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

PRESENTED BY:

ISOLA, KEHINDE OLUWAFISAYOMI

ID: 22185897

Presentation Content

- Introduction
- Research Question
- Definition of Terms
- Methodology
- Univariate Analysis
- Bivariate Analysis
- Statistical Inference
- Conclusion

INTRODUCTION

- Online shopping was facilitated by the rise of E-commerce and it has significantly boosted sales for many businesses worldwide

- It has also brought about notable changes in consumer market behaviour across various regions.

- The COVID - 19 pandemic has notably affected the behaviourial pattern of consumers across different demographics which also brought about a large demand in online shopping.

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 to respond win a significant way by enforcing some policies and restrictions that guarantee safety of individuals (Cucinotta and Vanelli, 2020) for the safety of the people which facilitates the stoppage of some business operations due to the sit at home policy and various social welfare provision amongst others.

- Following the restriction, the pandemic accelerated a shift in online shopping behaviours as most sales were forced to happen online, 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.

Research Question

What are the different factors that influence the online shopping pattern of different people?

AIMS AND OBJECTIVES

AIM

To discover the online shopping trends among different people – Post – COVID era.

OBJECTIVES

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.

DEFINITION OF TERMS

E-COMMERCE

E-commerce 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.

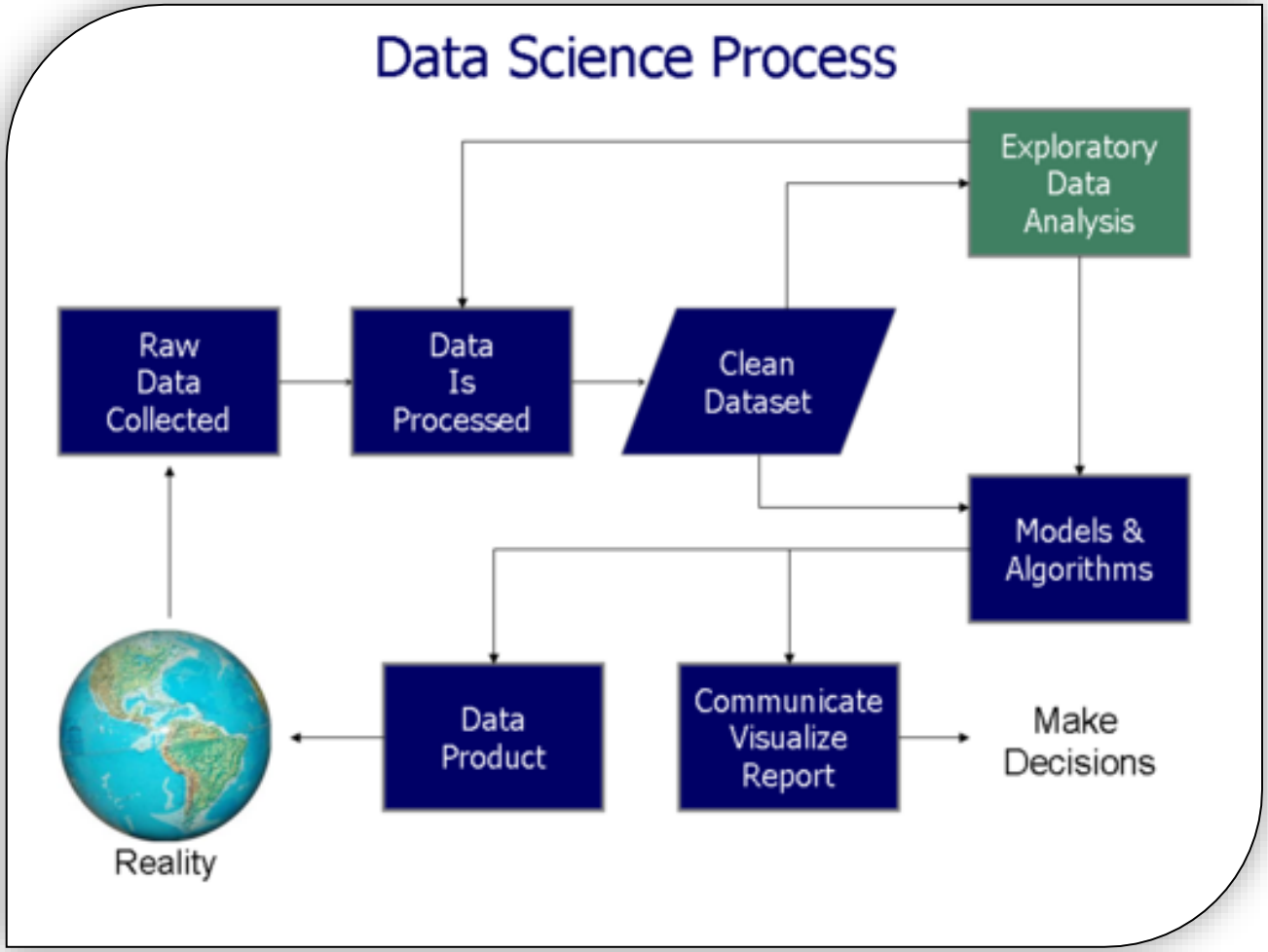
CONSUMER BEHAVIOUR

This is the study of how individuals make decisions about what they want, need or act in regards to a product or service

EXPLORATORY DATA ANALYSIS (EDA)

It is defined as a mechanism mainly used by data scientists to study and understand different datasets and also summarize their specific features, and visualize it with right pictorial representations.

