

Report of analysis on Market Customer's personal information

Preprocessing and Privacy of the dataset:

Working with sensitive customer information from the company database required implementing robust privacy protection measures. Multiple data sanitization techniques were applied to prevent security breaches while maintaining analytical value like (deleted, Masking, Binned, tekonzied) some of the data.

The following privacy techniques were systematically applied:

Column name	Privacy technique	Details of tech usage
Credit card Number	Deleteed	Sensitive data of the user and not important to analysis the data
credit_card_security_code	Deleted	Sensitive data of the user and not important to analysis the data
credit_card_expire	Deleted	Sensitive data of the user and not important to analysis the data
Employer	Deleted	Sensitive data of the user and not important to analysis the data
Residence	Deleted	sensitive data about the customer not used in the analysis
Current Location	Deleted	sensitive data about the customer not used in the analysis
Date registered	Deleted	Not used in the analysis process
Untitled column	Deleted	Default created column not used
Customer_id		It's a unique id for each customer so it's necessary to be there
UserName	Hided	It's sensitive and not used in the analysis but it may used in the future
name	Edited (Data Masked)	It's sensitive data so Masked the data to be initial name (first letter of each word)
Address	Edited (Data Masked)	It's a sensitive data so Masked data by check out just the state and post code of the address
Birthdate	Hided	It's a sensitive data and there is age column we can use it because some time Birthdate used as an Unique data
Job	Binned	I just category this column's data as (Technology, Health care, Education, Finance
Age	Binned	It's a sensitive data could classified the customer through the other information the so It's better to Binned the data as a range (18-25, 26-40, 41+) and useful for analysis process too (old age getting the Highest income (salary))
Salary	Binned	It's a sensitive data too, so after I know the Min, Max, and average of the data I could categorize it based on a range like (Low Income (salary <50k) Medium Income

		(50k<salary<150k), and High Income (Salary>150K)
Credit_card_provider	No change	I just calculated the most used types of card like (Visa, Master , and American cards)

This analysis focuses on understanding our market customer base through demographic and behavioral patterns. The goal is to identify target customer segments based on age, income, gender, location characteristics, and occupation to inform product development and marketing strategies.

Analysis Structure

Univariate Analysis Questions:

- Customer age distribution patterns
- Salary distribution across customer base
- Credit card provider preferences
- Gender composition of customer base
- Most common occupational categories

Bivariate Analysis Questions:

- Relationship between customer age and income levels
- Highest-paying job categories identification
- Gender distribution across occupational sectors

Multivariate Analysis Question:

- Comprehensive salary analysis considering age, gender, job category, and employer factors

Insights Discovered

1. Age-Income Correlation

Analysis reveals a strong positive relationship between customer age and income levels. Our older customers demonstrate significantly higher earning potential, suggesting premium products should target mature demographics.

Suggestion: Focus on Advirtisement for old customer as the age target.

2. Gender Distribution in Workforce

Male customers comprise a larger portion of our customer base compared to female customers, indicating potential opportunities for targeted female customer acquisition strategies.

Suggestion: focusing on Females more and their need will increase the sales

3. Payment Method Preferences

Visa credit cards dominate customer payment preferences, representing the most utilized payment method across all demographic segments.

4. Industry Concentration

Technology and management sectors show the highest customer concentration, with males representing the majority in these high-income professional categories.

Suggestion: Focus on Product related to Technology or management.

Moreover, Technology sector partnerships could yield significant customer acquisition opportunities.