

A PREDICTIVE ANALYSIS ON STUDENT BEHAVIOUR IN UTAR BY USING E-COMMERCE PLATFORM

BY

CHIN WAI TENG

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

**BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION
SYSTEMS**

Faculty of Information and Communication Technology

(Kampar Campus)

JUNE 2024

REPORT STATUS DECLARATION FORM

Title: A PREDICTIVE ANALYSIS ON STUDENT BEHAVIOUR IN UTAR BY USING E-COMMERCE PLATFORM

Academic Session: 2024

I

CHIN WAI TENG
(CAPITAL LETTER)

declare that I allow this Final Year Project Report to be kept in
Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.



(Author's signature)

Verified by,



(Supervisor's signature)

Address:

No 23, Selasar Tebing Kinta 1,

Taman Mirindy,

31400 Ipoh Perak

Ts Ahmad Subhi Bin Zolkafly

Supervisor's name

Date: 02/09/2024

Date: 12/9/2024

Universiti Tunku Abdul Rahman			
Form Title : Sample of Submission Sheet for FYP/Dissertation/Thesis			
Form Number: FM-IAD-004	Rev No.: 0	Effective Date: 21 JUNE 2011	Page No.: 1 of 1

FACULTY/INSTITUTE* OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TUNKU ABDUL RAHMAN

Date: 02/09/2024

SUBMISSION OF FINAL YEAR PROJECT /DISSERTATION/THESIS

It is hereby certified that Chin Wai Teng (ID No: 22ACB00233) has completed this final year project/ dissertation/ thesis* entitled “A PREDICTIVE ANALYSIS ON STUDENT BEHAVIOUR IN UTAR BY USING E-COMMERCE PLATFORM ” under the supervision of Ts Ahmad Subhi Bin Zolkafly (Supervisor) from the Department of _____, Faculty/Institute* of Information and Communication Technology , and _____ (Co-Supervisor)* from the Department of _____, Faculty/Institute* of _____.

I understand that University will upload softcopy of my final year project / dissertation/ thesis* in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,




(CHIN WAI TENG)

*Delete whichever not applicable

DECLARATION OF ORIGINALITY

I declare that this report entitled “**A PREDICTIVE ANALYSIS ON STUDENT BEHAVIOUR IN UTAR BY USING E-COMMERCE PLATFORM**” is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature : 

Name : CHIN WAI TENG

Date : 2/09/2024

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and appreciation to my supervisors, Ts Ahmad Subhi Bin Zolkafly who has given me this bright opportunity to engage in a research project. It is my first step to establish a career in the research field. A million thanks to you.

Next, I must say thanks to my parents and my family for their love, support, and continuous encouragement throughout the course.

ABSTRACT

People frequently use e-commerce platforms as their main daily shopping method during the global Covid-19 pandemic. E-commerce platforms are indispensable application software for university students. There will be a notable increase in their intention to purchase, particularly during promotional periods like Double 11. Decision-makers or the marketing department of a company may recommend or advertise users considering consumer characteristics like gender, age, needs, and product preferences. Their desire to purchase will grow more easily as a result, and the dashboard results can clearly highlight each point that is significant or might affect the sale. Thus, by utilizing suitable predictive models and dashboard tools like Power BI or Jupyter, the underlying factors can be examined. As a result, this study examines the variables affecting UTAR students' online buying behavior. Undergraduate students at UTAR's Kampar Campus are the research's intended respondents. Since all the data for this study is collected at once, it is cross-sectional in nature and uses a quantitative research methodology. The data is analyzed using the Data Science Life Cycle. Predictive models and dashboard technologies such as Power BI and Jupyter were used to investigate the underlying elements driving student behavior. The findings showed that perceived ease, availability, and time-saving benefits have a substantial impact on students' online buying selections. Furthermore, security issues, product quality worries, and the inability to verify things before purchasing were significant deterrents. Lastly based on the findings, it is advised that e-commerce platforms aimed at university students' transparency, improve transaction security, and provide flexible return policies to alleviate product-related concerns. Moreover, personalized marketing methods emphasizing convenience and availability benefits may boost student involvement with online buying platforms.

TABLE OF CONTENTS

TITLE PAGE	i
REPORT STATUS DECLARATION FORM	ii
FYP THESIS SUBMISSION FORM	iii
DECLARATION OF ORIGINALITY	iv
ACKNOWLEDGEMENTS	v
ABSTRACT	vi
TABLE OF CONTENTS	vii
LIST OF FIGURES	xi
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1	1
Introduction	1
1.1 Problem Statement and Motivation	4
1.2 Research Objectives	5
1.3 Project Scope and Direction	5
1.4 Contributions	5
1.5 Report Organization	5
CHAPTER 2	6
Literature Reviews	6
2.1 Previous Works on University Students' Online Shopping Behavior By Using E-Commerce Platform	6
2.1.1 Students' Online Shopping Behavior	6
2.1.2 Generation Z	7
2.1.3 E-commerce	7
2.1.4 Trust Environment for E-business	7
2.1.5 University students satisfaction of online shopping	8
2.2 Limitation of Previous Studies	8
2.2.1 Online shopping problems	8
2.2.2 Factors Influencing Consumer Buying Behavior Online	9
2.2.3 Scams on Consumer's Attitude towards Online Shopping	9
2.2.4 Online Shopping Hesitation	10
2.2.5 Problems faced by Online Customers on Ecommerce	10

2.3 Proposed Solutions	10
2.3.1 Advantages of Online Shopping	10
2.3.2 Trust of online purchase intention	11
2.3.3 Consumer Perception on Security	11
2.3.4 Customer Support	12
2.3.5 Product Information	12
2.4 Benefit of Online Shopping Behavior	12
2.4.1 Perceived Benefits	12
2.4.2 Convenience	13
2.4.3 Product Variety	13
2.5 Risk of Online Shopping Behaviour	13
2.5.1 Perceived Risk	13
2.5.2 Product Risk	14
2.5.3 Privacy Risk	14
2.6 Predictive Model	14
2.6.1 Linear Regression versus Random Forest	14
2.7 Summary of this Chapter	15
CHAPTER 3	16
SYSTEM MODEL (FOR RESEARCH-BASED PROJECT)	16
3.1 Proposed Method/Approach	16
3.1.1 Understanding the Business Problem	17
3.1.2 Preparing the Data	17
3.1.3 Exploratory Data Analysis (EDA)	17
3.1.4 Modeling the Data	18
3.1.5 Evaluating the Model	18
3.1.6 Deploying the Model	18
3.2 System Requirement	19
3.2.1 Hardware	19
3.2.2 Software	19
3.3 Activity Diagram	20
3.4 System Block Diagram	21
3.5 Summary of this Chapter	21
CHAPTER 4	22
EXPERIMENT/SIMULATION (FOR RESEARCH- BASED PROJECT)	22
4.1 Setting up	22
4.1.1 Software	22

4.2 Visualization and segmentation	22
4.3 Timeline	23
4.4 Data collection and preparation	23
4.5 Checking the Null Value, Duplicate Value	24
4.6 Visualize the results as dashboard by using Power BI	29
4.7 Visualize the results as dashboard by using Jupyter (Python)	41
4.8 Summary of this Chapter	49
CHAPTER 5	50
5.1 Evaluation and Discussion	50
5.1.1 Analysis of Figure 4.7.6 (Respondents' Attitude)	50
5.1.2 Analysis of Figure 4.7.7 (Respondents' Behavior)	51
5.1.3 Analysis of Figure 4.7.8 (Security Risk)	51
5.1.4 Analysis of Figure 4.7.9 (Product Risk)	52
5.1.5 Analysis of Figure 4.7.10 (Convenience Risk)	53
5.1.6 Analysis of Figure 4.7.11 (Non-Delivery Risk)	54
5.1.7 Analysis of Figure 4.7.12 (Convenience Benefit)	54
5.1.8 Analysis of Figure 4.7.13 (Availability Benefit)	55
5.1.9 Analysis of Figure 4.7.14 (Time Saving Benefit)	56
5.1.10 Analysis of Figure 4.7.15 (Predictive Model)	57
5.2 Discussion of Results	58
5.3 Summary of this chapter	60
CHAPTER 6	61
Conclusion and Future Recommendations	61
6.1 Conclusion	61
6.2 Future Recommendations	62
REFERENCES	64
APPENDIX	
Data	A-1
WEEKLY LOG	A-21

POSTER

A-25

PLAGIARISM CHECK RESULT

FYP2 CHECKLIST

LIST OF FIGURES

Figure Number	Title	Page
Figure 1. 1	Statista number of internet and social media users worldwide January 2024	1
Figure 1. 2	How Many Online Shoppers In 2023	3
Figure 1. 3	Prediction Number of Online Shoppers in Next Few Years	3
Figure 3. 1	Data Science Life Cycle	16
Figure 3.3. 1	Activity Diagram of Research	20
Figure 3.4. 1	System Block Diagram of Data Analysis	21
Figure 4.3. 1	Timeline	23
Figure 4.5. 1	Locate the database saved	25
Figure 4.5. 2	Select 'Python 3 (ipykernel)'	25
Figure 4.5. 3	Import the data and read it from located file	26
Figure 4.5. 4	Checking null values	27
Figure 4.5. 5	Checking duplicate values	28
Figure 4.6. 1	Get the data (Step 1 of Data Visualization)	29
Figure 4.6. 2	Select Excel Workbook (Step 2 of Data Visualization)	30
Figure 4.6. 3	Load the data (Step 3 of Data Visualization)	30
Figure 4.6. 4	Visualize the data (Step 4 of Data Visualization)	31
Figure 4.6. 5	Attitude	32
Figure 4.6. 6	Behavior	33
Figure 4.6. 7	Security Risk	34
Figure 4.6. 8	Product Risk	35
Figure 4.6. 9	Convenience Risk	36
Figure 4.6. 10	Non-Delivery Risk	37
Figure 4.6. 11	Convenience Benefit	38
Figure 4.6. 12	Availability Benefit	39
Figure 4.6. 13	Time Saving Benefit	40

Figure 4.7. 1 Launch Jupyter Notebook in Anaconda Prompt	41
Figure 4.7. 2 Locate the database saved	42
Figure 4.7. 3 Click on New and select 'Python 3 (ipykernel)'	42
Figure 4.7. 4 Read the data from the located file	43
Figure 4.7. 5 A Clustered Column Chart output of Online Shopping is Enjoyable (5)	44
Figure 4.7. 6 Dashboard overview Attitude using Python (Jupyter Notebook)	45
Figure 4.7. 7 Dashboard overview Behavior using Python (Jupyter Notebook)	45
Figure 4.7. 8 Dashboard overview Security Risk using Python (Jupyter Notebook)	46
Figure 4.7. 9 Dashboard overview Product Risk using Python (Jupyter Notebook)	46
Figure 4.7. 10 Dashboard overview Convenience Risk using Python (Jupyter Notebook)	47
Figure 4.7. 11 Dashboard overview Non-Delivery Risk using Python (Jupyter Notebook)	47
Figure 4.7. 12 Dashboard overview Convenience Benefit using Python (Jupyter Notebook)	48
Figure 4.7. 13 Dashboard overview Availability Benefit using Python (Jupyter Notebook)	48
Figure 4.7. 14 Dashboard overview Time Saving Benefit using Python (Jupyter Notebook)	49
Figure 4.7. 15 Dashboard overview of Predictive Model (Power BI)	49

LIST OF TABLES

Table Number	Title	Page
Table 3.2. 1	Specifications of laptop	19
Table 3.2. 2	System Requirements of Power BI	19

LIST OF ABBREVIATIONS

<i>FICT</i>	Faculty of Information and Communication Technology
<i>FBF</i>	Faculty of Business and Finance
<i>FAS</i>	Faculty of Arts and Social Science
<i>FSc</i>	Faculty of Science
<i>DSLCL</i>	Data Science Life Cycle

CHAPTER 1

Introduction

The global Internet population had been expanding quickly in this era of globalization. As of January 2024, there were 5.35 billion internet users worldwide, which was 66.2% of the world's population, according to The Statistics Portal. Of this total, 5.04 billion people (62.3%) used social media [1]. The growth of the Internet had completely changed both people's lives and traditional business practices.

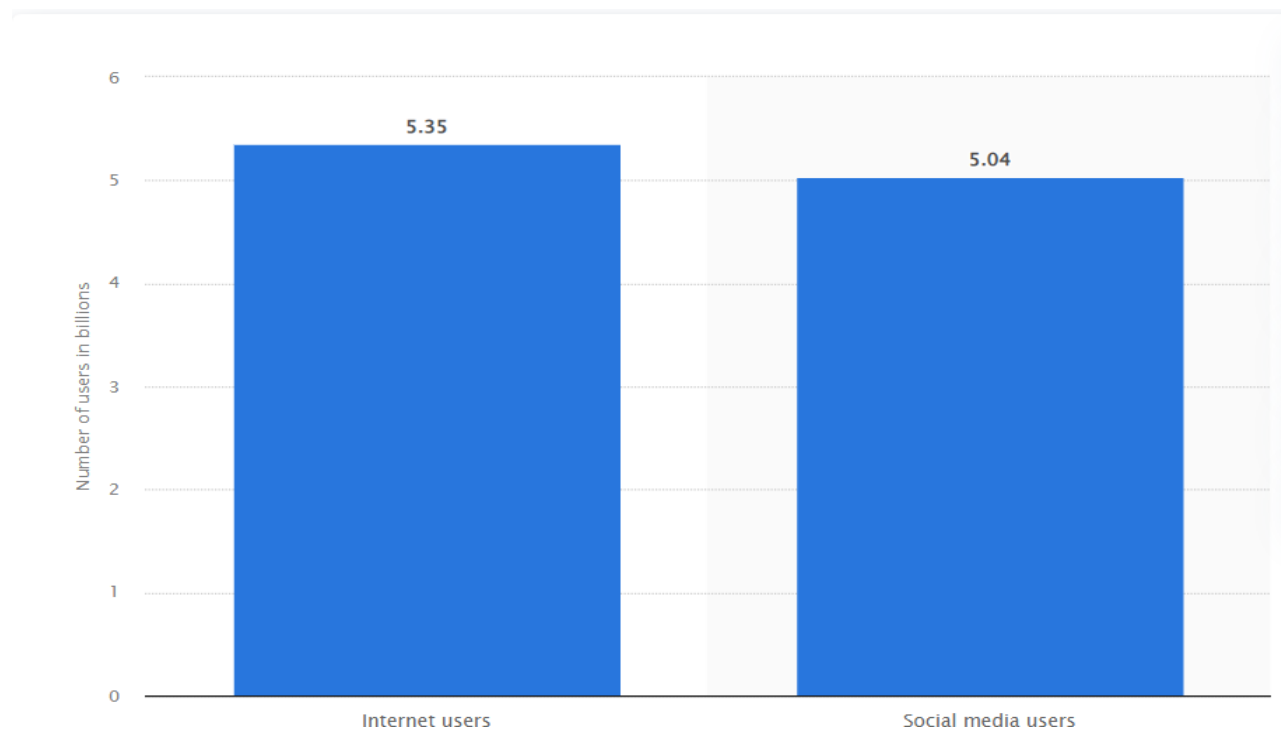


Figure 1. 1 Statista number of internet and social media users worldwide January 2024

CHAPTER 1

Online shopping was an action of making purchases of goods or services via the Internet and was commonly referred to as “E-commerce”. It offered a means for consumers to purchase goods and services from marketers online [2]. E-commerce also known as “E-business”, encompassed a very large online business space. Dealing online involved a number of parties, including suppliers, distributors, and customers. E-commerce was essential to extending trade relations globally via the Internet because it encompassed a variety of financial activities and trade-related business aspects in addition to transactions activities. By using the Internet, computers, information technology and communication to establish a relationship between customers and retailers, the entire e-commerce process could be completed manually.

Online shopping was becoming more and more popular worldwide thanks to E-commerce. By 2023, there were 2.64 billion online shoppers worldwide, predicts Yaguara.co. It represented over 33 percent of the global population, with an increase of 80 million online shoppers since 2022. That represented a 3.1% rise in online sales [3]. According to Datareportal, at the beginning of 2023, there were 33.03 million internet users in Malaysia, or 96.8% of the country’s total population, and 26.80 million social media users, or 78.5 percent of the country’s total population and there were 44.05 million active mobile phone connections, which translated to 129.1% of the country’s total population [4]. Aside from that, Malaysians used e-commerce at high rates. 15.63 million Malaysians were making online purchases of consumer goods at the beginning of 2023, up 8.3% from the previous year, according to Malaysia eCommerce Statistics [5]. It was discovered that the majority of Malaysians who shopped online were driven by the benefits of prices, the availability of reviews, and the selection of products offered by online retailers. In addition, the convenience, free shipping, and other special offers made by online retailers drew them in and encouraged them to make larger purchases there.

Depending on purchasing goods and services online had grown in popularity. It was not surprising that online shopping had become more popular given the advantages of doing it from home, comparing costs with other websites, and reading product evaluations. Consequently, according to the latest data, there were 2.64 billion online shoppers worldwide. In 2023, the number of online shoppers increased by 3.1% compared to the previous year [6].



Figure 1. 2 How Many Online Shoppers In 2023

According to e-commerce industry experts, the total number of online shoppers was expected to continue increasing worldwide from 2023 to 2025, with an anticipated growth of 130 million [6].

Year	The Number of Online Shoppers
2025*	2.77 billion
2024*	2.71 billion
2023	2.64 billion

Figure 1. 3 Prediction Number of Online Shoppers in Next Few Years

1.1 Problem Statement and Motivation

Online shopping had become popular worldwide. The advancement of Internet technology had made it possible for everyone to make online purchases of goods and services. Most often, customers used the Internet to research products and purchase the items they wanted. However, online shopping did bring certain problems, such as high abandonment rates and online scams. The factors or reasons that cause university students to spend time online shopping could have been influenced by societal values, cultural trends, politics, economy, family members and friends. But ultimately, it was up to them how much time they spend on it.

While technology and the Internet had greatly benefited people, they have also provided fraudsters with access to a multitude of venues and opportunities for both conventional and innovative forms of fraud. Internet fraud was the largest and most prevalent issue that many countries encountered. In Malaysia, there was an instance of internet buying fraud. As reported by New Straits Time, an international student enrolled at a Malaysian university emailed the National Consumer Complaints Centre (NCCC) with concerns about instances of online fraud. He stated that he used an online website to purchase a mobile phone from a merchant. He accepted the arrangement in exchange for a third of the smartphone's initial cost. Nevertheless, the phone was a fake, so they were unable to assist him. He was taken aback upon realizing that he had been duped by internet con artists [7].

Next, the other reason for conducting this study project was to find out if behaviour affected UTAR students' online shopping habits. In this case, studying internet shopping habits revealed the underlying problem. The importance of this research stemmed from the fact that during pandemics, a growing number of people, especially university students, began shopping online. There wasn't a single study that is specifically focused on students at private universities, despite earlier research being analyzed in relation to the students' online buying habits. To accurately establish whether UTAR students' online buying behaviour was influenced by perceived benefits and hazards, a thorough analysis of their online shopping behaviour was necessary as part of this project and with a more precise technique. This research aimed to determine how perceived risks and benefits influenced their online shopping habits.

1.2 Research Objectives

1. To investigate the online shopping behavior from students' attitude in UTAR, Perak Campus.
2. To develop a predictive e-commerce model by using Power BI.
3. To visualize the predictive e-commerce model using dashboard.

1.3 Project Scope and Direction

This study's target audience was UTAR students, specifically focusing on their online shopping habits through the use of e-commerce platforms. Because different faculty students and genders had varied demands when it came to online purchasing, students from a variety of faculties participated, which could have led to a more detailed and thorough outcome. Furthermore, Power BI was utilized as a dashboard tool to generate and present a clearer visualization of the results.

1.4 Contributions

The project's research utilized e-commerce websites such as Lazada, Shopee, Amazon and Taobao to help decision-makers better understand and gain insights into online shopping habits of UTAR students by identifying the factors that excited students and attracted them. Therefore, the results of this project helped decision-makers understand which commodities college students were interested in. As they learned more from the research, they were able to modify the strategies they used to market their products to the appropriate target market. When this study was completed, it demonstrated how many of the products captured students' attention, as well as the factors investigated and recommendations for problem-solving strategies for decision-makers.

1.5 Report Organization

The subsequent chapters presented the specifics of this study. Chapter 2 reviewed some prior studies in this area. Next, in Chapter 3, a suggested strategy or method like Data Science Life Cycle and Power BI was discussed. Next, Chapter 4 discussed the software's first installation work and some data dashboard. Moreover, in Chapter 5, the evaluation was discussed and a complete discussion was provided. Lastly, the Conclusion and Recommendations were the final chapter in this research.

CHAPTER 2

Literature Reviews

The previous chapter covered the Introduction, which included the following topics: Problem Statements and Motivation, Research Objectives, Project Scopes and Direction, Contributions and Report Organization. This chapter, Literature Reviews, discussed the Previous Works on University Students' Online Shopping Behavior by Using E-commerce Platform, Limitation of Previous Studies, Proposed Solutions, Benefit of Online Shopping Behavior, Risk of Online Shopping Behavior, and Predictive Model.

2.1 Previous Works on University Students' Online Shopping Behavior By Using E-Commerce Platform

2.1.1 Students' Online Shopping Behavior

Around the world, online shopping became more and more popular, especially in nations with highly developed internet infrastructure that could support online marketing. In recent years, the internet served as a global platform for consumer transactions in addition to being a networking tool. The use of the internet has increased significantly in recent years, and it became a widely used medium for trade, services, and the exchange of information. More than 627 million people worldwide were said to have shopped online in 2006. The e-commerce market was expected to grow from \$228 billion in 2007 to \$288 billion in 2009, according to Forrester. Ninety percent of university students used the internet every day, and they had \$200 billion in purchasing power in the US market annually. On average, a student's monthly available discretionary spending was \$287. Hence, given the influence that students had in the marketplace, it was critical that merchants and instructors of consumer behavior gained a deeper understanding of this demographic's perspective on internet purchasing. According to Sabrina et al., university students in Malaysia became regular consumers of various market segments due to the growth of educational services [8].

2.1.2 Generation Z

Generation Z or Millennial were defined as the generation that consisted of people who were born between 2000 and the present [9], in a world with the Internet, smartphones, laptops, freely available networks, and digital media. Because they were raised in the social media era, members of Generation Z were digital natives who identified with technology. They were also known as Gen Wii, Gen Tech, Digital Natives, and Generation I. Their existence was more intertwined with electronics and the digital world than previous generations, having been raised around the digital world from birth [10].

2.1.3 E-commerce

The term "e-commerce" described the trading of goods, services, and information between companies and between businesses, as well as the sale, purchase, transfer, or exchange of those commodities and services. It was not always possible to receive goods and services over the Internet. E-commerce development had two primary generations. American companies dominated the market during the initial wave of e-commerce, which spanned from 1995 to 2000. Global in reach, the second wave of e-commerce surfaced around 2000, with international companies transacting business in several countries and languages. Moreover, the development of technologies that reduced costs significantly while offering enhancements in functionality, usability, software availability, ease of site development, and enhanced security and accessibility was the main focus of current trends in e-commerce technology [11].

2.1.4 Trust Environment for E-business

Generally speaking, the degree of trust that placed in a specific person or organization within a network environment directly correlated with the security level of that environment. To ensure a secure environment, trust was essential. To create a trustworthy environment for e-business, security technology needed to be used in addition to physical security measures. Organizations could target and expand into new markets, develop business strategies, enhance customer satisfaction, and improve their business processes thanks to the Internet. On the other hand, not offering a secure Internet connection could result in lost revenue, missed business opportunities, a

damaged reputation for quality work, and diminished customer trust. Having a secure electronic environment was essential to using the Internet and realizing its full potential [12].

2.1.5 University students' satisfaction of online shopping

The general public was increasingly using the Internet. Because of this, customers could easily shop online whenever they want, day or night, and they do not have to waste time travelling, deciding what to buy, and making in person purchases. It was estimated that in recent years, 85% of people who used the Internet to order goods had done so. This study explored students' intention to make an online purchase as well as their satisfaction with that experience. Many consumers were interested in the ability to make purchases without having to leave their current location. Customers could also purchase the products they wanted at the best price using online tools for price comparison and searching. This gave them an extra advantage when making purchases [13].

2.2 Limitation of Previous Studies

2.2.1 Online shopping problems

Online shopping was more convenient than driving to a store since it provided an almost infinite number of products and the ability to compare prices. Despite the enormous advancements in online payment and security technology, users still occasionally encountered problems when shopping online. Typical issues included making the incorrect order or falling for online frauds. The fact that big retailers' online and retail divisions were run by entirely different companies seemed to be the root of many of the problems. Until lately, that was the case, and there were times when these organizations felt more like competitors than fellow project group members. When a customer recognized the brand name of a store, they naturally believed the two businesses were the same. However, this was not the case, and the outcome was immense confusion and annoyance [14].

2.2.2 Factors Influencing Consumer Buying Behavior Online

Due to the distinctive qualities and interactions between technology and culture, consumer behavior online differed slightly from traditional consumer behavior. Despite the commercial use of the Internet growing exponentially, little research had been done on how the multifaceted aspects of traditional consumer behavior changed in the context of e-commerce [15].

2.2.3 Scams on Consumer's Attitude towards Online Shopping

Although online platforms and applications claimed the highest level of security, using the Internet for online transactions that required the sharing of financial personal details raised concerns about security. The vulnerability to digital scams was still a major issue that needed to be resolved. The platforms created between the user and bank interface allowed fraudsters and hackers to take advantage of the customer base and lure them into becoming victims. The psychological impact of experiencing fraud during online transactions affected a customer's attitude towards digital marketing themes and online shopping. Security protocols were offered to bolster consumers' favorable perceptions of online shopping [16].

A review of prior research showed that consumers' perceptions of these activities had an impact on their attitudes towards their offline and online buying behaviors. If a customer viewed shopping as a useful activity, there was a greater chance that they would make actual online purchases [15]. Online customers decided to use online channels for shopping rather than traditional ones because they were lazy and thought it would take time to adopt them. They utilized e-commerce sites to find advantages they considered, such time and convenience savings. Furthermore, people's perceptions of the security dangers associated with online buying were impacted by the rise in online fraud cases caused by security concerns.

2.2.4 Online Shopping Hesitation

Online and traditional offline purchasing channels such as phone ordering, mail order, and physical storefronts were very different from each other. The perceived return on an individual's time, money, and effort inputs was more valuable to online buyers than it is to offline customers. Additionally, shopping in non-store environments and online carried additional hazards. There were several reasons for this elevated sense of perceived risk, including a lack of human contact, an inability to inspect the goods, a lack of security and confidence, the difficulty of returning the item, and others physical or visible issues. Because internet purchasing carried a bigger risk than conventional shopping, it was expected that consumers hesitated more when making purchases online [17].

2.2.5 Problems faced by Online Customers on Ecommerce

Because of their attitude towards saving time, price flexibility, and the availability of a variety of products on one platform, consumers became more and more drawn to online shopping. As more customers started to trust online retailers and have shifted a sizable portion of their shopping online, online shopping expanded over time. According to Savrul et al., two of the most recent issues preventing businesses from participating in e-commerce were security concerns and regulatory frameworks. One frequently cited reason why customers did not make purchases online was a lack of trust. The fundamental element that had always been crucial in the unpredictable online world of e-commerce was trust [18].

2.3 Proposed Solutions

2.3.1 Advantages of Online Shopping

Internet shoppers thought they can purchase goods not available in their local market and save time when they shop online. Meanwhile, some white-collar workers and students who were interested in fashion consumed it through online shopping. Online shopping offered certain benefits, such as easy payment processing and reduced shopping expenses. Online bank cards, such as those for construction, business, and industry, postal savings cards, and so forth, were the

primary means of convenient payment. Customers could place orders for the chosen goods using an email, feedback form, or any other method, and then waited for the goods to arrive at their homes. E-commerce also significantly raised the transaction costs for customers because they did not have to visit multiple stores to buy goods haggle with salespeople. Thanks to e-commerce, consumers could quickly compare product prices, which significantly increased product price transparency. The cost of the channel was greatly reduced and product prices were lowered with online marketing [19].

2.3.2 Trust of online purchase intention

In electronic commerce, trust was recognized as a critical component of buyer-seller relationships and online purchase intention. Within the realm of electronic commerce, trust encompassed the perceptions and anticipations of online buyers regarding the attributes of the virtual vendor. According to Kraeuter, the biggest long-term obstacle to consumers realizing the potential of e-commerce in online environments was trust. Consumers based significant purchasing decisions on how much they trusted the brand, the salesman, or the business. An alternative definition of trust was customers' willingness to perceive a potential loss while they were shopping; in this sense, trust could be thought of as a type of behavioral intention [12].

2.3.3 Consumer Perception on Security

Based on surveys conducted on online purchases, customers' top worries were security related. According to researchers, security was not just about technology but also about organizational and human aspects. This suggested that even with the best technical solutions and methods, a company's efforts could be for naught if it ignored how people perceived safe websites. The security perception of an individual was the extent to which they believed that a website or online retailer was secure. There were serious issues with transferring sensitive data, such as credit card numbers. They had to utilize security and give it thought in order to safeguard the information of their clients, as internet websites were the main target of hackers. Online transactions were not safe on the Internet [12].

2.3.4 Customer Support

Customer support offered a range of services to help customers solve problems pertaining to the appropriate and cost-effective use of products. Support for setup, training, maintenance, and updates was provided. Customer service was correlated with more positive perceptions regarding the website. Online customer loyalty and the inclination to shop online were positively impacted by customer service. The seven components of customer assistance were installation, user training, online help, online maintenance and repair, warranty or guarantee, and upgrade. The relative value of various customer support components had changed within the last fifteen years. Delivering exceptional customer service was essential to achieving customer happiness. This increased the success rate of new goods and provided firms with a competitive advantage [15].

2.3.5 Product Information

A brief description of the product that customers wished to buy online was called product information. This brief explanation needed to be packed with pertinent details so that readers could make wiser decisions. This was a result of our inability to hold items and inspect their quality when shopping online. Clear information encouraged customers to purchase the product and gave them a sense of satisfaction. Thus, a company could display product information through visually appealing video, audio, text, and images. Product details also showed improvement in attitudes towards internet buying. Better information influenced more satisfied customers and better purchasing decisions [15].

2.4 Benefit of Online Shopping Behavior

2.4.1 Perceived Benefits

The benefits of the Internet altered customer perceptions, leading them to make online purchases. Customers' belief in and happiness with their online purchases, as well as their view that online shopping was safer, easier, more convenient, and offered a wider selection of products than traditional retail, were the perceived benefits [20].

2.4.2 Convenience

The ability to quickly search for items customers were interested in and willing to buy made life easier thanks to the Internet. Access, search, possession, and transaction were the four alternatives that consumers had when purchasing convenience. Also, the main factor encouraging consumers to shop online was convenience. A beneficial link between convenience and online shopping was shown in previous research. Many investigations found that convenience had a big impact on consumers' online buying decisions. Another variable needed to be added because it was unclear how these two related in order to give a more thorough explanation [20].

2.4.3 Product Variety

When it came to advantages and convenience, online buying surpassed traditional retail for customers. The buyer had to physically visit stores to find the best product when buying conventionally, which wasted time. As a result, the Internet gave customers a wide range of options, enabling them to compare products from various brands and varieties without having to leave their homes and go shopping. Previous studies demonstrated a strong positive correlation between product variety and online shopping behaviour. Further research was necessary to provide a clearer explanation of the relationship between product variety and online buying behaviour, as it was currently unclear [20].

2.5 Risk of Online Shopping Behaviour

2.5.1 Perceived Risk

The definition of perceived risk was "the consumer's perception of the uncertainty and adverse consequences of buying a product or service". Research on the perception of risk in marketing and consumer behaviour dating back to the early 1960s, along with the advent of internet shopping in the 1990s and many more research since then, showed that customers' online shopping activity was adversely affected by perceived risk [20].

2.5.2 Product Risk

Product risk was the chance that a product wouldn't work as planned or wouldn't live up to performance expectations. Because consumers were unable to personally inspect and test products before making a purchase, internet shopping had a higher risk when it came to products. On the other hand, consumers could view and test the actual products when making a traditional purchase. This meant that compared to traditional retail, online shopping carried a higher risk. Prior research indicated that the risk of a product negatively affected consumers' online purchasing decisions. This was a result of the fact that buyers who made purchases online were limited to depending on the product's scant information [20].

2.5.3 Privacy Risk

The biggest risk was privacy risk, which also arose while making online purchases. Customers who shopped online had to divulge personal information during transactions. Additionally, as internet shopping grew, so did the risk to consumer privacy, which made people hesitant to make purchases online. This relationship was not entirely clear and required more data to support it and provide more specific details because only a little prior research has indicated that privacy risk considerably reduced online shopping behaviour [20].

2.6 Predictive Model

2.6.1 Linear Regression versus Random Forest

Random Forest was applied to classification, regression, and random survival forests. This research explored the regression scenario and compared recently proposed variable importance metrics from two different types of random forests to more traditional linear regression model tools. The emphasis was on inter-regressor correlation as a key factor influencing the behaviour of variable significance measures. The random forest variable significance technique could benefit from a more comprehensive understanding of linear model behaviour. Random forests, on the other hand, were nonparametric, which meant they could learn nonlinearities and interactions from data without having to represent them explicitly. They were shown to be effective not only in the

CHAPTER 2

$n > p$ setting, but also in data mining in the $p > n$ environment. The grounds for adopting varying importances were different in both circumstances.

Linear regression was a traditional parametric method that required explicit handling of nonlinearities and interactions. The model was considered robust when the number of observations (n) exceeded the number of variables (p). Linear regression failed when there were more variables than observations ($p > n$ or even $p \gg n$), unless downsizing methods such as ridge regression, the lasso, or the elastic net were used in tandem with one another [21].

