

Sprocket Central Pty Ltd

Identify which of the 1000 customers Sprocket Central Pty Ltd should target.

Problem

Problem1: Datasets provided by Sprocket Central Pty Ltd. We have to identify all data quality issues. Dataset have four sheets "Customer Demographic", "Customer Address", "Transactions" and "New Customer List" Once we've had a look at these datasets, draft an email to the client identifying all data quality issues.

Problem2: Create a PowerPoint presentation which outlines the approach we will be taking to identify which of the 1000 customers Sprocket Central Pty Ltd should target, based on the dataset. Explain the three phases: Data Exploration; Model Development and Interpretation.

Problem3: Develop a dashboard that we can present to the client at our next meeting. Display your data summary and results of the analysis in a dashboard (see tools/references for assistance). Specifically, your presentation should specify who Sprocket Central Pty Ltd' should be targeting out of the new 1000 customer list.

Tasks

Problem1: Data quality issue

- ❖ Analyse statistics summary from the datasets.
- ❖ Found outliers, reoccurrence of data and empty values in certain records.
- ❖ Found Inconsistent values for the same attribute.

Problem2: Data Exploration, Model Development and Interpretation

- ❖ Customer Analysis
- ❖ Customers Age distribution
- ❖ Job Industry category analysis
- ❖ Top bike model analysis
- ❖ Bike purchases last 3 years analysis
- ❖ Job Industry category analysis
- ❖ Top targeted customer

Problem3: Dashboards

- ❖ Purchase trends of customer
- ❖ Profit Insight
- ❖ Customer segment

Problem 1 Solution

Data Quality issue Solution

Table Name	No. of records	Distinct Customer IDs
Customer Demographic	3912	3912
Customer Address	3999	3999
Transaction Data	20000	19640
New Customer list	1000	1000

❑ Various columns, such as the brand of a purchase, or job title, have empty values in certain records

❑ Inconsistent values for the same attribute

e.g. -> *Victoria being represented as “V”, “Vic” and “Victoria”*

-> *Male being represented as “M”*

-> *Female being represented as “F”*

Mitigation: Use regular expression to replaced extended values into abbreviations to ensure consistency across addresses.

❑ Inconsistent data type for the same attribute

e.g. -> Change the date type to date format

-> Currency in currency format

-> Numbers in number format

❑ For key datasets, such as transactions, less than 1% of transactions (totalling less than 0.1% of revenue) have missing fields. These records have been removed from the training dataset.

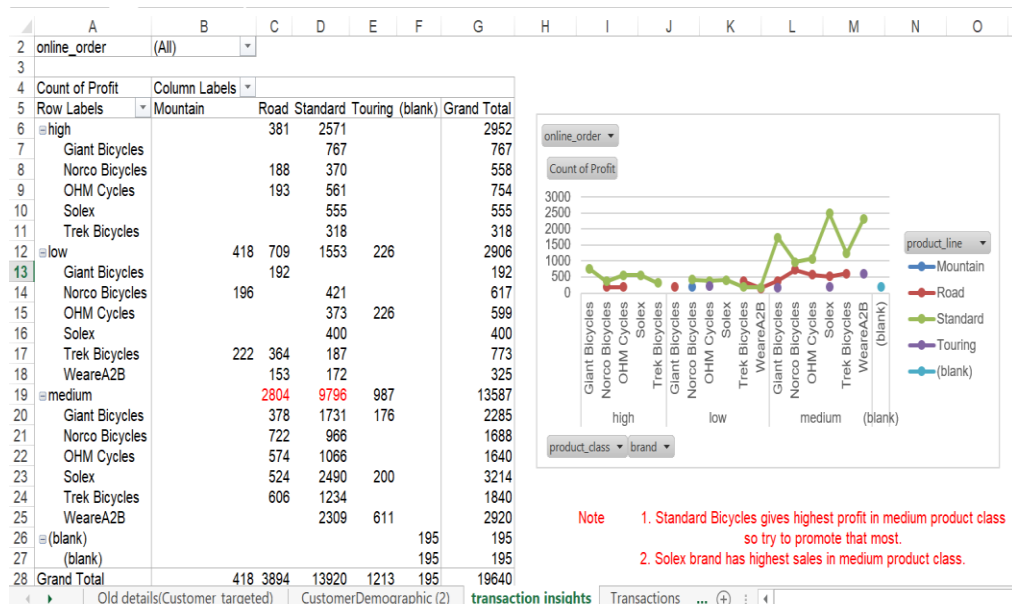
❑ Some Customer's age showing more that 120 years may be it outliers so filter the data.

