

INTERNAL VERIFICATION – ASSESSMENT DECISIONS			
Programme title	BTEC Higher National Diploma in Computing		
Assessor	Mr. Tharindu Wijethilake	Internal Verifier	Mr. Sahan Devinda
Unit(s)	Unit 16: Computing Research Project (Pearson Set)		
Assignment title	Final Research Report – Big Data		
Student's name	R.M. Nisitha Nethsilu		
List which assessment criteria the Assessor has awarded.	Pass	Merit	Distinction
INTERNAL VERIFIER CHECKLIST			
Do the assessment criteria awarded match those shown in the assignment brief?	Y/N		
Is the Pass/Merit/Distinction grade awarded justified by the assessor's comments on the student work?	Y/N		
Has the work been assessed accurately?	Y/N		
Is the feedback to the student: Give details: • Constructive? • Linked to relevant assessment criteria? • Identifying opportunities for improved performance? • Agreeing actions?	Y/N Y/N Y/N Y/N		
Does the assessment decision need amending?	Y/N		
Assessor signature			Date
Internal Verifier signature			Date
Programme Leader signature(if required)			Date

Confirm action completed

Remedial action taken Give details:			
Assessor signature		Date	
Internal Verifier signature		Date	
Programme Leader signature(if required)		Date	

Higher Nationals - Summative Assignment Feedback Form

Student Name/ID	R.M. Nisitha Nethsilu		
Unit Title	Unit 16: Computing Research Project (Pearson Set)		
Assignment Number		Assessor	Mr. Tharindu Wijethilake
Submission Date	2025/01/20	Date Received 1st submission	
Re-submission Date		Date Received 2nd submission	

Assessor Feedback:

LO2 Conduct and analyse research relevant to a computing research project

Pass, Merit & Distinction Descripts	P3 <input type="checkbox"/>	P4 <input type="checkbox"/>	M2 <input type="checkbox"/>	D1 <input type="checkbox"/>
-------------------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------

LO3 Communicate the outcomes of a research project to identified stakeholders

Pass, Merit & Distinction Descripts	P5 <input type="checkbox"/>	M3 <input type="checkbox"/>	D2 <input type="checkbox"/>
-------------------------------------	-----------------------------	-----------------------------	-----------------------------

LO4 Reflect on the application of research methodologies and concepts

Pass, Merit & Distinction Descripts	P6 <input type="checkbox"/>	P7 <input type="checkbox"/>	M4 <input type="checkbox"/>	D3 <input type="checkbox"/>
-------------------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------

Grade:	Assessor Signature:	Date:
--------	---------------------	-------

Resubmission Feedback:

Grade:	Assessor Signature:	Date:
--------	---------------------	-------

Internal Verifier's Comments:

Signature & Date:

* Please note that grade decisions are provisional. They are only confirmed once internal and external moderation has taken place and grades decisions have been agreed at the assessment board.

Assignment Feedback

Formative Feedback: Assessor to Student

Action Plan

Summative feedback

Feedback: Student to Assessor

Assessor signature		Date	
Student signature		Date	



Pearson Higher Nationals in Computing

Unit 16: Computing Research Project
(Pearson Set)
Final Research Report

General Guidelines

1. A Cover page or title page – You should always attach a title page to your assignment. Use previous page as your cover sheet and make sure all the details are accurately filled.
2. Attach this brief as the first section of your assignment.

3. All the assignments should be prepared using a word processing software.
4. All the assignments should be printed on A4 sized papers. Use single side printing.
5. Allow 1" for top, bottom, right margins and 1.25" for the left margin of each page.

Word Processing Rules

1. The font size should be **12** point and should be in the style of **Time New Roman**.
2. **Use 1.5 line spacing.** Left justify all paragraphs.
3. Ensure that all the headings are consistent in terms of the font size and font style.
4. Use **footer function in the word processor to insert Your Name, Subject, Assignment No, and Page Number on each page.** This is useful if individual sheets become detached for any reason.
5. Use word processing application spell check and grammar check function to help editing your assignment.

