Customer Retention Case Study

ACKNOWLEDGMENT

It gives me great delight and pleasure to deliver this report. Working on this project was a fantastic learning experience that provided me with a wealth of information on the data analysis process.

Flip Robo Technologies (Bangalore) provided all of the necessary information and datasets, which aided me in completing the project.

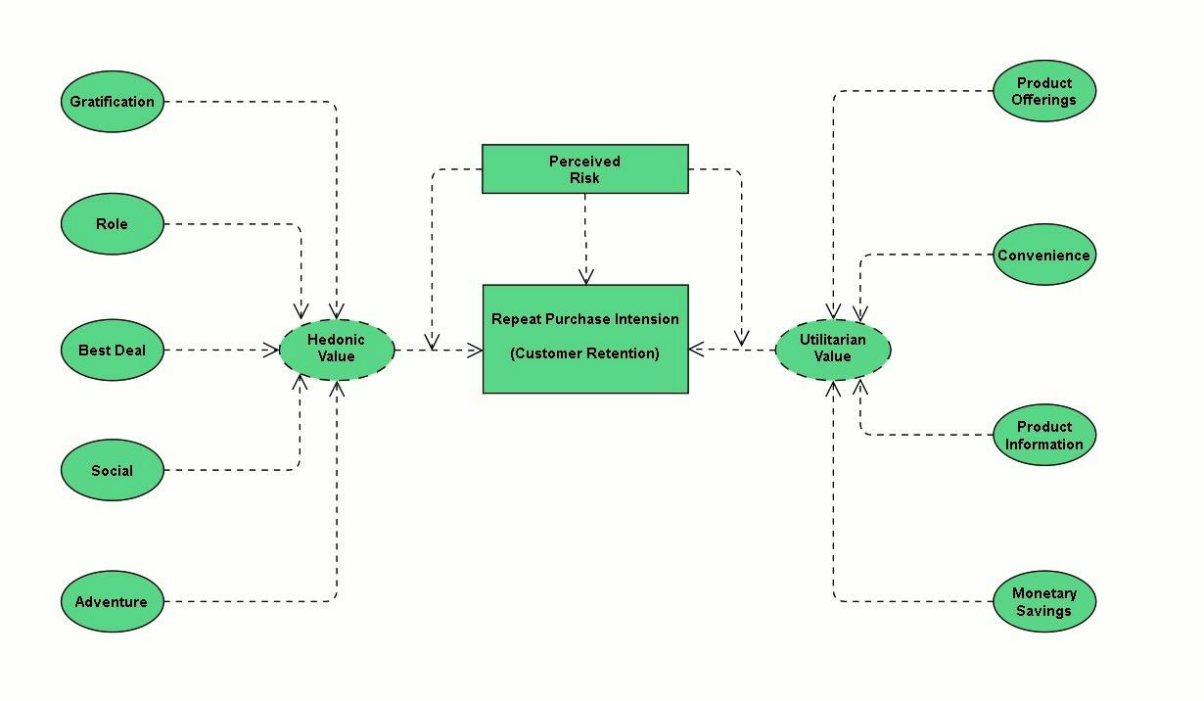
I'd want to express my gratitude to my SME, Khushboo Garg, for providing the dataset and directions for carrying out the entire case study procedure.

INTROUDCTION

Problem Statement

Customer satisfaction has emerged as one of the most important factors that guarantee the success of online store; it has been posited as a key stimulant of purchase, repurchase intentions and customer loyalty. A comprehensive review of the literature, theories and models have been carried out to propose the models for customer activation and customer retention. Five major factors that contributed to the success of an e-commerce store have been identified as: service quality, system quality, information quality, trust and net benefit. The research furthermore investigated the factors that influence the online customers repeat purchase intention. The combination of both utilitarian value and hedonistic values are needed to affect the repeat purchase intention (loyalty) positively. The data is collected from the Indian online shoppers. Results indicate the e-retail success factors, which are very much critical for customer satisfaction.

Use Case Diagram



We can see from the above use case diagram that our Customer Retention approach is based on Hedonic and Utilitarian value. We also see that our clients' buy and repurchase intentions are influenced by perceived risks. The Hedonic value consists of five primary components: satisfaction, role, best bargain, social aspect, and adventurous sensation. Product options, convenience, product knowledge, and monetary savings are all examples of utilitarian value.

Motivation for the Problem Undertaken

The primary goal of this study is to determine if consumers purchase for items on e-commerce websites. How did they provide feedback to these

websites based on a variety of good and negative characteristics, as well as user information such as age, gender, and location?

Benefits of Customer Retention

1. Retention is less expensive than acquisition
    - While the old cliché "it costs five times as much to acquire a new client" may not always be true, the core idea is correct: it is more cost-effective to maintain an existing customer than to bring in new ones.
    - Even yet, if it's data you're after, there's been plenty of studies on acquisition vs. retention, and every one of them has found that retention is the more economically feasible option.
    - One caveat, however: retention is less expensive than acquisition, but it isn't always simpler.

2. Customers that are loyal to you are more profitable.

- Loyalty is not only less expensive, but it also yields better results. According to studies, engaged customers buy 90% more frequently, spend 60% more each transaction, and are five times more likely to say it is the only brand they would buy in the future.
- They generate 23 percent more income and profit on average than the average consumer.
- While it's true that loyal consumers are more profitable, don't take it for granted.
- They'll be more receptive to price hikes, but don't boost prices only to test how long they'll stay.
- Consider the flip side: "Actively disengaged" consumers (those who are actively expressing their dissatisfaction with the brand) can cost a company up to 13% of its sales.

3. Your brand will stand out from the crowd.

- Put on your consumer hat and think about how many brands you encounter with that seem to appreciate your business.
- You probably just have one or two ideas.
- Most businesses prioritise acquisition, which makes those of us that prioritise retention stand out even more.
- People are exposed to 10,000 marketing messages every day, yet only engage with a small percentage of them.
- Those with whom they have an emotional connection on some level are the ones who gain continued involvement.
- The top brands have a distinctive retention proposition, not a unique selling proposition.

4. You'll get more referrals from word-of-mouth

- Your most reliable source of new business will be your loyal customers.
- People are still most strongly affected by referrals from friends and family, despite all of the efforts put into internet and mobile marketing and social media.
- Millennials, in particular, will spread the news about a company's accomplishments: 90% of millennials express their brand preferences online.

5. More Feedback from Engaged Customers

- Feedback is vital to any business's success.
- Why Customers who share feedback are frequently eager to give businesses a second chance.
- They continuously tell you how to win their business. According to studies, those who have complained and had their problem fixed are 84 percent less inclined to cut back on their spending.
- Do you need assistance dealing with clients who provide negative feedback?

