

TheAnalyticsTeam

Sprocket Central Pty Ltd

Data analytics approach

Vishal J Khot

Agenda

1. Introduction
2. Data Exploration
3. Dashboard
4. Model Development
5. Data Analysis

Content

Introduction:

- Outline of Problem

Data Exploration:

- Data quality assessment.
- Distribution of old and new customers based on:
 - 1 Age
 - 2 Gender
 - 3 jobs
 - 4 wealth
 - 5 state

Dashboard:

- Existing Customer Dashboard
- New Customer Dashboard
- Target Customer Dashboard

Model Development:

- RFM Analysis and Customer Segmentation.
- Scatter plot analysis based on RFM model.
- Customer Title Definition based on RFM assigned values.
- Customer Distribution.

Data Analysis:

- Customer segregation based on RFM condition.
- Top customers to target.

Introduction

Identify and Recommend TOP 1000 Potential Customers From Existing Datasets for Client.

Outline of Problem:

- Sprocket Central Pty Ltd (Client) specializes in high-quality bikes and accessible cycling accessories for riders.
- Their marketing team is looking to boost business by analyzing their existing customer dataset to determine customer trends and behavior.
- Using the 3 datasets, we need to recommend the TOP 1000 customers that shall target by Sprocket Central Pty Ltd to drive the most value for the organization.

Identify and Recommend High Value Customers From New Dataset for Client.

Outline of Problem:

- New list of 1000 potential customers do not have any prior transaction history with the organization.
- Their marketing team required analyzed data that could help optimize resource allocation for targeted marketing.
- Using this dataset, we need to recommend high-value customers to optimize resource allocation for targeted marketing.

Data Exploration

Data Quality Assessment :

Data extend to quality assessment. Several data quality issues were discovered which are summarized below:

	Accuracy	Completeness	Consistency	Currency
CustomerDemograhy	DOB: Inaccurate Age: Missing	JobTitle: Blank Customerid: incomplete	Gender: Data Inconsistency	Deceased Customer: filter out
CustomerAddress		Customerid: incomplete	State: Data Inconsistency	
Transactions	Profit: Missing	Customerid: incomplete Orderid:NULL records Productid:NULL records	Customerid,Productid,Transactionid: Data Type inconsistency	
NewCustomerList:	Age, Age category: Missing	Last_name, Job_title:NULL records	All columns has inconsistence datatype	

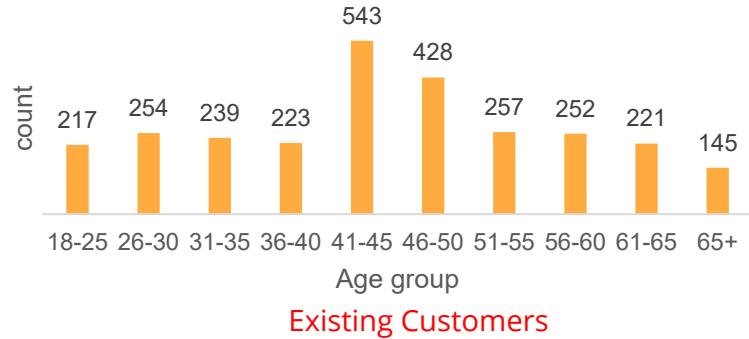
An In-depth analysis sent via Email.

Data Exploration

Age-wise Distribution

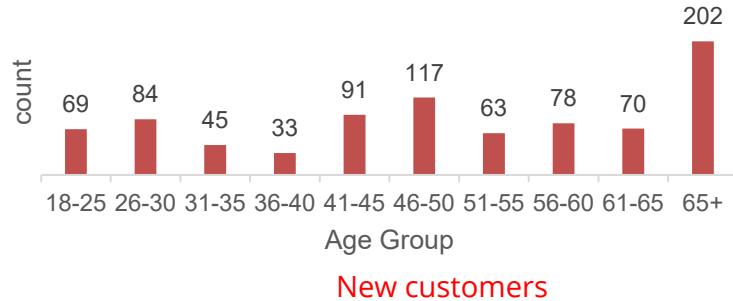
Existing Customers:

- Most customers belong to age between 41-50 years.
Followed by the 51-60 category.
- No Customer below the age of 18 years.
- Another Age group has an approximately equal share
except for the 65+ category.



