**Great Customer**

**Introduction:**

We have to work on three datasets in this project.

1. Great Customer Class
2. Mobile Price Range
3. Heart Disease

**Great Customer Class:**

The dataset contains information on individuals from a variety of backgrounds, including their age, workclass, education level, marital status, occupation, race, and sex. Additionally, it includes details about their lifestyle, such as the amount of time they spend drinking beer and exercising each year, as well as their tea and coffee consumption. The dataset also includes information on the number of hours they work each week and their salary. The objective is to find the relevant features to predict the great\_customer\_class.

**Mobile Price Range:**

The dataset contains mobile phone information, including hardware specifications like battery power, clock speed, memory, and software features like connectivity options. The objective is to identify the key features that can effectively predict the price category of mobile phones.

**Heart Disease:**

This dataset contains information about several risk factors, such as age, gender, smoking habits, blood pressure, diabetes, and cholesterol levels, for a group of individuals. The target column, TenYearCHD, indicates whether each individual developed coronary heart disease within ten years.

Our initial step involved addressing any null values present in the datasets. If any null values were found, we employed suitable methods to handle them appropriately.

Subsequently, we examined the categorical features to determine if they contained any string values. If such values were identified, we applied relevant encoding methods to convert the strings into numerical representations.

Next, we addressed the presence of outliers, if any, by employing appropriate methods to handle them effectively.

We then utilized hypothesis testing to identify the relevant independent features for predicting the target feature of each dataset. Since all target features were categorical, we employed the ANOVA (Analysis of Variance) test to calculate p-values between numerical and categorical features. Furthermore, for the association test, we utilized the Chi-square test of independence to determine the p-values.

Lastly, we selected the features with p-values less than 0.5 (at a significance level) as significant predictors.

For the Mobile Price Range dataset, we created a predictive model using the decision tree algorithm. This model aimed to predict the price range of mobile phones based on the relevant independent features identified in the previous steps.

**Data Preprocessing:**

**Handling Missing Values:**

The dataset contains several columns with null values. The percentage of missing values varies across columns, with some columns having only a small percentage of missing values, such as age and salary, while others have a much higher percentage of missing values, such as tea\_per\_year and coffee\_per\_year.

age: 3.095816%

workclass: 3.992941%

salary: 3.103169%

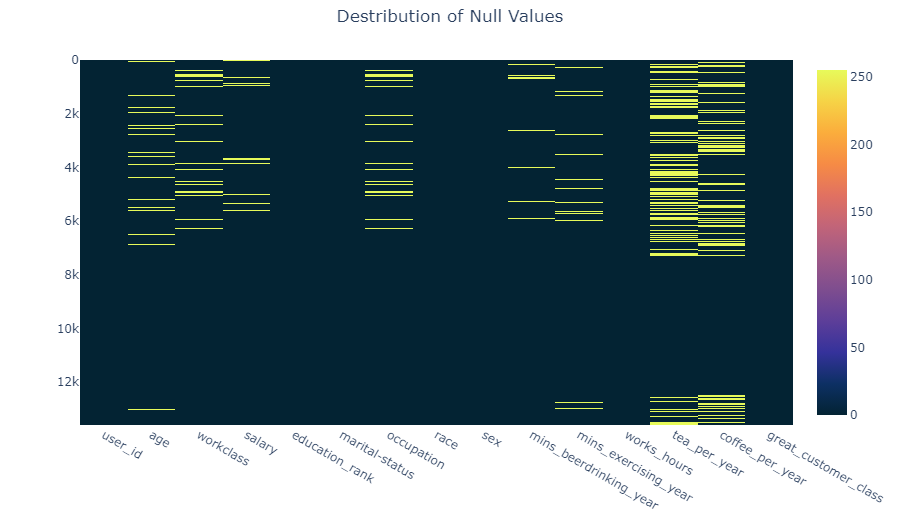
occupation: 3.992941%

mins\_beerdrinking\_year: 3.117876%

mins\_exercising\_year: 3.095816%

tea\_per\_year: 17.861607%

coffee\_per\_year: 17.729245%

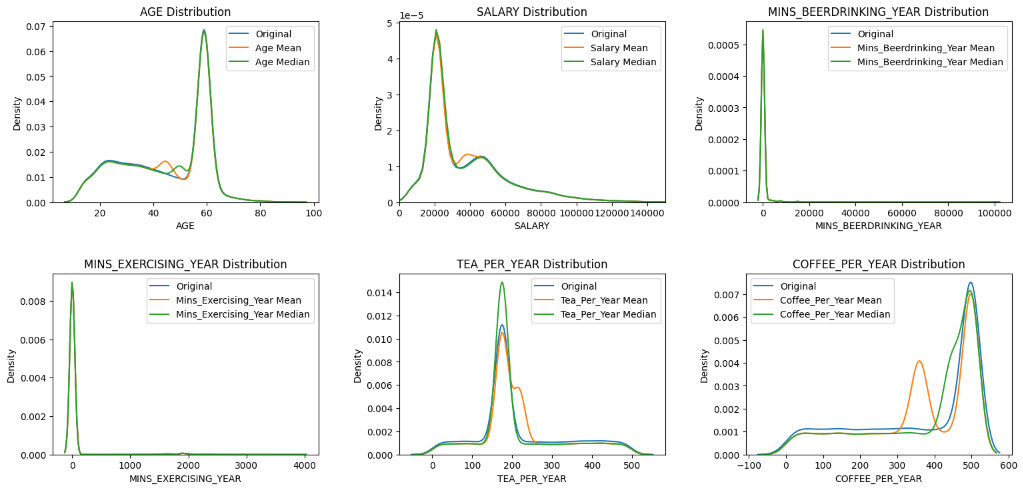


**Visualizing the Null Values**

This missing data can lead to biased or inaccurate results if not handled appropriately.