Methodology



Methodology

FIRST FIVE ROWS OF 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

Methodology

BASIC INFO OF 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
```

Methodology

FIRST FIVE ROWS OF THE PROCESSED DATASET

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

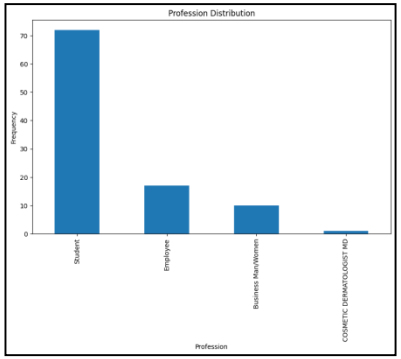
	Profession	Age	Do you prefer online over offline shopping during the current Covid-19 crises?	Do you think that scope for Online shopping has increased in this pandemic period?	Do u think that offline stores have incurred loss due to shift in trend to online Shopping in this pandemic?	Do you tend to 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

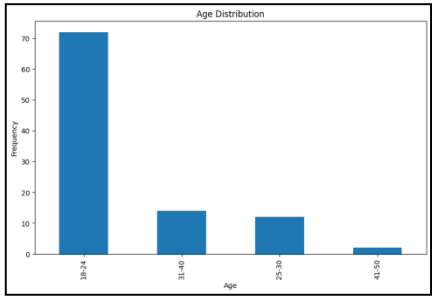
Univariate Analysis

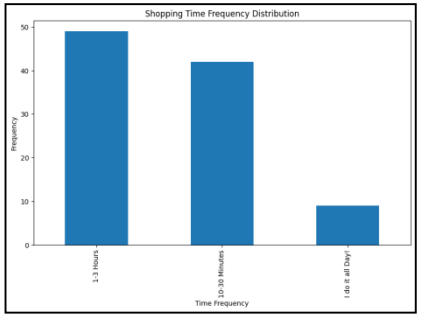