New customers:

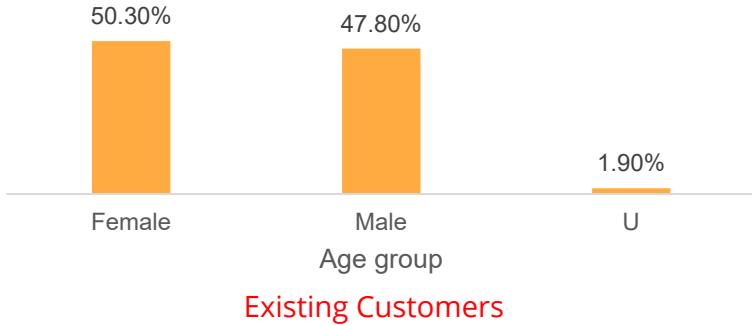
- Most customers belong to above 65+ age.
- Next target range shall be between 41-50 years.
- No Customer below the age of 18 years.



Data Exploration

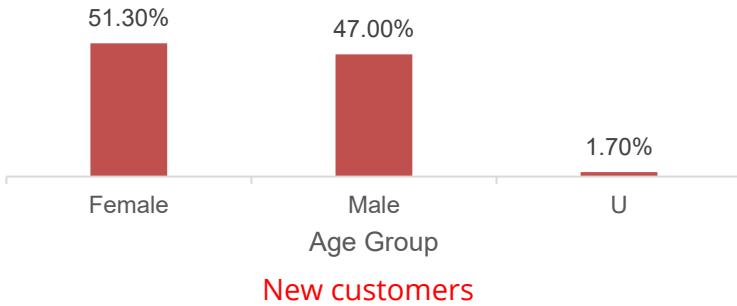
Gender-wise Distribution

- Female Customers bought more than 50% of bikes in existing Customers and new target Customers lists.
- Unknown gender has a minimal percentage impact which could ignore.



Numerically

- Females bought 4876 more than male customers in the last 3 years. whereas 43 more in new target customers.



Data Exploration

Job Category wise Distribution

Focus Area:

Manufacturing, Financial services, and health sector-related customers are high buyers of products.



The agriculture and Entertainment sector has a low number of Customers.

Data Exploration

- More than 50% of Customers are Mass Customers in both lists.
- Next are high net-worth Customers.

Wealth segment



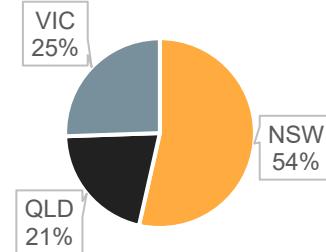
Existing Customers



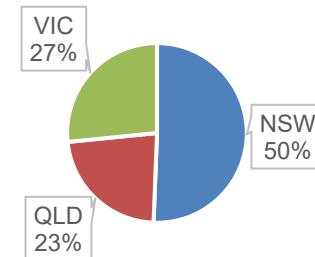
New customers

State wise who owns cars

- NSW has a higher number of Non-car owners which made them high-potential Customers.
- Below comes VIC and QLS state.