**Handling Numerical Features:**

Here age, salary, mins\_beerdrinking\_year, mins\_exercising\_year, tea\_per\_year, coffee\_per\_year this column are numeric columns. We filled up these columns with mean or median. We checked the variance for both mean and median compared it with the main variance.

****

**Visualizing the Variance Difference**

age Variance: 257.62 (0.00% Difference)

age\_mean Variance: 249.64 (**3.10%** Difference)

age\_median Variance: 250.44 (2.79% Difference)

salary Variance: 591275399.77 (0.00% Difference)

salary\_mean Variance: 572925773.45 (3.10% Difference)

salary\_median Variance: 576722530.07 (2.46% Difference)

mins\_beerdrinking\_year Variance: 22868806.57 (0.00% Difference)

mins\_beerdrinking\_year\_mean Variance: 22155733.03 (3.12% Difference)

mins\_beerdrinking\_year\_median Variance: 22161793.92 (3.09% Difference)

mins\_exercising\_year Variance: 86573.72 (0.00% Difference)

mins\_exercising\_year\_mean Variance: 83893.36 (3.10% Difference)

mins\_exercising\_year\_median Variance: 83956.94 (3.02% Difference)

tea\_per\_year Variance: 12983.33 (0.00% Difference)

tea\_per\_year\_mean Variance: 10664.13 (17.86% Difference)

tea\_per\_year\_median Variance: 10926.37 (15.84% Difference)

coffee\_per\_year Variance: 27119.69 (0.00% Difference)

coffee\_per\_year\_mean Variance: 22311.22 (17.73% Difference)

coffee\_per\_year\_median Variance: 23422.74 (13.63% Difference)

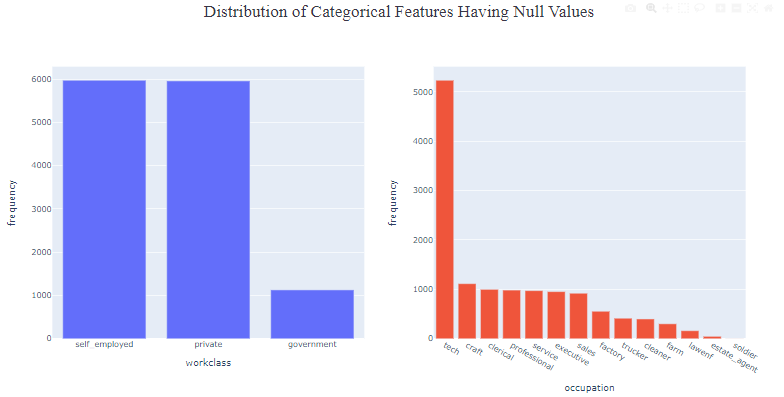
Based on the distribution analysis and the % of change in variance it is observed that there are certain columns that demonstrate less variation in distribution after filling the missing values with the mean. These columns include Mins\_BeerDrinking\_Year, Mins\_Exercising\_Year, and Coffee\_Per\_Year.

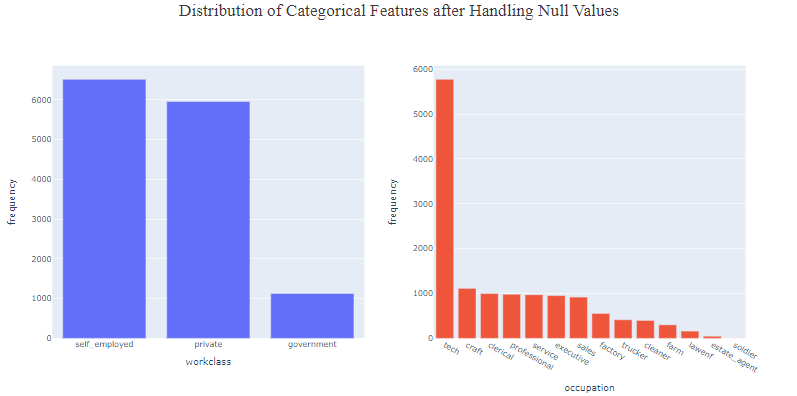
On the other hand, there are other columns such as Age, Salary, and Tea\_Per\_Year, where the distribution shows significantly less variation after filling the missing values with the median. Therefore, we have to use the median to fill the missing values in these columns.

