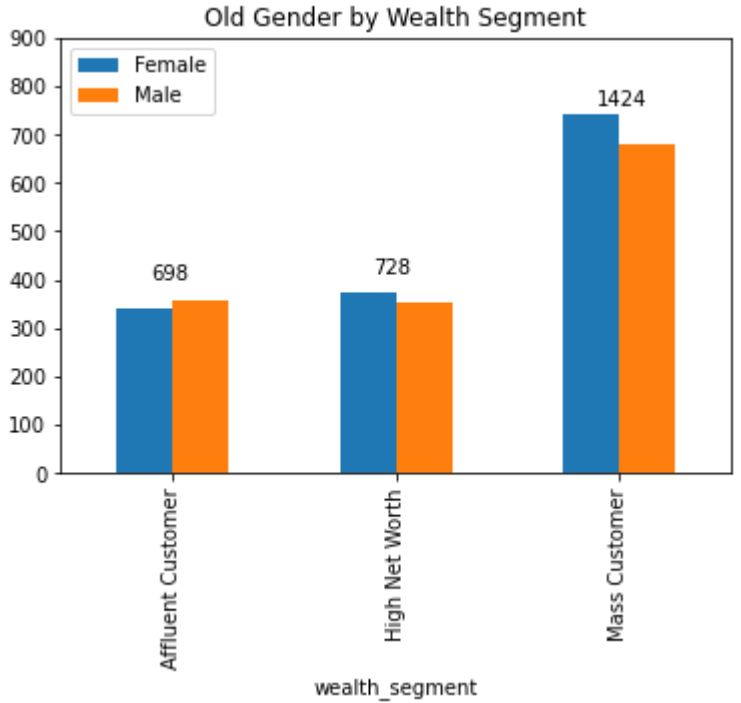
Gender with Age Distribution

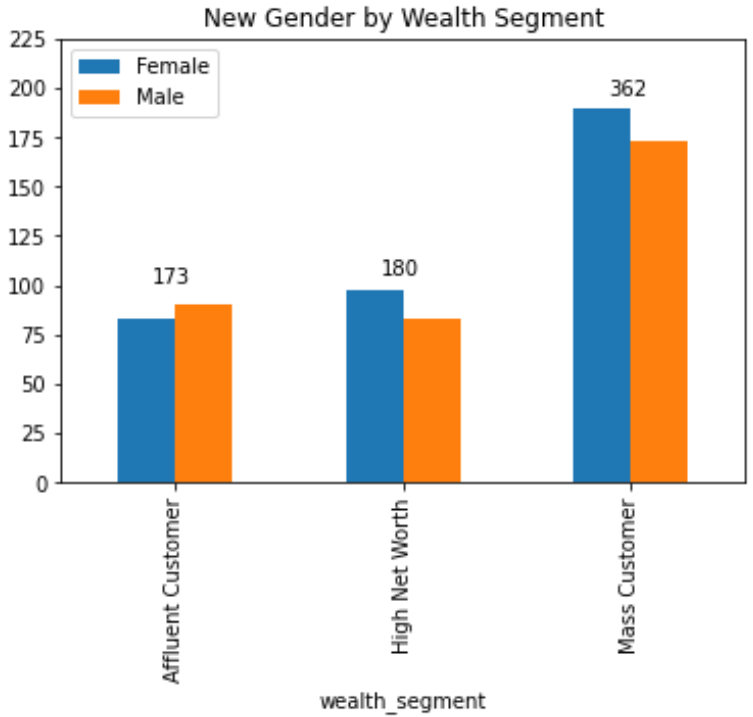




Data Exploration

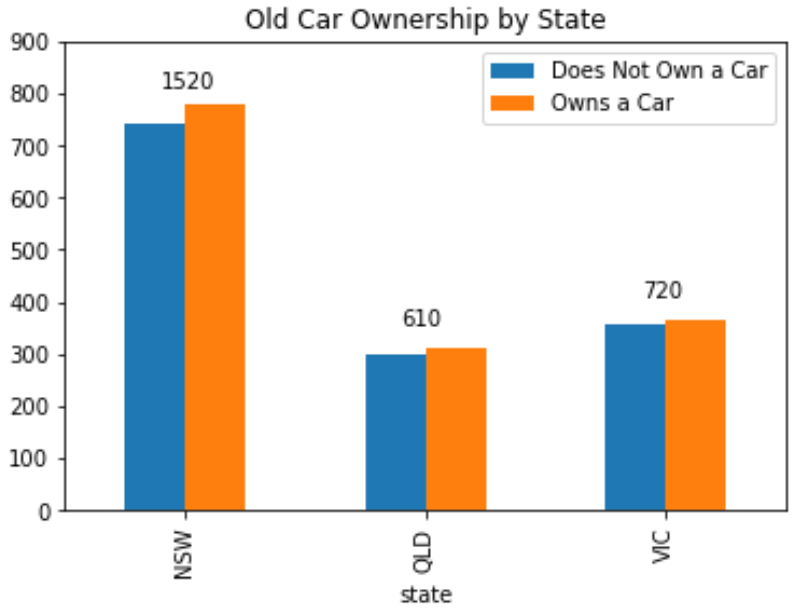
Gender with Wealth Segment

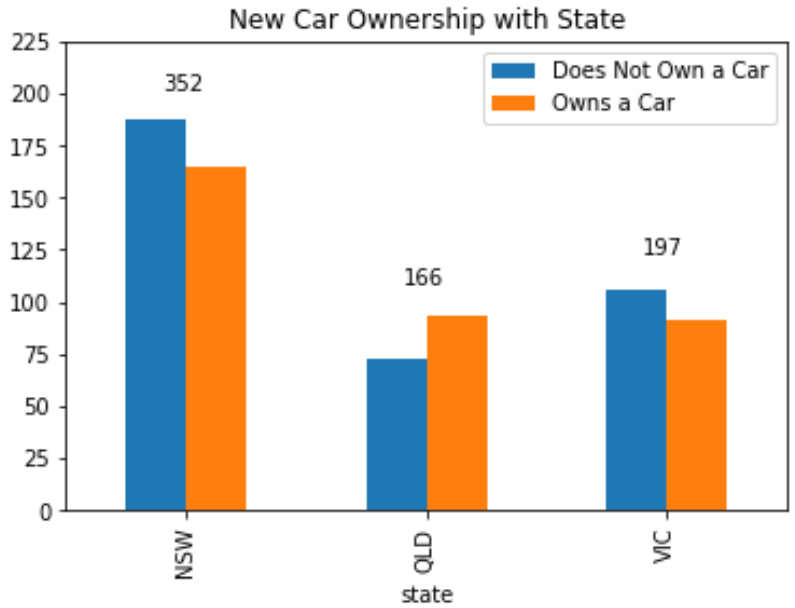




Data Exploration

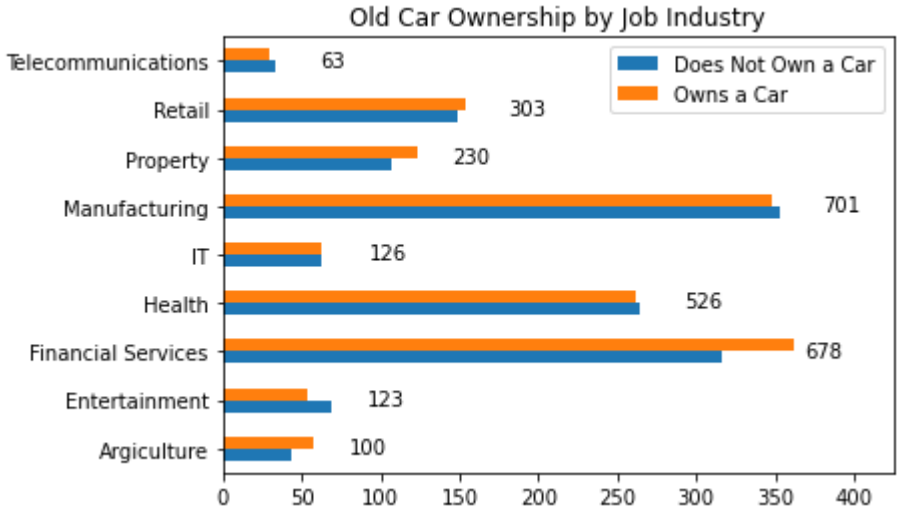
Car Ownership with State