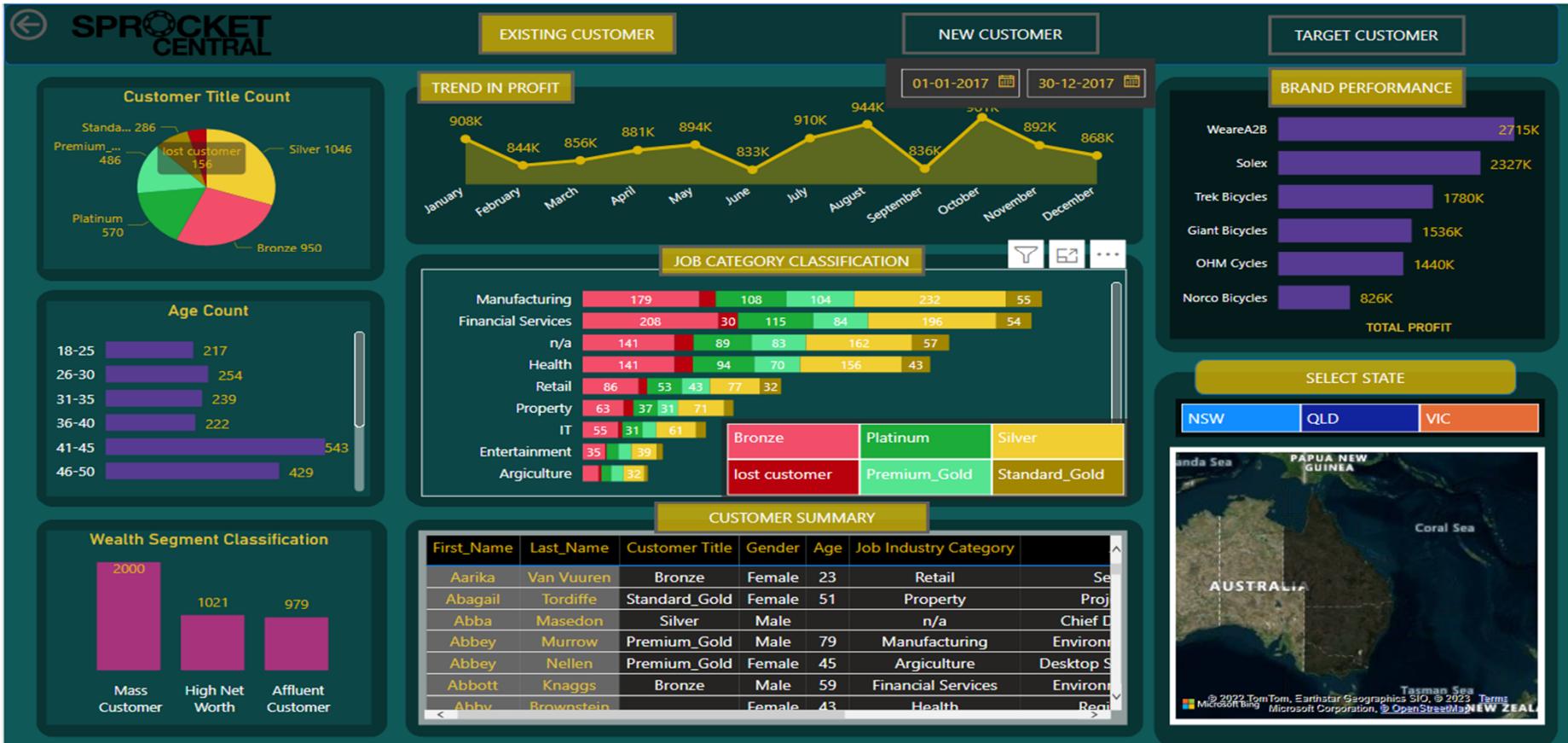
Existing Customers



New customers

Dashboard

Existing Customers Dashboard



Dashboard

New Customers Dashboard

SPROCKET CENTRAL

Customer Title Count

Title	Count
Bronze	122
Silver	373
Platinum	37
Premium_Gold	134
Standard_Gold	186

Age Count

Age Group	Count
18-25	69
26-30	84
31-35	45
36-40	33
41-45	90
46-50	118

Wealth Segment Classification

Segment	Count
Mass Customer	430
High Net Worth	217
Affluent Customer	205

EXISTING CUSTOMER

NEW CUSTOMER

TARGET CUSTOMER

JOB CATEGORY CLASSIFICATION

Category	Financial Services	Manufacturing	n/a	Health	Retail	Property	Entertainment	IT	Agriculture	Telecommunications
Customer Title	25	10	26	78	43					
Platinum		24	11	25	74	37				
Premium_Gold			20	29	54	30				
Silver				17	9	24	56	29		
Standard_Gold					16	34	13			
						5	5	10		
						5	15	9		
						6	13	5		
							15			
								5	6	6

OWN CAR

Industry	Count
Financial Services	91
Manufacturing	87
n/a	68
Health	68
Retail	41
Property	32
Entertainment	21
IT	13
Agriculture	12
Telecommunications	13

SELECT STATE

NSW QLD VIC

Microsoft Bing © 2023 Microsoft Corporation. All rights reserved.