**Handling Categorical Features:**

For categorical features with null values which are workclass and occupation, we used mode to fill up this column.

Distribution Difference before and after Handling Null Values

****

****

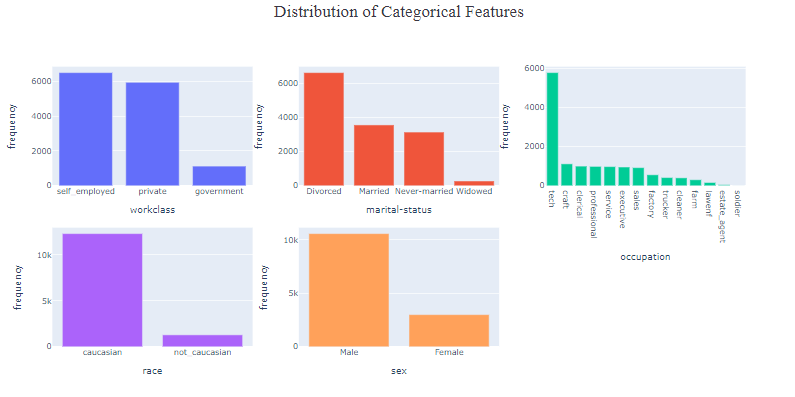
**Encoding:**

In the dataset, categorical columns such as workclass, marital-status, occupation, race, and sex contain string values that represent different categories or classes. In order to use this data for statistical analyses, it is necessary to encode these categorical string columns into numerical values first.

As our string columns are nominal, we used one hot encoding. In this method each category is represented by a binary vector, with a value of 1 indicating the presence of that category and a value of 0 indicating the absence of that category.

As an example,

There are four categories under marital-status feature. So, after encoding this feature convert into 4 new boolean features, which are 'Never-married' 'Divorced' 'Married' 'Widowed'. But we will consider any 3 from these 4 columns. Because it is possible to derive the fourth category from the other three, as it represents individuals who were previously married but are now widowed. By creating separate boolean columns for the remaining three categories, we can better understand the relationships between marital status and the target variable, "great\_customer\_class," in our analysis.



From the above plot we can see there are too many categories in occupation feature. If we use one hot encoding method directly in this feature it will generate a lot of new features resulting in a curse of dimensionality and increase computational complexity. As a result, we can’t use normal one hot encoding method directly to occupation.

To handle occupation feature we grouped the categories in 6 new features. Then we used binary encoding for this feature.

Grouping the categories for the occupation feature:

**Office jobs:** clerical, professional, executive, lawenf, estate\_agent

**Manual jobs:** farm, craft, factory, cleaner, soldier

**Sales jobs:** sales

**Service jobs:** service

**Driver jobs:** trucker

**Tech jobs**: tech

After encoding the dataset will have these new features shown below,

**work\_class:** workclass\_private, workclass\_self\_employed, workclass\_government,

**marital\_status:** marital-status\_Never-married, marital-status\_Divorced, marital-status\_Married, marital-status\_Widowed,

**race:** race\_not\_caucasian, race\_caucasian,

**sex:** sex\_Male, sex\_Female

As mentioned earlier we have to remove one new feature from each category that we encoded. So after this process the dataset will have these following features:

'user\_id', 'age', 'office\_jobs', 'manual\_jobs', 'sales\_jobs', 'service\_jobs', 'Tech\_jobs', 'workclass\_private',

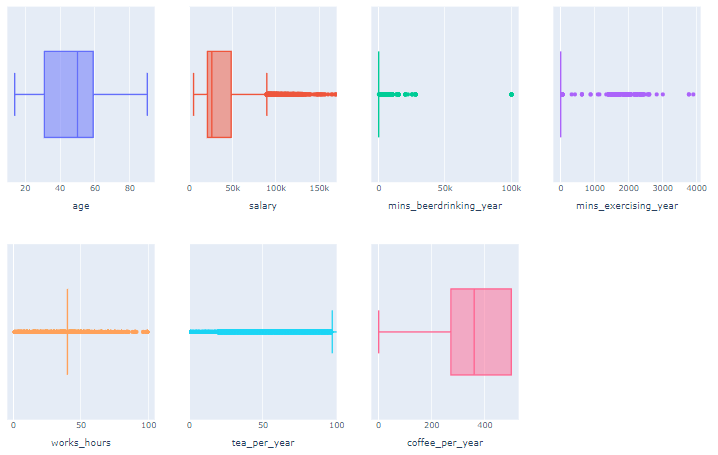
'workclass\_government', 'salary', 'education\_rank', 'marital-status\_Divorced', 'marital-status\_Married',

'marital-status\_Widowed', 'race\_caucasian', 'sex\_Male', 'mins\_beerdrinking\_year',

