Predict Customer Personality to boost marketing campaign by using Machine Learning

Supported by:
Rakamin Academy
Career Acceleration School
www.rakamin.com

Created by:
Saila Fikriyya
sailafs@gmail.com
linkedin.com/in/sailafs1203/

I'm a health psychology graduate turned data scientist, bridging the gap between human behavior and data-driven solutions. With a Master's in Health Psychology from the University of Stirling and ongoing training in data science, I combine expertise in research methods, data analysis, and understanding human behavior. My experience spans from studying post-traumatic growth in cancer survivors to applying data science techniques in digital healthcare and career development. I'm passionate about leveraging interdisciplinary skills to create evidence-based strategies that improve health outcomes and overall well-being, always prioritizing a human-centric approach in collaborative environments.

Overview

Rakamin Academy

“A company can grow rapidly when it understands its customer personality, allowing it to provide better services and benefits to customers who have the potential to become loyal customers. By analysing historical marketing campaign data, it helps to increase performance and target the right customers to make transactions on the company's platform, the focus from these data insights is to create a cluster prediction model to facilitate decision-making for the company."

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

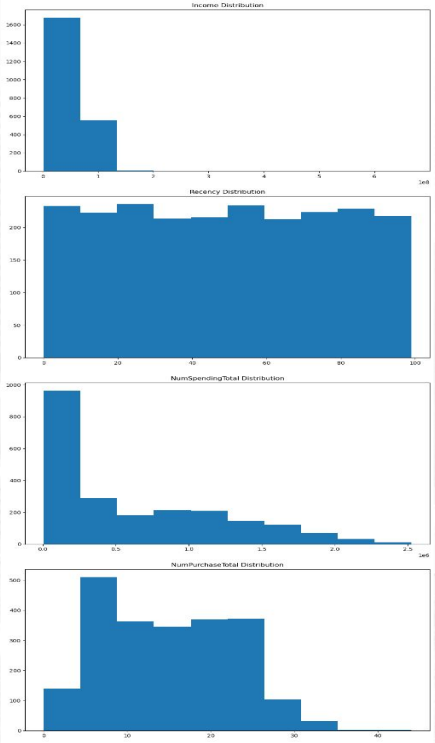
Feature engineering has been performed by calculating the conversion rate defined as (#response / #visit). Not only the conversion rate, but also other representative features such as,

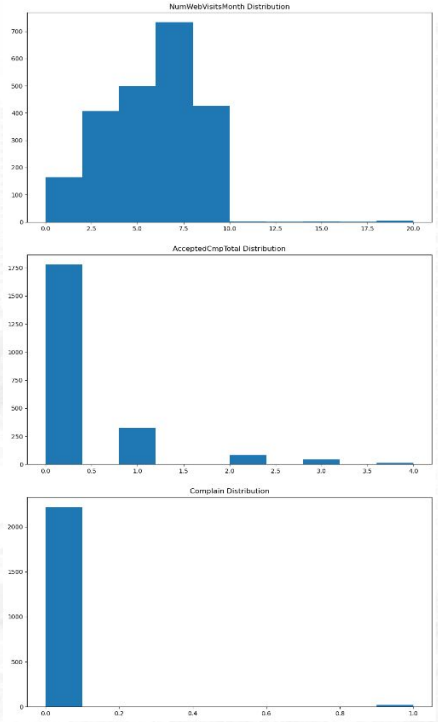
- Creating a new feature 'Age' from 'Year_Birth'
- Creating a new feature for age grouping
- Creating a new feature to sum up accepted campaigns
- Creating a new feature to sum up the number of children
- Creating a new feature to sum up total transactions
- Creating a new feature ‘Spending’ from some of products bought

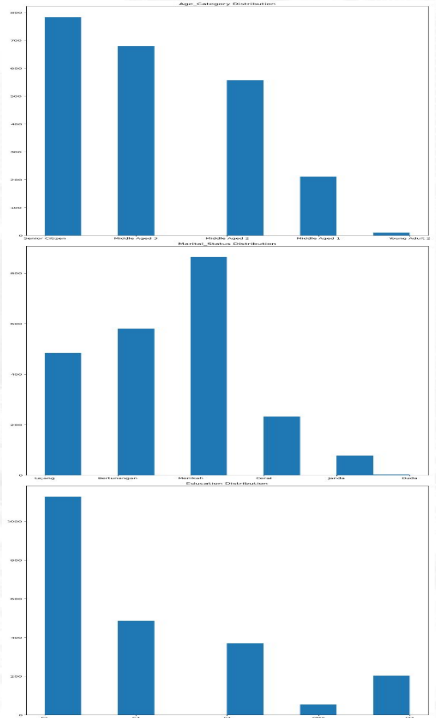
For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Univariate Analysis → most of the features are positively skewed except recency and total number of purchase which looked relatively uniform.