6. Customers will explore your brand

- That's a nice way of saying you'll be able to sell them more stuff.
- Once a brand has proven itself with one product or service, customers are six times more likely to say they'd try a new product or service from the brand as soon as it's available.

- This isn't just valuable for sales; these people can also be used to help with #5 above as beta testers, which is an important part of product development.

7. Customers that are loyal to you are more forgiving.

- According to an Accenture survey, clients who leave after a bad service experience lose nearly $1.6 trillion each year.
- We've even gone so far as to say that it's the most common cause for customers to abandon a brand.
- Customers who consider themselves loyal, on the other hand, will overlook certain mistakes if they don't happen too frequently.

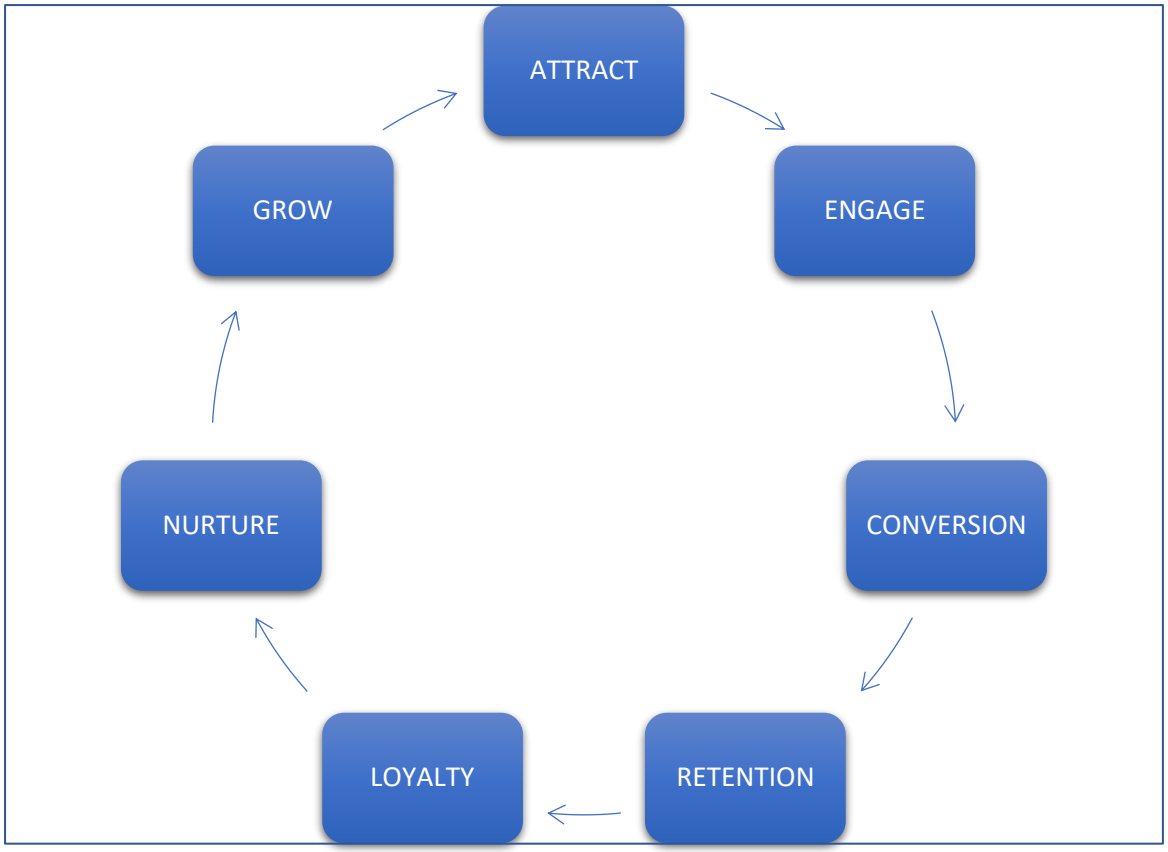
8. Customers will appreciate your marketing efforts.

- Except for devoted consumers, no one enjoys being sold to.
- They're four times more likely to say they "appreciate it when this brand contacts me" and seven times more likely to say they "always react to this brand's promotional offers."

9. You have the freedom to attempt new things by earning wiggle room.

- Because loyalty is fleeting, too many changes may drive individuals away.
- However, after you've built a solid foundation of loyal consumers, your business may grow.
- Perhaps fresh marketing, a new product line, or even a new logo is required. The bottom line is that as long as you preserve the fundamental principles that keep people on your side, they will remain with you through thick and thin. Some of them may even be interested in seeing what you can achieve.
- According to a survey, existing consumers are 50% more willing to try new items.

CLIENT LIFECYCLE STAGES



The eventual aim is lifetime revenue, not just today's revenue.

Need for Customer Retention:

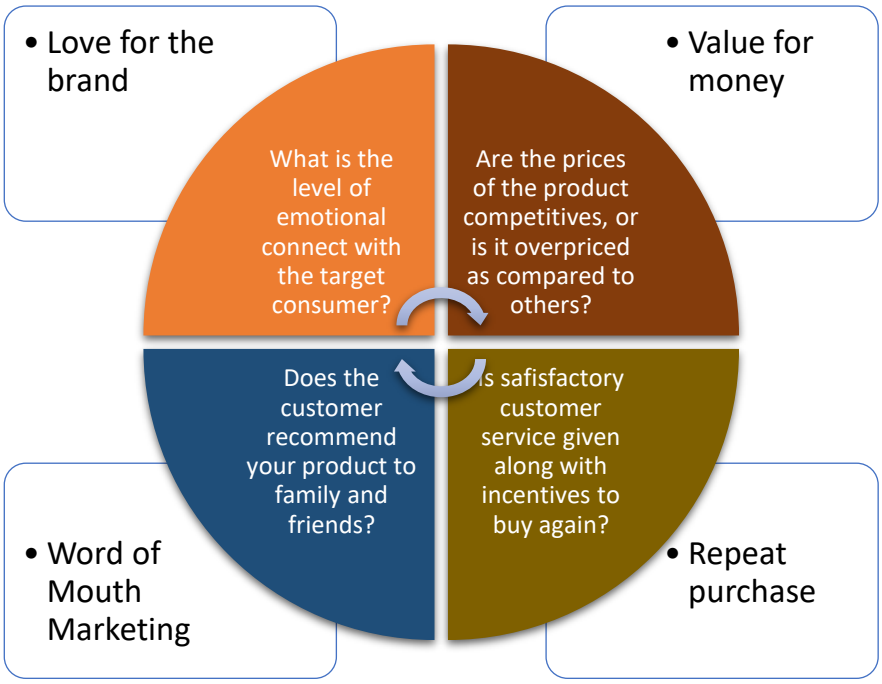
Keeping current customers satisfied is usually less expensive than recruiting new consumers. According to the Harvard Business Review, getting a new client might cost anywhere from five to twenty-five times more than keeping an existing one.

