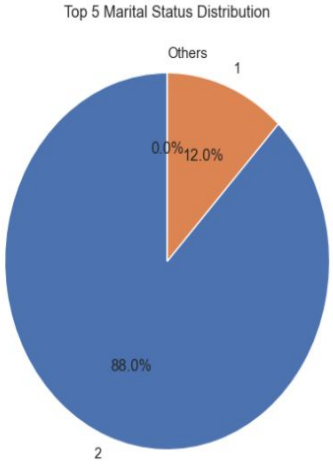
Bank Analysis

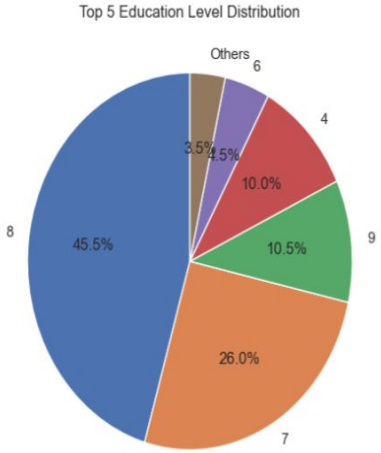
UIN- 66858810 Shanvi Mehta
UIN- 671284857 Akhil Juluru
UIN- 660136009 Shobhit Gopalakrishnan

Profiling of OCBC Customers

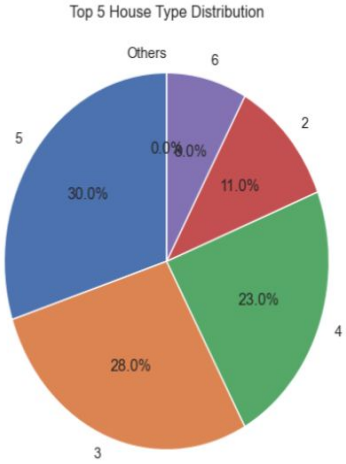
Based on Demographic



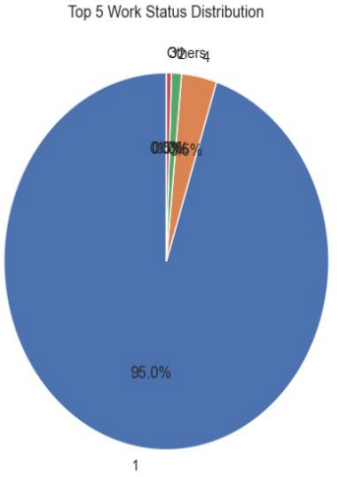
88% Married



University degree-45.5%

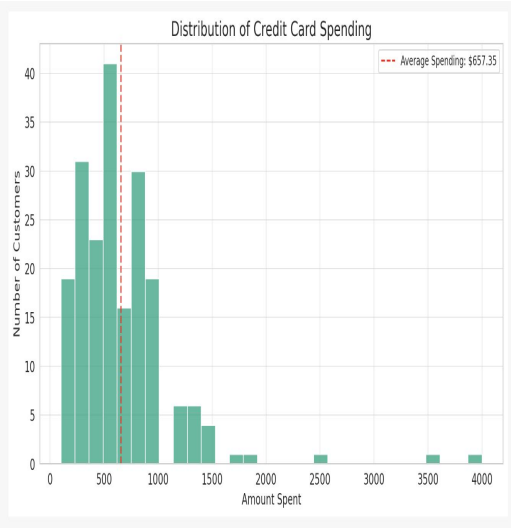


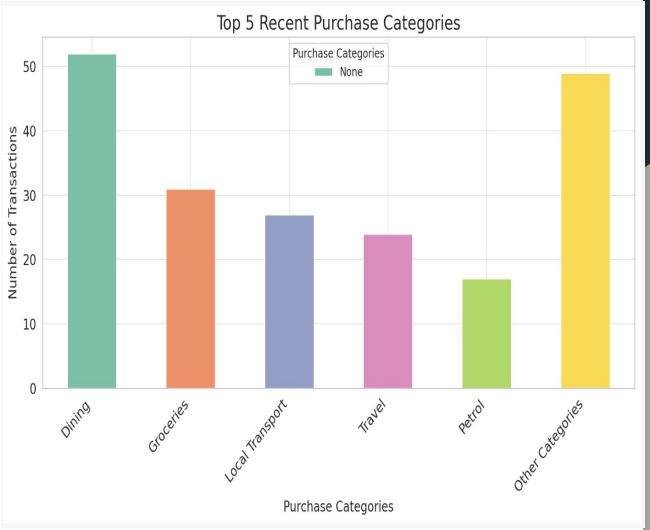
HDB 4/5 RM,Executive flat ,condo, pte apartment