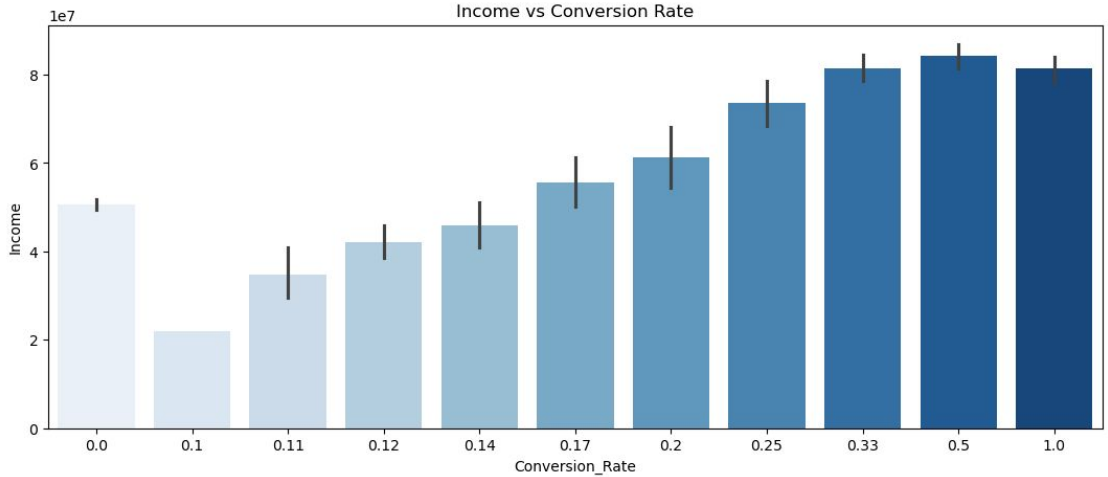




For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis (Income and Conversion Rate)

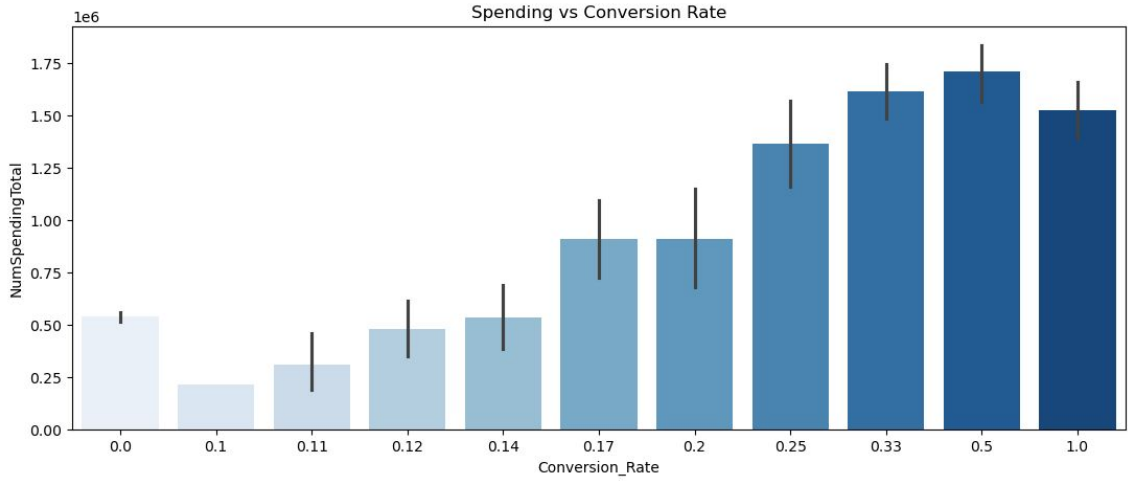


There's a slight positive trend, with income generally increasing as conversion rate increases. This implies that the higher the income, higher conversion rate is observed.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis (Spending and Conversion Rate)