Marketing, promotion, and sales outreach do not need to be expensive. Because they already trust your brand from prior purchases, it's simpler to convert existing consumers into repeat buyers. When it comes to the initial sale, new consumers, on the other hand, generally require more persuasion.

Customers that are loyal to a company are more likely to do business with them again. Customers that are loyal are more inclined to pass on free recommendations to their coworkers, friends, and family. One strategy for a firm to foster client loyalty for long-term success is to create that cycle of retained customers and buzz marketing.

Customer retention may be improved through enhancing the customer experience. In fact, according to a 2021 Customer Experience Trend Report, 77 percent of customers will be more loyal to a firm that provides a positive customer experience if they have a problem. 72 percent are prepared to pay more for a firm that provides excellent customer service. In addition, 50% feel that customer experience is more essential now than it was a year ago.

Customer Retention



Because it costs five to ten times more to acquire a new client than it does to maintain an existing one, nurturing loyal consumers is a potent method for organisations to expand.

Dataset Details:

To develop a complete data analysis in Python, I first loaded all of the essential libraries and dependencies.

```
#importing required libraries

import pandas as pd
import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt
%matplotlib inline

import missingno
import pandas_profiling

from sklearn.preprocessing import OrdinalEncoder

import warnings
warnings.filterwarnings('ignore')
warnings.simplefilter('ignore')
```

The sheets in our Excel spreadsheet were then segregated and placed in two independent dataframe variables.

```
# importing dataset and assiging it to xls
xls = pd.ExcelFile('customer_retention_dataset.xlsx')
```

```
# importing excel sheet_1 assigning it to df1
df1= pd.read_excel(xls, 'datasheet')
```

```
# importing excel sheet_2 assigning it to df2
df2= pd.read_excel(xls, 'codedsheet')
```

EDA:

After importing the dataset into our Jupyter Notebook, I saw that the data had been truncated owing to the enormous amount of rows and columns. As a result, I utilised the pandas code as given below to address this difficulty.

I made sure to rename any column names that were badly formed or too long for me to understand. I was able to use rename to adjust the titles of columns that were too long and might have been fit into a smaller format.

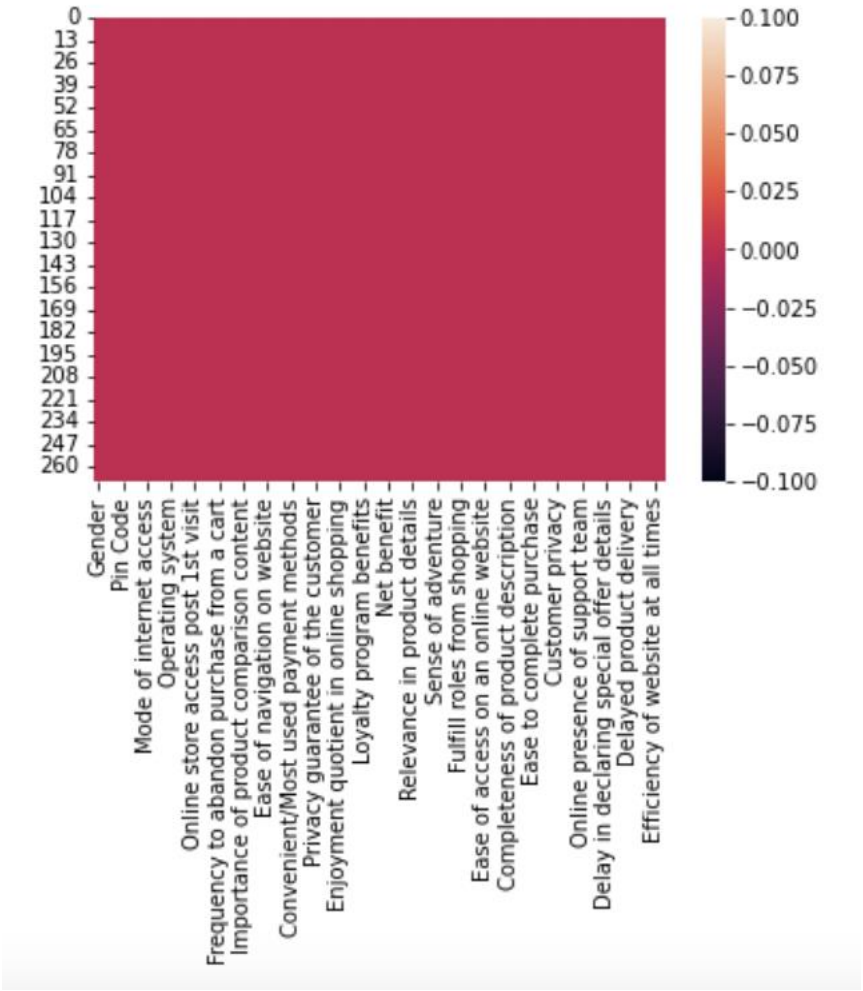
It was now time to check for any missing values or null values that could have been present in our dataset.

```
#checking for null values

df1.isnull().sum()
```

```
# checking nullvalue using heatmap
sns.heatmap(data=df1.isnull())
```

Fortunately, I was able to observe that our complete dataset had no missing values, as seen by the heatmap visual below.



I then proceeded to examine each record's information using the describe, info, and nunique functions.

```
df1.describe(include='all').T
# checking information of datatype using info()
df1.info()
#cheking unique values

df1.nunique().to_frame('Unique values')
```

I used a for loop to look at all of the unique values in the categorical columns, which covered the whole dataset's amount of rows.

```
# listing all Categorical data using for_loop method
for x in obj_columns:
    print(x)
    print(df1[x].value_counts())
    print("-"*70)
```

Visualization:

What is the definition of data visualisation?

A graphic depiction of information and data is referred to as data visualisation.

What are the advantages of good data visualisation?

Data visualisation is another visual art approach that piques our interest and retains our attention on the message collected by the eyes.

