Current Customer Segmentation

CurrentCustomer | Historical Transaction | NewCustomer

Targeted promotions and incentives resonate with each group's behaviour and preferences can be used for potentially increasing customer lifetime value and loyalty

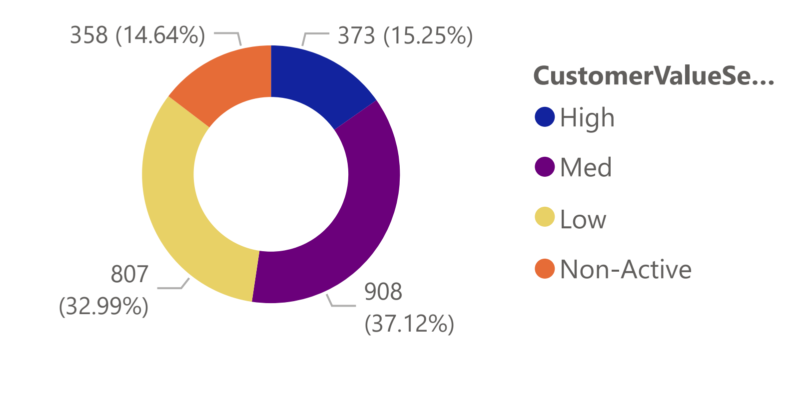
CustomerValueSegment: Non-Active | Low | Med | High

sum_profit: 15.08 – 11,668.95

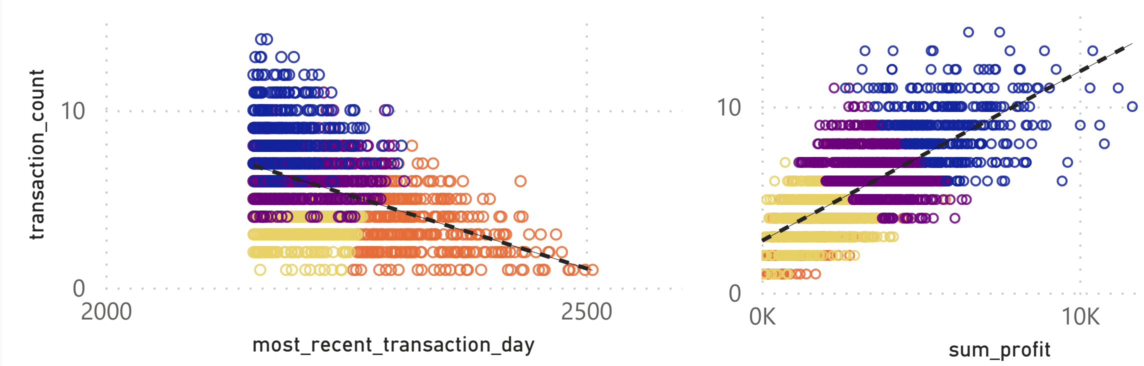
most_recent_transaction_day: 2154 – 2507