Problem 2 Solution

Data Exploration

Customer Analysis

- Age Distribution analysis
- Job Industry category analysis
- Top model analysis
- Wealth segment analysis
- Profit analysis
- Gender Distribution



Data Exploration

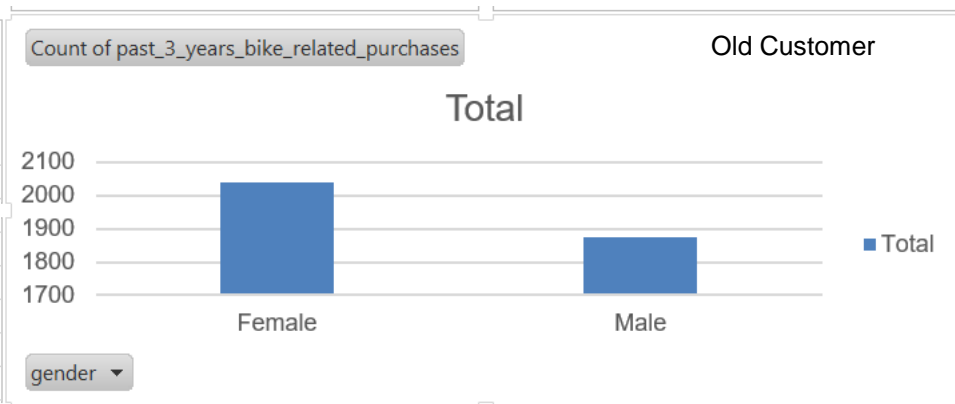
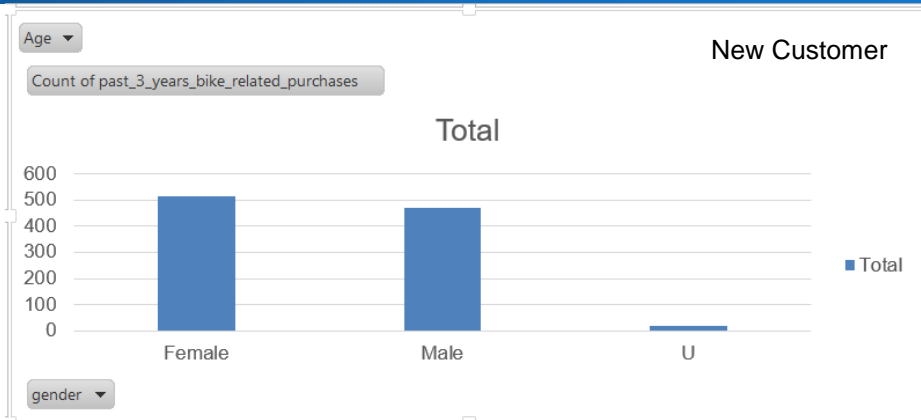
Customers Age distribution

1. According to new customer data we have to focus mostly on the customer females whose age is 41-51 and males whose age is between 21-31.

2. If you want more customers then target females whose age between 51-71.

3. According to old customers Females are more than males

Insights:	Age	Targeted Customer	
	41-51		210
	21-31		150
	51-61		149
	61-71		142
	Above 71		134
	31-41		95