The following are examples of several types of data visualisation analysis:

1. Univariate Analysis: In univariate analysis, we will study practically all of the attributes of a single feature.

2. Bivariate Analysis: Bivariate analysis is defined as a comparison of data between exactly two attributes.

3. Multivariate Analysis: We shall compare more than two variables in the multivariate analysis.

Univariate Analysis:

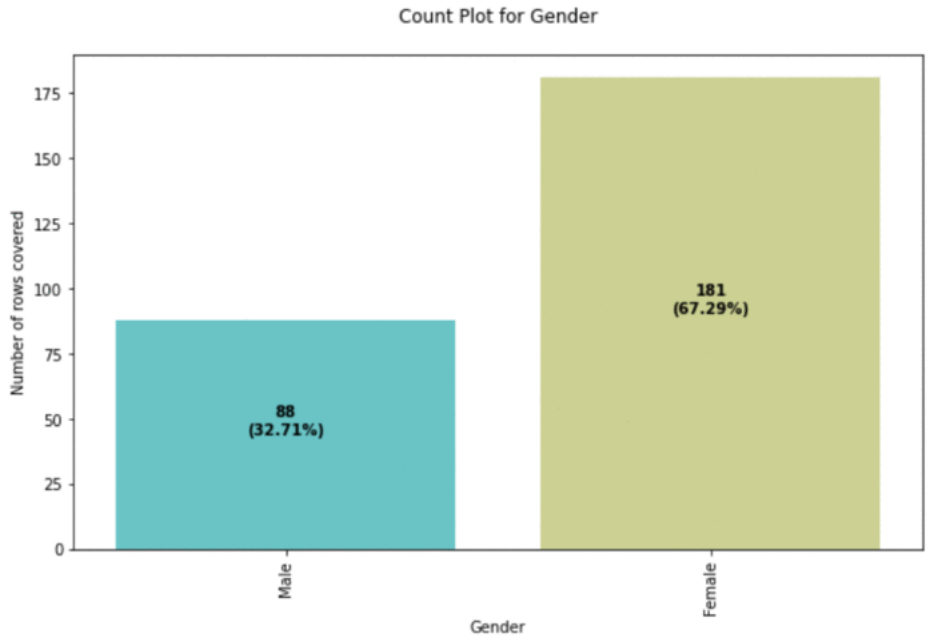
2 for loops were used to construct count plots for all of our columns, displaying the percentage of data coverage.

```
for C in df1[obj_columns]:
    plt.figure(figsize=(10,6))
    col_name = C
    values = df1[col_name].value_counts()
    index = 0
    ax = sns.countplot(df1[col_name], palette="rainbow")

    for i in ax.patches:
        h = i.get_height() #Count of each value
        t = len(df1[col_name]) # Total no of records using length
        s = f"{h}\n({round(h*100/t,2)}%)" #Displaying in count bar
        plt.text(index, h/2, s, ha="center", fontweight="bold")
        index += 1

    plt.title(f"Count Plot for {col_name}\n")
    plt.xlabel(col_name)
    plt.ylabel(f"Number of rows covered")
    plt.xticks(rotation=90)
    plt.show()
```

This code resulted in multiple count plot pictures, as seen below.



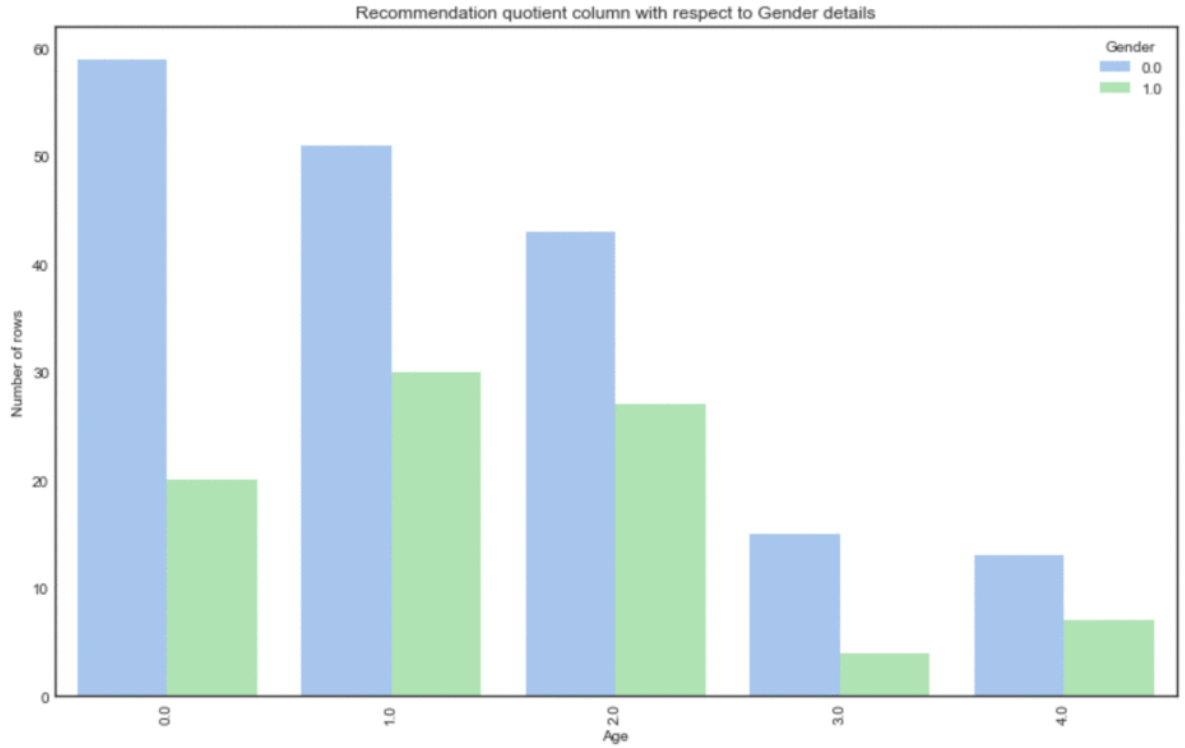
Bivariate Analysis:

I performed bivariate analysis again, this time using count plots and a different hue style. Please see the code and GIF outputs below for further information.