Important Points:

1. It is strictly prohibited to use textboxes to add texts in the assignments, except for the compulsory information. eg: Figures, tables of comparison etc. Adding text boxes in the body except for the before mentioned compulsory information will result in rejection of your work.
2. Carefully check the hand in date and the instructions given in the assignment. Late submissions will not be accepted.
3. Ensure that you give yourself enough time to complete the assignment by the due date.
4. Excuses of any nature will not be accepted for failure to hand in the work on time.
5. You must take responsibility for managing your own time effectively.
6. If you are unable to hand in your assignment on time and have valid reasons such as illness, you may apply (in writing) for an extension.
7. Failure to achieve at least PASS criteria will result in a REFERRAL grade.
8. Non-submission of work without valid reasons will lead to an automatic REFERRAL. You will then be asked to complete an alternative assignment.
9. If you use other people's work or ideas in your assignment, reference them properly using HARVARD referencing system to avoid plagiarism. You have to provide both in-text citation and a reference list.
10. If you are proven to be guilty of plagiarism or any academic misconduct, your grade could be reduced to A REFERRAL or at worst you could be expelled from the course

Student Declaration

I hereby, declare that I know what plagiarism entails, namely to use another's work and to present it as my own without attributing the sources in the correct way. I further understand what it means to copy another's work.

1. I know that plagiarism is a punishable offence because it constitutes theft.
2. I understand the plagiarism and copying policy of the Pearson UK.
3. I know what the consequences will be if I plagiarise or copy another's work in any of the assignments for this program.
4. I declare therefore that all work presented by me for every aspects of my program, will be my own, and where I have made use of another's work, I will attribute the source in the correct way.
5. I acknowledge that the attachment of this document signed or not, constitutes a binding agreement between myself and Pearson UK.
6. I understand that my assignment will not be considered as submitted if this document is not attached to the attached.

Student's Signature: E195991@esoft.academy
(Provide E-mail ID)

Date:2025/01/20
(Provide Submission Date)

Assignment Brief

Student Name /ID Number	R.M. Nisitha Nethsilu 00180125
Unit Number and Title	Unit 16: Computing Research Project (Pearson Set)
Academic Year	2022/2023
Unit Tutor	Mr. Tharindu Wijethilake
Assignment Title	Final Research Project Report -Big Data
Issue Date	2025/01/05
Submission Date	2025/01/20
IV Name & Date	Mr. Sahan Devinda
Submission Format:	
<ul style="list-style-type: none">• The submission is in the form of an individual written report.• This should be written in a concise, formal business style using single spacing and font size 12.• You are required to make use of headings, paragraphs and subsections as appropriate, and all work must be supported with research• referenced using the Harvard referencing system.• Please provide a referencing list using the Harvard referencing system.• The recommended word limit is minimum 4,500 words• Copy of the research proposal need to be attached to the final project report.	

Unit Learning Outcomes:

- LO2.** Conduct and analyse research relevant to a computing research project
- LO3.** Communicate the outcomes of a research project to identified stakeholders
- LO4.** Reflect on the application of research methodologies and concepts

Assignment Brief and Guidance:

Learner is now required to provide a comprehensive research project report based on the findings of secondary and primary researches carried out on the project proposal submitted in the previous section.

The Learner requires to produce a detailed research project report covering following areas

- Conduct primary and secondary research using appropriate methods for a computing research project that consider costs, access and ethical issues.
- Carry out your research and apply appropriate analytical tools to analyse research findings and data.
- Draw conclusion based on the research findings.
- Communicate the outcomes of your research project to the identified audience.
- Reflect on the success of your research project and your performance at the end of the project and analyse on the extent to which outcomes meet the research. Evaluate project outcomes and provide valid and justified recommendations.
- Discuss and analyze the research methods applied and their effectiveness to meet project objectives. Analyse results in recommended actions and reflect the research process to suggest future improvements and research considerations. Discuss alternative research methodologies and lessons learnt in view of the outcomes.