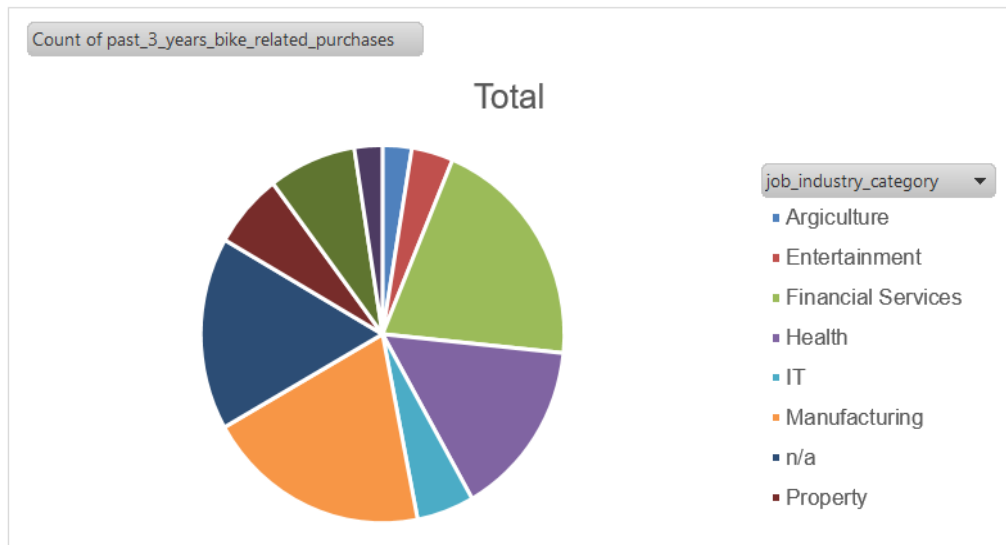


Data Exploration

Job Industry category analysis

Note: Financial Service, Manufacturing, and Health Customers are more so focus on that one

Row Labels	Count of past_3_years_bike_related_purchases
Agriculture	26
Entertainment	37
Financial Service	203
Health	152
IT	51
Manufacturing	199
n/a	165
Property	64
Retail	78
Telecommunica	25
Grand Total	1000

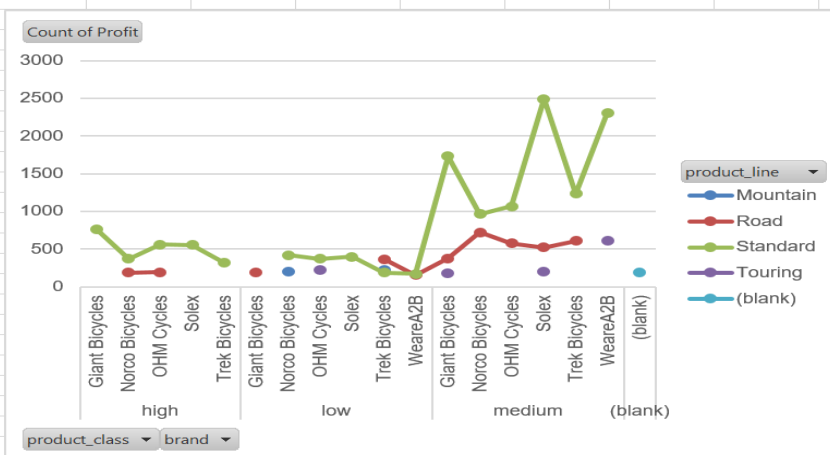


Data Exploration

Top bike model analysis

1. Standard Bicycles gives highest profit(9796) in medium product class so try to promote that most after that promote Road product line.
2. Solex brand has highest sales in medium product class.

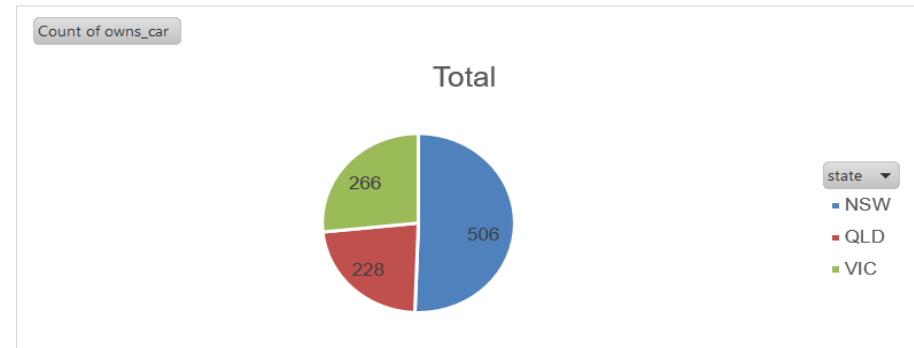
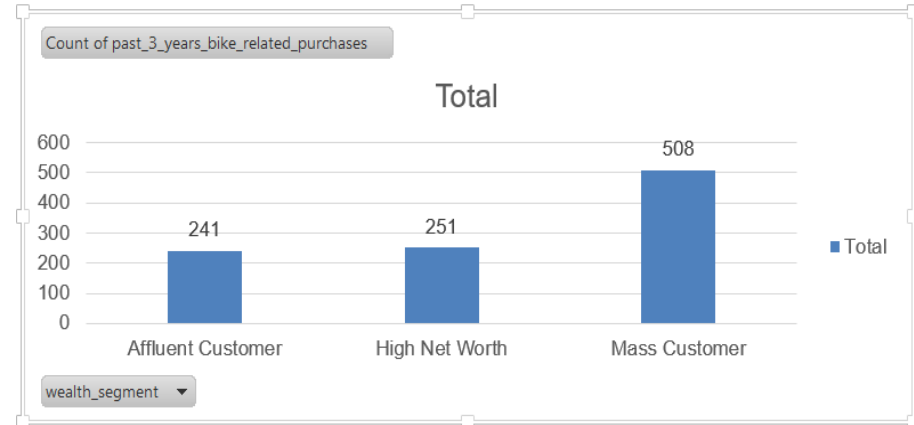
Count of Profit	Column Labels				
Row Labels	Mountain	Road	Standard	Touring	(blank)
high		381	2571		
Giant Bicycles			767		
Norco Bicycles		188	370		
OHM Cycles		193	561		
Solex			555		
Trek Bicycles			318		
low	418	709	1553	226	
Giant Bicycles		192			
Norco Bicycles	196		421		
OHM Cycles			373	226	
Solex			400		
Trek Bicycles	222	364	187		
WeareA2B		153	172		
medium		2804	9796	987	
Giant Bicycles		378	1731	176	
Norco Bicycles		722	966		
OHM Cycles		574	1066		
Solex		524	2490	200	
Trek Bicycles		606	1234		
WeareA2B			2309	611	
(blank)					195
(blank)					195
Grand Total	418	3894	13920	1213	195