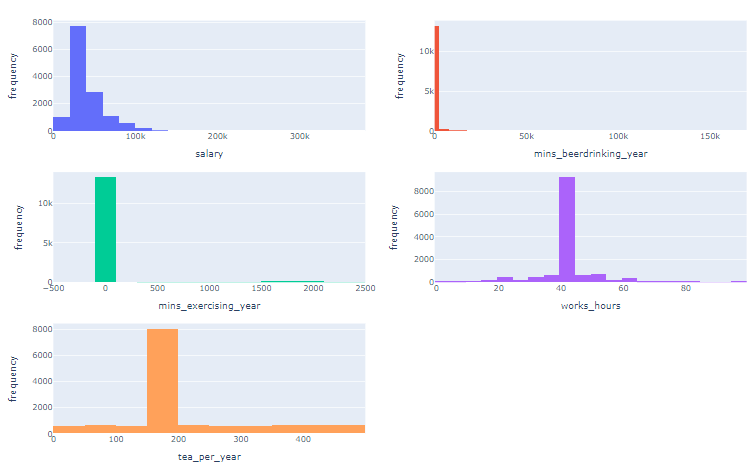
'mins\_exercising\_year', 'works\_hours', 'tea\_per\_year', 'coffee\_per\_year', 'great\_customer\_class'

**Outlier Handling:**

After handling missing values, we handled outliers.



**Visualization of the outliers of numerical features**



**Numerical Columns Distribution before Handling Outliers**

**Summary Statistics of features with outliers**:

salary Percentage : 3.75%

salary Mean : 111321.91

salary Median : 102429.62

salary Std : 28982.65

mins\_beerdrinking\_year Percentage : 7.09%

mins\_beerdrinking\_year Mean : 6318.75

mins\_beerdrinking\_year Median : 2174.00

mins\_beerdrinking\_year Std : 16596.03

mins\_exercising\_year Percentage : 5.51%

mins\_exercising\_year Mean : 835.81

mins\_exercising\_year Median : 46.03

mins\_exercising\_year Std : 928.96

works\_hours Percentage : 33.33%

works\_hours Mean : 39.75

works\_hours Median : 42.00

works\_hours Std : 17.04

tea\_per\_year Percentage : 26.83%

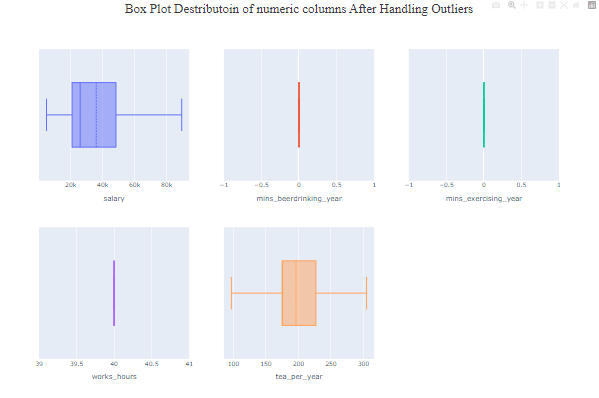
tea\_per\_year Mean : 286.60

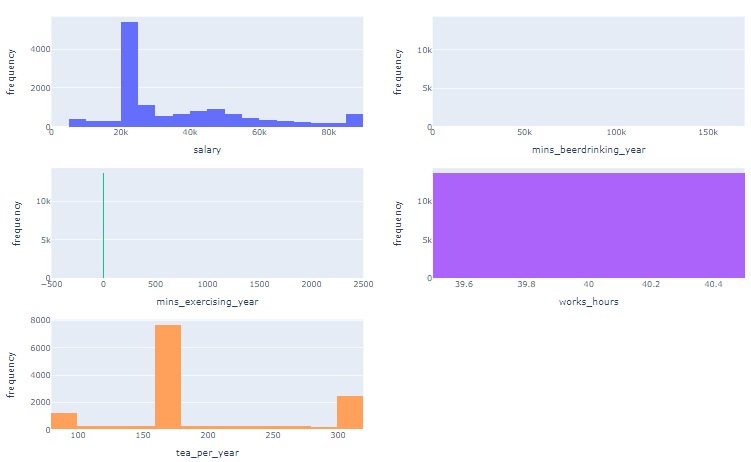
tea\_per\_year Median : 356.50

tea\_per\_year Std : 172.66

From the above analysis it is clear that there are a huge number of outliers in salary, mins\_beerdrinking\_year, mins\_exercising\_year, works\_hours, tea\_per\_year columns.

To handle outliers, we used a capping method.

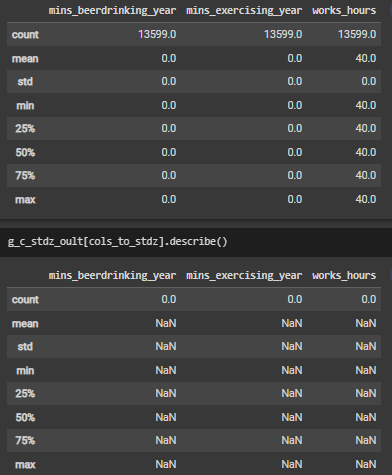




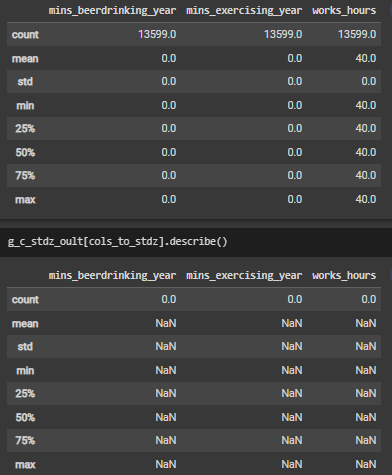
**Hist plot distribution of numerical feature after handling outliers**

As we can see there is extreme change in mins\_beerdrinking\_year, mins\_exercising\_year and works\_hour column. So, it is clear that we have to use other methods to determine if these features are important.