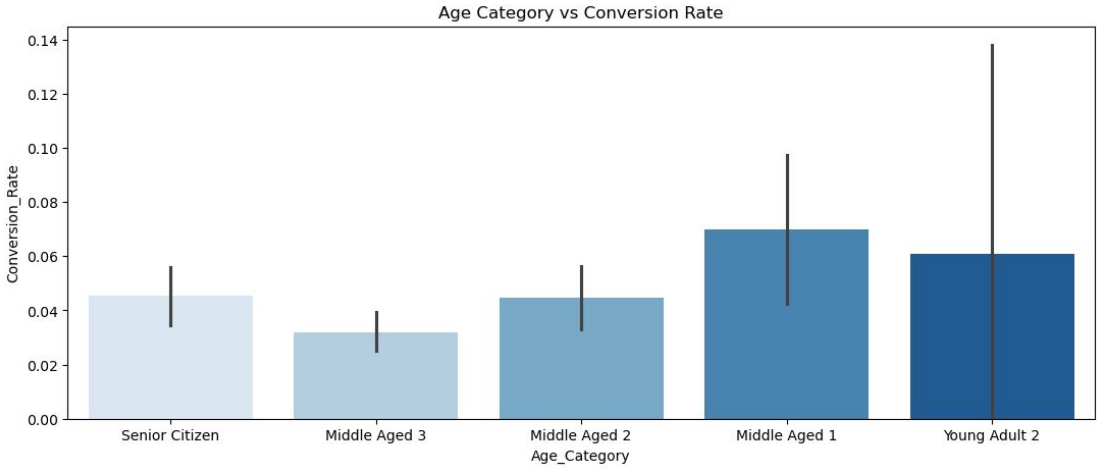


As the conversion rate increases from 0 to 1.0, there's a general trend of higher total spending, with the highest spending levels observed at conversion rates between 0.33 and 1.0. The relationship suggests that customers who are more likely to convert (make a purchase) tend to spend more overall, though there's some variability in spending amounts within each conversion rate category.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis (Age and Conversion Rate)



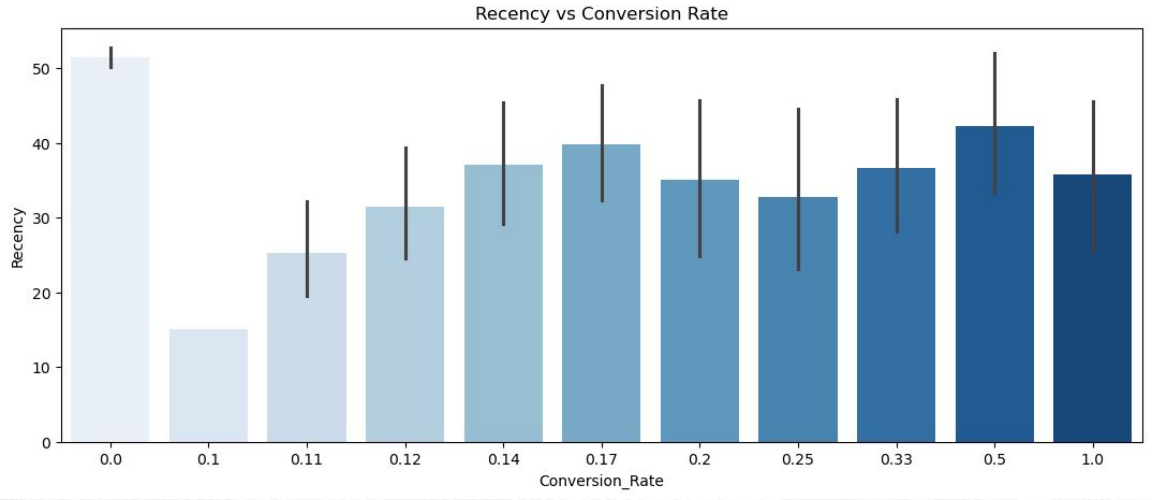
The middle-aged groups, particularly Middle Aged 1 (31-39), have the highest conversion rates, while Senior Citizen (above 60) and Middle Aged 3 (50-59) have the lowest. The younger customers are, the higher the conversion rate.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