AUSTRALIA

PAPUA NEW GUINEA

Tasman Sea

Coral Sea

Microsoft Bing © 2023 Microsoft Corporation. All rights reserved.

NEW ZEALAND

CUSTOMER SUMMARY

First_Name	Last_Name	Customer Title	Gender	Age	Job_Title
Abigale	Sives	Standard_Gold	Female	58	VP Marketing
Abner	Fraschetti	Bronze	Male	36	Staff Scientist
Adolpho	Bellerby	Platinum	Male	78	VP Sales
Adria	Van den Velde	Silver	Female	44	Nuclear Power Engineer
Adriane	Richardson	Silver	Female	46	Data Coordinator
Adriena	Giffin	Standard_Gold	Female	65	Analog Circuit Design manager
Afton	Andrassy	Silver	Female	24	Web Designer II
Agace	Hedge	Bronze	Female	25	Media Manager II
Agnella	Capener	Standard_Gold	Female	53	Teacher
Agnola	Batterson	Premium_Gold	Female	42	Physical Therapy Assistant
Ahmed	Pickthorne	Silver	Male	62	Marketing Manager
Alyna	Hawasta	Bronze	Female	21	Electrical Engineer

Dashboard

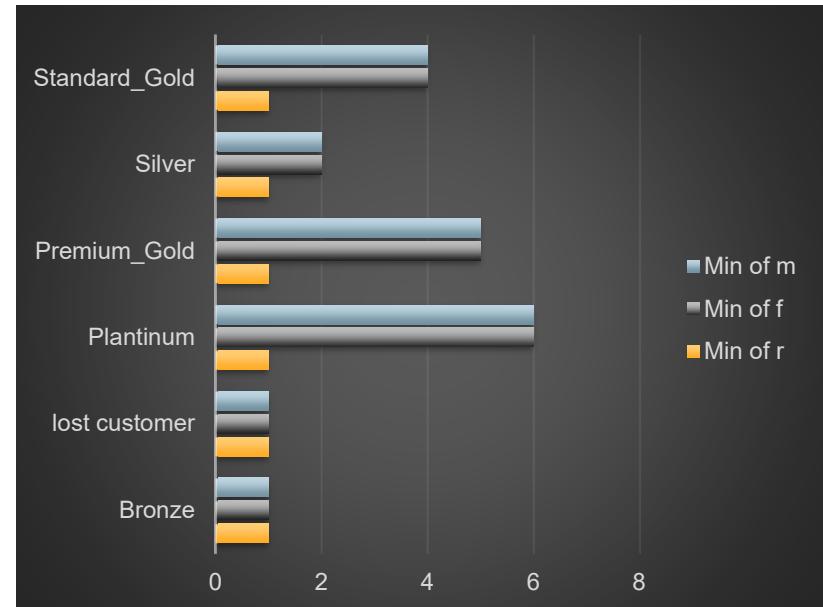
Target Customers Dashboard

SPROCKET CENTRAL							EXISTING CUSTOMER		TARGET CUSTOMER			NEW CUSTOMER			
TOP 1000 EXISTING CUSTOMERS							CAR OWN			TOP 20 % NEW CUSTOMERS			CAR OWN		
Platinum	Premium_Gold		No	Yes	Platinum	Premium_Gold	Standard_Gold		No	Yes		No	Yes		
Customer ID	First_Name	Last_Name	Customer Title	Gender	Age	Last_Purchase	Owns_Car	First Name	Last Name	Customer Title	Gender	Age	Past_Purchases	Owns_Car	
24	Kim	Skpsey	Premium_Gold	Female	45	2023-05-15	Yes	Skippie	Yearsley	Platinum	Male	66	6	No	
165	Aldon	Roelofs	Platinum	Male	38	2023-04-10	No	Sofie	Worsfold	Standard_Gold	Female	68	87	Yes	
340	Joshuah	Purvey	Premium_Gold	Male	42	2023-03-05	Yes	Gerik	Woodroof	Standard_Gold	Male	63	18	Yes	
377	Levy	Bonavia	Platinum	Male	23	2023-02-28	No	Tobias	Woodhams	Standard_Gold	Male	61	0	No	
417	Fernandina	Brambley	Platinum	Female	51	2023-01-15	Yes	Franklin	Wohler	Platinum	Male	84	55	Yes	
443	Lauree	Leason	Premium_Gold	Female	30	2022-12-10	No	Tamar	Windmill	Platinum	Female	83	92	No	
490	Lida	Boynton	Platinum	Female	30	2022-11-05	Yes	Cristen	Wiltshire	Standard_Gold	Female	61	59	Yes	