There comes our next choice standardizing the features which have extreme change. Standardizing is a data preprocessing method that scales the data to have a mean of 0 and standard deviation of 1. And the value range is -1 to 1.



Before standardization



After standardization

After standardization we can see that all the values are NaN which means data are missing due to division by standard deviation which was 0.00. So, it’s clear that we can’t use the standardization method.

After using capping method,

Variance of mins\_beerdrinking\_year : 0.0

Variance of mins\_exercising\_year : 0.0

Variance of works\_hours : 0.0

For these conditions we have excluded these three features from the dataset for further analysis.

**Hypothesis Testing:**

Hypothesis testing can be performed on this dataset to determine the statistical significance of the relationship between the target variable 'great\_customer\_class' and other features. One possible hypothesis test is to perform a chi-square test of the association for the categorical features and an ANOVA test for the numerical features to determine whether there is a significant difference in the mean target variable across different levels of the categorical features such as 'workclass', 'marital-status', 'occupation', 'race', and 'sex'.

The hypothesis test can be formulated as follows:

Null Hypothesis: There is no significant difference in the mean 'great\_customer\_class' across different levels of the categorical features in the dataset.

Alternative Hypothesis: There is a significant difference in the mean 'great\_customer\_class' across different levels of the categorical features in the dataset.

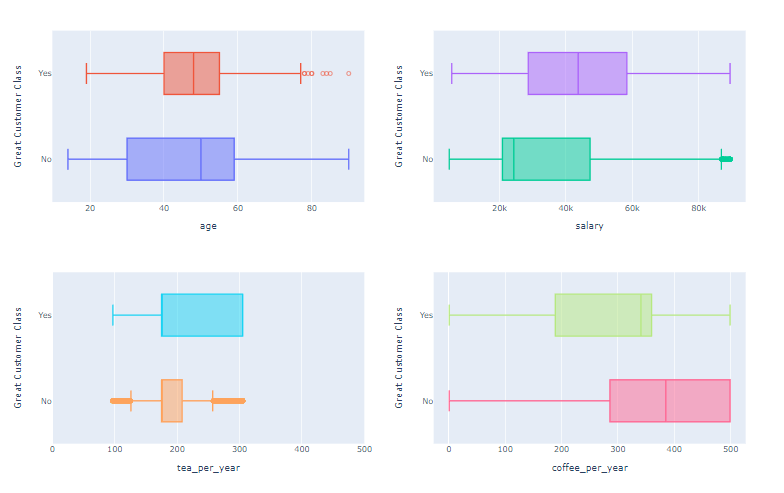
Anova test analyzes the variance between different categories or groups to check if the differences between the category means are significant. It is done by calculating the ratio of the variance between groups to the variance within groups, which is called F-statistic. If the F-statistic is greater than the critical value, then we can reject the null hypothesis, which means at least one group is significantly different than other groups.

**P-value, Test of association:**

A test of association means a test of independence or a chi-square test. It is a statistical test used to determine whether two categorical features are associated with each other or not. In chi-square the null hypothesis is that there is no association between the two variables. And the alternate hypothesis is that there is a significant association between the two variables.

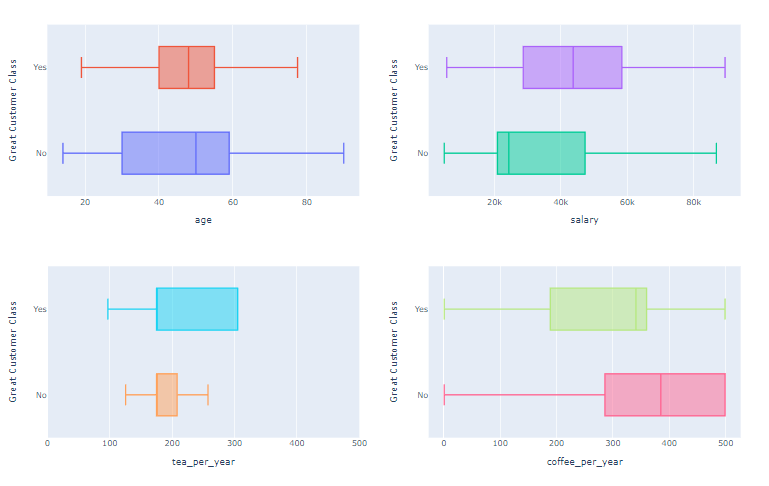
Chi-square test calculates a p-value, which helps to observe the association between the variables if the null hypothesis is true. If the p-values are low it means the null hypothesis is rejected.