transaction_count: 1 – 14

Current Customer Segment



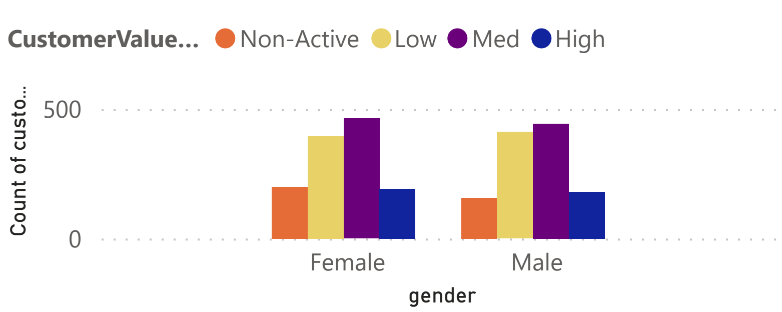
Recency, Frequency and Monetary Value



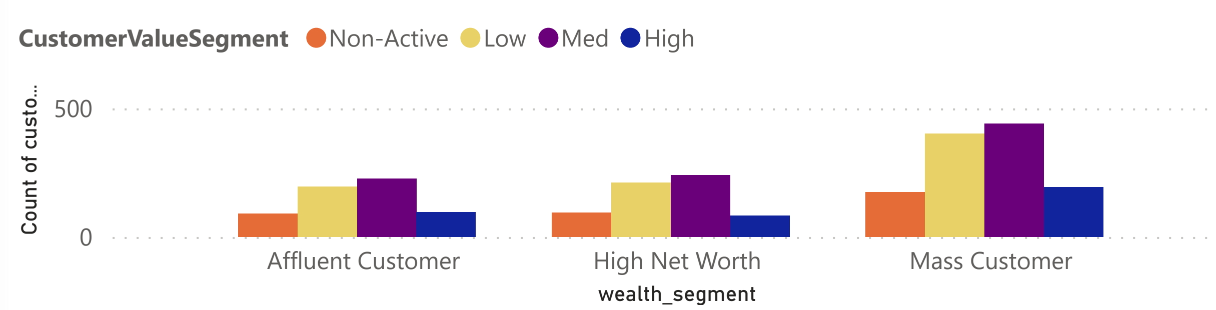
Python visuals are not supported with this operation. Learn more

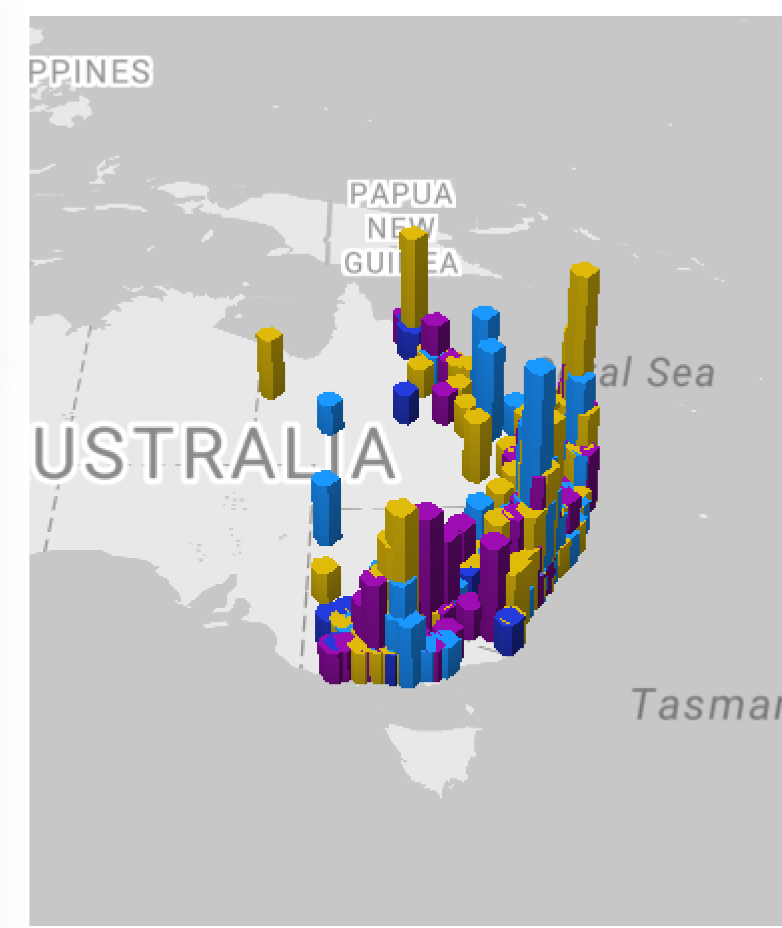
gender: All | wealth_segment: All | job_industry_cate...: All | state: All

