

United India Insurance Company

Business Understanding

Information about the Company

- United India Insurance Company is an Indian general insurance company, headquartered in Chennai, India. It was incorporated on 18 February 1938, and was nationalised in 1972.
- Previously it was a subsidiary of the General Insurance Corporation of India (GIC). But when GIC became a re-insurance company as per the IRDA Act 1999, its four primary insurance subsidiaries New India Assurance, United India Insurance, Oriental Insurance and National Insurance got autonomy.
- After nationalization, United India has 16385 nos. workforce spread across 2248 offices providing insurance cover to more than 10 million policyholders. The Company has a variety of insurance products to provide insurance cover from bullock carts to satellites ● It essentially runs as a Public Sector Undertaking(PSU).

Business Model

Cross-Selling Marketing Strategy:

Cross-selling is the action or practice of selling an additional product or service to an existing customer. In practice, businesses define cross-selling in many different ways. Elements that might influence the definition might include the size of the business, the industry sector it operates within and the financial motivations of those required to define the term.

Objective: of cross-selling can be either

1. To increase the income derived from the current clients or
 2. To protect the relationship with the client or clients. The approach to the process of crossselling can be varied.
- Unlike the acquiring of new business, cross-selling involves an element of risk that existing relationships with the client could be disrupted.
 - For that reason, it is important to ensure that the additional product or service being sold to the client or clients enhances the value the client or clients get from the organization. In practice, large businesses usually combine cross-selling and up-selling techniques to increase revenue.
 - For the vendor, the benefits are substantial. The most obvious example is an increase in revenue. There are also efficiency benefits in servicing one account rather than several.

- Most importantly, vendors that sell more services to a client are less likely to be displaced by a competitor. The more a client buys from a vendor, the higher the switching cost.

In the case of United India Insurance Company, there are two types of customers: Existing Customers and New Customers. Hence to maximise the benefits of cross-selling, existing customers are targeted.

For Existing Customers, three cases arise:

Case-1: When New Coverage is greater than the Current Coverage

- In this case, an existing customer who is already covered by one product insurance is offered an insurance coverage for another new product.
- The **current coverage** provided to a customer under a current product (1) may or (2) may not be equal to the coverage he/she receives under another new product
 - (1) Consider an example where a 30 year old customer's current product coverage is Rs. 10,00,000 and he is provided with Rs. 10,00,000 as coverage under another new product. The **New Coverage** for this customer is depicted as **Rs.20,00,000** i.e. the algebraic sum of the current product coverage and the coverage offered under another new product
 - (2) Consider an example where the same 30 year old customer's current product coverage is Rs. 10,00,000 and this time he is provided with Rs. 30,00,000 as coverage under another new product. The **New Coverage** for the same customer in this case is **Rs.40,00,000** i.e. the algebraic sum of the current product coverage and the coverage offered under another new product

Case-2: When New Coverage is equal to the Current Coverage:

- In this case, the validity or rather the lifetime period of a current product insurance under which a particular customer is covered is terminated or completed.
- Through the cross-selling strategy, the same customer, who is no longer covered under the current product insurance, is offered another new product insurance with a coverage equal to his/her expired or previous product coverage.
- Consider an example where the lifetime period of a particular customer's current product insurance ("ANS") with a current coverage of Rs. 10,00,000 has been completed or his current product insurance has expired.
- As part of cross-selling, the same customer is offered another new product insurance ("TLE") with a coverage equal to the previous expired product i.e. Rs 10,00,000. Hence, now the customer's **New Coverage** is Rs. 10,00,000

Case-3: When New Coverage is less than the Current Coverage:

- In this case, the validity or rather the lifetime period of a current product insurance under which a particular customer is covered is terminated or completed.

- Through the cross-selling strategy, the same customer, who is no longer covered under the current product insurance, is offered another new product insurance with a coverage not equal to his/her expired or previous product coverage.
- Consider an example where the lifetime period of a particular customer's current product insurance ("INV") with a **current coverage** of Rs. 40,00,000 has been completed or his current product insurance has expired.
- As part of cross-selling, the same customer is offered another new product insurance ("END") with a coverage equal to Rs 20,00,000. Hence, now the customer's **New Coverage** is Rs. 20,00,000.

Types of Cross-Selling: Broadly speaking, cross-selling takes three forms:

- a) First, while servicing an account, the product or service provider may hear of an additional need, unrelated to the first, that the client has and offer to meet it.
- b) Selling **add-on services** - That happens when a supplier convinces a customer that it can enhance the value of its service by buying another from a different part of the supplier's company.
- c) The third kind of cross-selling can be called selling a **solution**. In this case, the customer buying air conditioners is sold a package of both the air conditioners and installation services.

Benefits of Cross-Selling:

- Cross selling builds customer loyalty
- Strengthens the current customer relationship with the firm involved
- Increased sales revenue
- Improves customer and client satisfaction even in B2B businesses where business is conducted between companies rather than between a company and individual consumers
- Increases Customer Lifetime Value (CLV)

Customer Lifetime Value:

- The lifetime value of a customer, or customer lifetime value (CLV), represents the total amount of money a customer is expected to spend in a business, or on its products, during his/her lifetime.
- To calculate customer lifetime value you need to calculate average purchase value of a customer, and then multiply that number by the average purchase frequency rate to determine customer lifetime value.

Services:

- Motor Insurance
- Health Insurance
- Travel Insurance

- Personal Accident Insurance
- Householder's Insurance
- Shopkeeper's Insurance
- Fire,Marine,Industry,Liability,Micro and Credit Insurance

Data Understanding/Description

The data provided by the insurance company is spread across one dataset with 100000 records and 16 variables; the contents of which are:

- **Age:** Listing Customers' age ranging from 18 years to 60 years
- **Gender:** Listing whether the customer is a male or a female
- **Marital Status:** Listing whether a customer is married, divorced or single
- **Family Members:** Listing the number of members in the family of a customer i.e. ranging from 1 to 10 members
- **Education:** Listing the education background of customers
- **Occupation and Job Title:** Listing the occupations and job titles of customers
- **Income:** Listing the income of customers income in Lakhs of Rupees per Annum
- **Current Product:** Listing whether a customer has currently chosen a product insurance or not i.e. it indirectly implies whether a customer is new or already existing.
- **Current Product Type:** Listing the product type (code) of a current product
- **Current Coverage:** Listing the amount (in rupees) current product insurance covers under the current policy or scheme
- **New Product Type:** Listing the product type (code) of a new product
- **New Coverage:** - If a customer is already provided coverage for a current product and opts for a new product: **New Coverage** exhibits the cumulative total coverage(in rupees) for both the products chosen by a customer
- **Rating:** Listing the rating given by customers as Cold, Warm or Hot for the services provided by the insurance company
- **Converted:** Listing whether the cross-selling strategy of the insurance company has been successful or not i.e. whether the new and existing customers have converted or not
- **Status:** Listing whether the cross-selling strategy of the insurance company has been successful or not i.e. whether the new and existing customers have converted or not. If customers have not converted then various **codes** have been allocated by the company based on the insurance policy of a customer

