EXPLORATORY DATA ANALYSIS OF ONLINE SHOPPING TRENDS AMONG DIFFERENT DEMOGRAPHICS – POST COVID ERA

BY:

ISOLA, K.O.

(ID: 22185897)

A dissertation submitted as part of the requirements for the degree of Msc. Computer Science at the School of Computing, Engineering, and Built Environment, Birmingham City University, Birmingham, UK

SEPTEMBER, 2024

SUPERVISED BY:

Dr. ADEWUYI, Anuoluwapo Amarachukwu

ABSTRACT

The COVID - 19 era has notably affected the behaviourial pattern of consumers across different demographics which also brought about a large demand in online shopping. This research is based on exploratory data analysis of online shopping trends among different demographic in the post-covid 19 era with the focus on noting the key factors influencing their shopping trends. The main research question that this study focused on is to note the different factors that influence the online shopping patterns different individuals Post COVID - 19 era.

This research work provides background research and comprehensive literature review about consumer behaviours, online shopping of consumers before and during the COVID pandemic, the factors that contributed to such behaviours and the effects of COVID on businesses. The research analysis employed the use of a secondary dataset that checks consumer behaviour, frequency of shopping, preferences and spending habit and also identifies the importance of demographic factors, technological advancement, post covid-19 psychological shift and social influences.

The major tool used in carrying out this analysis is the Google Colab, which is the short form of Google Colaboratory, it is a tool majorly used by data analyst because of its unique features and advantages over other tools such as being able to connect the worksheet to a Google drive and being able to access the document from anywhere provided the availability of internet connect. Univariate and Bivariate analysis were carried out on the dataset to understand the frequency of distribution for each variable and also to understand the relationship between two variables. This relationship was evaluated using some statistical tools like Cross tabulation, Chi - Squared Test, p - value that tells if there is significant difference between both variables.

The research also suggests that some factors that plays important role in posing a shopping pattern may include advancement in payment technology, social media adverts, and multiple convenience of e-commerce platforms, personalized marketing and easy accessibility to social media platforms. Additionally, with the use of EDA techniques, the research suggest to business owners how they can adapt and adjust to the trends to developing trends by creating more enticing offers which imposes more attention to online shopping through creating better online shopping experiences.

ACKNOWLEDGMENT

My appreciation goes to my Heavenly Father, the Giver of Life, my source of Wisdom and Knowledge till this day. I also appreciate my Husband, my Parents, my Siblings and my Friends for their unwavering support throughout this journey to achieve this milestone. I also want to use this medium to appreciate my ever caring and supportive Supervisor, Dr. Adewuyi Anuoluwapo Amarachukwu for his immense help, mentoring and guidance to see this project comes to reality, and I pray that my Heavenly Father would reward his labour of Love, Amen.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE
INTRODUCTION

Since the development of E-commerce, there has been a rapid change in consumer buying behaviour and the pattern of this rapid change is what this research paper tends to discover as the major factors resulting in such changes. The act of buying products or services via the Internet is called online shopping or E-shopping which is a form of E-commerce. People have imbibed this character because of different reasons such as online discounts, enough time to make decisions, and a busy schedule amongst many others (Dr. V. Dheenadhayalan, 2021).

Statistically, the emergence and penetration of e-commerce in various regions in the United State America has experienced a wide and steady growth with an increased rate from 55% in 2017 to 81% in 2023, indicating that many people in the United State shop online. Globally, United State of America is said to hold one of the largest e-commerce, hereby influencing the scope of the market at large in various continent, leaving Amazon, Target, Walmart, Apple, Shein, Aliexpress and Costco amongst many others as the top rising ecommerce in the state (ecommercedb.com, 2024).

The COVID-19 was declared a pandemic by the World Health Organization on March 11, 2020, which gradually crept in various part of the world (Cucinotta and Vanelli, 2020). Following the lockdown that was experienced globally (Sirimanne, 2021), online sales grew widely in various western countries such as China by 5.2% between 2019 and 2020, 81% of Nigerians also increased their shopping online during the pandemic (Saleh, 2022) while in the United State of America experienced an increase in online shopping by 25% between 2019 and 2020 (Statista, 2021).

We are four years passed COVID-19 era and variants may continue to emerge. It is also fair to note that all restrictions are being fully lifted and shops are in operation again globally, however, consumers seems to be satisfied with their form of shopping online years after the pandemic, as continuous improved in technology has also influenced this decision widely (Cramer-Flood, 2022).

1.1 Background

Online shopping, facilitated by the rise of E-commerce, has significantly boosted sales for many businesses worldwide. Moreover, it has also brought about notable changes in consumer market behaviour across various regions. To fully comprehend the driving force behind this trend, it is crucial to understand the reasons why individuals choose this shopping method. The consumer market is composed of different age groups, male and female with younger people having a good population in the market space. Younger people are very vast in using technology so social media and peer groups contribute greatly to their shopping habits. The South Atlantic region of United State of America as a case study is highly populated with an estimated growing population of over sixty-eight million people as at 2023 and has an improved e-commerce infrastructure across the region which consist of Maryland, Washington DC, Delaware, Florida, Georgia, Virginia, South Carolina, North Carolina, and West Virginia (fred.stlouisfed.org, 2023).

1.2 Research Question

The major research question is: What are the different factors that influence the online shopping pattern of different people – Post – COVID Era?

1.3 Aims and Objectives

Aim: This research project aims to discover the online shopping trends among different people – Post – COVID era.

Objectives: The objectives of this research project are:

- Survey previous work on EDA on online shopping among young people
- Identify suitable datasets for this analysis
- Conduct Exploratory Data Analysis (EDA) on the dataset using suitable data analysis tools
- Evaluate and discuss the Post-COVID trends/patterns from the analyses and results.

1.4 Project Scope

The dataset employed for this research is a secondary dataset from collected recently for a study in United Kingdom and United State of America. The dataset will be compared and used where necessary.

1.5 Project Limitation

This research will employ the use of secondary data as there are some shortcomings as to getting primary data for the purpose of this work. The non-availability of a comprehensive secondary dataset for the United Kingdom cannot be undermined for achieving a reasonable and full picture in unveiling the trend pattern in the country. Unravelling the pattern of online shopping in the United States of America and any other similar western countries shows that similar policies were governing the operation of businesses globally during the COVID-19 especially in western countries where the pandemic is severe; such as the social distancing policy, the lockdown policy, quarantine policy, health concerns, personal protective measures, lock-up of some businesses that are not essentials to human and so many more will propel us to compare few dataset. The absence of funding for this research is another factor that must be noted which affects the scope of this research work as funds are needed to generate primary data from desired locations, mobilize resources needed and purchase some required materials.

1.6 Justification

The outbreak of COVID – 19 which was eventually declared a global pandemic by the World Health Organization (WHO) in March, 2020, has prompted the government globally (especially in the western countries where the death poll is significantly high) to respond in a significant way by enforcing some policies and restrictions that guarantee safety of individuals (Cucinotta and Vanelli, 2020). The policies and restrictions affected the various economic sectors globally not excluding the commerce, which includes closing of schools at all levels, travelling restrictions, public gathering and meeting ban, emergency investment in healthcare officials and facilities, stoppage of some business operations due to the sit at home policy, various social welfare provision amongst others.

People have tended to be forced to delve to shopping online as result of the development which restricted the on-site operations of some businesses. The pandemic accelerated the shift to online shopping behaviours as most sales were forced to happen online following the trends, the policies, the restrictions and various factors that came with the outbreak globally (Dalgleish, 2020). It is expedient to further research the continued effect years after the global experiences of the pandemic on businesses which this project work will focus on.

1.7 Assumption

The research study is subjected to the assumption that there are no comprehensive dataset on any previous study around this research in the United Kingdom and as such, various aligning dataset will be employed for this research study. In addition to this, it is also believed that the pattern of government policies, and other activities were the same globally especially in the western countries during COVID and as such, the lifting of the various policies were also globally, as countries began to fight the course and being free of the virus. This research will employ these bases to be true to carry out a further research on the effect of the pandemic to online shopping post COVID era.

CHAPTER TWO
LITERATURE REVIEW

2.1 The Emergence of E – Commerce

The oldest form of commerce is the “Barter System” which is a system of giving goods and services in exchange for other goods or services with the same value without having to use any form of legal tender (money); this system was based on mutual agreement between both parties (Kenton, 2019).

E-commerce then began with the emergence and development of the Internet and it involved the buying and selling of goods and services over the Internet and could be in different forms such as B2B (Business to Business), B2C (Business to Customers), B2G (Business to Government) amongst others. According to Coppola (2021), statistics show that the retail e-commerce sales in 2024 amounted to about 6.3 trillion US dollars and this resulted from the use and easy access to global internet over the world with over 5 billion internet users, hence the population of online purchases increases steadily.

E-commerce could be enhanced through various platforms such as mobile apps, online marketplaces, or websites (Hayes and Downie, 2024). It helps consumers to increase their purchase opportunities where they can have more time to see and select from a list of various goods and services on the internet and the fact that electronic payment method can be done easily is also an added advantage of buying online (Sree Saraswathi and College, 2020).

2.2 Understanding Consumer Behaviour

According to Saleem, Khalid, and Sadiq (2022), consumerism is the never-ending pattern from local trade in limited areas to the era of industrialisation and large-scale shopping. There has been a transformation in consumer behaviour where buying basic needs and lifestyles to luxuries, everything has now started online. Over 5.4 billion people are connected to the Internet which is a nowadays effective medium for the sale and purchase of goods and services. The main reason why consumers prefer shopping online is because of the ease of ordering and also fast delivery.

Consumer buying behaviour is influenced by so many factors which this research paper aims to discover ranging from the brand, price, convenience, other consumer reviews, and recommendations amongst others (Sree Saraswathi and College, 2020). According to Donthu and Gustafsson (2020), the COVID–19 pandemic period was an eye-opener for most people because people couldn’t move from one place to another as usual, they were compelled to stay at home, keeping themselves and their families safe from the COVID – 19 outbreak, people were getting bored and tired of the same routine for a long time. Due to this, they were able to achieve numerous activities online ranging from learning a course to baking and also online shopping wasn’t left out of these activities as there was high usage of internet and social media.

2.3 Exploratory Data Analysis

Exploratory Data Analysis (EDA) can be defined as a mechanism mainly used by data scientists to study and understand different datasets and also summarize their specific features, and this is majorly done using data visualisation methods (IBM, 2020). According to Santos (2023), it involves in-depth examination and characterization of data in order to discover more about the features, probable outliers, and hidden trends and correlations.

Exploratory Data Analysis (EDA) also helps to summarize data by understanding their main characteristics and visualizing it with right pictorial representations. It helps to describe the datasets, the rows and columns, the missing data, the data types and the review. It also involves cleaning the data to discover and handle missing or invalid data types (Sahoo et al., 2019).

2.4 Advantages of Exploratory Data Analysis

Exploratory Data Analysis has various advantages in analysing data and of which one of it that it helps to look at data before making assumptions, having a better understanding about the pattern of the data, detect any known error and also give a detailed relationship among variables (IBM, 2020). It can also be employed by both business owner and stakeholders to produce valid and applicable business goals and result through forecasting and projections. This result can be furthered used by data scientist for more sophisticated data analysis and modelling (IBM, 2020).

According to Ahmed (2022), EDA helps analysts to have better understanding of different variables and features of a dataset, and then different information can be deduced from the dataset. It also helps in discovering outliers, missing values and any anomaly in the dataset; as this is a very important stage before analysing any data. It also helps to discover hidden patterns which can be displayed graphically, for example by using bar charts, pie charts, scattered plots or histograms. It has a great advantage of helping organizations to understand their data, to efficiently analyse, facilitate report writing and support decision making.

2.5 Types of Exploratory Data Analysis

According to the research, (IBM, 2020) and (Nguyen, 2023), Exploratory Data Analysis is made up of four primary types which include Univariate Non-Graphical, Univariate Graphical, Multivariate Non – Graphical and Multivariate Graphical.

- ***Univariate Non-Graphical:*** This type of EDA deals with just one variable for analysis. The unit variables are being used to analyse the trend of an item that it is about. This is also the least form that a data can exist as it does not deal with a relationship. It can be used to understand the dispersion, median, range and other central tendencies which is also known as Descriptive Statistics.

- ***Univariate Graphical***: which is a graphical representation of a unit variable with relationship with its frequency i.e. it can be used to show the relationship between one variable and its dispersion graphically, such as the frequency and the distribution of the data using bar charts, box plots, or frequency tables.

- ***Multivariate non-Graphical:*** It is a type of analysis that has to do with two or more variables of a data. This relationship could be expressed in a tabular form and can be interpreted statistically.

- ***Multivariate Graphical:*** It uses a graphical representation to analyse two more variables in a given set of data. This analysis can also be expressed in form of a chart. It is mostly used for a complex analysis such as a grouped set of data. It can also be employed where a given set of data are scattered. Examples include scatter plot, heat map, bubbles charts.

2.6 Tools used for Exploratory Data Analysis

There are various tools used to implement exploratory data analysis such as Python, R Programming, Matlab. These tools vary in their use and also their efficiency (IBM, 2020), (Trantor, 2024) and (Nguyen, 2023).

- **Python:** It is majorly used for different tasks in exploratory data analysis such as describing the datasets, fetching out missing values, detecting outliers and also suggest decisions based on graphical representations. Libraries such as Pandas (provides data structures), Numpy (helps in large multi-dimensional arrays), Seaborn (a data visualising library based on Matplotlib), Matplotlib (used for designing static and interactive visualizations) and Scikit – Learn (helps in data pre-processing, dimensionality reduction and clustering) help yield straight forward answers, and they can also be implemented in the python tool.

- **R Programming:** This is a programming language and also a free software environment that can be used for statistical computing and graphics. It is widely used to develop statistical observations and data analysis. It makes use of libraries such as Leaflet, tidyr (used for tidying and reshaping data), Lattice, ggplot, SmartEDA, GGally (provides functions for visualizing multivariate data), Data Explorer to get more efficient results from the data analysis.

- **Visualization Tools:** According to Halladay (2023), they are software tools that help to produce graphical representation of data by processing the data and converting them to visual graphs that can be easily understood. Examples include Tableau (helps to create interactive dashboards and visualizations), Power BI (also helps in providing data visualization), Plotly (a charting library that supports visualizations in R, Python and JavaScript)

2.7 Impacts of Online Shopping

Online shopping has helped the normal retail business migrate into the global world. It has however made effective impacts and also helped many businesses to be more useful and relevant for themselves in terms of being able to make a living and also major profits. Online shopping has greatly transformed the retail business and also consumer behaviour in different ways such as:

- **Economic Impact:** According to Grewal, Roggeveen and Nordfalt (2017), the impact of e-commerce is so significant such that walk-in-store had to imbibe the modern culture into their business or either stays competitive. The shift is so exponential such that many stores generate income via e-commerce as compare to the walk-in-store. This impact may imply a growth in the e-commerce market while a significant impact on retail employment can be influenced where new job opportunities are created in terms of logistics, warehousing, and IT which have also caused a shift in skill acquisition or informing the walk-in-store owners to reskill.

- **Consumer Behaviour:** Zheng et al (2013) cannot overemphasize the relevance of e-commerce on consumers' behaviour. The ability for consumers to make their shopping online from their comfort zone (home) brings about conveniences and accessibility to buying their stuff which brings about a change in their shopping habit as it also helps consumers to go about their busy lifestyle and be assured that their shopping is being taken care of from the comfort of their home. It is also known that consumers are more likely to leverage on their ability to access various products at their own pace and also have ease in their online transaction also affects the purchasing pattern which is more of a personalised marketing strategy (Huang and Benyoucef, 2013).

- **Environmental Impact:** Edward, McKinnon and Cullinane (2010), discussed about the drastic minimal reduction in the need for shoppers to travel to a physical store, however, there is further need as to how goods and materials are being packaged and the emission of carbon by logistics can raise a concern with environmental sustainability. Also waste generated has also increased through online shopping, especially in respect to plastic and cardboards due to the need for packaging. This is well highlighted by Williams and Wikstrom, (2011).

- **Social and Cultural Impact:** The research done by Blazquez (2014) stated that there has been a shift in social interactions which caused a reduction in face – to – face interactions and also led to more confined consumer experience. Also, customers have gained more knowledge about product information, product price and comparism and product reviews which is an important factor that influences their final purchase decisions.

2.8 Impact of Technology on Shopping Behaviour

According to Goleckyte (2023), as technology advances in the world, the shopping habits of people have significantly increased. They have uninterrupted access to different types of websites that offer different products and services at the snap of their fingers, all credit to E-commerce; this in turn reduces the retail business place and also gives business owners more knowledge about their preferences and needs so they can know how to improve their products and marketing efforts and also make more profits to their businesses.

Recent studies show that app 96% of children and teenagers in the range of 3 – 17 in the United Kingdom, including London either own or have access to smartphones or devices (Anderson, Faverio, and Gottfried, 2023). According to a research work carried out by Ibrahim et al., (2023), there was a survey that found that young people enjoy purchasing online because it is easy, efficient, cheap, and more convenient for them, the fact that they could do from their comfort zones and still get what they want.

2.9 Importance of Conducting Analysis on Shopping Trends

There are so many beneficial reasons for analysing shopping trends such as market insights which are for helping businesses to identify the products and services needed by this age group and also improve their marketing strategies to meet their needs (McKinsey & Company, 2021). It also helps businesses to make certain and direct decisions such as promotions and discounts once they can understand the factors behind teenagers' online shopping behaviour. According to Kotler et al. (2016), strategic decision-making based on consumers' needs and preferences is very essential for the growth and success of a business amongst competitors.

Another importance of conducting analysis on shopping trends is the ability for businesses to be able to predict what the future trends would look like so they can start working towards achieving their goals by that set time. This gives them an advantage over their market competitors as they can make great use of the opportunities they have and also prepare ahead of time.

2.10 Factors That Tend to Influence Online Shopping Behaviour

Numerous factors influence consumer online shopping behaviour. For example, during the COVID–19 lockdown, people were at home and couldn't move in and out of their homes easily and this resulted into boredom for some people, while for some other people, they were able to maximize the period by going online to feed their eyes, do easy and seamless shopping and also learn different things online within their pace. Some of the factors that influence consumer online shopping are the type of products that are influenced by peer groups, social media ads, and whatever is trending at the moment.

Another factor that influences this behaviour is the fact that teenagers can work and make money for themselves, parental factors also, and additionally the advantage of making online payments easily. They can shop online frequently and they can also do this at any time, any location, and from their comfort spaces (Gulfraz et al., 2022).

Also, social media is another major factor that influences online shopping behaviour. There are different social media platforms ranging from TikTok to Instagram to YouTube and the likes; these platforms not only entertain them but also advertise different products and brands, giving them information about the products and how they can get them even at discounted prices. They also get influenced by the recommendations from their peer groups and maybe an influencer or role model they admire so much online (Duffett, 2017).

2.11 Effects of COVID-19 on shopping trends and consumer behaviour

COVID-19 Pandemic made a significant change in consumer behaviour worldwide in different sectors including the online shopping sector. Prior to the COVID-19 pandemic, online shopping was a normal activity globally but was more utilised in developed countries like US and UK than developing countries. In the United States, E-commerce accounted for 16% of the total retail and 19% for the United Kingdom (Statista, 2019). According to Olatokun & Bankole (2011), the utilization of e-commerce in developing countries, for example, Nigeria seems slower due to different factors like limited internet access, poor logistic infrastructure and lack of trust of online payment schemes.

When the COVID-19 pandemic started, there were government policies stated to control the spread of the virus and so resulting to forceful lockdowns and social distancing in different organizations, which is also one of the reasons consumers started shopping online. In 2020, the e-commerce sales of the United States increased by 44% amounting to $861 billion (Eight hundred and sixty-one billion US dollars) from their previous sales of $598 billion (Five hundred and ninety-eight billion US dollars) (US Department of Commerce, 2021). There was also a significant increase in the e-commerce sales of the United Kingdom which grew by 36% in 2020 (Office for National Statistics, 2021). The growth of e-commerce was more pronounced in Nigeria, as there was increase in necessities which made consumers turn to online shopping, to back this up; there was a report from Jumia that their transactions increased during the lockdown periods by 50% (Jumia, 2020).

One of the impacts of COVID-19 was the changes in consumer preference. Prior the lockdown, major online shopping were for fashion, entertainment, electronics and the likes but during the lockdown, as there were policies of staying indoors, families needed to get groceries for their daily life and so this led to a dynamic shift in consumer behaviour. It was recorded that in US and UK, there was a major increase in grocery shopping online with UK grocery sales increasing by 76% in year 2020 (Kantar, 2021).

Another impact of the COVID-19 was consumer trust and security concerns. Before the pandemic, the level of trust consumers had varied across regions especially in developing countries like Nigeria, but in US and UK, there were high levels of consumer trust established because of their governmental policies and also having an efficient payment system. But in Nigeria, consumers were more concerned about not being a victim of fraud, and because of this, they would rather walk into the store and see what they want to buy for them and ensure payment is accurate (Ayo, 2011). But during the lockdown, they had no choice and were forced to get used to the order of the day and this helped to build trust between vendors and consumers since that time till now.