Gender

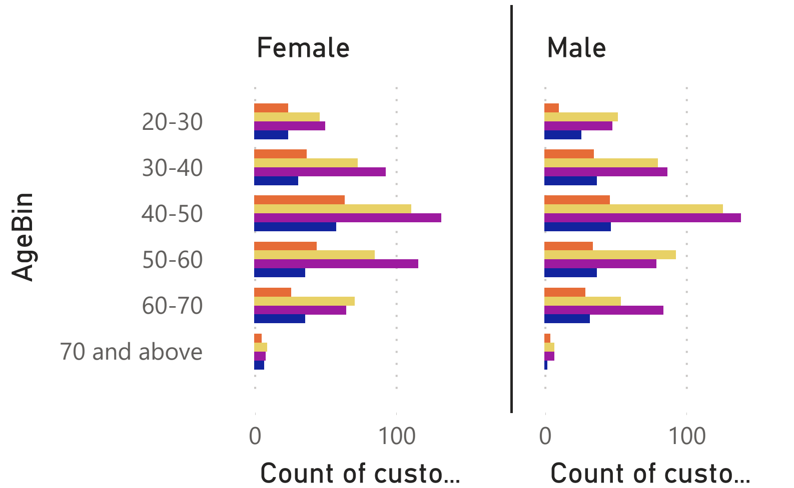


Gender

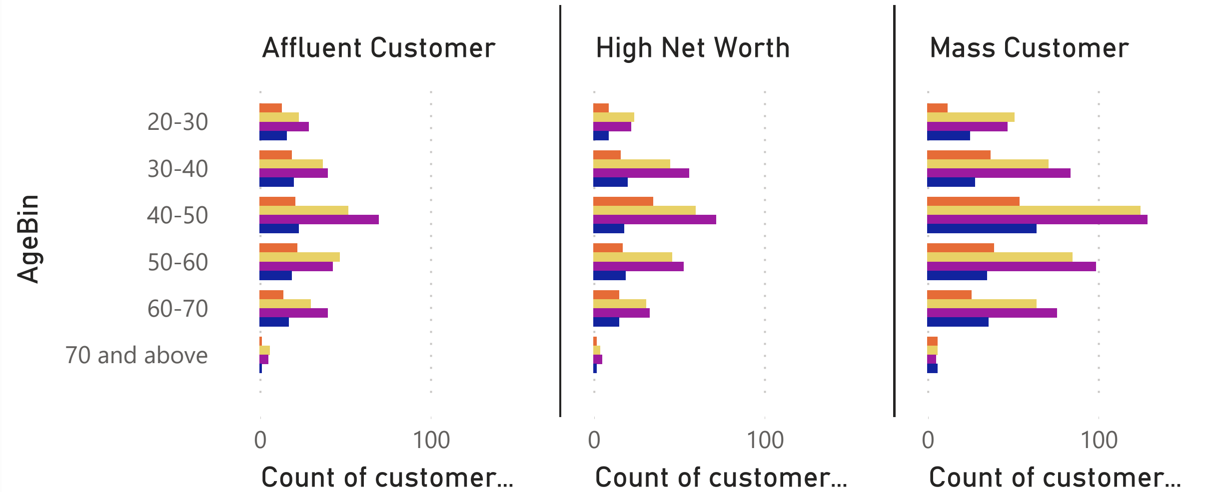




Gender by Age



Wealth Segment by Age



Clustering model identifies four customer behavior segments based on recency, frequency, and monetary of purchases:

- o **Non-active Customers:** No recent purchases or activity.
- o **Low-Value Customers**: New customers with either low profitability or recent but infrequent purchases.
- o **Medium-Value Customers:** Customers with medium profit levels who purchase frequently.
- o **High-Value Customers:** Customers generating high profits with the most recent and frequent purchases

Statistical analysis reveals significance only within RFM features. However, visual data analysis shows distinct distributions in age, tenure, property valuation, job industry, and job title clusters.
Facing challenges in identifying the underlying characteristics of the target RFM cluster, we implemented supervised learning algorithms to discern patterns and classify the target clusters.