In the dataset the target feature is "great\_customer\_class" and the categorical predictor features are ‘office\_jobs', 'manual\_jobs', 'sales\_jobs', 'service\_jobs', 'Tech\_jobs', 'workclass\_private', 'workclass\_government', 'education\_rank', 'marital-status\_Divorced', 'marital-status\_Married', 'marital-status\_Widowed', 'race\_caucasian', 'sex\_Male', 'great\_customer\_class'.



**Relationship between Numerical Features and Great Customer Class**

**After handling the outlier using capping method:**

****

**Relationship between Numerical Features and Great Customer Class After handling Outlier**

****

**Relationship between Categorical Features and Great Customer Class**

**Comparing the result of our Anova Test and Chi Square test to built in functions:**

**Anova:**

age

stat = 41.07733519698571, p\_value = 1.511501226582168e-10

my\_stat = 41.07733519698485, my\_p\_value = 1.511501226582168e-10

salary

stat = 287.8010578033135, p\_value = 6.799046308534331e-64

my\_stat = 287.80105780331127, my\_p\_value = 6.79904630853974e-64

tea\_per\_year

stat = 257.4107076172952, p\_value = 2.115958342152212e-57

my\_stat = 257.4107076172933, my\_p\_value = 2.115958342153896e-57

Chi Square:

office\_jobs

stat = 947.1691670031346, p\_value = 5.4713896157168446e-208

my\_stat = 947.1691670031346, my\_p\_value = 0.0

manual\_jobs

stat = 1.2597585965627767, p\_value = 0.2616967899077225

my\_stat = 1.2597585965627764, my\_p\_value = 0.2616967899077257

sales\_jobs

stat = 43.271700834306515, p\_value = 4.764243869439441e-11

my\_stat = 43.27170083430652, my\_p\_value = 4.764244554422703e-11

service\_jobs

stat = 62.48075234087651, p\_value = 2.690614379798907e-15

my\_stat = 62.480752340876506, my\_p\_value = 2.6645352591003757e-15

Tech\_jobs

stat = 696.202521882976, p\_value = 2.0021375665855697e-153

my\_stat = 696.202521882976, my\_p\_value = 0.0

workclass\_private

stat = 181.42132534949388, p\_value = 2.3718866118252305e-41

my\_stat = 181.42132534949386, my\_p\_value = 0.0

workclass\_government

stat = 126.21262757737627, p\_value = 2.762418448957248e-29

my\_stat = 126.21262757737627, my\_p\_value = 0.0

education\_rank

stat = 2028.2817524297695, p\_value = 0.0

my\_stat = 2028.2817524297698, my\_p\_value = 0.0

marital-status\_Divorced

stat = 881.8808654707948, p\_value = 8.526037551126789e-194

my\_stat = 881.8808654707948, my\_p\_value = 0.0

marital-status\_Married

stat = 2377.4212220155305, p\_value = 0.0

my\_stat = 2377.4212220155305, my\_p\_value = 0.0

bd268

marital-status\_Widowed

stat = 8.689087533219803, p\_value = 0.003201208453255956

my\_stat = 8.689087533219803, my\_p\_value = 0.0032012084532559992

race\_caucasian

stat = 1.1951838710199714, p\_value = 0.27428639646303465

my\_stat = 1.1951838710199716, my\_p\_value = 0.27428639646303776

sex\_Male

stat = 50.00942946582381, p\_value = 1.5300891364406532e-12

my\_stat = 50.00942946582381, my\_p\_value = 1.5301093725383907e-12

Feature Selection:

After performing Anova Test between each numerical feature and target feature(great\_customer\_class) and chi square test of association between each categorical feature and target feature we get the following p values in order (low - high):

office\_jobs : 0.0

Tech\_jobs : 0.0

workclass\_private : 0.0

workclass\_government : 0.0

education\_rank : 0.0

marital-status\_Divorced : 0.0

marital-status\_Married : 0.0

salary : 6.799046308534331e-64

tea\_per\_year : 2.115958342152212e-57

service\_jobs : 2.6645352591003757e-15

sex\_Male : 1.5301093725383907e-12

sales\_jobs : 4.764244554422703e-11

age : 1.511501226582168e-10

marital-status\_Widowed : 0.0032012084532559992

manual\_jobs : 0.2616967899077257

race\_caucasian : 0.27428639646303776

Based on the 0.5 significance level the analysis revealed that several features are highly significant predictors of the 'great\_customer\_class', including 'office\_jobs', 'Tech\_jobs', 'workclass\_private', 'workclass\_government', 'education\_rank', 'marital-status\_Divorced', and 'marital-status\_Married'.

Additionally, features such as 'salary', 'tea\_per\_year', 'service\_jobs', 'sex\_Male', 'sales\_jobs', and 'age' showed moderate significance.