Data Audit

Data Audit involves the following steps:

1. Formatting changes – Examining variables that need to be converted to an appropriate data type before analyzing them
2. Validation – Finding out whether the variables contain any missing values or not
3. Checking for Errors – Searching for Variables that contain any outliers or negative values
4. Subsetting – Removal of unnecessary or insignificant variables from the data set

Age:

- Missing Data Present - 44 missing records
- Ranging from 18 to 60
- No presence of any outliers
- Requires Data cleaning
- The Age Variable was converted from a double data type to a numeric data type
- Missing Values were imputed by considering the minimum value between the mean and median values which were calculated excluding the missing data
- In this case Median (39 years) was found to be lesser than Mean. Hence Median was used to impute missing values

Gender:

- Missing Data Present - 182 Missing Records
- Requires Data Cleaning
- The Gender Variable was converted to a character data type
- Females-39893 , Males-59925
- Missing Values were imputed with the mode(Male) calculated for the variable ●
After Imputation of missing values, datatype was converted to factor.

Marital Status:

- Missing Data Present - 42 Missing Records
- Requires Data Cleaning
- Data type: Character
- Divorced-23160 , Married-53950 , Single-22848
- Missing Values were imputed with the mode(Married) calculated for the variable
 - After Imputation of missing values, datatype was converted to factor.

Family Members:

- Missing Data Present - 22 Missing Records
- Ranging from 1 to 10
- Requires Data Cleaning
- No presence of any outliers

-
- The Age Variable was converted from a double data type to a numeric data type
Missing Values were imputed by considering the minimum value between the mean and median values which were calculated excluding the missing data
- Mean-4.659 , Median-4.0 (excluding Missing Values). Hence Median was used to impute missing values
- **Education:**
- Missing Data Present - 46 Missing Records
- Requires Data Cleaning
- Data type: Character
- BD-29586 LHS-13399 MD- 22543 NE-7232 PD-14746 UHS-12448 •
Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Occupation:

- Missing Data Present - 46 Missing Records
- Requires Data Cleaning
- Data type: Character
- SE-39443 SFT-8522 SPT- 51992
- Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Job Title:

- Missing Data Present - 123 Missing Records
- Requires Data Cleaning
- Data type: Character
- BA-7980 CB-5396 CF-2142 DD-9795 FH-23780 OC-1660 OM-5935
- OT-6530 PA-1609 PG-19623 PR-7666 RR-7761
- Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Income:

- Missing Data Present - 22 Missing Records
- Ranging from 1 to 10
- Requires Data Cleaning
- No presence of any outliers
- The Income Variable was converted from a double data type to a numeric data type
- Missing Values were imputed by considering the minimum value between the mean and median values which were calculated excluding the missing data

- Mean-4.96 , Median-3.0 (excluding Missing Values). Hence Median was used to impute missing values

Current Product:

- Missing Data Present - 18 Missing Records
- Requires Data Cleaning
- Data type: Character
- No-42862 Yes-57120
- Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Current Product Type:

- Missing Data Present - 33 Missing Records
- Requires Data Cleaning
- Data type: Character
- ANS-18461 END-5487 INV-10346 NA-42855 PMT-8477 TLE-14341 • Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Current Coverage:

- Missing Data Present - 49 Missing Records
- Ranging from 0 to 15000000
- Requires Data Cleaning
- Positive Outliers are Present
- The Current Coverage Variable was converted from a character data type to a numeric data type
- Missing Values were imputed by considering the minimum value between the mean and median values which were calculated excluding the missing data
- Mean-3633770 , Median-50000 (excluding Missing Values). Hence Median was used to impute missing values

New Product Type:

- Missing Data Present - 47 Missing Records
- Requires Data Cleaning
- Data type: Character
- ANS-29151 END-12718 INV-20429 PMT-20429 TLE-25067 • Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

-

New Coverage:

- Missing Data Present - 127 Missing Records
Ranging from 1000000 to 15000000
- Requires Data Cleaning
- No Outliers Present
- The New Coverage Variable was converted from a character data type to a numeric data type
- Missing Values were imputed by considering the minimum value between the mean and median values which were calculated excluding the missing data
- Mean-6106250 , Median-3000000 (excluding Missing Values). Hence Median was used to impute missing values

Rating:

- Missing Data Present - 44 Missing Records
- Requires Data Cleaning
- Data type: Character
- Cold-47551 Hot-22988 Warm-29417
- Missing Values were imputed with the mode calculated for the variable • After Imputation of missing values, datatype was converted to factor.

Converted:

- Missing Data Present - 49 Missing Records
- Data type: Character
- Converted- 38296 Not Converted -61655

Status:

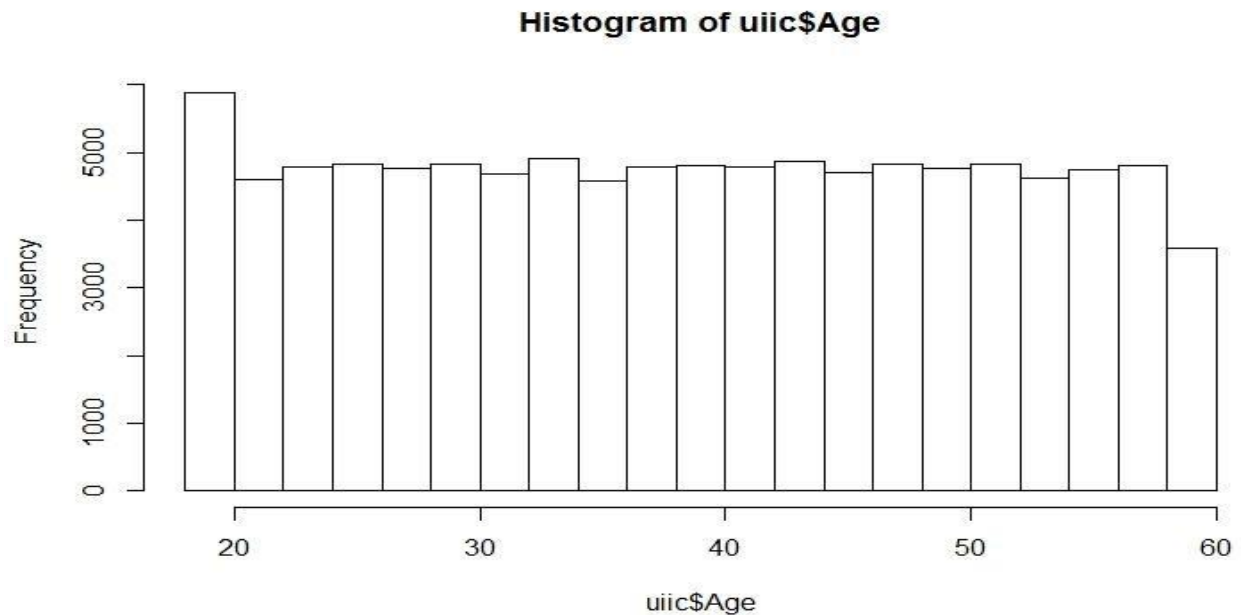
- No Missing Data Present
- To convert this variable to a target variable convenient for data modelling, all the data points written as codes are replaced with “Not Converted”
- The new variable called “**Status_new**” can now be used as a target variable for application of predictive modelling as it is a categorical variable with no missing values.
- **Status_new** was converted from a character data type to a factor data type.