Historical Transaction

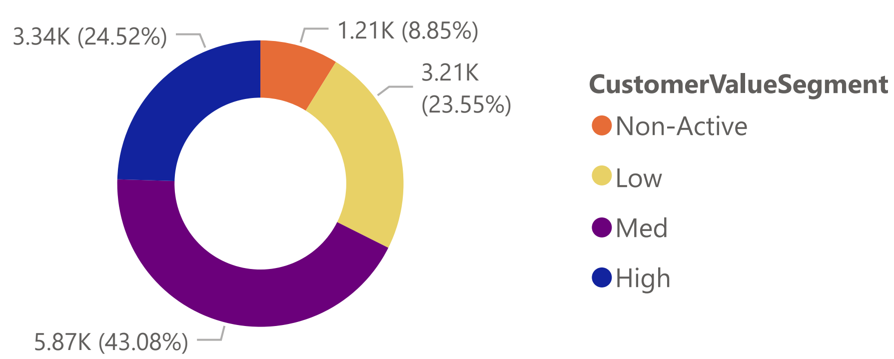
CurrentCustomer | Historical Transaction | NewCustomer

CustomerValueSegment: Non-Active | Low | Med | High

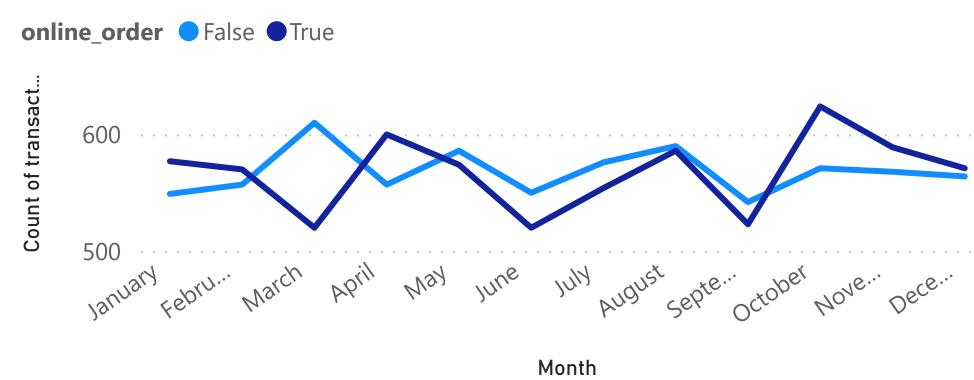
Quarter: All

Day: All

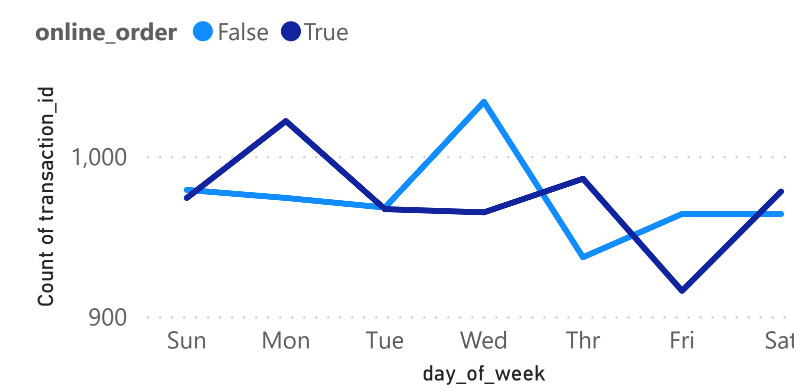
Transaction by Customer Segmentation



Order by Month



Order by Day of the Week

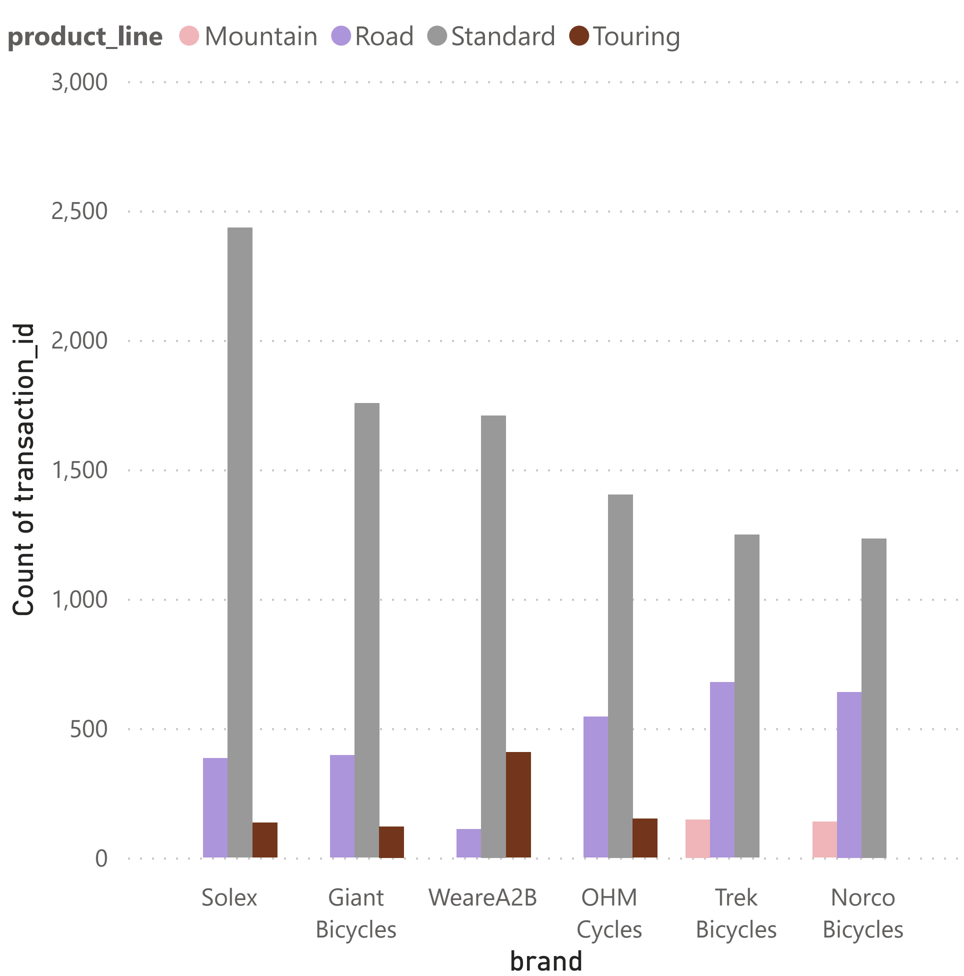