738	Bunnie	Whines	Platinum	Female	40	2022-10-01	No	Karly	Willavize	Premium_Gold	Female	68	2	No	
749	Andreas	Cosin	Platinum	Male	50	2022-09-15	Yes	Raff	Waycott	Platinum	Male	70	70	Yes	
768	Alexia	Wolford	Platinum	Female	48	2022-08-05	Yes	Ottolie	Wanless	Premium_Gold	Female	82	51	No	
993	Reine	Semken	Platinum	Female	45	2022-07-10	Yes	Latrena	Walklate	Premium_Gold	Female	79	97	Yes	
1181	Brita	Plant	Platinum	Female	44	2022-06-05	Yes	Packston	Wackett	Premium_Gold	Male	72	10	No	
1321	Bree	Domenge	Premium_Gold	Female	28	2022-05-01	No	Clari	Voas	Premium_Gold	Female	67	74	Yes	
1451	Courtney	Meas	Premium_Gold	Female	45	2022-04-15	Yes	Nancie	Vian	Standard_Gold	Female	49	74	Yes	
1493	Lisabeth	Fundell	Platinum	Female	35	2022-03-05	No	Elmira	Vasilyev	Standard_Gold	Female	71	33	No	
1581	Melessa	Delyth	Premium_Gold	Female	48	2022-02-10	Yes	Antony	Tuma	Premium_Gold	Male	68	11	Yes	
1686	Fremont	Trundle	Platinum	Male	42	2022-01-15	Yes	Suzy	Trounson	Standard_Gold	Female	75	13	Yes	
1729	Alla	Kyles	Premium_Gold	Female	35	2021-12-05	No	Dimitri	Tribbeck	Premium_Gold	Male	64	89	No	
1840	Sherman	Gullen	Premium_Gold	Male	45	2021-11-01	Yes	Gabril	Toopin	Premium_Gold	Female	59	7	No	
1901	Vail	Southernwood	Premium_Gold	Male	45	2021-10-05	Yes	Bryonna	Tivers	Standard_Gold	Female	48	81	No	
2058	Ronna	Wickersham	Premium_Gold	Female	44	2021-09-01	No	Loria	Tappington	Standard_Gold	Female	67	72	No	
2221	Tya	Scottini	Platinum	Male	27	2021-08-05	Yes	Hannah	Tompson	Platinum	Male	69	19	Yes	

Model Development

RFM Analysis and Customer Segmentation

- RFM analysis is utilized to determine which customers need to target to increase business.
- The RFM (Recency, Frequency, Monetary) model provides a list of potential buyers who have a high chance to purchase the product.
- Customers divide into a total of 6 categories as mentioned aside: **Platinum and premium_Gold** customers are potential targets for the business.
- All category customers have approximately equal recency behavior.