**Relevant Exploratory Data Analysis:**

1. Distribution of age among the great customers

A picture containing diagram, screenshot, plot, pixel

Description automatically generated

2. Is there a correlation between an individual's age and salary?

A black squares with white text

Description automatically generated with low confidence

A picture containing screenshot, text, colorfulness, line

Description automatically generated

3. Is there a significant difference in the average salary between males and females.

The t-test performed indicates a statistically significant difference between the average salaries of males and females. The obtained t-statistic of 23.66 and p-value of 7.33e-117 suggest that the difference in salary between males and females is not likely due to chance. Therefore, we reject the null hypothesis and conclude that there is a significant difference in the average salary between males and females.

4. What is the most common occupation among individuals who are self-employedA picture containing text, screenshot, diagram, number

Description automatically generated

Based on the bar plot, it appears that Tech\_jobs is the most common occupation among self-employed individuals. This may suggest that there is a growing trend towards self-employment in the technology industry.

5. Are individuals who work in government jobs more likely to consume coffee or tea, and how does this vary by race?

A picture containing text, screenshot, diagram, font

Description automatically generated

From above plot we can see that, it appears that both non-Caucasian and Caucasian individuals working in government jobs consume coffee more frequently than tea. However, the average coffee consumption among Caucasians appears to be slightly higher than non-Caucasians, with a mean value of 283.9708 compared to 260.54. On the other hand, non-Caucasians appear to consume slightly more tea on average, with a mean value of 204.3896 compared to 208.8661 for Caucasians. However, the difference in tea consumption between the two racial groups is much smaller than the difference in coffee consumption. Overall, it seems that both coffee and tea are popular beverages among government workers, regardless of their race.

6. Are individuals who are married more likely to have a higher education rank than those who are unmarried, and does this differ by race?

A picture containing text, screenshot, font

Description automatically generated

From the plot, it is difficult to draw any conclusions about the relationship between marital status and education rank, as all four data points are exactly the same. All four data points have a race value of either Caucasian or non-Caucasian, a marital status value of either married or unmarried, and an education rank value of 9.

Therefore, we can conclude that there is no observed difference in education rank based on marital status within each racial group in the data being presented.

7. Average salary across different job types (office, manual, sales, service, tech)

A picture containing text, screenshot, colorfulness, line

Description automatically generated

From above plot, we can see that the average salary varies slightly across different job types. Among the job types listed, the highest average salary is for manual jobs at 45118.59, followed closely by sales jobs at 45296.86 and office jobs at 44820.61. Service jobs have a slightly lower average salary at 43531.55.

The average salary for tech jobs is much lower than the other job types listed, at 24007.83. This could indicate that the tech jobs in this dataset are lower-skilled or entry-level positions with lower salaries, or that the dataset only includes a small number of tech jobs.

Overall, the plot suggests that there is some variation in average salaries across different job types, with manual and sales jobs generally having higher average salaries than service and tech jobs. However, it's important to note that the provided data only includes a limited number of job types, and that the salaries within each job type may also vary widely depending on factors such as experience level and location.

8. How does the great\_customer\_class vary across different categories such as race and sex?

A picture containing text, screenshot, plot, software

Description automatically generated

From the plot, we can see that the great\_customer\_class appears to vary across different categories of sex and race. Among non-Caucasian customers, the sum of the great\_customer\_class is 24 for females (sex\_Male=0) and 16 for males (sex\_Male=1). This suggests that female non-Caucasian customers may be more likely to have a higher great\_customer\_class than male non-Caucasian customers. However, it's important to note that the difference in great\_customer\_class between the two groups is relatively small.

Among Caucasian customers, the sum of the great\_customer\_class is much higher overall, with a sum of 930 for males (sex\_Male=1) and 138 for females (sex\_Male=0). This suggests that Caucasian male customers may be more likely to have a higher great\_customer\_class than female Caucasian customers.

Overall, it's important to note that the provided plot only shows the sum of the great\_customer\_class for each group, and does not provide information on the distribution or range of great\_customer\_class values within each group. Additionally, it's possible that other factors, such as age or income, could also be influencing the variation in great\_customer\_class across different categories.

9. Distribution of office\_jobs, manual\_jobs, sales\_jobs, service\_jobs, and Tech\_jobs across different categories such as race and sex?

A picture containing text, screenshot, diagram, line

Description automatically generated

A picture containing text, screenshot, diagram, line

Description automatically generated

A picture containing text, screenshot, diagram, line

Description automatically generated

A picture containing text, screenshot, rectangle, diagram

Description automatically generated

A picture containing text, screenshot, diagram, line

Description automatically generated

From the above plot, we can see the distribution of office\_jobs, manual\_jobs, sales\_jobs, service\_jobs, and Tech\_jobs across different categories such as race and sex.

For office\_jobs, we can see that the sum of office\_jobs is higher for males (sex\_Male=1) than females (sex\_Male=0) for both non-Caucasian and Caucasian groups, with a particularly large difference for the non-Caucasian group.

For manual\_jobs, we can see that the sum of manual\_jobs is also higher for males (sex\_Male=1) than females (sex\_Male=0) for both non-Caucasian and Caucasian groups, although the difference is not as large as for office\_jobs.

For sales\_jobs, we can see that the sum of sales\_jobs is higher for males (sex\_Male=1) than females (sex\_Male=0) for both non-Caucasian and Caucasian groups, although again the difference is not as large as for office\_jobs.