Additionally, a major impact of COVID-19 was the adjustment in payment methods. In US and UK, they were already used to making payment through digital wallets and contactless payments, and during this pandemic, there was an increase in that sector which aligned with the global trend towards cashless transactions (Accenture, 2020). However, in Nigeria, the reverse was the case because consumers were used to making payments over the counter or payments on delivery of items due to trust issues and also internet problems leading to delay in digital payments (Ayo, 2011). But the pandemic period helped in increasing trust between vendors and consumers, also the digital payment service providers also improved their services to ensure smooth and efficient payment online without hitches.

Amidst all these factors, there were major similarities between developed countries and developing countries in the sense that there was universal adaptability for all consumers during the pandemic, there were frequent online shopping during that time and even after because it was something every consumer could do from their comfort spaces and still have trust of their purchase. There were also improvements in digital infrastructures worldwide, improved logistics, and improved internet access.

CHAPTER THREE
RESEARCH METHODOLOGY

This chapter explains the research methodology utilized to explore the online shopping trends among people in the post – COVID era. The unexpected global impact of the pandemic has immensely affected consumer behaviours, especially among young people who are vast, adaptable and technologically savvy. This research aims to carry out a quantitative analysis on an existing dataset to have more understanding and provide comprehensive analysis on these shifts, focusing on patterns, preferences and the different influencing factors of online shopping among people.

3.1 Research Design

The research study is ideally suited to a quantitative research methodology for several compelling reasons. Primarily, this approach enables the objective measurement and statistical analysis of numerically quantifiable data, which is crucial for identifying trends, patterns, and correlations in large datasets. Such an approach is particularly relevant when examining online shopping behaviors among young people, as it can efficiently summarize the main characteristics of the data through statistical graphics and other visualization methods.

Additionally, quantitative methods excel in managing large sample sizes, enhancing the generalizability of the findings across a diverse population. This methodology not only supports robust, scalable research through automated data collection and analysis tools but also ensures precision and reliability through controlled measurements and established statistical techniques. The ability to handle vast amounts of secondary data from various sources like e-commerce platforms and consumer surveys makes it a potent tool for analysing trends and behavioral shifts in the dynamic domain of online shopping.

In essence, adopting a quantitative framework for this study allows for effective hypothesis testing and the exploration of relationships between various demographic factors such as age and profession. It provides a solid empirical foundation to affirm or reject hypotheses about changes in online shopping behaviors post-COVID, making it an invaluable approach for deriving accurate, scalable, and actionable insights about the shopping preferences and trends among young consumers in the contemporary digital landscape.

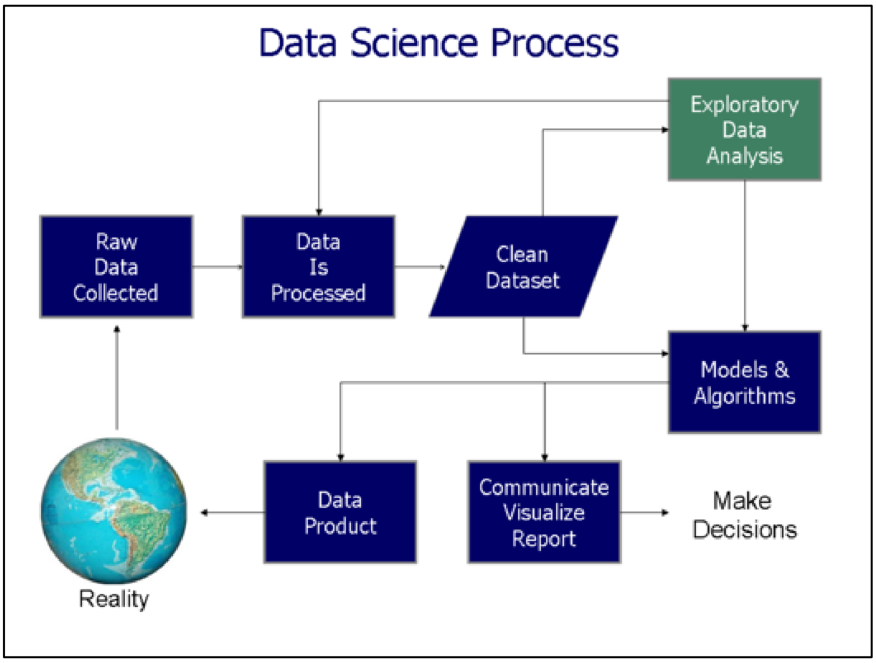


Figure 3.1: General Methodology of Data Science

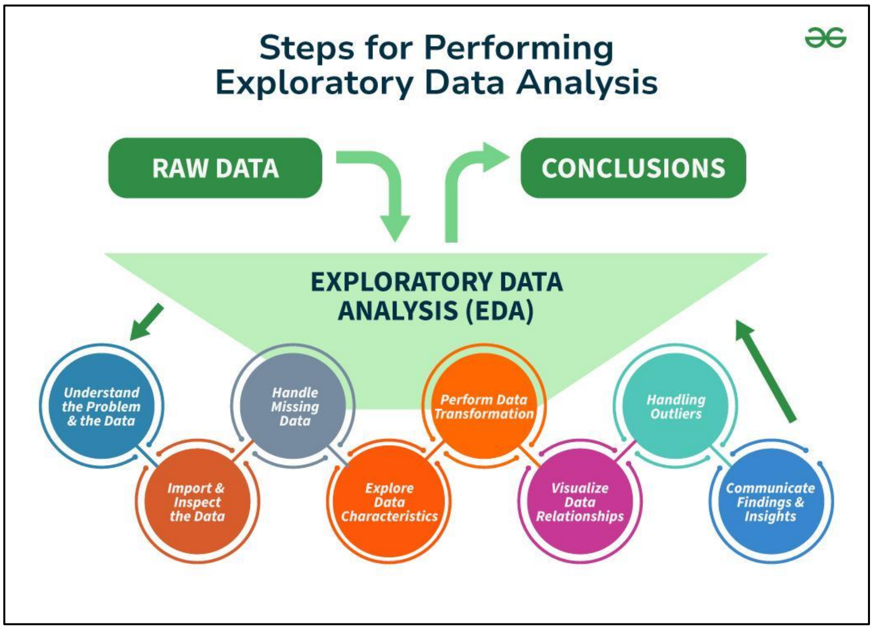


Figure 3.2: General Methodology for Exploratory Data Analysis

3.2 Data Collection

This research work uses a quantitative secondary dataset that contains the responses of different demographics about their preference for online shopping during and after the COVID-19 pandemic. The dataset was obtained from Youdata (*https://www.youdata.ai/m/datasets/662de3c57bb79dfcf0857bef*) which consists of 16 unique features: Timestamp, Name, Profession, Age, Do you prefer online over offline shopping during the current Covid-19 crises?, Do you think that scope for Online shopping has increased in this pandemic period?, Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?, Do you tend to spend more money if you are shopping online or in store?, How Often do you shop Online?, How much time do you spend in every visit?, What do you usually purchase online?, What do you usually purchase offline?, What is your satisfaction level when you do online shopping?, What is your satisfaction level when you do offline shopping?, What made you choose online shopping over offline shopping during the current crises? and What made you choose offline shopping over online shopping during the current crises? All these features were analysed and evaluated for better understanding of the shopping trend after the COVID-19 pandemic. This analysis and evaluation was done using the Google Colab notebook to have better visualization and understanding.

The major tool or interface that was used for this data analysis is Google Colab which is the short form of Google Colaboratory. It is an interface that helps in writing codes, running codes and sharing codes stored in the Google Drive. It is usually referred to as a “Notebook” because it consists of codes that work with texts, images, figures, graphs, tables and equations. Google Colab is majorly used by Data Scientists and Data Analysts due to its advantages and special features. Some of the advantages of Google Colaboratory are that users can access this notebook from a web browser without the need to install on a system, they already come with some pre-installed and necessary libraries that would be needed for the codes to run efficiently e.g. Pandas, Scikit, Numpy, Seaborn etc. and it can be easily integrated with Google Drive and Git hub whereby users can access the file from any location and has low risk of data loss (GeeksforGeeks, 2023).

To view this dataset, the following libraries were imported:

```
from google.colab import files
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Figure 3.3: Code for Importing needed Libraries

```
uploaded = files.upload()

Choose files  No file chosen          Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving Impact of COVID dataset.csv to Impact of COVID dataset.csv
```

```
covid_data = pd.read_csv('Impact of COVID dataset.csv')
covid_data.head()
```

Figure 3.4: Code Implemented to view the dataset

	Timestamp	Name	Profession	Age	Do you prefer online over offline shopping during the current Covid-19 crises?	Do you think that scope for Online shopping has increased in this pandemic period?	Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?	Do you tend to spend more money if you are shopping online or in store?	How Often do you shop Online ?	How much time do you spend in every visit?	What do you usually purchase online?	What do you usually purchase offline?	What is your satisfaction level when you do online shopping?	What is your satisfaction level when you do offline shopping?	What made you choose online shopping over offline shopping during the current crises?	What made you choose offline shopping over online shopping during the current crises?
0	2021/12/11 5:57:02 am GMT-8	Divyangi Hattarki	Student	18-24	Yes	Yes	Yes	Online	Once in 3-6 Months	1-3 Hours	Clothing	Groceries	Very satisfied	Satisfied	It is easier	Like to be hands on
1	2021/12/11 5:57:15 am GMT-8	Shreya	Student	18-24	Yes	Yes	No	Online	Daily	10-30 Minutes	Electronics	Electronics	Dissatisfied	Satisfied	More products online	Like in store offers only
2	2021/12/11 6:12:08 am GMT-8	varun	Student	18-24	Yes	Yes	Yes	In-Store	Monthly	10-30 Minutes	Clothing	Electronics	Neither satisfied nor dissatisfied	Satisfied	Time efficiency	Like to be hands on
3	2021/12/11 6:18:21 am GMT-8	Charmil jain	Student	18-24	Yes	Yes	Yes	Online	Once in 3-6 Months	10-30 Minutes	Electronics	Groceries	Very satisfied	Very satisfied	Online offers more discounts	Reliable
4	2021/12/11 6:50:21 am GMT-8	Vidisha singh	Student	18-24	No	Yes	Yes	In-Store	Once in 3-6 Months	1-3 Hours	Electronics	Clothing	Satisfied	Satisfied	Safer	Reliable

Figure 3.5: First five rows of the dataset

3.3 Data Pre-processing

3.3.1 Basic Info of the dataset

The basic information of the dataset explains the data type of each column if its either a string or an integer, it tells the total entries in the dataset and also indicates if there is a missing value in the dataset or not. The following codes were imputed to get the basic info of the dataset.

```
covid_data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 16 columns):
 #   Column                                                                                     Non-Null Count  Dtype
---  ------                                                                                     --------------  -----
 0   Timestamp                                                                                  100 non-null    object
 1   Name                                                                                       100 non-null    object
 2   Profession                                                                                 100 non-null    object
 3   Age                                                                                        100 non-null    object
 4   Do you prefer online over offline shopping during the current Covid-19 crises?             100 non-null    object
 5   Do you think that scope for Online shopping has increased in this pandemic period?         100 non-null    object
 6   Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?  100 non-null    object
 7   Do you tend to spend more money if you are shopping online or in store?                    100 non-null    object
 8   How Often do you shop Online ?                                                             100 non-null    object
 9   How much time do you spend in every visit?                                                 100 non-null    object
 10  What do you usually purchase online?                                                       100 non-null    object
 11  What do you usually purchase offline?                                                      100 non-null    object
 12  What is your satisfaction level when you do online shopping?                               100 non-null    object
 13  What is your satisfaction level when you do offline shopping?                              100 non-null    object
 14  What made you choose online shopping over offline shopping during the current crises?      100 non-null    object
 15  What made you choose offline shopping over online shopping during the current crises?      100 non-null    object
dtypes: object(16)
memory usage: 12.6+ KB
```

Figure 3.6: Code for viewing the basic info of the dataset

3.3.2 Data Cleaning and Filtering

This entails thorough checking for any missing values or duplicate entries in the dataset and also filtering the column and rows to the required ones, that is, either removing unnecessary column and rows or adding a new column or entry. The following codes were imputed for thorough data cleaning and filtering:

```
covid_data.isna().sum()

Timestamp                                                                                                       0
Name                                                                                                            0
Profession                                                                                                      0
Age                                                                                                             0
Do you prefer online over offline shopping during the current Covid-19 crises?                                  0
Do you think that scope for Online shopping has increased in this pandemic period?                              0
Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?    0
Do you tend to spend more money if you are shopping online or in store?                                         0
How Often do you shop Online ?                                                                                  0
How much time do you spend in every visit?                                                                      0
What do you usually purchase online?                                                                            0
What do you usually purchase offline?                                                                           0
What is your satisfaction level when you do online shopping?                                                    0
What is your satisfaction level when you do offline shopping?                                                   0
What made you choose online shopping over offline shopping during the current crises?                           0
What made you choose offline shopping over online shopping during the current crises?                           0
dtype: int64
```

Figure 3.7: Code to check for any missing value in the dataset

```
covid_new_data = covid_data.drop(columns = ['Timestamp', 'Name'],)
covid_new_data.head()
```

	Profession	Age	Do you prefer online over offline shopping during the current Covid-19 crises?	Do you think that scope for Online shopping has increased in this pandemic period?	Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?	Do you tend to spend more money if you are shopping online or in store?	How Often do you shop Online ?	How much time do you spend in every visit?	What do you usually purchase online?	What do you usually purchase offline?	What is your satisfaction level when you do online shopping?	What is your satisfaction level when you do offline shopping?	What made you choose online shopping over offline shopping during the current crises?	What made you choose offline shopping over online shopping during the current crises?
0	Student	18-24	Yes	Yes	Yes	Online	Once in 3-6 Months	1-3 Hours	Clothing	Groceries	Very satisfied	Satisfied	It is easier	Like to be hands on
1	Student	18-24	Yes	Yes	No	Online	Daily	10-30 Minutes	Electronics	Electronics	Dissatisfied	Satisfied	More products online	Like in store offers only
2	Student	18-24	Yes	Yes	Yes	In-Store	Monthly	10-30 Minutes	Clothing	Electronics	Neither satisfied nor dissatisfied	Satisfied	Time efficiency	Like to be hands on
3	Student	18-24	Yes	Yes	Yes	Online	Once in 3-6 Months	10-30 Minutes	Electronics	Groceries	Very satisfied	Very satisfied	Online offers more discounts	Reliable
4	Student	18-24	No	Yes	Yes	In-Store	Once in 3-6 Months	1-3 Hours	Electronics	Clothing	Satisfied	Satisfied	Safer	Reliable

Figure 3.8: Cleaned and Trimmed Dataset

```
covid_new_data.nunique()
```

```
Profession                                                                                                   4
Age                                                                                                          4
Do you prefer online over offline shopping during the current Covid-19 crises?                               2
Do you think that scope for Online shopping has increased in this pandemic period?                           2
Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic? 2
Do you tend to spend more money if you are shopping online or in store?                                      2
How Often do you shop Online ?                                                                               4
How much time do you spend in every visit?                                                                   3
What do you usually purchase online?                                                                         5
What do you usually purchase offline?                                                                        4
What is your satisfaction level when you do online shopping?                                                 4
What is your satisfaction level when you do offline shopping?                                                3
What made you choose online shopping over offline shopping during the current crises?                        6
What made you choose offline shopping over online shopping during the current crises?                        5
dtype: int64
```

Figure 3.9: Count of unique values in each feature

3.4 Descriptive Statistics

Descriptive statistics simplify large data sets, providing a clearer understanding of complex information through summaries such as means, medians, modes, and standard deviations. For example, in an online shopping trends study, these statistics can reveal average spending habits and the variability in shopping behaviors among young consumers, offering insights into patterns and trends. This simplification is essential for stakeholders to easily grasp the underlying patterns in data, particularly in sectors like e-commerce where understanding consumer behavior post-COVID is crucial.

By identifying trends and enabling comparisons between different demographic subgroups, descriptive statistics serve as a foundational tool for deeper analysis. Researcher will detect anomalies or significant deviations in data that may indicate unusual behaviors, which could suggest economic impacts on certain consumer segments post-pandemic. Furthermore, these statistics facilitate comparisons across different age groups, genders, or locations, enhancing targeted marketing strategies by understanding preferences for online versus in-person shopping.

Descriptive statistics not only help in providing framework for further statistical testing but it also important in data visualization. Charts like histogram, box plots, and bar charts can be used for visual representation of data, making large and complex data look more accessible and interpretable, especially to people that do not have a statistical background. Moreover, the ease of performing these statistical methods ensures that researcher can efficiently assess data quality and maintain the reliability and validity of the conclusions, thereby establishing a robust baseline for the quality of the collected data. The following codes were imputed to see the result:

```
covid_new_data.describe()
```

	Profession	Age	Do you prefer online over offline shopping during the current Covid-19 crises?	Do you think that scope for Online shopping has increased in this pandemic period?	Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?	Do you tend to spend more money if you are shopping online or in store?	How Often do you shop Online ?	How much time do you spend in every visit?	What do you usually purchase online?	What do you usually purchase offline?	What is your satisfaction level when you do online shopping?	What is your satisfaction level when you do offline shopping?	What made you choose online shopping over offline shopping during the current crises?	What made you choose offline shopping over online shopping during the current crises?
count	100	100	100	100	100	100	100	100	100	100	100	100	100	100
unique	4	4	2	2	2	2	4	3	5	4	4	3	6	5
top	Student	18-24	Yes	Yes	Yes	In-Store	Monthly	1-3 Hours	Clothing	Groceries	Satisfied	Very satisfied	Time efficiency	Like to be hands on
freq	72	72	84	97	86	55	37	49	50	45	58	48	35	38

Figure 3.10: Descriptive Statistics of the Dataset

3.5 Exploratory Data Analysis

3.5.1 Univariate Analysis:

Univariate analysis is the simplest and basic form of analysing and understanding data. The word “Uni” means “one”, which means the data has only one variable which can be analysed (Glen, 2014). Univariate analysis simply means the analysis of one variable and the purpose of univariate analysis is to understand the distribution of values for a single variable (Zach, 2021b). The description or visualization of univariate data can be done using central tendency (mean, median and mode) and dispersion, standard deviation, variance, range, frequency distribution tables, bar charts, histograms, pie charts etc.

For the purpose of this research, univariate analysis was carried out on each variable in the dataset by plotting bar charts to understand the frequency distribution of each column. The following codes were implemented to get the frequencies of each column:

```
prof_count = covid_new_data["Profession"].value_counts()
print("Profession Counts:\n", prof_count)

prof_data = covid_new_data['Profession'].value_counts()
plt.figure(figsize=(10,6))
prof_data.plot(kind='bar')
plt.title('Profession Distribution')
plt.xlabel('Profession')
plt.ylabel('Frequency')
plt.show()
```

```
Profession Counts:
 Profession
Student                       72
Employee                      17
Business Man/Women            10
COSMETIC DERMATOLOGIST MD      1
Name: count, dtype: int64
```

Figure 3.11: Code for Univariate Analysis (Profession)

```
count_age = covid_new_data['Age '].value_counts()
print("Age Counts:\n", count_age)

age_data = covid_new_data['Age '].value_counts()
plt.figure(figsize=(10,6))
age_data.plot(kind='bar')
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```

```
Age Counts:
 Age
18-24    72
31-40    14
25-30    12
41-50     2
Name: count, dtype: int64
```

Figure 3.12: Code for Univariate Analysis (Age)

```
online_imp_counts = covid_new_data["Do you think that scope for Online shopping has increased in this pandemic period?"].value_counts()
print("Online Impact Counts:\n", online_imp_counts)

online_data = covid_new_data["Do you think that scope for Online shopping has increased in this pandemic period?"].value_counts()
plt.figure(figsize=(10,6))
online_data.plot(kind='bar')
plt.title('Online Impact Distribution')
plt.xlabel('Online Impact')
plt.ylabel('Frequency')
plt.show()
```

```
Online Impact Counts:
 Do you think that scope for Online shopping has increased in this pandemic period?
Yes    97
No      3
Name: count, dtype: int64
```

Figure 3.13: Code for Univariate Analysis (Online Shopping Impact)

```
loss_rate_counts = covid_new_data["Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?"].value_counts()
print("Loss Rate Counts:\n", loss_rate_counts)

loss_rate_data = covid_new_data["Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?"].value_counts()
plt.figure(figsize=(10,6))
loss_rate_data.plot(kind='bar')
plt.title('Loss Rate Distribution')
plt.xlabel('Loss Rate')
plt.ylabel('Frequency')
plt.show()
```

```
Loss Rate Counts:
 Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?
Yes    86
No     14
Name: count, dtype: int64
```

Figure 3.14: Code for Univariate Analysis (Loss Rate Distribution)

```
spending_rate_counts = covid_new_data["Do you tend to spend more money if you are shopping online or in store?"].value_counts()
print("Spending Rate Counts:\n", spending_rate_counts)

spending_rate_data = covid_new_data["Do you tend to spend more money if you are shopping online or in store?"].value_counts()
plt.figure(figsize=(10,6))
spending_rate_data.plot(kind='bar')
plt.title('Spending Rate Distribution')
plt.xlabel('Spending Rate')
plt.ylabel('Frequency')
plt.show()

Spending Rate Counts:
 Do you tend to spend more money if you are shopping online or in store?
In-Store    55
Online      45
Name: count, dtype: int64
```