2.7 Summary of this Chapter

This chapter discussed about the Literature Reviews, and next chapter discussed the System Model/ Proposed Method or Approach.

CHAPTER 3

SYSTEM MODEL (FOR RESEARCH-BASED PROJECT)

The previous chapter covered on Literature Reviews, which included the following: Previous Work on University Students' Online Shopping Behavior by Using E-commerce Platform, Limitation of Previous Studies, Proposed Solutions, Benefit of Online Behavior, Risk of Online Shopping Behavior, and Predictive Model. The System Model, which included the Proposed Method/Approach, System requirement, Activity diagram, System block diagram was discussed in this chapter.

3.1 Proposed Method/Approach

The processes of this project were based on the Data Science Life Cycle which was divided into several phases in the development: (1) understanding the business problem, (2) preparing the data, (3) Exploratory Data Analysis (EDA), (4) modeling the data, (5) evaluating the model, (6) deploying the model. The figure below showed the DSLC methodology describing all the phases.

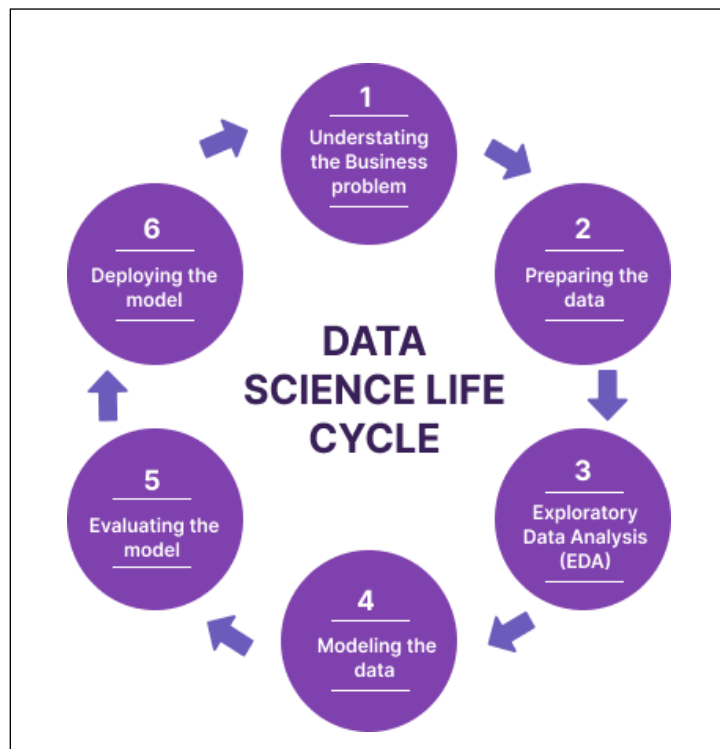


Figure 3. 1 Data Science Life Cycle

3.1.1 Understanding the Business Problem

Understanding the business problem was the first stage of the DSLC. A strong Data Science Life Cycle began with the question “why,” because many of the world’s achievements could be linked to the “why” question. The entire data science project was built upon this stage.

WHAT: To define the business problem or question that needed to be addressed.

HOW: To understand the objectives, requirements, and constraints of problem of this research.

3.1.2 Preparing the Data

The process of preparing the data came next, which was essential to comprehending the business issue and gathering information to address it. In order to ensure the quality and relevance of the data for analysis, it was acquired from multiple sources. The collected data was cleaned, transformed, and preprocessed to make it suitable for analysis.

WHAT: Gathered relevant data from various sources.

HOW: Identified data sources, collected the data from UTAR students, and ensured its quality and relevance to research for analysis.

3.1.3 Exploratory Data Analysis (EDA)

The third stage, also known as Exploratory Data Analysis (EDA), came before modelling, and involved understanding the problem and the variables that could have had an impact on it. To better comprehend the dataset, the collected data was cleaned, transformed, and preprocessed. At this point, the data was examined and analyzed to look for patterns, correlations, and outliers.

WHAT: Cleaned, preprocessed, transformed the collected data, and explored and analyzed it to extract insights.

HOW: Handle missing values, outliers, and normalized data using Python (Jupyter) to make the data suitable for analysis. Visualized data and identified patterns, correlations, and outliers to better understand it.

3.1.4 Modeling the Data

This phase involved building, validating, and optimizing the predictive model using appropriate algorithms to address the specified problem. Model types that were appropriate for the given problem included clustering, regression, and classification.

WHAT: Developed and validated predictive models.

HOW: Selected appropriate algorithms, split data into current type and predictive type, in order to evaluate model performance.

3.1.5 Evaluating the Model

A model was first be constructed, and in order to improve it, it was necessary to determine its current status.

WHAT: Assessed model performance and results.

HOW: Use metrics like accuracy, precision, recall, or AUC to evaluate models (Power BI) and determined if they met the research objectives.

3.1.6 Deploying the Model

Model deployment was the last phase of the DSLC. For practical use, the produced model was integrated into applications or production systems.

WHAT: Implemented the model into production.

HOW: Integrated the model into Power BI for real-world use, monitored performance, and maintained the model.

Every project was unique, so the Data Science Life Cycle could vary. It was crucial to carefully analyze each step of the process because a mistake at one would have impacted the subsequent phase and potentially ruined all the work.

3.2 System Requirement

3.2.1 Hardware

This project used a laptop as its hardware, which was used to perform data visualization using Power BI and data analysis utilizing prediction models like regression analysis. The following details were displayed for the laptop used for this project:

Description	Specifications
Model	Huawei MateBook D15
Processor	Intel Core i5-10210U
Operating System	Windows 11
Graphic	Intel UHD Graphics 620
Memory	8GB RAM
Storage	512GB

Table 3.2. 1 Specifications of laptop

3.2.2 Software

This project made use of Power BI software, which compiled all the data and displayed it on a dashboard. As a result, the Power BI system requirements for completing this project were as follows:

Description	System Requirements
Operating system	Windows 10 or Windows Server 2016 or later
Memory (RAM)	At least 2 GB available, 4 GB or more recommended
Display	At least 1440x900 or 1600x900 (16:9) required
CPU	1 gigahertz (GHz) 64-bit (x64) processor or better recommended

Table 3.2. 2 System Requirements of Power BI

3.3 Activity Diagram

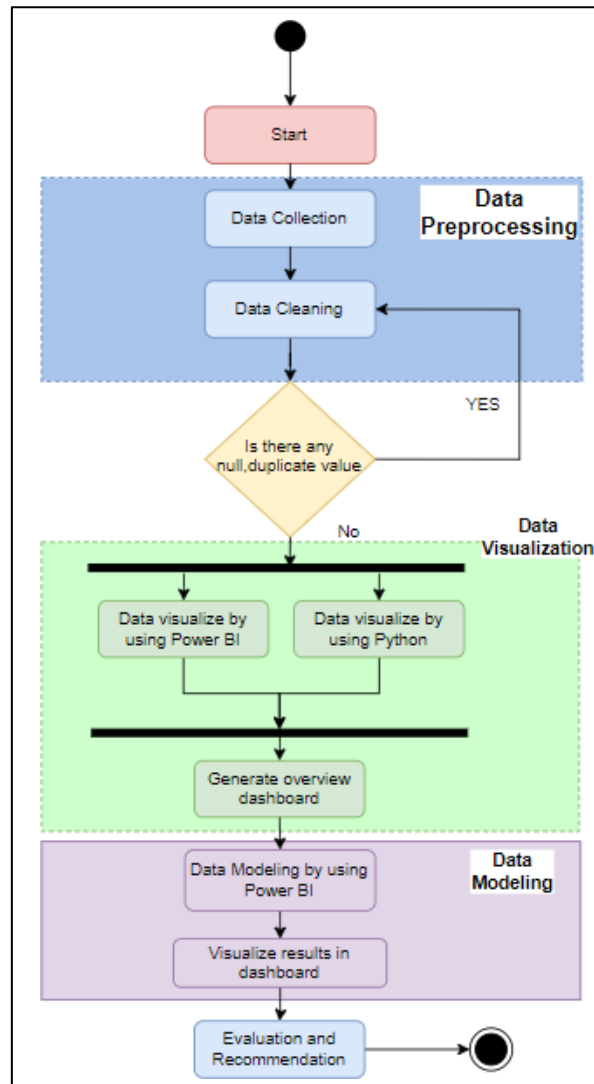


Figure 3.3. 1 Activity Diagram of Research

3.4 System Block Diagram

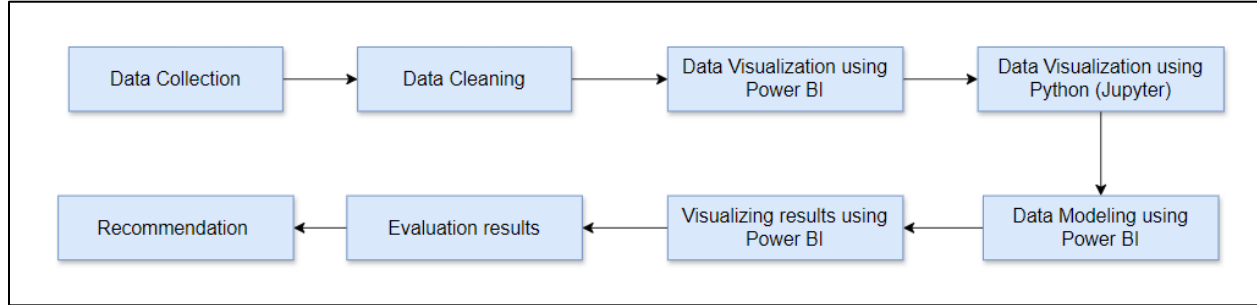


Figure 3.4 1 System Block Diagram of Data Analysis

Initially, data was collected from a group of students regarding their majors, gender, attitudes, behaviors, and various pros and cons of online shopping. This group of students came from several colleges, including the Faculty of Arts and Society (FAS), the Faculty of Business and Finance (FBF), the Faculty of Information and Communication Technology (FICT), and the Faculty of Science (FSc). The next step was data cleaning, which involved determining if any values were empty or duplicated, and if so, replacing or removing them. Because before starting any data analysis, it was crucial to confirm the reliability and correctness of the data. The data was moved and merged into Power BI, and then the overview was visualized through two application types, Power BI and Python, and the results were modeled using the Power BI tool. Finally, the findings were evaluated, and some recommendations were provided for future decisions.

3.5 Summary of this Chapter

This chapter discussed the System Model, and next chapter continued with the Experiment or Simulation.

CHAPTER 4

EXPERIMENT/SIMULATION (FOR RESEARCH- BASED PROJECT)

Last chapter summarized the system model or proposed method/ approach. In this chapter will discuss the Experiment/Simulation which involves Setting up, Visualization and segmentation, Timeline, Data collection and preparation, Checking the null value, Duplicate value, Visualize the results as dashboard by using Python and Jupyter.

4.1 Setting up

4.1.1 Software

There was some software that had to be installed on the laptop before the data collection can be visualized:

1. Power BI
2. Python
3. Anaconda Prompt (Jupyter)

4.2 Visualization and segmentation

In this project, a laptop is used as hardware, while Power BI software is used to visualize data and Python is used for data cleaning. The data of a sample of 224 students with different course backgrounds were converted into raw data and then presented in the form of a dashboard and visualized using Python programming tools. In addition, analysis was performed based on the processed results.

The following are the steps of data visualization process:

- I. Data collection and preparation
- II. Checking the null value, duplicate value by using Python (Jupyter)
- III. Visualize the results by using Power BI and Python

CHAPTER 4

- IV. The graphical presentation will be analyzed and discussed
- V. Provide recommendations to improve or solve the private students' online shopping behaviour

4.3 Timeline



Figure 4.3. 1 Timeline

4.4 Data collection and preparation

The process of obtaining and examining information or data from various sources in order to evaluate findings, predict patterns and likelihood, and pinpoint answers to research issues was known as data collection. In many forms of analysis, research, and decision-making, including business, social science, and medical work, it was an essential stage. Thus, reliable data collection was crucial to provide quality control, uphold research integrity, and make wise business judgements. During the data gathering process, researchers had to decide on the data kinds, data sources, and procedures [22].

The process of acquiring, compiling, organizing, and arranging data to make it suitable for use in business intelligence (BI), analytics, and data visualization applications was known as data preparation. Data preparation included tasks including profiling, cleansing, validation,

transformation, and preprocessing. Another common duty was gathering data from multiple internal and external sources.

As this ensured the validity of the results from BI and analytics applications, one of the primary goals of data preparation was to ensure that the raw data being processed and analyzed was accurate and consistent. Reconciling the various formats of two or more combined data sets was necessary because data frequently contained errors, missing numbers, and other issues. Resolving data issues, verifying data accuracy, and combining data sets were significant parts of data preparation job descriptions [23].

In this process, we collected a handful of students' responses based on each of their respective faculty. Students' responses were divided into several categories which are attitude, behaviours, perceived benefits, and risks. The respondents who answered it were from different faculty of study which are Faculty of Arts and Social Science (FAS), Faculty of Business and Finance (FBF), Faculty of Information and Communication Technology (FICT), and Faculty of Science (FSc).

4.5 Checking the Null Value, Duplicate Value

It was critical to ensure that there were no null or duplicate values in your data, as these issues could have made the information incomplete and meaningless. The analysis lacked completeness and accuracy in identifying any patterns or problems if NULL values were not checked. Additionally, this could have resulted in poor decision-making [24]. Therefore, Python (Jupyter) was used to check the null value and duplicate value.

CHAPTER 4

The first step was to download the database and store it in the Downloads folder.

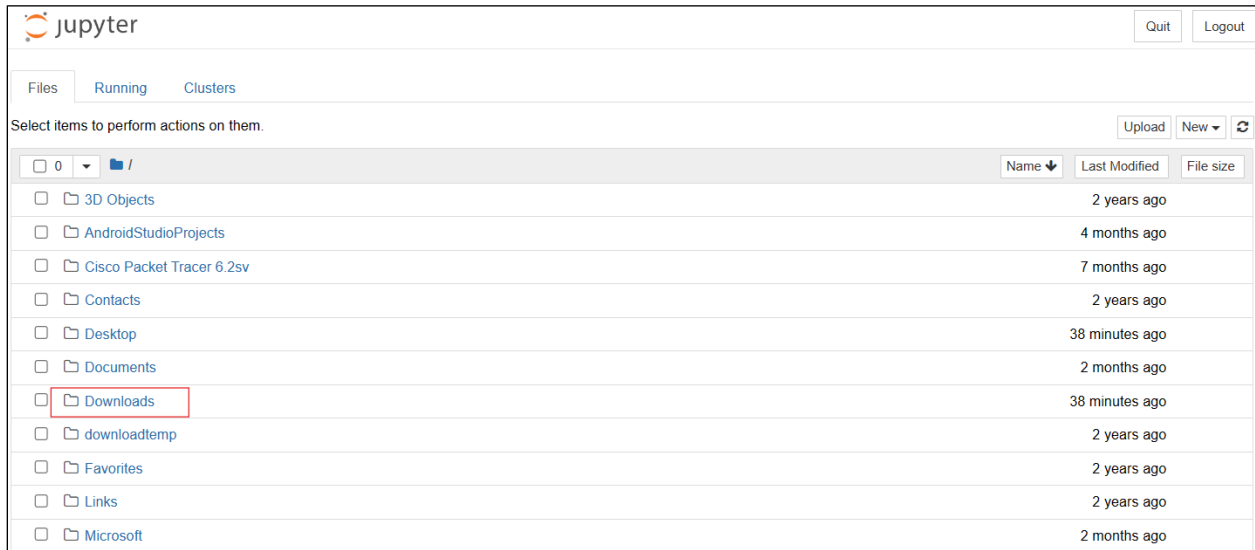


Figure 4.5. 1 Locate the database saved

The next stage involved changing the data source by purging the data to provide consistent, error- and anomaly-free data that facilitates data visualization and yields more insightful outcomes. To begin data cleaning, clicked New and selected "Python 3 (ipykernel)".



Figure 4.5. 2 Select 'Python 3 (ipykernel)'

CHAPTER 4

Reading the data file and importing pandas were the next steps.

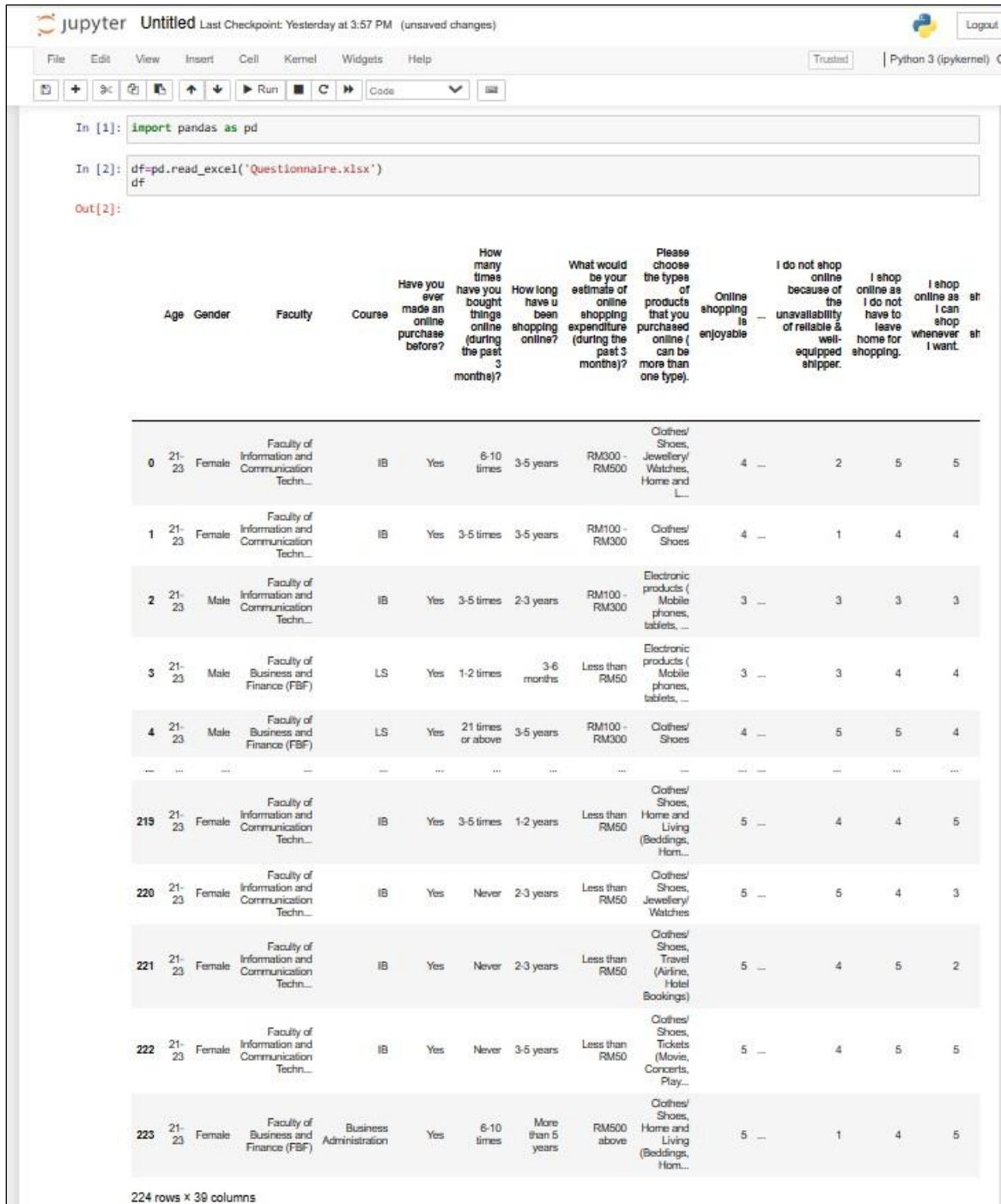
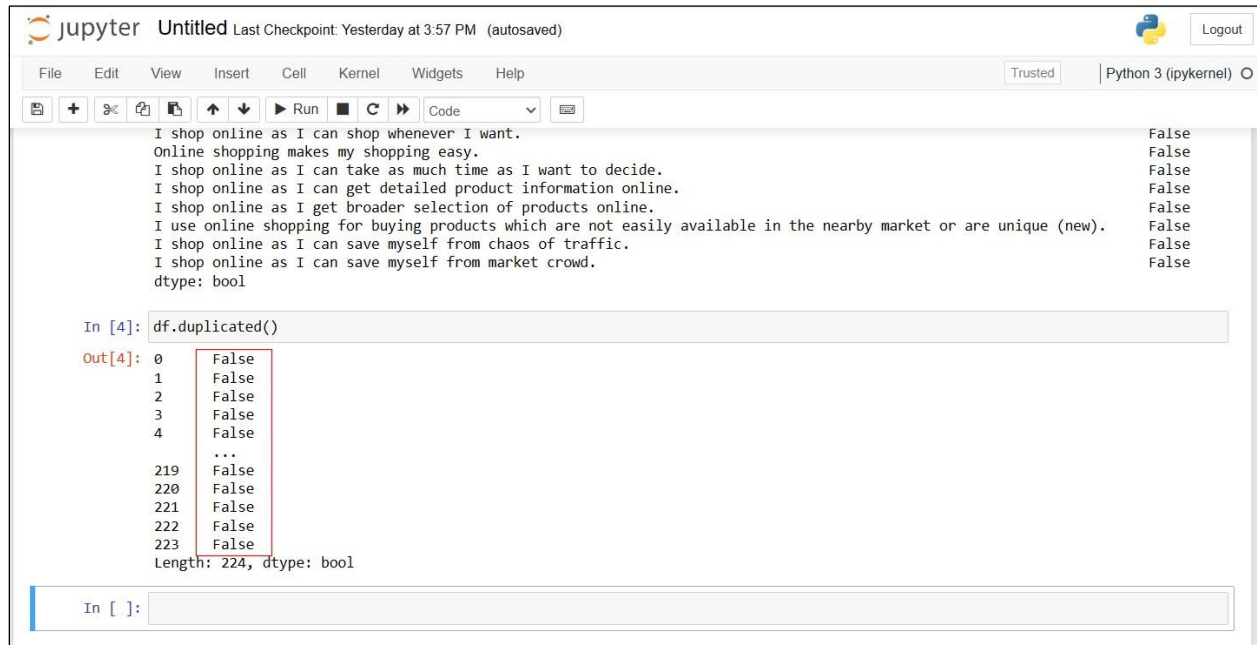


Figure 4.5. 3 Import the data and read it from located file

CHAPTER 4

Next, checked if there were any duplicate values. **True** indicated that there was a duplicate value, whereas **False** indicated that there wasn't one.



```

I shop online as I can shop whenever I want.
Online shopping makes my shopping easy.
I shop online as I can take as much time as I want to decide.
I shop online as I can get detailed product information online.
I shop online as I get broader selection of products online.
I use online shopping for buying products which are not easily available in the nearby market or are unique (new).
I shop online as I can save myself from chaos of traffic.
I shop online as I can save myself from market crowd.
dtype: bool
False
False
False
False
False
False
False
False
False

In [4]: df.duplicated()
Out[4]: 0    False
        1    False
        2    False
        3    False
        4    False
        ...
        219  False
        220  False
        221  False
        222  False
        223  False
Length: 224, dtype: bool

In [ ]:
```

Figure 4.5. 5 Checking duplicate values

4.6 Visualize the results as dashboard by using Power BI

The business intelligence tool Power BI was utilized in this process to visualize the results. In the first stage, Excel-formatted sample data on students' online shopping behaviors was merged with Power BI. First collected information from the Excel workbook to locate the database, which had to be integrated into the application.

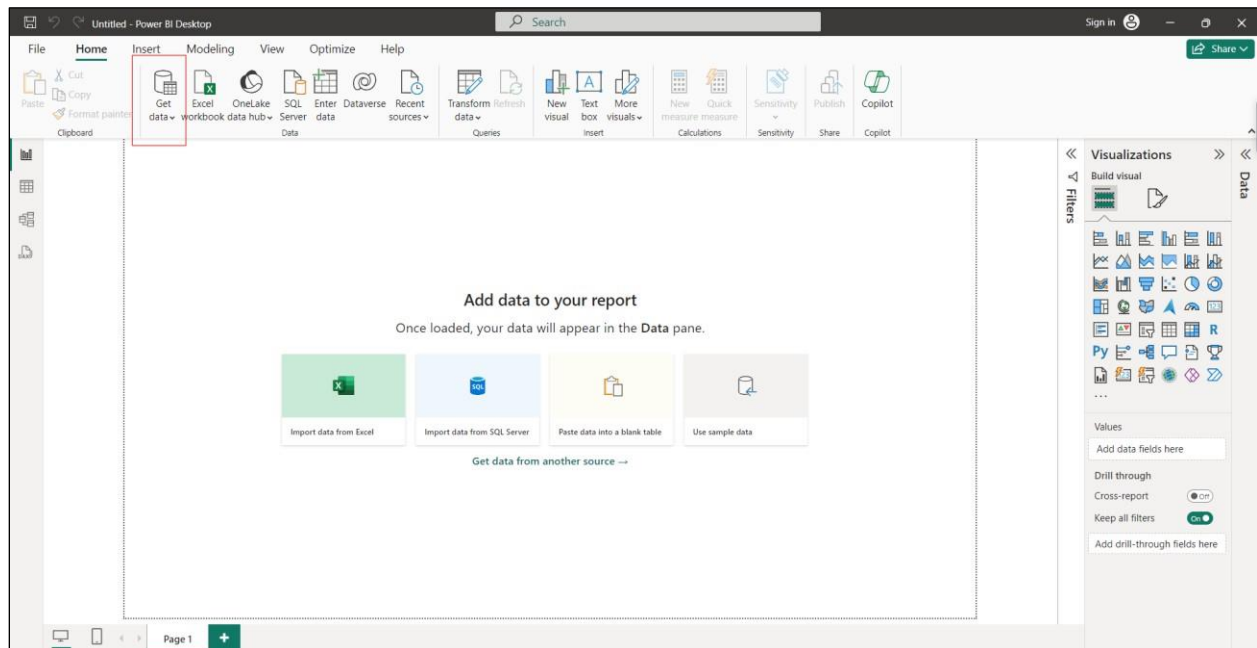


Figure 4.6. 1 Get the data (Step 1 of Data Visualization)

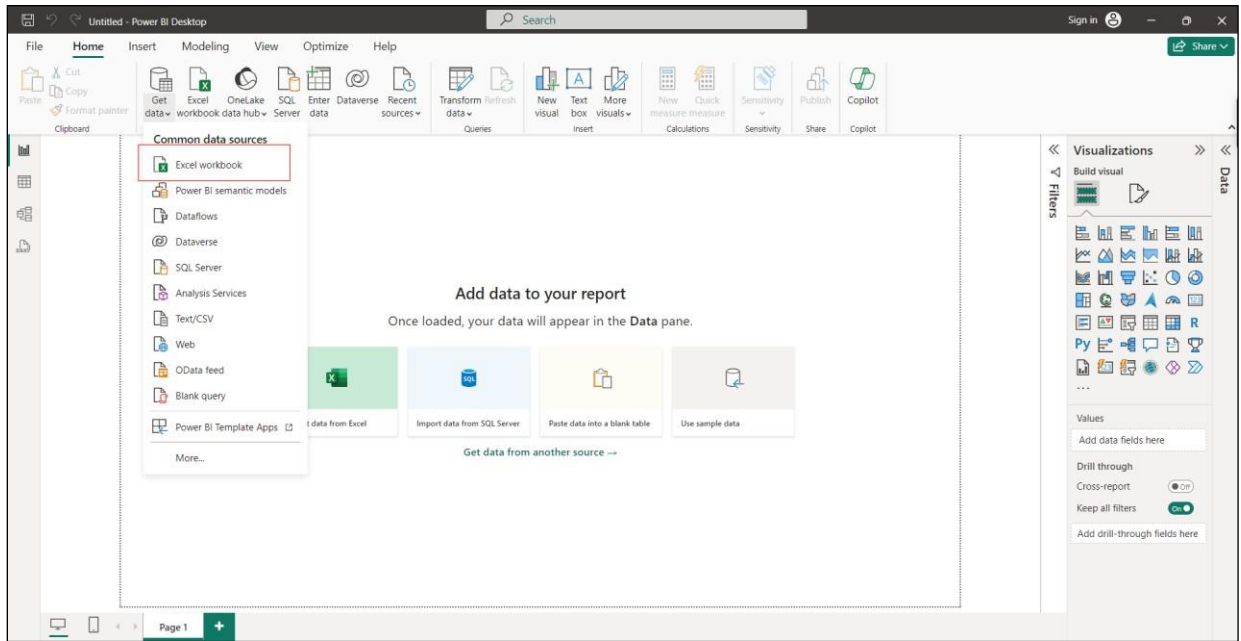


Figure 4.6. 2 Select Excel Workbook (Step 2 of Data Visualization)

In the following phase, the data was transformed to make certain modifications to the queries, such as changing the names of the columns, eliminating null fields, and filtering out only the important data that required analysis, before being loaded. This was done in the hopes that they could be combined into a pertinent and insightful query.

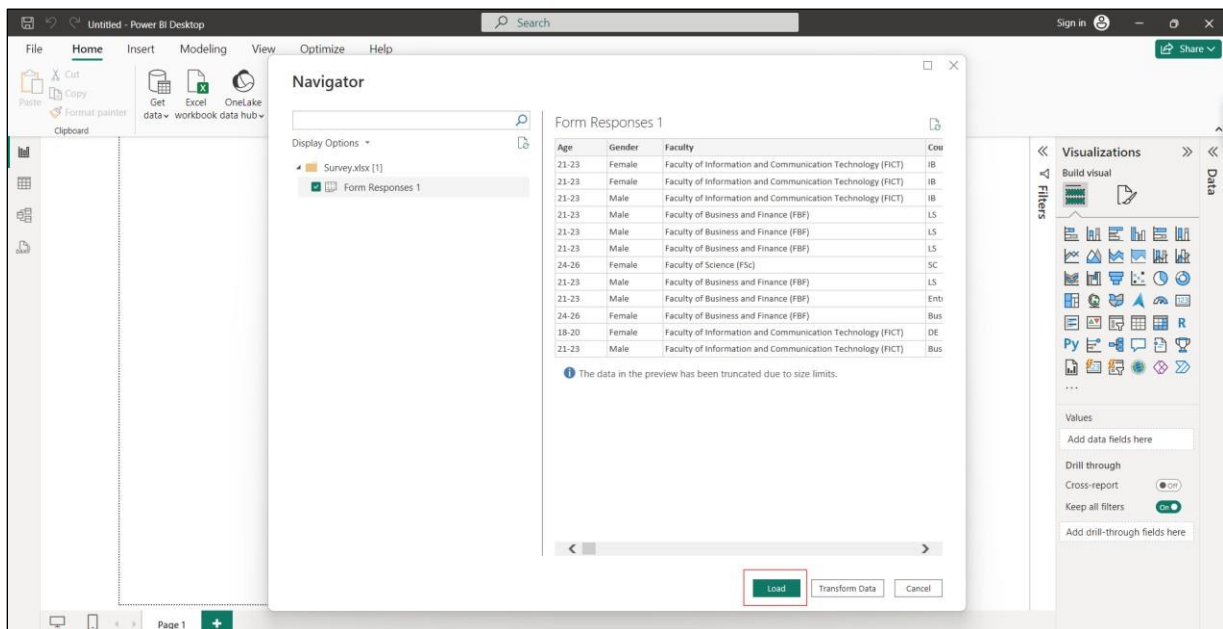


Figure 4.6. 3 Load the data (Step 3 of Data Visualization)

CHAPTER 4

Subsequently, and possibly as a last step, the data was visualized by dragging or selecting the required data from the right side into the visualization fields to display it as a stacked column chart, pie chart, bar chart, table, area chart, tree map, and line chart.

