# Assignment 1

## EDA on Hufty Bikes

## Dataset Information

Unique Transaction IDs - `20510`

## Data validation with great expectation

1. Unique IDs check on transaction_id

Unexpected_count : `0`

2. Dates validations

```
{
    "observed_value" : "datetime64"
}
```

3. Null values check on order_status and brand
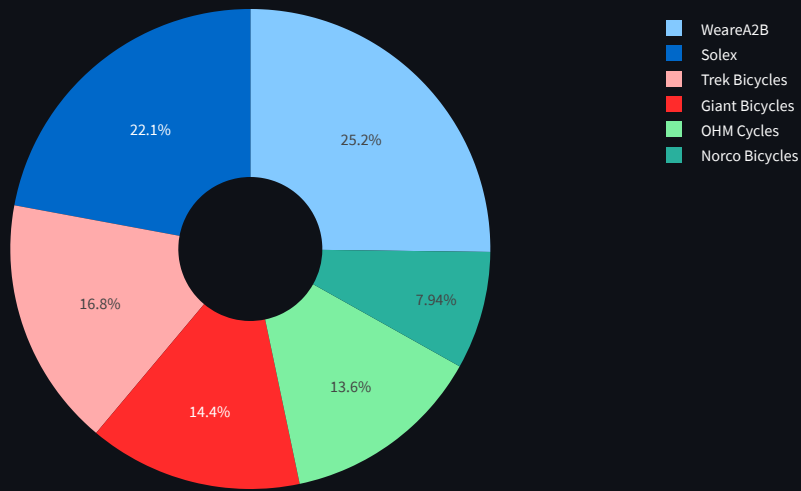
Unexpected_count : `2`

4. Prices are relavant and in between a tight range of 10-2100 $USD

Unexpected_count : `0`

## 1. Brand - Units sold



## 2. Brand - Profits Contributed in $USD

Legend:
- WeareA2B
- Solex
- Trek Bicycles
- Giant Bicycles
- OHM Cycles
- Norco Bicycles

| brand | profit |
|---|---|
| Giant Bicycles | 1,573,840.38 |
| Norco Bicycles | 867,683.77 |
| OHM Cycles | 1,483,038.84 |
| Solex | 2,413,851.605 |
| Trek Bicycles | 1,837,974.2 |
| WeareA2B | 2,753,895.17 |

Total profit generated - 10,930,283 $

Insight - Although the units sold by the brands are almost equal, Solex and WeAreA2B together contribute 47% of the total profit generated.
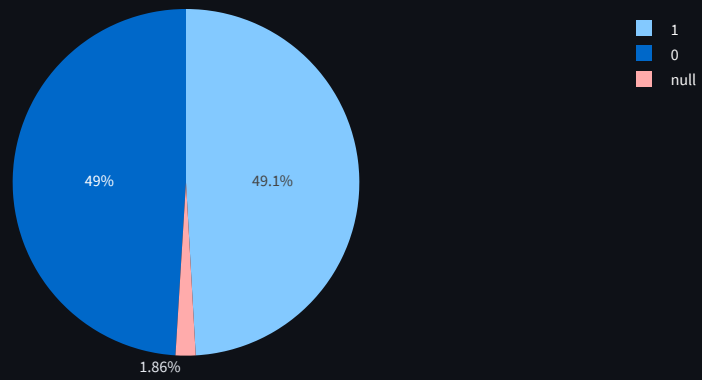
## 3. Customer attributes - Revenue