Data Exploration

Age:

- Minimum Age - 18 years , Maximum Age - 60 years
- Average Age of Customers is 39 years
- Maximum no. of Customers (2480) fall in the category of 33 years and 51 years
- Customers who are either 60 years of age (1189) or 18 years of age (1177) are the least when compared to customers of other ages
- All other age categories have about 2300-2400 customers each

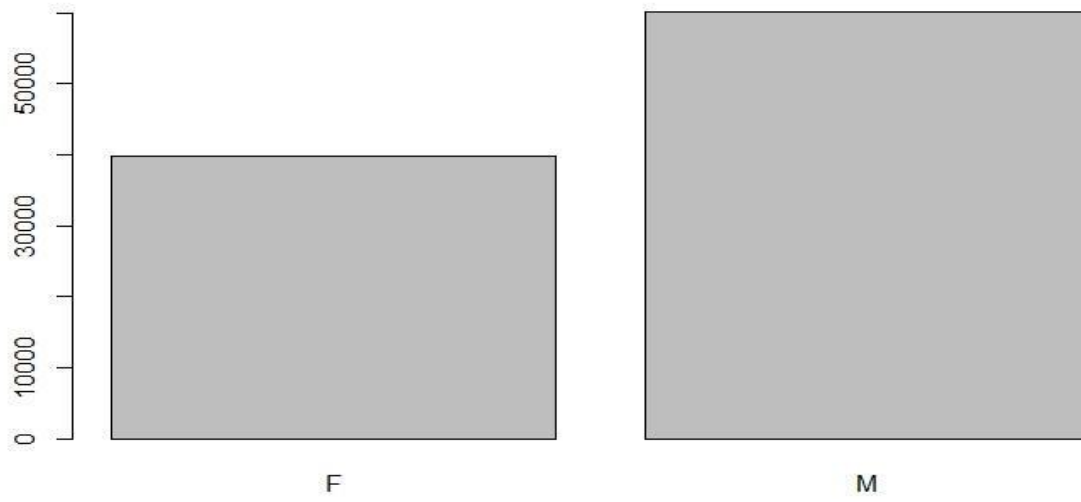
Histogram showing Number of Customers (Y-Axis) vs Age of Customers (XAxis)



Gender:

- Out of 1,00,000 customers- 39,893 are Females and 60,107 are Males
- Among Females, those who are aged 51 years (1004) are maximum in number when compared to other age groups.
- Among Males, those who are aged 50 years (1512) are maximum in number when compared to other age groups

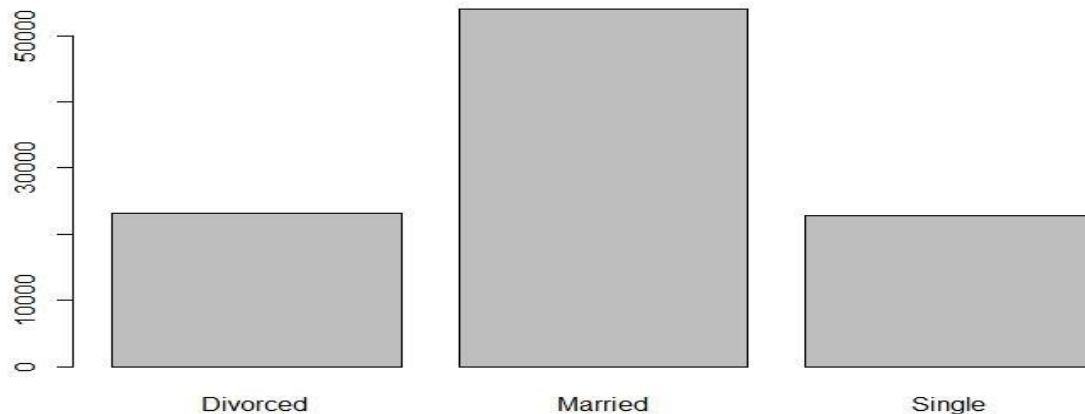
Bar Plot showing Number of Customers (Y-Axis) vs Gender of Customers (X-Axis)



Marital Status:

- Out of 1,00,000 customers: Divorced-23160 Married-53992 Single-22848
- 21,519 Females and 32473 Males are Married
- 9075 Females and 13773 Males are Single
- A Total of 5,602 Customers who belong to category of 19-21 years are Single
- None of the Customers who belong to the age group 18-21 years are Divorced

Bar Plot showing Number of Customers (Y-Axis) vs Marital Status of Customers (X-Axis)



Family Members:

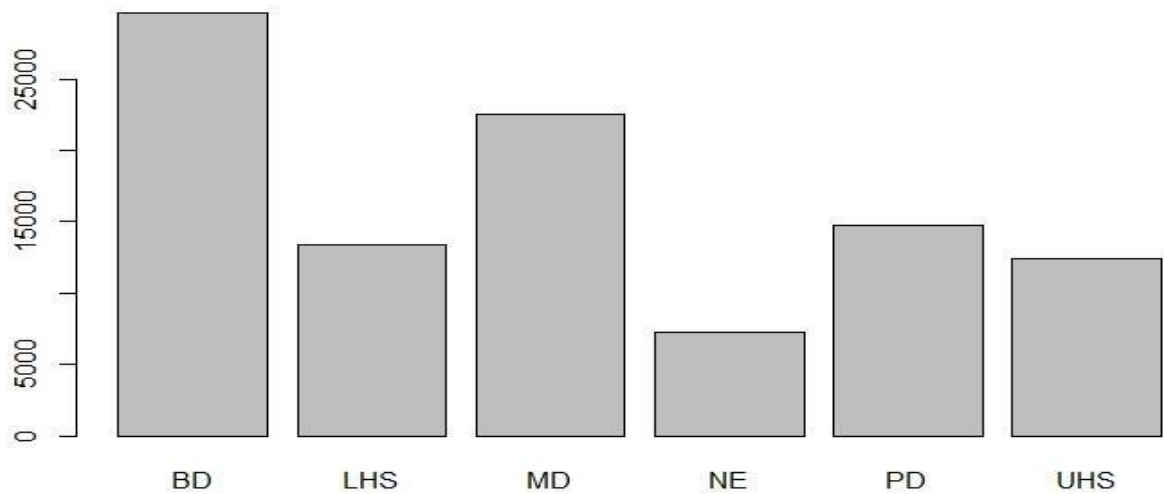
- Out of 1,00,000 customers: 1 Member-8654 2 Members-19011
- 3 Members-11795 4 Members-13322 5 Members-11734 6 Members-10376
- 7 Members-7414 8 Members- 5850 9 Members-7487 10 Members- 4357
- 118 Customers who are aged 34 have 10 members each in their families
- 373 Customers who are aged 21 have only one member each in their families
- Most Customers have at least 2 members each in their families
- 11929 Customers who are Married have only 2 members each in their families
- Only 2 Customers who are Married have only one other member in their families

Education:

- Out of 1 Lakh Customers:
- Bachelor's Degree-29586 LHS-13399 Master's Degree-22543 NE-7232
Professional Degree-14746 UHS-12448
- 1207 Customers aged 19 years are not enrolled in any academic program
- Only 2 twenty year olds and 1 eighteen year old have a bachelor's degree
- None of the Customers belonging to 18-20 years age group have either a master's degree or a professional degree
- Among Males, 17863 have a Bachelor's Degree and 13592 have a Master's Degree and 4309 Males are not enrolled or don't have a degree.
- Among Females, 11769 have a Bachelor's Degree and 8951 have a Master's Degree

- Among Married Customers, 16868 have a Bachelor's Degree and 12653 have a Master's Degree
- 905 Divorced customers are not enrolled in any academic program or do not have any educational qualification

Bar Plot showing Number of Customers (Y-Axis) vs Education Background of Customers (X-Axis)



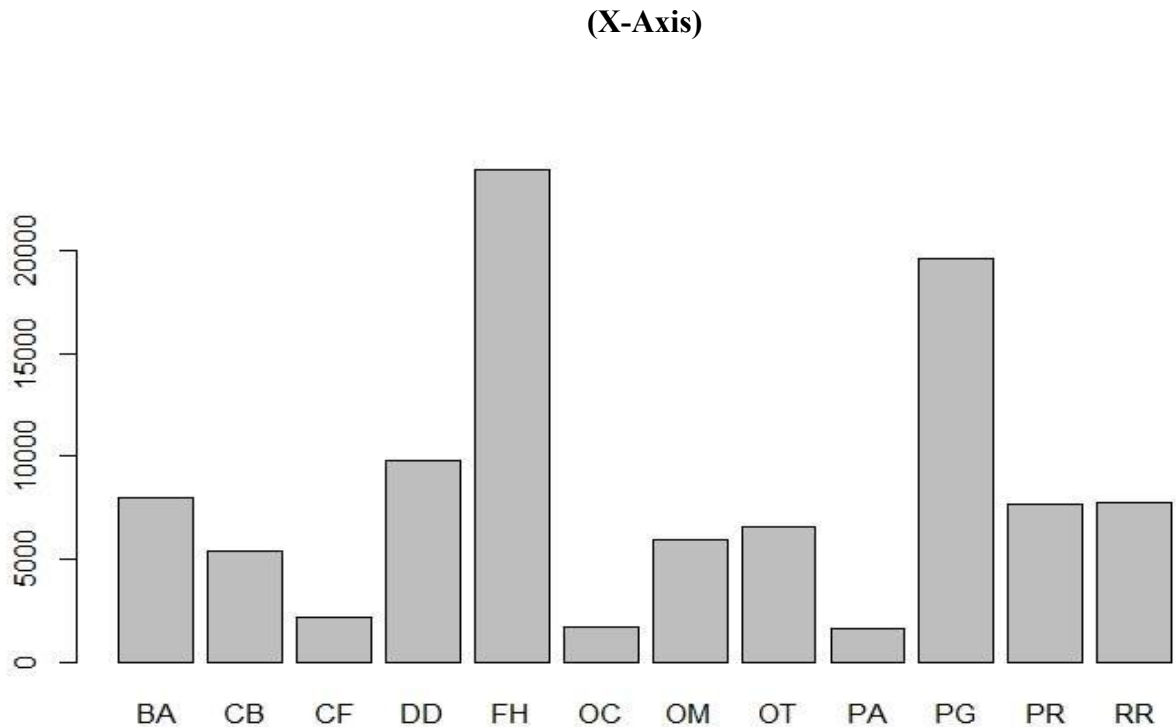
Occupation:

- Out of 1 Lakh Customers: SE-39443 , SFT-8522 and SPT-52035
- Among Females, SE-15731, SFT-3458 and SPT-20704
- Among Males, SE-23712 , SFT-5064 and SPT-31331
- Among Divorced Customers: SE-9483 SFT-1665 SPT-12012
- Among Married Customers: SE-21887 SFT-4228 SPT-27877
- Among Single Customers: SE-8073 SFT-2629 SPT-12146
- Among Customers who have a Bachelor's Degree: SE-12310 SFT-5 SPT-17317
- Among Customers who have their Education as LHS, only 4 have their Occupation as SPT
- Among Customers who have their Education as an MD,NE,PD or UHS, none have their Occupation as SFT
- Only 4 Customers who have their qualification as LHS, have their occupation as SPT

Job Title:

- Out of 1,00,000 Customers:
- BA-7980 CB-5396 CF-2142 DD-9795 FH-23903 OC-1660 OM-5935
- OT-6530 PA-1609 PG-19623 PR-7666 RR-7761
- Among the Female Customers, most work as a FH (9517) and 7762 work as a PG
- Among the Male Customers, most work as a FH (14386) and 11861 work as a PG
 - Among the Customers who have a Bachelor's Degree: 10,629 work as a FH, 3 work as an OT, 2 work as a CB, 1 works as an OC while none of them work as a CF,OM or as a PA
- Among Customers who have their Qualification as LHS: 3948 of them work as a DD while none of them work as a BA,PG,PR or as a RR
- Among the Customers who have a Master's Degree, 8,052 of them work as a FH while none of them work as CB,CF,DD,OC,OM,OT or as a PA
- Among the Customers who have a Professional Degree, 5196 work as a FH while none of them work as a BA,PG,PR or as a RR
- Among the Customers who have their occupation as SE; 23831 work as a FH, while none of them work as a BA,OC,PA,PR or RR
- Among the Customers who have their occupation as SFT; 2626 work as a DD, while none of them work as a BA,CF,PG,PR or RR
- Among the Customers who have their occupation as SPT; Most work as a PG, while none of them work as a CF

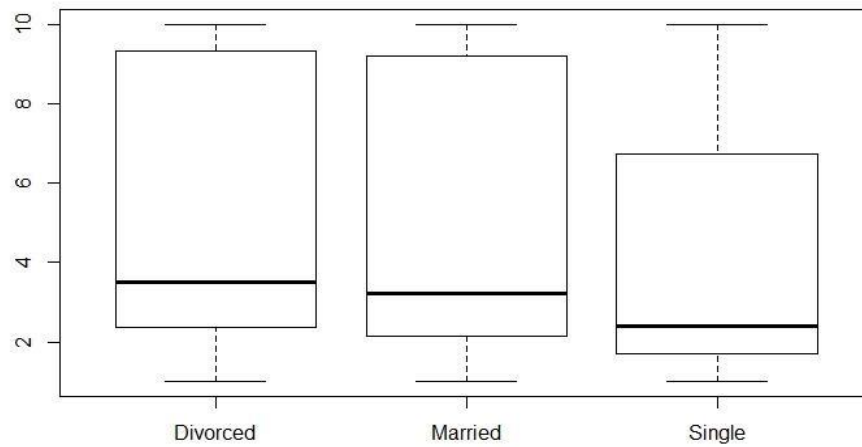
Bar Plot showing Number of Customers (Y-Axis) vs Job Title of Customers



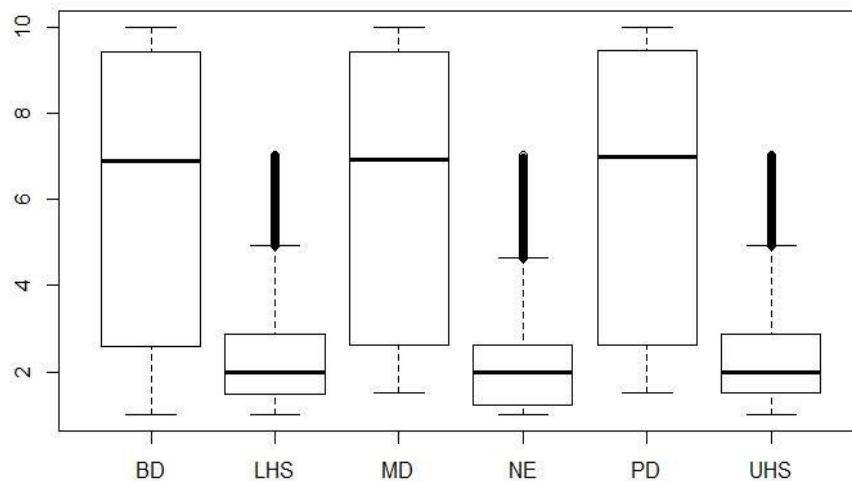
Income:

- Minimum Income of any Customer is One Lakh Rupees per annum; while Maximum is Ten Lakh Rupees per annum.
- Average Income is nearly 5 Lakhs Per Annum among all Customers
- Average income for Customers aged 18 years is 1.83 Lakhs per Annum and people aged 34 years and above earned at least an average of 5 Lakhs Per Annum
- Males have a higher income than Females in general but the average income for both groups is nearly equal (4.9 Lakhs Per Annum)
- Although Married Customers earn more than Divorced Customers in general, their average salary is slightly less than the divorced customers.
- The average income of Customers with a Bachelor's, Master's or Professional degree is nearly equal i.e. approximately 6.2 Lakhs per Annum; Customers with other educational qualifications earn nearly 2.35 Lakhs Per Annum
- Average Income of customers whose occupation is SPT is 6.89 Lakhs per annum; this is higher than than the average income of customers with other occupations

Box Plot showing Income of Customers (Y-Axis) vs Marital Status of Customers (X-Axis)



Bar Plot showing Income of Customers (Y-Axis) vs Educational Qualification of Customers (X-Axis)



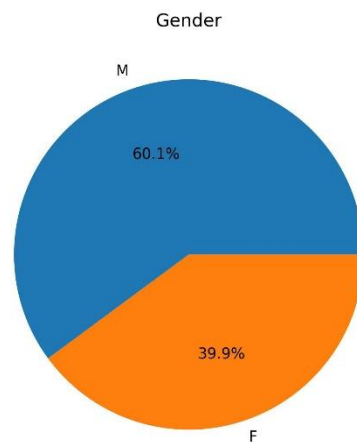
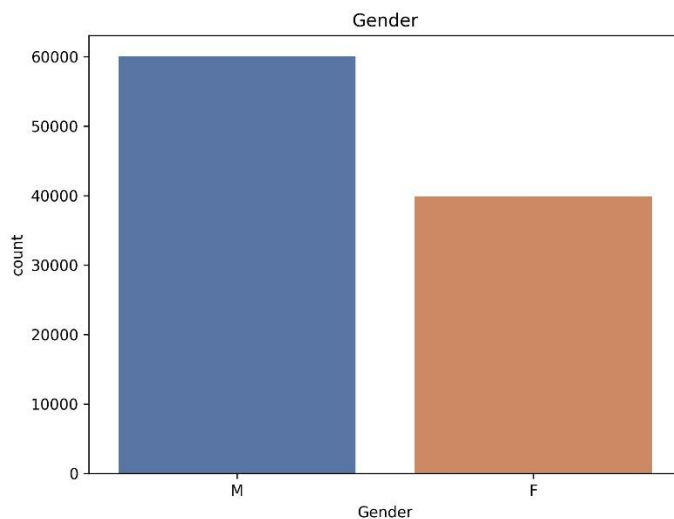
Current Product:

- Out of 1 Lakh customers, 57,129 are existing customers i.e. who are already covered through a product insurance and 42,871 are new customers i.e. who are not covered through any kind of product insurance.
- Out of all the existing customers, 17174 are females and 25697 are males

- Out of all the new customers, 22719 are females and 34410 are males
- Out of all the existing customers: 32437 are married, 9253 are single and the rest divorced
- Out of all the new customers: 21555 are married, 13595 are single and the rest divorced
- Among both existing and new customers, most customers have a bachelor's degree as their academic qualification
- For Existing Customers, both cumulative income (Algebraic sum of incomes of all existing customers) and average income are greater than those of their new counterparts
- Average income of existing customers is 5.2 Lakhs per Annum and for new customers it is 4.64 Lakhs per Annum

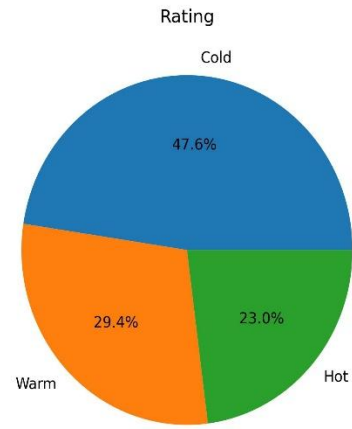
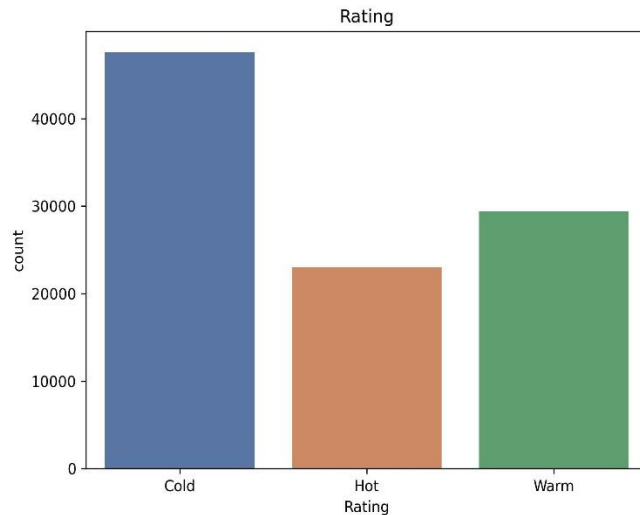
Data Exploration:

Gender:



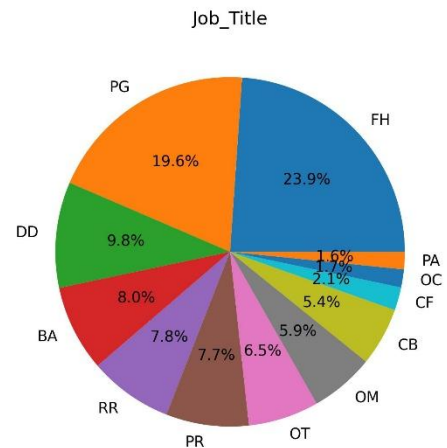
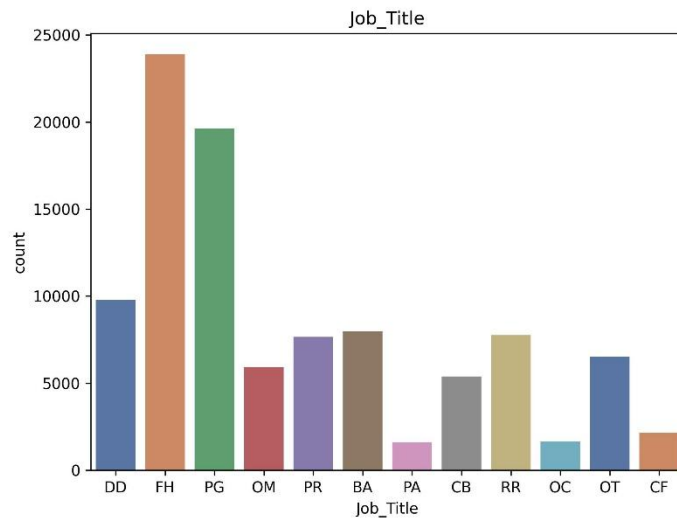
- Male (M) count is higher than Female (F).
- Male 60.1% and Female 39.9%
- Total Male = 60092 and Total Female = 39875
- There are 27,840 males and 19,100 females who are 20 years old and single

Rating:



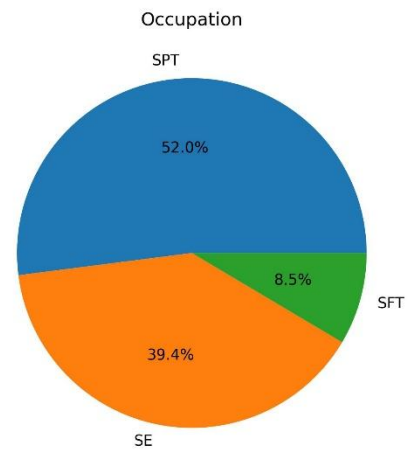
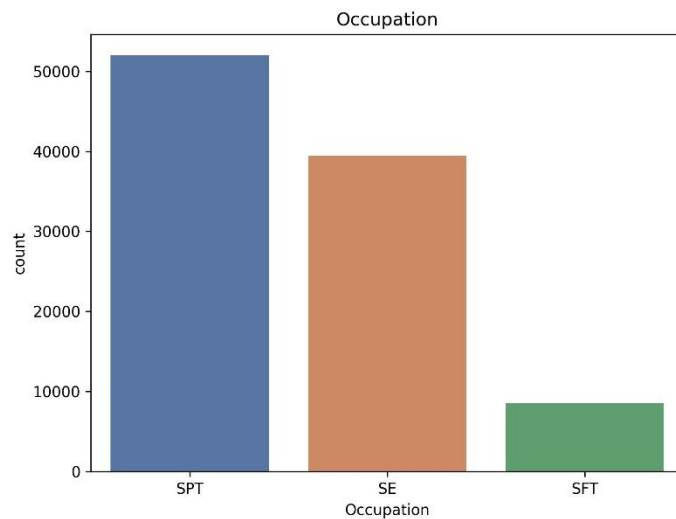
- There are 27,840 males and 19,100 females who are 20 years old and single.
- Cold has the highest number of entries, accounting for 47.6% of the total data.
- This means that almost half of the dataset falls under the Cold rating.
- Warm represents 29.4% of the total.
- This shows that nearly one-third of the customers have a medium engagement or conversion potential.
- **Hot** is the **smallest category**, making up only **23.0%** of the dataset.
- that there are 47,576 Cold, 29,401 Warm, and 22,974 Hot ratings.