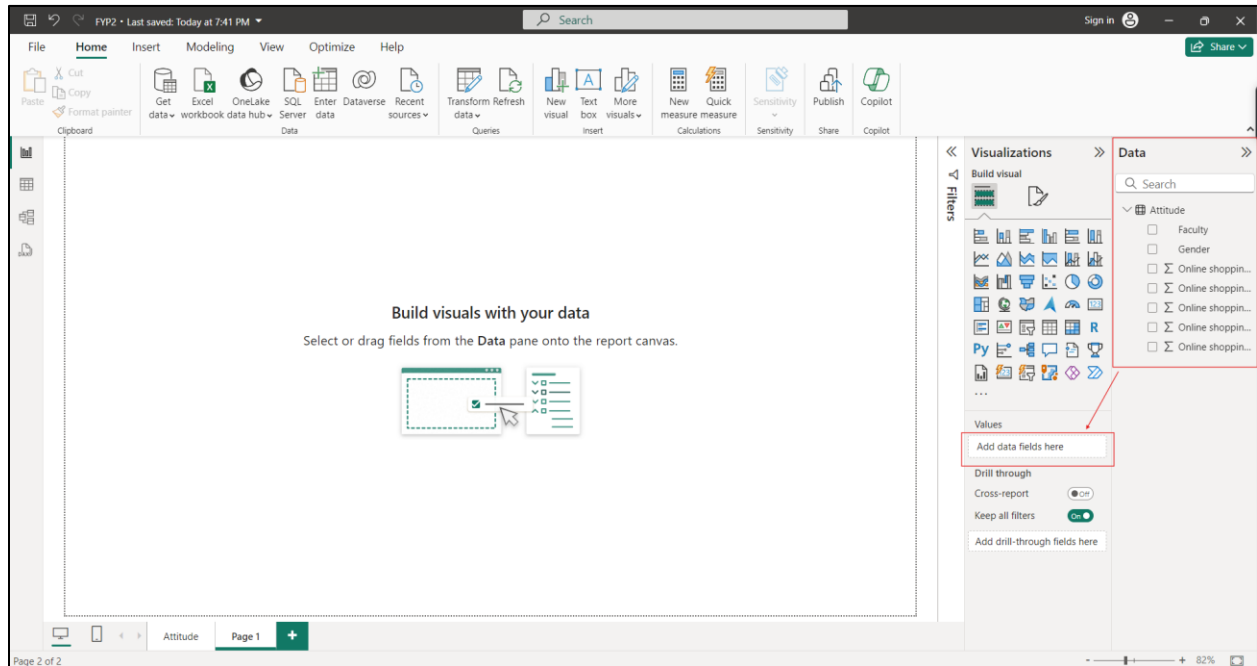


Figure 4.6. 4 Visualize the data (Step 4 of Data Visualization)

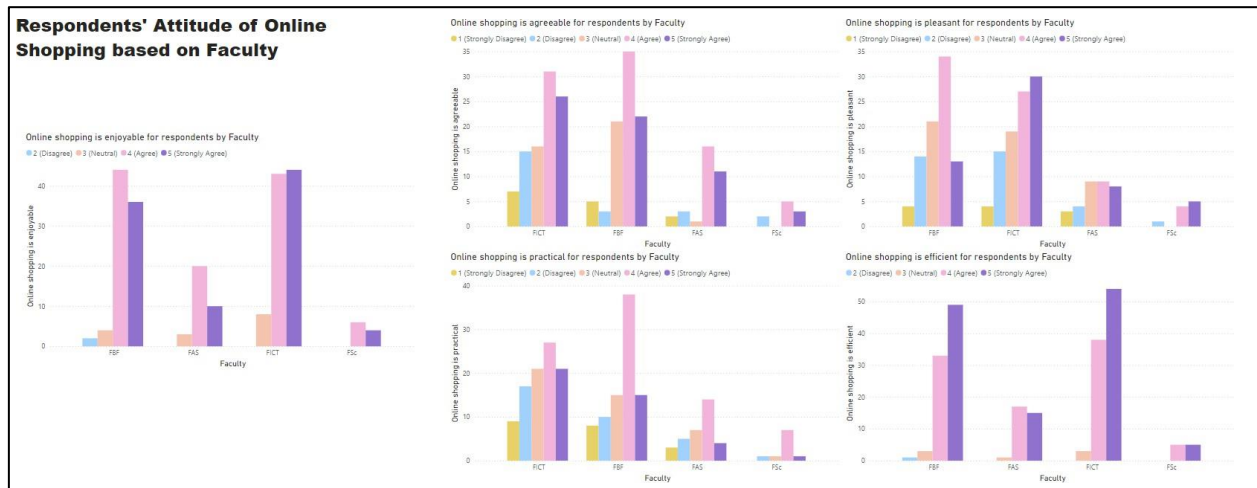


Figure 4.6. 5 Attitude

The bar charts showed how students felt about purchasing online in various faculties. In general, students from the Faculty of Information and Communication Technology (FICT) had the most positive sentiments overall, with a sizable portion strongly concurring that purchasing online was efficient, practical, agreeable, pleasant, and pleasurable. Strongly positive comments were also shown by the Faculty of Business and Finance (FBF), especially when it came to how enjoyable and effective internet purchasing was. Mixed feelings were indicated by a balanced distribution of viewpoints across the Likert scale, which was more variable among students from the Faculty of Arts and Social Science (FAS). Ultimately, fewer respondents strongly agreed with every aspect of online buying, indicating that Faculty of Science (FSc) students generally exhibited the least excitement for it.

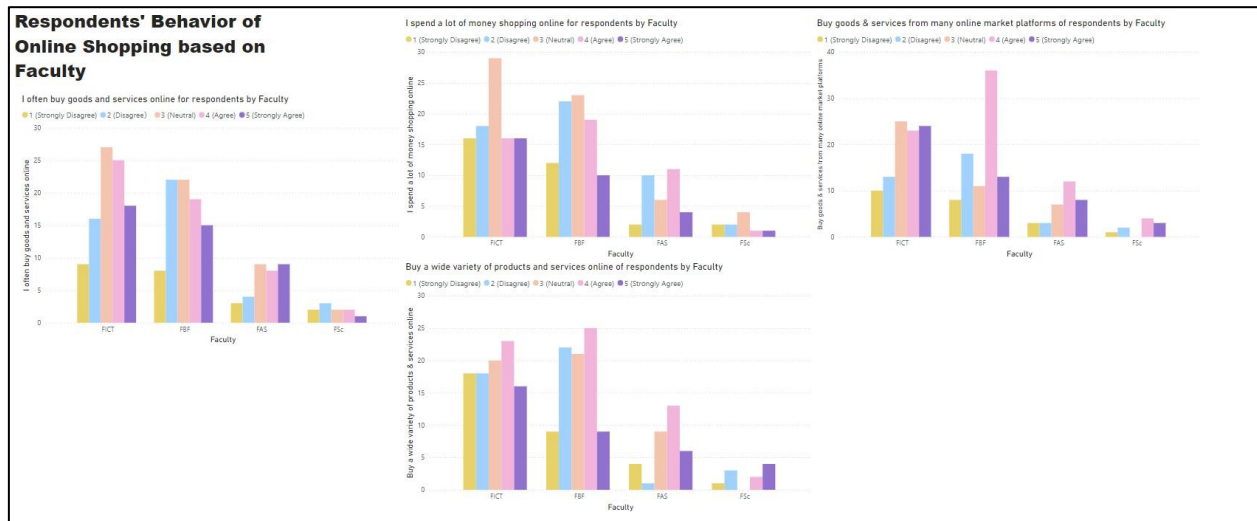


Figure 4.6. 6 Behavior

The bar charts shed light on how students from various faculties behaved when they shop online. Online purchases were common among FICT students, many of them agreed or strongly agreed that they spent a lot of money, used a variety of platforms, and bought a wide range of goods and services. Even while their purchasing patterns varied slightly, they typically engaged in a variety of shared activities, including internet buying. FBF students also engaged in active online purchasing, particularly when utilizing numerous platforms, but their spending was more diverse, with a sizable portion exhibiting slight agreement or neutrality. With a mixture of agreement, neutrality, and disagreement across all elements, FAS students showed a more balanced and less consistent online shopping behavior, indicating less frequent and diverse online purchase. When it came to frequency of transactions, diversity of products purchased, and use of numerous platforms, FSC students exhibited the lowest levels of engagement with online shopping. Their opinions tended to be neutral or disagreeable.

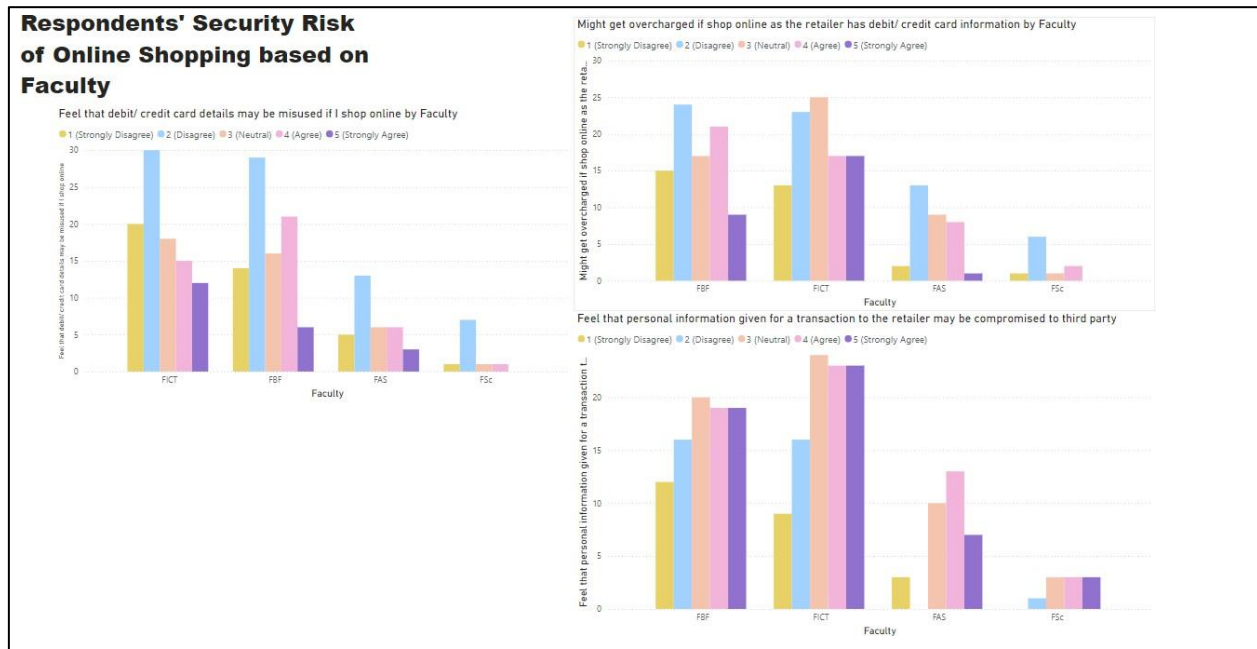


Figure 4.6. 7 Security Risk

The bar charts showed the differing levels of worry about online shopping security issues across faculties. FICT students were the most concerned, with many strongly agreeing that their debit or credit card information could be exploited, that they could be overcharged because businesses had their card information, and that their personal information could be stolen by other parties. FBF students had comparable worries, notably about overcharging and personal information security, but their replies were more equally distributed, with a mix of strong agreement, neutrality, and dissent. FAS students showed moderate concern, with more balanced replies indicating that, while they were concerned about these hazards, they were less nervous than FICT and FBF students. In contrast, FSc students were the least concerned about online shopping security, with the majority of responses leaning towards neutrality or disagreement, indicating little fear about money or personal information being exposed.

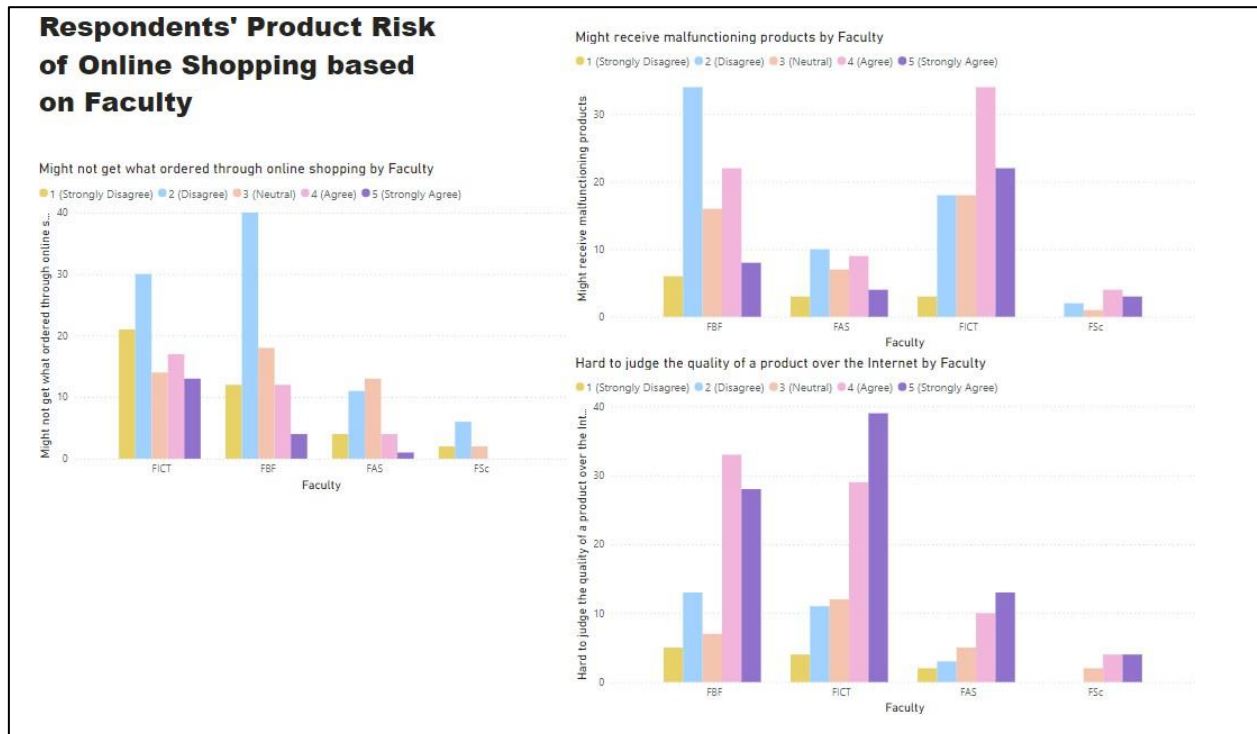


Figure 4.6. 8 Product Risk

The bar charts showed various levels of worry among faculty about the product dangers linked with online buying. FBF students were the most concerned about not receiving what they requested, with many strongly agreeing that this was an issue. FICT students indicated great worry, albeit their replies were more evenly distributed between agreement and disagreement. FAS students were moderately concerned, whereas FSc students were the least concerned about the matter. When it came to the possibility of receiving defective products, FICT students were the most concerned, with many strongly agreeing, whilst FBF students had comparable concerns but with more diverse opinions. FAS students were moderately concerned, while FSc students exhibited minor fear. Finally, FICT students found determining product quality online to be the most difficult, with many strongly agreeing. FBF students also struggled with this, but their comments were more balanced. FAS students were moderately concerned, while FSc students were least worried by this difficulty. Overall, FICT and FBF students were more concerned about online buying hazards than FAS students, who were moderately concerned.

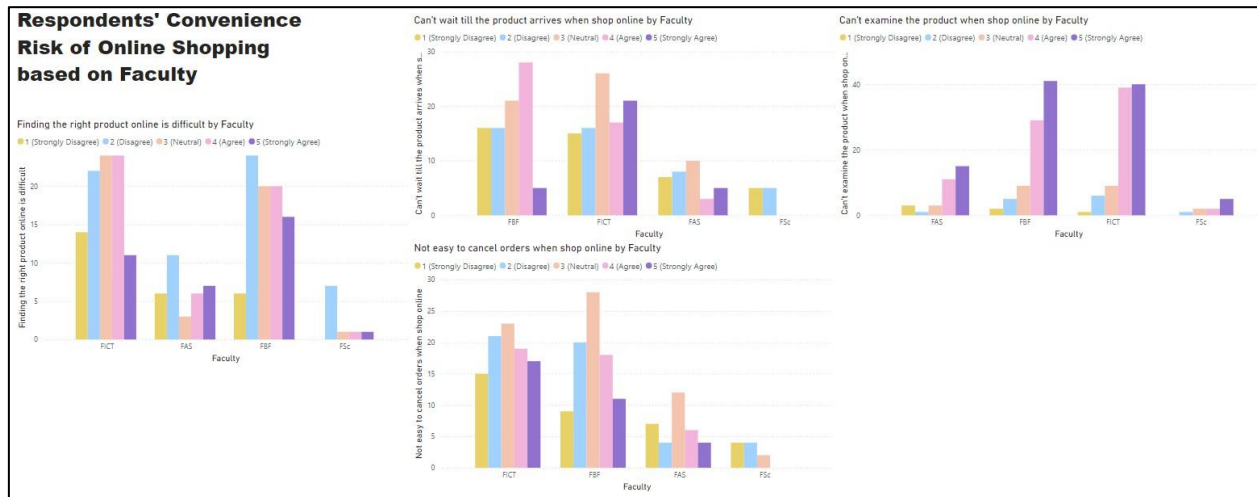


Figure 4.6. 9 Convenience Risk

The bar charts depicted the numerous convenience hazards connected with internet buying, as evaluated by students from various faculties. FBF students had the most trouble selecting the proper things online and the highest degrees of irritation when waiting for their orders to arrive. They were also concerned about being unable to see things before purchasing and found it difficult to cancel orders once completed. FICT students had comparable worries, particularly about the inability to inspect things and the difficulties of cancelling orders, albeit their experiences with discovering products and waiting for delivery differed. FAS students were moderately concerned about all of these hazards, showing a variety of experiences but less intensity in their anxieties. On the other hand, FSc students exhibited the least amount of anxiety regarding product discovery, shipping wait periods, product inspection, and purchase cancellation. All things considered, the students in FBF and FICT were most affected by these convenience hazards.

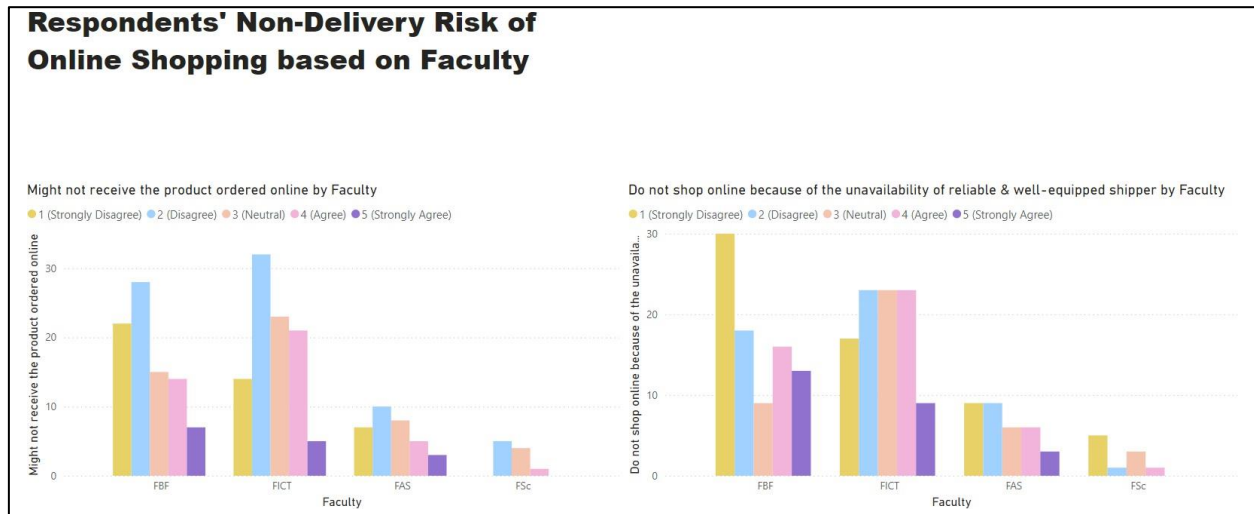


Figure 4.6. 10 Non-Delivery Risk

The two bar charts depicted the non-delivery risks connected with online purchase, as evaluated by students across faculties. In the first graphic, which depicted concerns about not receiving the product ordered online, FBF students showed the greatest level of worry, with many firmly agreeing that this was a big danger. A significant number of students agreed or were ambivalent, showing that they were concerned about non-delivery. FICT students exhibited great concern, with many strongly agreeing, albeit their opinions were more evenly spread. FAS students were moderately concerned about non-delivery, with an even split of agreement, indifference, and disagreement. Meanwhile, FSc students were the least concerned, with most responses leaning towards disagreement or indifference, showing that non-delivery was not a huge issue for them. The second chart focused on students' hesitation to shop online due to a lack of reputable and well-equipped shippers. FBF students were again the most concerned, with a sizable number strongly believing that this was a disincentive to online buying. However, there was a sizable proportion who disagreed or were neutral, implying that while many students were concerned, others were less affected by this issue. FICT students were also concerned, with many agreeing or strongly agreeing, although their views varied. FAS students were moderately concerned, with a mixed bag of replies indicating that although some found this a hindrance, others didn't. FSc students were the least concerned about the availability of dependable shippers, with most responses being neutral or disagreeing, indicating that this is not a serious issue for them. Overall, non-delivery risks had the greatest impact on FBF and FICT students, while FAS students were moderately concerned and FSc students were the least concerned.

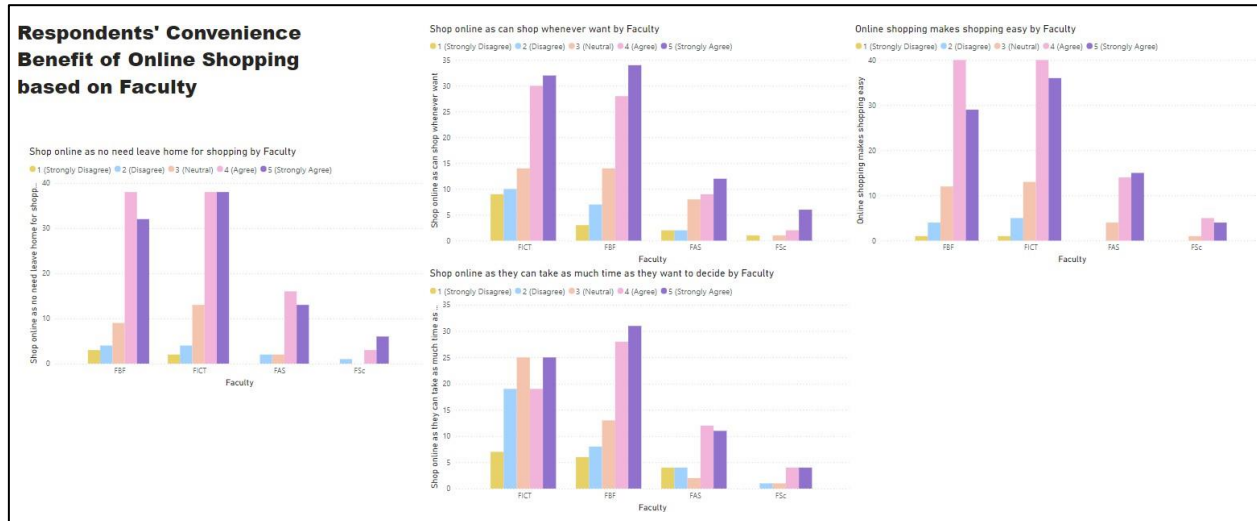


Figure 4.6. 11 Convenience Benefit

The bar charts depicted the convenience benefits of internet buying as viewed by students across faculties. In the first chart, which covered the convenience of not having to leave the house to shop, FICT students largely agreed, with a strong preference for "Strongly Agree," showing that they placed a high value on this element. FBF students similarly valued this benefit, however their replies were more evenly distributed throughout agreement levels. FAS and FSc students had intermediate agreement, with fewer strongly agreeing. The second chart emphasized the convenience of being able to shop when customers wished. FICT students were once again unanimous, with many strongly stating that this flexibility was a big benefit. FBF students paid great attention and appreciated the convenience. FAS and FSc students again demonstrated reasonable agreement, with fewer strongly agreeing. The fourth figure demonstrated students' appreciation for the time flexibility that internet shopping provided, allowing them to take as long as they needed to decide. FICT students once again demonstrated great agreement, with FBF students trailing closely behind. FAS and FSc students showed moderate agreement, as do the other charts. Overall, FICT and FBF students were the most excited about the convenience benefits of online purchasing, particularly the flexibility and ease it provided, whilst FAS and FSc students were moderately positive but not completely convinced.

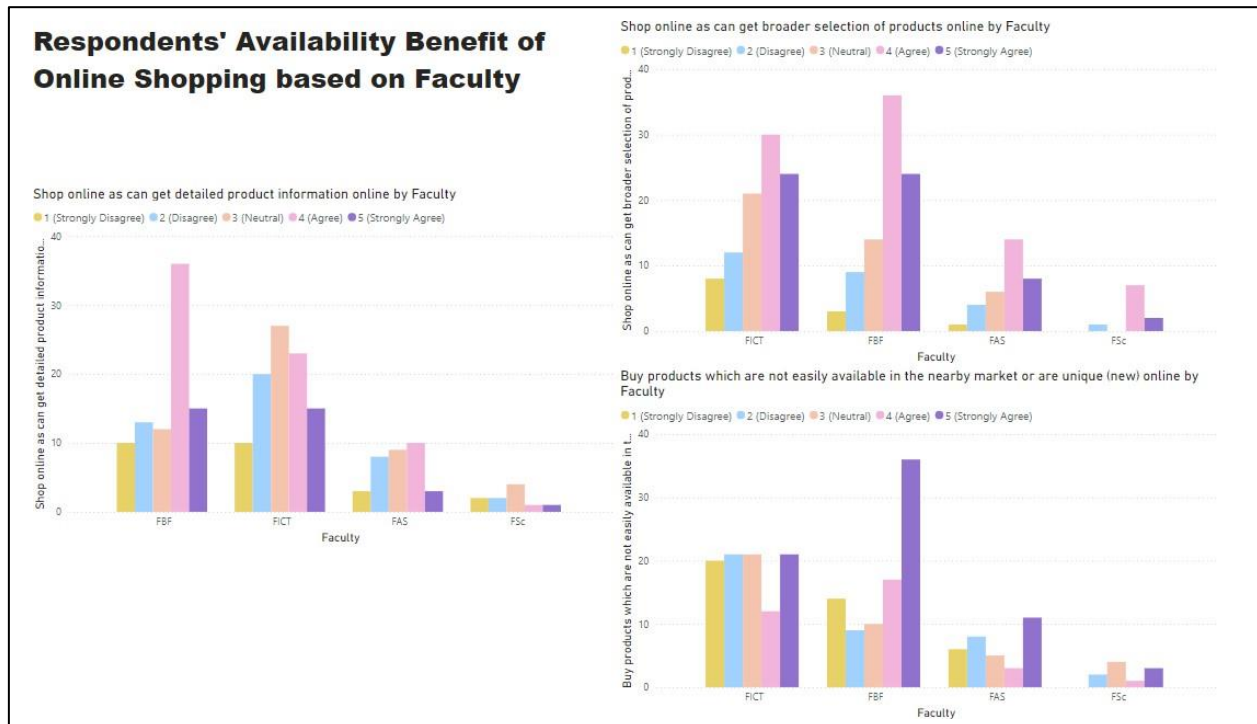


Figure 4.6. 12 Availability Benefit

The bar charts depicted the availability and benefits of internet purchasing as seen by students from various faculties. In the first bar chart, which covered the benefit of getting extensive product information online, FBF students had the highest degree of agreement, with a high concentration of responses in the "Strongly Agree" group, showing they placed a high value on this feature. FICT students similarly valued this benefit, however their replies were more evenly split throughout agreement levels. FAS and FSc students had more moderate agreement, with fewer strongly agreeing. The second chart emphasized the benefits of having access to a larger assortment of products online. FBF students again showed considerable agreement, with a sizable number strongly agreeing that this was a significant advantage. FICT students valued this feature as well, albeit they were less enthusiastic about it. FAS and FSc students exhibited moderate agreement, with some leaning towards neutrality. In the third bar chart, which evaluated the ability to purchase things that were not widely available in adjacent marketplaces or were unique, FBF students once again demonstrated the highest degree of agreement, with many strongly believing that this is an essential benefit. FICT students agreed, albeit to a lesser level. FAS and FSc students again demonstrated more intermediate levels of agreement, with fewer students completely convinced of the benefit. Overall, FBF students are the most enthused about the availability benefits of online

buying, particularly detailed product information, a wider selection, and access to rare or hard-to-find items. FICT students share this enthusiasm to a significantly lower extent, whilst FAS and FSc students are moderately optimistic but less sure.

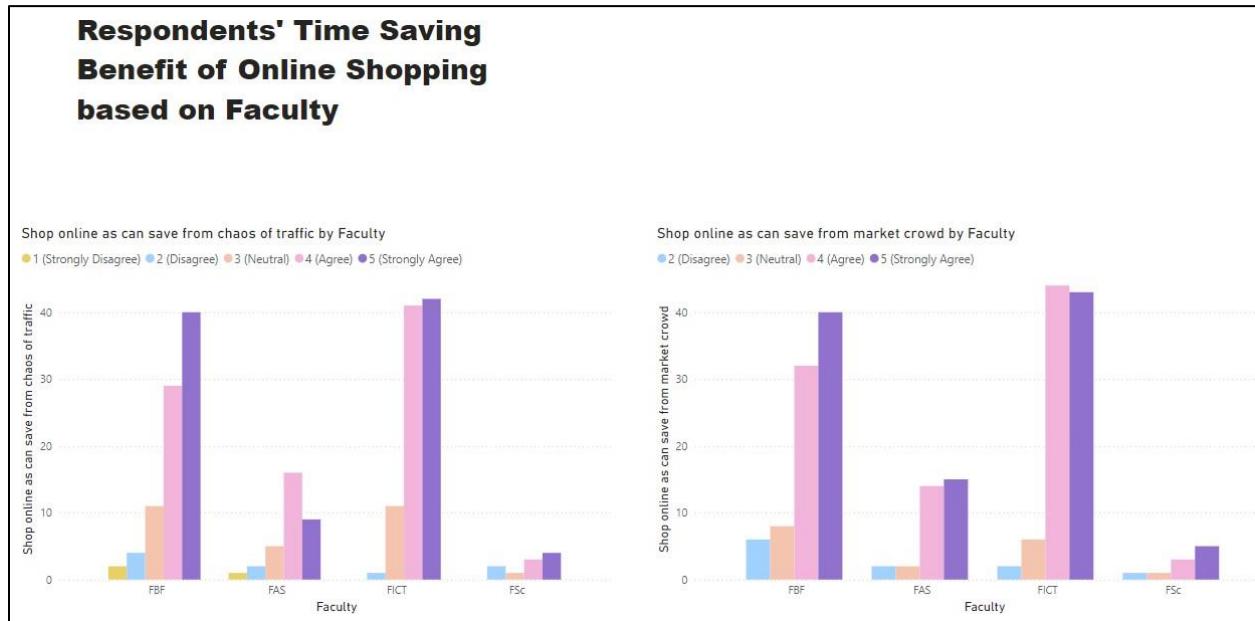


Figure 4.6. 13 Time Saving Benefit

The bar charts depicted the time-saving benefits of internet purchasing as viewed by students across faculties. Students from the Faculty of Business and Finance (FBF) had the highest degree of agreement in the first chart, which discussed the benefit of avoiding traffic mayhem by purchasing online. A significant number of these students fell into the "Strongly Agree" group, showing that they placed a high importance on time savings. Students in the Faculty of Information and Communication Technology (FICT) likewise valued this benefit, albeit their replies were more evenly split across the agreement levels. Students in the Faculty of Arts and Social Sciences (FAS) and the Faculty of Science (FSc) had more moderate agreement, with fewer strongly agreeing. The second graphic showed how shopping online allows you to avoid crowded markets. Again, FBF students indicated considerable agreement, with a sizable majority strongly agreeing that this was a big advantage. FICT students valued this feature as well, although they are less enthusiastic about it. Students in FAS and FSc exhibited moderate agreement, with some leaning towards neutrality. Overall, FBF students were the most enthused about the time-saving benefits of online shopping,

CHAPTER 4

Next, the database that had been saved in the Downloads folder was located.

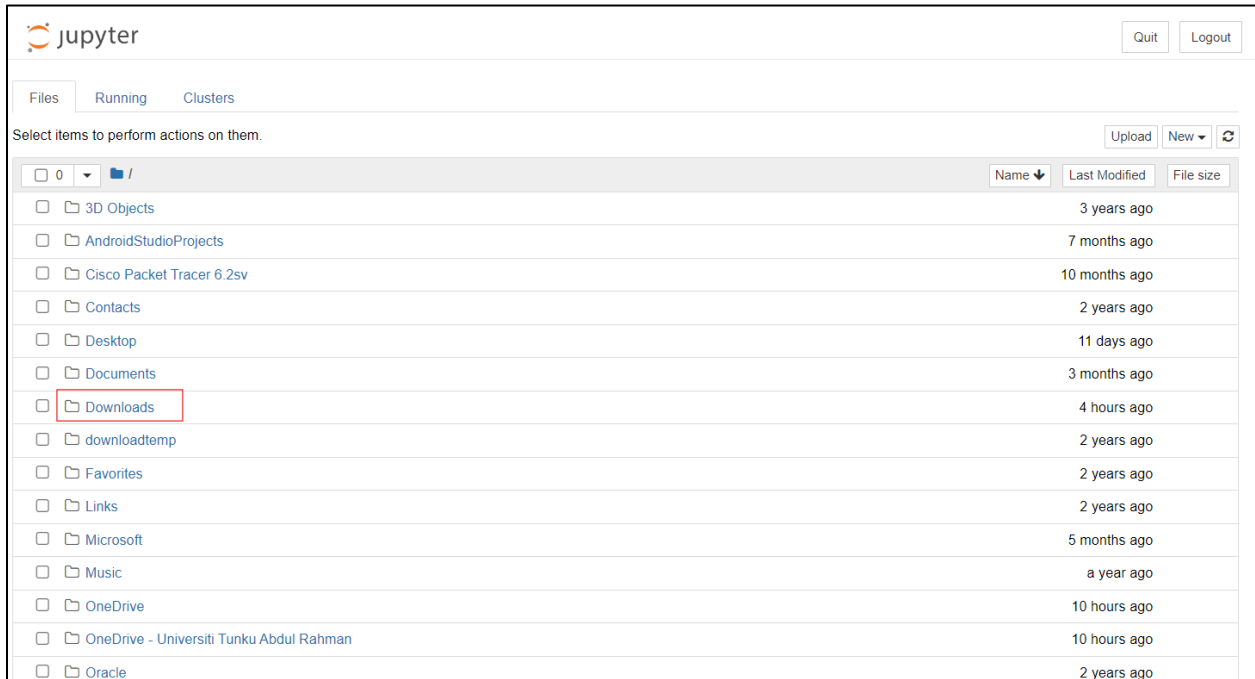


Figure 4.7. 2 Locate the database saved

After locating the dataset, data visualizing could begin by clicking on New and selecting 'Python 3 (ipykernel)'.