95% Full time

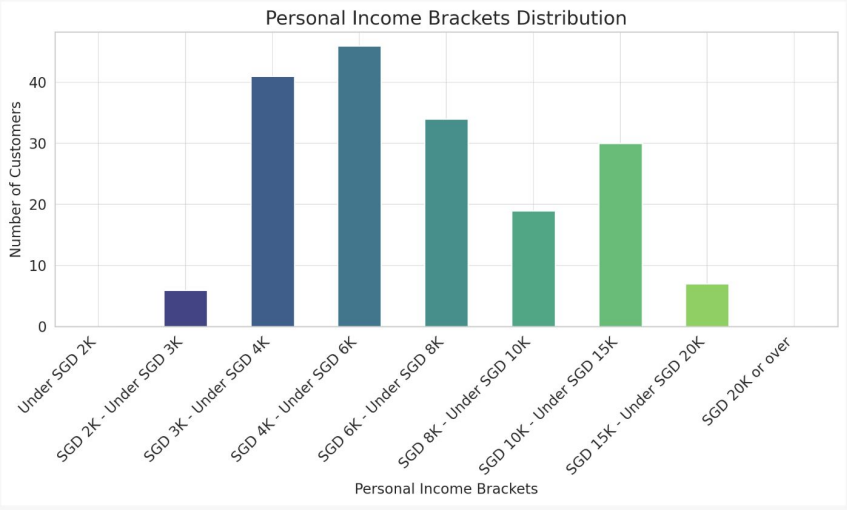
Based on Purchase Behaviour





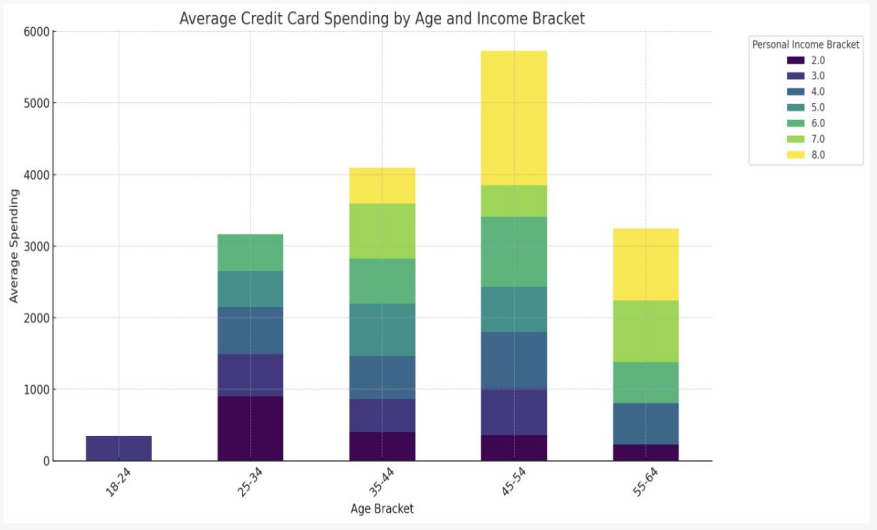
Here we can see Number of customers satisfied in dining privileges are the most and the number of transaction count for dining is maximum. OCBC bank main customers were thus mainly for 365 credit card giving 6% dining discount.

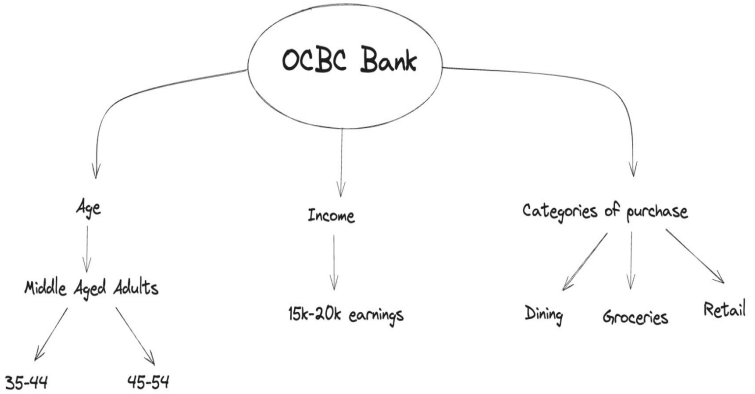
Average spending per customer $657. The distribution is right skewed.