Model Development

Bike purchases last 3 years analysis

- ❖ As we can see, our new customers mostly Female with 51.30% purchases and Male contributed to 47% purchases. So, we should focus on advertises on Female customers than Male customers.
- ❖ Focus on Mass customer more.
- ❖ Focus on QLD and VIC state more because they don't own cars so high chance they will purchase bikes.



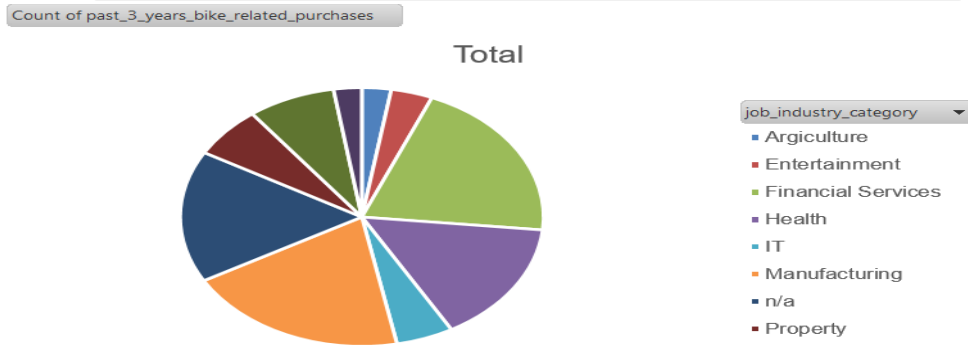
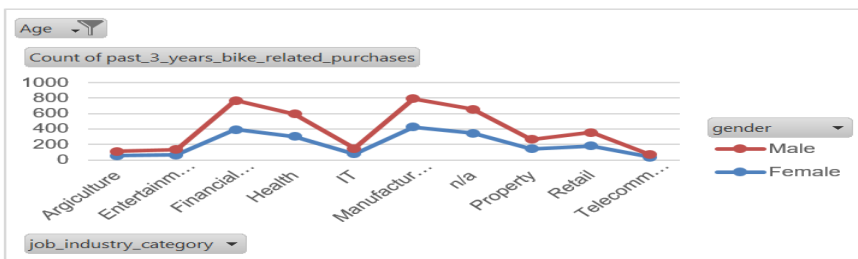
Interpretation

Job Industry category analysis

- ❖ Mostly our new customers are on Finance industries, Manufacturing and Health are still on top3.
- ❖ The rest industries is still same.