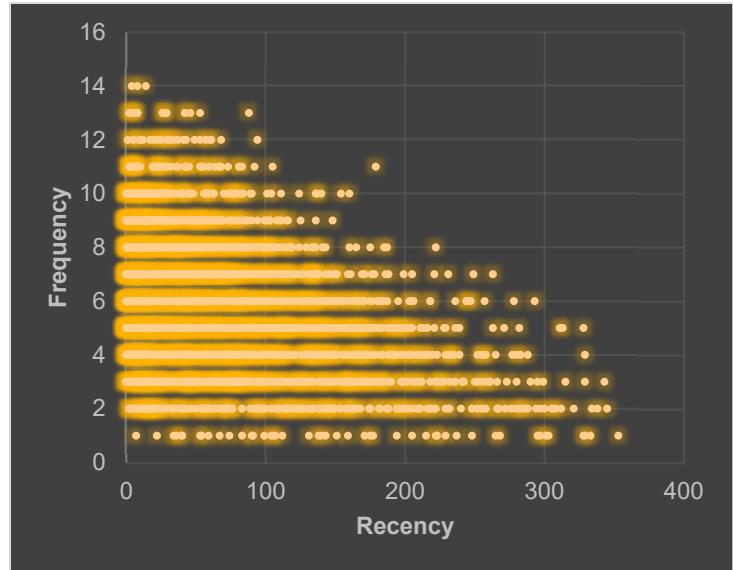


Existing Customers

Model Development

Scatter plot analysis based on RFM model

- Customers that buy within 100 days have high frequency (2-9).
- if high-frequency buyers do not buy within 200 days. then there is a high chance of fewer orders.
- Most customers buy within the (2-10) token range.
- Recency has an inverse relationship with frequency means customers that bought recently will have high tokens purchased.

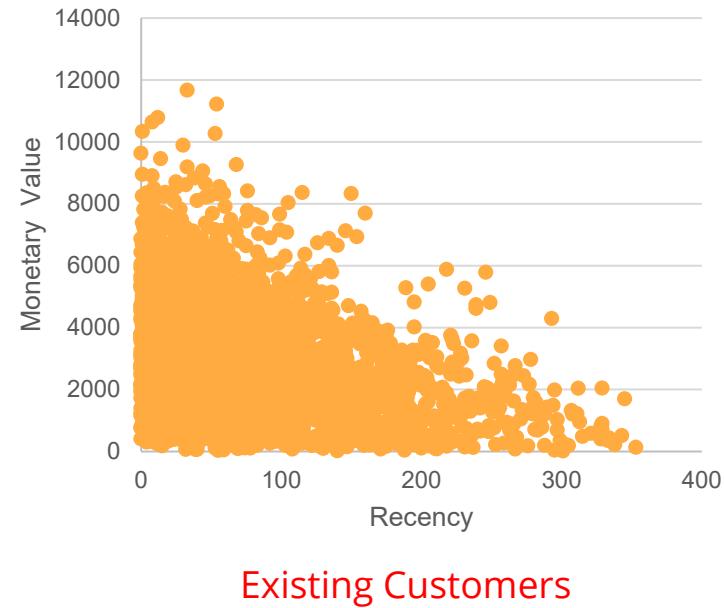


Existing Customers

Model Development

Scatter plot analysis based on RFM model

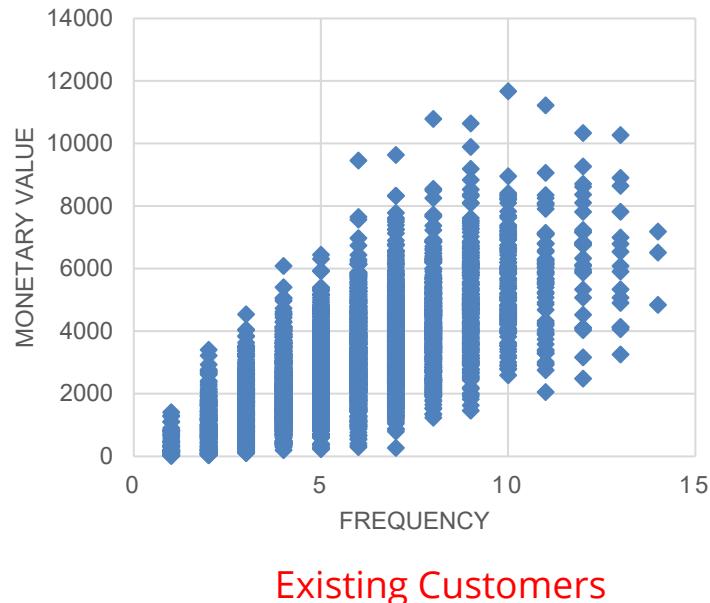
- Customers who purchase more recently have generated more revenues.
- Up to 100 days, customers purchased with high monetary value (approximately 6000\$). From onward both monetary and recency value decreases.
- From 200 days onwards, monetary value sharply decreases.