brand: All

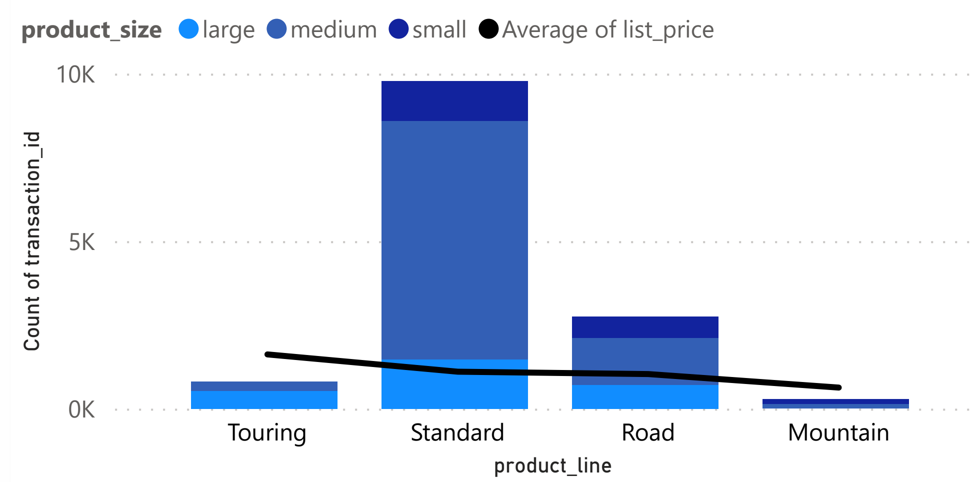
product_line: All

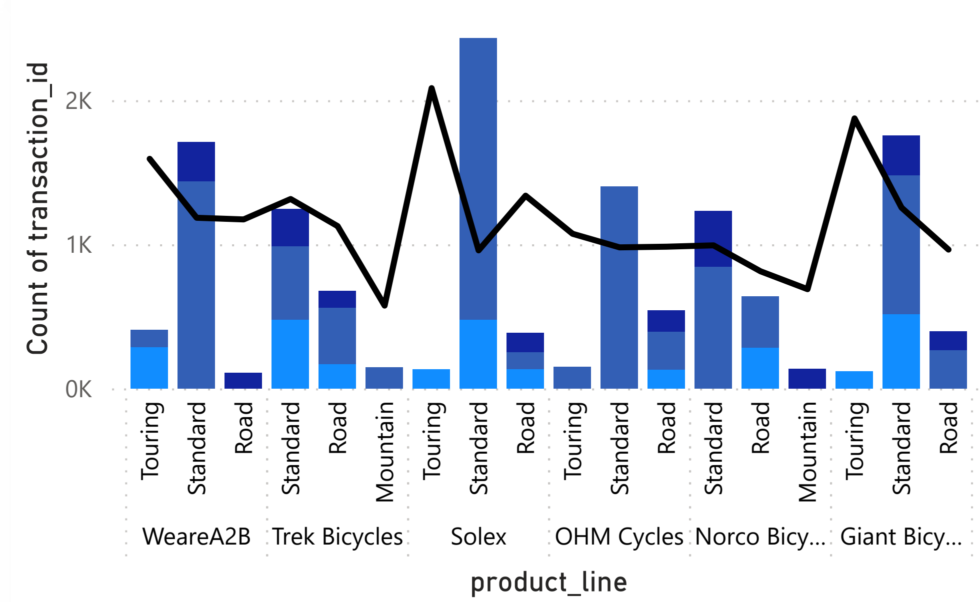
product_size: All

Transaction by Brands



Transaction by Product Lines





state: All

Transactions: 19.45k total, 48.6% by women, 45.85% by men
Transaction Peak: Peak in October and August, followed by July, May, and January.
Online/offline: equally, online Transaction high on Tuesdays/Fridays, offline high on Thursdays/Mondays
Seasonality: online high on Aug, Oct, Jan/ offline high on July, Oct, March

Brands: Popular - Solex, Wearea2b, Giant, OHM; Least - Trek, Norco
Product Lines: Standard, Road, Touring; Least - Mountain
Sizes: Medium most popular, then large, small
Prices: Highest - WeareA2B; Lowest - Norco
Types: Expensive - Touring, then Road, Mountain