- Univariate analysis is the simplest and basic form of analysing and understanding data. The word “Uni” means “one”, which means the data has only one variable which can be analysed (Glen, 2014).

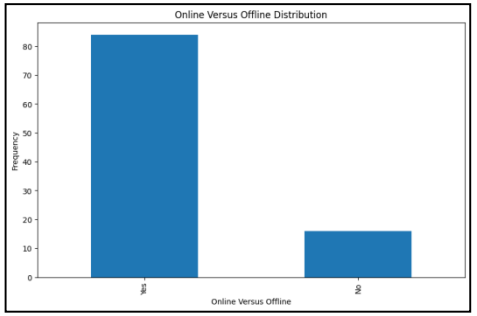
- For the purpose of this research, univariate analysis was carried out on each variable in the dataset by plotting bar charts to understand the frequency distribution of each column.

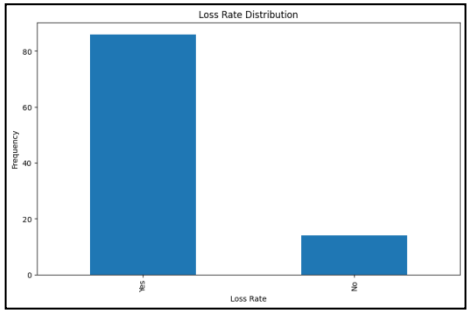
Univariate Analysis

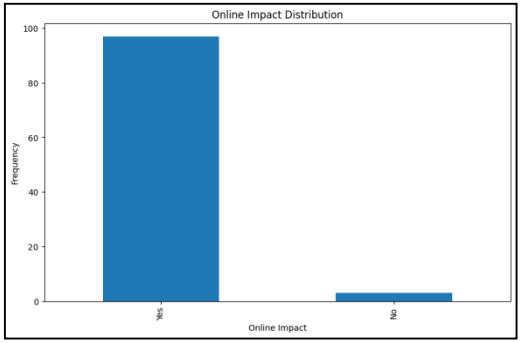


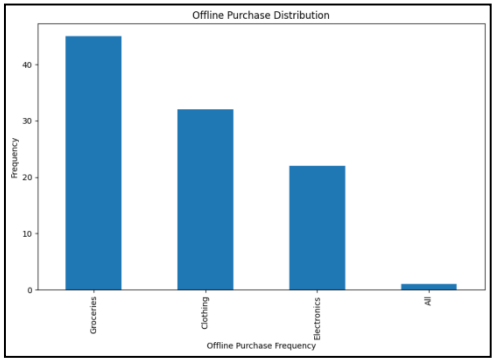






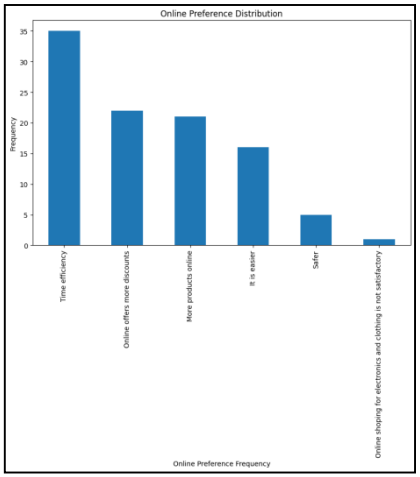


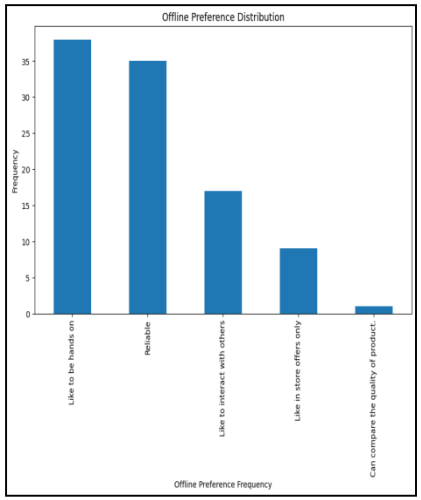


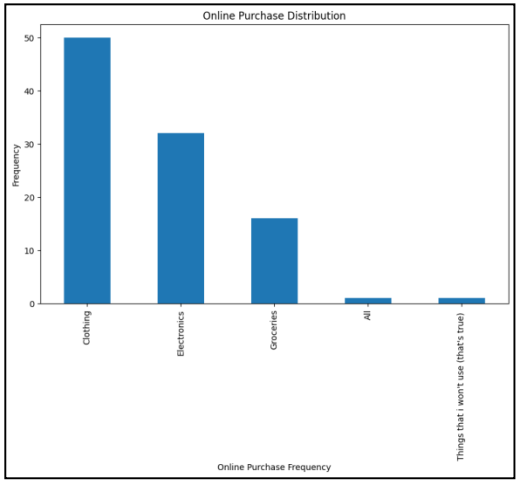


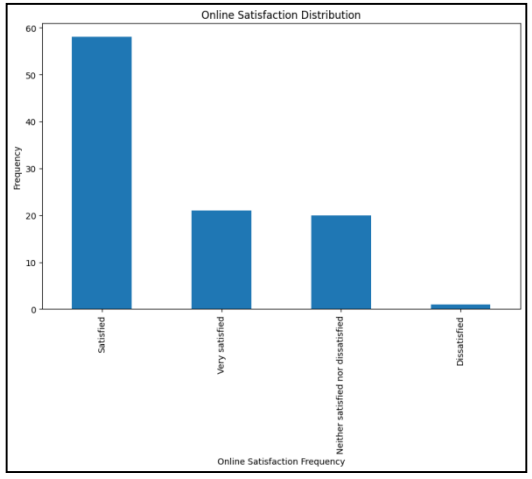
BIRMINGHAM CITY University

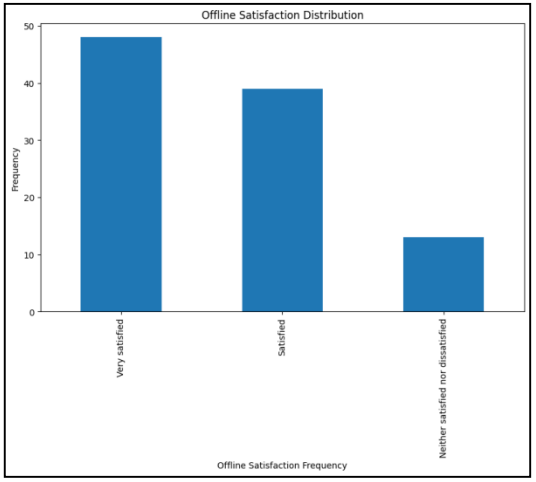
Univariate Analysis











BIRMINGHAM CITY University