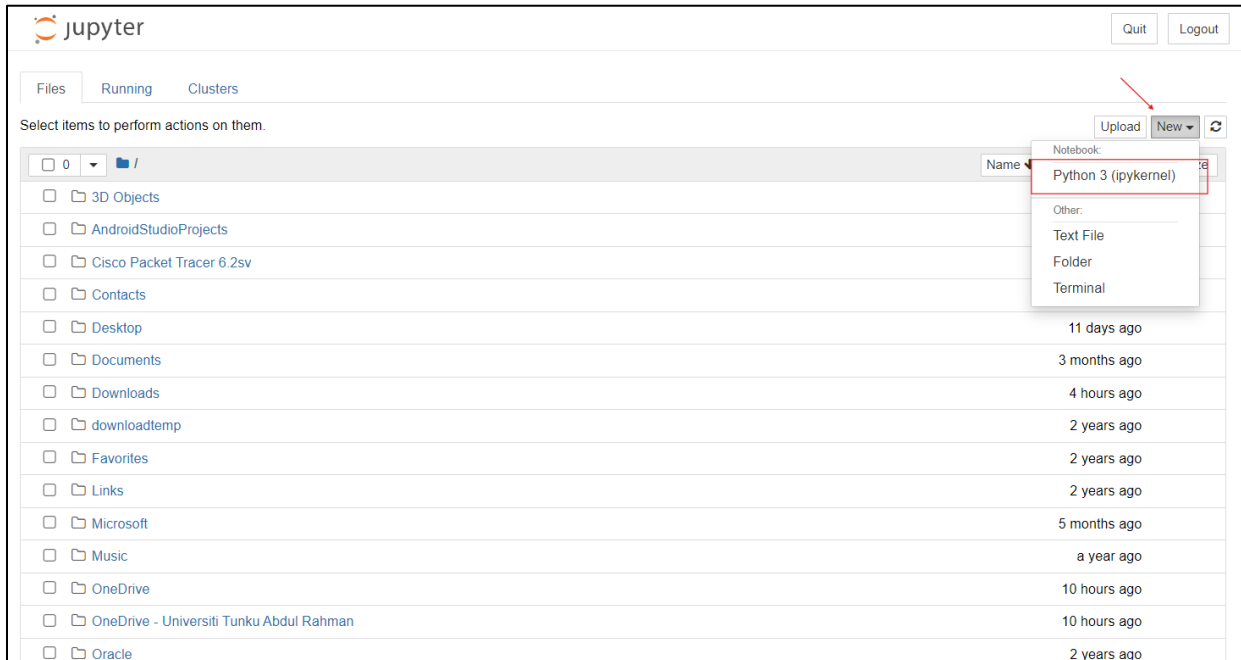
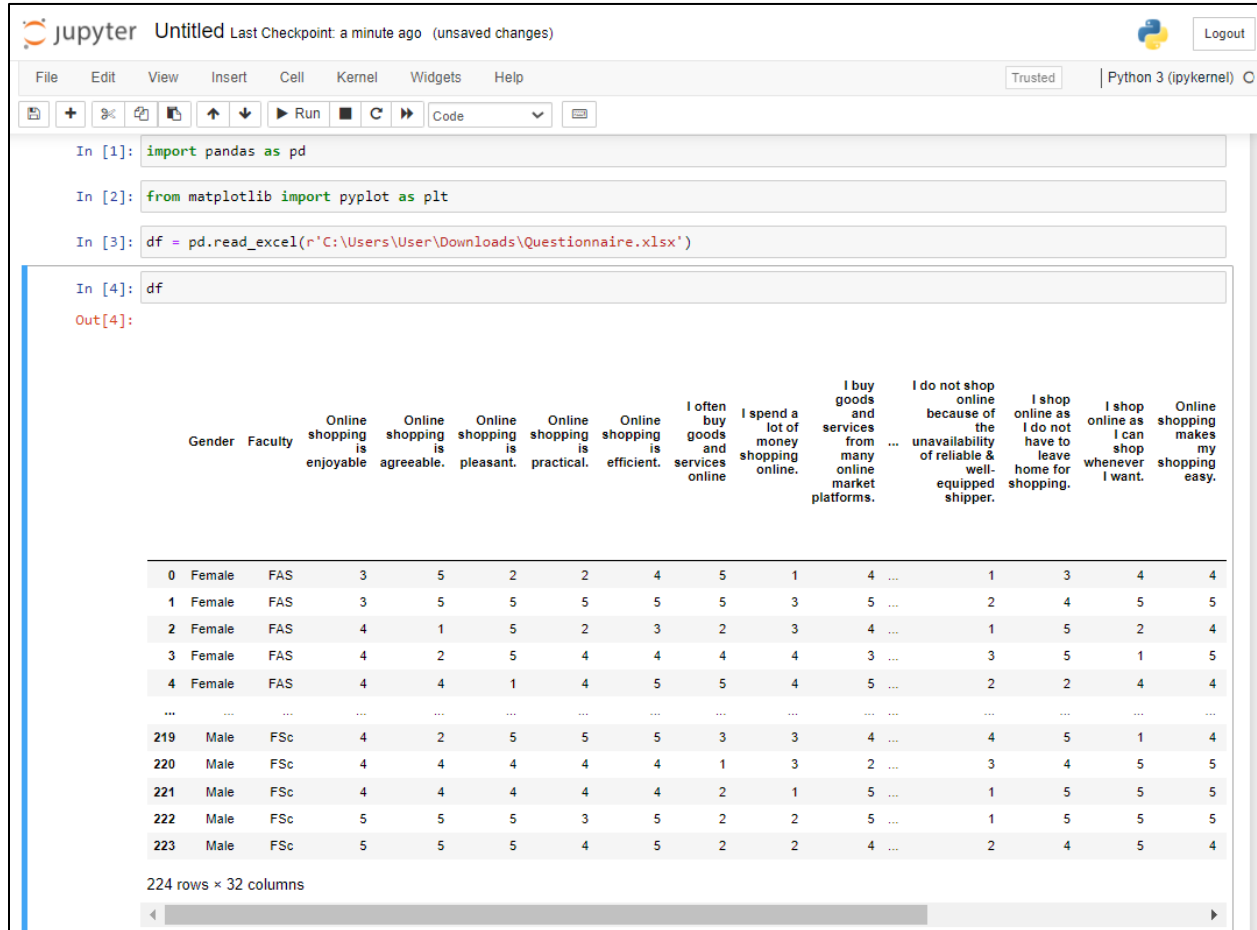


Figure 4.7. 3 Click on New and select 'Python 3 (ipykernel)'

CHAPTER 4

The matplotlib library and pandas were then imported to produce static, animated, and interactive visualizations. The sample data is transferred from the file location into the web application, where it was displayed and ready for further visualization.



The screenshot shows a Jupyter Notebook interface with the following code cells and output:

```
In [1]: import pandas as pd
In [2]: from matplotlib import pyplot as plt
In [3]: df = pd.read_excel(r'C:\Users\User\Downloads\Questionnaire.xlsx')
```

Out[4]:

	Gender	Faculty	Online shopping is enjoyable	Online shopping is agreeable.	Online shopping is pleasant.	Online shopping is practical.	Online shopping is efficient.	I often buy goods and services online	I spend a lot of money shopping online.	I buy goods and services from many online market platforms.	I do not shop online because of the unavailability of reliable & well-equipped shipper.	I shop online as I do not have to leave home for shopping.	I shop online as I can shop whenever I want.	Online shopping makes my shopping easy.
0	Female	FAS	3	5	2	2	4	5	1	4	1	3	4	4
1	Female	FAS	3	5	5	5	5	5	3	5	2	4	5	5
2	Female	FAS	4	1	5	2	3	2	3	4	1	5	2	4
3	Female	FAS	4	2	5	4	4	4	4	3	3	5	1	5
4	Female	FAS	4	4	1	4	5	5	4	5	2	2	4	4
...
219	Male	FSc	4	2	5	5	5	3	3	4	4	5	1	4
220	Male	FSc	4	4	4	4	4	1	3	2	3	4	5	5
221	Male	FSc	4	4	4	4	4	2	1	5	1	5	5	5
222	Male	FSc	5	5	5	3	5	2	2	5	1	5	5	5
223	Male	FSc	5	5	5	4	5	2	2	4	2	4	5	4

224 rows × 32 columns

Figure 4.7. 4 Read the data from the located file

CHAPTER 4

To visualize the data, the matplotlib and pandas libraries were implementing, which produced the clustered bar chart in descending order. This bar chart was visualized based on the count of attitude (y-axis) and faculty (x-axis) and with the gender (legend).

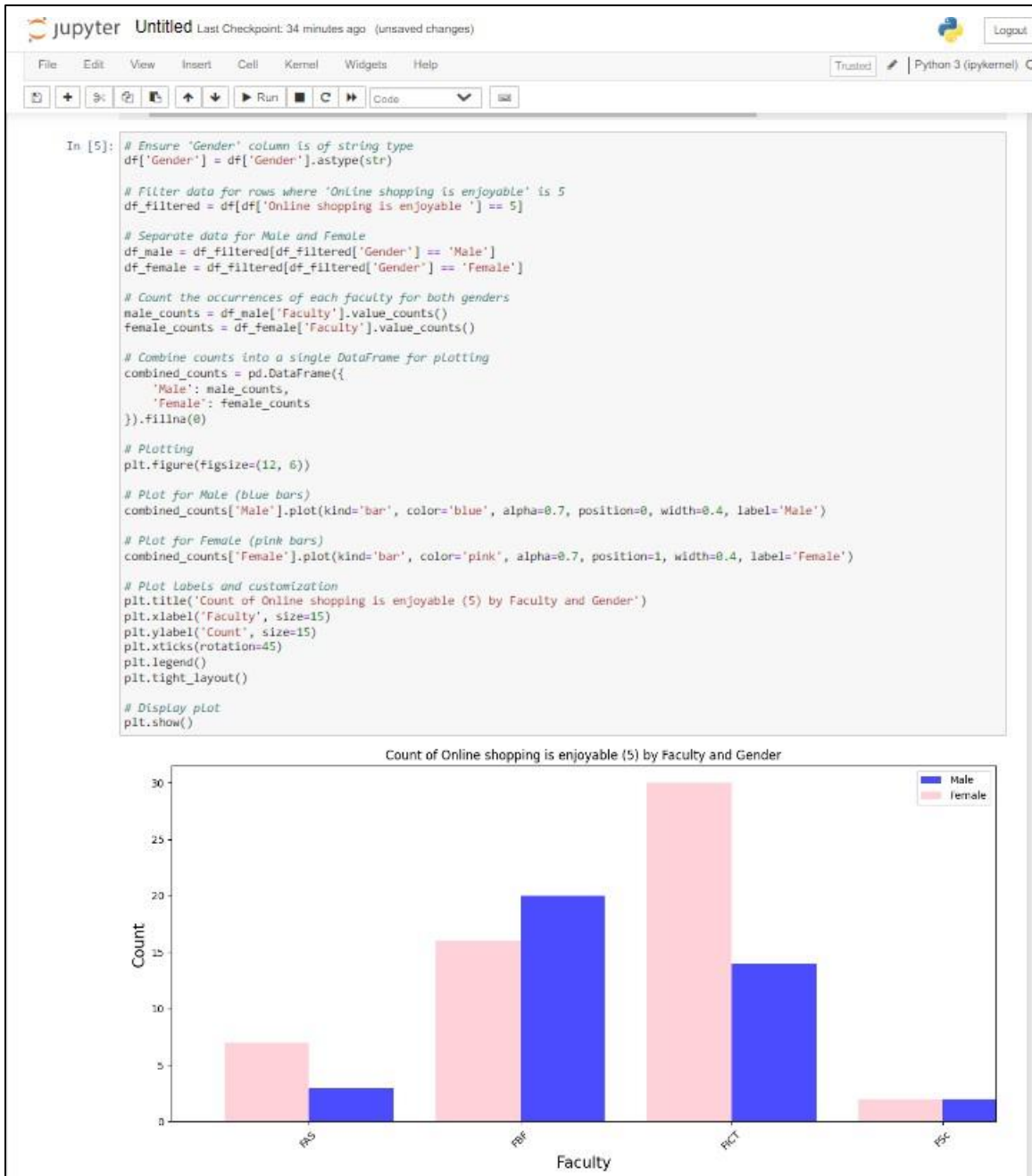


Figure 4.7. 5 A Clustered Column Chart output of Online Shopping is Enjoyable (5)

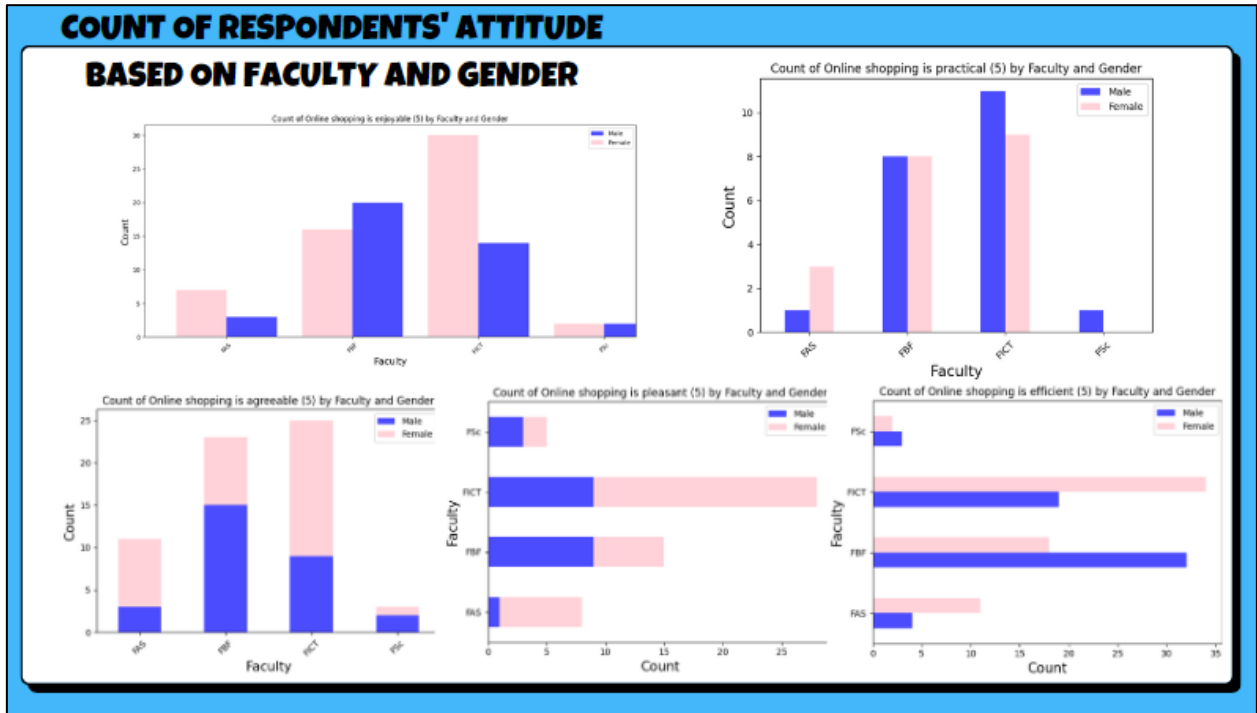


Figure 4.7. 6 Dashboard overview Attitude using Python (Jupyter Notebook)

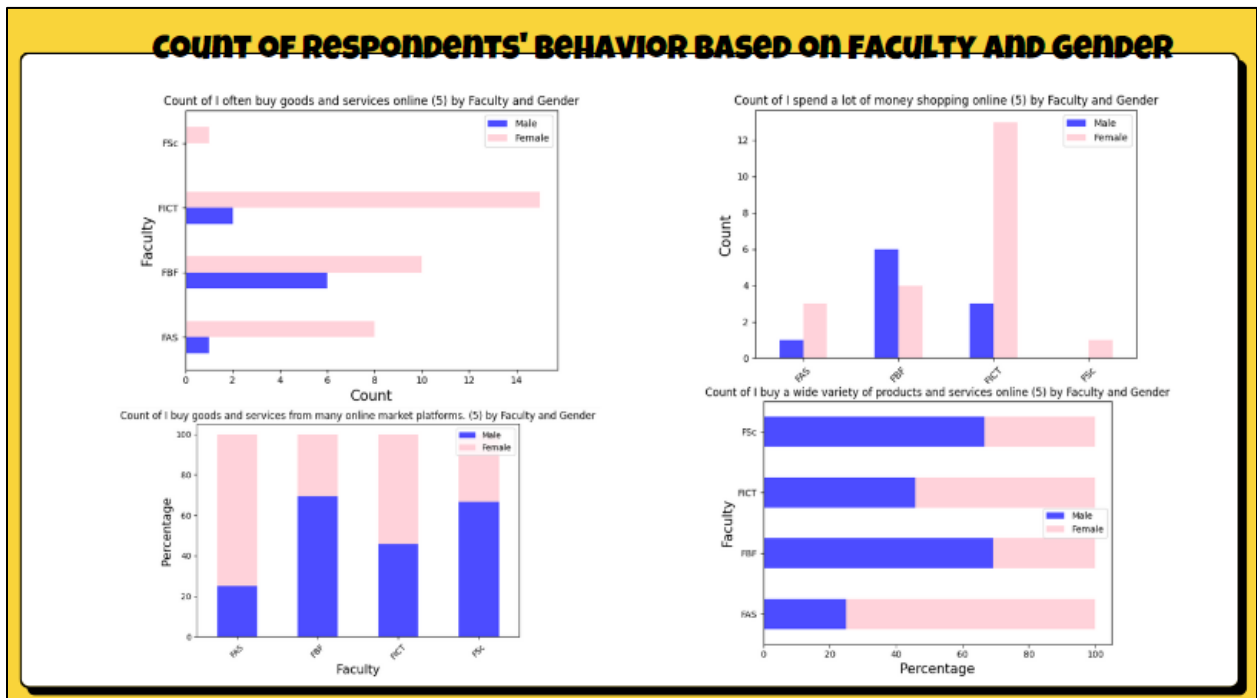


Figure 4.7. 7 Dashboard overview Behavior using Python (Jupyter Notebook)

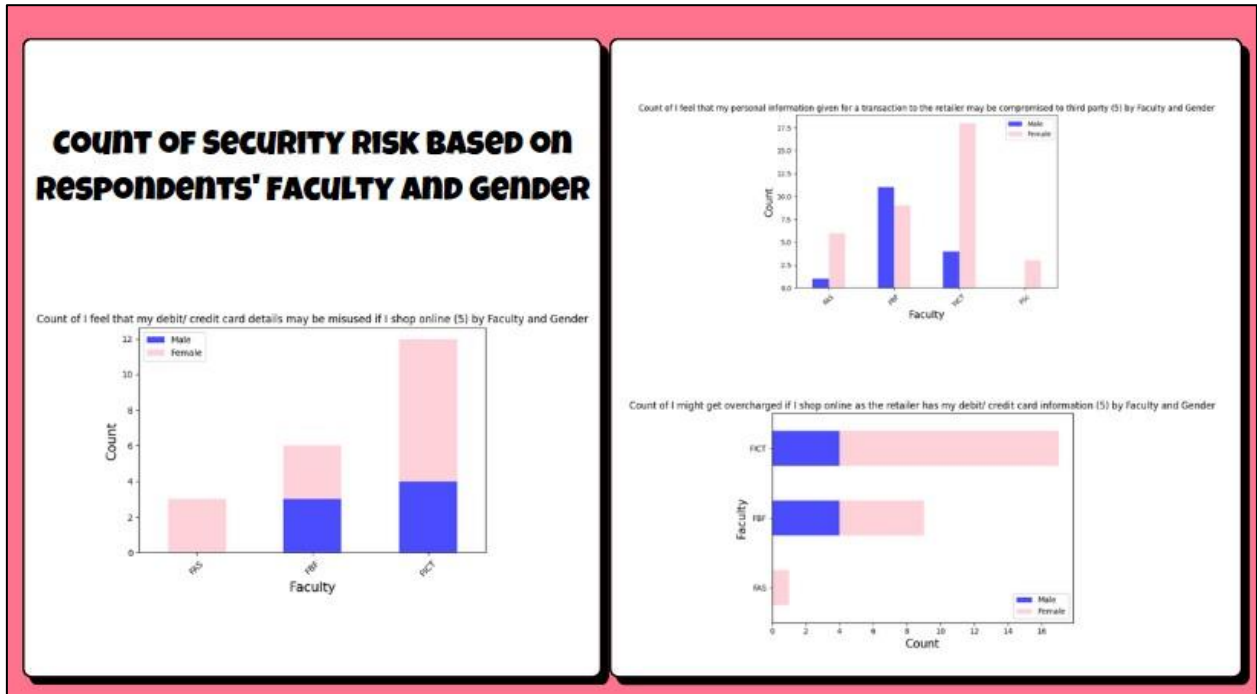


Figure 4.7. 8 Dashboard overview Security Risk using Python (Jupyter Notebook)

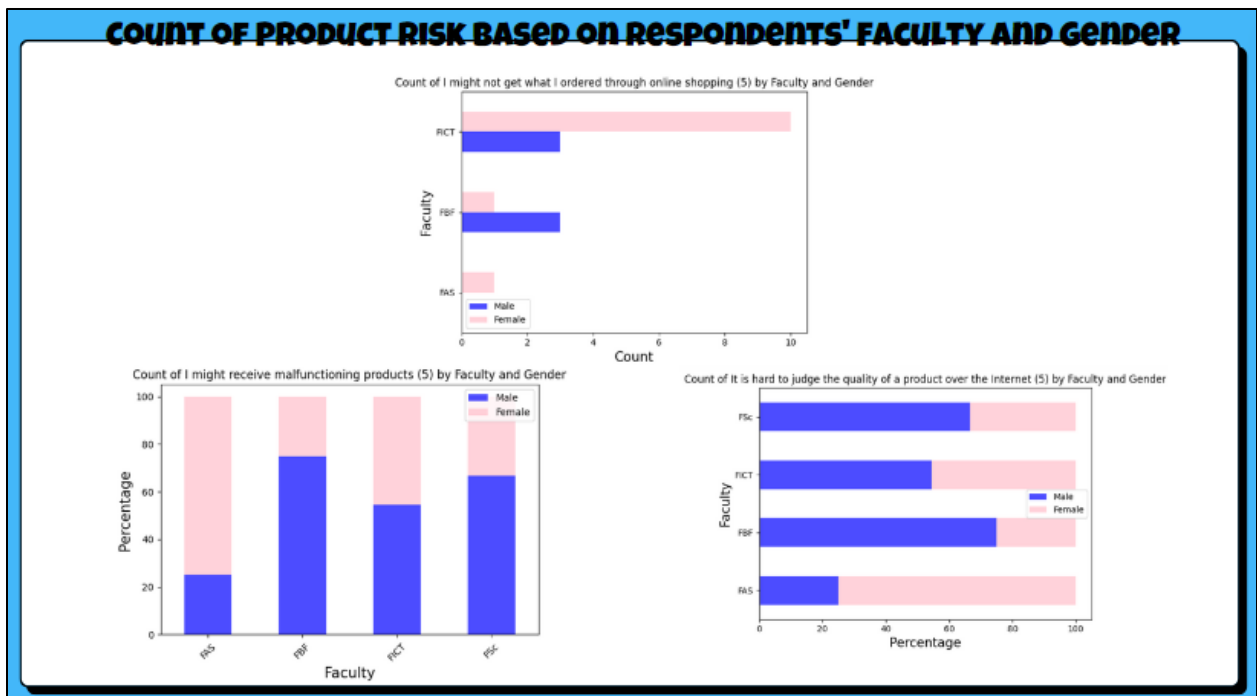


Figure 4.7. 9 Dashboard overview Product Risk using Python (Jupyter Notebook)

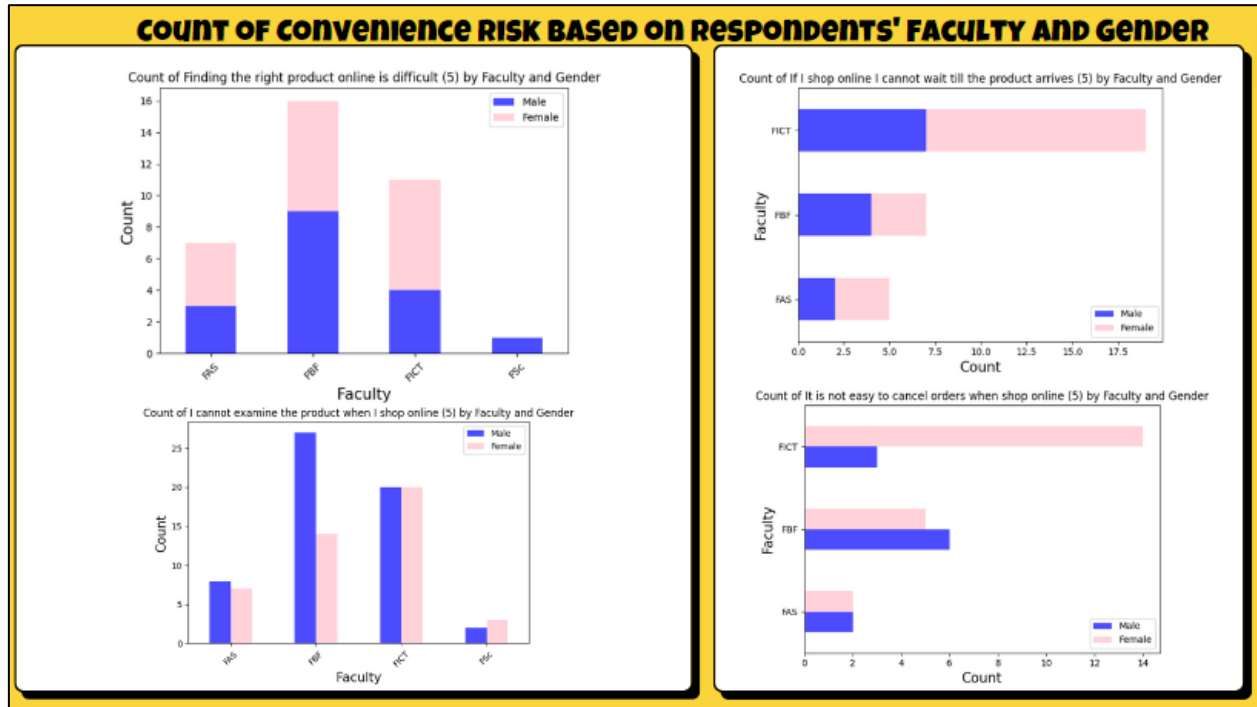


Figure 4.7. 10 Dashboard overview Convenience Risk using Python (Jupyter Notebook)

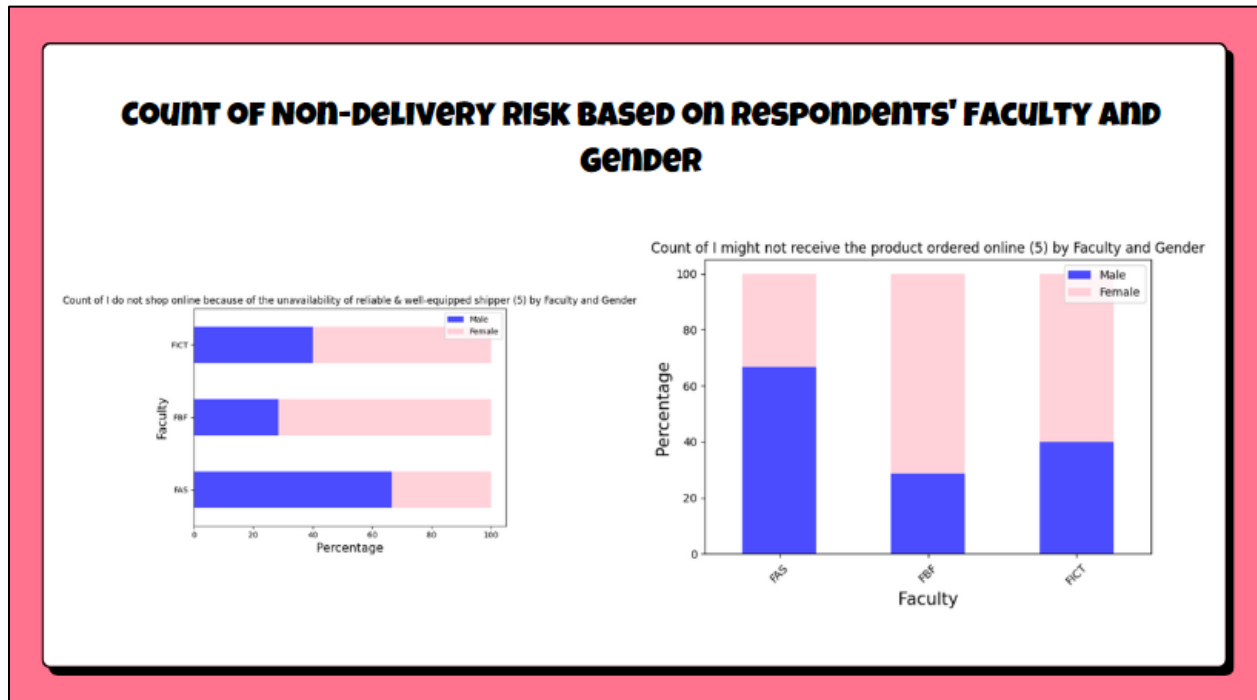


Figure 4.7. 11 Dashboard overview Non-Delivery Risk using Python (Jupyter Notebook)

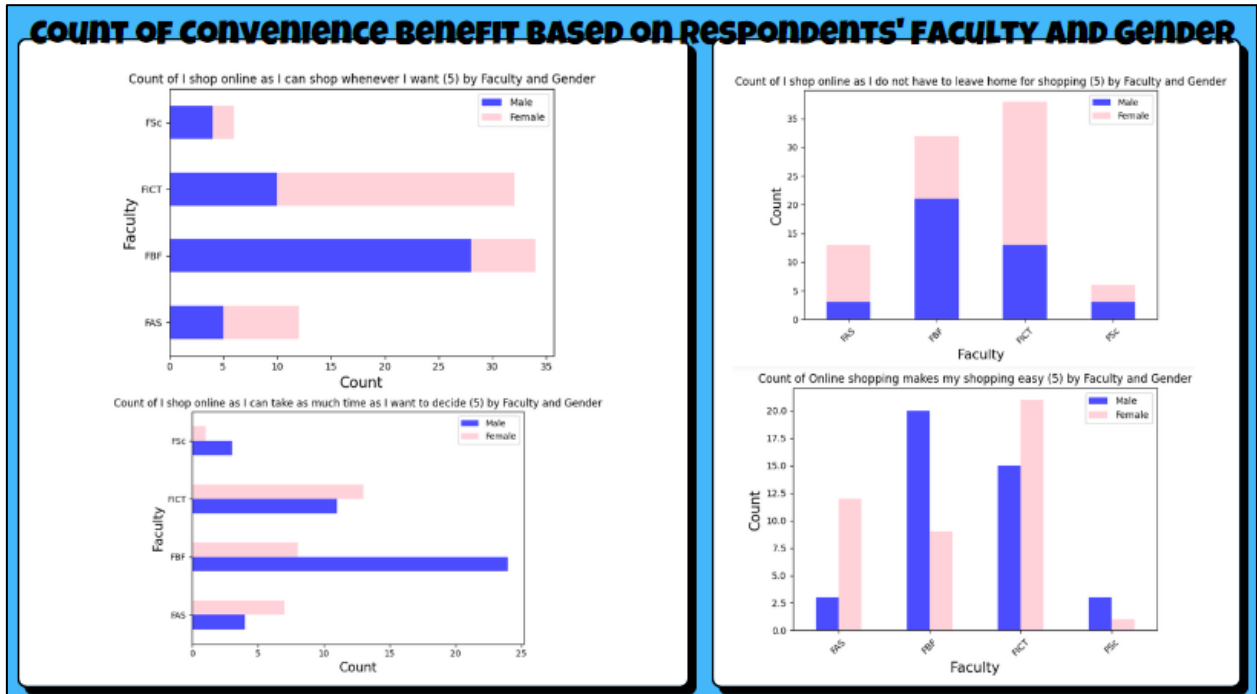


Figure 4.7. 12 Dashboard overview Convenience Benefit using Python (Jupyter Notebook)

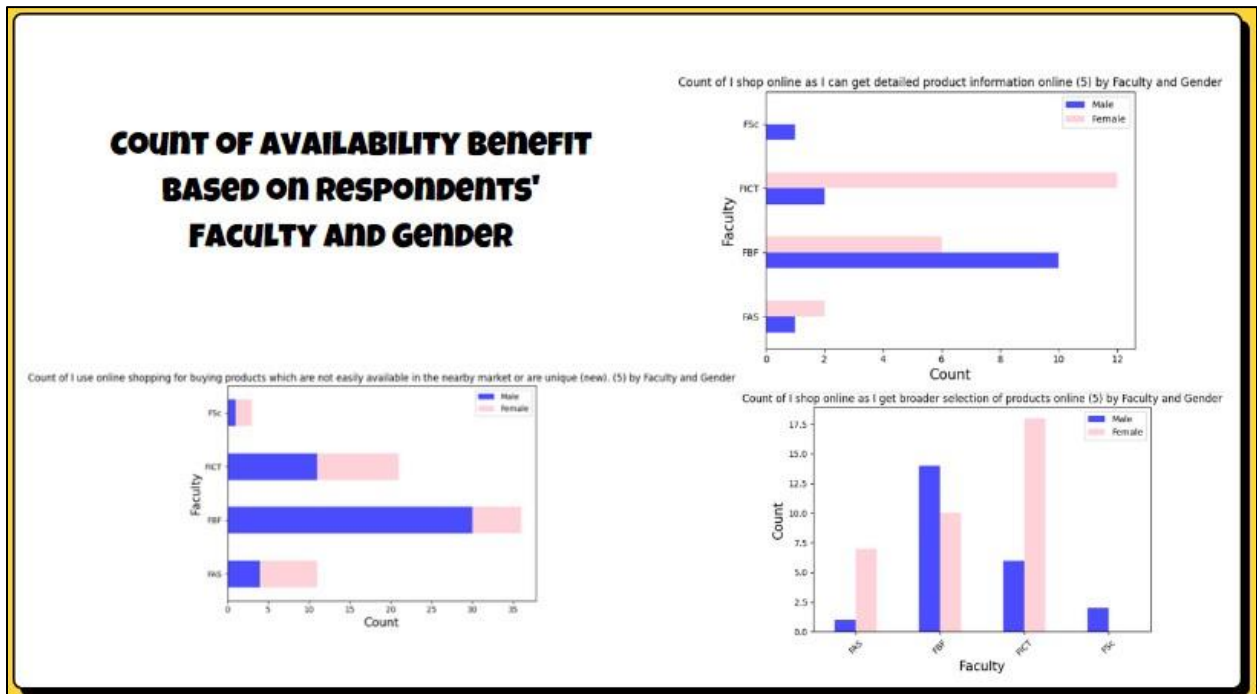


Figure 4.7. 13 Dashboard overview Availability Benefit using Python (Jupyter Notebook)