Job Title:



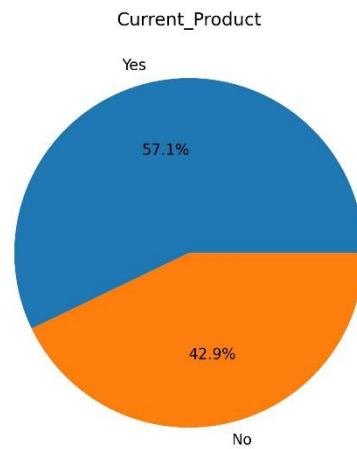
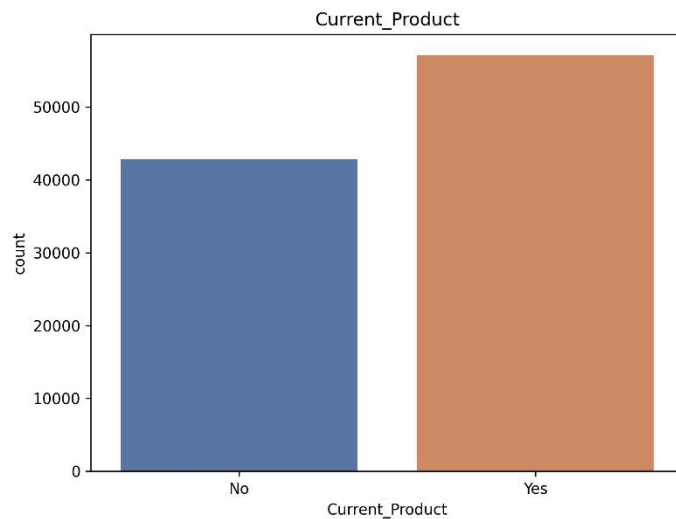
- The job titles and their respective percentages are shown in the graph
- The most frequent job title is FH (23.9%), followed by PG (19.6%).
- These two categories together account for **over 43%** of the total dataset.
- This indicates that the majority of individuals belong to these two job types.
- DD (9.8%), BA (8.0%), RR (7.8%), and PR (7.7%) have moderate representation.
- These job titles together form another **~33%** of the total.
- **OM (6.5%)** and **CB (5.4%)** represent smaller groups.
- CF (2.1%), OC (1.7%), and PA (1.6%) are the least represented categories, contributing very little to the total count.
- the number of job title are FH – 23,893, PG – 19,614, DD – 9,789, BA – 7,977, RR – 7,758, PR – 7,655, OT – 6,529, OM – 5,932, CB – 5,395, CF – 2,141, OC – 1,659, and PA – 1,609.

Occupation:



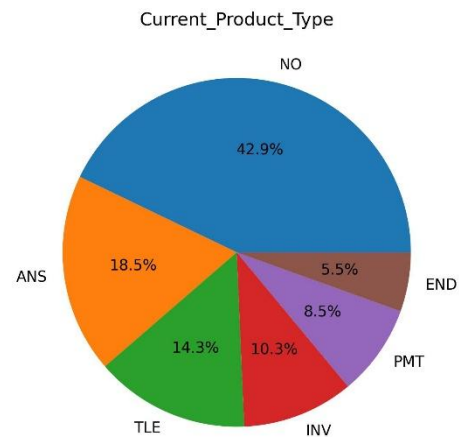
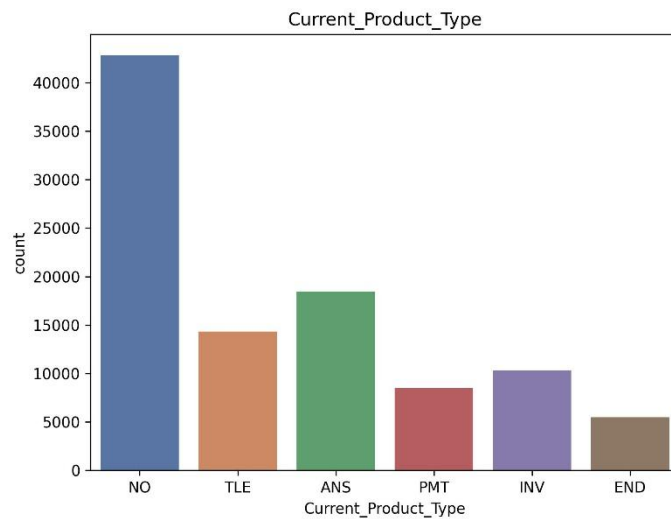
- The graph shows the number of occupations and their respective percentages.
- There are three occupation types SPT (52.0%), SE (39.4%), and SFT (8.5%).
- the bar chart, it is clear that SPT has the highest number of people, followed by SE, while SFT has the lowest count.
- the occupation distribution suggests that SPT and SE are the major occupation types in this dataset.
- The number of Occupation are SPT – 52,004, SE – 39,426, and SFT – 8,521 available in this data set

Current Product:



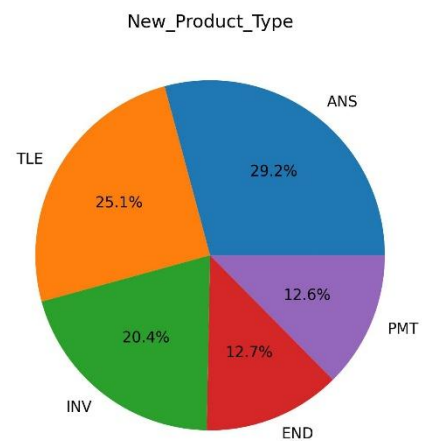
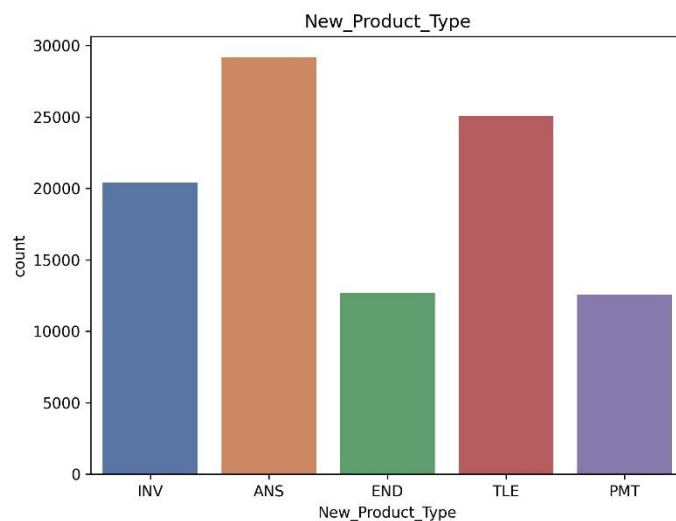
- The variable Current Product represents whether customers currently hold a product or not
- There are two categories Yes (57.1%) and No (42.9%).
- The pie chart confirms that 57.1% of customers already own a product, whereas 42.9% do not.
- This indicates that a majority of customers are existing product holders, showing good retention or existing engagement.
- However, the 42.9% “No” group represents a potential target audience for marketing or new product offerings, as they currently do not have any product.
- The total number of s 99,951 (Yes – 57,104 and No – 42,847).