Figure 3.15: Code for Univariate Analysis (Spending Rate)

```
shopping_freq_counts = covid_new_data["How Often do you shop Online ? "].value_counts()
print("Shopping Frequency Counts:\n", shopping_freq_counts)

shopping_freq_data = covid_new_data["How Often do you shop Online ? "].value_counts()
plt.figure(figsize=(10,6))
shopping_freq_data.plot(kind='bar')
plt.title('Shopping Frequency Distribution')
plt.xlabel('Spending Frequency')
plt.ylabel('Frequency')
plt.show()

Shopping Frequency Counts:
 How Often do you shop Online ?
Monthly               37
Once in 3-6 Months    31
Weekly                24
Daily                  8
Name: count, dtype: int64
```

Figure 3.16: Code for Univariate Analysis (Spending Frequency)

```
time_freq_counts = covid_new_data["How much time do you spend in every visit? "].value_counts()
print("Time Frequency Counts:\n", time_freq_counts)

time_freq_data = covid_new_data["How much time do you spend in every visit? "].value_counts()
plt.figure(figsize=(10,6))
time_freq_data.plot(kind='bar')
plt.title('Shopping Time Frequency Distribution')
plt.xlabel('Time Frequency')
plt.ylabel('Frequency')
plt.show()

Time Frequency Counts:
 How much time do you spend in every visit?
1-3 Hours         49
10-30 Minutes     42
I do it all Day!   9
Name: count, dtype: int64
```

Figure 3.17: Code for Univariate Analysis (Shopping Time)

```
online_pur_counts = covid_new_data["What do you usually purchase online?"].value_counts()
print("Online Purchase Counts:\n", online_pur_counts)

online_pur_data = covid_new_data["What do you usually purchase online?"].value_counts()
plt.figure(figsize=(10,6))
online_pur_data.plot(kind='bar')
plt.title('Online Purchase Distribution')
plt.xlabel('Online Purchase Frequency')
plt.ylabel('Frequency')
plt.show()

Online Purchase Counts:
 What do you usually purchase online?
Clothing                                    50
Electronics                                 32
Groceries                                   16
All                                          1
Things that i won't use (that's true)        1
Name: count, dtype: int64
```

Figure 3.18: Code for Univariate Analysis (Online Purchase)

```
offline_pur_counts = covid_new_data["What do you usually purchase offline?"].value_counts()
print("Offline Purchase Counts:\n", offline_pur_counts)

offline_pur_data = covid_new_data["What do you usually purchase offline?"].value_counts()
plt.figure(figsize=(10,6))
offline_pur_data.plot(kind='bar')
plt.title('Offline Purchase Distribution')
plt.xlabel('Offline Purchase Frequency')
plt.ylabel('Frequency')
plt.show()

Offline Purchase Counts:
 What do you usually purchase offline?
Groceries      45
Clothing       32
Electronics    22
All             1
Name: count, dtype: int64
```

Figure 3.19: Code for Univariate Analysis (Offline Purchase)

```
online_sat_counts = covid_new_data["What is your satisfaction level when you do online shopping?"].value_counts()
print("Online Satisfaction Counts:\n", online_sat_counts)

online_sat_data = covid_new_data["What is your satisfaction level when you do online shopping?"].value_counts()
plt.figure(figsize=(10,6))
online_sat_data.plot(kind='bar')
plt.title('Online Satisfaction Distribution')
plt.xlabel('Online Satisfaction Frequency')
plt.ylabel('Frequency')
plt.show()

Online Satisfaction Counts:
 What is your satisfaction level when you do online shopping?
Satisfied                             58
Very satisfied                        21
Neither satisfied nor dissatisfied    20
Dissatisfied                           1
Name: count, dtype: int64
```

Figure 3.20: Code for Univariate Analysis (Online Satisfaction)

```
offline_sat_counts = covid_new_data["What is your satisfaction level when you do offline shopping?"].value_counts()
print("Offline Satisfaction Counts:\n", offline_sat_counts)

offline_sat_data = covid_new_data["What is your satisfaction level when you do offline shopping?"].value_counts()
plt.figure(figsize=(10,6))
offline_sat_data.plot(kind='bar')
plt.title('Offline Satisfaction Distribution')
plt.xlabel('Offline Satisfaction Frequency')
plt.ylabel('Frequency')
plt.show()

Offline Satisfaction Counts:
 What is your satisfaction level when you do offline shopping?
Very satisfied                        48
Satisfied                             39
Neither satisfied nor dissatisfied    13
Name: count, dtype: int64
```

Figure 3.21: Code for Univariate Analysis (Offline Satisfaction)

```
online_pref_counts = covid_new_data["What made you choose online shopping over offline shopping during the current crises?"].value_counts()
print("Online Preference Counts:\n", online_pref_counts)

online_pref_data = covid_new_data["What made you choose online shopping over offline shopping during the current crises?"].value_counts()
plt.figure(figsize=(10,6))
online_pref_data.plot(kind='bar')
plt.title('Online Preference Distribution')
plt.xlabel('Online Preference Frequency')
plt.ylabel('Frequency')
plt.show()

Online Preference Counts:
 What made you choose online shopping over offline shopping during the current crises?
Time efficiency                                              35
Online offers more discounts                                 22
More products online                                         21
It is easier                                                 16
Safer                                                         5
Online shoping for electronics and clothing is not satisfactory    1
Name: count, dtype: int64
```

Figure 3.22: Code for Univariate Analysis (Online Preference)

```
offline_pref_counts = covid_new_data["What made you choose offline shopping over online shopping during the current crises?"].value_counts()
print("Offline Preference Counts:\n", offline_pref_counts)

offline_pref_data = covid_new_data["What made you choose offline shopping over online shopping during the current crises?"].value_counts()
plt.figure(figsize=(10,6))
offline_pref_data.plot(kind='bar')
plt.title('Offline Preference Distribution')
plt.xlabel('Offline Preference Frequency')
plt.ylabel('Frequency')
plt.show()

Offline Preference Counts:
 What made you choose offline shopping over online shopping during the current crises?
Like to be hands on               38
Reliable                          35
Like to interact with others      17
Like in store offers only          9
Can compare the quality of product.    1
Name: count, dtype: int64
```

Figure 3.23: Code for Univariate Analysis (Offline Preference)

3.5.2 Bivariate Analysis:

Bivariate analysis means the analysis of two variables. This analysis is done to understand the relationship between two variables in a dataset. There are three ways to perform bivariate analysis such as scatterplots, correlation coefficients and simple linear regression (Zach, 2021a). Bivariate analysis involves analysing the relationship between two variables in a dataset, and it can be used to uncover patterns, correlations and associations between variable, it helps to understand simple cause and effect relationship between two variables in a dataset. Bivariate analysis helps in identifying relationships between two variables, predictive modelling, variable selection and hypothesis testing to determine the significance of relationships between two variables. Bivariate analysis can be done using various tools such as R, Python, MATLAB, SPSS, SAS, Minitab, Excel, STATA etc (Arshad, 2024).

For the purpose of this research, bivariate analysis was carried out using Python and different variables in the dataset were compared by using the scatterplots to understand their relationship and afterwards, they underwent evaluation testing. The following codes were implemented to show the relationship between different variables:

- ***Bivariate analysis between Age and Online over Offline Shopping Preference:*** This analysis was done to check if different age groups have certain preferences for shopping online than shopping offline during the post – COVID pandemic. The code below was implemented to run the analysis:

```
plt.figure(figsize=(6, 3))
sns.scatterplot(x = 'Do you prefer online over offline shopping during the current Covid-19 crises?', y = 'Age ', data = covid_new_data)
plt.title('Scatter Plot between Do you prefer online over offline shopping during the current Covid-19 crises? and Age')
plt.xlabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.ylabel('Age ')
plt.show()
```

Figure 3.24: Code for Bivariate Analysis (Preferred Shopping Medium and Age)

- ***Bivariate analysis between Online Shopping Satisfaction Level and Online Shopping Frequency:*** This analysis was done to understand the relationship between people's online shopping frequency and their level of satisfaction after shopping online. The code below was implemented to run the analysis:

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'What is your satisfaction level when you do online shopping?', y = 'How Often do you shop Online ? ', data = covid_new_data)
plt.title('Scatter Plot between Frequency of Online Shopping and Online Satisfaction Level')
plt.xlabel('What is your satisfaction level when you do online shopping?')
plt.ylabel('How Often do you shop Online ?')
plt.show()
```

Figure 3.25: Code for Bivariate Analysis (Online Shopping Frequency and Online Shopping Satisfaction)

- ***Bivariate Analysis between Online Shopping Scope and Preferred Shopping Medium:*** This analysis was done to determine if the increase in online scope during the pandemic is related with the preferred shopping medium. The code below was implemented to run the analysis:

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'Do you think that scope for Online shopping has increased in this pandemic period?', y = 'Do you prefer online over offline shopping during the current Covid-19 crises?',
plt.title('Scatter Plot between Percieved Scope and Preferred Shopping Medium')
plt.xlabel('Do you think that scope for Online shopping has increased in this pandemic period?')
plt.ylabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.show()
```

Figure 3.26: Code for Bivariate Analysis (Online Shopping Scope and Preferred Shopping Medium)

- ***Bivariate Analysis between Preferred Shopping Medium and Spending Habits:*** This analysis was done to determine if people spend more when they shop online or offline. The code below was implemented to run the analysis:

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'What made you choose online shopping over offline shopping during the current crises?', y = 'Do you tend to spend more money if you are shopping online or in store?', data = covid_new_data)
plt.title('Scatter Plot between Percieved Scope and Spending Habits')
plt.xlabel('What made you choose online shopping over offline shopping during the current crises?')
plt.ylabel('Do you tend to spend more money if you are shopping online or in store?')
plt.show()
```

Figure 3.27: Code for Bivariate Analysis (Preferred Shopping Medium and Spending Habit)

- ***Bivariate Analysis between Preferred Shopping Medium and Products Purchased:*** This analysis was done to determine if the preferred shopping medium affected the types of products being purchased. The code below was implemented to run the analysis:

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'Do you prefer online over offline shopping during the current Covid-19 crises?', y = 'What do you usually purchase online?', data = covid_new_data)
plt.title('Scatter Plot between Preferred Shopping Medium and Product Purchased')
plt.xlabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.ylabel('What do you usually purchase online?')
plt.show()
```

Figure 3.28: Code for Bivariate Analysis (Online over Offline Shopping Preference and Age)

3.6 Statistical Tests on Bivariate Analysis

3.6.1 Cross Tabulation

Cross tabulation can also be referred to as crosstabs or contingency table analysis; it is a statistical method that can be used to analyse the relationship between two or more categorical features or variables. Cross tabulation consists of visualizing the frequencies of entries in different categories in a tabular form and then examines how the categories intersect. This cross tabulation helps to uncover trends, associations, explore relationships, summarize data and in decision making between categorical variables in a dataset (Appinio Research, 2024). According to Clark (2021), it can also be referred to as a process used for bivariate tables to show the relationship between nominal level variables.

3.6.2 Heatmap Correlation

Heatmap is one of the interesting data visualization tool used for providing visual representations of data but in a matrix form. It makes understanding of a data easier through the use of colours from light shades to darker shades. They are majorly used in statistics, user experience design (UX) and data analysis. Heatmap works in colab notebooks by importing seaborns, they operate in mapping out data values to a certain colour scale where light colours indicates higher values and dark colours indicate lower values. They offer benefits like visual clarity, data exploration, correlation between two variables, data – driven decision making, effective and efficient problem solving and performance evaluation (Appinio Research, 2023)

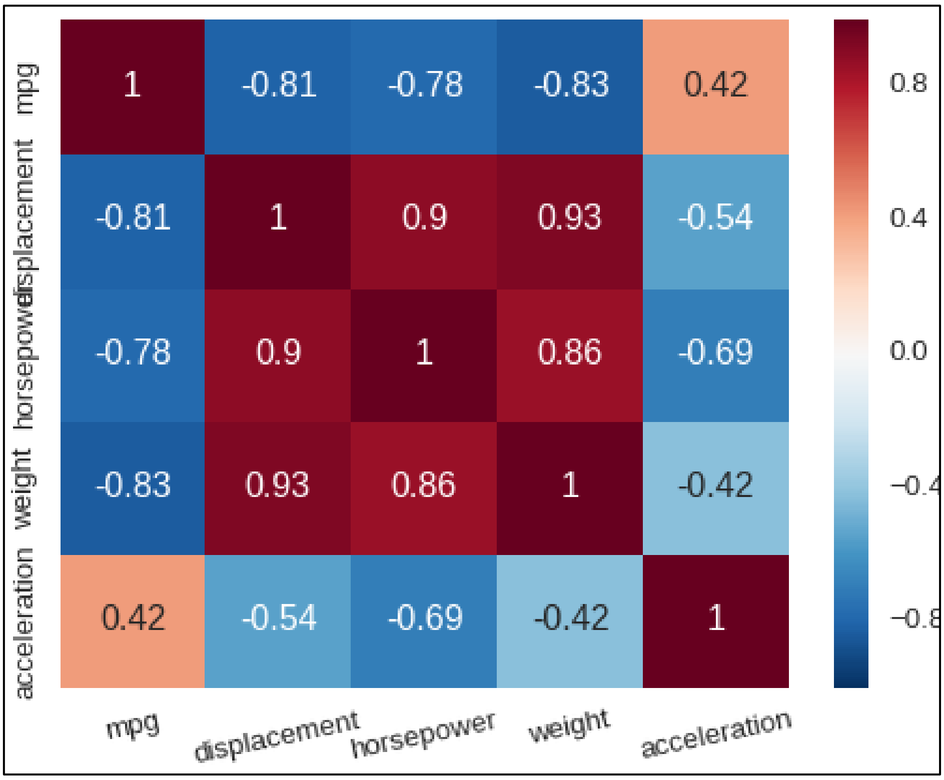


Figure 3.29: Overview of a Heatmap Correlation

3.6.3 Chi-Square Test (X^2)

Chi – square test (X^2) is a statistical test which is mainly used to compare both observed and expected result. It helps to determine the correlation between two categorical variables in a dataset. It helps a researcher in understanding and being able to interpret the connection between two categorical variables (Biswal, 2023). According to Hayes (2023), Chi-square tests are used to test hypothesis by comparing the results from the observed results and the expected results.

The formula for calculating chi – square:

$$x^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

Equation 3.1: Chi - Squared Formula

Where,

x^2 = Chi – Square Value

O_i = Observed frequency

E_i = Expected frequency if there is no relationship between the variables

Expected value (EV) is usually seen if there is no relationship between two categorical variables; it is the theoretical frequency of occurrences in each cell of a cross table.

Formula for calculating expected values:

$$EV = \frac{\sum R \times \sum C}{GT}$$

Equation 3.2: Expected Values Formula

Where:

EV = Expected value

$\sum R$ = Row Total

$\sum C$ = Column Total

GT = Grand Total

3.6.4 p – value

p – value can be referred to as a statistical measurement used to validate a hypothesis against an observed data. Assuming that a null hypothesis is true, P-value helps in measuring the probability of obtaining the observed results; and the lower the p-value, the greater the statistical significance of the observed data and vice versa. A p-value of 0.5 or lower is usually considered statistically significant (Beers, 2023). According to Bevans (2020), P-values can be interpreted as the risk of rejecting the null hypothesis of the test carried out when the null hypothesis is actually true but it is important to note that p-value can only tell if or not a null hypothesis is accepted.

Hypothesis:

H_o (Null Hypothesis) – means there is no significant difference between the variables (p – value ≥ 0.05)

H_1 or H_a (Alternative Hypothesis) – means there is significant difference between the variables (p – value ≤ 0.05)

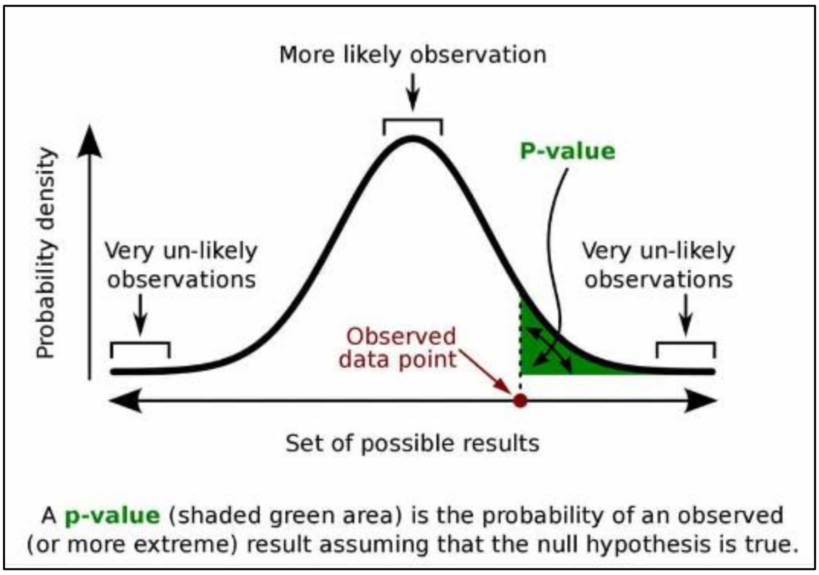


Figure 3.30: P - Value Analysis

3.6.5 Degree of Freedom (df)

This can be termed as the maximum number of logically independent values and this varies in data sample, and they are usually calculated by subtracting one from the number of items within the dataset sample (Ganti, 2023). They are usually represented by v or df, and it can referred to as the number of independent pieces of information usually used to calculate a statistic (Turney, 2022).

The formula for calculating degree of freedom:

$$D_f = N - 1$$

Equation 3.3: Degree of Freedom Formula

Where:

D_f = degree of freedom

N = sample size

3.6.6 Implementation of Test Codes

```
#statistical test
#crosstabulation
crosstab = pd.crosstab(covid_new_data['Do you prefer online over offline shopping during the current Covid-19 crises?'], covid_new_data['Age '])
crosstab

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium and Age')
plt.xlabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.ylabel('Age ')
plt.show()

from scipy.stats import chi2_contingency
chi2, p, dof, expected = chi2_contingency(crosstab)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', round(chi2, 2))
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

Figure 3.31: Code for Statistical Inference (Preferred Shopping Medium and Age)

```
#crosstabulation
crosstab_2 = pd.crosstab(covid_new_data['What is your satisfaction level when you do online shopping?'], covid_new_data['How Often do you shop Online ? '])
crosstab_2

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_2, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Frequency of Online Shopping and Online Satisfaction Level')
plt.xlabel('What is your satisfaction level when you do online shopping?')
plt.ylabel('How Often do you shop Online ? ')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_2)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', round(chi2, 2))
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

Figure 3.32: Code for Statistical Inference (Online Shopping Frequency and Online Shopping Satisfaction)

```
#crosstabulation
crosstab_3 = pd.crosstab(covid_new_data['Do you think that scope for Online shopping has increased in this pandemic period?'], covid_new_data['Do you prefer online over offline shopping during the current Covid-19 crises?'])
crosstab_3

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_3, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Percieved Scope and Preferred Shopping Medium')
plt.xlabel('Do you think that scope for Online shopping has increased in this pandemic period?')
plt.ylabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_3)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

Figure 3.33: Code for Statistical Inference (Online Shopping Scope and Preferred Shopping Medium)

```
#crosstabulation
crosstab_4 = pd.crosstab(covid_new_data['What made you choose online shopping over offline shopping during the current crises?'],
                         covid_new_data['Do you tend to spend more money if you are shopping online or in store?'])
crosstab_4

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_4, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium and Spending Habits')
plt.xlabel('What made you choose online shopping over offline shopping during the current crises?')
plt.ylabel('Do you tend to spend more money if you are shopping online or in store?')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_4)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

Figure 3.34: Code for Statistical Inference (Preferred Shopping Medium and Spending Habit)

```
#crosstabulation
crosstab_5 = pd.crosstab(covid_new_data['Do you prefer online over offline shopping during the current Covid-19 crises?'],
                         covid_new_data['What do you usually purchase online?'])
crosstab_5

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_5, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium and Product Purchased')
plt.xlabel('Do you prefer online over offline shopping during the current Covid-19 crises?')
plt.ylabel('What do you usually purchase online?')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_5)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

Figure 3.35: Code for Statistical Inference (Preferred Shopping Medium and Products Purchased)

CHAPTER FOUR

RESULTS AND DISCUSSION OF FINDINGS

The COVID-19 pandemic has precipitated unprecedented changes across various sectors, significantly altering consumer behaviors and preferences. This study seeks to understand the shift from offline to online shopping during the pandemic, with a focus on how different age groups have adapted to these changes. The data collected provides insights into the frequency of shopping, types of products purchased, and overall consumer satisfaction with both online and offline shopping experiences.

4.1 Descriptive Statistics

The descriptive statistics of the dataset couldn't capture the mean, median, mode and standard deviation of the dataset because they weren't numerical columns, but the analysis gave the count of each column entry, the unique answers in each column, the maximum unique entry in each column and the frequency of each entries. This result wasn't enough to give the needed understanding of dataset, hence the proceeding to carry out the univariate analysis.