Grading Rubric

Grading Criteria	Achieved	Feedback
P3 Conduct primary and secondary research using appropriate methods for a computing research project that consider costs, access, and ethical issues.		
P4 Apply appropriate analytical tools, analyse research findings and data.		
M2 Discuss merits, limitations, and pitfalls of approaches to data collection and analysis.		
D1 Critically evaluate research methodologies and processes in application to a computing research project to justify chosen research methods and analysis.		
P5 Communicate research outcomes in an appropriate manner for the intended audience.		
M3 Analyse the extent to which outcomes meet set research objectives and communicate judgements effectively for the intended audience		
D2 Evaluate outcomes and make valid, justified recommendations.		
P6 Discuss the effectiveness of research methods applied, for meeting objectives of the computing research project.		
P7 Discuss alternative research methodologies and lessons learnt in view of the outcomes.		
M4 Analyse results in recommended actions for improvements and future research considerations.		
D3 Demonstrate reflection and engagement in the resource process, leading to recommended actions for future improvement.		

Enhancing Customer Experience in E-Commerce Using Big Data Sentiment Analysis in Sri Lanka

By
R.M. Nisitha Nethsilu

E195991

Submitted in accordance with the requirements for the
COMPUTING RESEARCH PROJECT MODULE OF
PEARSON'S HND IN COMPUTING PROGRAMME

at the
ESOFT METRO CAMPUS

Name of research Tutor: Mr. Tharindu Wijethilake

2025/01/20

DECLARATION

Name of Research Candidate: R.M. Nisitha Nethsilu

Pearson Registration Number: E195991

Program Name: HND in Computing

Research Title: Enhancing Customer Experience in E-Commerce Using Big Data Sentiment Analysis in Sri Lanka

Field of Study: Data Science and Bigdata

I do solemnly and sincerely declare that:

- i. I'm the sole author of this study
- ii. This work is original
- iii. In case of any use if any information from other sources references of copyright with its ownership have been acknowledged in this document
- iv. I do not have any actual knowledge, nor do I ought reasonably to know that the making of the work constitutes an infringement of any copyright work
- v. I know that plagiarism is a punishable offence because it constitutes theft, I understand the plagiarism and copying policy of the Edexcel UK, I know what the consequences will be if I plagiarise or copy another's work in this research for this program.

Candidate Signature: E195991@esoft.academy

Date: 2025/01/20

Subscribed and solemnly declared before,

Supervisor's Name: Mr. Tharindu Wijethilake

Designation:

Supervisor's Signature:

Date:

ACKNOWLEDGMENT

First and foremost, I would like to extend my heartfelt gratitude to my research supervisor, Mr. Tharindu Wijethilake, for his invaluable guidance, encouragement, and support throughout the course of this research. His expertise, insightful feedback, and consistent mentorship have been instrumental in shaping the direction and quality of this study. I am deeply appreciative of the time and effort he dedicated to providing me with the necessary knowledge and skills to carry out this research successfully.

Additionally, I would like to express my gratitude to all the lecturers and staff members of the ESOFT Metro Campus, Piliyandala branch, for their unwavering support, resources, and the learning environment they fostered. Their collective efforts have played a significant role in broadening my understanding of the subject matter and enhancing my academic journey.

Special thanks to my parents and friends, whose encouragement, patience, and moral support have been my strength throughout this research. Their belief in my abilities and willingness to stand by me during challenging times have been a source of inspiration.

Finally, I would like to extend my thanks to everyone who contributed, directly or indirectly, to the completion of this research. Your guidance, assistance, and encouragement are sincerely appreciated, and I am truly grateful for the role you have all played in helping me achieve this milestone.

ABSTRACT

The swift growth of e-commerce in Sri Lanka has heightened rivalry among online retailers, making the customer experience a crucial factor for achieving success. This study examines how big data sentiment analysis contributes to improving customer experience in the e-commerce industry of Sri Lanka. Using sentiment analysis methods on customer feedback, companies can obtain important insights into consumer preferences, satisfaction rates, and engagement trends. The research intends to assess the influence of sentiment-based insights on customer satisfaction, buying behavior, and retention tactics.

A quantitative research method was used, gathering survey answers from 63 e-commerce users in Sri Lanka. Data analysis was conducted with IBM SPSS, utilizing correlation and regression analyses to discover significant connections between sentiment scores and important customer experience metrics. The results demonstrate a significant relationship between sentiment analysis and customer satisfaction, with insights derived from sentiment being essential in shaping engagement levels and retention rates. The regression models indicate that positive sentiment greatly boosts satisfaction and raises the chances of repeat purchases and recommendations, whereas negative sentiment underscores aspects needing service enhancements.