Code:

```
for C in df1:
    if C == "Gender":
        pass
    elif C == "Pin Code":
        pass
    else:
        plt.style.use('seaborn-white')
        plt.figure(figsize=(11,7))
        sns.countplot(x=C, data=df1, hue="Gender")
        plt.title("{} column with respect to Gender details".format(col))
        plt.tight_layout()
        plt.xticks(rotation=90)
        plt.ylabel("Number of rows")
        plt.show()
```

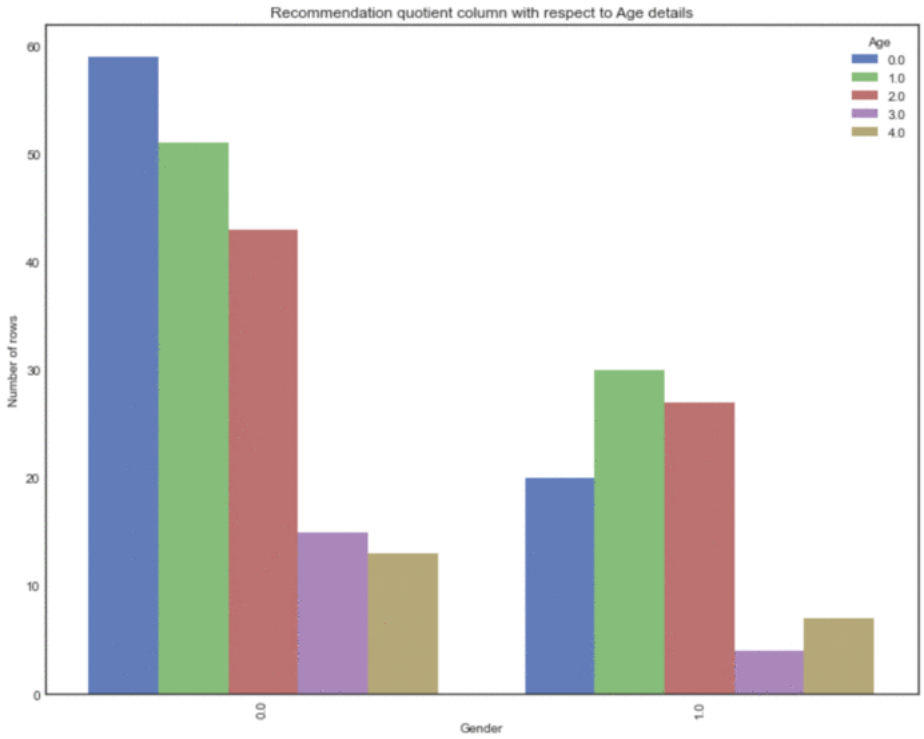
Output:



Code:

```
for C in df1:
    if C == "Age":
        pass
    elif C == "Pin Code":
        pass
    else:
        plt.style.use('seaborn-muted')
        plt.figure(figsize=(11,7))
        sns.countplot(x=C, data=df1, hue="Age")
        plt.title("{} column with respect to Age details".format(col))
        plt.tight_layout()
        plt.xticks(rotation=90)
        plt.ylabel("Number of rows")
        plt.show()
```

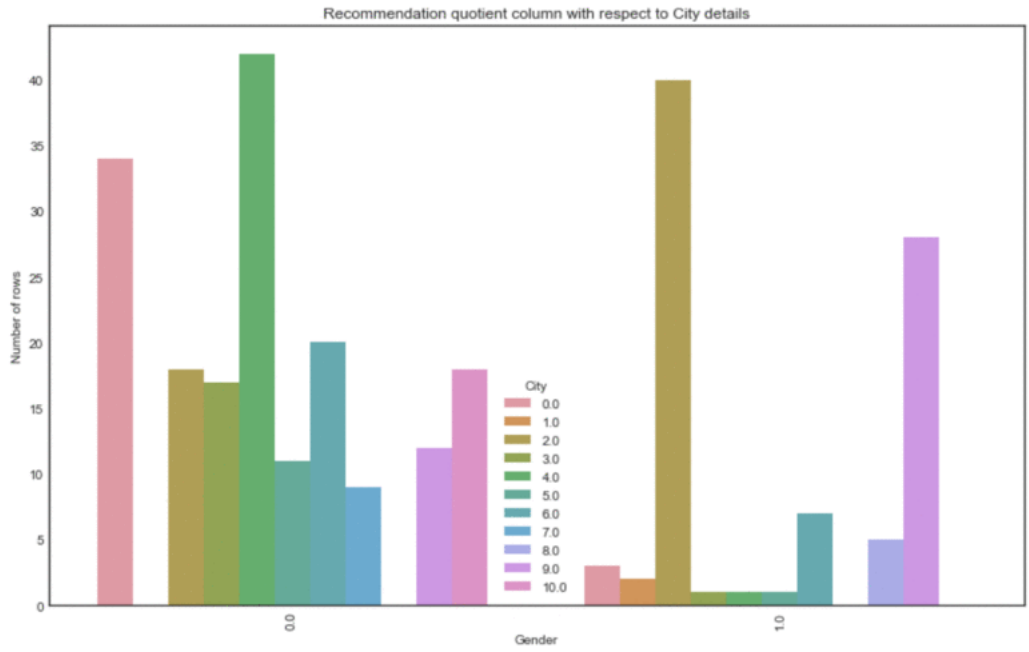
Output:



Input:

```
for C in df1:
    if C == "City":
        pass
    elif C == "Pin Code":
        pass
    else:
        plt.style.use('seaborn-colorblind')
        plt.figure(figsize=(11,7))
        sns.countplot(x=C, data=df1, hue="City")
        plt.title("{} column with respect to City details".format(col))
        plt.tight_layout()
        plt.xticks(rotation=90)
        plt.ylabel("Number of rows")
        plt.show()
```

Output:



Before I could do any kind of multivariate analysis, I had to execute Ordinal Encoding on all of the object datatype columns.

Code:

```
# Ordinal Encoding

oe = OrdinalEncoder()

def ordinal_encode(df, column):
    df[column] = oe.fit_transform(df[column])
    return df

oe_col = df1.columns
df=ordinal_encode(df1, oe_col)
df.head()
```

Using the Ordinal Encoding technique, I was able to convert object datatypes to numeric datatypes.

Multivariable Analysis:

For the purpose of multivariate analysis In my Jupyter Notebook, I used Pandas Profiling. pandas-profiling is a free Python module that allows us to perform exploratory data analysis with just a few lines of code. It creates interactive online reports that may be delivered to anybody, even if they have no programming experience.

It also provides report production for the dataset, with a variety of features and customizations. In other words, pandas-profiling saves us the time and effort of seeing and comprehending each variable's distribution. It creates a report with all of the data in one place.