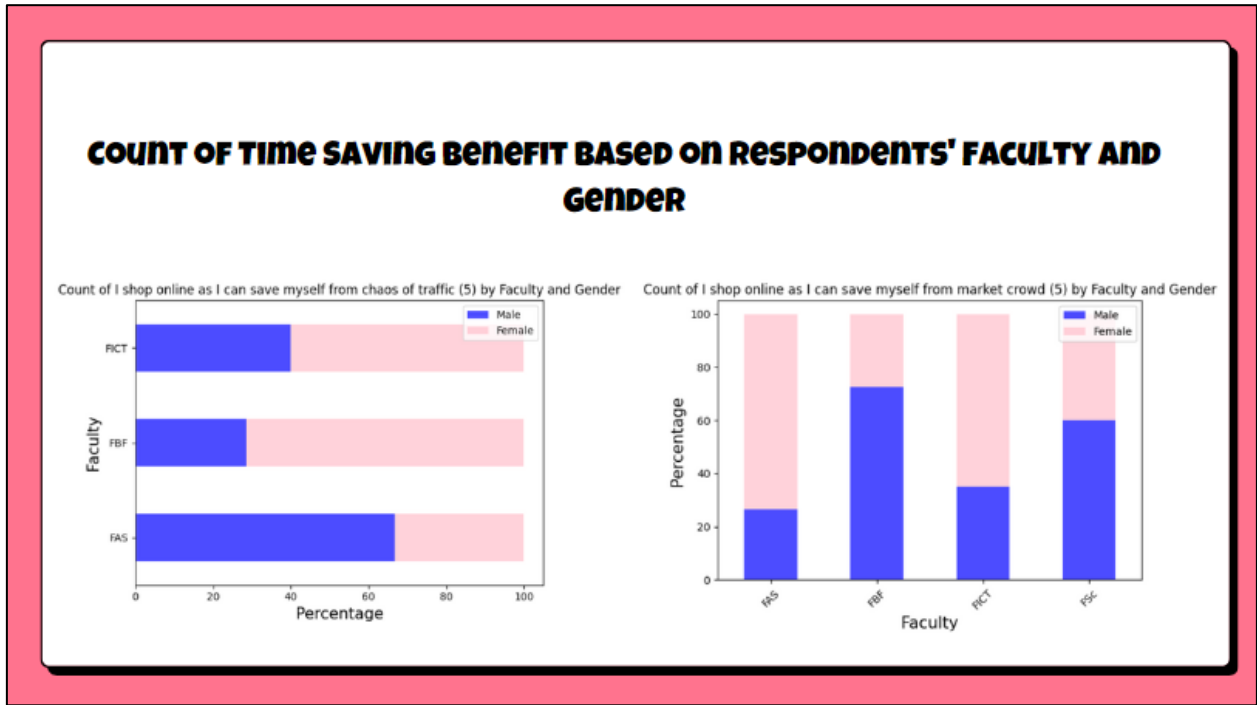


Figure 4.7. 14 Dashboard overview Time Saving Benefit using Python (Jupyter Notebook)

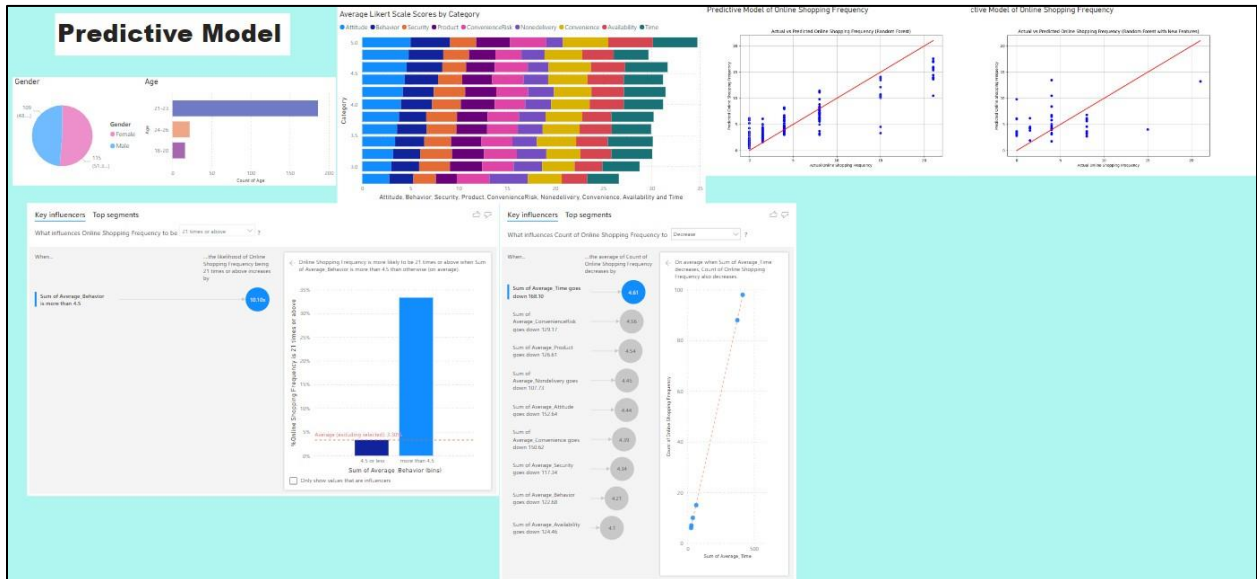


Figure 4.7. 15 Dashboard overview of Predictive Model (Power BI)

4.8 Summary of this Chapter

This chapter discussed the Experiment/Simulation, and the next chapter continued with the Evaluation and Discussion.

Chapter 5

5.1 Evaluation and Discussion

The previous chapter covered the Experiment or Simulation, which included the following: Setting up, Visualization and segmentation, Timeline, Data collection and preparation, Checking the null value, duplicate value, and visualizing the results as dashboard using Python and Jupyter. Thus, this Evaluation and Discussion chapter discussed all the dashboard that had been developed.

5.1.1 Analysis of Figure 4.7.6 (Respondents' Attitude)

Based on the faculty and gender of the respondents, the figure offered insights into their perspectives regarding online shopping. Numerous trends were shown by the data. In general, female respondents from FICT, FBF, and FAS had a more positive attitude towards online shopping, particularly regarding how acceptable, enjoyable, and pleasant it was. More specifically, in the FICT faculty, males viewed online shopping as more efficient and practical, while females found it more agreeable, entertaining, and pleasant. Similarly, both genders in the FBF faculty regarded online shopping as efficient and enjoyable, with females marginally favoring its agreeableness and enjoyment more. Across a wide range of indicators, females at FAS consistently held a more positive opinion of online shopping.

Male respondents viewed online shopping as more practical and efficient, especially those from FICT and FBF faculties. However, when it came to how enjoyable and satisfactory the experience was, their opinions differed, with females typically viewing it more favorably. The response rate from the FSc faculty was low, which made it difficult to reach firm conclusions. However, the available information indicated that while male viewed online shopping as more effective and practical, females found it more acceptable and pleasurable.

In conclusion, the dashboard showed that attitudes regarding online shopping varied according on gender, with males placing a higher importance on its efficiency and practicality, while female typically expressed a more positive perspective in terms of enjoyment, agreeableness, and pleasantness. These attitudes were also significantly shaped by faculty-

specific trends, emphasizing the importance of considering both demographic factors when attempting to explain consumer behaviour.

5.1.2 Analysis of Figure 4.7.7 (Respondents' Behavior)

The dashboard contained faculty- and gender- specific information on respondents' online shopping behaviors. Compared to both gender in FICT faculty, females were more likely to purchase a variety of goods and services from several online platforms. In contrast, FICT males spent less and used the Internet less. FBF faculty mirrored this trend, with males reporting higher spending but females showing a wider range of purchasing behavior.

Female respondents in FAS faculty showed greater engagement in the range of goods and services purchased and the frequency of online shopping. Notably, despite high engagement, gender differences in consumer behavior persisted. The data showed that for FSc faculty, which had the fewest respondents, online shopping behavior was evenly distributed between genders. Male and female buyers showed similar trends in frequency and variety.

Taken together, the evidence suggested that while males tended to spend more money online, females were more likely to shop online frequently and purchase a wider variety of goods and services. When studying consumer behavior in an academic setting, it was important to consider the characteristics and demographics of specific departments. These insights highlighted gender-specific patterns in online shopping behavior across faculties.

5.1.3 Analysis of Figure 4.7.8 (Security Risk)

The dashboard displayed respondents' security concerns about online purchasing by faculty and gender. The information showed that various groups' views on security followed distinct patterns. Compared to males, female respondents in FICT faculty were more concerned about credit or debit card information being misused. Their concerns about the compromise of personal information during transactions also followed this trend. Furthermore, FICT females were more concerned about being overcharged when businesses obtained their credit card information.

Within the FBF faculty, worries about security risks were evident in both genders; however, females seemed to be slightly more concerned about the compromise of personal information

and being overcharged. Male respondents expressed significant anxiety about information security and card details being misused, although they were less concerned than female respondents in other areas.

In this dashboard, FAS had fewer overall worries than those at FICT and FBF. Concerns about security dangers were moderately expressed by female respondents and even less so by male respondents. With the fewest replies, the FSc faculty showed modest security concerns for both genders, suggesting either a lower perceived risk or a small level of online purchasing activity.

Overall, the data pointed to a greater degree of security concerns among female respondents in all faculties, but especially in FICT and FBF. These worries included worries included potential overcharging, personal information compromise, and card detail abuse. Even though they were less concerned overall, male respondents still showed a noticeable level of anxiety, particularly in FBF. Understanding these security issues was essential for removing obstacles to online buying and improving customer satisfaction through increased security and trust.

5.1.4 Analysis of Figure 4.7.9 (Product Risk)

The dashboard displayed faculty- and gender-specific product risk concerns regarding online buying. The information showed that various groups' perceptions of product risk followed distinct patterns.

Both male and female respondents in FICT faculty voiced serious worries about product risks. Male respondents were equally concerned about the challenge of determining product quality online, while female respondents were especially anxious about receiving defective products and not getting what they requested. In the FBF faculty, significant concerns about product risks were displayed, but males are more concerned about receiving defective products and finding it difficult to assess product quality. Female respondents also expressed significant fear about these risks, along with a supplementary concern about not receiving what they ordered.

When comparing FAS to FICT and FBF, the data showed a decrease in overall concerns. While male respondents expressed similar concerns, though at a lesser level, female

respondents exhibited moderate levels of worry about product quality and receiving what they ordered. In the FSc faculty, concerns about product risk were minimal for both genders, suggesting either a lower perceived risk or a less online purchasing activity.

Overall, the data indicates that respondents, both male and female, had strong concerns about product risks, particularly in FICT and FBF. These worries include the possibility of ordering the wrong item, receiving defective products, and having trouble determining the quality of products online. Understanding these concerns was essential for removing barriers to online purchasing and boosting customer confidence.

5.1.5 Analysis of Figure 4.7.10 (Convenience Risk)

The dashboard displayed respondents' convenience risk concerns about online shopping by faculty and gender. The information showed that various groups' views on convenience risks followed distinct patterns.

Regarding convenience risks, both male and female respondents in the FBF faculty voiced serious concerns. Males are especially worried about not being able to inspect products before buying them and having trouble finding the perfect product online. Female respondents were also worried about these issues and found it difficult to cancel orders. FICT faculty respondents displayed significant concerns about convenience risks, with males being more concerned about not wanting to wait for a product to arrive and having trouble selecting the ideal goods online. Female respondents also expressed significant anxiety about these risks and further concern about unreliable shipping services.

When comparing FAS to FICT and FBF, the data showed a decrease in overall concerns. While male respondents expressed similar concerns, though at lower levels, female respondents exhibited moderate levels of fear about convenience issues. With fewer replies, the FSc faculty showed both genders to be less concerned about convenience risks, suggesting a decreased perceived risk or a reduced online shopping activity.

Overall, convenience risk concerns were evident among both male and female respondents, particularly in FBF and FICT, according to the data. These concerns included inability to inspect merchandise, difficulties in finding the correct product, and trouble cancelling orders.

Addressing these concerns and improving the user experience through increased convenience and reliability would require understanding these worries.

5.1.6 Analysis of Figure 4.7.11 (Non-Delivery Risk)

The dashboard displayed respondents' concerns about non-delivery risk when they shop online, broken down by faculty and gender. Different groups' views of non-delivery risk showed unique patterns in the data.

Significant concerns over non-delivery risks were expressed by both male and female respondents in the FBF faculty. The primary concern among male respondents was the lack of reputable and adequately prepared shippers, whilst female respondents were more worried about not receiving the merchandise ordered online. Regarding non-delivery risks, both male and female FICT faculty were highly concerned, with males being more concerned about the lack of reliable shippers. Female respondents also expressed strong concerns about these dangers, with one prominent concern being non-delivery of items ordered online. Comparing FAS to FBF and FICT, the data showed a decrease in overall concerns. While male respondents expressed similar concerns, albeit to a lesser degree, female respondents showed moderate concerns about non-delivery risks. Additionally, FSc had no such concerns.

In summary, the data indicated that respondents, both male and female, had significant concerns about non-delivery risks, especially in FBF and FICT. These concerns ranged from fear of not receiving ordered items to fear of not having a trustworthy shipper. By understanding these concerns was critical to removing barriers to online shopping and improve the customer experience by increasing delivery reliability and confidence.

5.1.7 Analysis of Figure 4.7.12 (Convenience Benefit)

The dashboard offered faculty - and gender-specific insights regarding the convenience benefit of online shopping as experienced by users. The information revealed several patterns in the opinion of respondents of all genders and faculties regarding the ease of online shopping.

Most female respondents from FBF stated that they bought online because it gave them the freedom to shop whenever they want and the ability to take as much time as necessary to make decisions. Males exhibited this tendency as well, albeit to a low degree. The notion that convenience was a key consideration for female respondents from FBF was supported by the fact that they found online shopping to be substantially easier. The tendency was fairly similar for the FICT, as both male and female respondents expressed a high value on the flexibility to purchase whenever it was convenient for them. Male respondents, however, were more likely to say that they bought online because it saved them from having to leave their houses; this suggested that males in FICT were motivated to avoid physical businesses. In FICT, both genders agreed that shopping online was easier, albeit women tended to say this slightly more.

Although the percentages were smaller than in FBF and FICT, FAS exhibited a balanced appreciation for the convenience of online shopping across genders. FAS demonstrated that male marginally appreciated being able to shop from home more than women do. Next, FSc is exhibited less noticeable trends, despite having the fewest responders. The convenience of not having to leave house to shopping was valued by both genders, but generally online shopping for convenience was less popular than it was in other academic fields.

According to the findings, ease was a major factor in online shopping behavior across all faculties, albeit male and female perceptions of these benefits differed noticeably. Male respondents placed more weight on the opportunity to avoid physical establishment, while female respondents generally valued the flexibility and ease of online shopping more. These findings highlighted how crucial convenience was while shopping online, especially for female students in FBF and FICT faculties.

5.1.8 Analysis of Figure 4.7.13 (Availability Benefit)

Based on respondents' gender and faculty, the dashboard analyzed the benefits of online shopping as seen by availability. The analysis showed significant patterns in the preferences of respondents from various categories for a wider range of products and online access to comprehensive product details.

Both male and female respondents from FBF strongly valued the wider range of products that were offered online. Particularly, female valued this benefit more than male did; a greater percentage of them agreed that they shopped online is for this reason. FBF also demonstrated that both genders placed a high importance on being able to locate uncommon or hard-to-find things, with male expressing a marginally higher preference. Both genders benefited greatly from the comprehensive product information readily available online, albeit female tended to mention this more. The tendency was similar in FICT, with both genders appreciating the availability of innovative products and a wider assortment. In contrast to females, males in FICT had a greater preference for comprehensive product information. This suggested that although both genders valued the accessibility advantages, males might have prioritized information more when making online purchases. For FSc, who had the fewest responses, there was the least amount of interest in the accessibility advantages of e-commerce. All metrics showed low counts for both genders, suggesting either a decreased demand for these benefits or a general decline in online shopping activity.

Overall, the evidence pointed to the benefits of availability, such as access to unusual items, a larger selection of products, and through product information as important motivators for online shopping in all faculties. Male respondents also expressed considerable gratitude, particularly for comprehensive product information, but female respondents, notably those in FBF and FAS, placed a high value on these perks. Comprehending these inclinations could facilitate customization of virtual retail encounters to satisfy the demands of diverse demographic cohorts more effectively.

5.1.9 Analysis of Figure 4.7.14 (Time Saving Benefit)

The dashboard examined the advantages of online shopping in terms of time saving, broken down by the respondents' gender and faculty. The information shed light on why people chose online shopping over traditional retail, which could be time-consuming due to factors like traffic and crowded stores.

Online shopping was highly valued by both male and female respondents in the FICT because it allowed them to avoid the traffic chaos. The number almost matched for both genders, suggesting that they both valued this perk equally. Furthermore, the advantage of avoiding

market crowds was valued similarly by both genders in FICT, indicating that time saving features were important drivers of their online shopping behavior. Similar findings were seen in the FBF, where both genders respondents valued being able to avoid traffic and saving time. But the evidence suggested that male marginally valued this benefit more than female do. The gender distribution was nearly equal when it comes to avoiding market crowds, indicating that these advantages were equally significant to all genders in FBF.

The time saving advantages were also noticeable for FAS respondents, though they were not as pronounced as they were in FICT and FBF. In FAS, female placed a somewhat higher priority on avoiding traffic than do male, while both genders placed an equal value on avoiding crowded markets. This suggested that FAS respondents had a moderate appreciation for the advantages of time saving. Next is FSc, which had the fewest responses, exhibited the least interest in the time saving advantages of online shopping. The low figures for both genders suggested that either these benefits were not as important, or that there was less overall online shopping in FSc.

Overall, the data showed that a major factor driving online shopping across faculties, especially in FICT and FBF, was time savings from avoiding transportation and market congestion. The importance of efficiency and ease in their shopping choices was highlighted by the nearby identical valued placed on these benefits by male and female respondents in these faculties. Understanding these incentives could facilitate customization of e-commerce platforms to cater to the time-constrained requirements of diverse consumer segments more effectively.

5.1.10 Analysis of Figure 4.7.15 (Predictive Model)

The figure depicted a detailed study of a predictive model for online purchasing frequency that included demographic, behavioral, and attitudinal variables. The top left area indicated that most responses were from female, with the majority falling between the ages of 21 and 23. The Likert scale chart in the center assessed average scores in numerous categories, including attitude, behavior, security, and convenience, to determine how respondents perceived these aspects of online buying. The Random Forest model was featured in the upper right area of prediction models. The model aimed to forecast online purchasing frequency, using scatter plots comparing real and anticipated values and a red line demonstrating the model's fit. The comparison showed which model better predicted internet shopping patterns.

The Key Influencers section, which appeared at the bottom, revealed crucial elements influencing the frequency of online purchase. It demonstrated that having a greater "Sum of Average Behavior" considerably increased the likelihood of shopping online. Furthermore, the "Sum of Average Time" was a significant predictor, with a resulting in reduced shopping frequency. The study concluded that demographic characteristics, together with certain behaviors and attitudes, had an important influence in predicting how frequently people shop online. The visualization presented a complete overview of the elements that influenced online buying behavior, as well as the success of various predictive models in comprehending these patterns.

5.2 Discussion of Results

Based on the evaluation and analysis of the predictive model dashboard, the results showed that some elements had a substantial influence on UTAR students' online buying behavior. The demographic study suggested that most respondents were males between the ages of 21 and 23, implying that these segments may have had a significant impact on general online shopping patterns.

Likert scales provided an organized method for academics to collect varied ideas and attitudes. They allowed respondents to indicate whether they agreed, disagreed, or were neutral on statements or questions. A Likert scale was a rating system for evaluating survey respondents' opinions, attitudes, motivations, and other qualities. It employed several answer possibilities that ranged from one extreme attitude to the next, with the rare moderate or neutral option thrown in. However, the most usual scales ranged from four to seven points. Likert scales (named after their founder, American social scientist Rensis Likert) were widely used because they were one of the most accurate ways to examine attitudes, perceptions, and behaviors. Likert scales were simple to build and could be used for a wide range of purposes, including quantitative surveys. Unlike binary questions, which only offered two possible answers, Likert-type questions gave more specific feedback on whether your product was "good enough" or (preferably) "excellent." This strategy allowed differences of opinion to emerge, helping researchers better understand feedback. It could also help identify areas for improvement in the customer or employee experience. Using the same Likert scales in longitudinal or recurrent surveys can help track seasonal variations in customer views. For example, quarterly survey could gauge customer loyalty to a business. With this information,

researchers could track quarterly and annual changes in client loyalty, understanding the brand's current and past health and predicting how it might trend in the future [25]. The Likert Scale analysis provided insight into student perspectives, with higher ratings in categories such as convenience, time and product satisfaction indicating that these were important drivers of online shopping frequency.

The Key Influencers visual in Microsoft Power BI was a useful tool that helped users identify and understand the factors that had a substantial impact on a given measure or outcome in their data. It had two key features: the Key Influencers page, which ranked and showed factors based on their impact on the selected outcome, and the Top Segments tab, which found different groups within the data that differed significantly in relation to the outcome. This graphic was very interactive, allowing users to investigate how changes in many parameters affected the likelihood of a result, making it easier to extract actionable insights without the need for complex statistical models. The Key Influencers graphic was especially useful for data analysts and business users because it allowed them to quickly surface key insights, analyze various segments of data, and utilize the results to inform decision-making. It was frequently used in a variety of fields, including customer behavior analysis, operational efficiency, financial performance, marketing effectiveness, and healthcare results. By making advanced analytics more accessible and intuitive, the Key Influencers visual enabled businesses to make data-driven decisions based on empirical evidence and predictive analysis [26]. The Key Influencers in the Figure 4.7.15 confirmed that Average Time, Average Convenience Risk, and Average Product were the best predictors of shopping frequency. These data implied that students who prioritized time efficiency, find online shopping comfortable, and were satisfied with their purchases were more likely to shop online regularly.

Random Forest was a powerful and versatile machine learning algorithm that is widely used for classification and regression. It operated by training many decision trees and then displays the mean prediction (for regression) or mode of the classes (for classification). This ensemble technique increased the model's accuracy and robustness by reducing the possibility of overfitting, which could happen with single decision trees. Each tree in a Random Forest was trained on a distinct portion of the data, with some attributes chosen at random at each split point, resulting in various trees that contributed to more accurate and generalized predictions. The algorithm was well-known for its high accuracy, ability to handle huge datasets with higher

CHAPTER 5

dimensionality, and efficacy in determining feature importance, which aided in comprehending the significance of each variable in the prediction process [27]. The scatter graphs comparing real and anticipated frequencies using the Random Forest model showed a fair alignment between the two, although some fluctuation indicated that the model could be improved. The distribution of points around the line of perfect prediction suggested that, while the model captures the overall trend, some deviations might be addressed by fine-tuning the model or including more relevant information.

In summary, the findings showed that the predictive model efficiently identified critical characteristics impacting online shopping behavior among UTAR students. While the model predicted online shopping frequency fairly well, there was room for improvement, and it identified areas where students' attitudes and perceptions played a significant role in their online shopping decisions.

5.3 Summary of this chapter

This chapter discussed the Evaluation and Discussion, and the next chapter continued with the Conclusion and Recommendation.

Chapter 6

Conclusion and Future Recommendations

The previous chapter covered the Discussion and Recommendation, which analyzed all the dashboard that had been developed. In this final chapter, the Conclusion was discussed, and some Recommendation were provided.

6.1 Conclusion

This study sought to examine UTAR students' online shopping behaviors during the global Covid-19 pandemic, with a focus on the factors influencing their purchasing decisions via e-commerce platforms. The study found that convenience, availability, and time-saving benefits had a significant impact in encouraging students to shop online. These considerations were especially crucial during promotional times like Double 11, when the desire to take advantage of discounts and deals pushed students to shop online.

The study also revealed concerns that kept students from fully embracing internet buying. Students, particularly female respondents, expressed concerns about security issues such as personal information theft and the abuse of credit or debit card information. Furthermore, product-related hazards, such as receiving defective items or not receiving the purchased items at all, posed major impediments. The inability to personally inspect items before purchasing, as well as the difficulty of cancelling orders, contributed to student's cautious approach to online shopping.

To address these issues, e-commerce platforms needed to establish stronger security measures, increase product transparency, and offer flexible return policies. These enhancements would have helped to increase trust and lower the perceived risks associated with online purchasing. Furthermore, personalized marketing methods emphasizing the convenience, availability, and time-saving advantages of online buying could have boosted student engagement and satisfaction.

In conclusion, this study provided useful insight into UTAR students' online shopping behavior, highlighting both the opportunities and constraints that e-commerce platforms faced when catering to this group. By addressing the identified concerns and exploiting the factors

that motivated online purchase, e-commerce platforms could have better fulfilled the needs of university students, resulting in increased engagement and satisfaction.

6.2 Future Recommendations

Future researchers were advised to broaden the scope of study beyond UTAR students. By integrating students from other universities or even various age groups, the findings could provide a more complete picture of online shopping behaviour. This would have assisted future academics in identifying any broader tendencies that may apply to university students in general, rather than focusing just on one institution. Broader data collection could have improved the study's generalizability, making the findings more applicable to a wider range of student populations.

It was also advised that future research should include longitudinal studies to track changes in online shopping behaviour over time. A longitudinal method would have enabled researchers to monitor how external influences, such as technical improvements or changes in economic conditions, influenced student behaviour. This would have provided more detailed insights into how trends evolve, particularly during major occurrences such as the worldwide pandemic. Long-term statistics could provide a better understanding of how these characteristics affect online shopping patterns over time.

Another recommendation was to investigate the usage of additional machine learning methods to improve the prediction models generated. While this study used Random Forest, future research could compare different methods such as Gradient Boosting or Neural Networks. Testing multiple algorithms would have increased the accuracy and resilience of the predictive models, potentially leading to more accurate predictions of online shopping behaviour. This would have helped future academics determine the best model for analyzing e-commerce developments.

Furthermore, future research should investigate the impact of social media on online shopping behaviour, as it is becoming an increasingly essential factor in consumer decision-making. By focusing on the impact of social media marketing and peer influence, researchers might have investigated how social platforms influence student purchasing patterns. This

CHAPTER 6

would have provided significant insights for firms looking to improve their e-commerce strategy via social media channels, resulting in more engagement with university students.

Finally, future researchers were encouraged to incorporate psychological elements such as customer motivation and impulse buying into their investigations. Investigating the underlying psychological motives of Internet purchasing may have revealed deeper reasons that were not completely explored in this study. Incorporating psychological theories would have offered a new layer of study, providing a more comprehensive understanding of the factors influencing students' online shopping decisions.

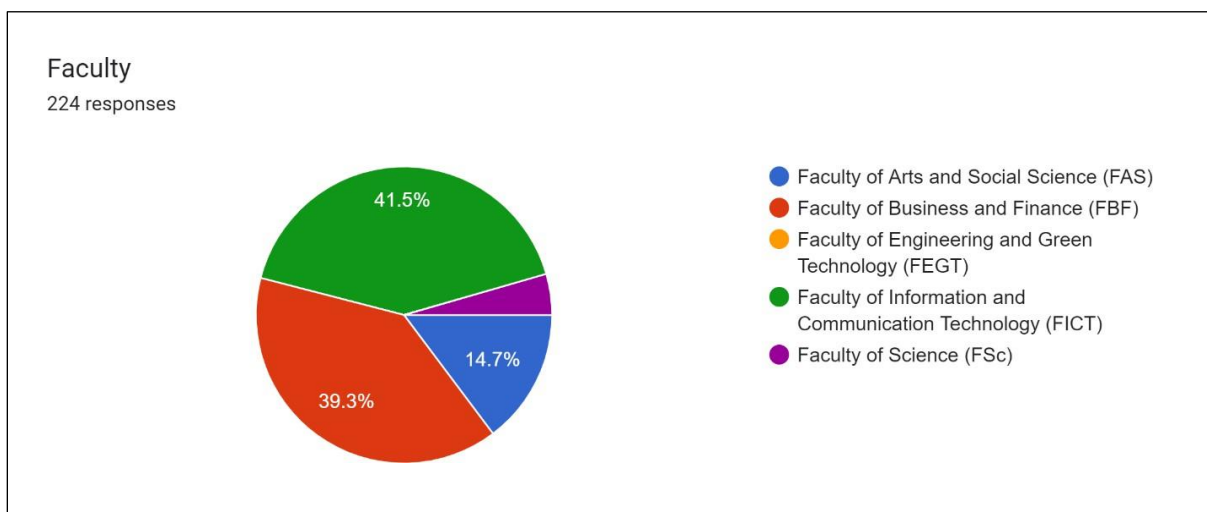
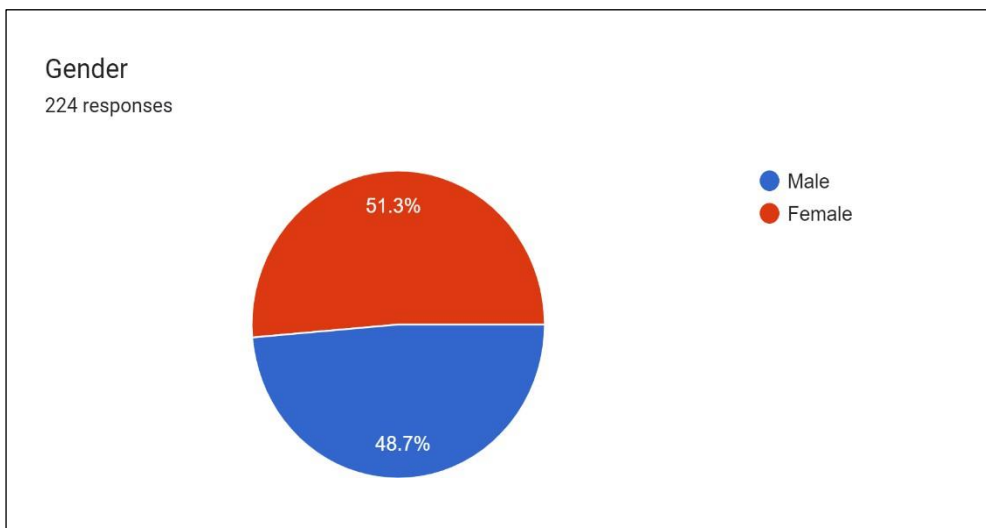
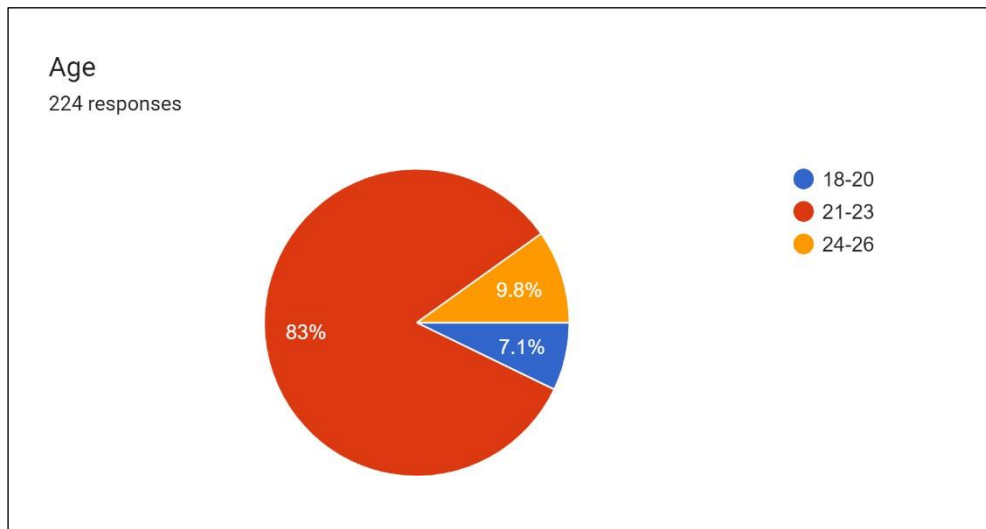
References

- [1] "“Internet and social media users in the world 2024,” Statista.," [Online]. Available: <https://www.statista.com/statistics/617136/digital-population-worldwide/>.. [Accessed 21 02 2023].
- [2] M. M. R. & S. Mudaa, "Online purchase behavior of generation Y in Malaysia: procedia economics and finance," 2016, pp. 37, 292 – 298.
- [3] A. Mosby, "100+ online shopping statistics 2023 (data & latest trends)," Yaguara," [Online]. Available: <https://www.yaguara.co/online-shopping-statistics/>.. [Accessed 05 10 2023].
- [4] " Ieee.org.," [Online]. Available: https://ieeexplore.ieee.org/stamp.jsp?arnumber=5660278&casa_token=G0eoCwB0fKYAAAAA:WeXsT20_N6vA1sw-dXTaq0YZ-cIWYLj7_p6LW_oruf33ES1jqjTGlmzxfwqVbboPPXlga7RuJ5M&tag=1 . [Accessed 04 12 2023].
- [5] "Commission Factory, “Malaysia eCommerce statistics and trends in 2023,” Commissionfactory.com," [Online]. Available: <https://blog.commissionfactory.com/ecommerce-marketing/malaysia-ecommerce-statistics..> [Accessed 09 12 2023].
- [6] S. Rose, "“How many people shop online in 2024 (latest trends & data),” GrowthDevil,," [Online]. Available: <https://growthdevil.com/how-many-people-shop-online.> . [Accessed 11 09 2023].
- [7] S. Pillay, "Online shoppers in a web of scammers," New Straits Times," [Online]. Available: <https://www.nst.com.my/news/exclusive/2017/06/245502/online-shoppers-web-scammers.> [Accessed 09 12 2023].
- [8] "Researchgate.net," [Online]. Available: https://www.researchgate.net/profile/Danijela-Tuljak-Suban/publication/282149945_Maritime_education_institutions_a_review_of_quality_assessment_methods_and_the_development_of_a_fuzzy_assessment_tool/links/5cb86fb8a6fdcc1d499de4bc/Maritime-education-institut. [Accessed 04 12 2023].
- [9] A. S. Zolkafly, "The dominant of values and lifestyles (VALS) between generations in Malaysia," Edu.my.," [Online]. Available: <https://ir.uitm.edu.my/id/eprint/55231/1/55231.pdf.> . [Accessed 02 01 2024].
- [10] Y. & B. L. R. Seock, "The influence of college students shopping orientations and gender differences on online information searches and purchase behaviours", International Journal of Consumer Studies," 2008, pp. 113-121.
- [11] A. Manzoor, "E-Commerce: An Introduction.," LAP Lambert Academic Publishing, 2010.
- [12] N. M. D. a. M. Z. Jamaluddin, "BUILDING A TRUSTER ENVIRONMENT FOR E-BUSINESS : A MALAYSIAN PERSPECTXVE," Edu.my.," [Online]. Available: https://repo.uum.edu.my/id/eprint/376/1/N.M._Din.pdf. [Accessed 04 12 2023].
- [13] H. M. e. al., "Psu.edu.," [Online]. Available: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e9ed56e5c28117d4083fcf867e5cd0ad26b71561#page=154.> [Accessed 04 12 2023].