Model Development

Scatter plot analysis based on RFM model

- Frequency and Monetary Value has a correlated relationship between them.
- Customers who purchase between (5-10) times have higher Monetary value with consistency.
- Customers that purchased more than 10 times have generated high monetary value but with inconsistency.



Model Development

Customer Title Definition based on RFM assigned values

RANK	CUSTOMER TITLE	DESCRIPTION	RFM VALUE RANGE
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	101-126
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	85-100
3	Standard_Gold	RECENT BUYER,BUY INCONSISTENTLY,AVERAGE SPEND	75-84
4	Silver	PURCHASE LONG TIME AGO,BUY INCONSISTENTLY,AVERAGE SPEND	51-74
5	Bronze	LOW RECENCY,SEDOMELY BUY,BELOW AVERAGE SPEND	26-50
6	Lost customer	LOW RECENCY,SEDOMELY BUR,LOW SPEND	1-25

Existing Customer

Model Development

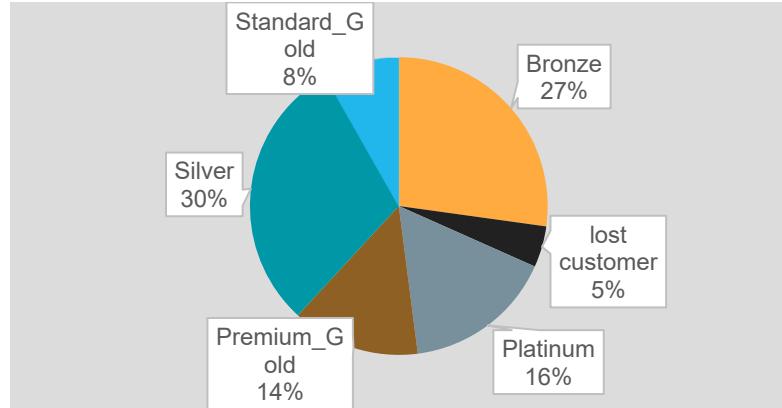
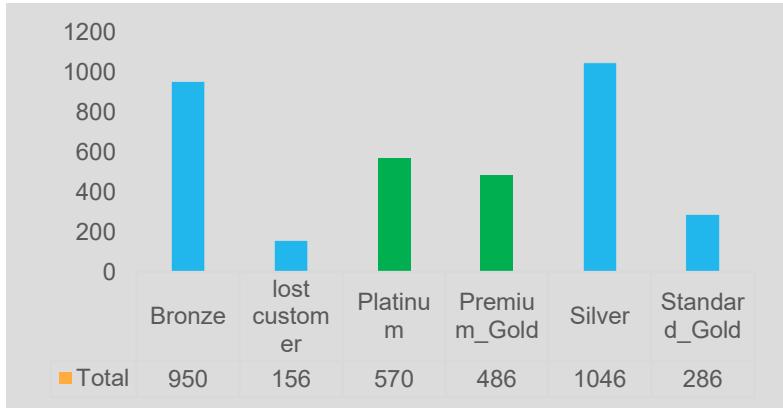
Customer Title Definition based on RFM assigned values

RANK	CUSTOMER TITLE	DESCRIPTION	RFM VALUE RANGE
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	1500-1650
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	1400-1499
3	Standard_Gold	RECENT BUYER,BUY INCONSISTENTLY,AVERAGE SPEND	1250-1399
4	Silver	PURCHASE LONG TIME AGO,BUY INCONSISTENTLY,AVERAGE SPEND	1000-1249
5	Bronze	LOW RECENCY,SELDOMELY BUY,BELOW AVERAGE SPEND	800-999