```
covid_new_data.describe()
```

	Profession	Age	Do you prefer online over offline shopping during the current Covid-19 crises?	Do you think that scope for Online shopping has increased in this pandemic period?	Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?	Do you tend to spend more money if you are shopping online or in store?	How Often do you shop Online ?	How much time do you spend in every visit?	What do you usually purchase online?	What do you usually purchase offline?	What is your satisfaction level when you do online shopping?	What is your satisfaction level when you do offline shopping?	What made you choose online shopping over offline shopping during the current crises?	What made you choose offline shopping over online shopping during the current crises?
count	100	100	100	100	100	100	100	100	100	100	100	100	100	100
unique	4	4	2	2	2	2	4	3	5	4	4	3	6	5
top	Student	18-24	Yes	Yes	Yes	In-Store	Monthly	1-3 Hours	Clothing	Groceries	Satisfied	Very satisfied	Time efficiency	Like to be hands on
freq	72	72	84	97	86	55	37	49	50	45	58	48	35	38

Figure 36: Descriptive Analysis of Dataset

4.2 Univariate Analysis

4.2.1 Demographic Analysis

Table 4.1: Demographic Information

Profession	Business Man/Women	Cosmetic Dermatologist MD	Employee	Student
	10 (10%)	1 (1%)	17 (17%)	72% (72%)
Age range	18 – 24	25 – 30	31 – 40	41 – 50
	72 (72%)	12 (12%)	14 (14%)	2 (2%)
How much time do you spend in every visit?	10 – 30 Minutes	1 – 3 Hours	I do it all Day!	
	42 (42%)	49 (49%)	9 (9%)	

The demographic data in the study categorizes participants by profession, age range, and duration of visit, likely in a COVID-19-related context. The predominant representation of students (72%) may skew the findings and limit their relevance to a broader audience. Additionally, the minuscule representation of cosmetic dermatologists (1%) may not yield sufficient data for substantial subgroup analysis. The age distribution also exhibits a significant bias towards the 18-24 age groups (72%), with very limited data on older age groups, such as the 41-50 range (2%), which might hinder robust conclusions across a more diverse population.

The table 4.1 indicates a diverse range of visit durations, with a notable split between individuals spending 10-30 minutes and those spending 1-3 hours, and a smaller fraction reporting "all day" visits. However, the categorization of "all day" lacks specificity, and the purpose of these visits remains unclear whether they pertain to work, healthcare facilities, or other COVID-19-related activities. The inclusion of more precise time brackets and a clearer definition of visit types would enhance the table's clarity and utility.

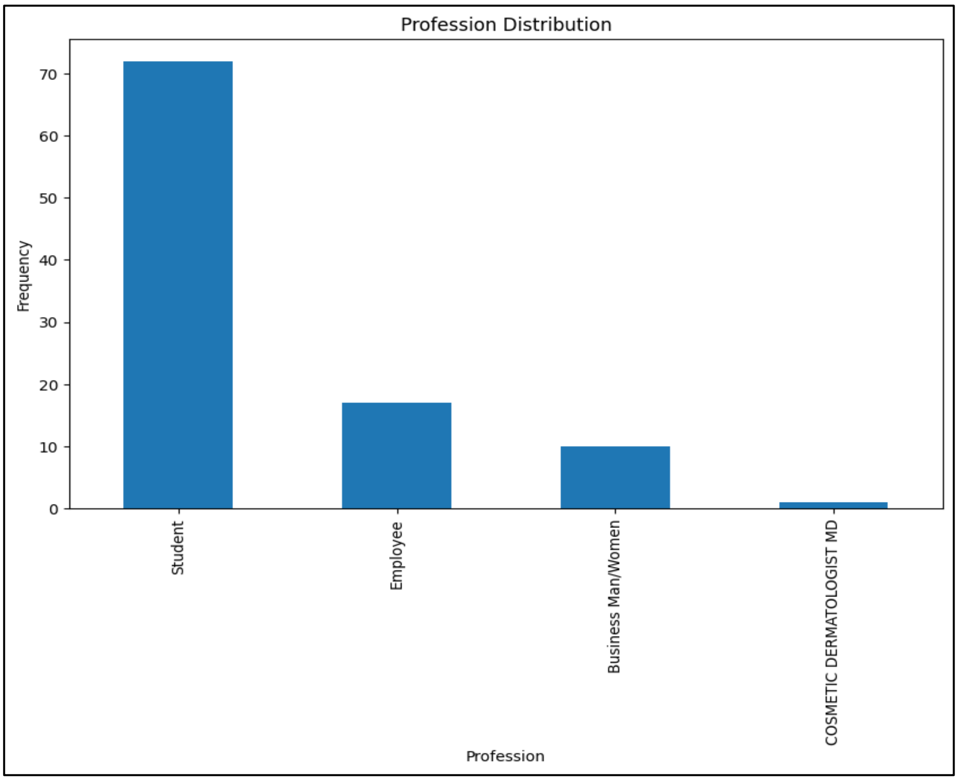


Figure 372: Univariate Analysis (Profession Frequency)

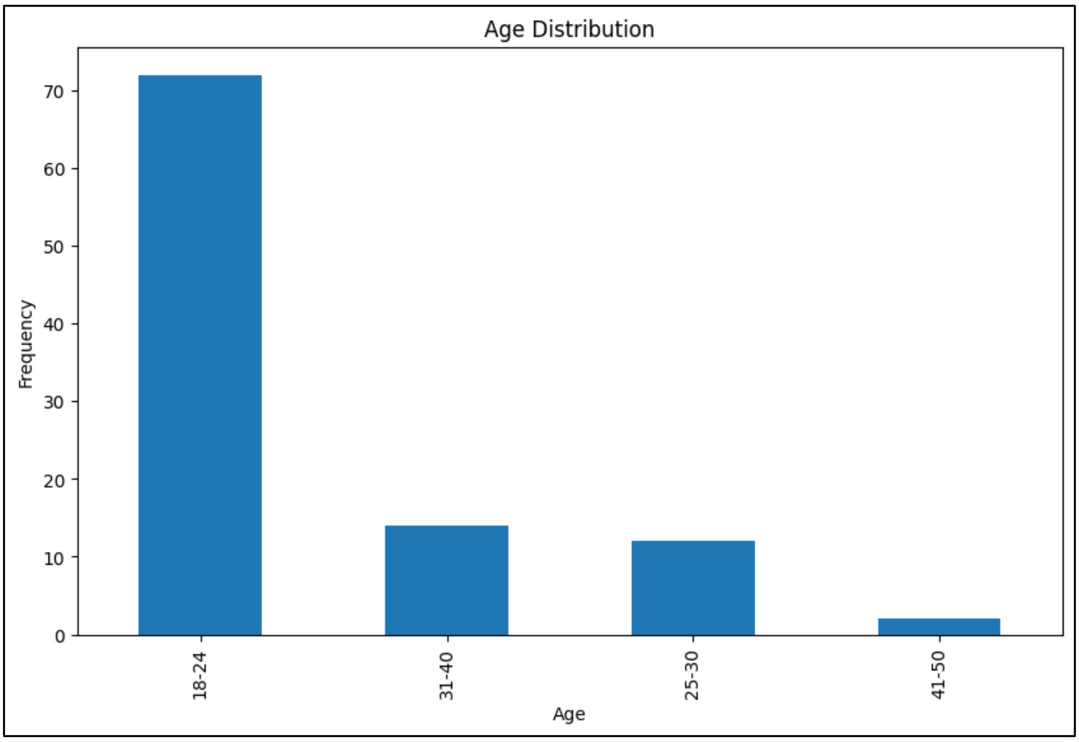


Figure 383: Univariate Analysis (Age Frequency)

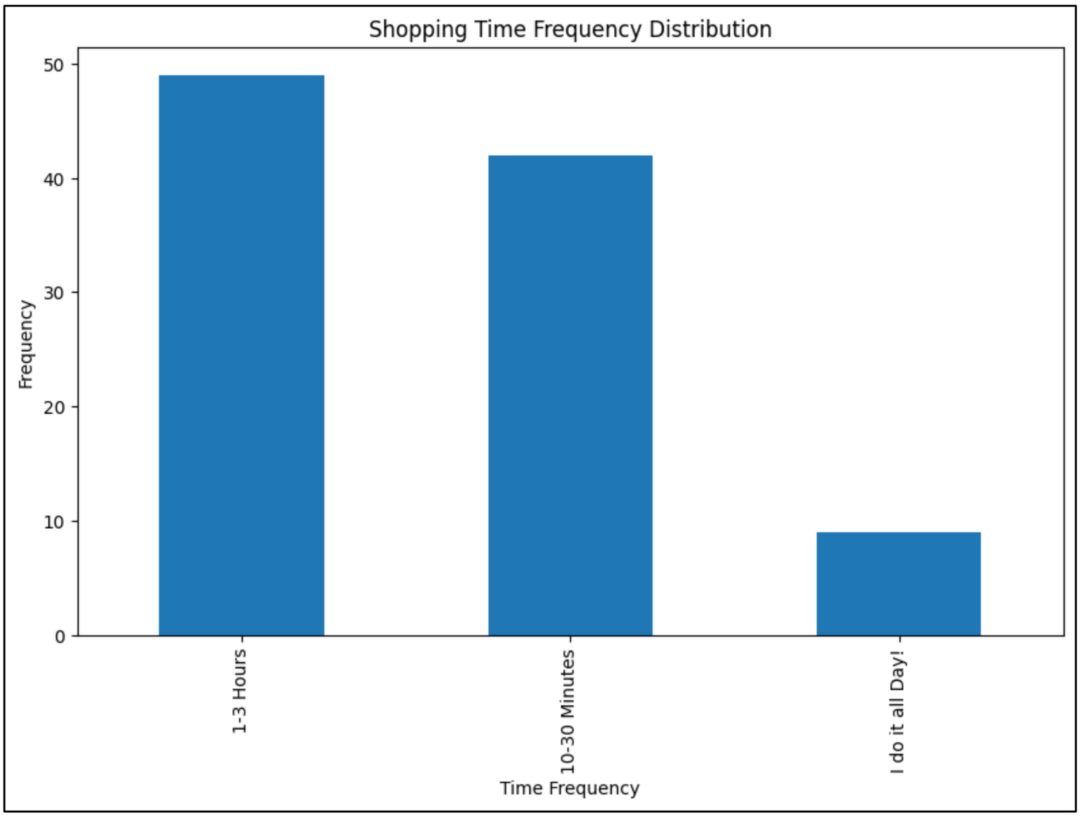


Figure 394: Univariate Analysis (Shopping Time Frequency)

While the data offers a glimpse into certain demographic patterns, its skewed distribution concerning profession and age calls for cautious interpretation of any statistical analyses. The absence of variables such as gender, socioeconomic status, and geographic location further complicates the analysis. To improve representativeness and applicability, it would be advantageous to collect more balanced data and possibly apply weighting techniques in the analysis. Providing additional context regarding the nature of the visits and expanding demographic details would also enrich the interpretation of COVID-19's impact on various population segments.

4.2.2 Shift in consumer preference during covid-19 pandemic

Table 4.2: Shift in Consumer Shopping Preference during COVID - 19

S/N	Statement	YES	NO
1	Do you prefer online over offline shopping during the current Covid-19 crises?	84(84%)	16(16%)
2	Do you think that scope for online shopping has increased in this pandemic period?	97(97%)	3(3%)
3	Do you think that offline stores have incurred loss due to the shift in trend to online shopping in this pandemic?	86(86%)	14(14%)
	Statement	**In-Store**	**Online**
4	Do you tend to spend more money if you are shopping online or in store?	55(55%)	45(45%)

The table 4.2 contains survey results that analyze consumer shopping preferences during the COVID-19 pandemic. It reveals that 84% of participants favored online shopping, attributing this trend to the safety and convenience provided by digital platforms during the crisis. Moreover, a significant 97% agree that the pandemic has expanded the possibilities for online shopping, likely due to increased service availability and lockdown protocols. Conversely, 86% believe that physical stores have suffered losses, underscoring the adverse impacts such as reduced customer presence and spending.

However, the data on spending habits shows a more balanced view, with 55% of respondents indicating they spend more in physical stores compared to 45% online, suggesting the continued appeal of direct shopping experiences possibly due to immediate product access and impulsive purchases. This points to a complex consumer behavior that might not fully lean towards online shopping despite its apparent preference.

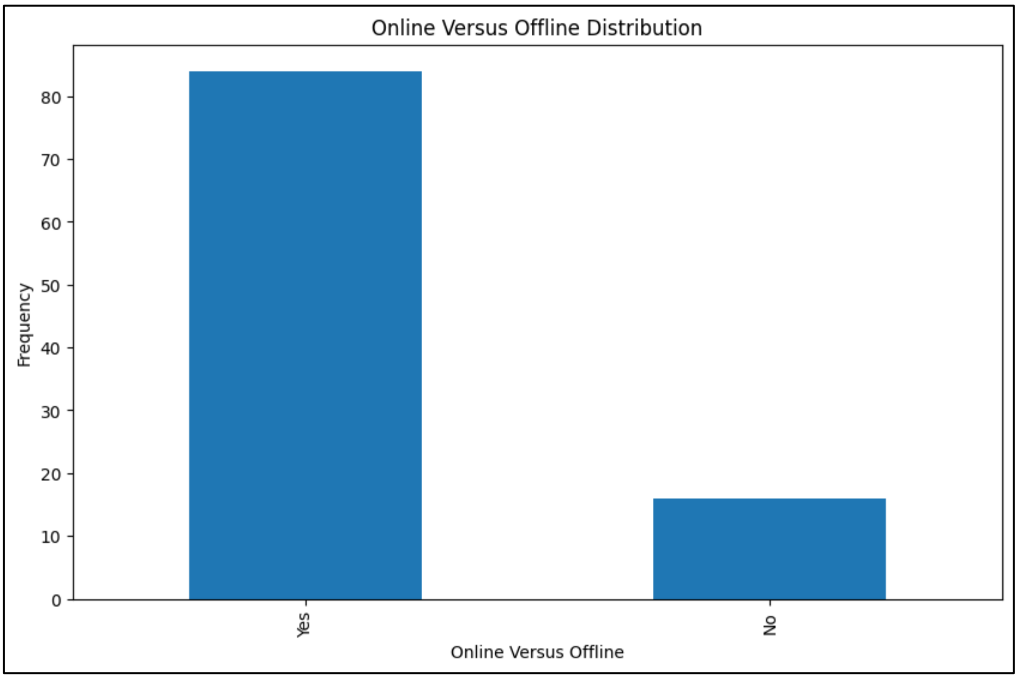


Figure 405: Univariate Analysis (Preferred Shopping Medium Frequency)

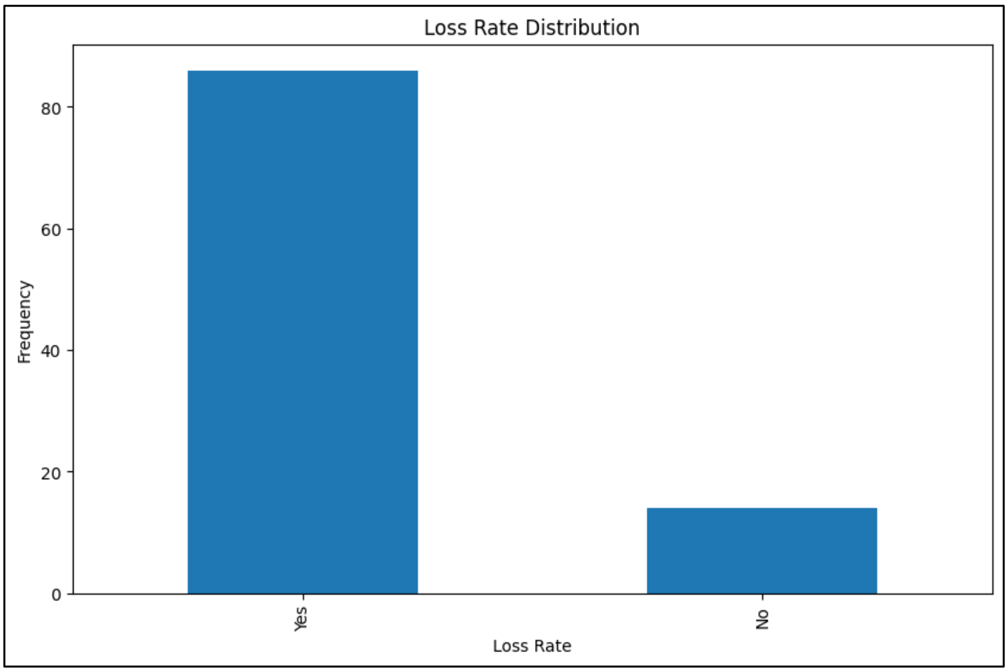


Figure 416: Univariate Analysis (Loss Rate Frequency)

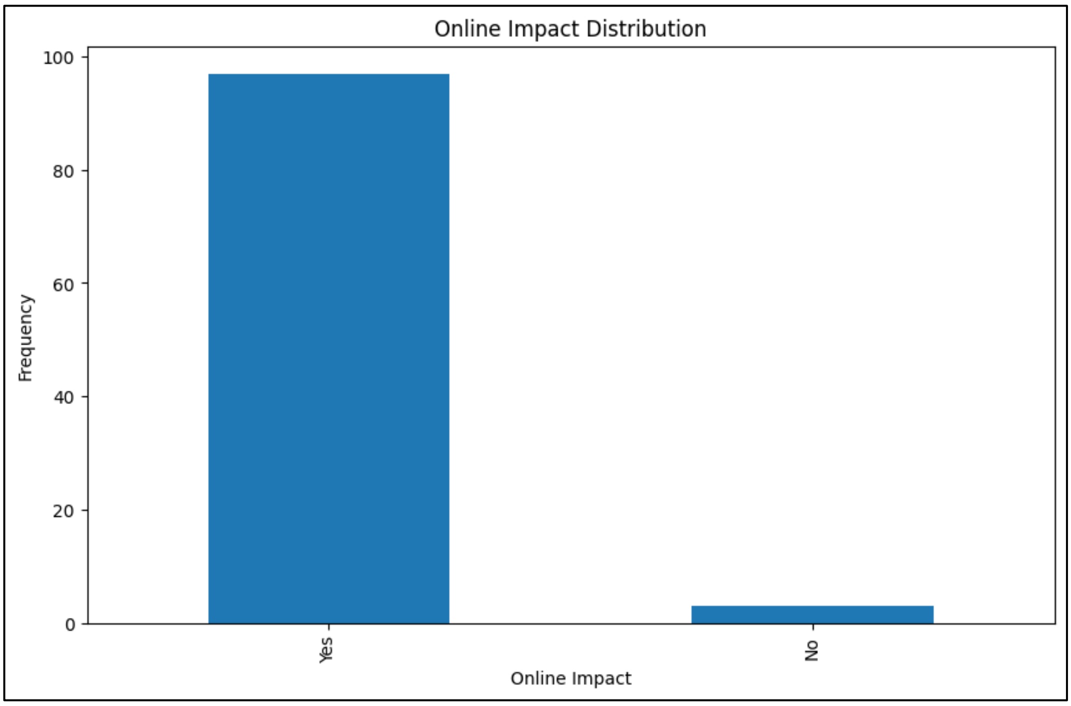


Figure 427: Univariate Analysis (Online Impact Frequency)

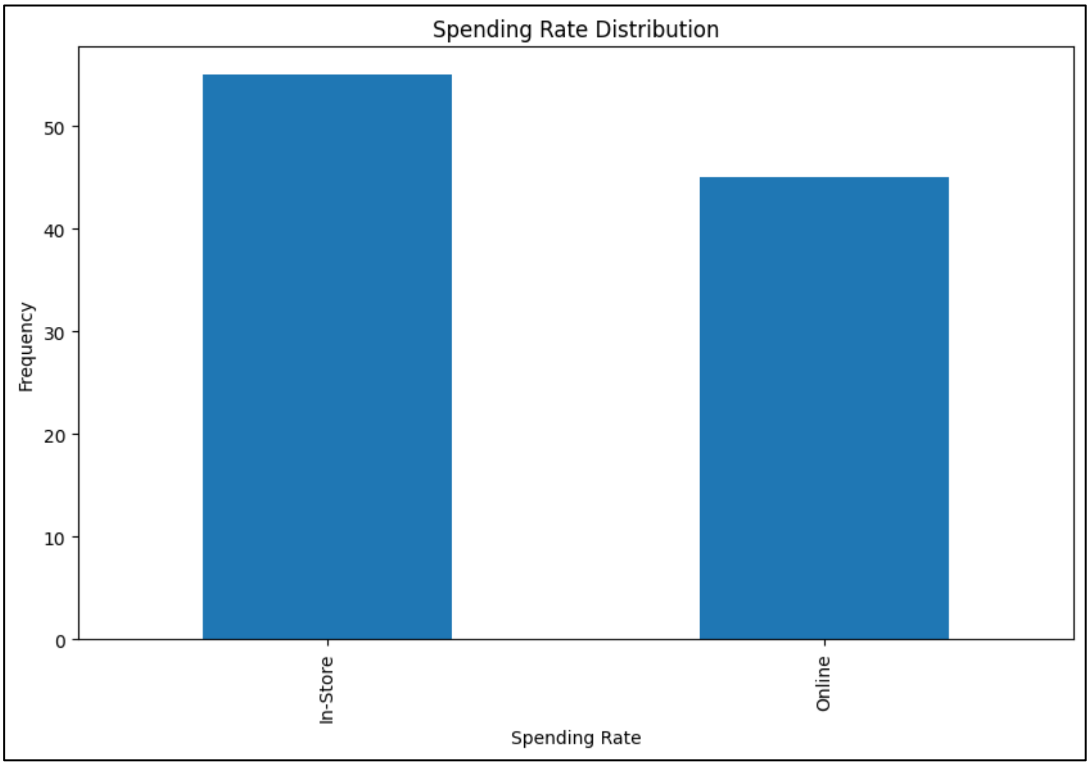


Figure 438: Univariate Analysis (Spending Rate Frequency)

The simplicity of the survey questions, particularly the binary format, may not capture the nuanced consumer attitudes effectively. Additionally, the data would be more insightful with more detailed options, such as the specific nature of losses faced by offline stores and the types of products influencing higher spending in-store. Enhancing the survey with these elements could provide a more accurate depiction of shopping behaviors during the pandemic.

4.2.3 Online and Offline Shopping Preference

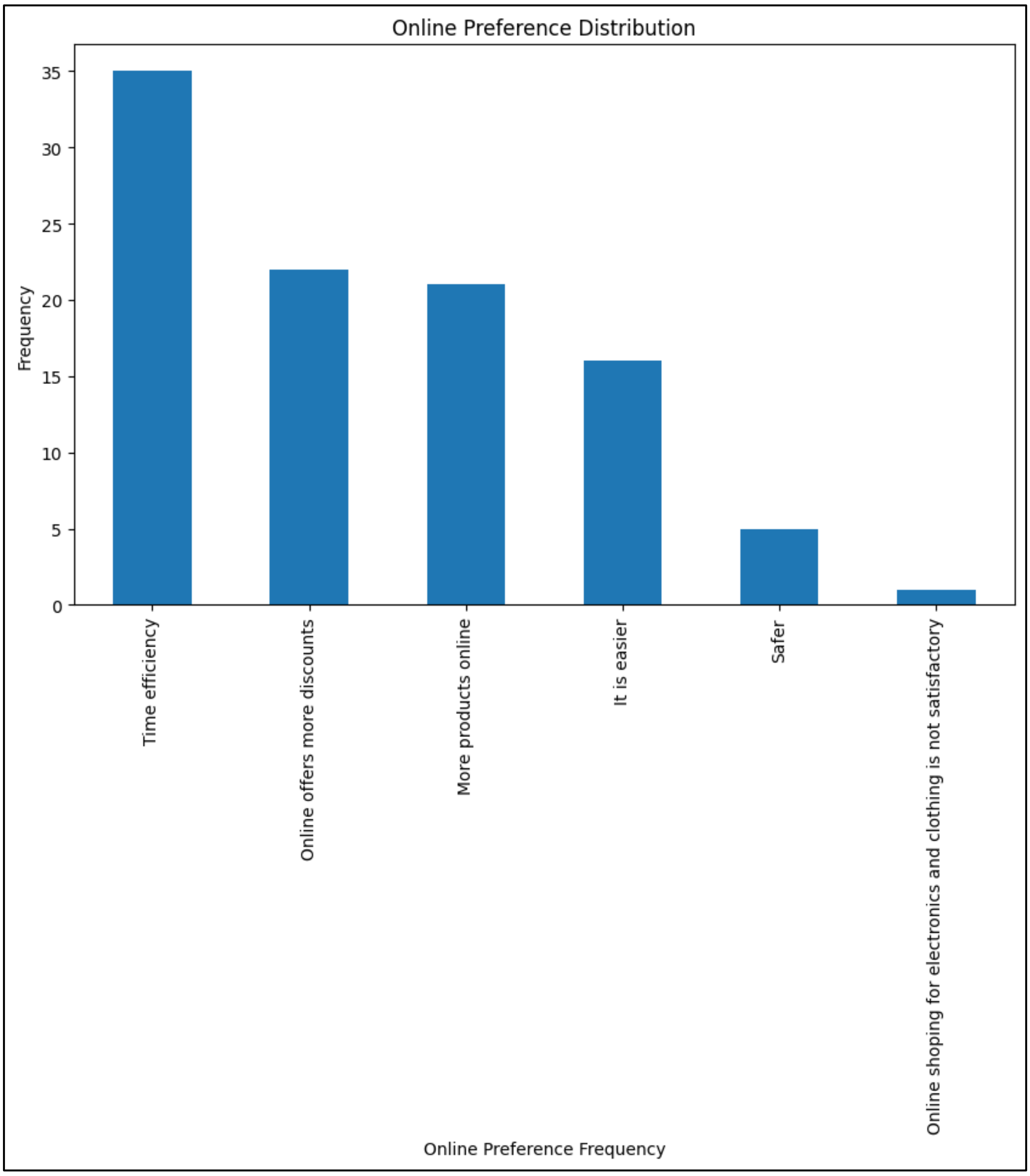


Figure 449: Univariate Analysis (Online Preference Frequency)

The survey results shed light on various consumer preferences and concerns about shopping during a crisis, revealing significant insights into behaviors and choices. Safety emerged as the primary concern for 35% of respondents, who prefer offline shopping to avoid potential health risks associated with package delivery during crises. This emphasis on safety suggests a possible misunderstanding or overestimation of the risks tied to online shopping and deliveries, highlighting a need for further education and data to clarify these perceptions. Additionally, 22% and 20% of participants pointed out that offline shopping often lacks the discounts and product variety available online, indicating that cost and selection are crucial factors in their shopping decisions.

Critiques of these preferences suggest that the survey might not fully capture the nuances of consumer behavior. For instance, the preference for the tactile experience of offline shopping could be influenced by demographic variables like age or internet accessibility, which were not detailed in the survey. Moreover, a mere 3% of respondents appreciated the time efficiency of offline shopping, a surprising find as online shopping is generally regarded as more convenient. This could be due to specific crisis-related issues such as delayed shipping. The minimal responses for categories like 'None' or 'All of the above' suggest that the survey questions may not have captured all consumer preferences accurately, indicating a potential area for improvement in survey design to better understand and address consumer needs during crises.

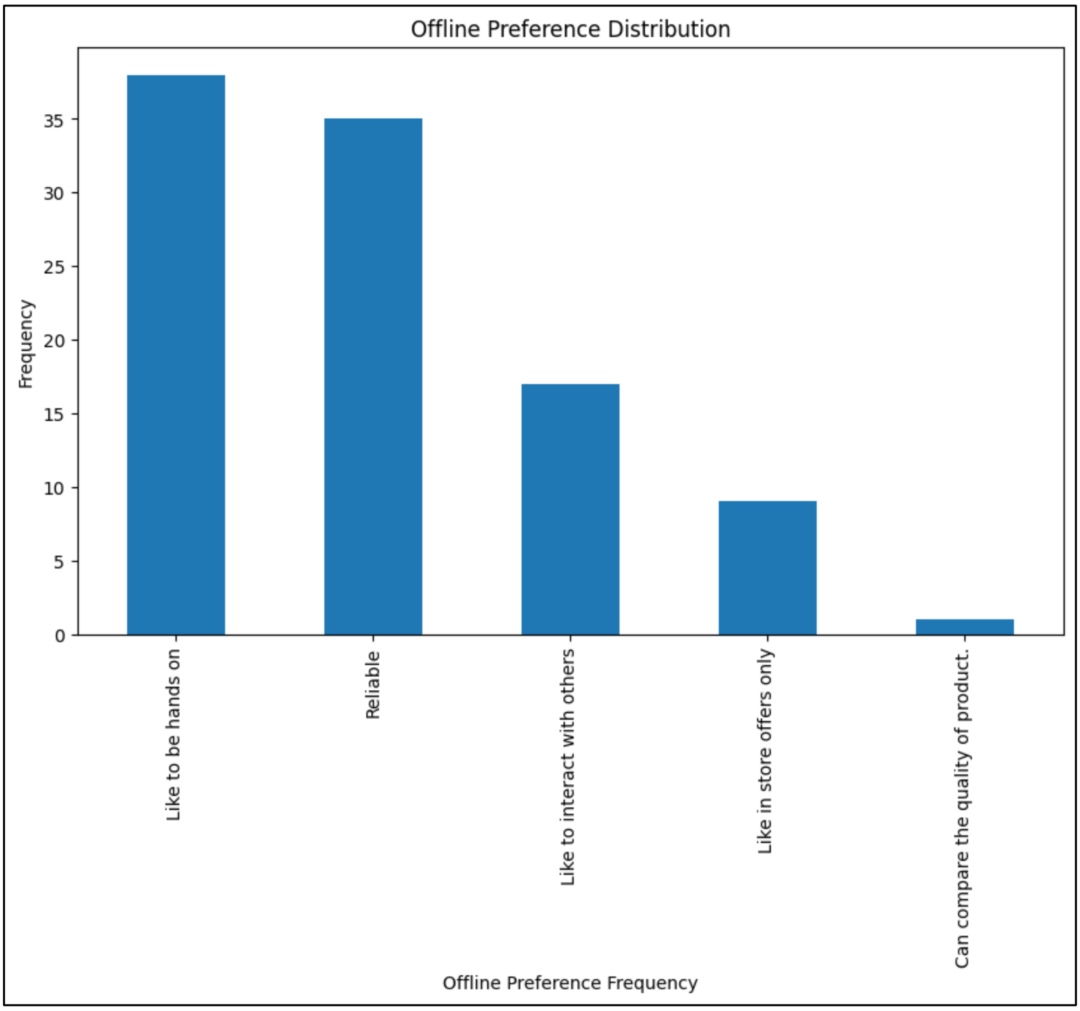


Figure 450: Univariate Analysis (Offline Preference Frequency)

The survey results offer valuable insights into consumer behavior during crises, underscoring the importance of detailed online interactions and platform reliability. About 38% of respondents favor a hands-on approach during online shopping, highlighting a preference for in-depth product information and comparison features like enhanced visualizations, comprehensive descriptions, and user reviews. Additionally, 35% value reliability, emphasizing the need for dependable product availability and delivery during crises, which bolsters the perceived trustworthiness of e-commerce platforms. Notably, 17% of consumers still seek social interaction, suggesting opportunities for more interactive online shopping experiences, such as live chats or video assistance.

However, the survey's credibility faces limitations due to its small sample size of 100 respondents and the lack of specific demographic data, which are vital for a comprehensive understanding of consumer preferences across different groups. The low emphasis on product quality and in-store offers, marked by 1% and 9% respectively, raises questions about whether factors like safety and convenience during the crisis might overshadow these aspects. Furthermore, the ambiguity of terms such as "reliable" and the survey's failure to specify the crisis context whether health-related or economic restrict the interpretation of the data, suggesting the need for a more robust methodology and clearer categorization to accurately capture consumer behavior nuances.

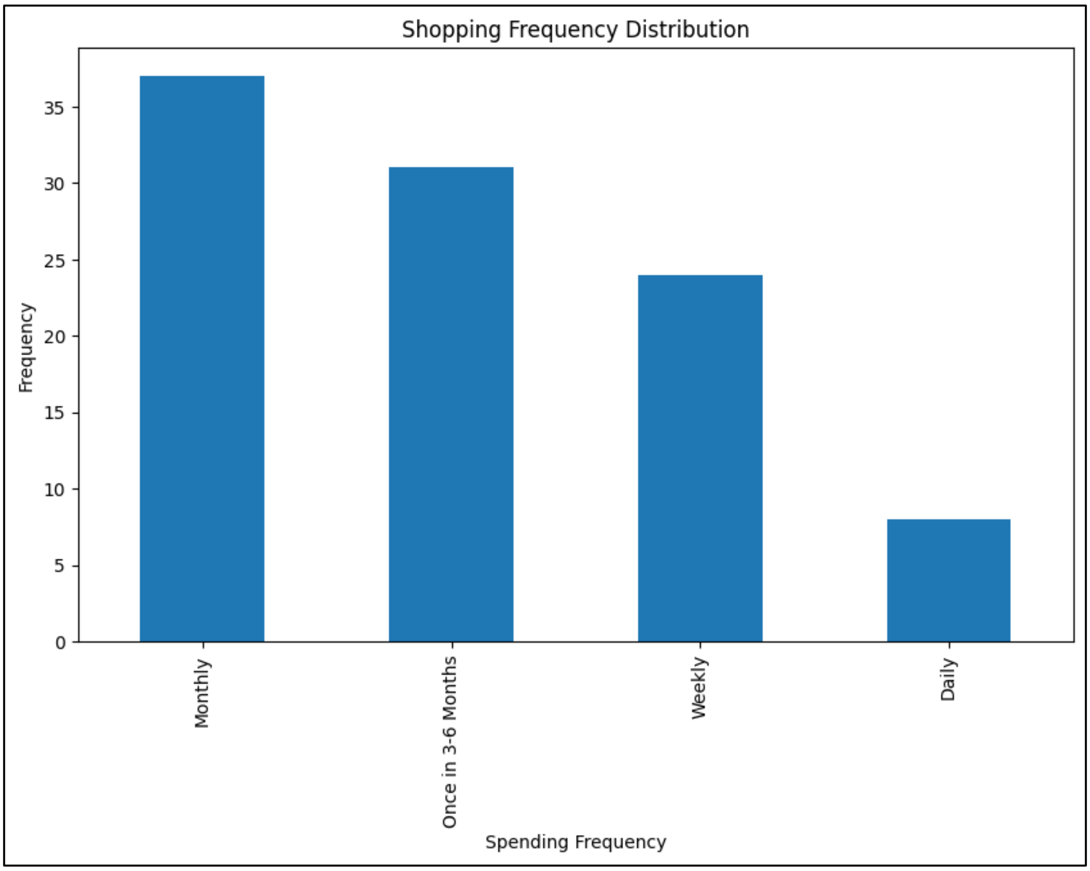


Figure 46: Univariate Analysis (Spending Habit Frequency)

The survey data reveals various frequencies of online shopping among respondents: 8% shop daily, suggesting that while daily online shopping is uncommon, it may cater to those with specific needs or who find it highly convenient. Weekly shopping is more prevalent, with 24% of participants engaging in it, highlighting its role in meeting regular, albeit non-urgent, needs. The majority, 37%, shop monthly, possibly reflecting alignment with pay cycles or routine stock replenishment. Lastly, 31% shop once every three to six months, indicating purchases are likely for durable or non-essential items that do not necessitate frequent buying.

4.2.4 Online and Offline Satisfaction Level

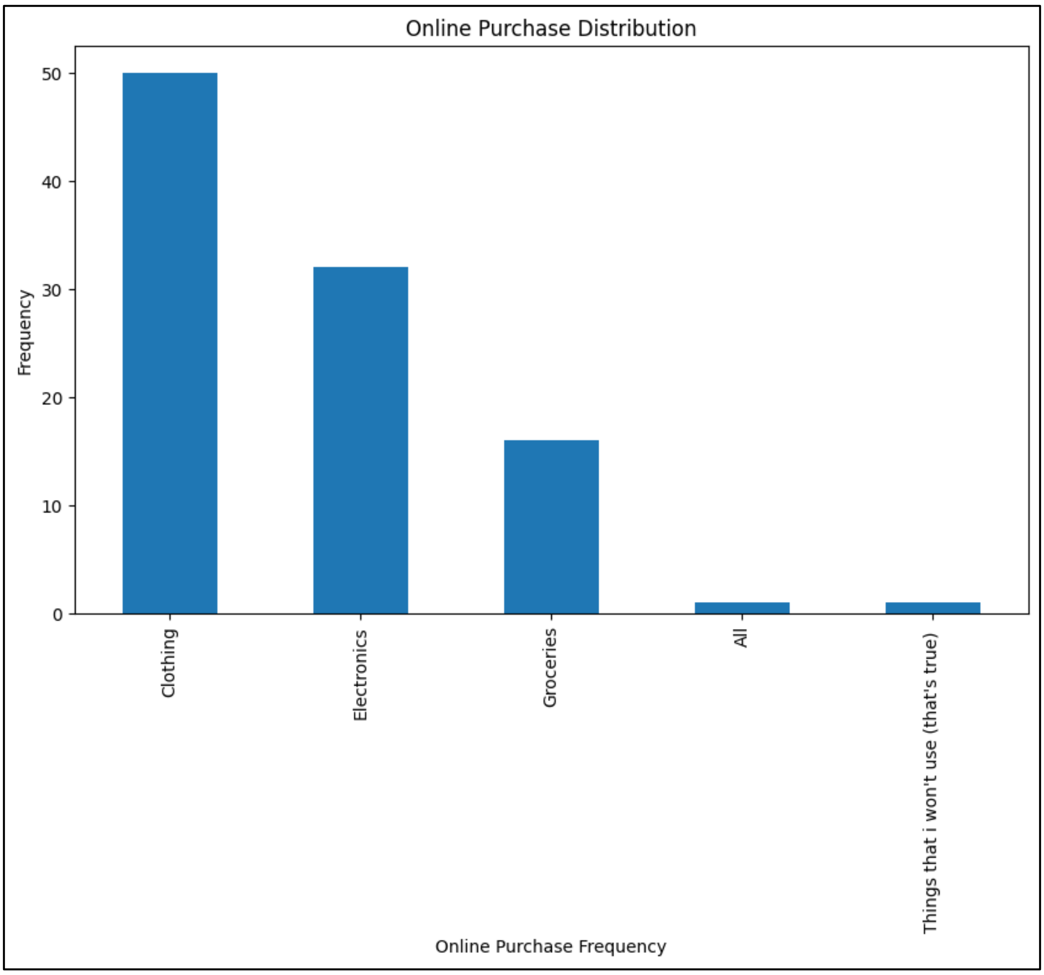


Figure 47: Univariate Analysis (Online Purchase Frequency)

The bar charts present survey results concerning online shopping habits across various categories. Clothing dominates the online shopping sphere with 50% of respondents favoring this category, which might highlight the allure of online platforms offering extensive choices, ease of navigation, and accommodating return policies for sizing issues. Electronics follow with 32% preference, possibly due to the comprehensive information available online including detailed specifications, user reviews, and competitive prices that facilitate well-informed purchasing decisions from home.

Conversely, only 16% of participants reported purchasing groceries online, likely due to a preference for hand-picking fresh produce or challenges linked to delivering perishable goods. Interestingly, a mere 1% of the survey participants admitted to buying items online that they do not intend to use, reflecting occasional impulse purchases or mistaken orders in the easy-click digital shopping environment. These findings illustrate a varied pattern in consumer online shopping behaviors, offering valuable insights for e-commerce marketing strategies to capitalize on prevailing trends and preferences.

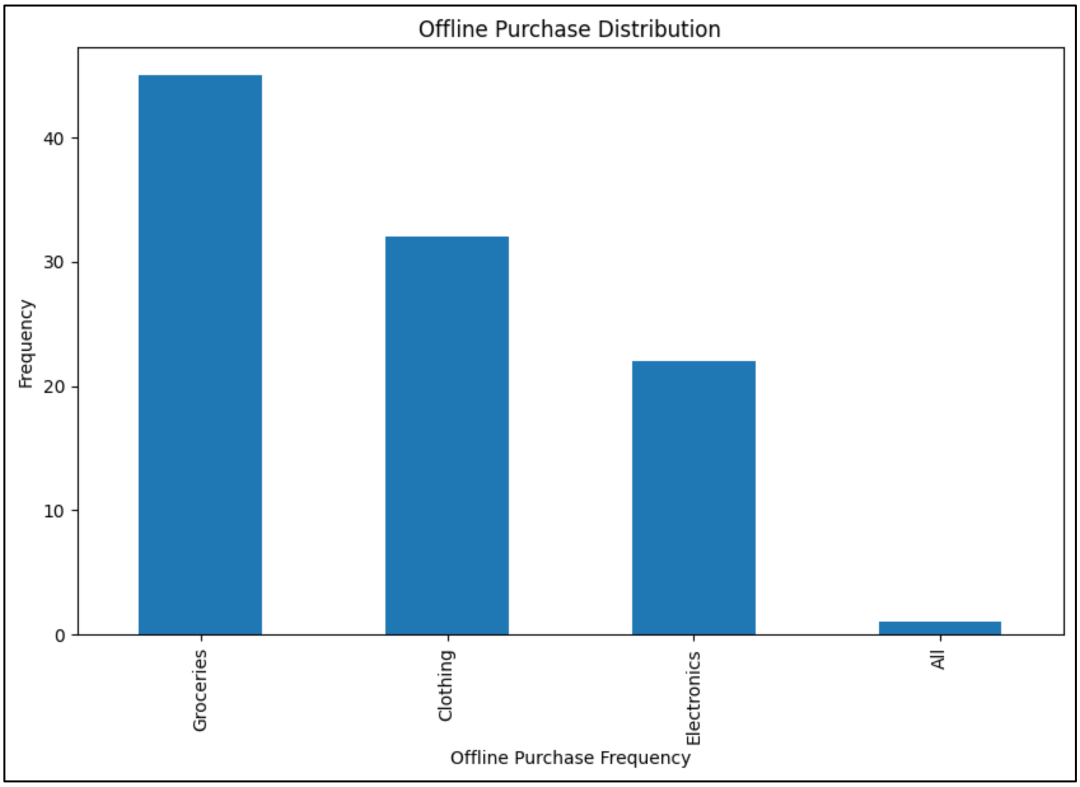


Figure 48: Univariate Analysis (Offline Purchase Frequency)

The figure 4.13 details the offline purchase preferences across various product categories, with the following percentage breakdown from respondents: 45% usually buy groceries offline, 32% purchase clothing, and 22% acquire electronics in physical stores. This suggests that groceries, being perishable, are most often bought in person, likely because consumers prefer to assess their quality directly. Clothing also has a significant offline presence, though it competes with online shopping, which offers more convenience, variety, and potential savings. Electronics, meanwhile, have the lowest offline purchase rate, possibly due to the extensive product information and reviews available online.

Table 4.3: Online and Offline Satisfaction Level

Statement	**Dissatisfied**	**Neither satisfied nor dissatisfied**	**Satisfied**	**Very satisfied**
What is your satisfaction level when you do online shopping?	1(1%)	20(20%)	58(58%)	21(21%)
Statement	**Very satisfied**	**Satisfied**	**Neither satisfied nor dissatisfied**	
What is your satisfaction level when you do offline shopping?	48(48%)	39(39%)	13(13%)	

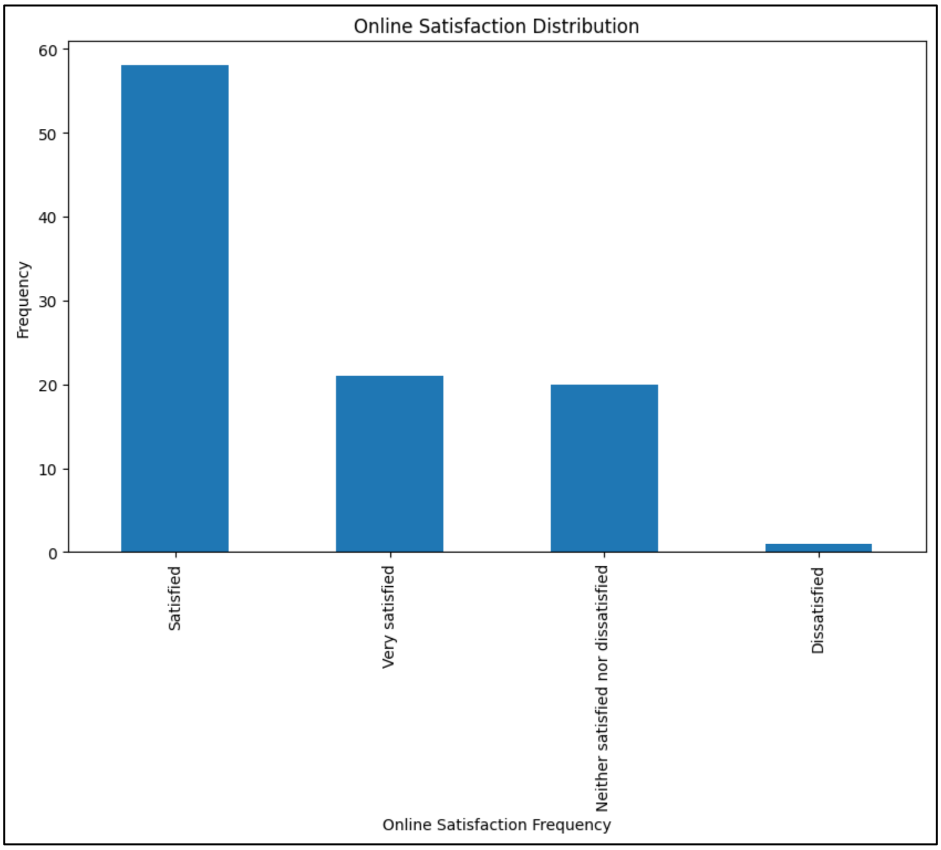


Figure 49: Univariate Analysis (Online Satisfaction Frequency)

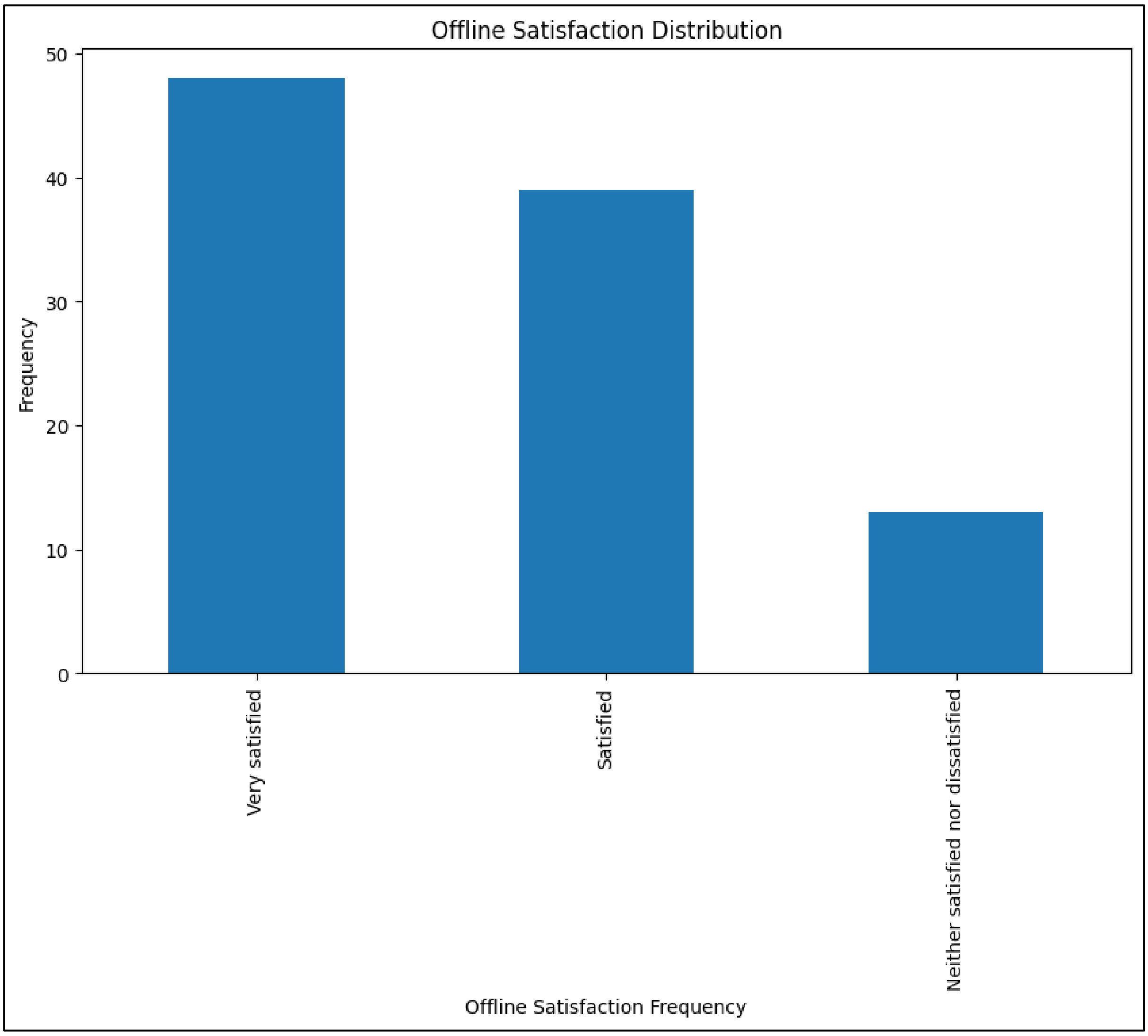


Figure 50: Univariate Analysis (Offline Satisfaction Frequency)

The survey provides insights into consumer satisfaction with both online and offline shopping experiences. For online shopping, only a small fraction (21%) report being very satisfied, highlighting potential improvement areas such as the user interface, delivery times, product quality, or customer service. The majority (58%) feel satisfied, likely appreciating the convenience and variety of online platforms, while a neutral stance by 20% of respondents suggests mixed experiences or indifference. Only 1% express dissatisfaction, indicating a positive but crucial area for businesses to explore and resolve specific complaints.

In contrast, offline shopping shows higher satisfaction, with 48% of respondents feeling very satisfied, likely due to immediate product access, the tangible nature of shopping, and more personalized service. Another 39% are satisfied, underscoring the appeal of the physical shopping experience that includes instant gratification and the sensory aspects of shopping. However, 13% remain neutral, pointing to potential areas that might not be exceptionally negative or positive but could be optimized for better consumer engagement.

Analyzing these findings, it is clear that while offline shopping currently provides a more satisfying experience, online shopping also holds significant satisfaction among consumers. The high rate of neutral responses in both modes calls for further investigation into the factors contributing to this ambivalence, with the aim of transforming these experiences into more engaging and memorable ones. Enhancing the online shopping experience through technological advancements like augmented reality and improving customer service could bridge the satisfaction gap between the two shopping modes, thereby increasing overall consumer satisfaction.

4.3 Bivariate Analysis

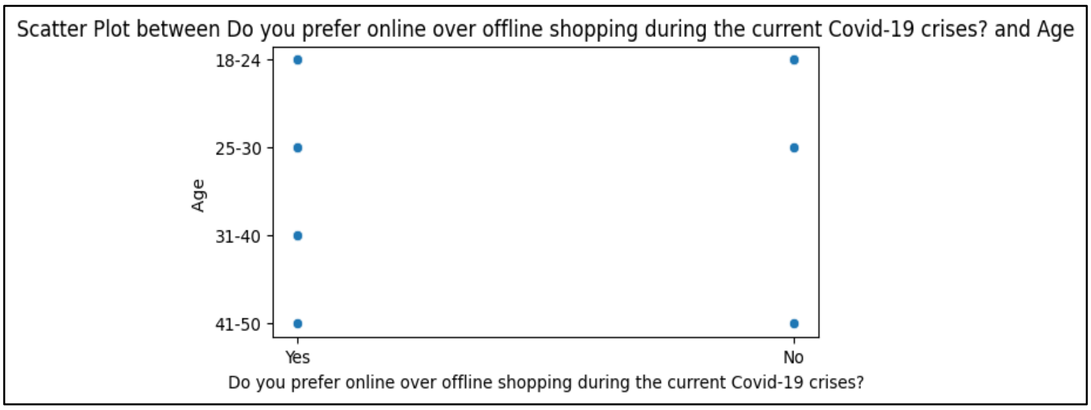


Figure 51: Bivariate Analysis (Age and Shopping Preference)

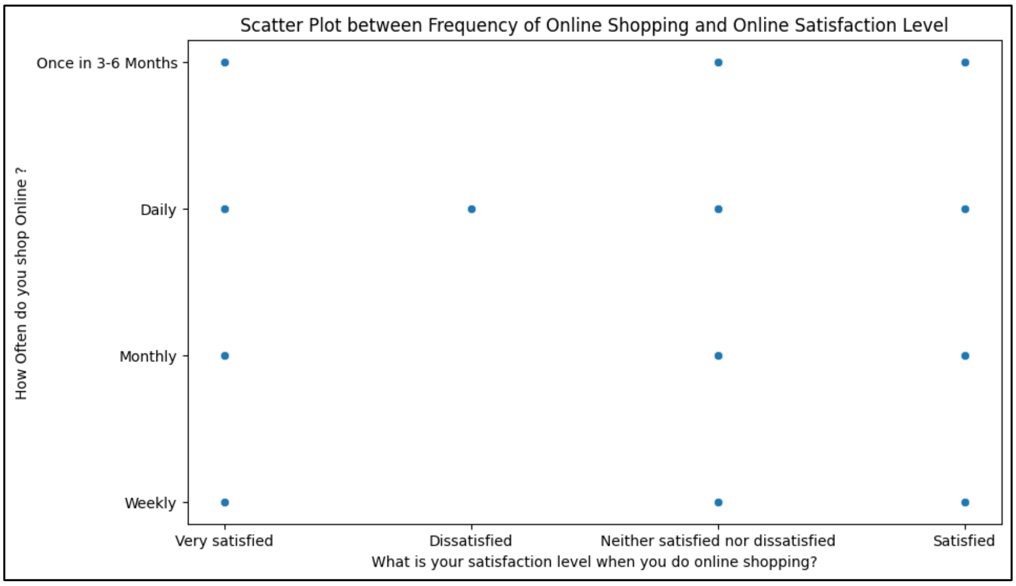


Figure 52: Bivariate Analysis (Online Shopping Satisfaction and Online Shopping Frequency)

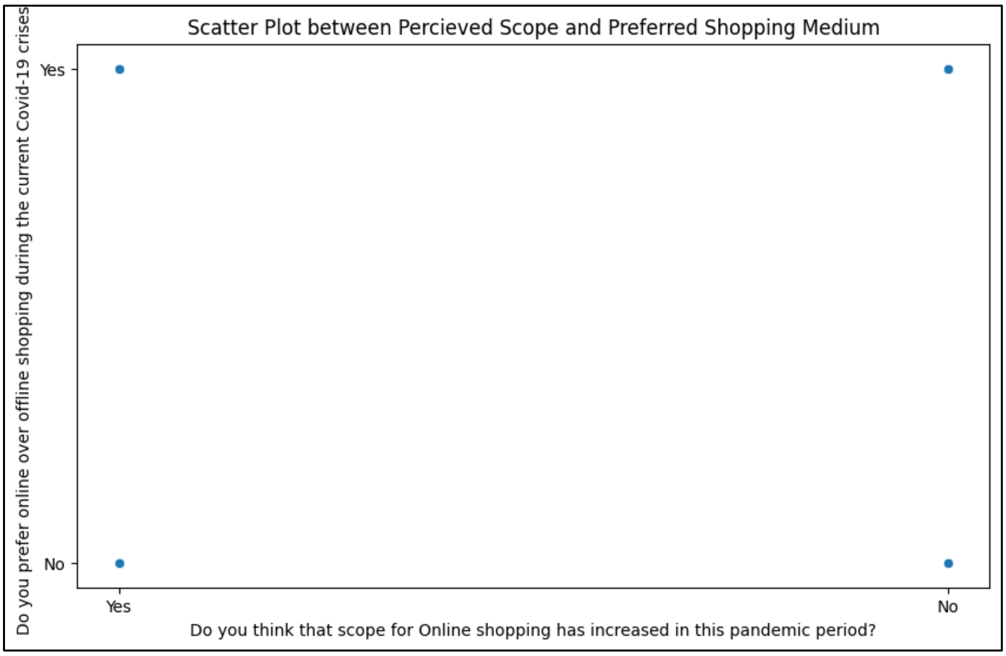


Figure 53: Bivariate Analysis (Perceived Scope and Preferred Shopping Medium)

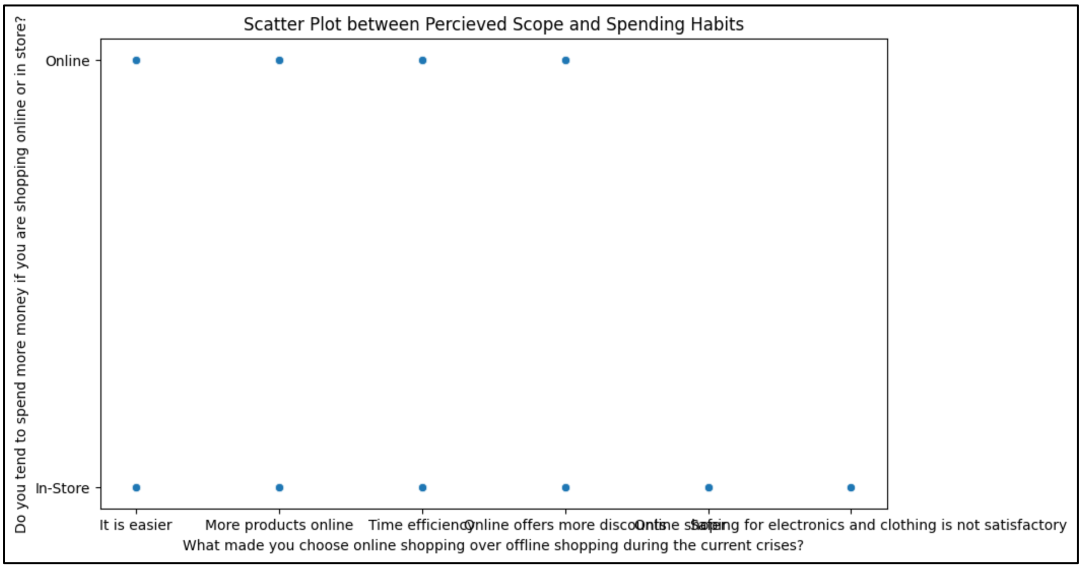


Figure 54: Bivariate Analysis (Preferred Shopping Medium and Spending Habit)

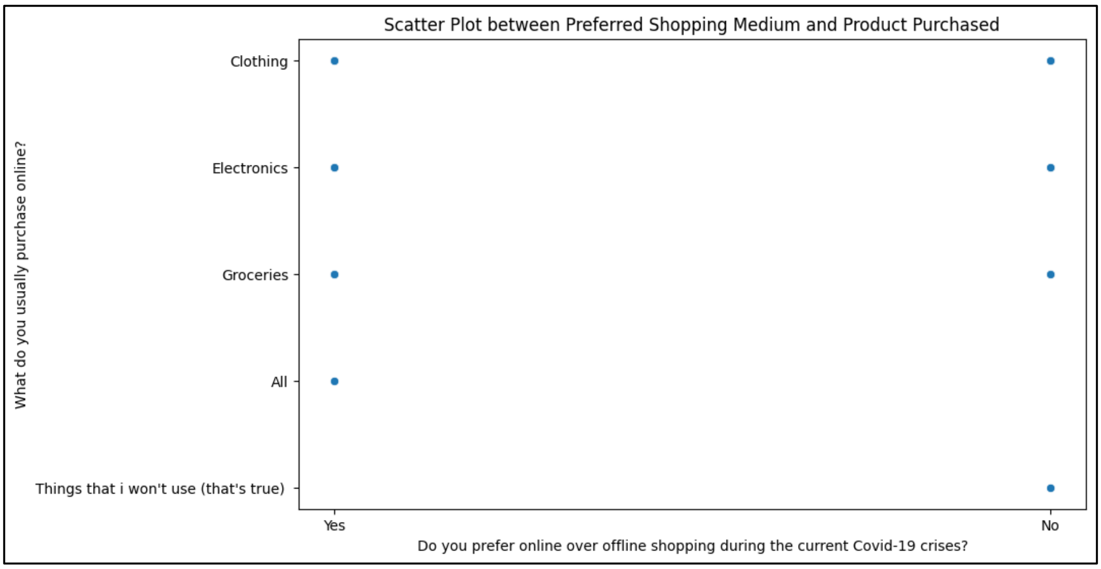
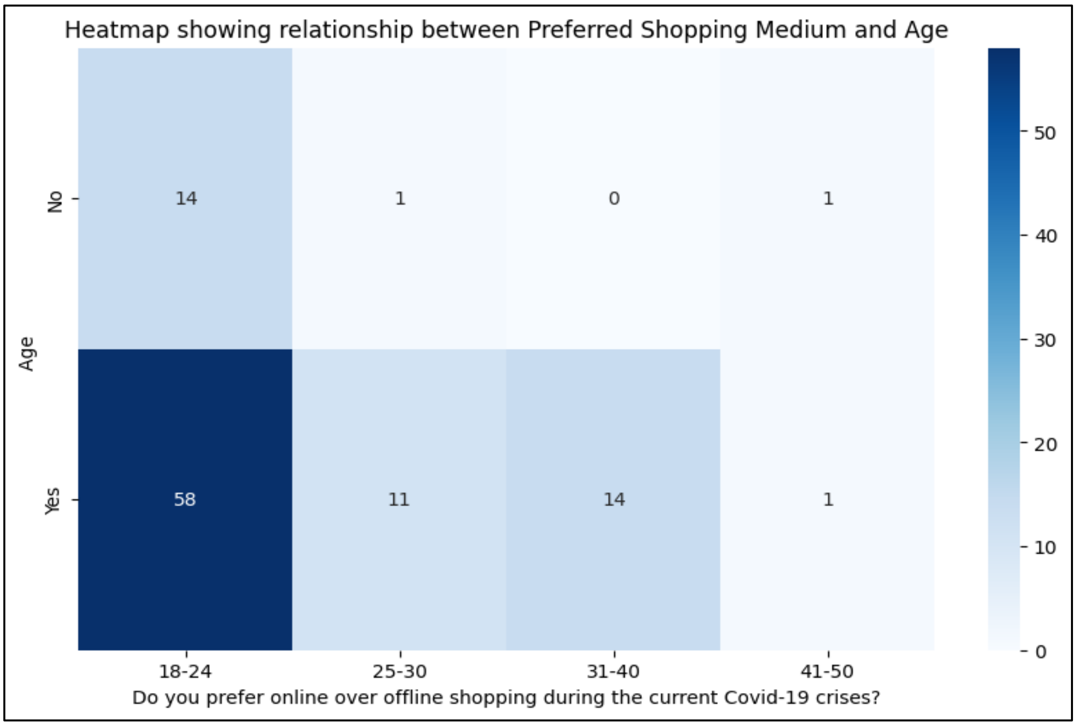


Figure 55: Bivariate Analysis (Preferred Shopping Medium and Product Purchased)

4.4 Statistical Test on Bivariate Analysis

4.4.1 Analysis between Preferred Shopping Medium and Age

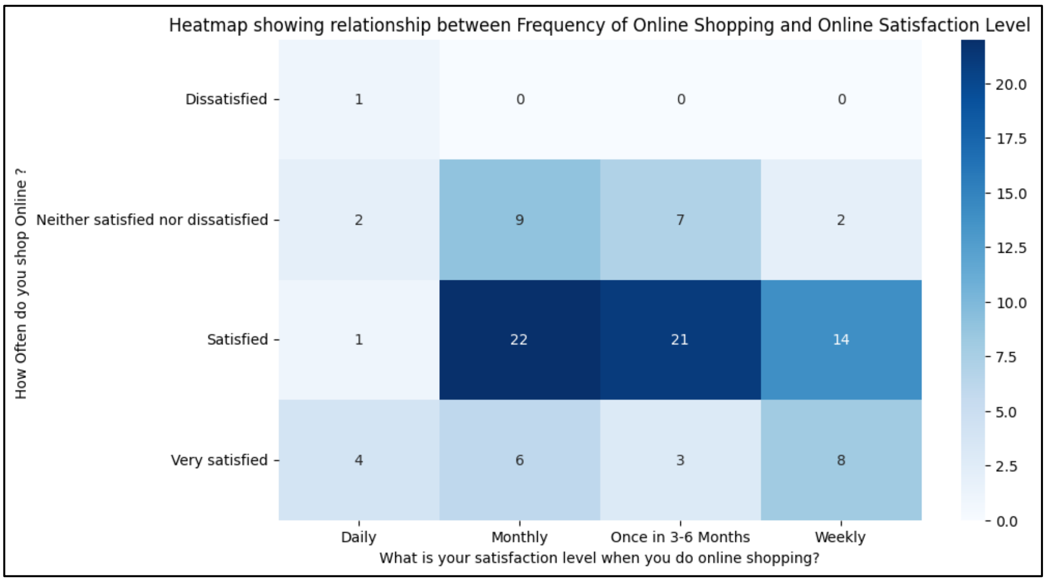