Although we could scroll through for a detailed analysis report on our dataset while browsing through different tabs, I took a screenshot of the initial output for pandas-profiling. The embedded report can be obtained with a single line of code as follows:

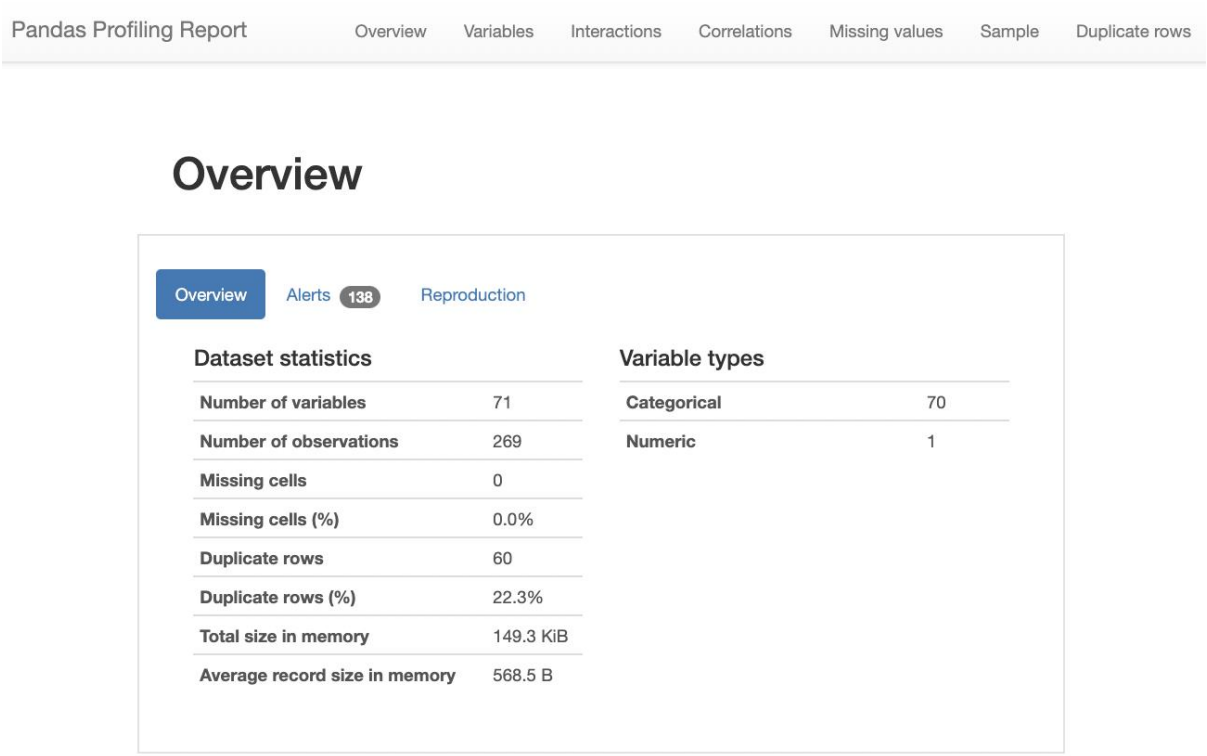
```
pandas_profiling.ProfileReport(df1)
```

Summarize dataset: 100% 85/85 [00:20<00:00, 3.31it/s, Completed]

Generate report structure: 100% 1/1 [00:09<00:00, 9.68s/it]

Render HTML: 100% 1/1 [00:01<00:00, 1.50s/it]

Output:



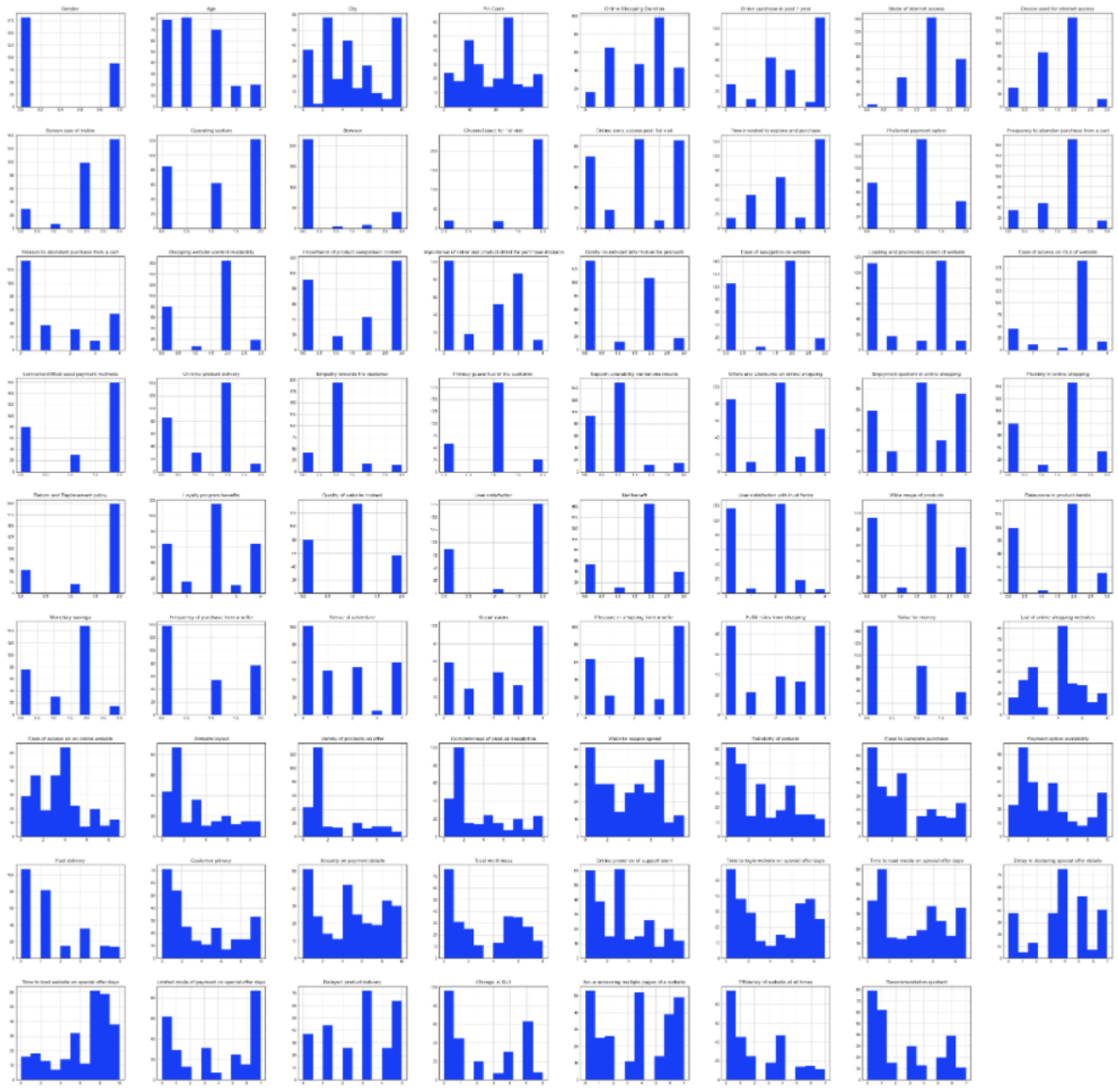
I created a histogram post encoding all of my column data using the pandas-profiling technique.