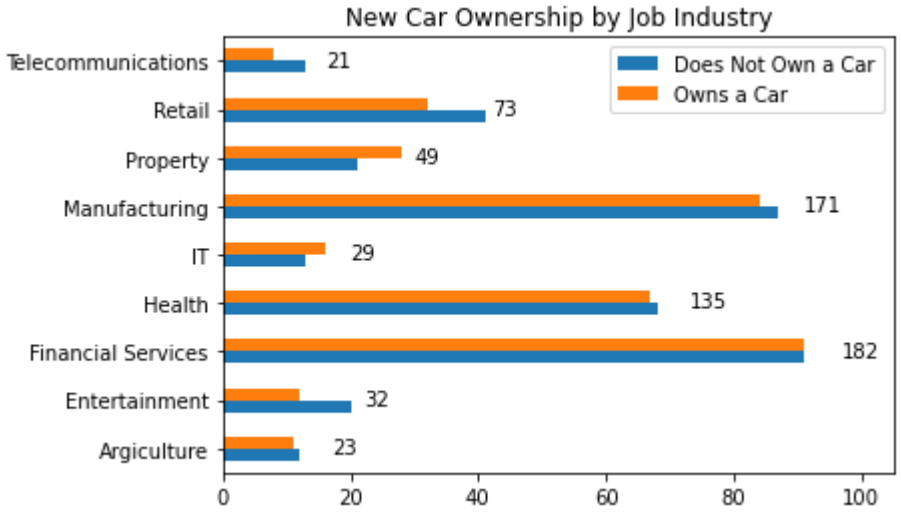




Data Exploration

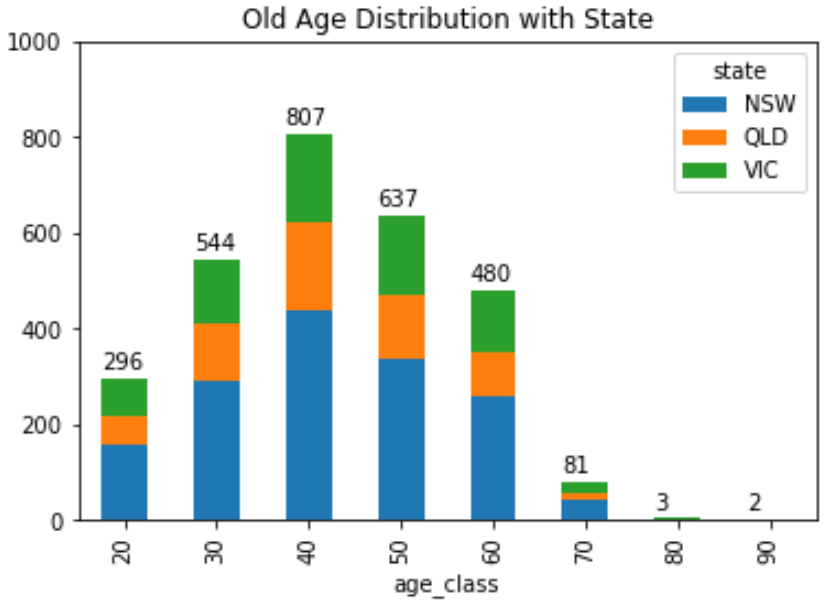
Car Ownership with Job Industry

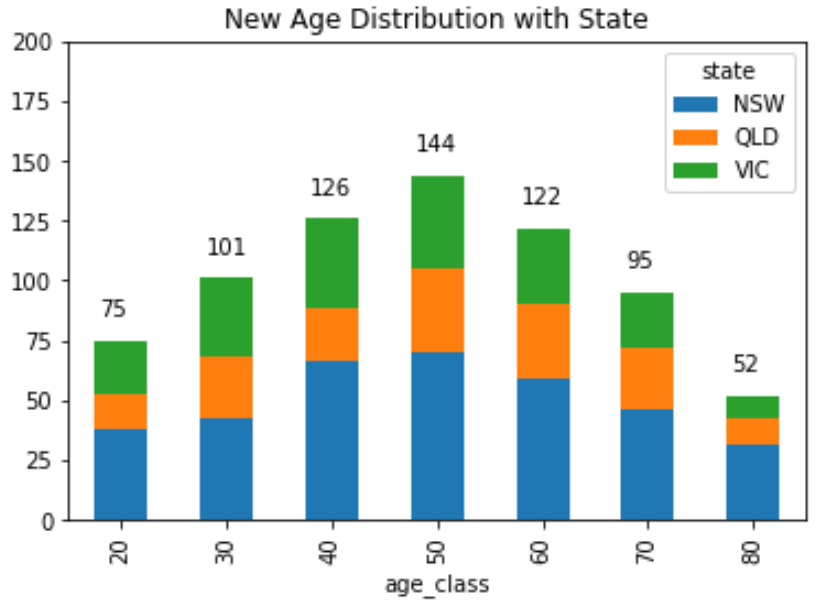




Data Exploration

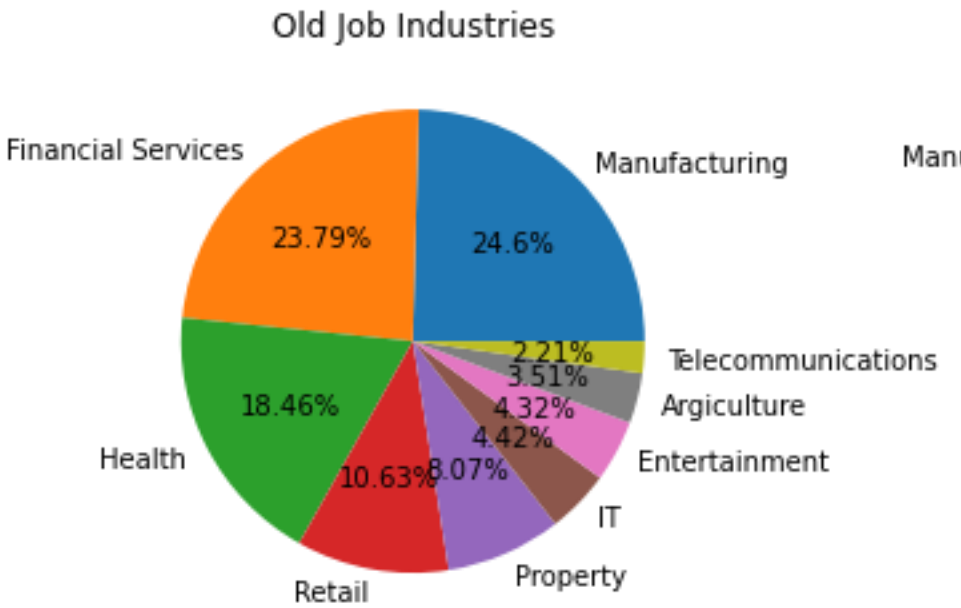
Age Distribution with State

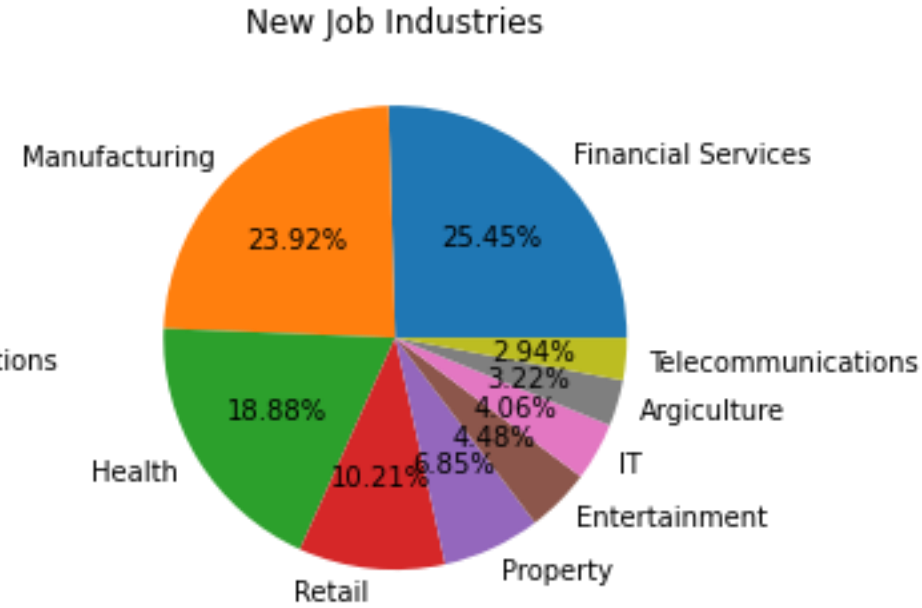




Data Exploration

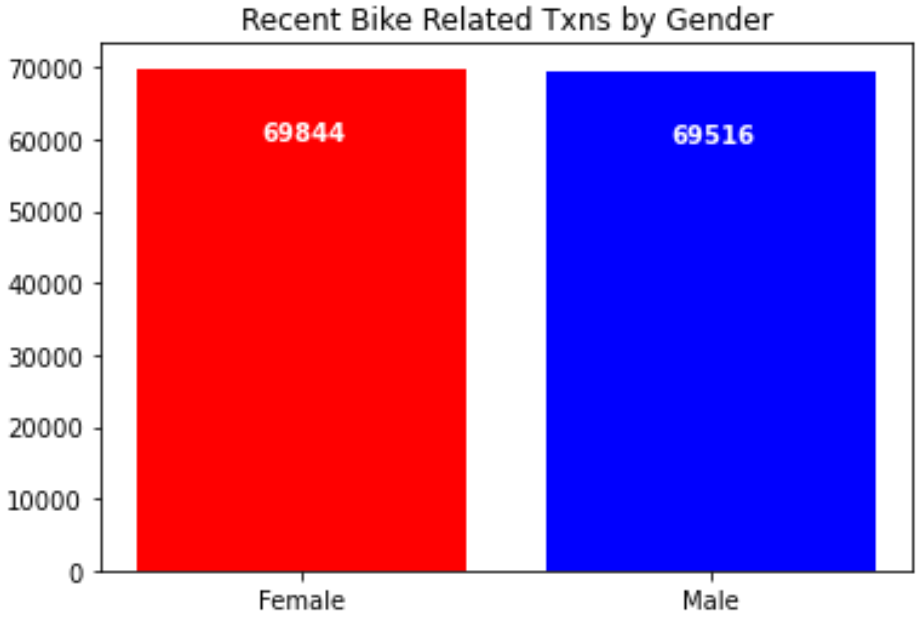