Age	(Multiple Items)		
Count of past_3_years_bike_related_purchases	Column Labels		
Row Labels	Female	Male	Grand Total
Agriculture	57	56	113
Entertainment	67	69	136
Financial Services	392	375	767
Health	303	293	596
IT	79	72	151
Manufacturing	425	369	794
n/a	349	307	656
Property	143	123	266
Retail	183	173	356
Telecommunications	39	33	72
Grand Total	2037	1870	3907

Row Labels	Count of past_3_years_bike_related_purchases
Agriculture	26
Entertainment	37
Financial Services	203
Health	152
IT	51
Manufacturing	199
n/a	165
Property	64
Retail	78
Telecommunications	25
Grand Total	1000



Data Insight Conclusion

Top targeted customer

New Customer insights-

- ❑ Focus on females whose age is 41-61 and males whose age is between 21-31.
- ❑ Focus on mass customer followed by High Net Worth
- ❑ Focus on QLD and VIC state more because they don't own cars so high chance they will purchase bikes.
- ❑ Financial Service, Manufacturing, and Health Customers are more so focus on that one

Customer Demographic insights-

- ❑ According to previous data Manufacturing, Financial Service and Health Customers are more
- ❑ Mass customer are 51 % means almost half so focus more on this wealth segment customers .
- ❑ Female customers are more in each field according to previous data

Transaction insights-

- ❑ Standard Bicycles gives highest profit in medium product class so try to promote that most.
- ❑ Solex brand has highest sales in medium product class.

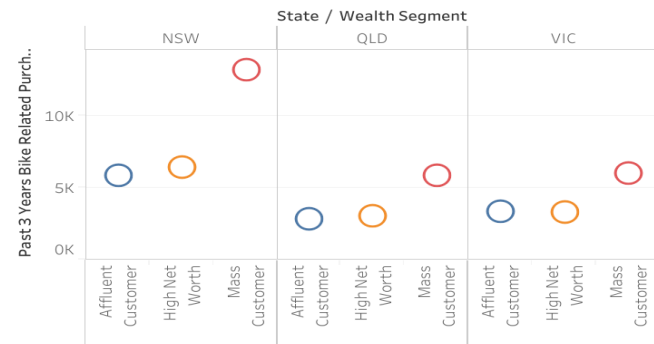
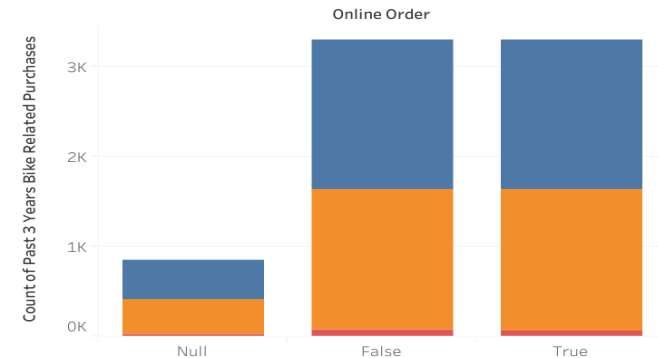
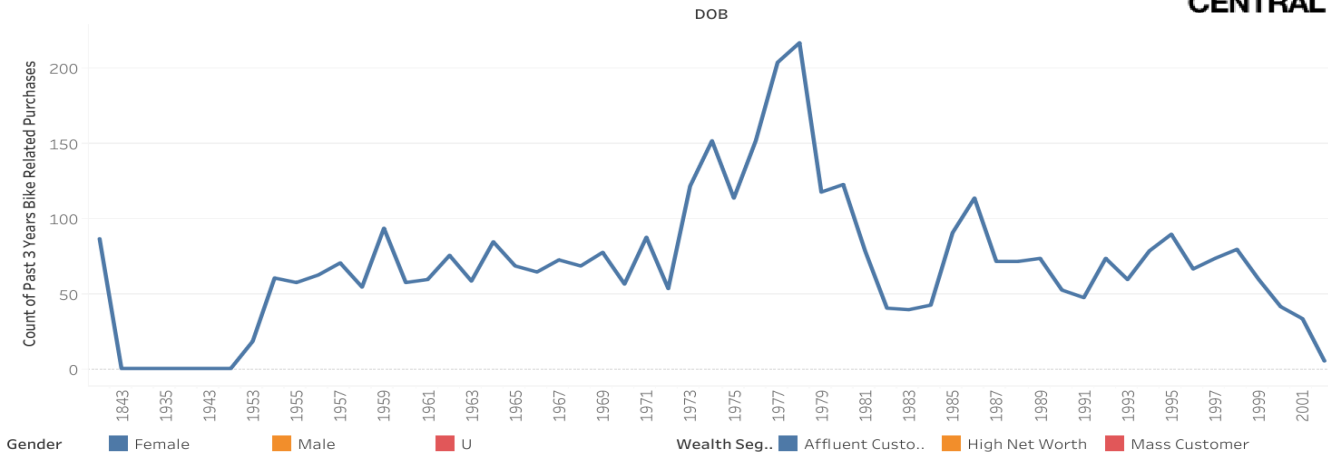
Top 1000 customers		
New customer		61
old customer		416
Top customer who gives you highest profit		588