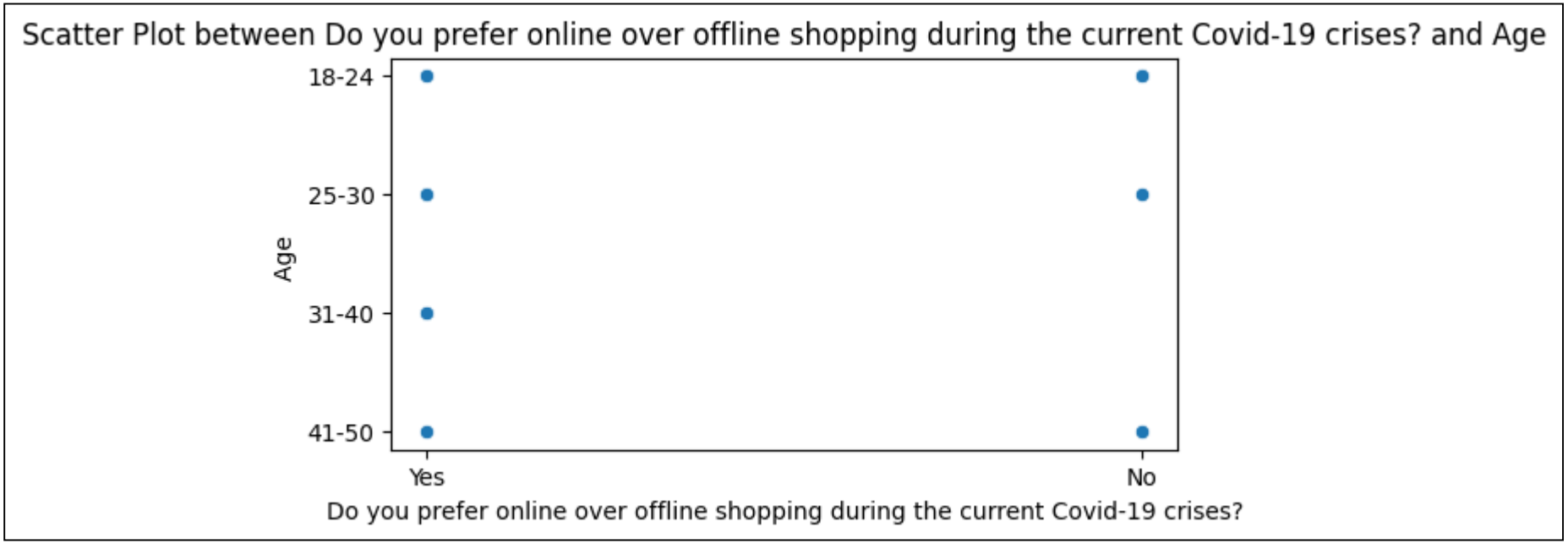
Bivariate Analysis

- 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.
- There are three ways to perform bivariate analysis such as scatterplots, correlation coefficients and simple linear regression (Zach, 2021a).
- Bivariate analysis can be done using various tools such as R, Python, MATLAB, SPSS, SAS, Minitab, Excel, STATA (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.

Bivariate Analysis

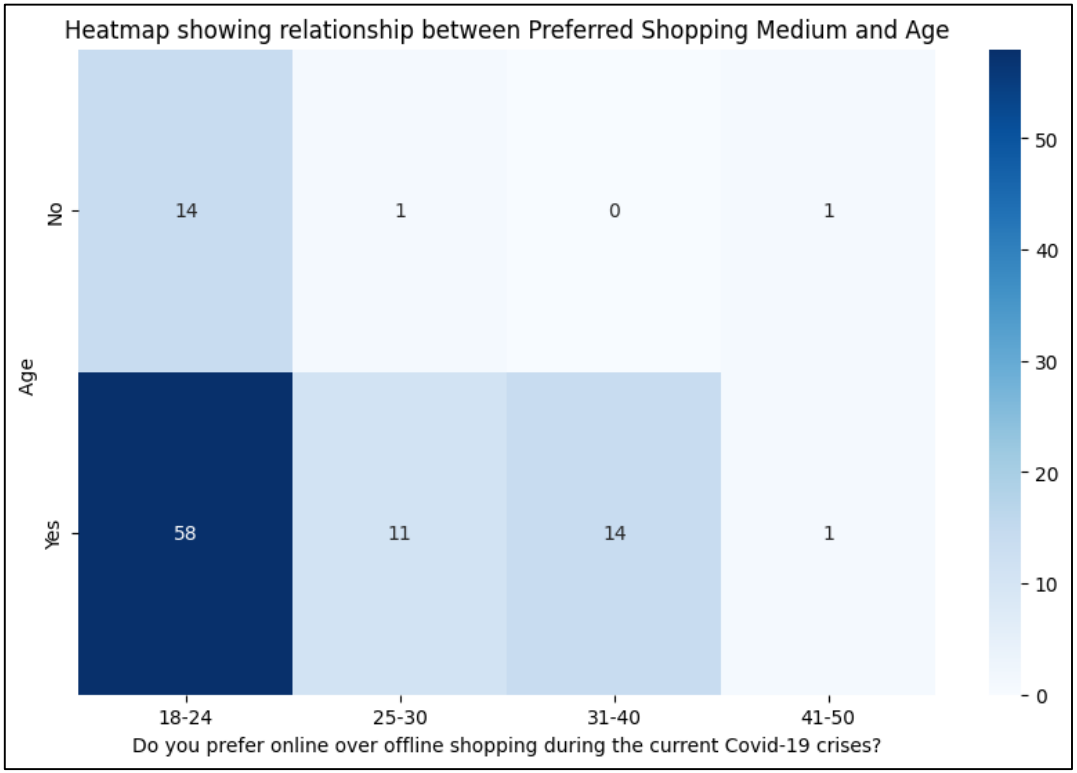
Bivariate analysis between Age and Online over Offline Shopping Preference

This analysis was done to check if different age groups have certain preferences for shopping online than shopping offline during the post – COVID pandemic.



Statistical Inference

Bivariate analysis between Age and Online over Offline Shopping Preference