REFERENCES

- [14] "“Online shopping problems and solutions,” CORE Reader,," [Online]. Available: <https://core.ac.uk/reader/234652510>. [Accessed 04 12 2023].
- [15] T. B. a. M. Bhatt, "“A study of intrinsic and extrinsic factors influencing consumer buying behaviour online,”," in *Asia-Pac, J. Manag. Res. Innov.*, 2013, p. 77–90.
- [16] " Researchgate.net," [Online]. Available: https://www.researchgate.net/profile/Sneha-Mittal-5/publication/365853239_In_The_Future_of_E-Commerce_Editors_Traditional_to_Digital_Commerce_Impact_of_Scams_on_Consumer's_Attitude_towards_Online_Shopping_Complimentary_Copy/links/63872347148d2362a4b1b0ab/. [Accessed 04 12 2023].
- [17] J. K. a. H. J. C. C.-H. Cho, "“Online shopping hesitation,” *Cyberpsychol. Behav.*, 2006, p. 261–274.
- [18] B. A.-P. a. L. A. Z. S. Sabou, "“The impact of the problems faced by online customers on ecommerce,”," in *Stud. Univ. Babe-Bolyai Econ.*, 2017, p. 77–88.
- [19] "Ieee.org," [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5660278&casa_token=G0eoCwB0fKYAAAAA:WeXsT20_N6vA1sw-dXTaq0YZ-cIWYLj7_p6LW_oruf33ES1jqjTGlmzxfwqVbboPPXlga7RuJ5M&tag=1. [Accessed 04 12 2023].
- [20] S. U. R. Anam Bhatti, "PERCEIVED BENEFITS AND PERCEIVED RISKS EFFECT ON ONLINE SHOPPING BEHAVIOR WITH THE MEDIATING ROLE OF CONSUMER PURCHASE INTENTION IN PAKISTAN," *International Journal of Management Studies*, pp. 35-38, 2020.
- [21] U. Grömping, "Variable Importance Assessment in Regression: Linear Regression versus Random Forest," 11 2009. [Online]. Available: https://prof.bht-berlin.de/fileadmin/prof/groemp/downloads/tast_2E2009_2E08199.pdf. [Accessed 9 08 2024].
- [22] "What Is Data Collection: Methods, Types, Tools," [Online]. Available: <https://www.simplilearn.com/what-is-data-collection-article>. [Accessed 20 03 2024].
- [23] "What is data preparation? An in-depth guide to data prep," [Online]. Available: <https://www.techtarget.com/searchbusinessanalytics/definition/data-preparation>. [Accessed 20 03 2024].
- [24] T. Osborn, "7 Data Quality Checks in ETL Every Data Engineer Should Know," [Online]. Available: <https://www.montecarlodata.com/blog-data-quality-checks-in-etl/>. [Accessed 10 04 2024].
- [25] SurveyMonkey, "Likert Scale: What It Is & How to Use It | SurveyMonkey," 2024. [Online]. Available: <https://www.surveymonkey.com/mp/likert-scale/>. [Accessed 23 08 2024].
- [26] aphilip94, "Key influencers visualizations tutorial - Power BI," [Online]. Available: <https://learn.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-influencers?tabs=powerbi-desktop>. [Accessed 23 08 2024].
- [27] IBM, "What is Random Forest? | IBM," 2023. [Online]. Available: <https://www.ibm.com/topics/random-forest>. [Accessed 23 08 2024].

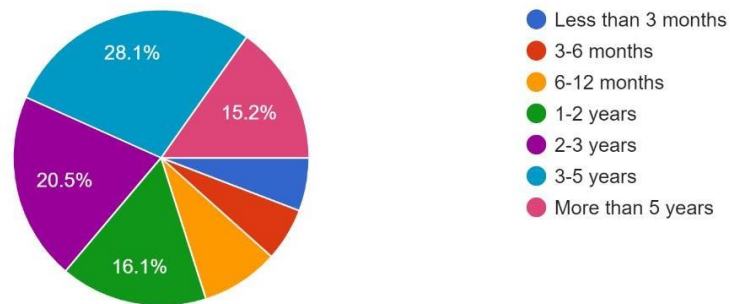
Questionnaire Data



APPENDIX

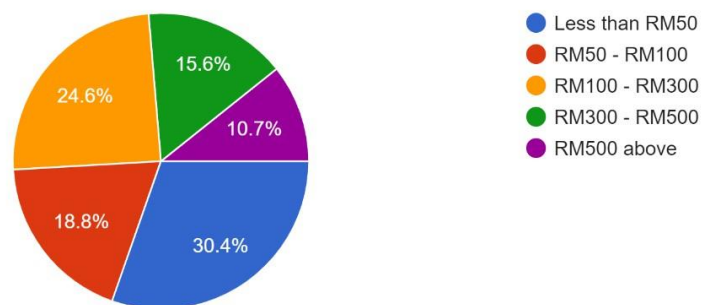
How long have u been shopping online?

224 responses



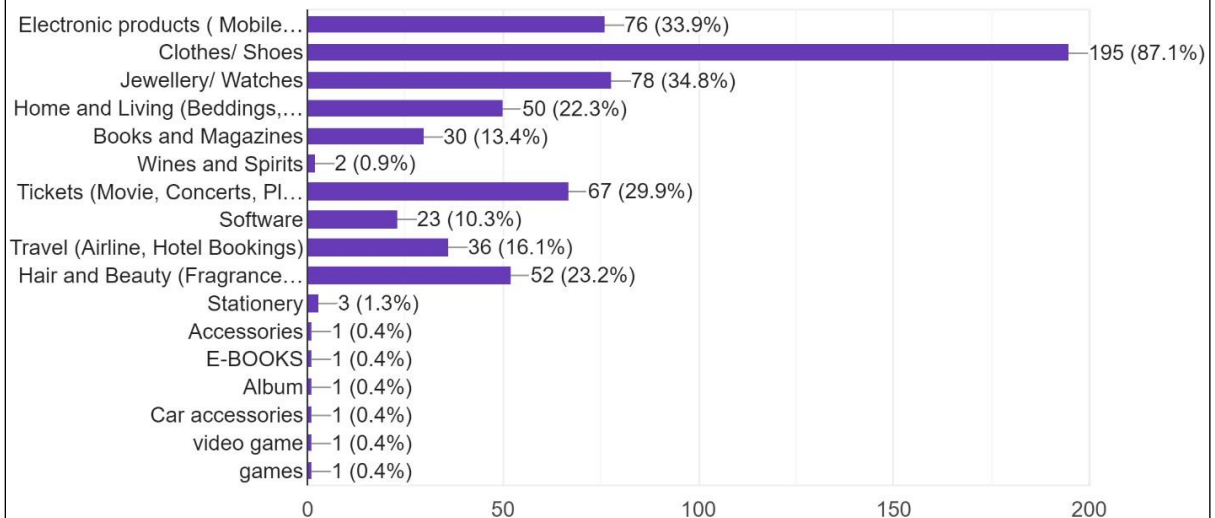
What would be your estimate of online shopping expenditure (during the past 3 months)?

224 responses

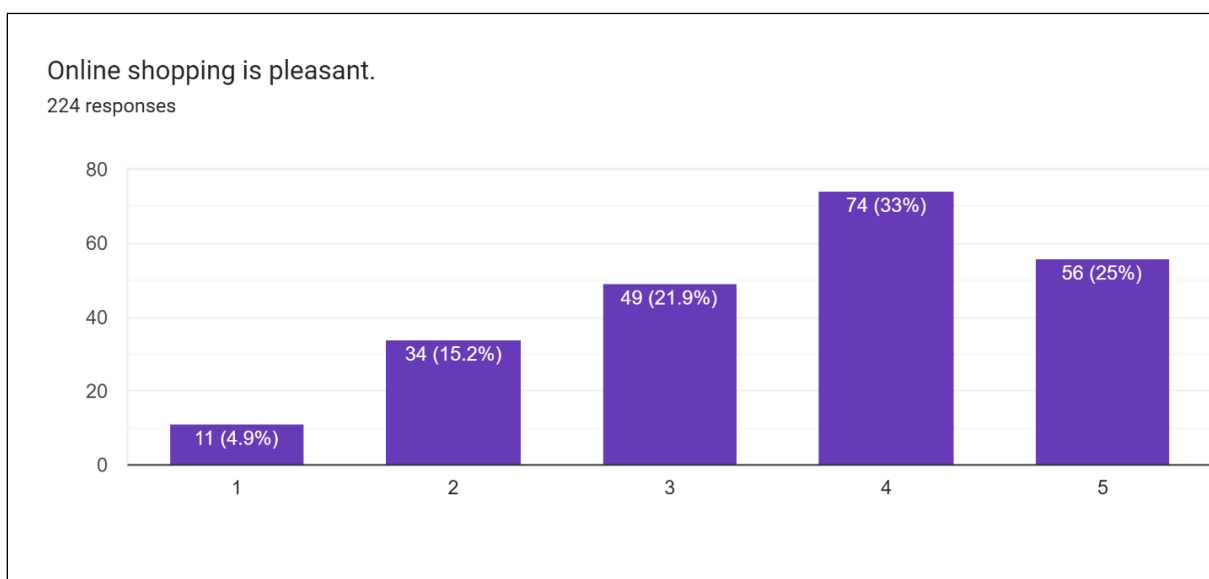
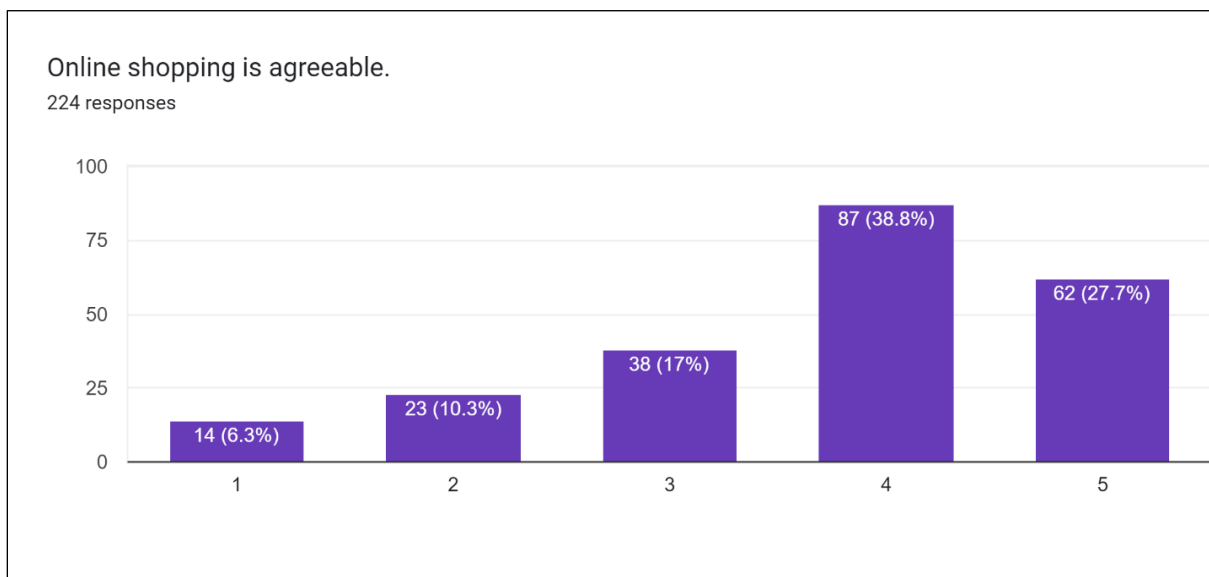
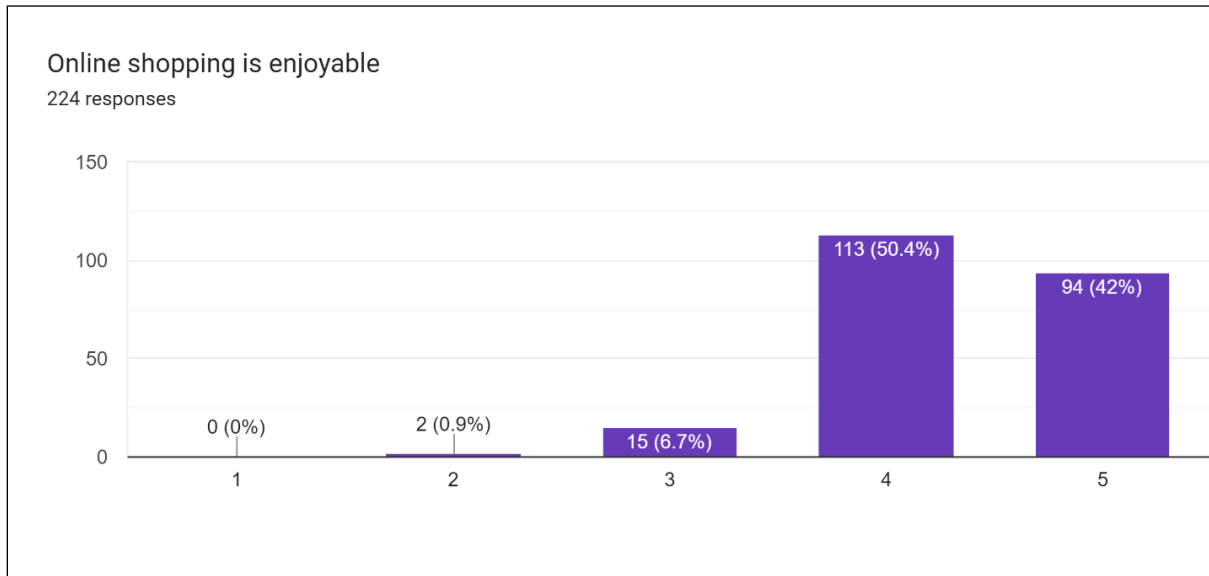


Please choose the types of products that you purchased online (can be more than one type).

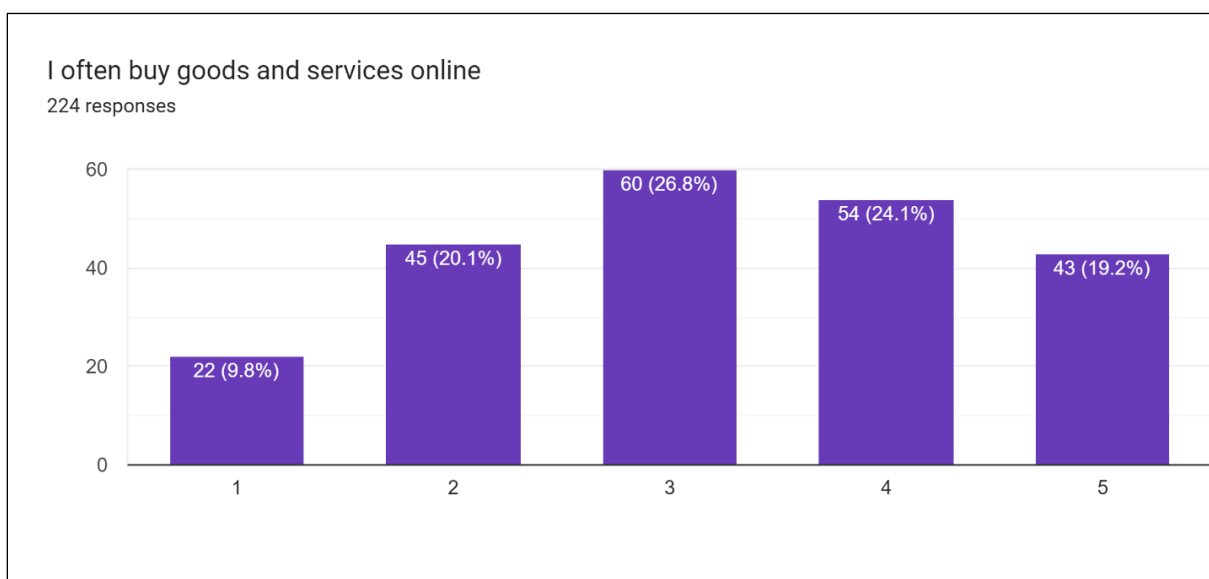
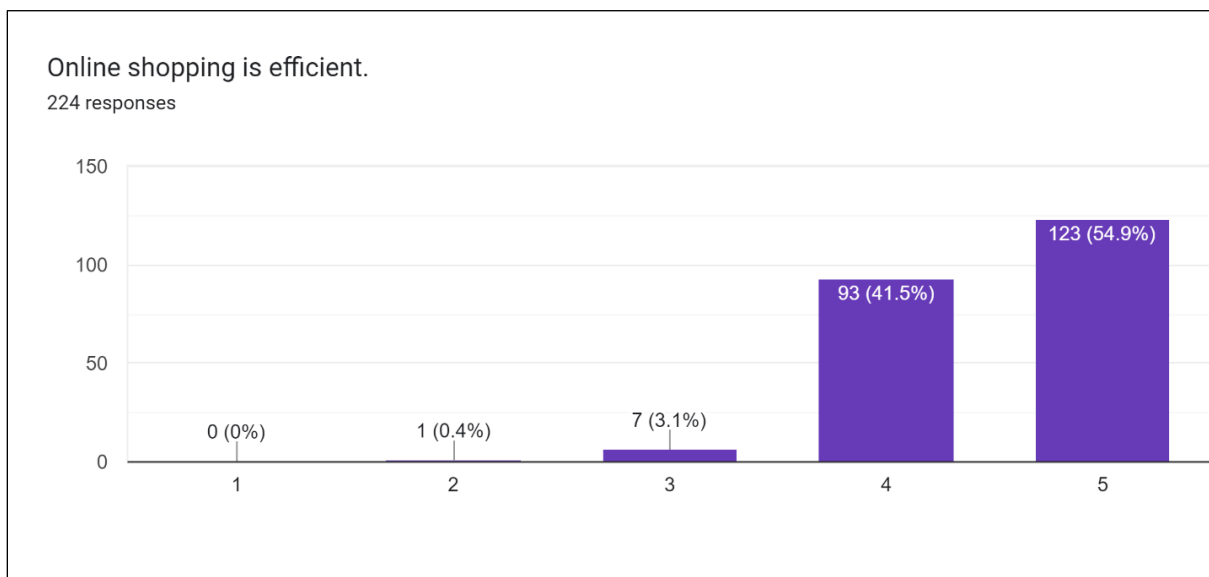
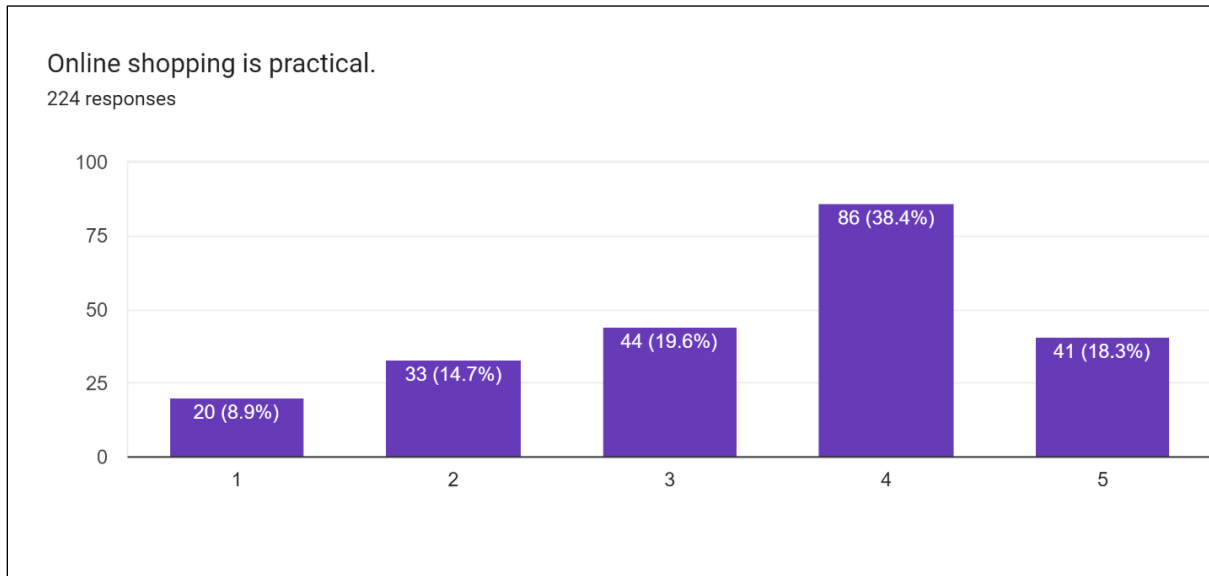
224 responses



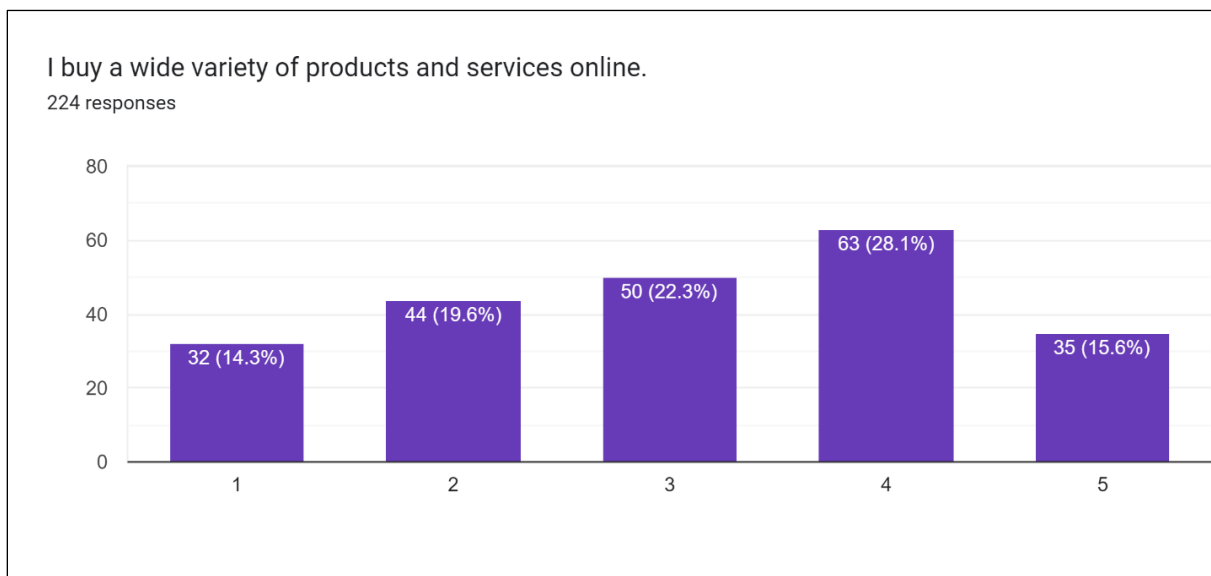
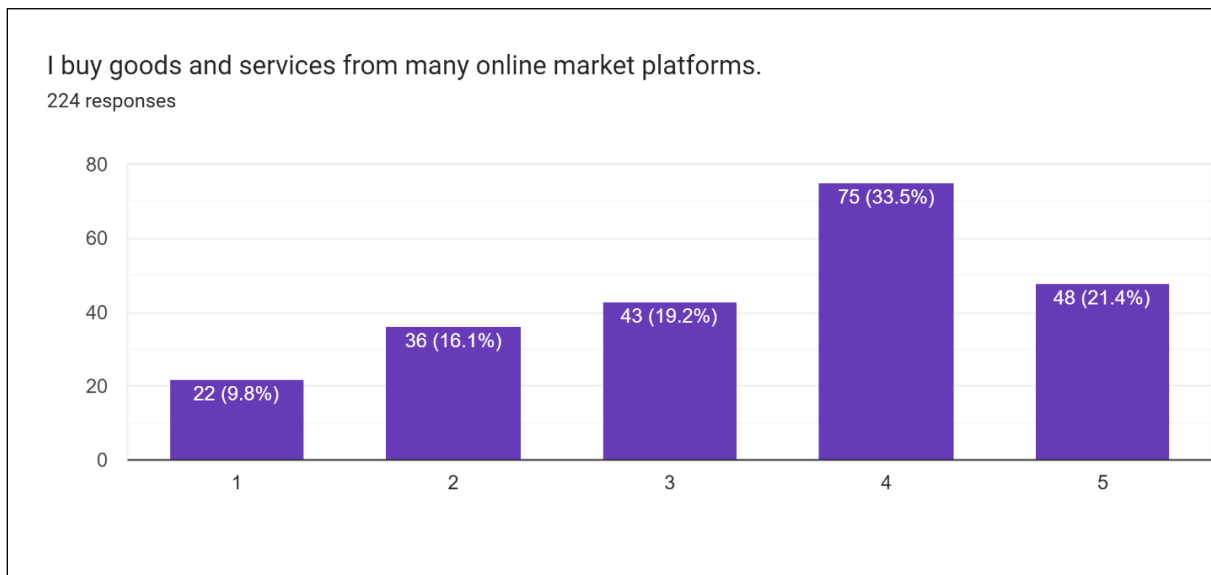
APPENDIX



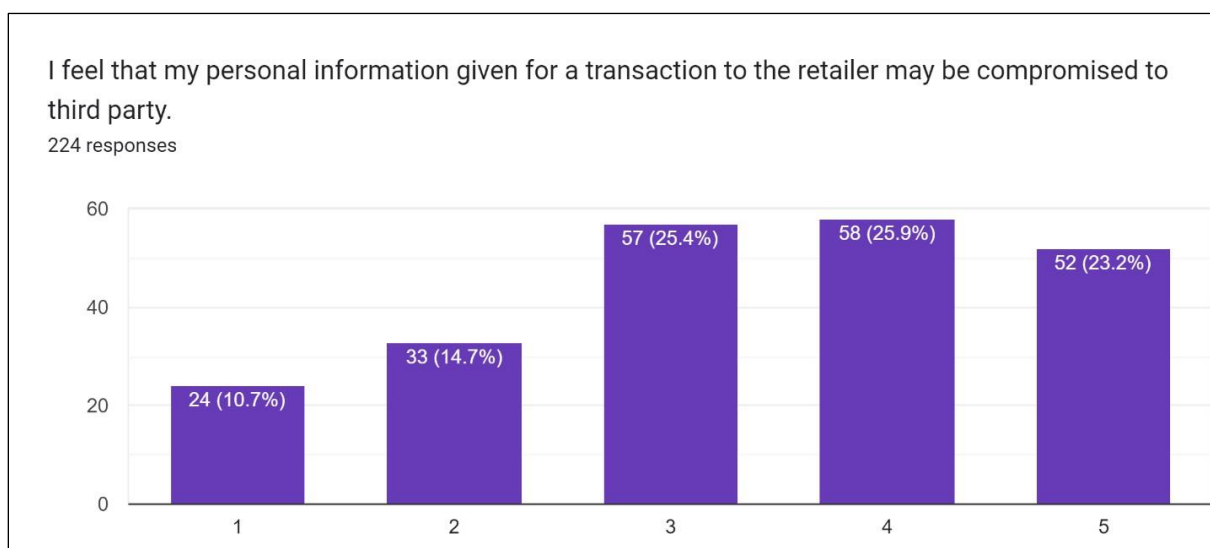
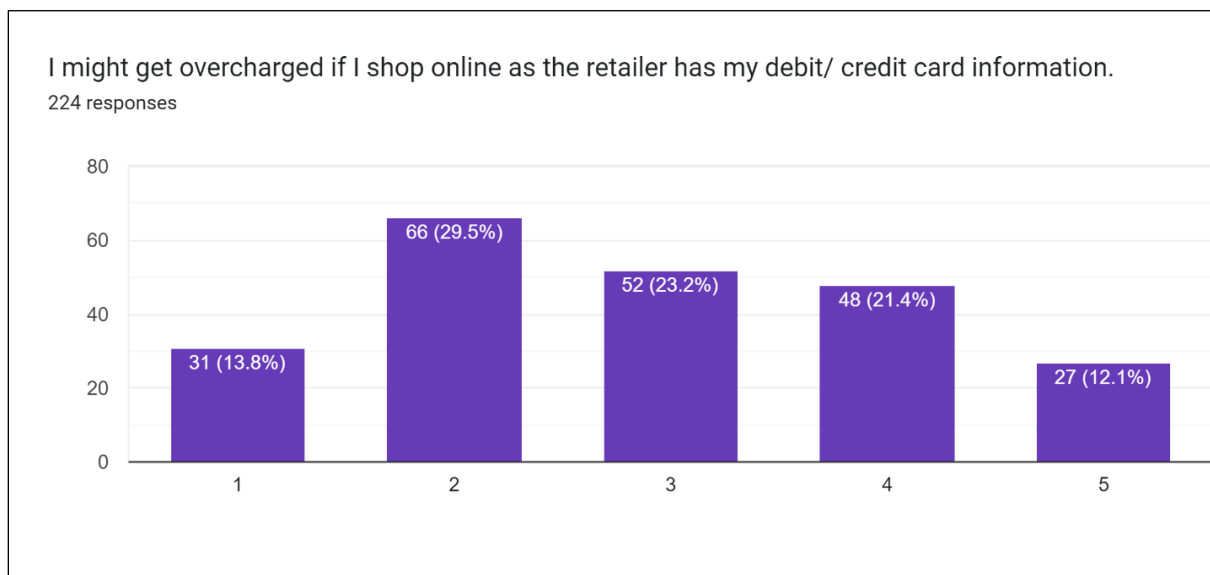
APPENDIX