Job Industry Diversity





Data Exploration

Recent Transactions



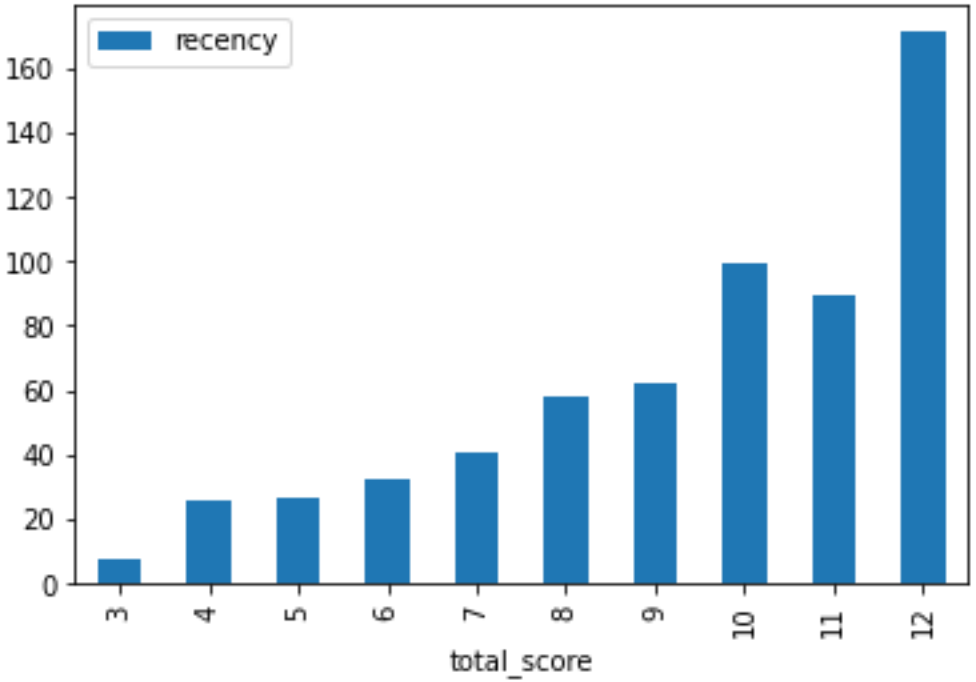
Model Development

RFM Analysis

- **Recency**
  - The last day on which a customer performed a transaction was taken as the recency parameter.
  - Customers were divided into 4 quartiles and given a R_Score.
- **Frequency**
  - The frequency of transactions done by a particular customer was taken as the frequency parameter.
  - Customers were divided into 4 quartiles and given a F_Score.
- **Monetary Value**
  - The average profit per customer was taken as the monetary value parameter.
  - Customers were divided into 4 quartiles and given a M_Score.