```
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]]
```

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.

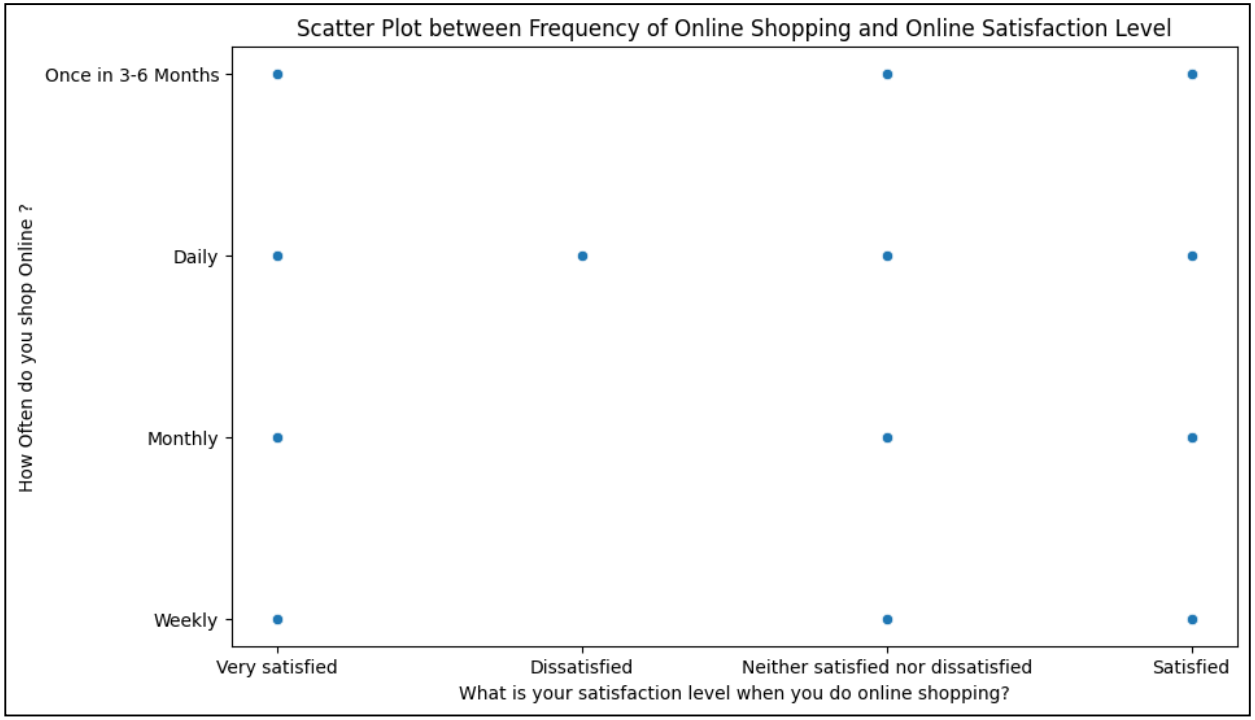
The p – value of 0.136 is higher than the significance level of 0.05 (5%) and this suggests that there is no significant difference in the age of customers and their preferred shopping medium hence the null hypothesis (H_o) is accepted, which indicates that age was not a determinant factor of the shopping medium of individuals during the pandemic.

BIRMINGHAM CITY University

Bivariate Analysis

Bivariate analysis between Online Shopping Satisfaction Level and Online Shopping Frequency

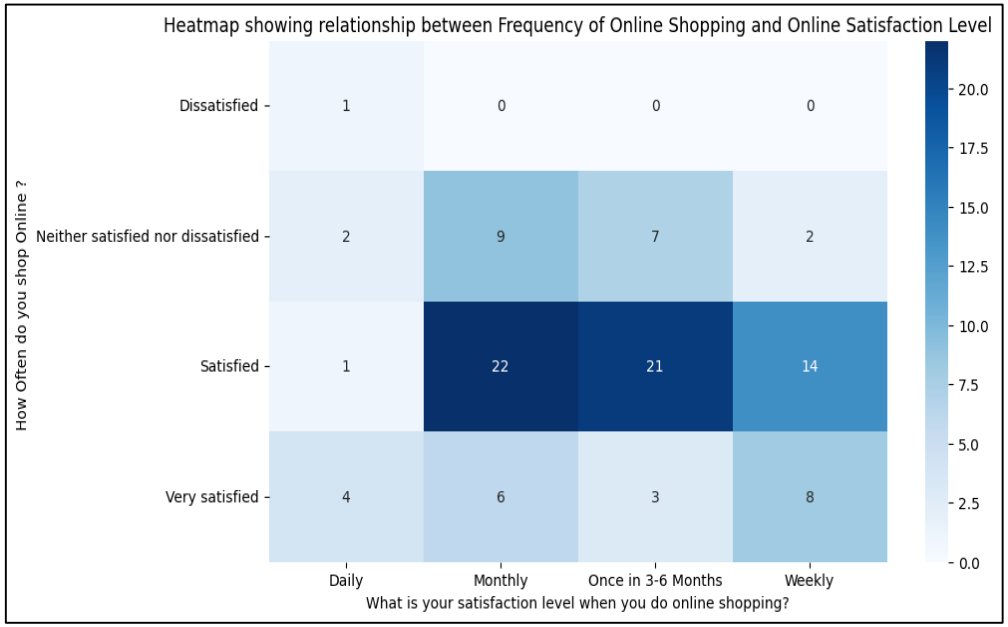
This analysis was done to understand the relationship between people's online shopping frequency and their level of satisfaction after shopping online.



BIRMINGHAM CITY University

Statistical Inference

Bivariate analysis between Online Shopping Satisfaction Level and Online Shopping Frequency