3.1 Owning cars



Legend:
- Yes
- No
- null

3.2 Gender

Legend:
- Female
- Male
- U
- null

49.7%
47.8%
2.41%
0.00686%
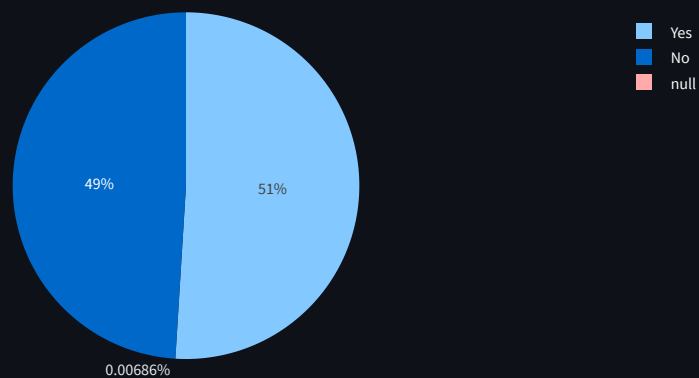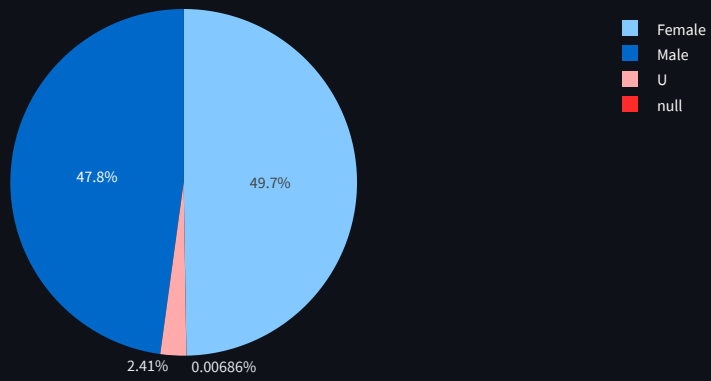
3.3 Online / Offline channel
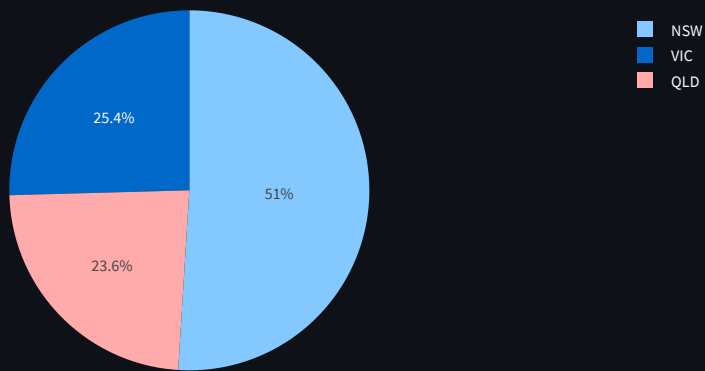


Legend:
- 1
- 0
- null

49.1%
49%
1.86%

Insights - These customer attributes do not affect sales. Gender, Online/Offline Channel, Owning cars

## 4. Customer location analysis

**Existing customer data**



Legend:
- NSW
- Victoria
- VIC
- QLD
- New South Wales

51.5%
22.5%
21.3%
2.48%
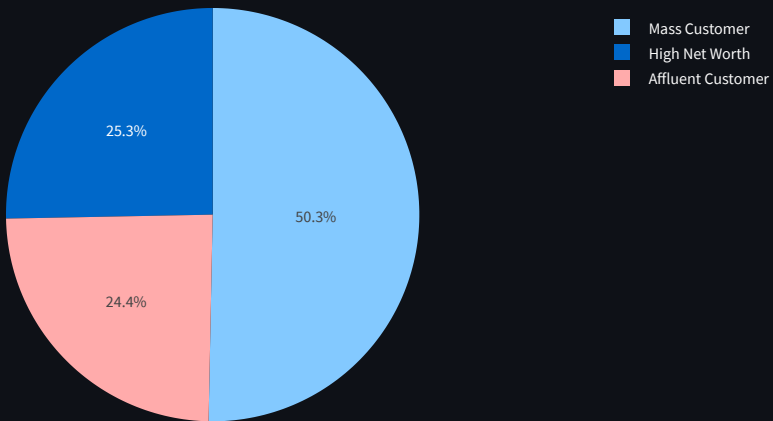2.25%
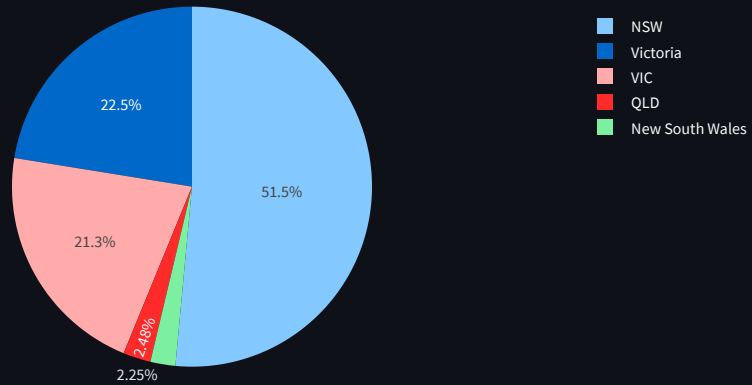
**New customer data**