```
Statistical Inferences:
Chi-squared statistic: 5.55
p-value: 0.1358374491537132
Degrees of freedom: 3
Expected Values: [[11.52  1.92  2.24  0.32]
 [60.48 10.08 11.76  1.68]]
```

Figure 56: Statistical Inference (Age and Shopping Preference)

The chi – squared value of 5.55 from this analysis is not big enough to state if there is a significant deviation from what is normally expected if there is no significant relationship between the two features. The p – value of 0.136 is higher than the significance level of 0.05 (5%) and this suggests that there is no significant difference in the age of customers and their preferred shopping medium during the COVID – 19 pandemic hence the null hypothesis (H_o) is accepted, which indicates that age was not a determinant factor of the shopping medium of individuals during the pandemic. For the variables to have relationship, they need to have the expected values as seen in the test result above.

4.4.2 Analysis between Frequency of Online Shopping and Online Shopping Satisfaction Level

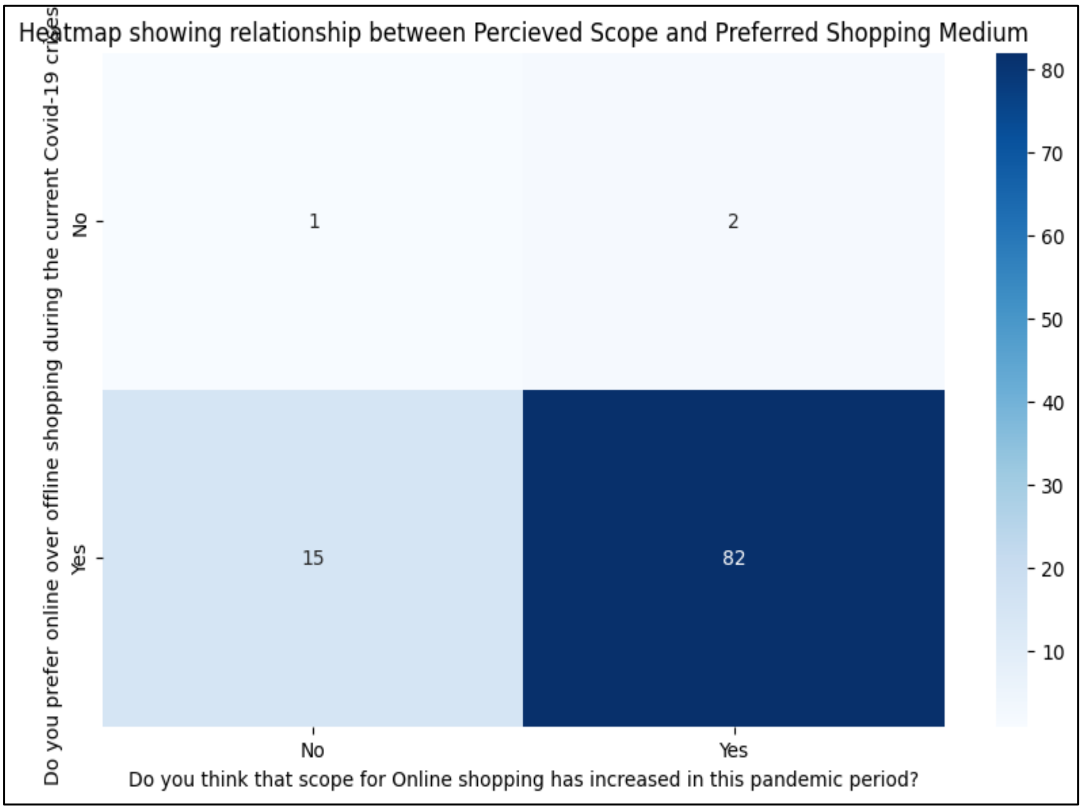