Input:

```
plt.style.use('seaborn-bright')

df.hist(figsize=(45,45))
plt.show()
```

Output:



Using the correlation values between the dataset columns, I created a heatmap. The positive and negative halves of the correlation information are divided.

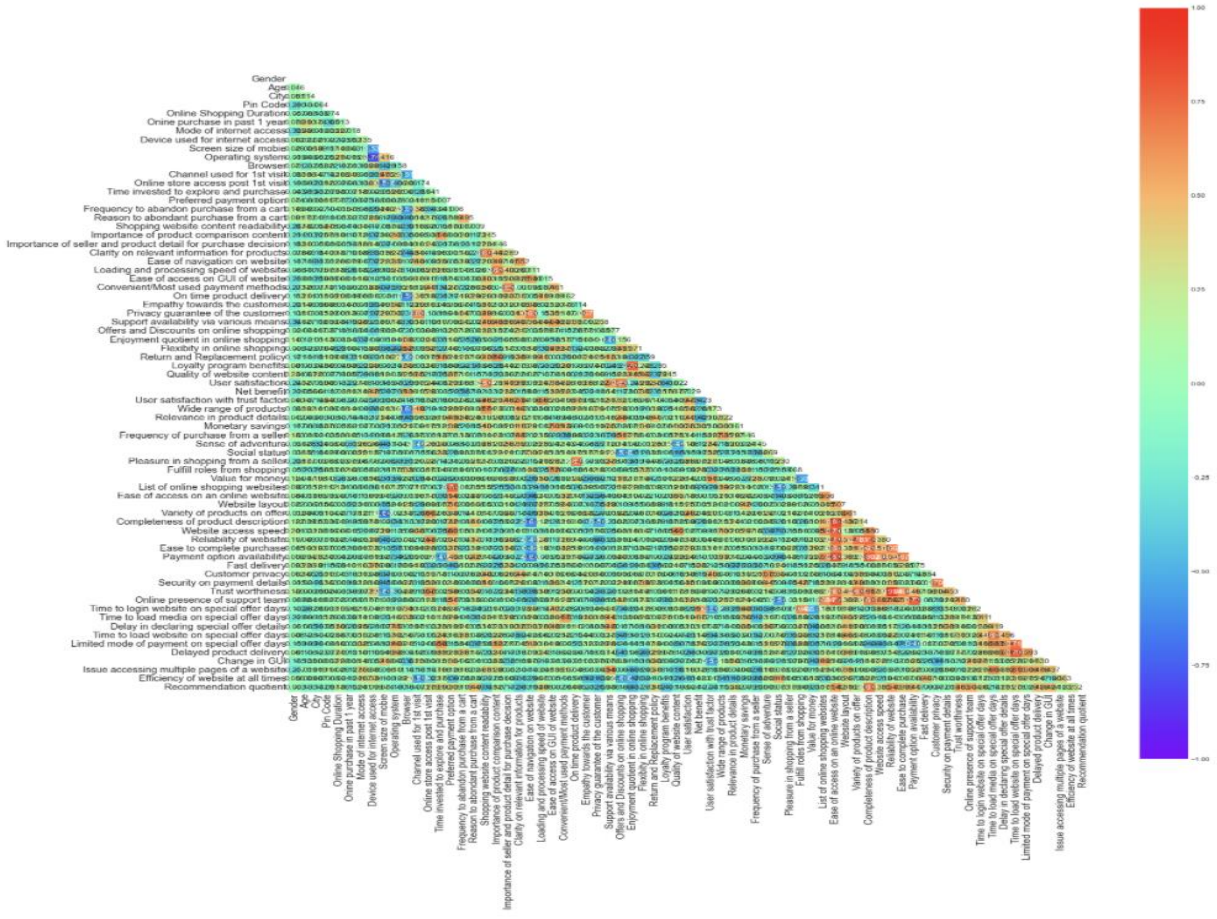
A perfect positive correlation is shown by a correlation of +1, which means that both variables move in the same direction.

A negative correlation of –1 implies a complete negative correlation, which means that when one measure rises, the other falls.

Input:

```
# using heatmap
upper_triangle = np.triu(df1.corr())
plt.figure(figsize=(24,24))
sns.heatmap(df1.corr(), vmin=-1, vmax=1, annot=True, square=True, fmt='0.3f',
            annot_kws={'size':11}, cmap="rainbow", mask=upper_triangle)
plt.xticks(fontsize=14)
plt.yticks(fontsize=14)
plt.show()
```

Output:



We can't see the correlation details in the above heatmap owing to the large number of columns, but we can see the color-coding details and get a signal that there is no multicollinearity worry between the column values.