```
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]]
```

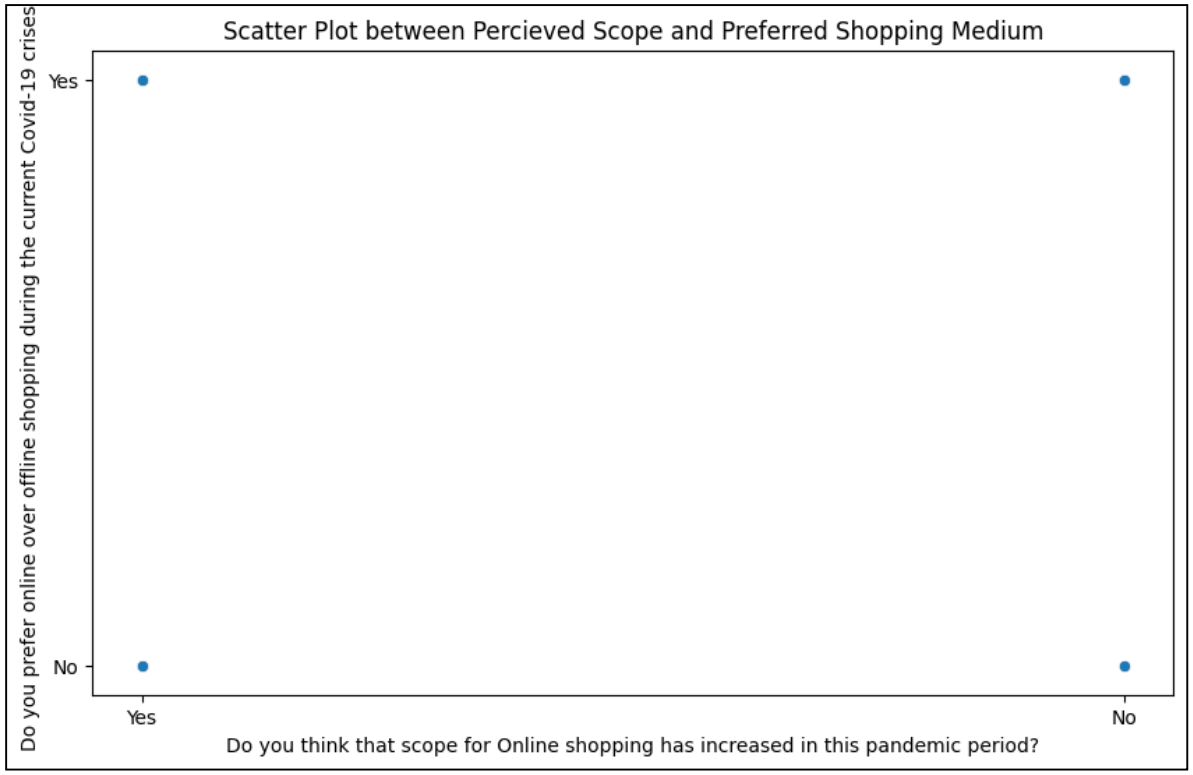
The chi – squared value of 24.3 with a degree of freedom 9, indicates that there is a greater difference between the observed data and the data expected under the null hypothesis.

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.

Bivariate Analysis

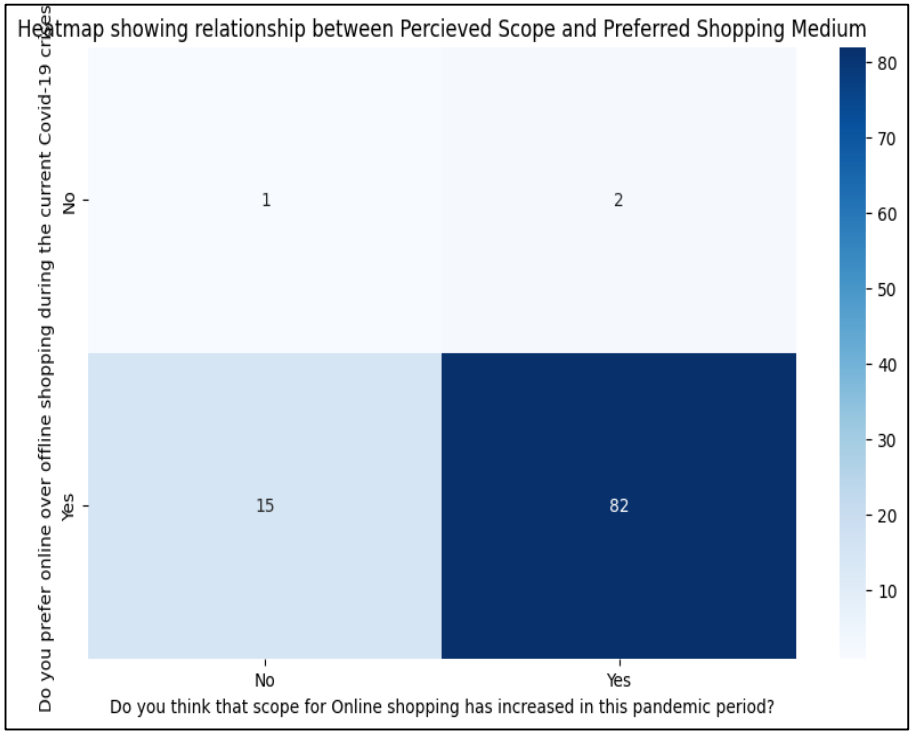
Bivariate Analysis between Online Shopping Scope and Preferred Shopping Medium

This analysis was done to determine if the increase in online scope during the pandemic is related with the preferred shopping medium.



Statistical Inference

Bivariate Analysis between Online Shopping 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]]
```