Problem 3 Solution

Purchase trends of customer

- According to Age Distribution analysis 1978 DOB customer purchase highest bikes.
- In NSW state mass customer Wealth segment purchase highest.
- Females purchase more than males.

Purchase trends of customer



Profit Insight

1 In Quarter 4 we got highest purchase and profit too.

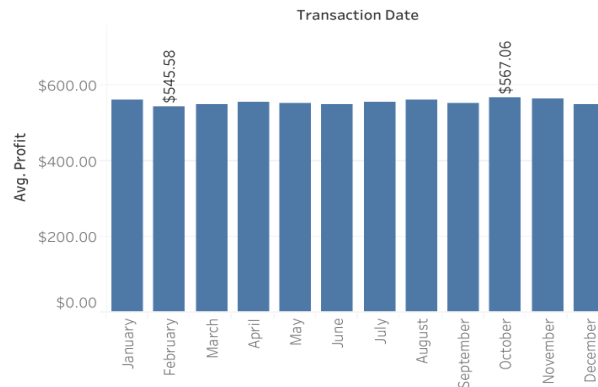
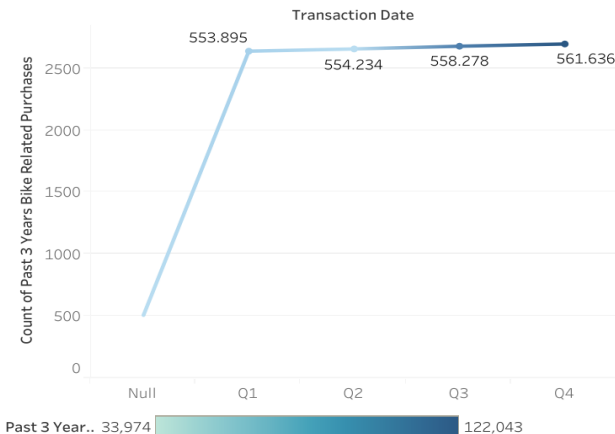
2. In Feb minimum profit we got 545.58\$ and in Oct we got highest profit 567.06\$.

3. Highest purchase brand was Solex and highest profit given by WeareA2B.

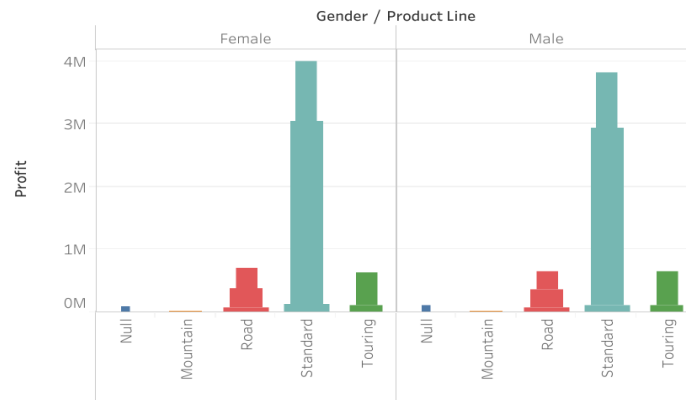
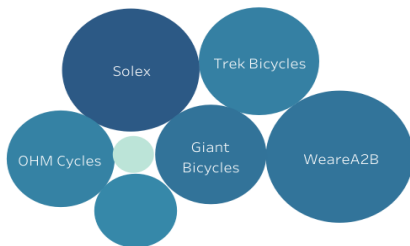
4. Highest profit by female, medium product size bike and standard product line.