For service\_jobs, we do not have any information in the provided plot.

For Tech\_jobs, we can see that the sum of Tech\_jobs is higher for males (sex\_Male=1) than females (sex\_Male=0) for both non-Caucasian and Caucasian groups, although again the difference is not as large as for office\_jobs.

Overall, it appears that there may be differences in the distribution of jobs across different categories such as race and sex, with males generally having a higher sum of jobs than females, particularly for office\_jobs.

10. The proportion of great customers in each marital status category

A picture containing text, screenshot, plot, number

Description automatically generated

The plot shows the proportion of great customers in each marital status category. Among the four categories, married customers have the highest proportion of great customers with 86.13%, followed by divorced customers with 7.19%. Unmarried customers have a proportion of 5.82%, while windowed customers have the lowest proportion with 0.86%. This information could be useful for businesses that target customers based on their marital status to adjust their marketing strategies to attract more customers who are more likely to be great customers. However, it's important to keep in mind that this data is based on a specific sample and may not be representative of the population as a whole.

11. The proportion of not great customers in each marital status category

A picture containing text, screenshot, diagram, rectangle

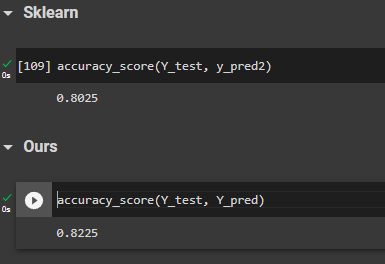
Description automatically generated

The plot shows the proportion of not great customers in each marital status category. Among the four categories, divorced customers have the highest proportion of not great customers with 52.62%, followed by unmarried customers with 24.72%. Married customers have a proportion of 20.54% of not great customers, while widow customers have the lowest proportion with 2.12%. This information could be useful for businesses that target customers based on their marital status to adjust their marketing strategies to avoid attracting customers who are less likely to be great customers. However, it's important to keep in mind that this data is based on a specific sample and may not be representative of the population as a whole.

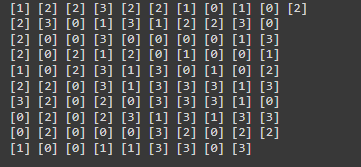
**Model:**

We create a model using the decision tree algorithm to predict the price range of mobile phones using relevant independent features.

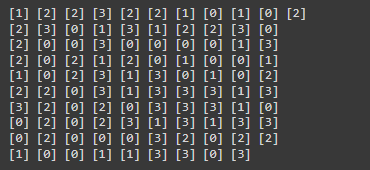
Here is the accuracy of our model:



Sampel of Y\_Test:



Sample of Y Predict:



**Discussion:**

Great Customer Class:

Based on our analysis using a significance level of 0.5, we have identified several highly significant predictors of the 'great\_customer\_class' variable. These predictors include 'office\_jobs', 'Tech\_jobs', 'workclass\_private', 'workclass\_government', 'education\_rank', 'marital-status\_Divorced', and 'marital-status\_Married'.

Furthermore, we found that certain features, such as 'salary', 'tea\_per\_year', 'service\_jobs', 'sex\_Male', 'sales\_jobs', and 'age', exhibited moderate significance in predicting the 'great\_customer\_class' variable.

These results suggest that the identified features play a crucial role in determining the classification of a customer as a "great\_customer\_class."

Mobile Price Range

Following a comprehensive analysis of the dataset and the determination of p-values, we have reached the following conclusions. The features "Ram" and "battery\_power" demonstrate the smallest p-values, indicating a robust relationship with the target variable. Similarly, "Px\_width" and "px\_height" exhibit statistically significant p-values, albeit to a slightly lesser extent. Conversely, the remaining features have larger p-values, suggesting a weaker association with the target variable.

Consequently, we have selected "ram," "battery\_power," "px\_width," "px\_height," "mobile\_wt," "int\_memory," "n\_cores," and "sc\_h" as the top features for training our model. These features are deemed to hold the most predictive power in relation to the target variable.

Heart Disease:

After conducting an analysis on this dataset, we have concluded that, with the exception of 'heartRate' and 'currentSmoker', all the features have p-values less than 0.05. This indicates that these features are statistically significant and possess a substantial relationship with the target variable, TenYearCHD.

Consequently, the features identified as statistically significant, based on the provided p-values, are as follows: prevalentHyp, age, sysBP, diaBP, diabetes, male, BPMeds, totChol, education, BMI, prevalentStroke, and cigsPerDay.

**Conclusion:**

This project has been a valuable learning experience for us in the field of data analysis. We have gained knowledge and skills in handling missing values, outliers, encoding techniques, exploratory data analysis, hypothesis testing, and model creation.

Throughout the project, we encountered challenges and obstacles, but with the help of online documentation and resources like Stack Overflow, we were able to overcome them and find solutions.

Overall, this project has increased our familiarity and comfort with various data analysis techniques. We have gained practical experience and a better understanding of how to apply these techniques effectively. We feel more confident in our abilities and are excited to apply what we have learned in future projects.

# Appendix