```
Statistical Inferences:
Chi-squared statistic: 24.3
p-value: 0.0038553116210475213
Degrees of freedom: 9
Expected Values: [[ 0.08  0.37  0.31  0.24]
 [ 1.6   7.4   6.2   4.8 ]
 [ 4.64 21.46 17.98 13.92]
 [ 1.68  7.77  6.51  5.04]]
```

Figure 57: Statistical Inference (Online Shopping Satisfaction and Online Shopping Frequency)

The chi – squared value of 24.3 with a degree of freedom 9, indicates that there is a greater difference between the observed data and the data expected under the null hypothesis. The low p – value of 0.003 is lower than the significance level of 0.05 and this indicates that there is a significant difference in the frequency of online shopping and the level of satisfaction from the online shopping; hence the alternative hypothesis (H_a) will be accepted.

4.4.3 Analysis between Perceived Scope and Preferred Shopping Medium

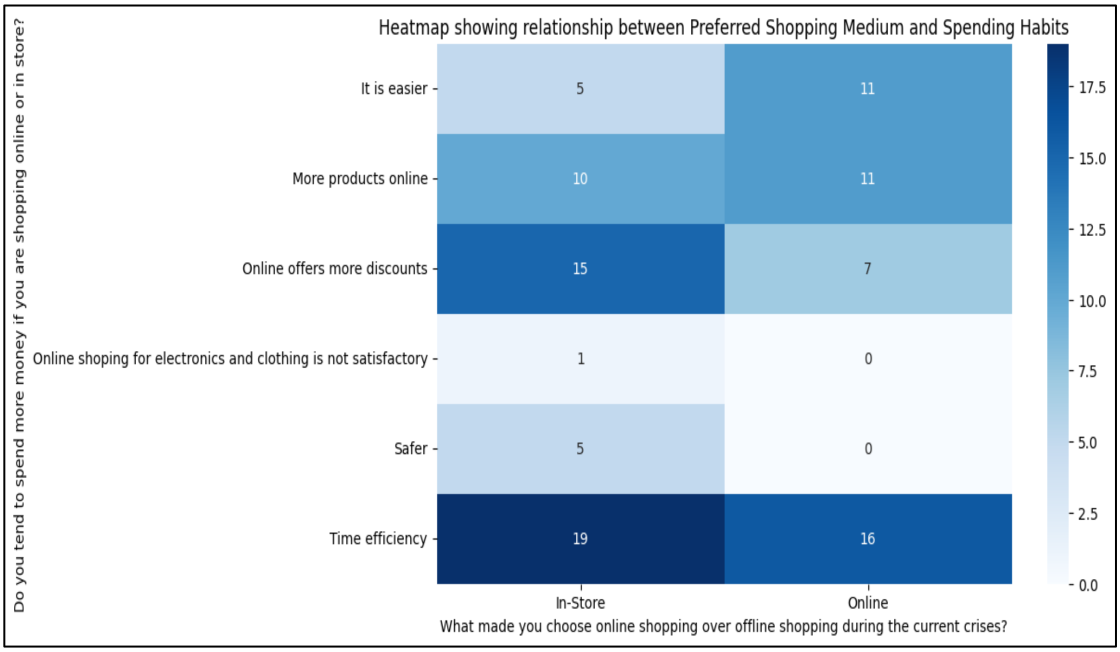