- OCBC bank had more customers with Personal income from range SGD 3k to SGD 8k
- Average spending was more by the customers whose personal income is above SGD 15k-20k

Segmentation and Prioritization

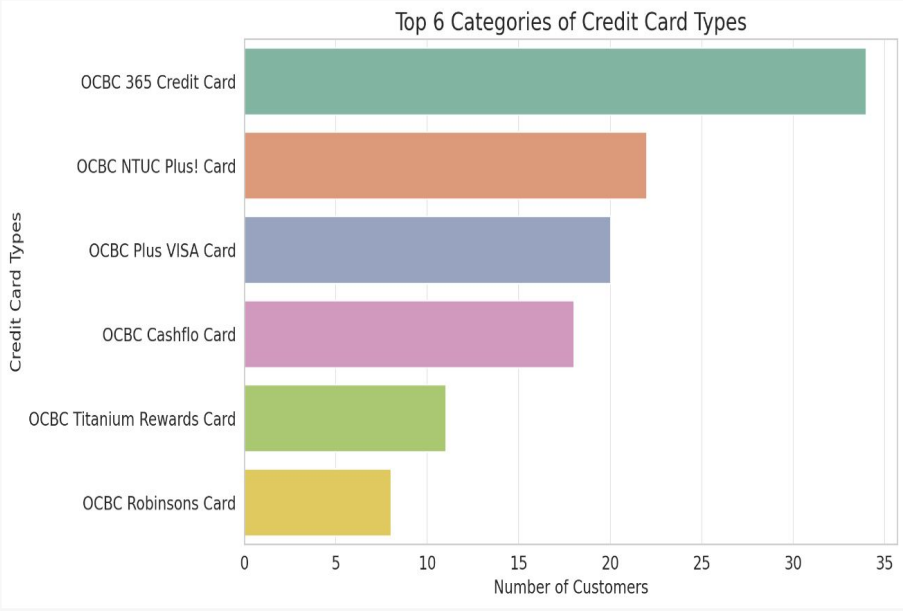


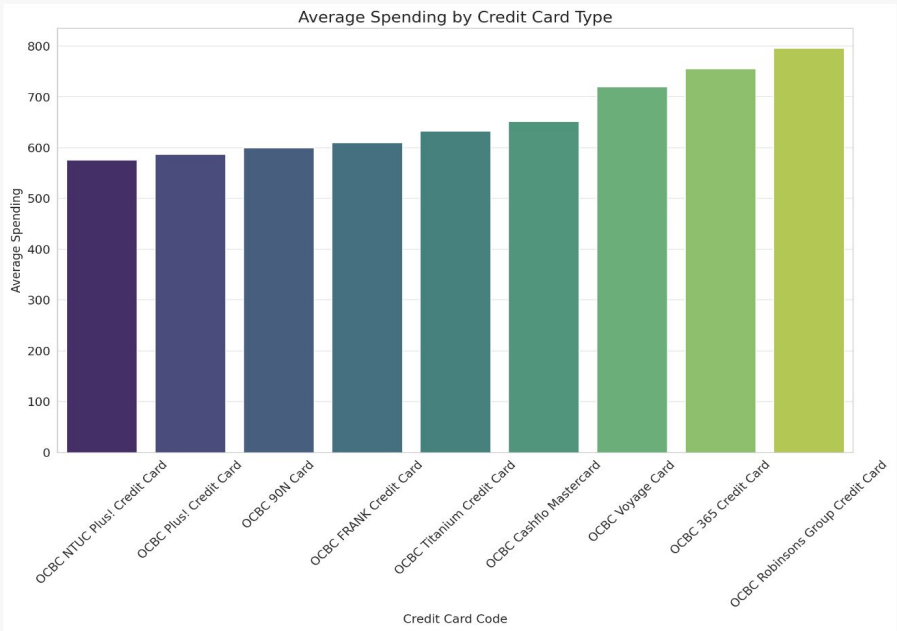


Prioritization

- The company should focus on income group which is between 15k - 20k since the average spending by these customers are high
- The company should target ages between 35-44 and 45-54 since they have an established spending pattern
- OCBC should focus on customers who spend on Dining and Retail since as per data, the customers have spent highly on these categories.

Exploratory Data Analysis for Credit Card types





Analysis

- H0: There is no relationship between customer spending and the type of credit card
- Ha: There is a relationship between customer spending and the type of credit card