EDA: Bivariate Analysis

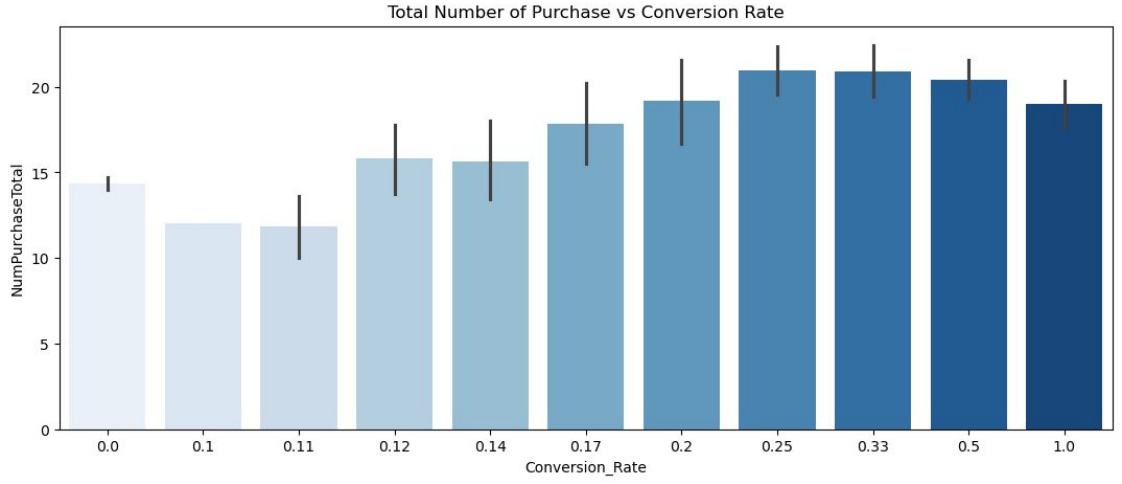


In terms of recency, trend fluctuated across conversion rate. The sweet spot for higher conversion rates seems to be customers with moderate Recency values. This could imply that nurturing relationships over time, rather than focusing solely on the most recent or oldest customers, might be beneficial for improving conversion rates.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis

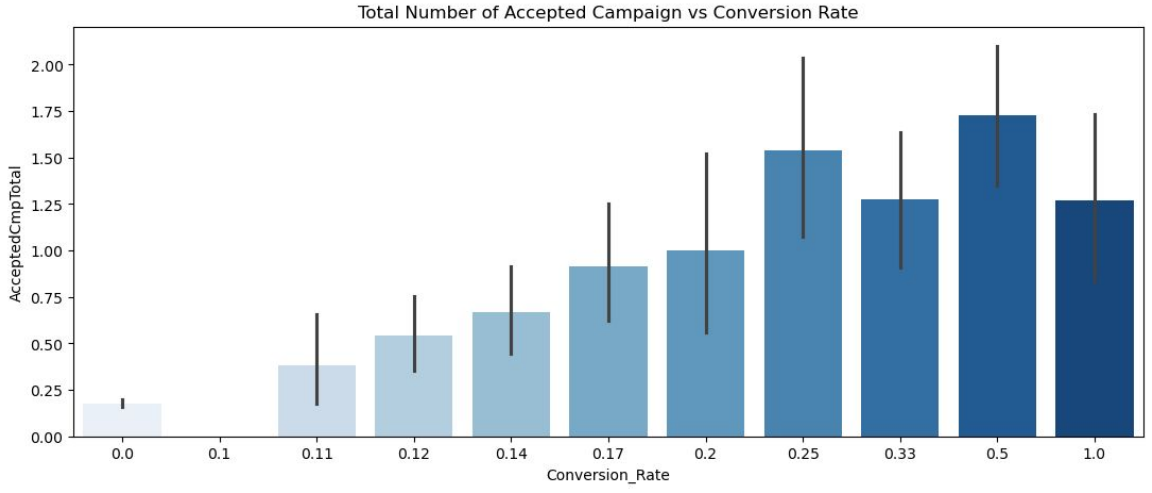


In terms of total purchase, a plateau or slightly decrease for very high conversion rates (0.33 to 1.0) is observed despite steadily increasing. This could suggest that beyond a certain point, higher conversion rates don't necessarily lead to more total purchases. Maximum purchase is between 20-21 and minimum purchase is about 12.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis



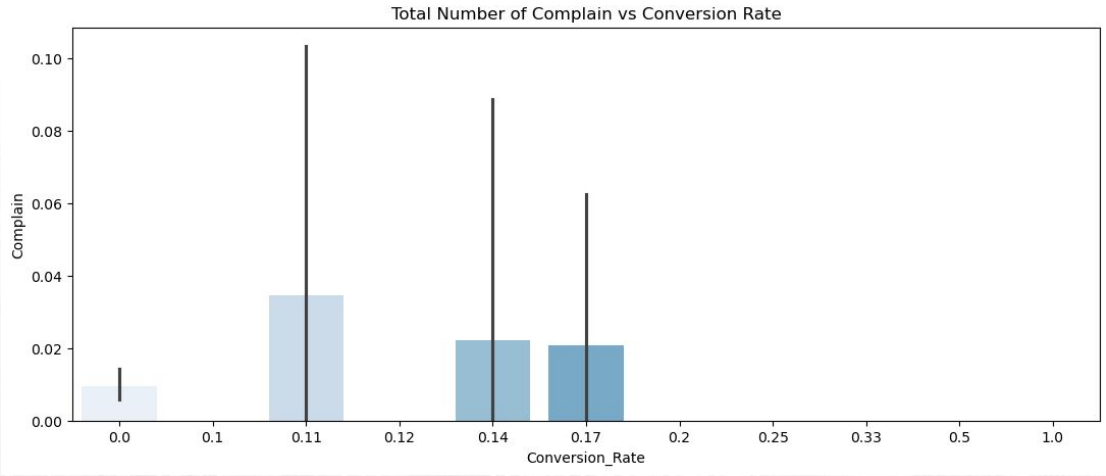
Similar to total purchase, total campaign increased steadily and fluctuate from 0.25 to 1. However, it can be concluded that higher accepted campaigns associated with higher conversion rate.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

EDA: Bivariate Analysis



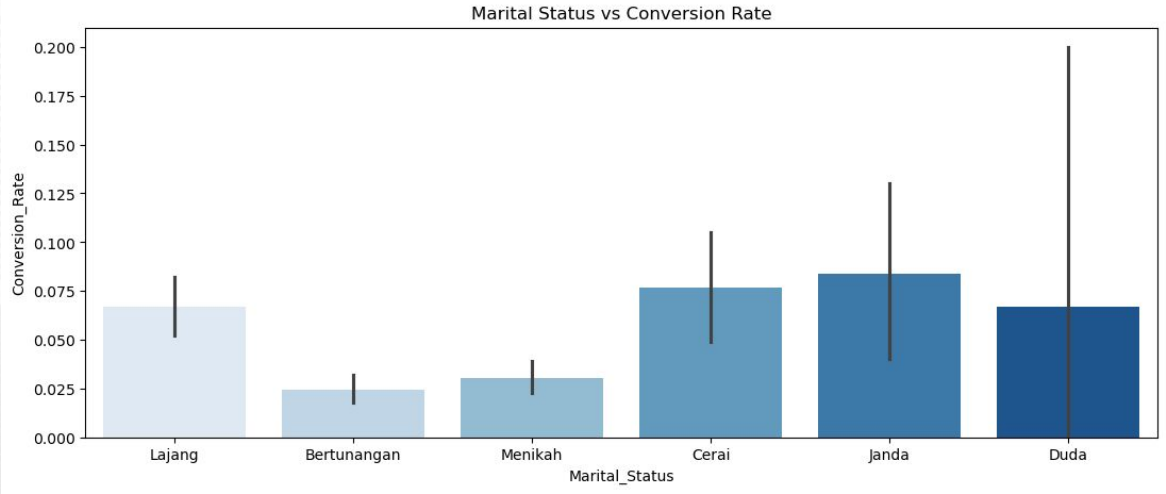
There is a slight negative correlation between complain and conversion rate, suggesting that campaigns with higher conversion rates tend to have fewer complaints.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