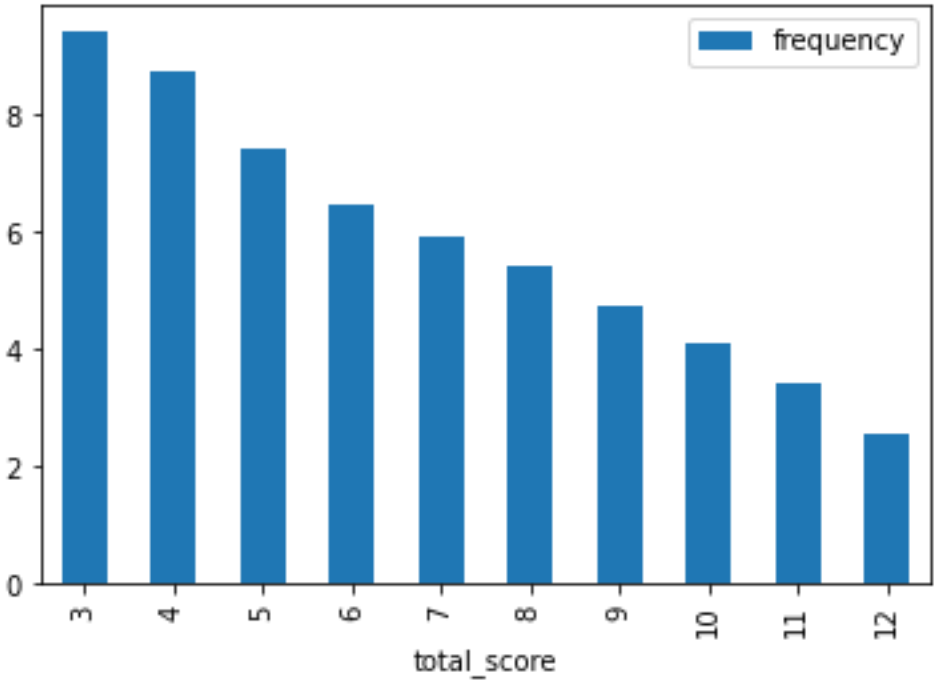
Model Development

RFM Analysis



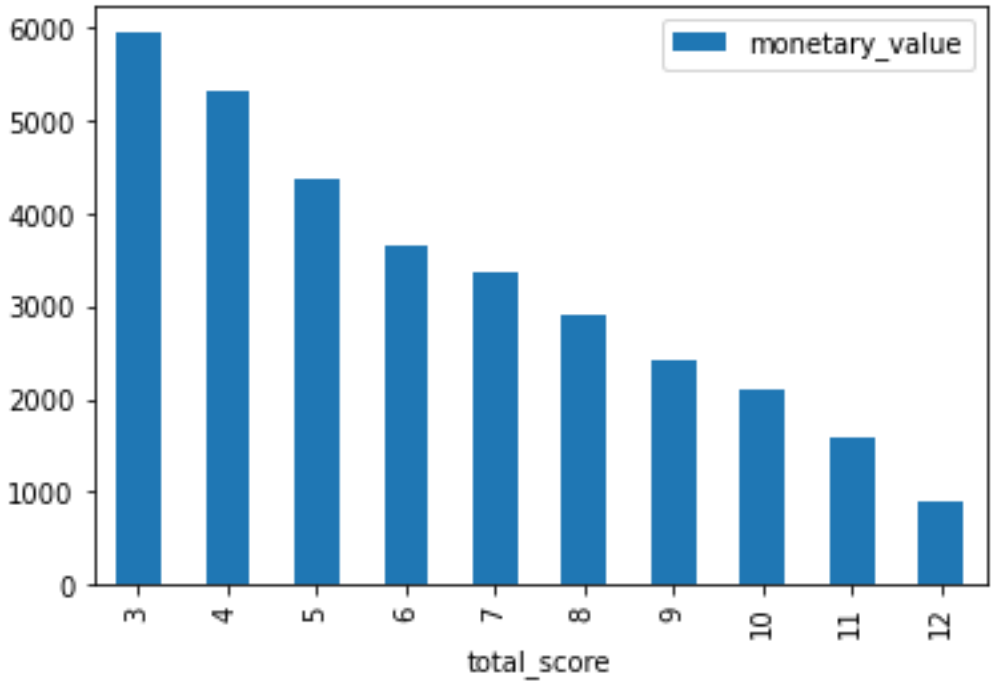
Model Development

RFM Analysis



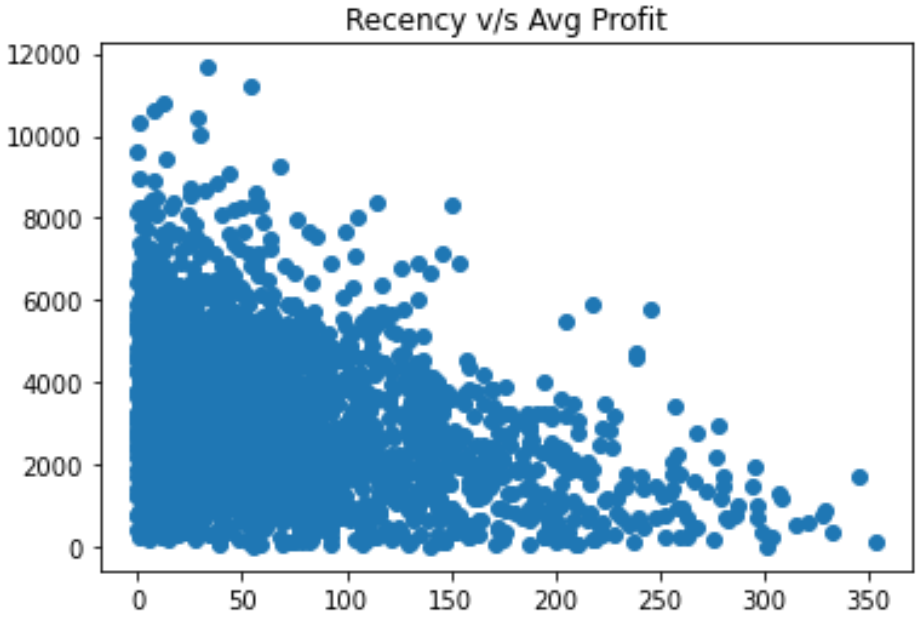
Model Development

RFM Analysis



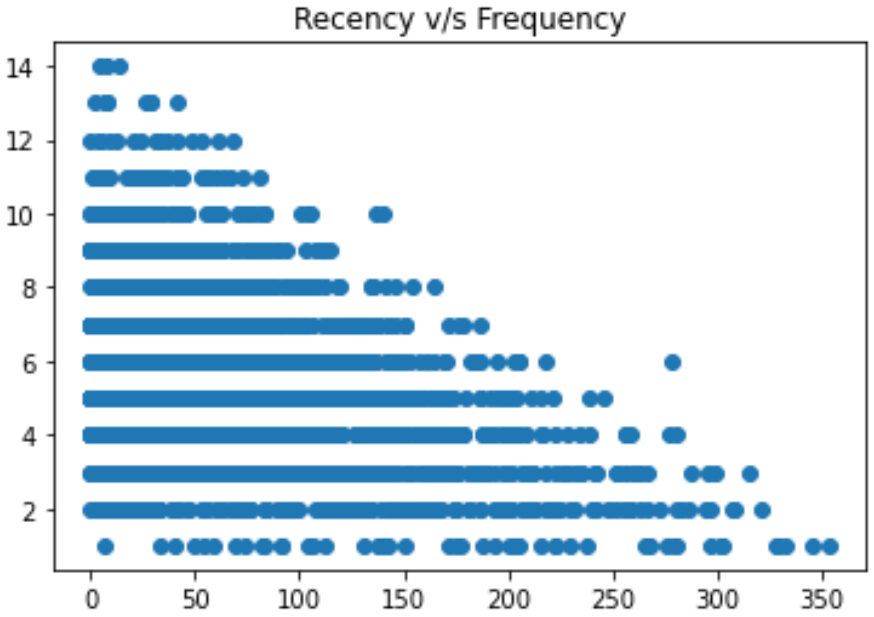
Model Development

RFM Analysis



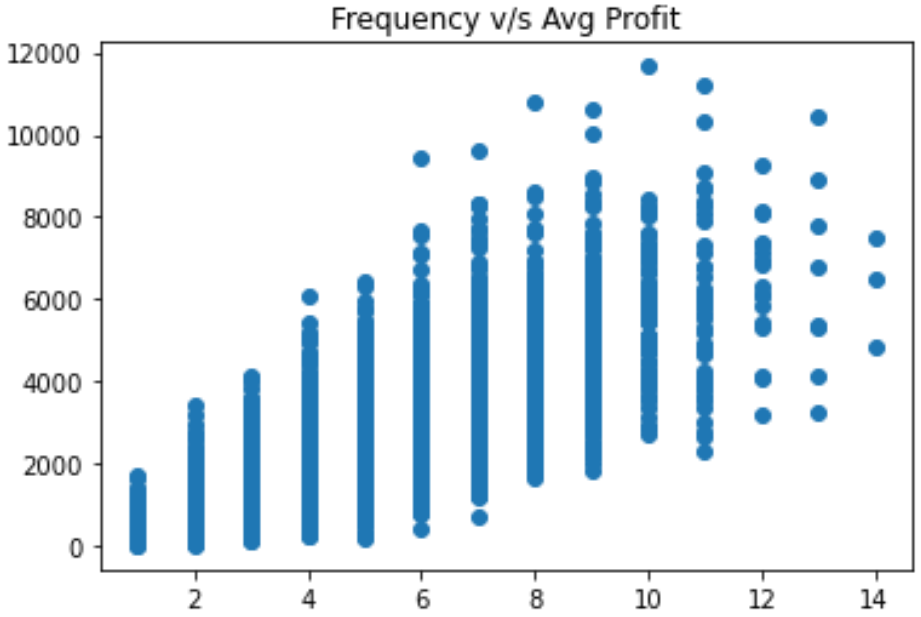
Model Development

RFM Analysis



Model Development

RFM Analysis

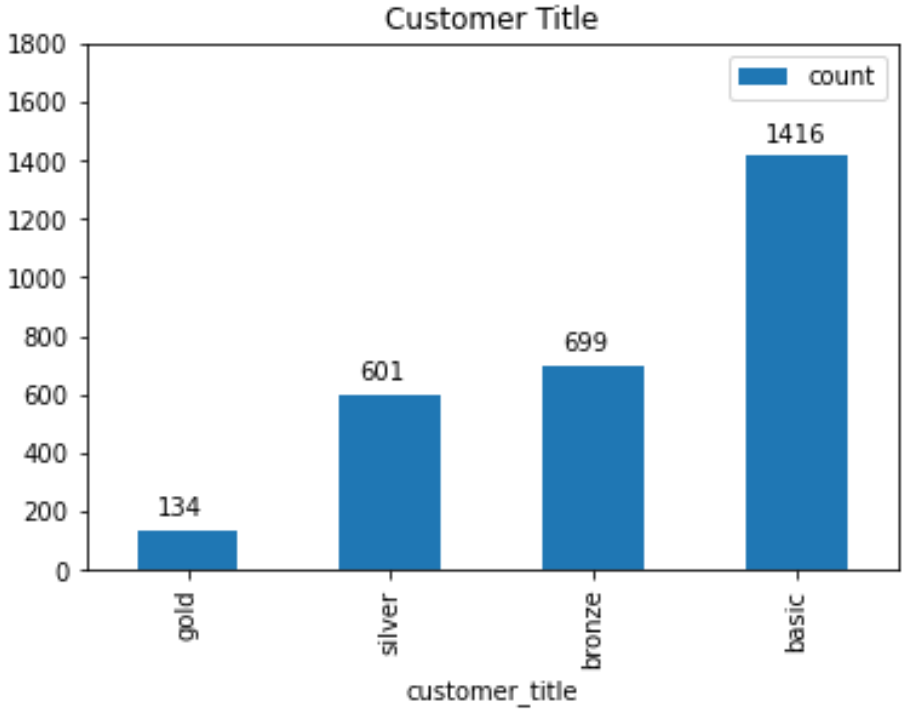


Model Development

RFM Analysis

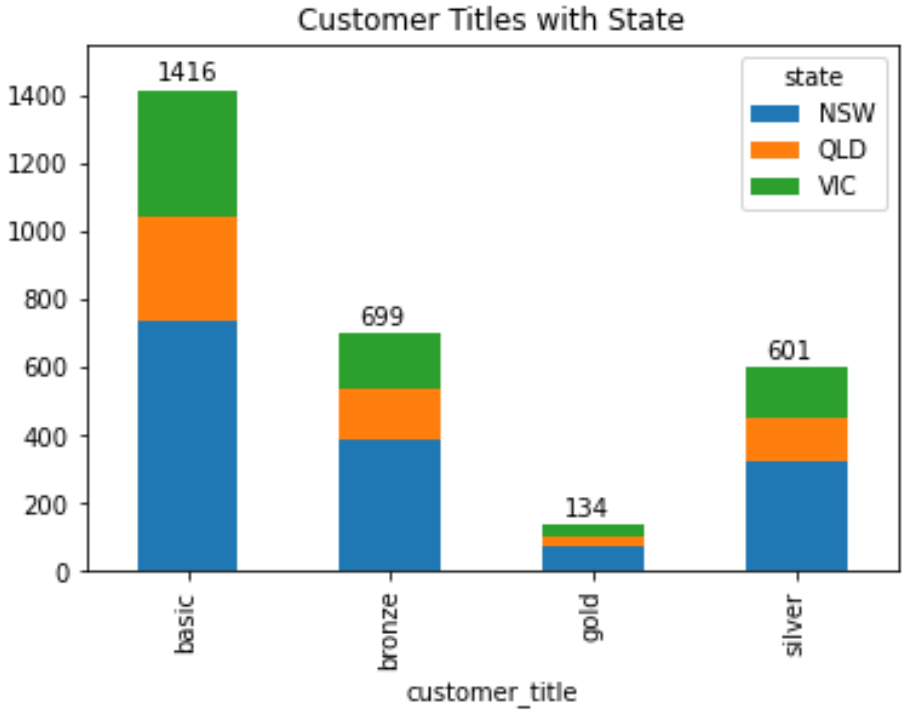