```
Statistical Inferences:
Chi-squared statistic: 0.0010227458681066918
p-value: 0.9744876828830116
Degrees of freedom: 1
Expected Values: [[ 0.48  2.52]
 [15.52 81.48]]
```

Figure 58: Statistical Inference (Perceived Scope and Preferred Shopping Medium)

The chi – squared value of 0.00102274 is relatively small for the analysis of these variables and this indicates a minimal difference between the observed and the expected values. The p-value of 0.9745 is higher than the level of significance (0.05) and this indicates that the null hypothesis (H_0) is accepted which means that there is no significant difference between the perceived scope of shopping and the preferred shopping medium of people during the COVID-19 pandemic.

4.4.4 Analysis between Preferred Shopping Medium and Spending Habit

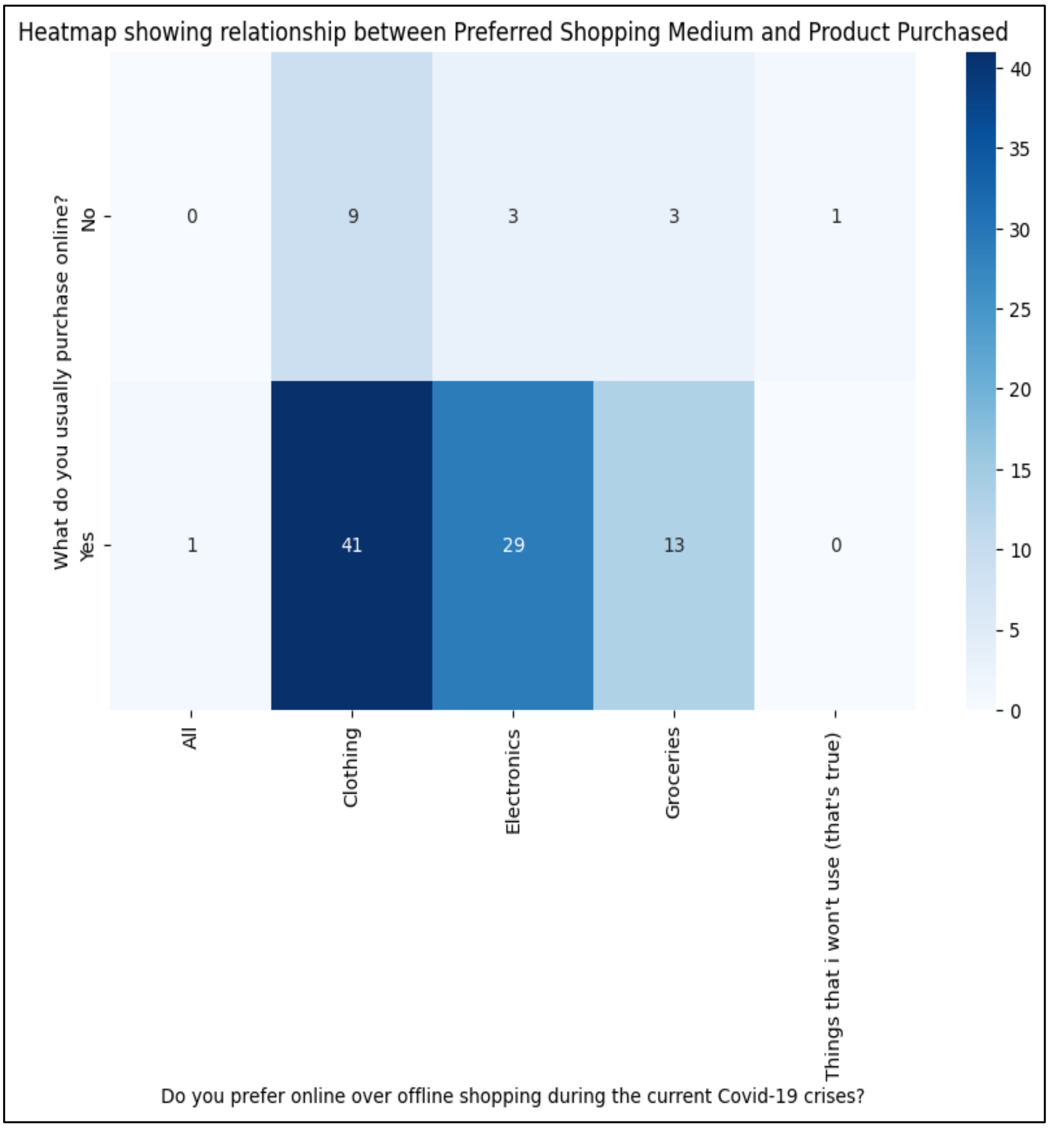



```
Statistical Inferences:
Chi-squared statistic: 10.569548296821026
p-value: 0.060615033004222996
Degrees of freedom: 5
Expected Values: [[ 8.8   7.2 ]
 [11.55  9.45]
 [12.1   9.9 ]
 [ 0.55  0.45]
 [ 2.75  2.25]
 [19.25 15.75]]
```

Figure 59: Statistical Inference (Preferred Shopping Medium and Spending Habit)

The result from this analysis shows that the chi – squared value equals 6.724 and the p-value equals 0.151 and this indicates that there is no significant difference between the variables, hence, the null hypothesis (H_o) is accepted which means that the variables are independent of each other. The chi – squared value explains the difference between the observed and expected values of the variables.

4.4.5 Analysis between Preferred Shopping Medium and Product Purchased