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.

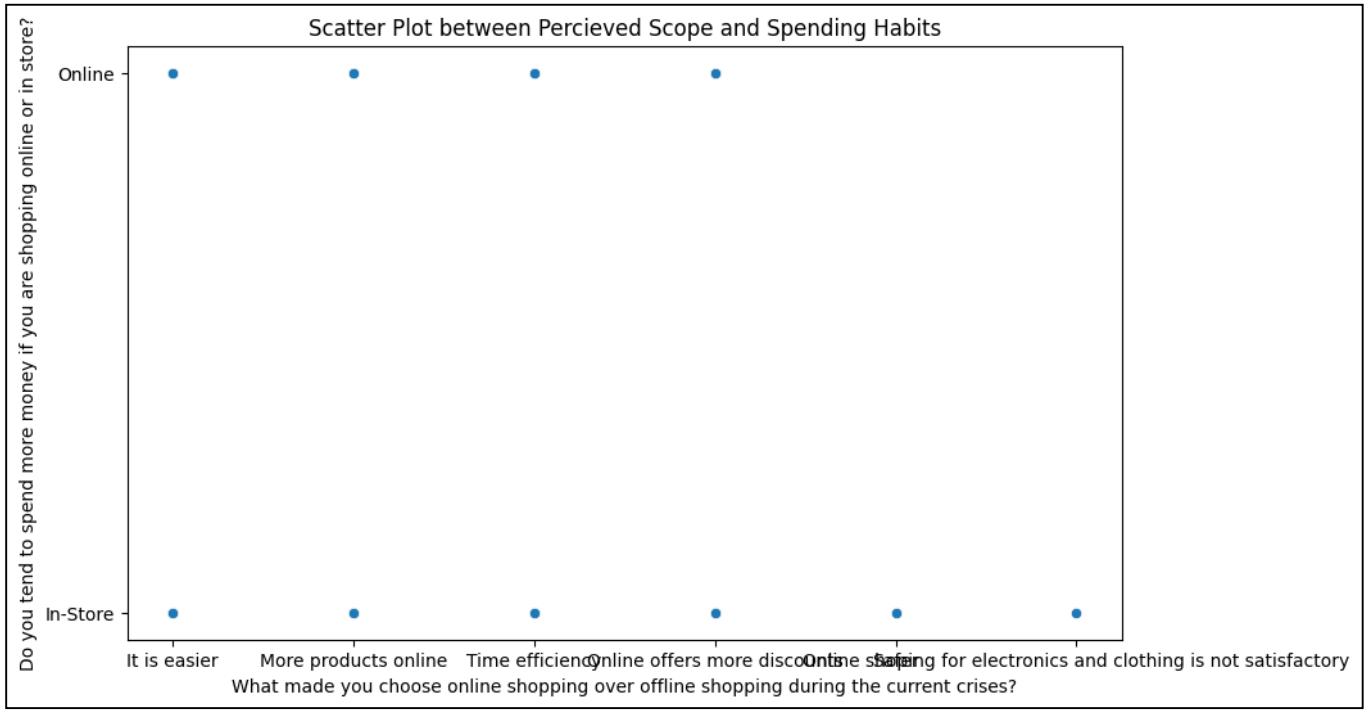
The p-value of 0.9745 is higher than the level of significance (0.05) and this indicates that the null hypothesis (H_o) is accepted which means that there is no significant difference between the perceived scope of shopping and the preferred shopping medium of people during the COVID-19 pandemic.

BIRMINGHAM CITY University

Bivariate Analysis

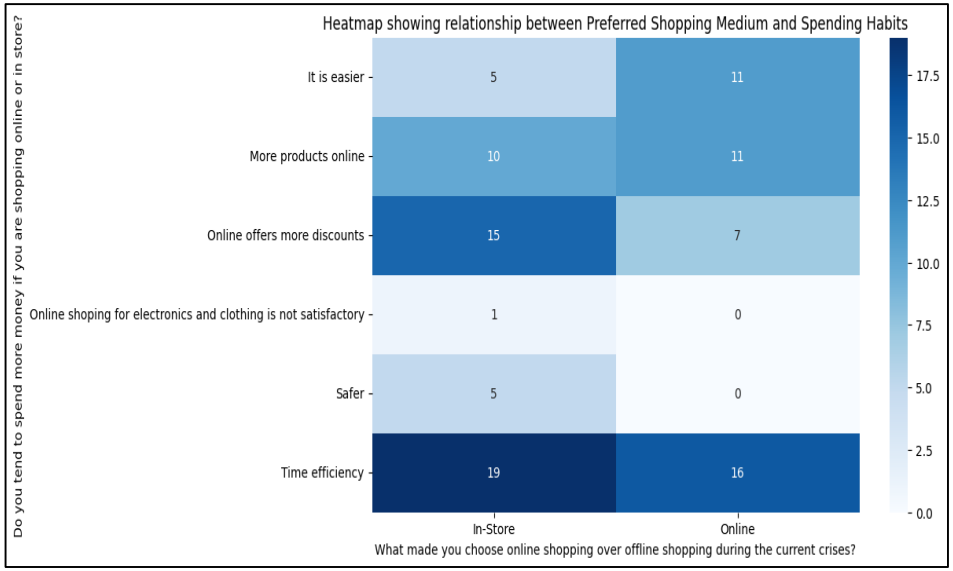
Bivariate Analysis between Preferred Shopping Medium and Spending Habits

This analysis was done to determine if people spend more when they shop online or offline.