The research determines that large-scale data sentiment analysis is an effective means for enhancing e-commerce tactics, enabling companies to customize marketing initiatives, boost customer assistance, and enhance service excellence. By incorporating sentiment analysis into their decision-making strategies, Sri Lankan e-commerce platforms can enhance customer relationships and attain sustainable growth in a competitive online market. The study offers practical suggestions for e-commerce companies to adopt sentiment-based strategies, highlighting the necessity of monitoring sentiment in real-time and utilizing AI-driven tools for customer interaction.

CONTENTS

DECLARATION	i
ACKNOWLEDGMENT	ii
ABSTRACT	iii
CONTENTS	iv
LIST OF ABBREVIATIONS	vi
LIST OF TABLES	vii
LIST OF FIGURES	vii
CHAPTER 1 – INTRODUCTION	1
1.1. Introduction	1
1.2. Purpose of research	2
1.3. Significance of the Research	3
1.4. Research objectives	5
1.5. Research Sub objectives	8
1.6. Research questions	10
1.7. Hypothesis	15
1.8. Thesis structure	17
CHAPTER 1 - Introduction	17
CHAPTER 2 - Literature Review	18
CHAPTER 3 - Methodology	18
CHAPTER 4 - Presentation of Results	19
CHAPTER 5 - Conclusions and Recommendations	20
CHAPTER 2 - LITERATURE REVIEW	21
2.1. Literature Review	21
2.2. Conceptual framework	25
CHAPTER 3 – METHODOLOGY	29
3.1. Research philosophy	29
3.2. Research approach	32
3.3. Research strategy	35
3.4. Research Choice	40
3.5. Time frame	43
3.6. Data collection procedures	46
3.6.1. Type of Data	46
3.6.2. Data Collection Method	47
3.6.3. Data Collection and Analyze Tools	48
3.6.4. Questionnaire structure	50
3.6.5. Data Storage	54
3.7. Target population and sampling	58
3.7.1. Sampling Strategy	59

3.7.2. Sample Size	61
3.8. The selection of participants	63
3.9. Reliability, Validity, and Generalizability.....	64
3.10. Ethical issues of the research study	69
CHAPTER 4 - PRESENTATION OF RESULTS	73
4.1. Demographic Analysis	74
4.2. Correlation Analysis.....	77
4.2.1. Correlation Between RO2 (Customer Sentiment) and SO1 (Customer Satisfaction).....	78
4.2.2. Correlation Between RO3 (Sentiment Analysis) and SO2 (Customer Engagement & Purchase Behavior)	79
4.2.3. Correlation Between RO4 (Sentiment-Based Marketing Strategies) and SO3 (Customer Retention)	81
4.3. Regression Analysis	83
4.3.1. RO1 / Main Objective (The Role of Sentiment Analysis in Enhancing Customer Experience).....	84
4.3.2. RO2 / SO2 (Evaluating the Role of Sentiment Analysis)	87
4.3.3. RO3 / SO3 (Explore Big Data Tools for Analysis).....	90
4.3.4. RO4 / SO4 (Benchmark Customer Experience)	94
4.3.5. RO5 / SO5 (Provide Recommendations for Implementation)	98
CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS	103
5.1. Conclusion.....	103
5.1.1. RO1	103
5.1.2. RO2	104
5.1.3. RO3	105
5.1.4. RO4	105
5.1.5. RO5	106
5.2. Recommendations.....	106
5.3. Limitations	110
5.4. My Final Decision.....	110
5.5. Future Improvements.....	112
5.6. Personnel Reflection.....	113
5.6.1. Benefits for the researcher.....	113
5.6.2. Benefits for the Industry/organization.....	114
Referencing	115
Annexures.....	116
Annexures A: Glossary of Terms	116
Annexures B: Sample SPSS Charts/ Table.....	118
Annexures C: Feedback Form / Question list	138

LIST OF ABBREVIATIONS

1. Key performance indicator (KPI)
2. Net Promoter Score (NPS)
3. Correlation coefficients (r-values)
4. Significance levels (p-values)

LIST OF TABLES

Table 1 Research Strategy Table	40
Table 2 Research Timeline Summary	45
Table 3 Frequency Analysis of Categorical Variables	76
Table 4 Descriptive Statistics for Continuous Variables (Age Distribution)	76