SPROCKET
CENTRAL

Profit Insight



Past 3 Year.. 33,974 122,043

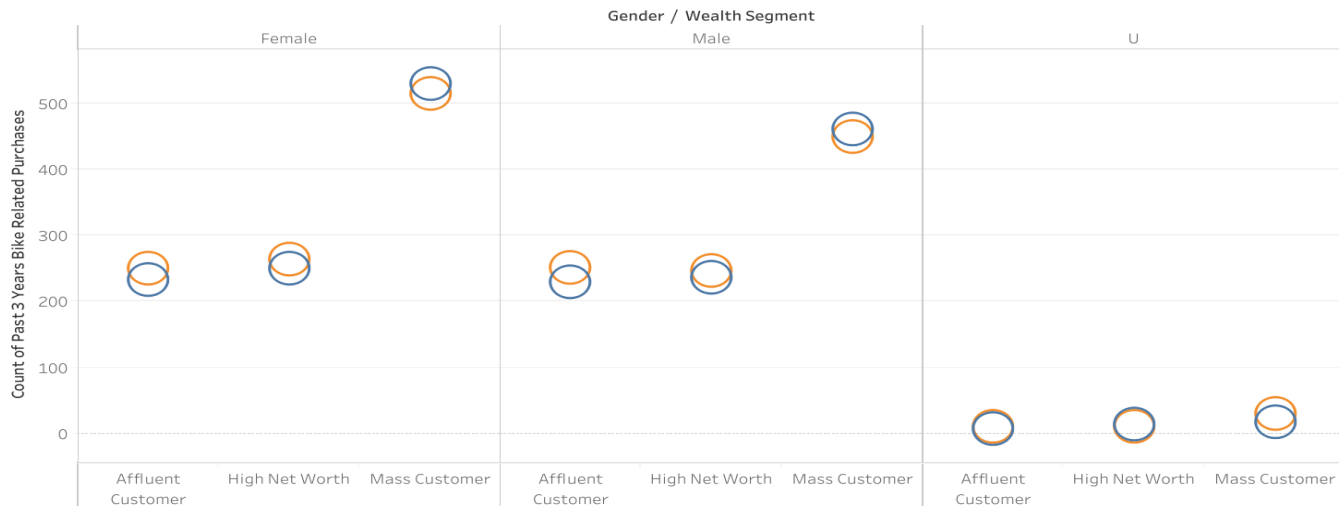
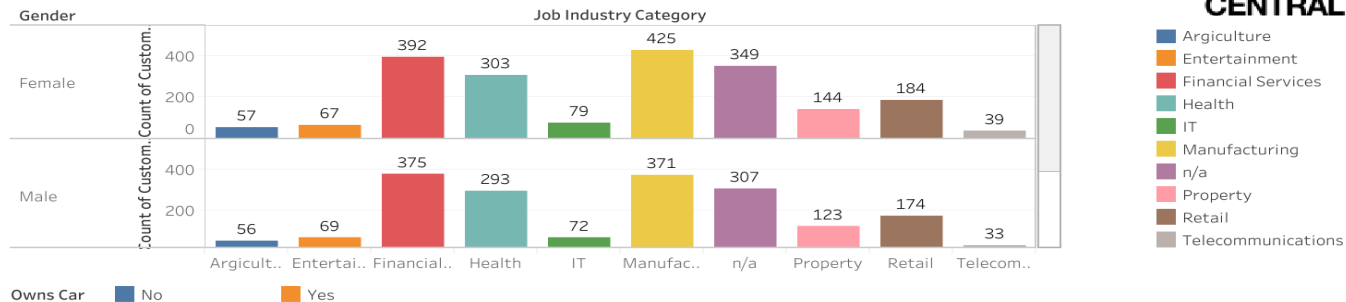


Customer Segment Details

- Highest customers from
→ Financial Service- 392 females and 375 males
→ Manufacturing- 425 females and 371 males

- Mass customer Females are more in number as compare to others.

Customer segment details



Dashboard Conclusion

Purchase trends of customer-

- ❑ According to Age Distribution analysis 1978 DOB customer purchase highest bikes.
- ❑ In NSW state mass customer Wealth segment purchase highest.
- ❑ Females purchase more than males.

Profit Insight-

- ❑ In Quarter 4 we got highest purchase and profit too.
- ❑ In Feb minimum profit we got 545.58\$ and in Oct we got highest profit 567.06\$.
- ❑ Highest purchase brand was Solex and highest profit given by WeareA2B.
- ❑ Highest profit by female, medium product size bike and standard product line.

Customer Segment Details-

- ❑ Highest customers from
 - Financial Service- 392 females and 375 males
 - Manufacturing- 425 females and 371 males
- ❑ Mass customer Females are more in number as compare to others.

Thank You!

- Khushi Yadav

14-04-2023