Statistical Inference

Bivariate Analysis between Preferred Shopping Medium and Spending Habits



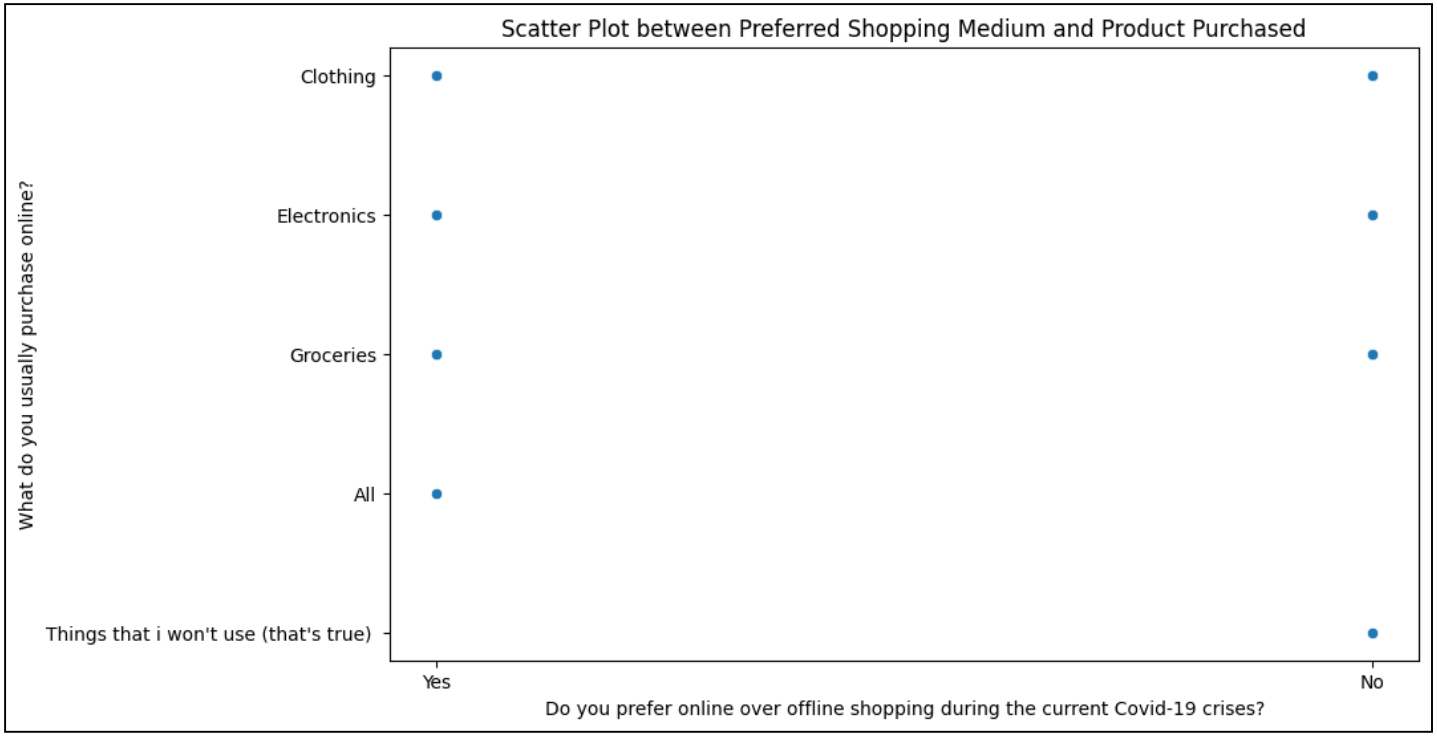

```
Statistical Inferences:
Chi-squared statistic: 10.569548296821026
p-value: 0.060615033004222996
Degrees of freedom: 5
Expected Values: [[ 8.8   7.2 ]
 [11.55  9.45]
 [12.1   9.9 ]
 [ 0.55  0.45]
 [ 2.75  2.25]
 [19.25 15.75]]
```

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.

Bivariate Analysis

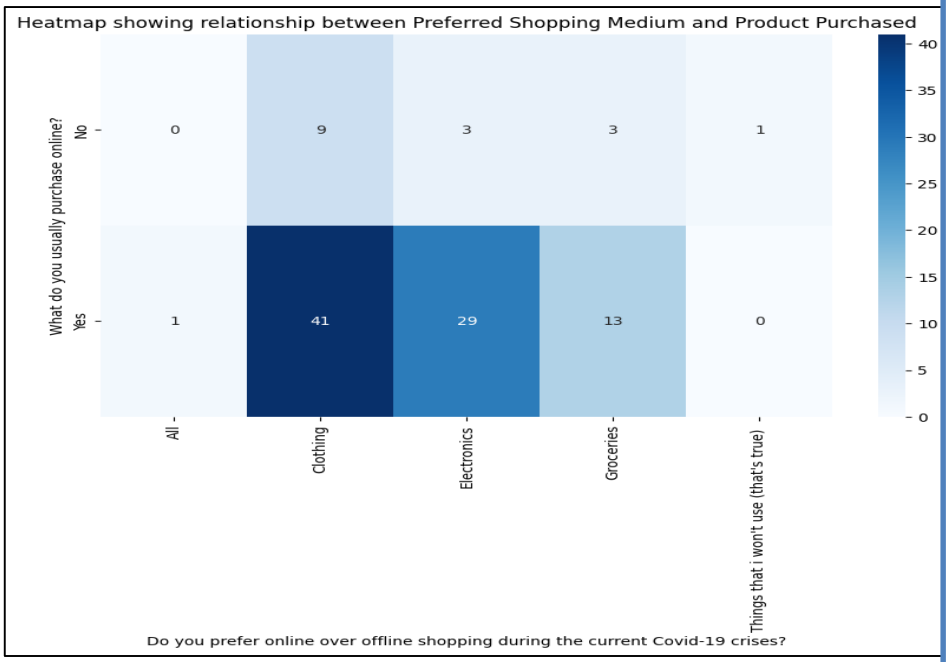
Bivariate Analysis between Preferred Shopping Medium and Products Purchased

This analysis was done to determine if the preferred shopping medium affected the types of products being purchased.



Statistical Inference

Bivariate Analysis between Preferred Shopping Medium and Products 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]]
```

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, hence the alternative hypothesis (H_a) would be accepted meaning that there is a significant difference between the perceived scope and spending habits of individuals during the COVID – 19 pandemic.

LIMITATION AND RECOMMENDATION

LIMITATIONS

- Use of secondary data as there were some shortcomings as to getting primary data for the purpose of this work such as time frame for gathering primary data and also for a particular location.
- The absence of funding for this research.

RECOMMENDATIONS

- This research work can be used for further research work incorporating advanced analytics to understand consumer preferences and tailor online content to bridge the gap between online convenience and offline satisfaction.
- Further research should focus to host a broad range of consumer behavior and choices by expanding the survey questions to hold more refined consumer attitudes and extending the research beyond the pandemic scope to give deeper perspective into the long-term inferences of these changes in consumer behavior.

CONCLUSION

The 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 offline to online shopping in particular. This change, affected greatly by the pandemic, shows a comprehensive pattern in consumer behavior and modern adaptation.

BIRMINGHAM CITY University

THANKS FOR LISTENING

BIRMINGHAM CITY University