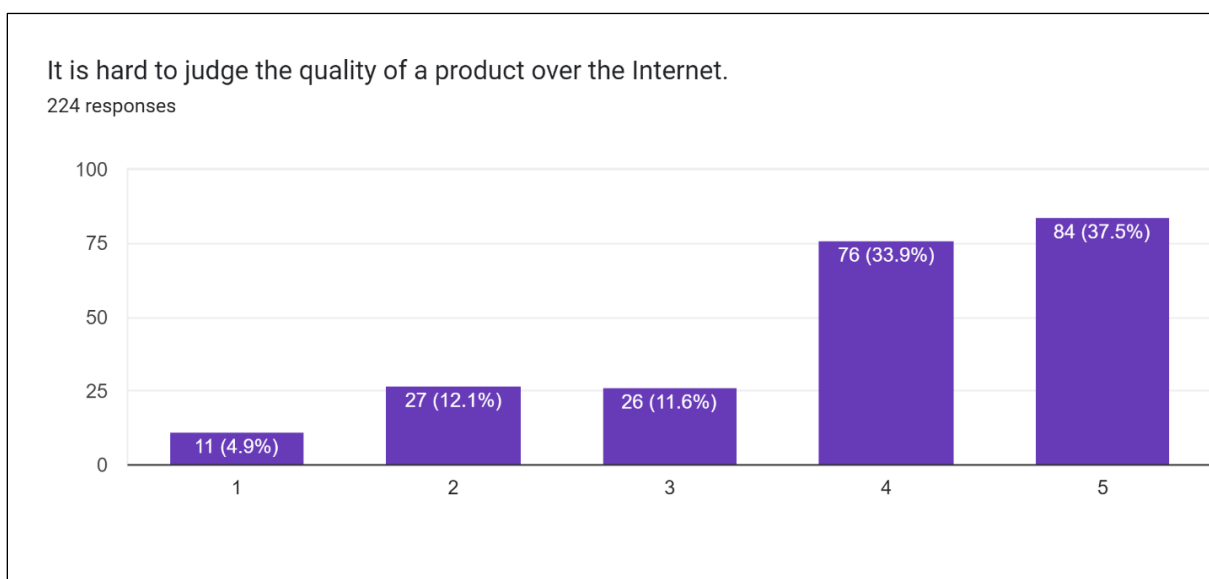
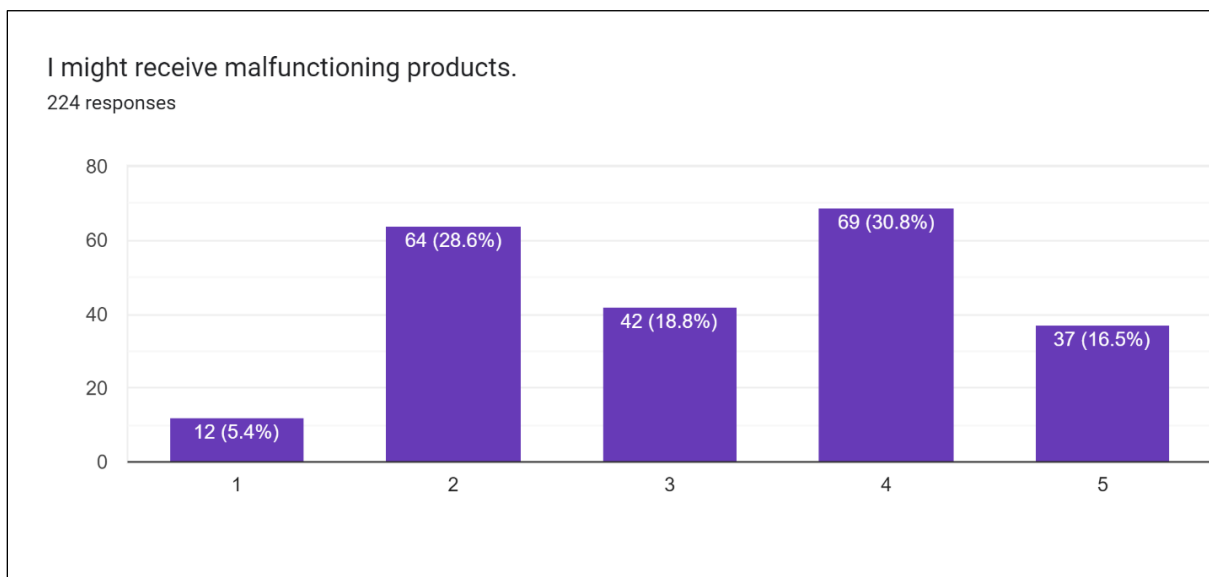
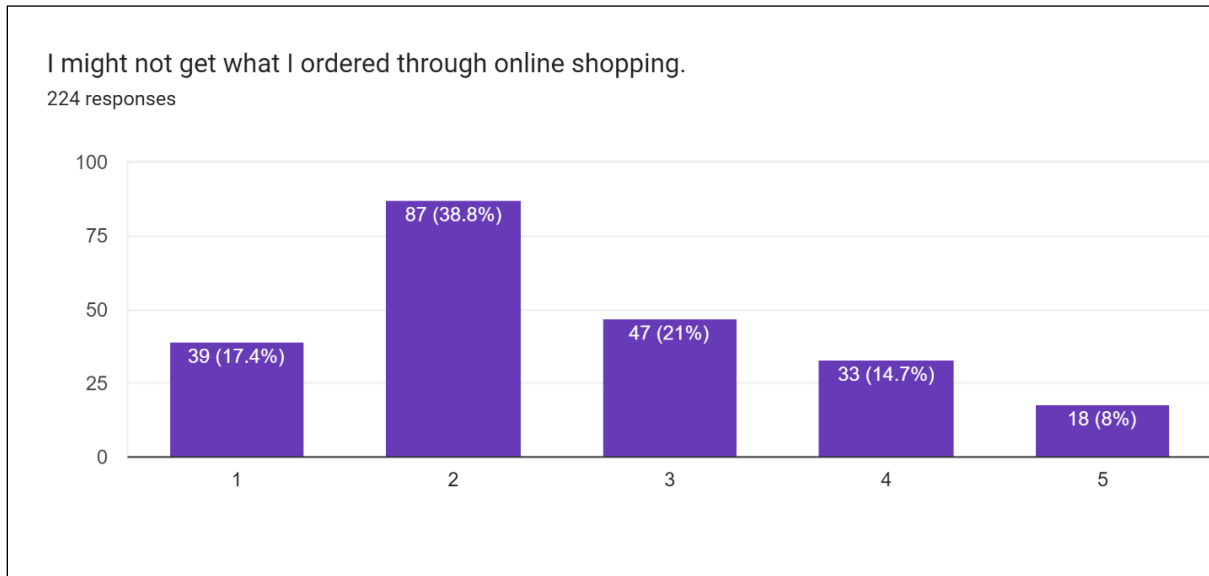
APPENDIX



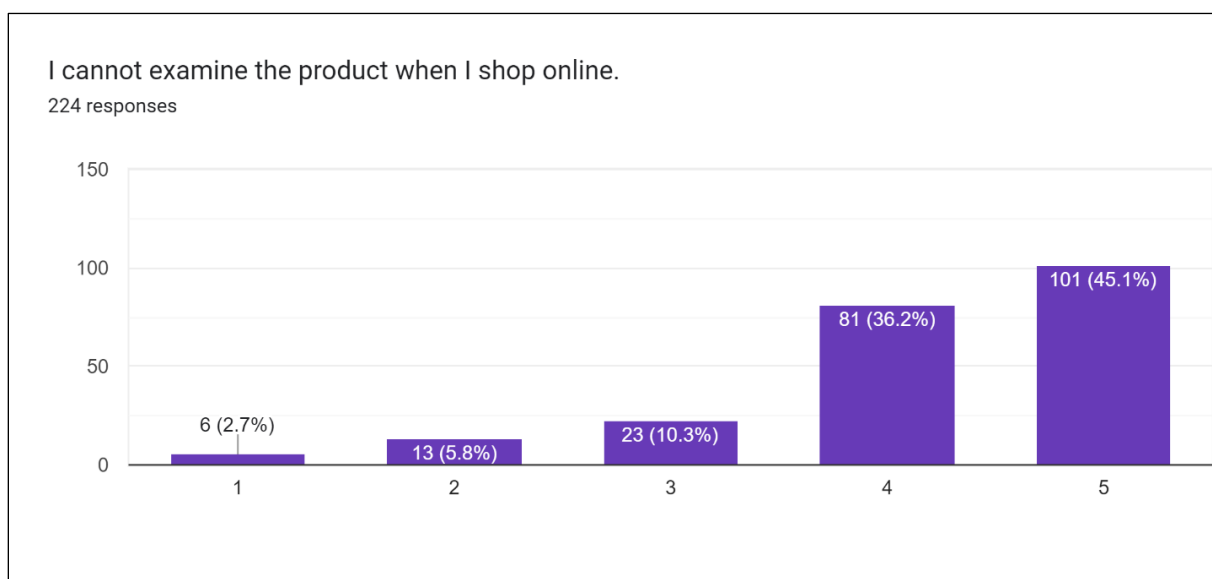
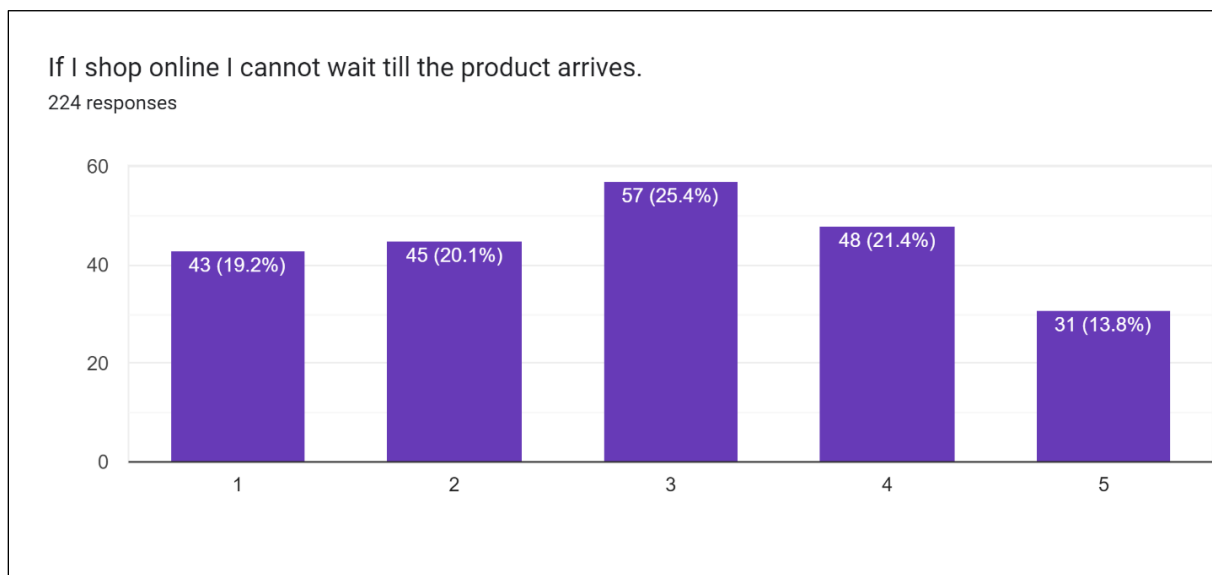
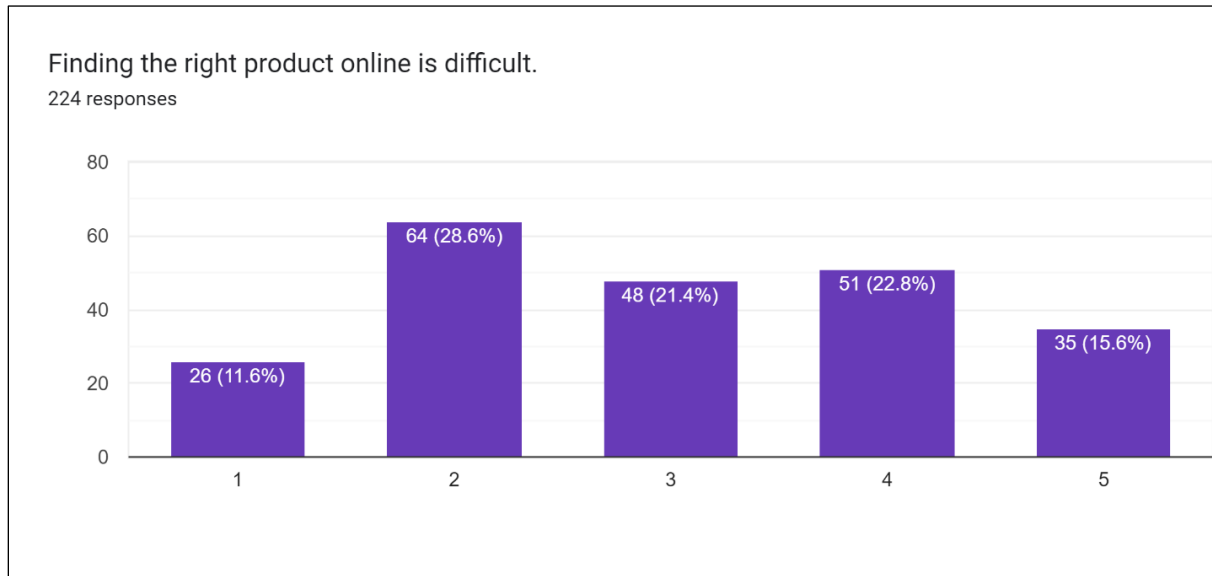
APPENDIX



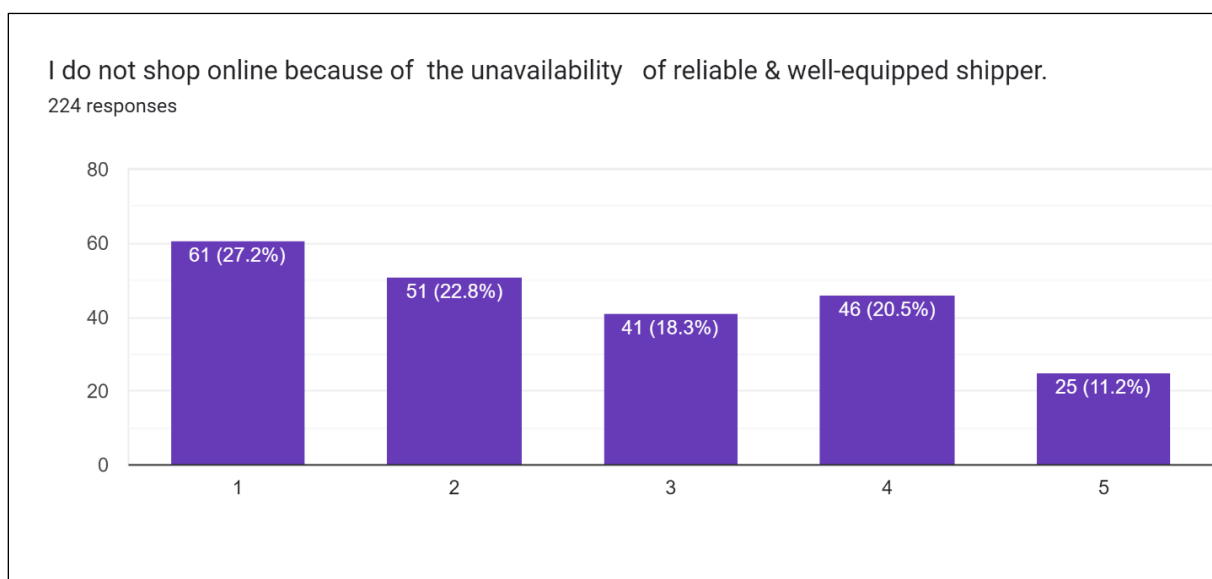
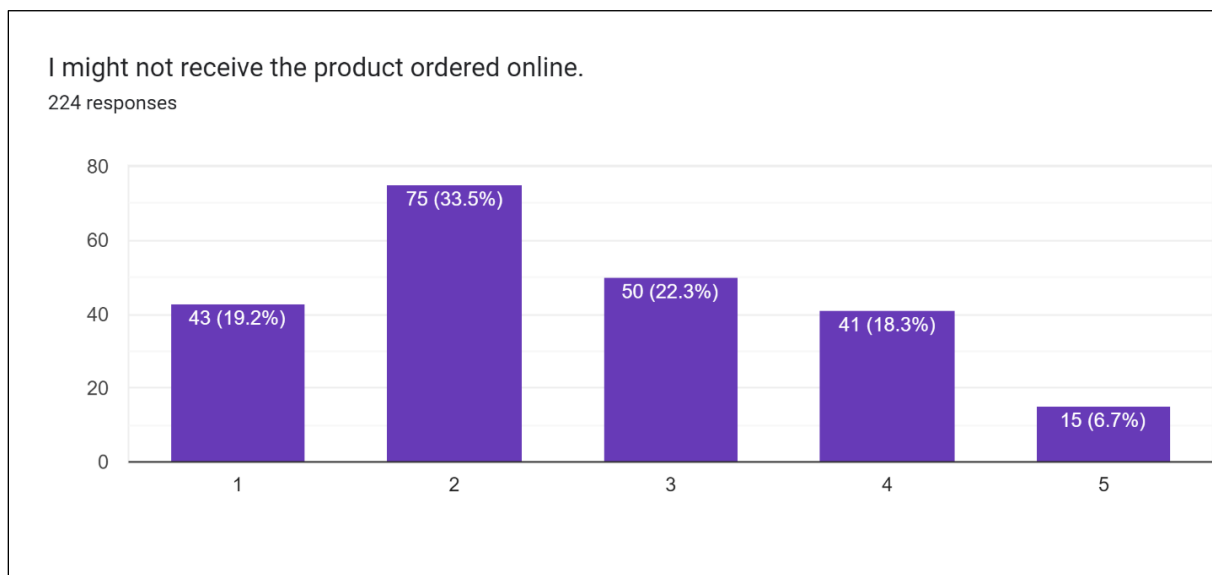
APPENDIX



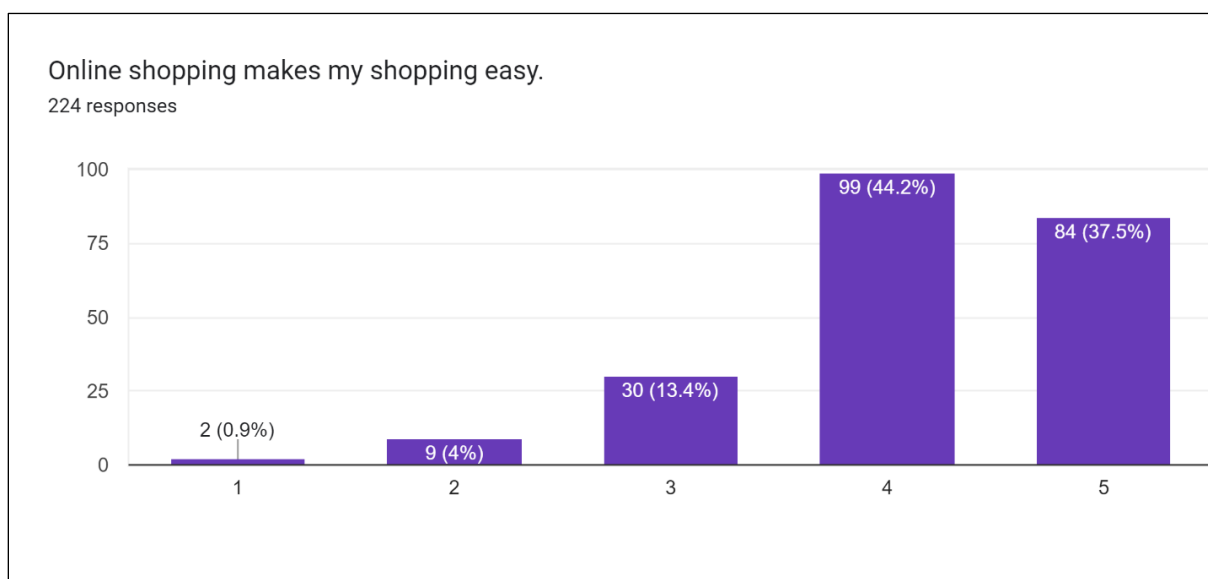
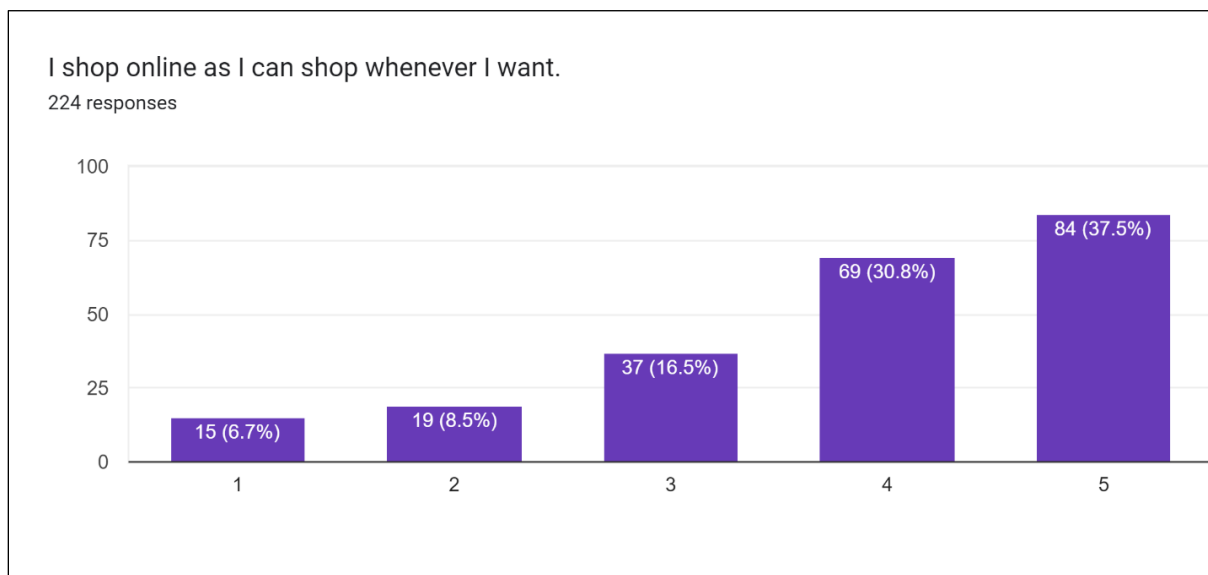
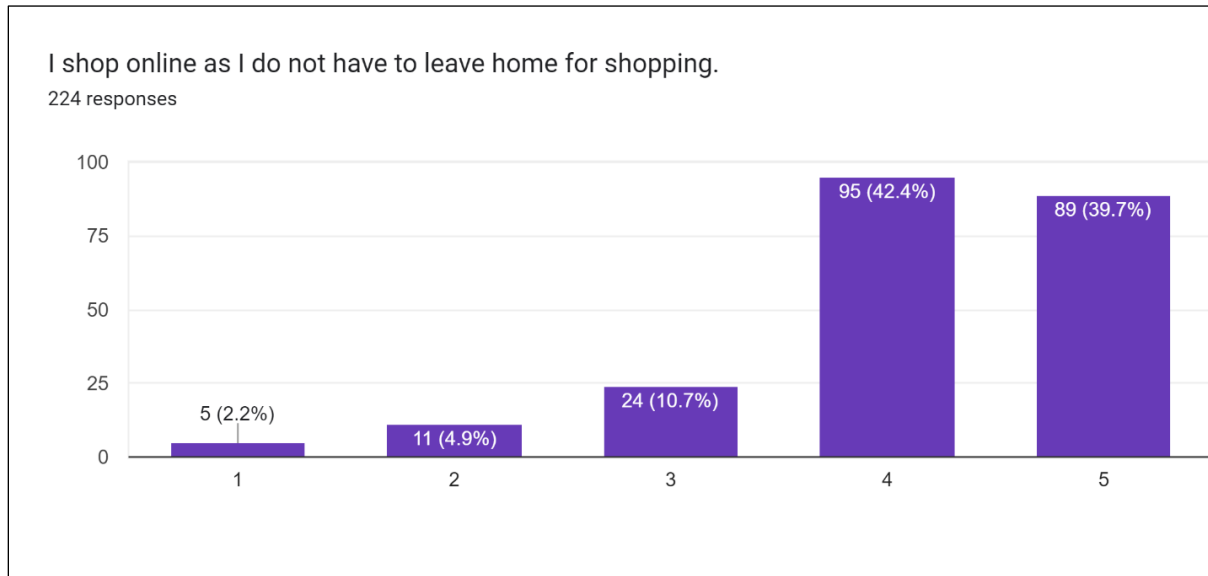
APPENDIX



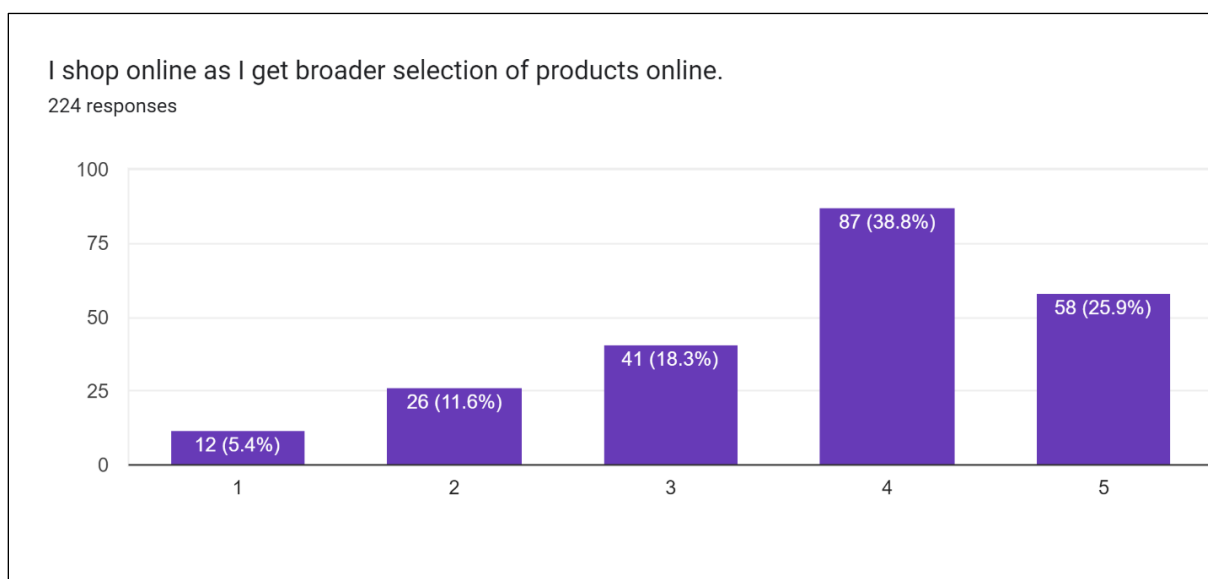
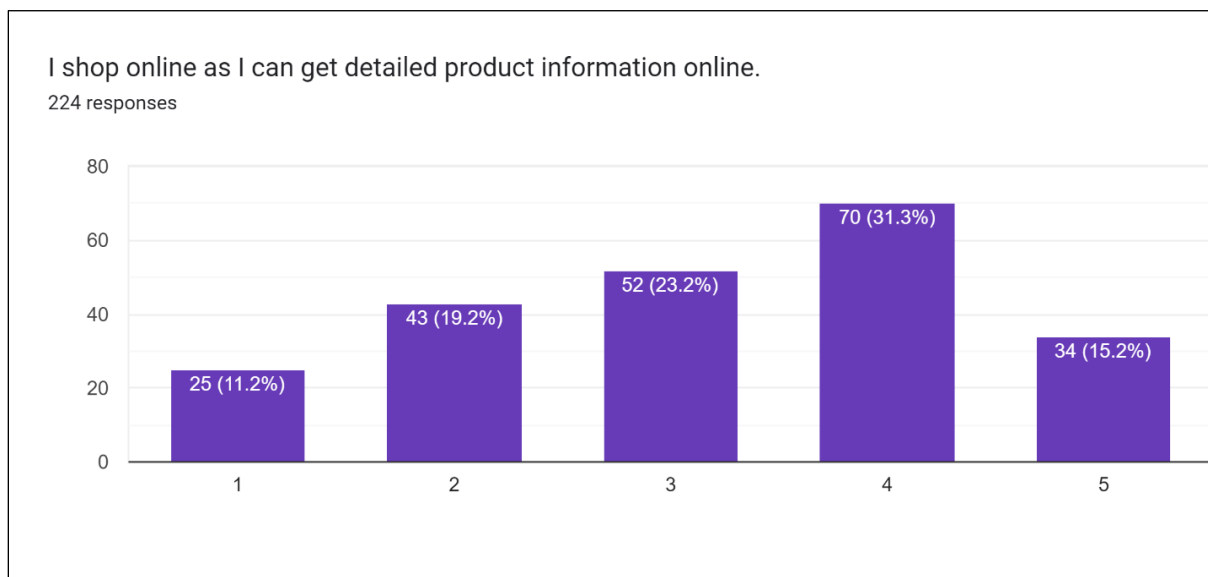
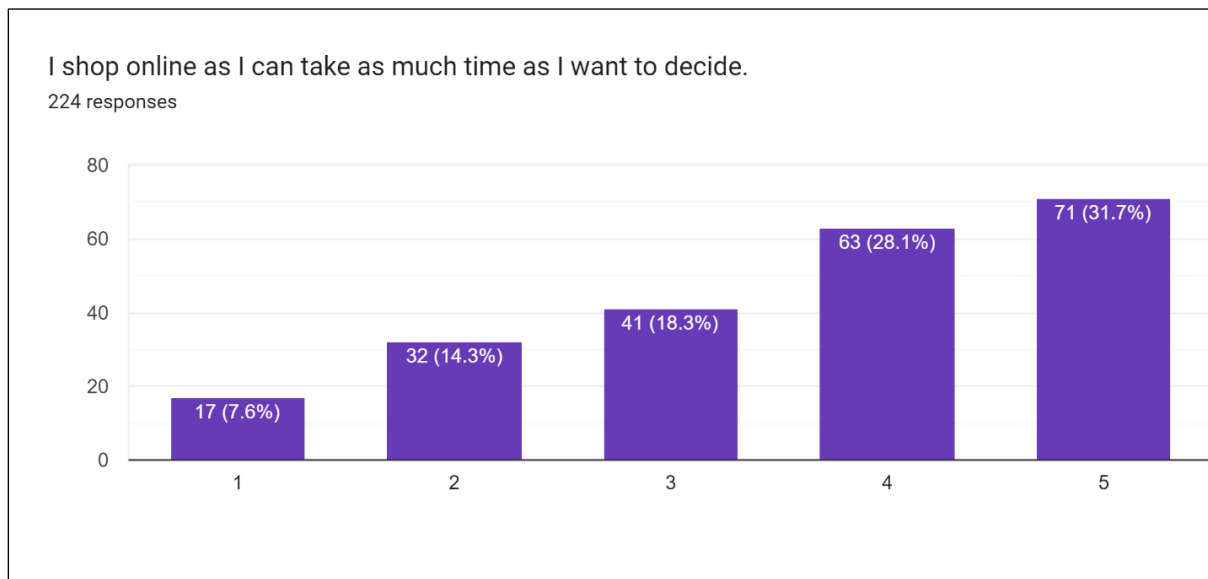
APPENDIX



APPENDIX



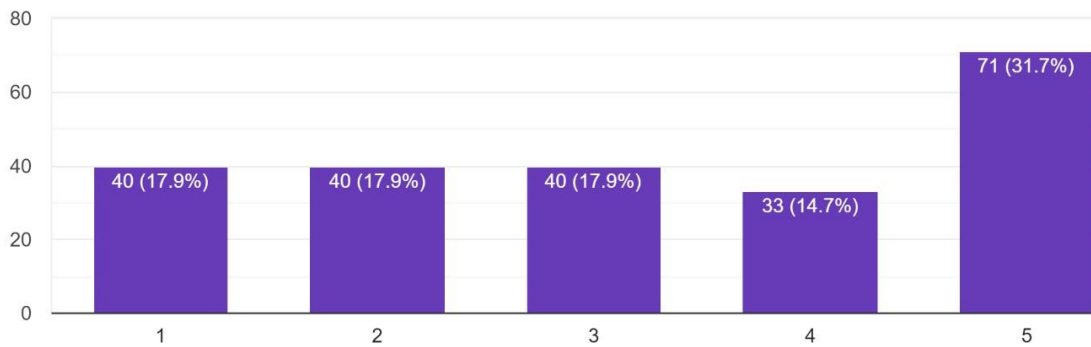
APPENDIX



APPENDIX

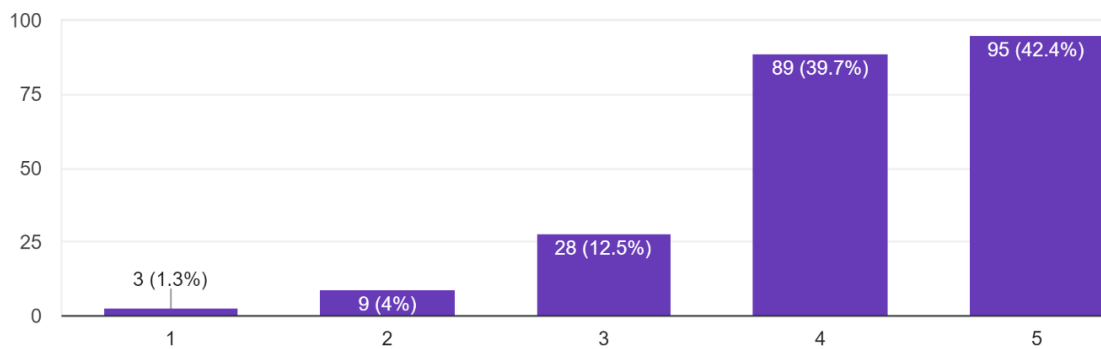
I use online shopping for buying products which are not easily available in the nearby market or are unique (new).

224 responses



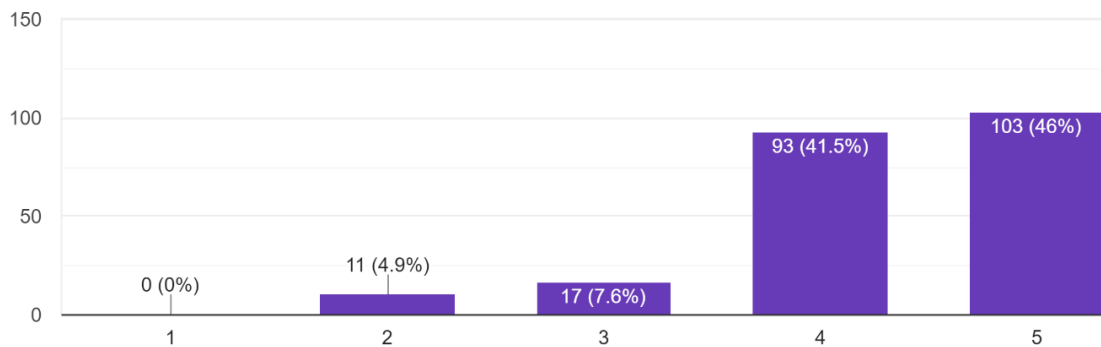
I shop online as I can save myself from chaos of traffic.