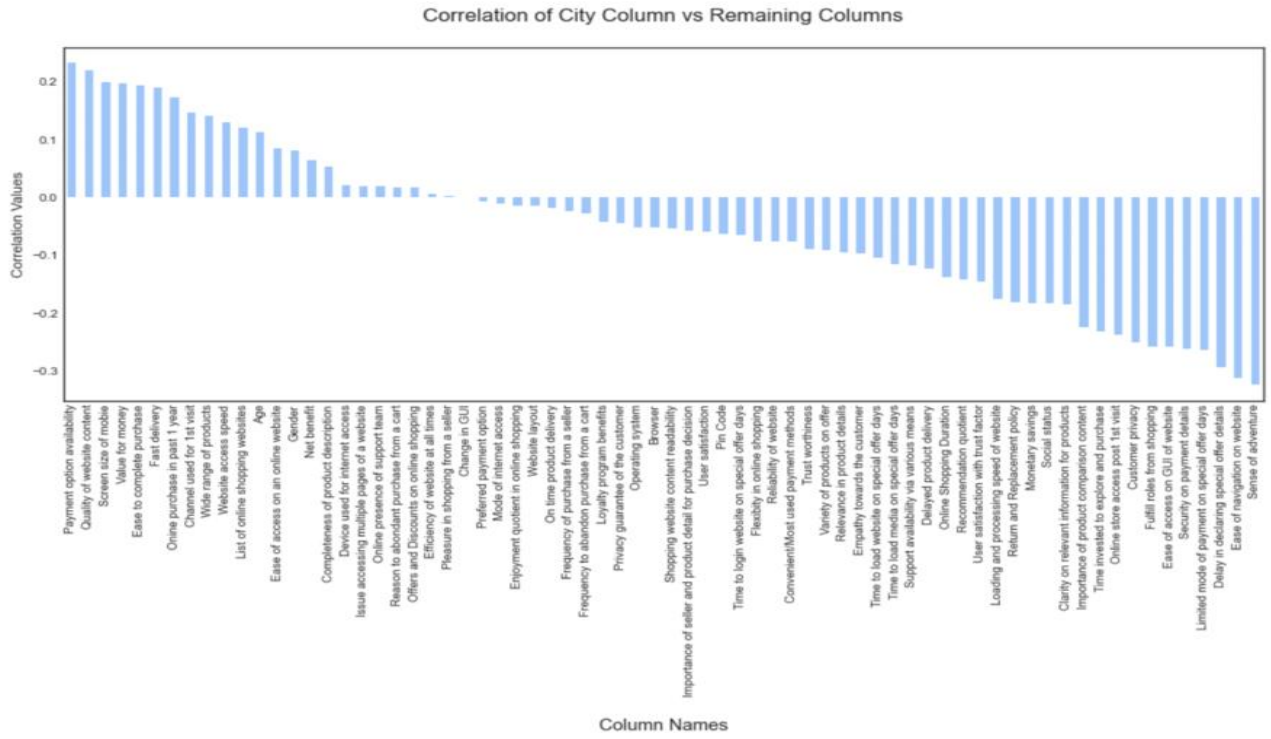
Input:

```
#Comparing the Gender column to the other columns

plt.style.use('seaborn-pastel')

column_names = df1.columns
for col in df1[column_names]:
    df_corr = df1.corr()
    plt.figure(figsize=(16,6))
    df_corr[col].sort_values(ascending=False).drop(col).plot.bar()
    plt.title("Correlation of {} Column vs Remaining Columns\n".format(col), fontsize=16)
    plt.xlabel("\nColumn Names", fontsize=14)
    plt.ylabel("Correlation Values", fontsize=12)
    plt.show()
```

Output:



I created this bar plot for each column versus remaining column to demonstrate the positive and negative correlation data because the heatmap was not apparent in terms of its values.

Inference:

1. **Amazon**

To be improved:

- During promotions, attempt to provide clients with a stress-free shopping experience.
- Provide customers with extra payment alternatives.
- Give the pricing as soon as possible during the offer.
- Shorten the time it takes for things to arrive.

Positive feedback summary:

- It's easy to use, and it's also a fantastic buying site.
- Products are delivered quickly.
- Availability of comprehensive product information.
- Multi-channel online assistance is available.
- Trustworthy website or app, as well as perceived trustworthiness.

2. **Flipkart.com**

To be improved:

- During promotions, attempt to provide clients with a stress-free shopping experience.
- Provide customers with extra payment alternatives.
- Give the pricing as soon as possible during the offer.
- Shorten the time it takes for things to arrive.
- The only difference between Flipkart and Amazon is that their feedbacks are practically identical, with varied percentages.

Positive feedback summary:

- It's easy to use, and it's also a fantastic buying site.
- Products are delivered quickly.
- Availability of comprehensive product information.
- Multi-channel online assistance is available.
- Trustworthy website or app, as well as perceived trustworthiness.
- A wide range of items are available.

3. **Myntra.com**

To be improved:

- During promotions, attempt to provide clients with a stress-free shopping experience.

- During promotions, try to offer the pricing as soon as possible.
- During promotions, reduce the time it takes for things to arrive.

Positive feedback summary:

- It's easy to use and has an excellent webpage.
- A variety of payment alternatives are available.
- Products are delivered more quickly.
- Detailed information on all available goods.
- Trustworthy website or app, as well as perceived trustworthiness.
- A wide range of products to choose from

4. Paytm.com

To be improved:

- During promotions, reduce the time it takes for things to arrive.
- Give the pricing as soon as possible during the offer.
- During promotions, attempt to provide clients with a stress-free shopping experience.
- Price and discount announcements are made late.
- When moving from one page to the next, there is a lot of noise.

Positive feedback summary:

- It's easy to use and has an excellent webpage.
- The ability to finish a transaction in a timely manner.
- About 64% of consumers believe that either the web or the app is trustworthy.
- Paytm is thought to have a wide range of items, according to 20% of consumers.

5. Snapdeal.com

To be improved:

- During promotions, reduce the time it takes for things to arrive.
- Give the pricing as soon as possible during the offer.
- During promotions, attempt to provide clients with a stress-free shopping experience.
- Price and discount announcements are made late.
- No one has indicated a desire to promote Snapdeal to a contact since it has the most number of negative reviews of any other website.

Positive feedback summary:

- It's simple to use.
- A total of 54% of consumers are satisfied with the provision of financial data security.

Conclusion:

The first 47 characteristics, based on general observations, give insights into how the e-tailer is helping and expanding depending on client suggestions. The data revealed which CITY, PIN CODE, AGE, and other variables were utilized most frequently on the web platform. It also revealed that some factors are given less weight in contributing to the success of an e-commerce store; as a result, we could eliminate those factors while keeping all of the important ones, as well as improve on some of the factors that influence online customers' repeat purchase intentions.

Aside from the first 47 elements, the remaining attributes demonstrated which online platform has been used more frequently based on success factors. According to the case study for client activation and retention, Amazon is the most dependable and has met the customer's needs. Data indicated that Flipkart was utilized more for online buying after Amazon.

According to a case study of Indian e-commerce clients, Amazon and Flipkart are the most frequently utilized for online shopping and are the most frequently suggested by friends. So, based on the research findings, Amazon and Flipkart are e-commerce platforms that combine utilitarian and hedonistic values to favorably influence repeat purchase intention (loyalty).

Future work:

- I'll have to do some preparation with the data, such as utilizing scaling algorithms.
- I'm not going to bother about reducing outliers or skewness because the dataset contains largely categorical data.
- Unsupervised machine learning models must be built.

• Will need to double-check the specifics of the clustering or association method that may be employed on the dataset.

• K-means clustering, k-nearest neighbours for unsupervised machine learning, hierarchal clustering, apriori algorithm, and neural networks are some of the methods I plan to work on.