LIST OF FIGURES

Figure 1 Conceptual Framework	29
Figure 2 Time Frame	38
Figure 3 Visualizing Demographics Using Graphs	77
Figure 4 Correlation Between RO2 (Customer Sentiment) and SO1 (Customer Satisfaction)	78
Figure 5 Correlation Between RO3 (Sentiment Analysis) and SO2 (Customer Engagement & Purchase Behavior)	80
Figure 6 Correlation Between RO4 (Sentiment-Based Marketing Strategies) and SO3 (Customer Retention)	82
Figure 7 Analysing 1	84
Figure 8 Analysing 2	85
Figure 9 Analysing 3	85
Figure 10 Analysing 4	86
Figure 11 Analysing 5	87
Figure 12 Analysing 6	88
Figure 13 Analysing 7	88
Figure 14 Analysing 8	89
Figure 15 Analysing 9	89
Figure 16 Analysing 10	90
Figure 17 Analysing 11	91
Figure 18 Analysing 12	92
Figure 19 Analysing 13	92
Figure 20 Analysing 14	93
Figure 21 Analysing 15	93
Figure 22 Analysing 16	94
Figure 23 Analysing 17	95
Figure 24 Analysing 18	96
Figure 25 Analysing 19	96
Figure 26 Analysing 20	97
Figure 27 Analysing 21	99
Figure 28 Analysing 22	99
Figure 29 Analysing 23	100
Figure 30 Analysing 24	100
Figure 31 Analysing 25	101
Figure 32 Analysing 26	102

CHAPTER 1 – INTRODUCTION

1.1. Introduction

In the fast-changing digital economy of today, e-commerce has transformed both the shopping habits of consumers and the operations of businesses. In Sri Lanka, the rise of e-commerce has been significant, fueled by greater internet access, the prevalent use of mobile devices, and a change in consumer habits favoring online shopping. Nonetheless, the competitive e-commerce environment requires ongoing innovation to draw in, keep, and please customers. An exceptional customer experience is fundamental to success, as favorable interactions and contentment directly affect customer loyalty and profit margins. Grasping and responding to customer requirements, tastes, and feelings have become essential for companies to succeed in the digital market.

Sentiment analysis, driven by big data, has become an essential instrument for understanding customer views and deriving practical insights. Through the examination of extensive amounts of unstructured data like customer feedback, social media engagement, and support inquiries, companies can discover patterns, trends, and sentiments that uncover true customer opinions. This sophisticated analytical method allows e-commerce platforms to tailor their strategies, refine marketing campaigns, and improve service quality, leading to increased customer satisfaction. However, Sri Lanka's e-commerce industry has not completely leveraged sentiment analysis to enhance customer experiences and maintain competitiveness internationally.

Worldwide, e-commerce leaders such as Amazon and Alibaba have established standards by utilizing big data and sentiment analysis to improve customer interaction and contentment. These industry pioneers employ advanced algorithms and analytical tools to gauge customer sentiment instantly, allowing them to provide personalized user experiences, suggest products, and tackle complaints proactively. Conversely, numerous Sri Lankan e-commerce sites depend on conventional, manual feedback analysis techniques that tend to be labor-intensive, inefficient, and less adept at adjusting to the evolving nature of customer expectations. Closing this gap by utilizing big data sentiment analysis might transform local e-commerce businesses, enabling them to compete more efficiently on a global level.

This research seeks to investigate how sentiment analysis powered by big data can enhance customer experience in the e-commerce sector of Sri Lanka. Through examining customer feedback and recognizing essential sentiments, companies can make educated choices to enhance

their operations, marketing approaches, and customer interaction initiatives. The study further investigates how implementing sentiment analysis offers a competitive advantage by contrasting domestic platforms with international e-commerce giants. Additionally, it assesses the tools and technologies that facilitate sentiment analysis, providing insights into their relevance and efficiency within the Sri Lankan framework.

To conclude, this research aims to offer practical suggestions for incorporating sentiment analysis into e-commerce practices in Sri Lanka. The goal is not just to improve customer satisfaction but also to promote innovation and development within the sector. This research highlights the transformative potential of big data sentiment analysis, stressing its importance as a strategic asset for achieving lasting success in Sri Lanka's growing e-commerce industry.