```
Statistical Inferences:
Chi-squared statistic: 6.724330357142857
p-value: 0.15119322206015454
Degrees of freedom: 4
Expected Values: [[ 0.16  8.    5.12  2.56  0.16]
 [ 0.84 42.   26.88 13.44  0.84]]
```

Figure 60: Statistical Inference (Preferred Shopping Medium and Product Purchased)

The chi – squared value of 10.5695 indicates that there are some level of difference between the observed data and the expected data if there is no relationship between the two variables. The p – value of 0.060 is slightly higher than the significance level of 0.05 (5%) which might be due to some errors like measurement error, and so it is not enough reason to accept the null hypothesis. For the purpose of this research, giving room for systematic/margin error (± 0.02), the alternative hypothesis (H_1) would be accepted meaning that there is a significant difference between the perceived scope and spending habits of individuals during the COVID – 19 pandemic, meaning the variables are not independent of themselves.

CHAPTER FIVE
CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

There COVID-19 has noticeably changed the consumer behavior, particularly in the aspect of shopping preferences. The result achieved in this study points the different behavioral strategies among different demographic groups, concentrating on the changes from online to offline shopping in particular. This change, affected greatly by the pandemic, shows a comprehensive pattern in consumer behavior and modern adaptation.

The data shows a notable illustration of younger individuals (age 18 – 24), occupying 72% of the sample. This population changes has possibly skewed the research towards the choice and behaviors of younger consumers, who are literally more minded towards digital platforms. The choice for online shopping is notably embraced by 84% of respondents, highlighting the pandemic’s role in accelerating the acceptance of e-commerce. This choice is pushed by the sensed safety, convenience and expanded possibilities provided by online shopping during the pandemic.

Nevertheless, despite the increasing lean towards online shopping, a significant percentage of consumers whose choice is still offline shopping remains eminent, firstly due to the tactile and instant nature of such occurrences. This preference is evident with 55% of respondents indicating they spend more money in physical stores. This suggests that physical stores still hold a significant appeal, possibly due to the direct access to products and the sensory experience of shopping, which online platforms can't fully replicate.

The survey results give a subtle perspective of consumer satisfaction. As online shopping gives comfortable and a broad range of products, only 21% of respondents are said to be “very satisfied” with their online shopping practice. This denotes possible areas of growth in e-commerce, such as promoting user interfaces, enhancing delivery periods and strengthening customer services.

Contrarily, offline shopping tends to give a level of reasonable satisfaction, with 48% of participant seemed very satisfied. The immediacy of acquiring goods, coupled with the physical interaction with products, contributes significantly to consumer contentment. Yet, 13% neutrality in satisfaction levels suggests there is room for enhancing the physical shopping experience, perhaps by integrating digital tools to streamline service and providing a more personalized shopping journey.

In conclusion, this analysis shows the lofty landscape of consumer choices during the COVID-19 pandemic, unveiling a notable change towards online shopping while also identifying the enduring interest of offline stores. By tackling the highlighted breach and incorporating technological innovations, businesses can well provide to advance consumer needs and promote both online and offline shopping practices.

5.2 Objective Achievement

Relating to the stated objectives of this research work, all objectives have been completed. The first objective was to survey previous works on exploratory data analysis and customer online shopping behaviour before and during the pandemic era. The analysis of this research was done by obtaining a dataset from Kaggle that consisted of information from people as regarding their online shopping habits during the pandemic era and both univariate and bivariate analysis was carried out on this dataset, including some statistical inferences. The results were then evaluated and discussed.

5.3 Technological Integration and Future Opportunities

The data underscores a significant opportunity for blending technological advancements with traditional shopping to enhance consumer experiences. For instance, augmented reality (AR) could simulate the tactile experience of offline shopping online, potentially increasing satisfaction levels. Moreover, this research work can be used for further research work incorporating advanced analytics to understand consumer preferences and tailor online content could bridge the gap between online convenience and offline satisfaction.

5.4 Limitations and Recommendations for Future Research

This study's limitations include a skewed demographic profile, focusing predominantly on younger consumers, and a small sample size, which may not provide a comprehensive overview of broader consumer behaviors. Further research should focus to involve a different population group and host a broad range of consumer behavior and choices. In addition, expanding the survey questions to hold more refined consumer attitudes and extending the research beyond the pandemic scope could give deeper perspective into the long-term inferences of these changes in consumer behavior.

REFERENCE LIST

Ahmed, N. (2022). *Exploratory Data Analysis Simplified 101 - Learn | Hevo*. [online] Available at: https://hevodata.com/learn/exploratory-data-analysis/.

Anderson, M., Faverio, M. and Gottfried, J. (2023). *Teens, Social Media and Technology 2023*. [online] Pew Research Center: Internet, Science & Tech. Available at: https://www.pewresearch.org/internet/2023/12/11/teens-social-media-and-technology-2023/.

Appinio Research (2023). *What Is a Heatmap? Types, How It Works, Examples | Appinio Blog*. [online] www.appinio.com. Available at: https://www.appinio.com/en/blog/market-research/heatmap.

APPINIO RESEARCH (2024). *Cross-Tabulation Analysis: A Full Guide (+ Examples) Appinio Blog*. [online] www.appinio.com. Available at: https://www.appinio.com/en/blog/market-research/cross-tabulation-analysis.

Arshad, T. (2024). *Bivariate Analysis in Data Science: Theory, Tools and Practical Use Cases - Data Science For Bio*. [online] Data Science For Bio. Available at: https://datascienceforbio.com/bivariate-analysis/#:~:text=It%20involves%20analyzing%20the%20relationship%20between%20two%20variables.

Beers, B. (2023). *What P-Value Tells Us*. [online] Investopedia. Available at: https://www.investopedia.com/terms/p/p-value.asp.

Bevans, R. (2020). *Understanding P-values | Definition and Examples*. [online] Scribbr. Available at: https://www.scribbr.com/statistics/p-value/.

Biswal, A. (2023). *What is a Chi-Square Test? Formula, Examples & Uses | Simplilearn*. [online] Simplilearn.com. Available at: https://www.simplilearn.com/tutorials/statistics-tutorial/chi-square-test.

Chevalier, S. (2021). *U.S. teen online shopping reach 2016*. [online] Statista. Available at: https://www.statista.com/statistics/266836/percentage-of-online-shoppers-among-american-teenagers/.

Clark, R. (2021). Bivariate Analyses: Crosstabulation. *pressbooks.ric.edu*. [online] Available at: https://pressbooks.ric.edu/socialdataanalysis/chapter/bivariate/.

Coppola, D. (2021). *E-commerce Worldwide*. [online] www.statista.com. Available at: https://www.statista.com/topics/871/online-shopping/.

Coppola, D. (2024). *UK e-commerce Revenue 2021*. [online] Statista. Available at: https://www.statista.com/statistics/282162/e-commerce-annual-sales-in-the-united-kingdom-uk/#:~:text=The%20United%20Kingdom%20is%20the%20leading%20market%20in.

Creswell, J.W. and Piano Clark, V.L. (2017). *Designing and Conducting Mixed Methods Research*. [online] psycnet.apa.org. Available at: https://psycnet.apa.org/record/2006-11884-000.

Cucinotta, D. and Vanelli, M. (2020). *WHO Declares COVID-19 a Pandemic*. [online] Acta bio-medica: Atenei Parmensis. Available at: https://pubmed.ncbi.nlm.nih.gov/32191675/.

Dalgleish, R. (2020). *How the COVID-19 Pandemic has accelerated the shift to online spending | National Statistical*. [online] Office for National Statistics. Available at: https://blog.ons.gov.uk/2020/09/18/how-the-covid-19-pandemic-has-accelerated-the-shift-to-online-spending/.

Donthu, N. and Gustafsson, A. (2020). Effects of COVID-19 on Business and Research. *Journal of Business Research*, 117(1), pp.284–289.

Dr. V.Dheenadhayalan, S.A. (2021). Impact of E-Commerce on the Changes in Consumer's Buying Behaviour in Malappuram District. *Annals of the Romanian Society for Cell Biology*, [online] pp.3441–3452. Available at: http://annalsofrscb.ro/index.php/journal/article/view/4991/3985.

Duffett , R.G. (2017). *Influence of Social Media Marketing Communications on Young Consumers' attitudes.* Young Consumers 18(1), 19 - 39.

Elida, T., Rahardjo, W., Raharjo, A. and Sukirman, E. (2019). Online Shopping: What Factors Determine Consumers to Buy? *Management Studies*. [online] Available at: https://www.semanticscholar.org/paper/Online-Shopping%3A-What-Factors-Determine-Consumers-Elida-Rahardjo/f08e844df0bc6b462cde4a4770a66086a0027541

Ganti, A. (2023). *Degrees of Freedom Definition*. [online] Investopedia. Available at: https://www.investopedia.com/terms/d/degrees-of-freedom.asp.

GeeksforGeeks (2023). *Getting Started with Google Colab*. [online] GeeksforGeeks. Available at: https://www.geeksforgeeks.org/getting-started-with-google-colab/.

Glen, S. (2014). *Univariate Analysis: Definition, Examples*. [online] Statistics How To. Available at: https://www.statisticshowto.com/univariate/.

Goleckyte, E. (2023). *Student Shopping Habits in the Digital Age: E-commerce Trends and Insights*. [online] Size-Charts.com - When size matters. Available at: https://size-charts.com/blog/ecommerce/student-shopping-habits-in-the-digital-age/.

Gulfraz, M.B., Sufyan, M., Mustak, M., Salminen, J. and Srivastava, D.K. (2022). Understanding the Impact of Online Customers' Shopping Experience on Online Impulsive buying: a Study on Two Leading E-commerce Platforms. *Journal of Retailing and Consumer Services*, [online] 68(103000), p.103000. doi: https://doi.org/10.1016/j.jretconser.2022.103000.

Halladay, K. (2023). *The 12 Best Data Visualization Tools for Professionals | Built In*. [online] builtin.com. Available at: https://builtin.com/data-science/data-visualization-tools.

Hayes, A. (2023). *How the Chi Square Statistic Works*. [online] Investopedia. Available at: https://www.investopedia.com/terms/c/chi-square-statistic.asp.

Hayes, M. and Downie, A. (2024). *What is Ecommerce? | IBM*. [online] www.ibm.com. Available at: https://www.ibm.com/topics/ecommerce.

IBM (2020). *What Is Exploratory Data Analysis? | IBM*. [online] www.ibm.com. Available at: https://www.ibm.com/topics/exploratory-data-analysis.

Ibrahim, N.A.N., Husseini, F., Aisyah, N., Khazlyn, N., Faiz, W. and Kahar, N. (2023). Online Shopping Behaviour in Youth: a Systematic Review of the Factors Influencing Online Shopping in Young Adults. *International Journal of Academic Research in Business and Social Sciences*, [online] 13(2). Available at:

https://www.academia.edu/98884669/Online_Shopping_Behaviour_in_Youth_A_Systematic_Review_of_The_Factors_Influencing_Online_Shopping_in_Young_Adults.

Jackson, T. and Wang, Y. (2013). *Cultural Differences and Consumer Behavior: A Study of Tourism and Online shopping. International Journal of Business and Management*. 8(1) ed.

Kenton, W. (2019). *Barter (or Bartering) Definition*. [online] Investopedia. Available at: https://www.investopedia.com/terms/b/barter.asp.

Kotler, P., Keller, K., Brady, M., Goodman, M. and Hansen, T. (2016). *Marketing Management*. [online] elibrary.pearson.de. Available at: https://elibrary.pearson.de/book/99.150005/9781292106083.

McKinsey & Company (2021). *The Value of Getting Personalization right--or wrong--is Multiplying*. [online] McKinsey & Company. Available at: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-value-of-getting-personalization-right-or-wrong-is-multiplying.

Nguyen, H. (2023). *Exploratory Data Analysis: Everything you need to know*. [online] TECHVIFY Software. Available at: https://techvify-software.com/exploratory-data-analysis/.

Sahoo, K., Samal, A.K., Pramanik, J. and Pani, S.K. (2019). *Volume-8 Issue-12, October 2019*. [online] International Journal of Innovative Technology and Exploring Engineering (IJITEE). Available at: https://www.ijitee.org/download/volume-8-issue-12/

Saleem, N., Khalid, A. and Sadiq, S. (2022). Impact of e-commerce on Consumer Behavior: A Study of Online Shopping Trend among Youth in Lahore City. *Indian Journal of Economics and Business*, [online] 21(1). Available at: https://www.ashwinanokha.com/resources/ijeb%20v21-1-2022-68.pdf.

Santos, M. (2023). *A Data Scientist's Essential Guide to Exploratory Data Analysis*. [online] Medium. Available at: https://towardsdatascience.com/a-data-scientists-essential-guide-to-exploratory-data-analysis-25637eee0cf6.

Sree Saraswathi, G. and College, T. (2020). Exploring the Influential Factors of Online Consumer Shopping Habits and Intention in GCC Special Reference to Kingdom of Bahrain. *International Journal of Innovative*

Science and Research Technology, [online] 5(4). Available at: https://www.ijisrt.com/assets/upload/files/IJISRT20APR480.pdf.

Trantor (2024). *Exploratory Data Analysis (EDA): Types, Tools, Process*. [online] Trantorinc. Available at: https://www.trantorinc.com/blog/exploratory-data-analysis.

Turney, S. (2022). *How to Find Degrees of Freedom* | *Definition & Formula*. [online] Scribbr. Available at: https://www.scribbr.com/statistics/degrees-of-freedom/.

Wikipedia Contributors (2019). *Exploratory data analysis*. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Exploratory_data_analysis.

Zach (2021a). *How to Perform Bivariate Analysis in Python (With Examples)*. [online] Statology. Available at: https://www.statology.org/bivariate-analysis-in-python/.

Zach (2021b). *What Is Univariate Analysis? (Definition & Example)*. [online] Statology. Available at: https://www.statology.org/univariate-analysis/.

APPENDIX

Link to Google Colab

https://colab.research.google.com/drive/1v73lkujmZGa79XZaLOKsQ6_5HQ2LF9F4#scrollTo=pEqrcIXmjN4h

```
#to import necessary libraries needed
from google.colab import files
import pandas as pd
import numpy as np
import scipy.stats as stats
import matplotlib.pyplot as plt
import seaborn as sns
```

[ ]

```
uploaded = files.upload()
```

[ ]

```
#to view the data in the dataset
covid_data = pd.read_csv('Impact of COVID dataset.csv')
covid_data.head()
```

[ ]

```
 #to check basic info of the dataset
 covid_data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 16 columns):
 #                                                Column
Non-Null Count  Dtype
0                                                 Timestamp
100 non-null    object
 1                                                Name
100 non-null    object
```

2 Profession
100 non-null object
3 Age
100 non-null object
4 Do you prefer online over offline shopping during the current Covid-19 crises? 100 non-null object
5 Do you think that scope for Online shopping has increased in this pandemic period? 100 non-null object
6 Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic? 100 non-null object
7 Do you tend to spend more money if you are shopping online or in store?
100 non-null object
8 How Often do you shop Online ?
100 non-null object
9 How much time do you spend in every visit?
100 non-null object
10 What do you usually purchase online?
100 non-null object
11 What do you usually purchase offline?
100 non-null object
12 What is your satisfaction level when you do online shopping?
100 non-null object
13 What is your satisfaction level when you do offline shopping?
100 non-null object
14 What made you choose online shopping over offline shopping during the current crises? 100 non-null object
15 What made you choose offline shopping over online shopping during the current crises? 100 non-null object
dtypes: object(16)
memory usage: 12.6+ KB

[ ]

```
#to check for any missing value
covid_data.isna().sum()
```

[ ]

```
#to drop none needed columns
covid_new_data = covid_data.drop(columns = ['Timestamp', 'Name'],)
covid_new_data.head()
```

[ ]

```
#to describe the dataset
covid_new_data.describe()
```

[ ]

```
#to count the unique values in each column
covid_new_data.nunique()
```

[ ]

```
#Univariate analysis
#to count each unique profession
prof_count = covid_new_data["Profession"].value_counts()
print("Profession Counts:\n", prof_count)

prof_data = covid_new_data['Profession'].value_counts()
plt.figure(figsize=(10,6))
prof_data.plot(kind='bar')
plt.title('Profession Distribution')
plt.xlabel('Profession')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count each unique age
count_age = covid_new_data['Age '].value_counts()
```

```
print("Age Counts:\n", count_age)

age_data = covid_new_data['Age '].value_counts()
plt.figure(figsize=(10,6))
age_data.plot(kind='bar')
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count each online and offline preference
preference_counts = covid_new_data["Do you prefer online over offline shopping during the current Covid-19 crises?"].value_counts()
print("Preference Counts:\n", preference_counts)

pref_data = covid_new_data["Do you prefer online over offline shopping during the current Covid-19 crises?"].value_counts()
plt.figure(figsize=(10,6))
pref_data.plot(kind='bar')
plt.title('Online Versus Offline Distribution')
plt.xlabel('Online Versus Offline')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the online impact
online_imp_counts = covid_new_data["Do you think that scope for Online shopping has increased in this pandemic period?"].value_counts()
print("Online Impact Counts:\n", online_imp_counts)

online_data = covid_new_data["Do you think that scope for Online shopping has increased in this pandemic period?"].value_counts()
plt.figure(figsize=(10,6))
```

```
online_data.plot(kind='bar')
plt.title('Online Impact Distribution')
plt.xlabel('Online Impact')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the loss rate
loss_rate_counts = covid_new_data["Do u think that offline stores have incur
red loss due to shift in trend to online Shopping in this pandemic?"].value_co
unts()
print("Loss Rate Counts:\n", loss_rate_counts)

loss_rate_data = covid_new_data["Do u think that offline stores have incurre
d loss due to shift in trend to online Shopping in this pandemic?"].value_cou
nts()
plt.figure(figsize=(10,6))
loss_rate_data.plot(kind='bar')
plt.title('Loss Rate Distribution')
plt.xlabel('Loss Rate')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the spending rate
spending_rate_counts = covid_new_data["Do you tend to spend more mone
y if you are shopping online or in store?"].value_counts()
print("Spending Rate Counts:\n", spending_rate_counts)

spending_rate_data = covid_new_data["Do you tend to spend more money i
f you are shopping online or in store?"].value_counts()
plt.figure(figsize=(10,6))
spending_rate_data.plot(kind='bar')
plt.title('Spending Rate Distribution')
```

```
plt.xlabel('Spending Rate')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the shopping frequency
shopping_freq_counts = covid_new_data["How Often do you shop Online ?
"].value_counts()
print("Shopping Frequency Counts:\n", shopping_freq_counts)

shopping_freq_data = covid_new_data["How Often do you shop Online ? "].
value_counts()
plt.figure(figsize=(10,6))
shopping_freq_data.plot(kind='bar')
plt.title('Shopping Frequency Distribution')
plt.xlabel('Spending Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the time frequency
time_freq_counts = covid_new_data["How much time do you spend in every
 visit? "].value_counts()
print("Time Frequency Counts:\n", time_freq_counts)

time_freq_data = covid_new_data["How much time do you spend in every vi
sit? "].value_counts()
plt.figure(figsize=(10,6))
time_freq_data.plot(kind='bar')
plt.title('Shopping Time Frequency Distribution')
plt.xlabel('Time Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the online purchase frequency
online_pur_counts = covid_new_data["What do you usually purchase online
?"].value_counts()
print("Online Purchase Counts:\n", online_pur_counts)

online_pur_data = covid_new_data["What do you usually purchase online?"]
.value_counts()
plt.figure(figsize=(10,6))
online_pur_data.plot(kind='bar')
plt.title('Online Purchase Distribution')
plt.xlabel('Online Purchase Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the offline purchase frequency
offline_pur_counts = covid_new_data["What do you usually purchase offline
?"].value_counts()
print("Offline Purchase Counts:\n", offline_pur_counts)

offline_pur_data = covid_new_data["What do you usually purchase offline?"]
.value_counts()
plt.figure(figsize=(10,6))
offline_pur_data.plot(kind='bar')
plt.title('Offline Purchase Distribution')
plt.xlabel('Offline Purchase Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the online satisfaction frequency
online_sat_counts = covid_new_data["What is your satisfaction level when y
ou do online shopping?"].value_counts()
print("Online Satisfaction Counts:\n", online_sat_counts)

online_sat_data = covid_new_data["What is your satisfaction level when you
 do online shopping?"].value_counts()
plt.figure(figsize=(10,6))
online_sat_data.plot(kind='bar')
plt.title('Online Satisfaction Distribution')
plt.xlabel('Online Satisfaction Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the offline satisfaction frequency
offline_sat_counts = covid_new_data["What is your satisfaction level when y
ou do offline shopping?"].value_counts()
print("Offline Satisfaction Counts:\n", offline_sat_counts)

offline_sat_data = covid_new_data["What is your satisfaction level when you
 do offline shopping?"].value_counts()
plt.figure(figsize=(10,6))
offline_sat_data.plot(kind='bar')
plt.title('Offline Satisfaction Distribution')
plt.xlabel('Offline Satisfaction Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the online preference frequency
online_pref_counts = covid_new_data["What made you choose online shop
ping over offline shopping during the current crises?"].value_counts()
print("Online Preference Counts:\n", online_pref_counts)
```

```
online_pref_data = covid_new_data["What made you choose online shoppin
g over offline shopping during the current crises?"].value_counts()
plt.figure(figsize=(10,6))
online_pref_data.plot(kind='bar')
plt.title('Online Preference Distribution')
plt.xlabel('Online Preference Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#to count the offline preference frequency
offline_pref_counts = covid_new_data["What made you choose offline shoppi
ng over online shopping during the current crises?"].value_counts()
print("Offline Preference Counts:\n", offline_pref_counts)
offline_pref_data = covid_new_data["What made you choose offline shopping
 over online shopping during the current crises?"].value_counts()
plt.figure(figsize=(10,6))
offline_pref_data.plot(kind='bar')
plt.title('Offline Preference Distribution')
plt.xlabel('Offline Preference Frequency')
plt.ylabel('Frequency')
plt.show()
```

[ ]

```
#bivariate analysis
plt.figure(figsize=(6, 3))
sns.scatterplot(x = 'Do you prefer online over offline shopping during the curr
ent Covid-19 crises?', y = 'Age ', data = covid_new_data)
plt.title('Scatter Plot between Do you prefer online over offline shopping duri
ng the current Covid-19 crises? and Age')
plt.xlabel('Do you prefer online over offline shopping during the current Covi
d-19 crises?')
```

```
plt.ylabel('Age ')
plt.show()
```

[ ]

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'What is your satisfaction level when you do online shopp
ing?', y = 'How Often do you shop Online ? ', data = covid_new_data)
plt.title('Scatter Plot between Frequency of Online Shopping and Online Sati
sfaction Level')
plt.xlabel('What is your satisfaction level when you do online shopping?')
plt.ylabel('How Often do you shop Online ?')
plt.show()
```

[ ]

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'Do you think that scope for Online shopping has increas
ed in this pandemic period?', y = 'Do you prefer online over offline shopping
during the current Covid-19 crises?', data = covid_new_data)
plt.title('Scatter Plot between Percieved Scope and Preferred Shopping Medi
um')
plt.xlabel('Do you think that scope for Online shopping has increased in this
pandemic period?')
plt.ylabel('Do you prefer online over offline shopping during the current Covi
d-19 crises?')
plt.show()
```

[ ]

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'Do you prefer online over offline shopping during the curr
ent Covid-
19 crises?', y = 'What do you usually purchase online?', data = covid_new_d
ata)
plt.title('Scatter Plot between Preferred Shopping Medium and Product Purc
```

```
hased')
plt.xlabel('Do you prefer online over offline shopping during the current Covi
d-19 crises?')
plt.ylabel('What do you usually purchase online?')
plt.show()
```

[ ]

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x = 'What made you choose online shopping over offline sho
pping during the current crises?', y = 'Do you tend to spend more money if y
ou are shopping online or in store?', data = covid_new_data)
plt.title('Scatter Plot between Percieved Scope and Spending Habits')
plt.xlabel('What made you choose online shopping over offline shopping duri
ng the current crises?')
plt.ylabel('Do you tend to spend more money if you are shopping online or in
 store?')
plt.show()
```

[ ]

```
#statistical test
#crosstabulation
crosstab = pd.crosstab(covid_new_data['Do you prefer online over offline sh
opping during the current Covid-19 crises?'], covid_new_data['Age '])
crosstab

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium
 and Age')
plt.xlabel('Do you prefer online over offline shopping during the current Covi
d-19 crises?')
plt.ylabel('Age ')
plt.show()
```

```
from scipy.stats import chi2_contingency
chi2, p, dof, expected = chi2_contingency(crosstab)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', round(chi2, 2))
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

[ ]

```
#crosstabulation
crosstab_2 = pd.crosstab(covid_new_data['What is your satisfaction level w
hen you do online shopping?'], covid_new_data['How Often do you shop Onl
ine ? '])
crosstab_2

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_2, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Frequency of Online Shoppi
ng and Online Satisfaction Level')
plt.xlabel('What is your satisfaction level when you do online shopping?')
plt.ylabel('How Often do you shop Online ? ')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_2)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
```

```
print('Chi-squared statistic:', round(chi2, 2))
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

[ ]

```
#crosstabulation
crosstab_3 = pd.crosstab(covid_new_data['Do you think that scope for Onlin
e shopping has increased in this pandemic period?'], covid_new_data['Do yo
u prefer online over offline shopping during the current Covid-19 crises?'])
crosstab_3

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_3, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Percieved Scope and Prefer
red Shopping Medium')
plt.xlabel('Do you think that scope for Online shopping has increased in this
pandemic period?')
plt.ylabel('Do you prefer online over offline shopping during the current Covi
d-19 crises?')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_3)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

[ ]

```
#crosstabulation
crosstab_4 = pd.crosstab(covid_new_data['What made you choose online shopping over offline shopping during the current crises?'],
                    covid_new_data['Do you tend to spend more money if you are shopping online or in store?'])
crosstab_4

#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_4, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium and Spending Habits')
plt.xlabel('What made you choose online shopping over offline shopping during the current crises?')
plt.ylabel('Do you tend to spend more money if you are shopping online or in store?')
plt.show()

#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_4)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)

print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```

[ ]

```
#crosstabulation
crosstab_5 = pd.crosstab(covid_new_data['Do you prefer online over offline shopping during the current Covid-19 crises?'],
```

```
                    covid_new_data['What do you usually purchase online?'])
crosstab_5
#heatmap
plt.figure(figsize=(10,6))
sns.heatmap(crosstab_5, annot=True, cmap='Blues')
plt.title('Heatmap showing relationship between Preferred Shopping Medium
and Product Purchased')
plt.xlabel('Do you prefer online over offline shopping during the current Covid
-19 crises?')
plt.ylabel('What do you usually purchase online?')
plt.show()
#ch-squared test
chi2, p, dof, expected = chi2_contingency(crosstab_5)
rounded_expected = np.round(expected, 2)
np.set_printoptions(suppress=True)
print("Statistical Inferences:")
print('Chi-squared statistic:', chi2)
print('p-value:', p)
print('Degrees of freedom:', dof)
print('Expected Values:', rounded_expected)
```