Based on the RFM Class, four customer tiers were identified:

1. Gold Class: These customers have recently made a purchase, are frequent and are most profitable.
2. Silver Class
3. Bronze Class
4. Basic Class: These customers have not made any recent purchase, are not frequent and do not contribute majorly.



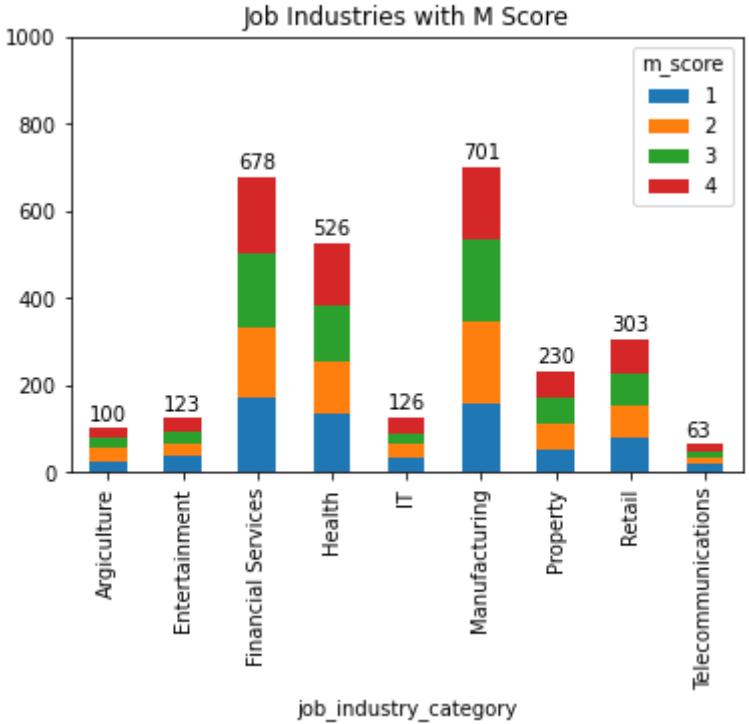
Interpretation

Customer Tiers by State



Interpretation

Profitable Job Industries



Interpretation

Customer Segments

Segment	RFM Score	Cumulative
Platinum	3	134
Very Loyal	4	296
Becoming Loyal	5	558
Recent	6	891
Average	7	1300
High Risk	8	1712
Evasive	9	2110
Losing	10	2417
Inactive	11	2595
Lost	12	2850

Interpretation

Targeting Methodology

- Customers having high RFM Scores can be filtered and targeted.
- The customers have made recent purchases, are frequent, and drive the most profits.

Appendix

Appendix

S4. SELECTED DISTRIBUTION INDICATORS, Household net worth and gross household income—2011–12

		Household net worth(a)	*Gross household income*(a)
Ratio of values at top of selected percentiles			
P90/P10	ratio	53.87	9.06
P80/P20	ratio	11.61	4.45
P80/P50	ratio	2.36	1.90
P20/P50	ratio	0.20	0.43
Percentage share received by households in			
Lowest net worth quintile	%	0.9	12.1
Middle net worth quintile	%	12.0	17.4
Highest net worth quintile	%	60.8	31.3
Percentage share received by households in			
Lowest gross income quintile	%	12.4	4.3
Middle gross income quintile	%	16.4	15.7
Highest gross income quintile	%	36.8	46.5

Source: Australian Bureau of Statistics