224 responses



I shop online as I can save myself from market crowd.

224 responses



APPENDIX

21-23	Male	8	1	5 FBF	Yes	6-10 times	3-6 months	4	4	2	4	5	3	1	1	5	5	1	4	3	3	2	5	2	3	1	1	5	5	2	4	1	1	3	3	5	4		
21-23	Male	2	1	5 FBF	Yes	1-2 times	1-2 years	5	1	3	3	5	4	1	5	2	5	3	5	3	2	4	5	3	3	1	3	5	5	2	5	5	5	1	5	5			
21-23	Male	0	1	5 FBF	Yes	Never	1-2 years	5	4	2	3	5	3	3	1	3	2	5	1	3	5	2	4	1	5	5	5	3	5	3	3	3	3	5	5				
21-23	Male	0	1	5 FBF	Yes	Never	1-2 years	4	4	4	4	5	3	1	3	4	4	4	4	4	5	3	4	3	5	3	2	5	2	5	2	3	4	4	4	4			
21-23	Female	2	0	5 FBF	Yes	1-2 times	More than 5 years	5	5	1	3	4	2	5	1	5	2	1	3	2	3	3	1	4	4	5	1	4	5	5	3	3	3	5	4	4			
21-23	Female	8	0	5 FBF	Yes	6-10 times	2-3 years	5	1	2	4	5	5	2	2	3	4	4	2	1	1	3	3	4	4	5	2	5	1	5	3	4	5	5	3	5	5		
21-23	Female	15	0	5 FBF	Yes	11-20 times	3-5 years	5	4	2	4	5	2	1	5	4	4	5	3	4	3	1	5	3	3	4	5	5	4	3	5	5	3	5	5	2	3		
21-23	Female	30	0	5 FBF	Yes	21 times or above	3-5 years	4	5	2	4	4	2	4	1	4	3	2	3	3	1	5	1	5	3	4	5	5	3	4	4	1	2	1	5	5	5		
21-23	Female	8	0	5 FBF	Yes	6-10 times	2-3 years	5	2	3	4	5	4	1	2	2	4	2	2	1	3	1	5	4	4	3	2	4	5	5	4	4	1	4	4	4	4		
21-23	Female	15	0	5 FBF	Yes	11-20 times	More than 5 years	4	3	2	4	5	4	2	3	3	3	4	4	2	2	5	3	5	4	4	3	2	4	5	5	3	4	3	1	5	5		
21-23	Female	0	0	4 FICT	Yes	Never	3-5 years	4	3	5	1	5	5	3	3	4	3	1	5	4	4	4	3	4	2	2	3	2	4	5	4	2	5	1	5	4	4		
21-23	Female	4	0	4 FICT	Yes	3-5 times	1-2 years	5	4	3	2	5	5	3	1	3	3	1	4	4	2	5	5	5	5	5	3	5	5	5	3	4	3	1	5	2	2		
21-23	Female	2	0	5 FBF	Yes	1-2 times	2-3 years	5	5	4	5	4	5	1	1	1	3	2	5	1	4	4	4	4	5	3	3	3	4	1	4	5	3	3	4	4	4		
21-23	Female	0	0	4 FICT	Yes	Never	6-12 months	4	2	2	5	5	1	3	3	4	2	3	5	1	2	3	4	2	4	2	2	3	5	5	4	2	3	3	2	4	5		
21-23	Female	4	0	4 FICT	Yes	3-5 times	6-12 months	4	4	5	1	4	1	4	1	4	1	2	2	3	5	4	4	4	4	5	3	2	4	5	5	4	1	2	5	4	4		
21-23	Female	8	0	4 FICT	Yes	6-10 times	More than 5 years	5	4	4	3	5	4	5	3	5	4	4	5	5	5	5	5	5	5	3	4	5	5	4	3	2	1	5	5	3	4	4	
21-23	Male	8	1	4 FICT	Yes	6-10 times	Less than 3 months	5	5	5	5	5	4	4	3	2	2	1	2	2	2	5	3	5	5	4	1	1	4	4	5	4	4	5	5	5	5	5	
21-23	Male	4	1	4 FICT	Yes	3-5 times	6-12 months	5	4	5	4	5	4	5	5	5	1	1	1	2	2	5	2	5	5	2	1	1	5	5	5	4	5	5	5	5	5	5	
24-26	Male	30	1	4 FICT	Yes	21 times or above	Less than 3 months	5	5	5	5	5	5	5	5	5	1	2	1	1	2	5	2	5	5	1	1	1	5	5	5	5	5	5	5	5	5	5	
24-26	Male	15	1	3 FAS	Yes	11-20 times	3-5 months	5	5	5	5	5	5	5	5	5	1	1	1	1	3	5	4	5	5	4	1	1	5	5	5	5	5	5	5	5	5	5	
24-26	Male	8	1	5 FBF	Yes	6-10 times	3-5 months	5	4	4	4	4	4	4	4	4	2	2	2	2	3	4	4	4	4	4	1	1	4	4	4	4	4	4	4	4	4	3	3
21-23	Female	8	0	5 FBF	Yes	6-10 times	More than 5 years	5	5	5	5	5	2	4	2	1	1	2	2	5	5	2	5	5	1	1	4	5	1	4	4	4	4	4	4	4	4	3	3
21-23	Male	8	1	5 FBF	Yes	6-10 times	2-3 years	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	1	1	1	4	4	4	4	4	4	4	4	4	4
21-23	Male	15	1	5 FBF	Yes	11-20 times	3-5 years	4	5	4	4	5	4	2	2	1	1	2	2	2	4	4	3	4	1	1	1	4	4	5	5	4	4	4	4	4	4	4	
21-23	Male	15	1	5 FBF	Yes	11-20 times	3-5 years	4	5	4	4	5	4	2	2	1	1	2	2	2	4	4	1	4	1	2	1	1	5	5	5	4	4	4	4	4	4	4	
21-23	Male	15	1	5 FBF	Yes	11-20 times	More than 5 years	5	5	5	5	5	4	5	4	4	1	1	1	2	4	4	1	1	1	1	1	4	4	5	5	4	4	4	4	4	4	4	5
18-20	Male	8	1	5 FBF	Yes	6-10 times	3-5 years	4	5	4	5	5	3	3	4	2	1	1	3	2	2	5	4	1	5	4	1	5	5	5	5	4	4	4	4	4	4	5	4
21-23	Male	2	1	5 FBF	Yes	1-2 times	2-3 years	4	4	4	4	4	2	2	3	3	1	1	2	3	2	4	4	2	5	4	1	1	5	5	5	4	4	4	4	4	4	4	
21-23	Male	8	1	5 FBF	Yes	6-10 times	3-5 years	4	5	4	4	5	4	4	4	4	1	1	2	2	2	5	2	2	5	2	1	1	4	5	4	4	4	4	4	4	4	5	5
21-23	Male	2	1	5 FBF	Yes	1-2 times	3-5 years	5	4	4	5	1	2	4	4	5	1	2	4	2	2	5	4	5	4	2	1	1	4	5	4	4	4	4	4	4	4	4	4
24-26	Male	0	1	5 FBF	Yes	Never	1-2 years	4	3	3	4	4	3	2	2	2	4	4	2	3	4	4	3	4	2	1	2	4	5	4	4	4	4	4	4	4	4	4	5

21-23	Male	2	1	5 FBF	Yes	1-2 times	2-3 years	4	3	3	4	5	2	3	3	2	3	3	2	2	5	4	4	5	4	1	1	4	4	4	5	2	4	5	5	5	5		
21-23	Female	8	0	4 FICT	Yes	6-10 times	6-12 months	5	3	5	4	4	4	5	1	3	5	2	4	3	5	4	2	4	3	4	5	2	2	3	2	2	2	4	4	5	5		
21-23	Female	0	0	4 FICT	Yes	Never	2-3 years	3	2	3	5	5	3	4	1	5	5	3	2	5	1	4	2	3	4	5	2	3	4	2	2	2	1	3	3	4	4		
21-23	Female	8	0	4 FICT	Yes	6-10 times	More than 5 years	5	3	5	5	3	5	1	5	1	3	1	3	1	2	1	4	5	4	4	2	3	4	2	5	2	5	3	3	4	4	4	
21-23	Female	0	0	4 FICT	Yes	Never	1-2 years	5	5	3	2	5	3	5	2	1	4	4	2	2	2	4	4	4	2	1	1	5	4	2	2	2	5	1	5	5	5	5	
21-23	Female	2	0	4 FICT	Yes	1-2 times	Less than 3 months	4	3	3	2	5	3	5	4	5	2	5	1	4	2	4	4	5	3	5	1	4	4	4	5	3	2	1	4	4	4	5	
21-23	Female	2	0	4 FICT	Yes	1-2 times	2-3 years	5	3	3	4	4	2	5	1	1	5	4	1	3	5	5	3	5	4	3	2	3	5	4	4	3	3	2	5	4	4		
21-23	Female	15	0	4 FICT	Yes	11-20 times	1-2 years	4	5	1	3	5	2	2	1	3	1	5	2	5	3	4	2	2	5	1	1	5	5	1	3	3	3	2	5	4	4	4	
21-23	Female	0	0	4 FICT	Yes	Never	Less than 3 months	4	2	4	4	5	3	2	3	1	3	3	5	2	5	3	1	3	1	3	3	1	5	3	4	5	3	4	4	4	4	5	
21-23	Female	8	0	4 FICT	Yes	6-10 times	6-12 months	5	3	5	5	5	5	1	4	4	1	1	2	3	1	2	5	4	3	3	1	4	4	5	4	3	4	1	4	4	4	4	
21-23	Female	8	0	4 FICT	Yes	6-10 times	3-6 months	4	2	5	1	5	1	1	5	3	4	1	3	4	4	3	1	3	4	3	4	4	2	2	2	5	4	4	4	4	4	5	5
21-23	Male	0	1	4 FICT	Yes	Never	Less than 3 months	5	3	5	3	5	2	4	3	1	4	1	2	1	4	3	4	1	3	4	2	1	2	4	4	5	4	4	4	4	4	4	5
21-23	Male	8	1	4 FICT	Yes	6-10 times	2-3 years	4	4	4	2	4	4	4	4	5	3	3	1	1	5	5	4	4	5	5	2	3	2	3	2	1	2	5	4	4	4	4	
21-23	Male	0	1	4 FICT	Yes	Never	2-3 years	4	5	2	5	5	4	2	3	3	1	1	2	5	5	1	4	4	4	1	2	1	5	2	4	5	3	5	5	4	4	4	
21-23	Male	2	0	4 FICT	Yes	1-2 times	6-12 months	5	2	5	2	5	5	3	3	2	2	2	4	1	5	2	2	4	4	4	2	2	5	5	5	3	4	4	4	4	4	4	4
21-23	Male	8	1	4 FICT	Yes	6-10 times	1-2 years	5	2	4	4	4	3	3	5	1	2	3	3	5	1	2	3	3	2	1	4	5	3	3	5	5	4	4	4	4	4	4	
21-23	Female	15	0	4 FICT	Yes	11-20 times	3-5 years	4	5	3	5	5	4	5	5	2	4	2	5	5	4	4	2	1	4	4	5	1	3	3	4	3	2	4	4	4	4	4	
21-23	Female	8	0	4 FICT	Yes	6-10 times	6-12 months	5	1	4	2	5	4	4	2	5	4	4	2	1	2	4	5	1	2	5	4	4	5	4	4	5	4	4	4	4	4	4	4
2																																							

APPENDIX

21-23	Male	8	1	4	FICT	Yes	6-10 times	More than 5 years	5	2	2	5	5	2	1	4	4	2	2	4	2	3	3	2	1	5	4	4	4	4	4	5	4	4	5	4	4				
21-23	Male	0	1	4	FICT	Yes	Never	2-3 years	5	5	3	4	4	4	3	2	1	5	5	1	1	5	5	2	2	5	3	2	4	4	5	4	3	1	2	2	4	4			
21-23	Female	0	0	4	FICT	Yes	Never	1-2 years	4	1	4	3	5	1	3	1	1	2	2	2	1	4	4	4	2	4	4	2	2	5	4	5	5	2	3	1	4	5			
21-23	Female	0	0	4	FICT	Yes	Never	3-5 years	4	2	2	3	4	3	3	5	2	2	3	5	4	4	4	2	4	4	1	3	2	5	4	4	3	2	2	1	5	4			
21-23	Female	4	0	4	FICT	Yes	3-5 times	More than 5 years	5	4	4	2	4	3	4	3	1	3	3	5	2	3	3	5	3	1	1	3	3	5	1	4	5	3	5	2	4	4			
18-20	Male	0	1	4	FICT	Yes	Never	3-5 years	4	1	5	2	5	1	4	1	1	3	2	4	3	4	4	3	3	4	4	2	3	5	4	4	3	2	4	3	5	5			
18-20	Female	0	0	3	FAS	Yes	Never	3-5 years	4	2	5	4	4	4	4	3	3	4	4	5	3	4	4	1	2	3	3	1	3	5	1	5	4	3	5	5	5	5			
18-20	Male	0	1	3	FAS	Yes	Never	3-5 years	5	2	1	2	5	3	4	3	1	2	4	4	2	3	3	2	3	5	5	4	4	4	3	1	1	2	1	5	5	5			
21-23	Female	4	0	3	FAS	Yes	3-5 times	2-3 years	5	5	2	1	4	5	2	3	1	1	1	3	3	4	4	1	1	4	4	2	2	5	5	5	2	3	1	2	5	5			
18-20	Female	2	0	3	FAS	Yes	1-2 times	6-12 months	5	1	3	5	4	4	4	3	4	5	4	4	3	3	5	2	1	4	4	4	5	4	4	5	2	5	2	4	4	4			
21-23	Male	4	1	4	FICT	Yes	3-5 times	6-12 months	5	3	5	3	4	3	1	4	2	2	3	1	2	5	5	3	3	4	2	3	3	4	5	5	1	3	2	4	4	4			
21-23	Female	2	0	4	FICT	Yes	1-2 times	3-5 years	5	4	1	1	4	5	5	1	4	1	5	5	2	3	5	3	4	4	5	4	4	5	4	5	5	3	5	5	5	5			
21-23	Female	4	0	4	FICT	Yes	3-5 times	1-2 years	5	5	3	5	5	5	1	5	3	4	4	3	1	4	4	2	3	4	3	2	4	4	5	5	3	3	4	2	4	4	4		
21-23	Female	8	0	4	FICT	Yes	6-10 times	2-3 years	4	3	4	3	5	1	5	3	2	5	5	1	5	4	1	3	5	1	2	3	4	2	4	2	3	2	5	5	5	5	5		
21-23	Female	2	0	4	FICT	Yes	1-2 times	More than 5 years	5	5	5	3	5	2	2	4	3	4	2	5	1	3	2	1	3	5	4	1	3	5	4	5	3	3	2	5	5	5			
21-23	Female	0	0	4	FICT	Yes	Never	1-2 years	4	5	2	4	5	3	1	3	5	1	4	5	2	2	5	4	3	4	4	3	1	4	5	5	1	3	3	4	4	4			
21-23	Female	2	0	4	FICT	Yes	1-2 times	More than 5 years	4	2	4	3	4	5	3	5	3	1	3	4	2	5	5	1	2	4	3	3	3	5	5	5	2	2	1	1	5	5	5		
21-23	Male	4	1	4	FICT	Yes	3-5 times	3-5 years	5	3	2	5	5	3	2	3	3	2	3	3	2	4	5	4	4	1	5	4	2	4	4	5	3	2	3	5	5	5	5		
21-23	Female	8	0	6	FSc	Yes	6-10 times	1-2 years	4	2	2	4	5	3	4	2	2	2	3	3	1	5	5	2	1	2	3	3	5	4	4	5	3	5	5	5	5	5	5		
21-23	Female	2	0	3	FAS	Yes	1-2 times	1-2 years	4	5	2	3	5	1	5	4	2	2	2	2	5	3	3	3	1	2	4	2	2	5	4	5	3	2	5	4	4	4	4		
21-23	Female	0	0	3	FAS	Yes	Never	3-5 years	5	2	3	4	5	3	5	1	1	4	3	3	1	5	5	4	1	4	2	1	1	5	1	4	3	2	5	5	5	5	5		
24-26	Male	8	1	3	FAS	Yes	6-10 times	More than 5 years	5	5	4	4	5	4	4	5	5	2	2	3	2	2	5	2	3	5	4	1	1	5	5	5	4	4	5	3	3	5	3		
24-26	Male	4	1	3	FAS	Yes	3-5 times	More than 5 years	4	3	3	4	4	4	4	4	4	2	2	4	2	2	4	2	3	4	3	2	2	5	4	4	4	4	5	4	4	5	4	4	
24-26	Male	4	1	3	FAS	Yes	3-5 times	More than 5 years	4	4	3	4	4	4	4	4	2	2	4	2	4	3	4	5	3	2	2	4	4	5	4	4	4	4	4	4	4	5	4	4	
24-26	Female	4	0	3	FAS	Yes	3-5 times	More than 5 years	4	4	4	4	5	3	3	4	4	3	3	4	2	2	4	2	3	4	3	2	2	4	4	5	4	4	4	4	4	4	5	4	4
24-26	Female	2	0	3	FAS	Yes	1-2 times	More than 5 years	4	4	4	4	4	2	2	4	4	2	2	3	3	3	4	4	4	3	2	1	4	5	5	4	4	5	4	4	5	4	4	5	4
24-26	Male	2	1	5	FBF	Yes	1-2 times	More than 5 years	5	4	4	4	5	3	3	4	4	2	2	3	3	2	4	3	4	4	3	2	2	4	5	5	4	4	5	4	4	5	4	4	
24-26	Male	2	1	5	FBF	Yes	1-2 times	3-5 years	4	3	3	5	3	3	4	2	2	2	3	2	2	4	3	4	4	3	2	2	4	4	4	4	4	4	4	4	5	5	5	5	
24-26	Male	4	1	5	FBF	Yes	3-5 times	3-5 years	5	4	4	5	5	3	3	4	4	3	3	4	2	2	4	3	4	4	3	2	2	4	5	5	4	4	5	5	5	5	5	5	
24-26	Male	4	1	5	FBF	Yes	3-5 times	3-5 years	4	3	4	3	5	2	2	3	3	2	2	3	2	2	4	2	3	4	4	2	3	5	4	5	4	4	5	5	5	5	5	5	

21-23	Male	4	1	5	FBF	Yes	3-5 times	3-5 years	5	4	4	4	5	4	4	5	4	2	2	3	2	2	4	3	4	4	3	2	2	4	5	4	4	4	5	5	5	5	5			
21-23	Male	4	1	5	FBF	Yes	3-5 times	3-5 years	4	4	4	4	5	3	3	4	3	3	4	2	1	4	3	4	4	4	3	2	3	4	5	5	4	4	5	5	5	5	5	5		
21-23	Female	8	0	4	FICT	Yes	6-10 times	3-5 years	4	5	5	2	4	2	3	2	1	1	3	5	3	4	4	3	5	5	2	2	5	2	4	2	3	4	4	4	4	4	4	4		
21-23	Female	2	0	4	FICT	Yes	1-2 times	1-2 years	4	5	1	5	4	2	1	2	2	1	4	3	2	1	1	1	4	3	2	3	4	3	5	3	5	5	1	5	5	5	5	5		
21-23	Female	2	0	4	FICT	Yes	1-2 times	6-12 months	5	2	5	2	5	2	1	3	4	2	3	3	2	5	5	2	3	4	3	2	2	4	3	4	5	1	3	5	4	5	4	5		
21-23	Female	4	0	4	FICT	Yes	3-5 times	3-5 years	5	5	2	1	5	5	5	3	3	2	2	3	1	3	3	4	4	5	3	2	5	4	3	2	5	5	4	3	2	5	4	5		
21-23	Female	0	0	4	FICT	Yes	Never	less than 3 months	4	1	3	2	4	1	3	3	1	1	1	3	1	2	1	3	1	5	5	4	5	2	3	5	4	1	5	2	4	3	4	3		
21-23	Female	0	0	4	FICT	Yes	Never	2-3 years	5	4	3	4	5	3	1	2	4	4	2	3	5	5	2	2	4	3	4	5	4	4	3	5	5	4	4	3	5	5	4	5	5	
21-23	Female	4	0	4	FICT	Yes	3-5 times	3-5 years	4	3	4	5	4	3	2	2	3	3	4	2	4	5	2	3	4	4	3	4	4	4	5	3	4	5	1	5	4	4	5	4		
21-23	Female	8	0	4	FICT	Yes	6-10 times	2-3 years	4	4	4	3	4	5	1	1	2	2	5	4	3	4	4	3	3	5	2	1	5	4	4	4	2	2	1	4	4	4	5	4	5	
21-23	Female	4	0	4	FICT	Yes	3-5 times	1-2 years	5	2	5	2	4	3	2	3	4	4	1	2	2	1	5	3	1	5	4	4	5	4	4	3	3	2	5	5	5	5	5	5		
21-23	Female	0	0	4	FICT	Yes	Never	2-3 years	5	3	5	4	4	3	4	2	1	2	2	5	3	4	4	1	3	4	5	4	3	4	4	3	4	2	5	5	4	3	4	2	5	5
21-23	Female	0	0	4	FICT	Yes	Never	2-3 years	5	5	4	4	4	2	3	4	3	5	5	2	4	4	2	1	5	4	4	4	5	2	5	4	2	2	2	5	4	4	5	5	5	
21-23	Female	0	0	4	FICT	Yes	Never	3-5 years	5	3	5	2	5	4	1	5	4	3	2	3	3	4	4	3	5	5	2	4	4	5	2	4	5	4	2	3	2	5	5	5		
21-23	Female	8	0	5	FBF	Yes	6-10 times	More than 5 years	5	5	5	5	5	5	5	5	5	5	4	4	5	4	5	2	1	4	2	4	1	4	5	5	5	5	5	5	5	5	5	5		

Code (Predictive Model 1)

```

from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score

# Assuming your data is already prepared
df = dataset

# Split data into features (X) and target (y)
# Define the features (X) based on the key influencers
X = df[['Gender (Numeric)', 'Faculty (Numeric)', 'Mid_Age',
        'Average_Attitude', 'Average_Availability', 'Average_Behavior',
        'Average_Convenience', 'Average_ConvenienceRisk', 'Average_Nondelivery',
        'Average_Product', 'Average_Security', 'Average_Time']]

# Define the target variable (y)
y = df['Frequency_Numeric']

# Split into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Train the model
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)

# Make predictions
df['Predicted Online Shopping Frequency'] = model.predict(X)

# Calculate R^2 and MSE
r2 = r2_score(y, df['Predicted Online Shopping Frequency'])
mse = mean_squared_error(y, df['Predicted Online Shopping Frequency'])

print(f'R-Squared: {r2}')
print(f'Mean Squared Error: {mse}')

from sklearn.metrics import mean_squared_error, r2_score

# Assuming df now has 'Predicted Online Shopping Frequency' and 'Online Shopping
Frequency Numeric'
r2 = r2_score(df['Frequency_Numeric'], df['Predicted Online Shopping Frequency'])
mse = mean_squared_error(df['Frequency_Numeric'], df['Predicted Online Shopping
Frequency'])

print(f'R-Squared: {r2}')
print(f'Mean Squared Error: {mse}')
print
import matplotlib.pyplot as plt

```


APPENDIX

```
# Plot the predicted vs actual values for the entire dataset
plt.figure(figsize=(10, 6))
plt.scatter(y, df['Predicted Online Shopping Frequency'], color='blue')
plt.plot([y.min(), y.max()], [y.min(), y.max()], color='red', linewidth=2)
plt.xlabel('Actual Online Shopping Frequency')
plt.ylabel('Predicted Online Shopping Frequency')
plt.title('Actual vs Predicted Online Shopping Frequency (Random Forest)')
plt.grid(True)
plt.show()
```

Code (Predictive Model 2)

```
import pandas as pd
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.preprocessing import LabelEncoder
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score
import matplotlib.pyplot as plt
import numpy as np

# Load your dataset (update the path and filename as necessary)
data = pd.read_excel(r'C:\Users\User\Desktop\Dataset.xlsx')

# If 'Online Shopping Frequency Numeric' doesn't exist, create it
if 'Online Shopping Frequency Numeric' not in dataset.columns:
    def convert_frequency_to_value(frequency):
        if frequency == "1-2 times":
            return 1.5
        elif frequency == "3-5 times":
            return 4
        elif frequency == "6-10 times":
            return 8
        elif frequency == "11-20 times":
            return 15
        elif frequency == "21 times or above":
            return 21
        return 0

    dataset['Online Shopping Frequency Numeric'] = dataset['Online Shopping
Frequency'].apply(convert_frequency_to_value)

# Encode categorical variables
label_encoder = LabelEncoder()
dataset['Gender Encoded'] = label_encoder.fit_transform(dataset['Gender'])
dataset['Faculty Encoded'] = label_encoder.fit_transform(dataset['Faculty'])

# If there is no grouping variable like 'CustomerID', compute standard deviation and range
without grouping
```

APPENDIX

```
dataset['Time_StdDev'] = dataset['Average_Time'].std()
dataset['ConvenienceRisk_StdDev'] = dataset['Average_ConvenienceRisk'].std()

dataset['Time_Range'] = dataset['Average_Time'].max() - dataset['Average_Time'].min()
dataset['ConvenienceRisk_Range'] = dataset['Average_ConvenienceRisk'].max() -
dataset['Average_ConvenienceRisk'].min()

# Create interaction terms
dataset['Time_ConvenienceRisk_Interaction'] = dataset['Average_Time'] *
dataset['Average_ConvenienceRisk']
dataset['Attitude_Product_Interaction'] = dataset['Average_Attitude'] *
dataset['Average_Product']

# Create polynomial features
dataset['Time_Squared'] = dataset['Average_Time'] ** 2
dataset['ConvenienceRisk_Squared'] = dataset['Average_ConvenienceRisk'] ** 2

# Define features and target based on new and existing features
X =
dataset[['Average_Attitude','Average_Convenience','Average_Product','Average_Behavior','F
aculty Encoded','Gender Encoded','Mid_Age'
]]
y = dataset['Online Shopping Frequency Numeric']

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Initialize and train the Random Forest Regressor
rf_model = RandomForestRegressor(n_estimators=100, random_state=42, max_depth=10,
min_samples_split=2)
rf_model.fit(X_train, y_train)

# Make predictions on the test set
y_pred_test_rf = rf_model.predict(X_test)

# Calculate performance metrics
mse_rf = mean_squared_error(y_test, y_pred_test_rf)
r2_rf = r2_score(y_test, y_pred_test_rf)

print(f"Random Forest with New Features - Mean Squared Error on Test Set: {mse_rf}")
print(f"Random Forest with New Features - R-Squared on Test Set: {r2_rf}")

# Perform 5-fold cross-validation
cv_scores_rf = cross_val_score(rf_model, X, y, cv=5, scoring='r2')

print(f"Random Forest with New Features - Cross-Validation R-Squared Scores:
{cv_scores_rf}")
print(f"Random Forest with New Features - Mean Cross-Validation R-Squared:
{np.mean(cv_scores_rf)}")
```

APPENDIX

```
print(f'Random Forest with New Features - Standard Deviation of Cross-Validation R-Squared:  
{np.std(cv_scores_rf)}')
```

```
# Plot the predicted vs actual values for the test set  
plt.figure(figsize=(10, 6))  
plt.scatter(y_test, y_pred_test_rf, color='blue')  
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], color='red', linewidth=2)  
plt.xlabel('Actual Online Shopping Frequency')  
plt.ylabel('Predicted Online Shopping Frequency')  
plt.title('Actual vs Predicted Online Shopping Frequency (Random Forest with New Features)')  
plt.grid(True)  
plt.show()
```

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 2, Year 3	Study week no.: 6
Student Name & ID: Chin Wai Teng 22ACB00233	
Supervisor: Ts Ahmad Subhi Bin Zolkafly	
Project Title: A predictive analysis on student behaviour in UTAR by using e-commerce platform	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- **Visualize the demographic, Likert scale chart, important features.**

2. WORK TO BE DONE

- **Refer the research objective in FYP1**

3. PROBLEMS ENCOUNTERED

- **There are no problems have encountered so far upon completing this project.**

4. SELF EVALUATION OF THE PROGRESS

- **On the beginning stage of FYP2**



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 2, Year 3	Study week no.: 8
Student Name & ID: Chin Wai Teng 22ACB00233	
Supervisor: Ts Ahmad Subhi Bin Zolkafly	
Project Title: A predictive analysis on student behaviour in UTAR by using e-commerce platform	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- **Develop predictive model**

2. WORK TO BE DONE

- **Select a model that are most suitable and can predict each point well for the data**

3. PROBLEMS ENCOUNTERED

- **Hard to choose one model that suitable for the data because of the data type that did not have hypothesis.**

4. SELF EVALUATION OF THE PROGRESS

- **Hard to find but keep progress.**



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 2, Year 3	Study week no.: 10
Student Name & ID: Chin Wai Teng 22ACB00233	
Supervisor: Ts Ahmad Subhi Bin Zolkafly	
Project Title: A predictive analysis on student behaviour in UTAR by using e-commerce platform	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- **Success develops a predictive model and visualize it by scatter plot chart**

2. WORK TO BE DONE

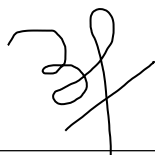
- **Develop a predictive model and visualize it**

3. PROBLEMS ENCOUNTERED

- **There are no problems have encountered so far upon completing this project.**

4. SELF EVALUATION OF THE PROGRESS

- **OK**



Supervisor's signature



Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 2, Year 3	Study week no.: 11
Student Name & ID: Chin Wai Teng 22ACB00233	
Supervisor: Ts Ahmad Subhi Bin Zolkafly	
Project Title: A predictive analysis on student behaviour in UTAR by using e-commerce platform	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- **Finalized FYP 2 report**

2. WORK TO BE DONE

- **Evaluate all the figure or dashboard to complete the report**

3. PROBLEMS ENCOUNTERED

- **There are no problems have encountered so far upon completing this project.**

4. SELF EVALUATION OF THE PROGRESS

- **Good**



Supervisor's signature



Student's signature

POSTER

UTAR
UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF INFORMATION COMMUNICATION AND TECHNOLOGY
A predictive analysis on student behaviour in UTAR by using e-commerce platform

01 INTRODUCTION
This research is to determine UTAR Students Online Shopping Behaviour by using proposed method

02 OBJECTIVE
To investigate online shopping behaviour from UTAR Students' attitude therefore develop and evaluate predictive ecommerce model

03 PROPOSED METHOD

- POWER BI
- DSLC
- PYTHON (JUPYTER)

04 CONCLUSION
Conduct a more complete and detailed study of this topic by gathering feedback from UTAR students

Project Developer:
Chin Wai Teng

Project Supervisor:
Ts Ahmad Subhi Bin Zolkafly

PLAGIARISM CHECK RESULT

PLAGIARISM CHECK RESULT

FYP2.docx			
ORIGINALITY REPORT			
7 %	3 %	1 %	4 %
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS
PRIMARY SOURCES			
1	eprints.utar.edu.my Internet Source		2 %
2	Submitted to Universiti Tunku Abdul Rahman Student Paper		1 %
3	Submitted to Victorian Institute of Technology Student Paper		1 %
4	Ulrike Grömping. "Variable Importance Assessment in Regression: Linear Regression versus Random Forest", The American Statistician, 2009 Publication		<1 %
5	forum.lowyat.net Internet Source		<1 %
6	Submitted to Hogeschool Rotterdam Student Paper		<1 %
7	Submitted to Taylor's Education Group Student Paper		<1 %
8	Submitted to Stadio Holdings Student Paper		<1 %

PLAGIARISM CHECK RESULT

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 0	Effective Date: 01/10/2013	Page No.: 1 of 1



FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Full Name(s) of Candidate(s)	Chin Wai Teng
ID Number(s)	22ACB00233
Programme / Course	BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION SYSTEMS
Title of Final Year Project	A predictive analysis on student behaviour in UTAR by using e-commerce platform

Similarity	Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)
Overall similarity index: <u>7</u> % Similarity by source Internet Sources: <u>3</u> % Publications: <u>1</u> % Student Papers: <u>4</u> %	
Number of individual sources listed of more than 3% similarity: <u>0</u> %	
Parameters of originality required and limits approved by UTAR are as Follows: (i) Overall similarity index is 20% and below, and (ii) Matching of individual sources listed must be less than 3% each, and (iii) Matching texts in continuous block must not exceed 8 words <i>Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.</i>	

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Signature of Supervisor

Signature of Co-Supervisor

Name: **Ts Ahmad Subhi Bin Zolkafly**

Name: _____

Date: 12/09/2024

Date: _____

FYP 2 CHECKLIST



UNIVERSITI TUNKU ABDUL RAHMAN

**FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY
(KAMPAR CAMPUS)**

CHECKLIST FOR FYP2 THESIS SUBMISSION

Student Id	22ACB00233
Student Name	Chin Wai Teng
Supervisor Name	Ts Ahmad Subhi Bin Zolkafly

TICK (√)	DOCUMENT ITEMS
	Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item.
√	Title Page
√	Signed Report Status Declaration Form
√	Signed FYP Thesis Submission Form
√	Signed form of the Declaration of Originality
√	Acknowledgement
√	Abstract
√	Table of Contents
√	List of Figures (if applicable)
√	List of Tables (if applicable)
√	List of Symbols (if applicable)
√	List of Abbreviations (if applicable)
√	Chapters / Content
√	Bibliography (or References)
√	All references in bibliography are cited in the thesis, especially in the chapter of literature review
√	Appendices (if applicable)
√	Weekly Log
√	Poster
√	Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005)
√	I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report.

*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

Y. Teng

(Signature of Student)

Date: 02/09/2024