New Customer

Model Development

Customer Distribution

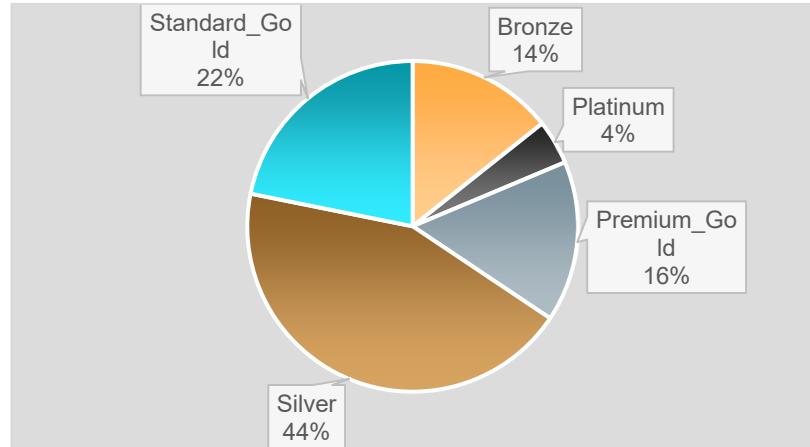


Existing Customer

- About 57% of customers belong to the silver and bronze categories.
- Approximately about 30% of customers are from platinum and premium_Gold categories.

Model Development

Customer Distribution



New Customer

- 44% of customers belong to the silver category.
- Approximately about TOP 20% of customers are from platinum and premium_Gold categories.

Data Analysis

Customer segregation based on RFM condition

RANK	CUSTOMER TITLE	DESCRIPTION	No of Customers	Selected customers
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	570	570
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	486	430
3	Standard_Gold	RECENT BUYER,BUY INCONSISTENTLY,AVERAGE SPEND	286	0
4	Silver	PURCHASE LONG TIME AGO,BUY INCONSISTENTLY,AVERAGE SPEND	1046	0
5	Bronze	LOW RECENCY,SELDOMELY BUY,BELOW AVERAGE SPEND	950	0
6	Lost customer	LOW RECENCY,SELDOMELY BUR,LOW SPEND	156	0

Existing Customer

Data Analysis

Customer segregation based on RFM condition

RANK	CUSTOMER TITLE	DESCRIPTION	No of Customers	Selected customers
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	37	37
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	134	134
3	Standard_Gold	RECENT BUYER,BUY INCONSISTENTLY,AVERAGE SPEND	186	29
4	Silver	PURCHASE LONG TIME AGO,BUY INCONSISTENTLY,AVERAGE SPEND	373	0
5	Bronze	LOW RECENCY,SELDOMELY BUY,BELOW AVERAGE SPEND	122	0

New Customer

Data Analysis

Top 1000 customers to target in Existing Customers

RANK	CUSTOMER TITLE	DESCRIPTION	No of Customers	Selected customers
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	570	570
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	486	430

Top 20% customers to target in New Customers

RANK	CUSTOMER TITLE	DESCRIPTION	No of Customers	Selected customers
1	Platinum	MOST RECENT BUYER,BUY FREQUENTLY,MOST SPEND	37	37
2	Premium_Gold	RECENT BUYER,BUY FREQUENTLY,HIGH SPEND	134	134
3	Standard_Gold	RECENT BUYER,BUY INCONSISTENTLY,AVERAGE SPEND	186	29

- This customers are selected with the RFM model and filtered through to obtain the top 1000 and 20% of the Customers respectively.
- This are the most frequent buyer with high revenue generators in recent times

Data Analysis

High value Customers that should be target:

Criteria	Existing Customers	New Customers
Age	Between 41-50	Above 65+
Job Category	Manufacturing, Financial services, and health sector	Manufacturing, Financial services, and health sector
Wealth	Mass Customers	Mass Customers
State	NSW>VIC>QLD	NSW>VIC>QLD



THANK YOU