Insight - New customers are mostly in the same region as the old customers validating that the locations are targeted accurately for the buisiness
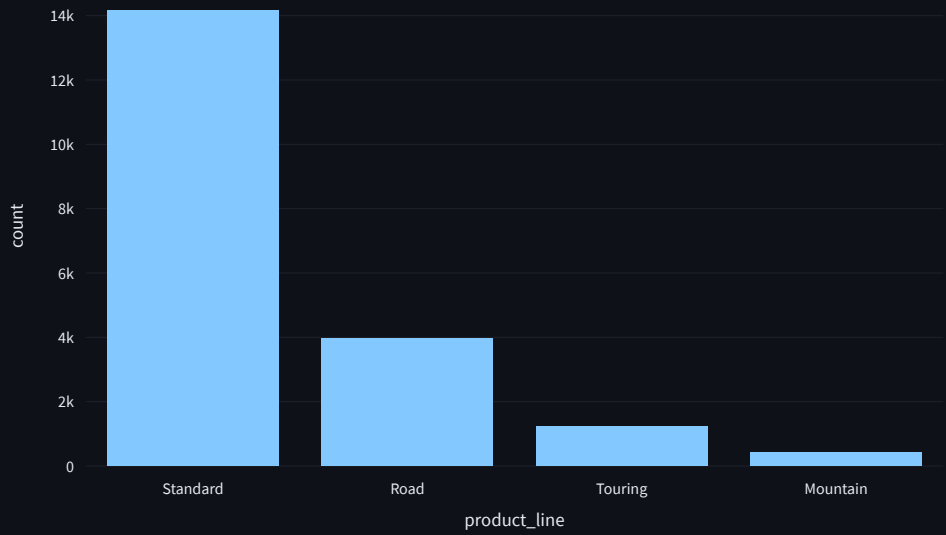
## 5. Customer class & Product class sales analysis
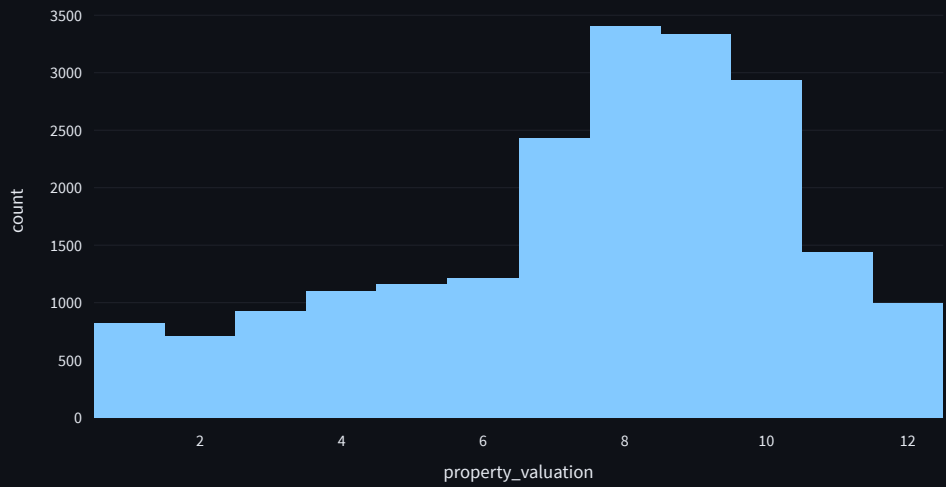
5.1 Wealth segment

| wealth_segment | list_price |
|---|---|
| Affluent Customer | 5,401,437.24 |
| High Net Worth | 5,606,418.48 |
| Mass Customer | 11,147,213.34 |



5.2 Product line sales

## 5.3 Customer wealth (based on residance area)



## 5.4 Customer age
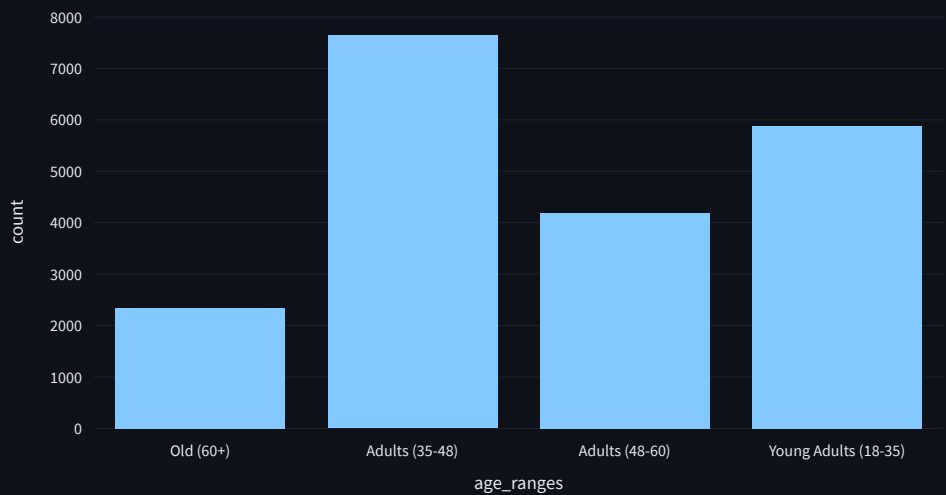
| product_class | profit |
| --- | --- |
| high | 1,115,725.61 |
| low | 1,204,527.66 |
| medium | 8,610,030.695 |

**Product class revenue**



- medium
- low
- high

11%

10.2%

78.8%

| product_class | profit |
| --- | --- |
| high | 1,115,725.61 |
| low | 1,204,527.66 |
| medium | 8,610,030.695 |