New Customer Segmentation

CurrentCustomer | Historical Transaction | NewCustomer

PredictedCustomerSegment: All

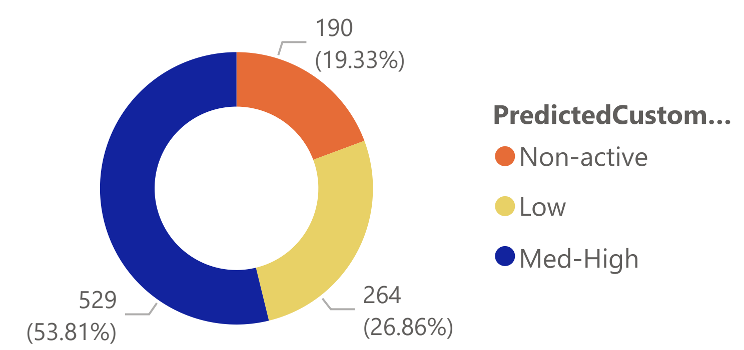
Non-active | Low | Med-High

Rank: 1 – 1000

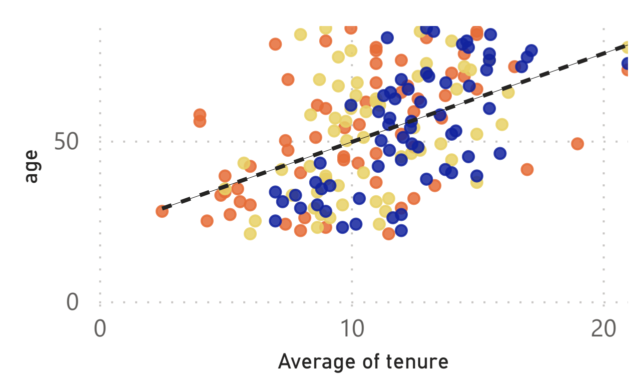
Value: 0.34 – 1.72

state: All

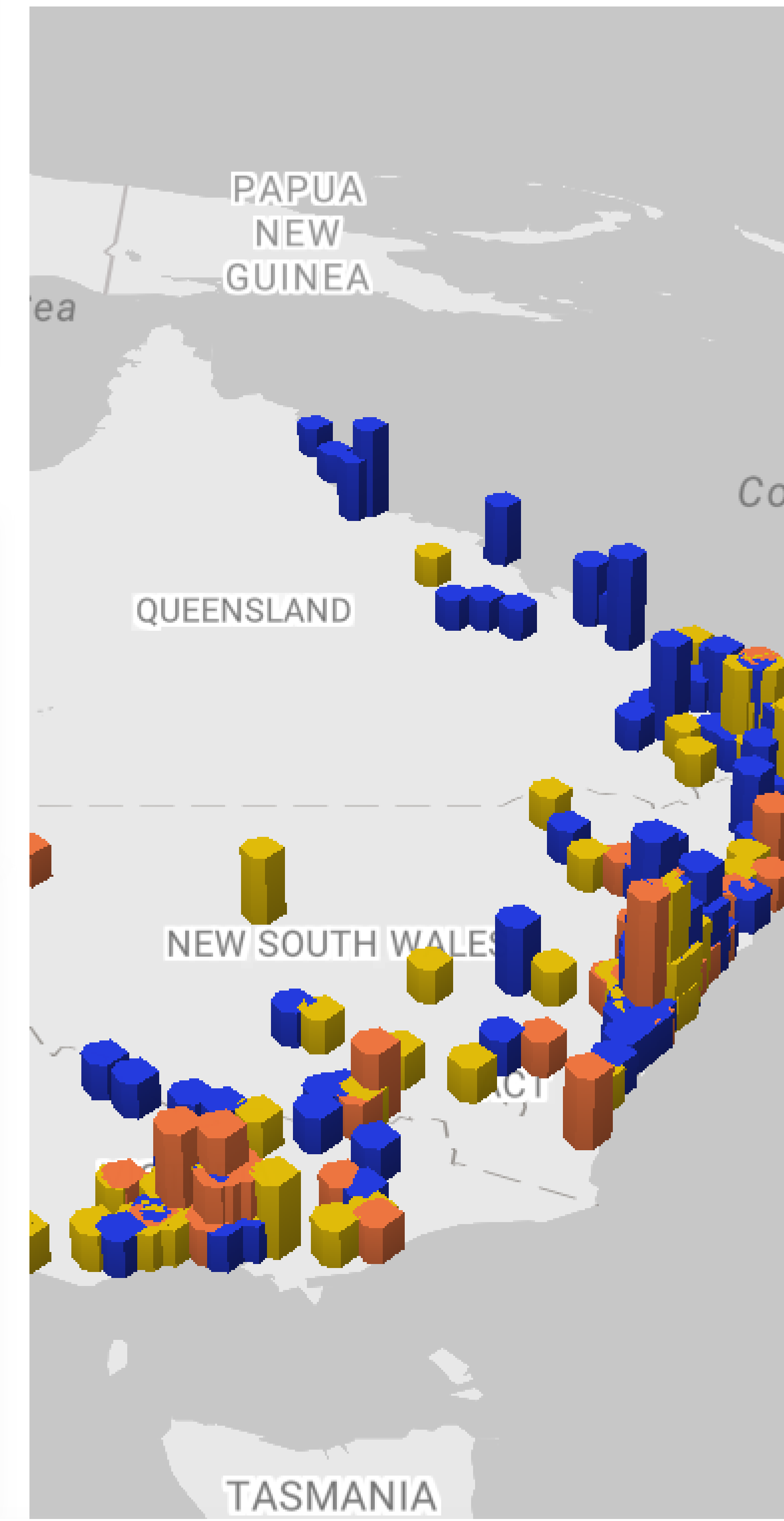
Predicted Customer Segment



Average of Tenure and Age

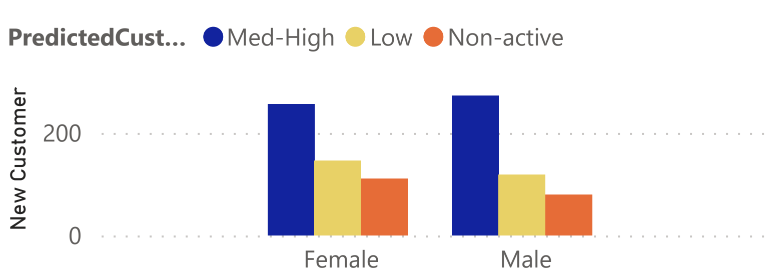


Average of PropertyValuation and Age

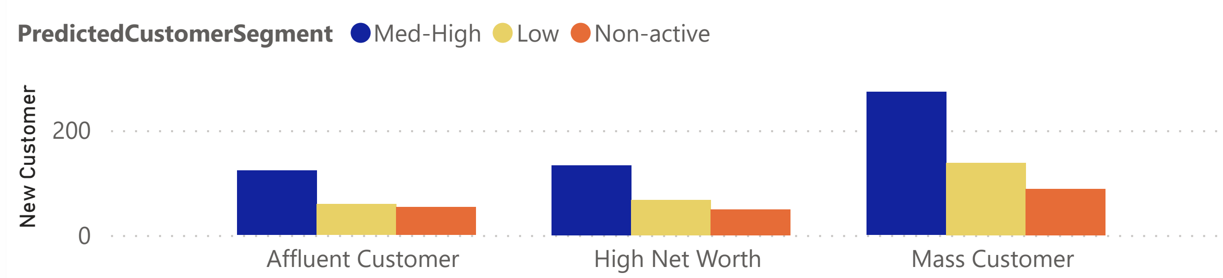


gender: All | wealth_segment: All | job_industry_cate...: All

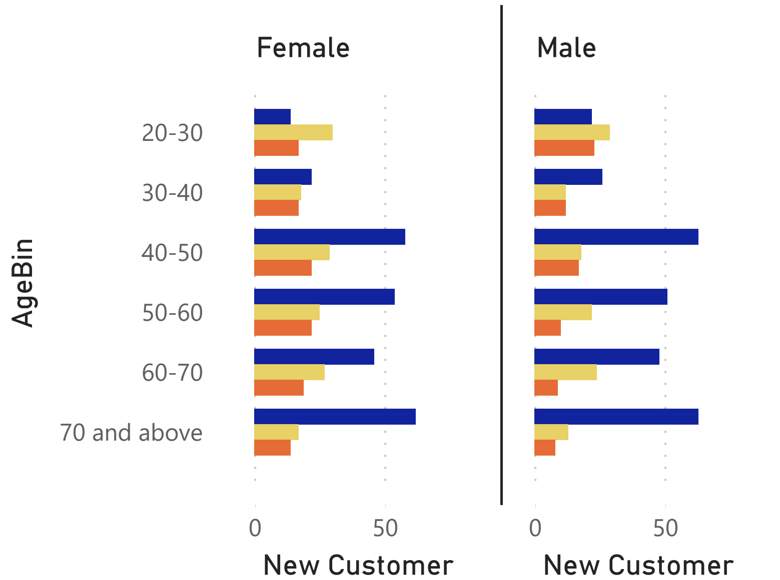
Predicted Customer Segment by Gender



Predicted Customer Segment by Wealth Segment



AgeBin and Gender



Wealth Segment and AgeBin

The result of two-stage binary classification model can identify the target class of customer value, including Non-active, Low-value, and Med/High-value customer. The models achieve performance at AUC score pf 0.92 for Inactive customer identification and 0.69 for high-value customer identification, respectively.

In cases where a new customer's rank and value display a negative correlation, the predicted segment class does not show a relationship with either of these features. This suggests that the segmentation model may be relying on different characteristics or that these particular features do not have a significant impact on the classification within this model. It indicates that the segmentation process may be capturing more complex patterns that are not directly explained by rank and value alone, or there could be other overriding features that have a greater influence on the predicted segment.