EDA: Bivariate Analysis

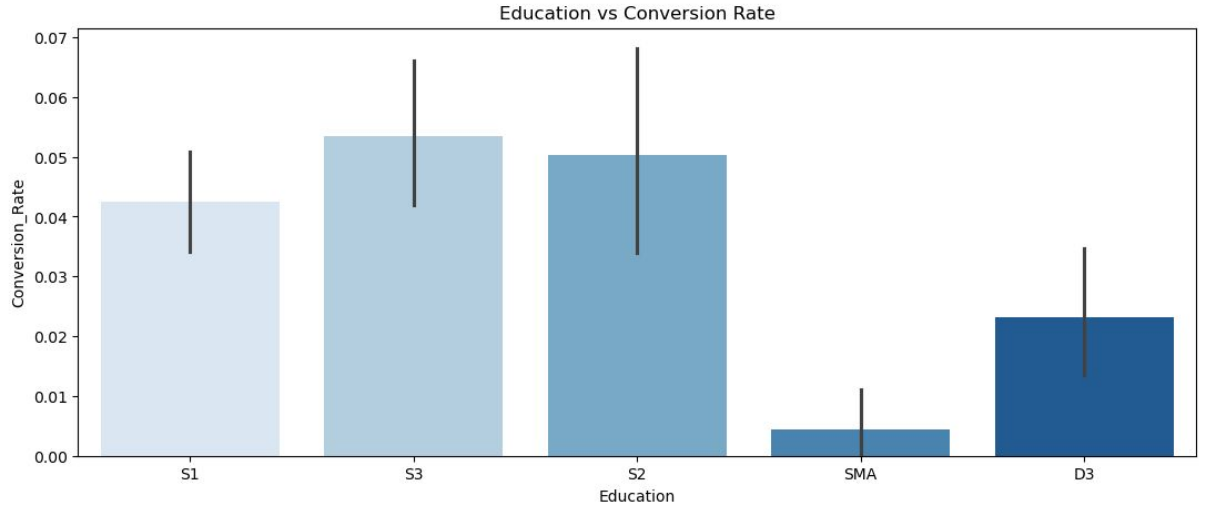


In terms of relationship between marital status and conversion rate. The highest conversion rates are observed among those who are married and widowed, while the lowest rates are seen among the single and divorced individuals. However, there is significant overlap among the groups, suggesting that marital status alone may not be the sole determinant of conversion rate.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

EDA: Bivariate Analysis



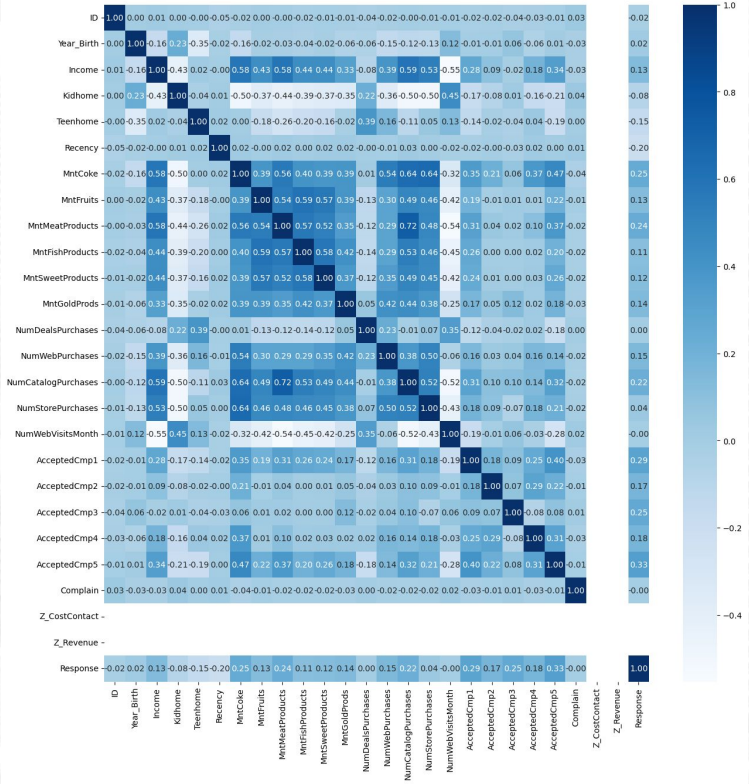
In terms of relationship between education level and conversion rate. The highest conversion rates are observed among individuals with S1 and S3 education levels, while the lowest rates are seen among those with SMA and D3 education levels. Alas, those who are having tertiary education are more likely to convert.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

EDA: Multivariate Analysis



Strong positive correlations are observed among different product spending categories (e.g., MntCoke, MntFruits, MntMeatProducts), suggesting that customers who spend more in one category tend to spend more in others as well. There are also notable correlations between income and various spending categories, and between different types of purchases (web, catalog, store), indicating consistent purchasing behavior across channels.

For more information, see jupyter notebook here

Conversion Rate Analysis Based on Income, Spending and Age

Rakamin Academy

Insights

- In terms of income, higher income levels are associated with higher conversion rates, indicating that wealthier customers are more likely to make purchases.
- In terms of spending, higher conversion rates are associated with increased total spending, indicating that customers who convert more frequently are also likely to be higher-value customers in terms of overall expenditure.
- In terms of age, younger to middle-aged adults tend to have higher conversion rates compared to older age groups, suggesting that marketing strategies may need to be tailored differently for various age demographics to maximise conversions.
- The heatmap reveals strong interconnections between different types of customer spending, with positive correlations across product categories and purchase channels, suggesting that high-value customers tend to engage more across all aspects of the business.

For more information, see jupyter notebook here