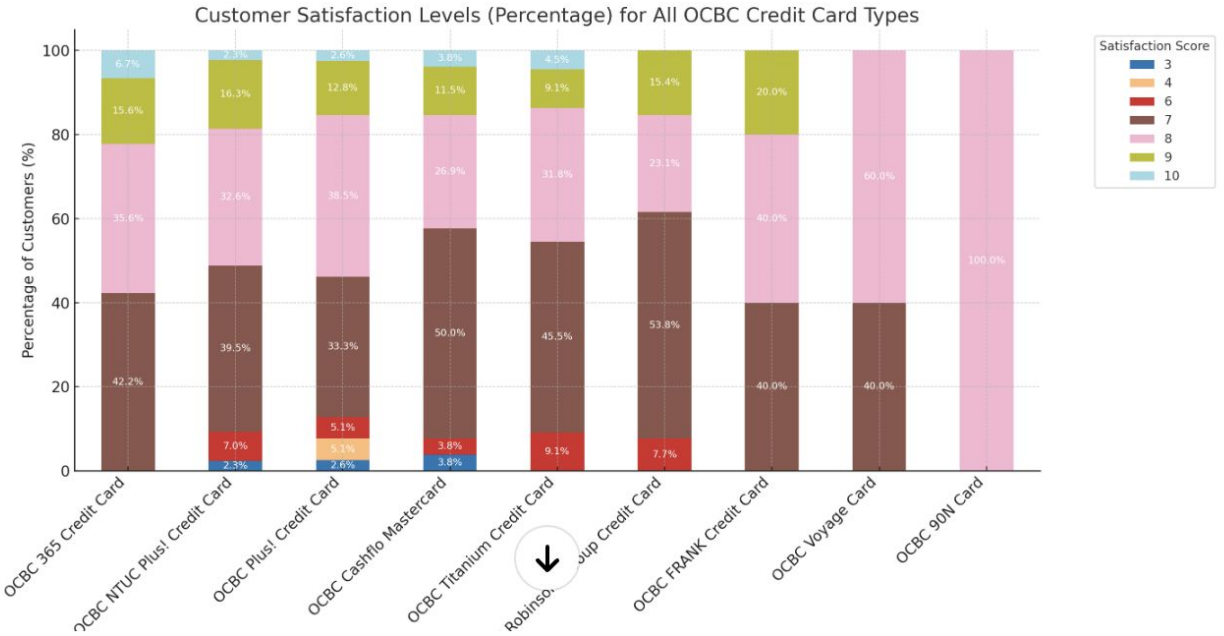
Descriptive Statistics

```
(count      200.000000
 mean       657.350000
 std        474.666217
 min        100.000000
 25%        387.500000
 50%        600.000000
 75%        800.000000
 max       4000.000000
 Name: vn_7002_T23, dtype: float64,
 count                        200
 unique                         9
 top        OCBC 365 Credit Card
 freq                          45
 Name: Q34_Creditcard_code, dtype: object)
```

- Using Chi Square test for the above variables we found that
- Chi Square: 16.251
- P value is 0.436

```
(16.251038601989976,
 0.4355816727495333,
 16,
 array([[11.25,  0.5 ,  6.5 ,  1.25, 10.75,  9.75,  3.25,  5.5 ,  1.25],
        [24.75,  1.1 , 14.3 ,  2.75, 23.65, 21.45,  7.15, 12.1 ,  2.75],
        [ 9.  ,  0.4 ,  5.2 ,  1.  ,  8.6 ,  7.8 ,  2.6 ,  4.4 ,  1.  ]]))
```

Interpretation: Since p value is not less than alpha(0.05) we cannot reject Ho for Ha which means there is no relationship between customer spending and type of credit card



As per the graph we can see that OCBC_90N Card has the highest satisfaction followed by Voyage Card and Frank Credit Card. OCBC should concentrate on these credit cards to increase the loyalty and retention rate

Factors that are significantly associated with repur_code

1. overallx - higher quality of expectations (+ve)
2. pq - price given the quality of the card(+ve). Higher rating associated with reuse of card
3. qp - quality given the price(+ve)
4. repur - likelihood to use the card again(+ve)
5. vn_7002_T01 - feeling of comfort and safety(+ve)
6. vn_7002_T09 - ease of reward redemption(+ve)
7. vn_7002_T25_99 - respondent has not used or interacted with the services are more likely to reuse
8. vn_T002_T22_4 - having a credit card from issuers that the respondent has a complaint with
9. Age - young respondents are more likely to reuse
10. Work - full-time

These are the statistically significant variables affecting the rating of likelihood of reuse of creditcard

	Feature	Coefficient
15	repur	3.410882
31	vn_7002_T01	0.624442
51	vn_7002_T22_4	0.566924
44	vn_7002_T25_99	0.563148
3	overallx	0.535415
65	age	-0.444333
13	qp	0.440666
12	pq	0.423716
67	work	-0.410170
39	vn_7002_T09	0.375683)

Logistic Regression Analysis Accuracy

The model's accuracy on the test set is 95%, with a precision of 97% for predicting the class '0' (low inclination to reuse the card) and 89% for predicting the class '1' (high inclination to reuse).

The confusion matrix indicates that:

- 30 out of 40 instances of class '0' were correctly predicted (with 1 false positives).
- 8 instances of class '1' were correctly predicted (with 1 false negative).

```
('              precision    recall  f1-score   support\n\n           0       0.97      0.97      0.97        31\n
1       0.89      0.89      0.89         9\n\n    accuracy                           0.95        40\n   macro avg
0.93      0.93      0.93        40\nweighted avg       0.95      0.95      0.95        40\n',
 array([[30,  1],
        [ 1,  8]]))
```