Current Product Type:



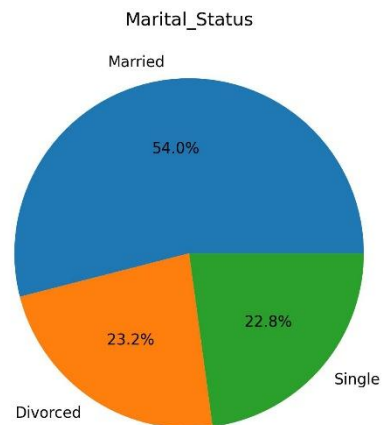
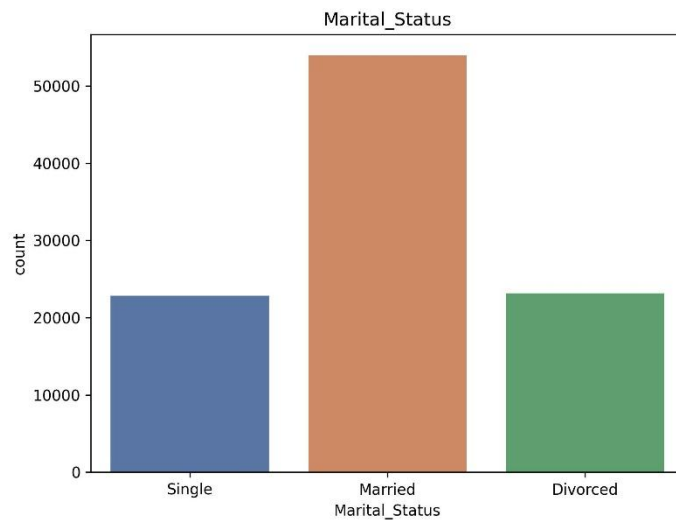
- The variable Current Product Type represents the types of products currently held by customers.
- There are six product types NO (42.9%), ANS (18.5%), TLE (14.3%), INV (8.5%), PMT (10.3%), and END (5.5%).
- “NO” has the highest count, showing that nearly 43% of customers don’t currently hold any product.
- ANS (18.5%) and TLE (14.3%) are the next most common product types, indicating customer preference for these two categories.
- ANS and TLE segments can be considered strong performing product lines with high engagement potential.
- The total number of 99,951 (NO – 42,847, ANS – 18,468, TLE – 14,336, INV – 10,343, PMT – 8,472, and END – 5,485).

New Product Type :



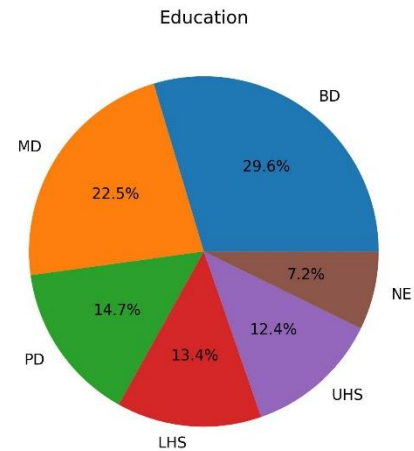
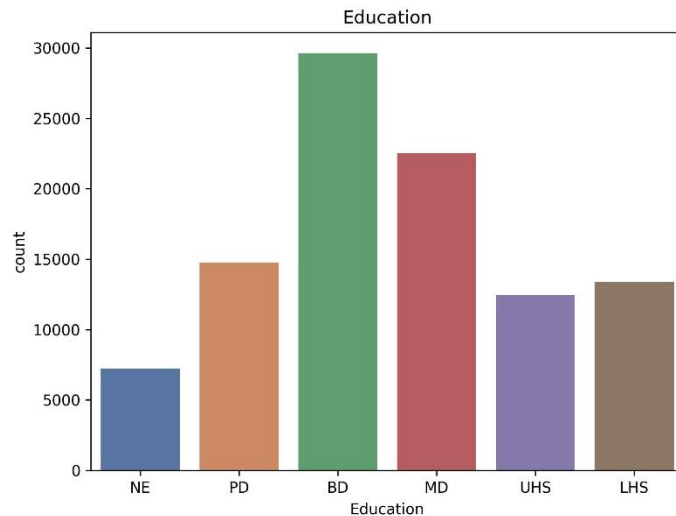
- The variable New Product Type shows the types of products that customers are interested in or have recently taken.
- There are five product categories shown NS (29.2%) is the most preferred, followed by TLE (25.1%) and INV (20.4%), while END (12.7%) and PMT (12.6%) are the least preferred.
- we can see ANS and TLE together cover over 54% of all selections — indicating that most customers are drawn toward these two product types.
- INV has a reasonable share, suggesting customers are also interested in investment products.
- The total number of **99,951** (ANS – 29,181, TLE – 25,060, INV – 20,420, END – 12,710, and PMT – 12,580).

Marital Status:



- The variable Marital_Status shows the distribution of customers based on their marital status — Single, Married, and Divorced
- Married individuals form the majority group (54%), indicating that more than half of the customers are married.
- Single (22.8%) and Divorced (23.2%) individuals together make up the remaining 46%.
- The total number of is 99,951 (Married – 53,966, Divorced – 23,151, and Single – 22,834).

Education:



- The Education variable shows the distribution of customers across different education levels
- **The (BD)** group forms the **largest portion (29.6%)**, indicating that most customers are well-educated with a qualification.
- The (MD) group follows with 22.5%, showing a significant number of highly educated individuals
- (PD) and Lower High School (LHS) have moderate representation (around 14–13%)
- **(NE)** group is the smallest, making up only **7.2%** of the total.
- The total number of is 99,951 (BD – 29,616, MD – 22,530, PD – 14,739, LHS – 13,395, UHS – 12,441, and NE – 7,230)