1.2. Purpose of research

The main objective of this research is to investigate how big data sentiment analysis can be effectively applied to improve customer experiences in Sri Lanka's e-commerce industry. In a time when customer satisfaction is crucial for business achievement, grasping the emotions and feelings expressed through customer feedback has become essential for e-commerce platforms. Using sentiment analysis, companies can obtain valuable insights from large unstructured data sources like customer feedback, ratings, and social media engagement. These insights allow businesses to achieve a greater grasp of customer preferences, challenges, and satisfaction rates, helping them to make educated choices that correspond with market needs. This study seeks to demonstrate the transformative power of sentiment analysis as a strategic means to connect customer expectations with business products.

The research also seeks to help e-commerce businesses comprehend customer feelings to improve their services and engagement tactics. Recognizing what customers appreciate and what annoys them is crucial for providing personalized and outstanding experiences. For example, customer input frequently points out recurring problems like late deliveries, product discrepancies, or insufficient customer service. By conducting extensive analysis of this feedback, companies can identify ongoing patterns and tackle these issues efficiently. Furthermore, sentiment analysis facilitates immediate evaluation of customer satisfaction, allowing businesses to react proactively to developing concerns. This proactive strategy not only establishes customer trust but also nurtures loyalty by ensuring customers feel recognized and appreciated.

Another important aim of this study is to assess the effect of sentiment analysis on marketing approaches and customer loyalty. In the current digital environment, successful marketing goes beyond promoting products; it focuses on creating significant connections with customers. Sentiment analysis enables companies to identify the emotional factors influencing customer actions, like their likes for products or their involvement with certain campaigns. Armed with this understanding, businesses can create focused marketing approaches that connect with their audience. Promotional efforts can emphasize positive sentiment patterns, whereas negative sentiment trends can inform enhancements in product features or service quality. This detailed comprehension guarantees that marketing efforts are pertinent, effective, and able to enhance customer engagement and loyalty.

Moreover, this research seeks to offer actionable suggestions for implementing cutting-edge big data tools and technologies in Sri Lanka's e-commerce industry. Although global e-commerce leaders have effectively incorporated these technologies to achieve a competitive advantage, Sri Lankan companies remain in the initial phases of adoption. The research will determine appropriate tools and platforms for sentiment analysis, considering the specific challenges and opportunities in the local context. Resources including Python libraries (e.g., NLTK, VADER) and cloud services like Google Cloud Natural Language API will be assessed for their practicality and effectiveness in sentiment data analysis. The suggestions will act as a roadmap for local companies to incorporate these technologies smoothly, guaranteeing that they meet operational requirements and meet customer demands.

To sum up, this research aims to promote innovation and boost competitiveness in Sri Lanka's e-commerce industry by harmonizing local practices with international benchmarks. By revealing crucial insights into customer feelings and providing practical approaches for utilizing big data tools, this study seeks to assist companies in adjusting to changing customer needs. It highlights the significance of welcoming technological innovations and adopting customer-focused strategies to succeed in a more competitive market. In conclusion, this research emphasizes the importance of sentiment analysis as a driver for advancement, excellence, and sustainable success in the e-commerce sector.

1.3. Significance of the Research

This research focuses on an essential element of contemporary e-commerce: improving customer experiences through the tactical use of big data sentiment analysis. In the current competitive online market, customer satisfaction is not just a measure but the cornerstone for ongoing business growth and achievement. For Sri Lanka's emerging e-commerce industry falling behind global leaders such as Amazon and Alibaba grasping customer sentiment is crucial to stay competitive and fulfill the demands of a technology-oriented consumer group. This study seeks to show how sentiment analysis can serve as a powerful instrument, allowing companies to capture the "voice of the customer" and utilize these findings to innovate, adjust, and succeed.

A significant contribution to this research is connecting customer expectations with business offers. Customer feedback, an important information source, frequently goes underused because of insufficient analytical skills. Sentiment analysis allows businesses to derive usable insights from this information, revealing customer discontent, likes, and new trends. For example, recognizing ongoing negative feelings regarding delivery lags could lead to operational enhancements, whereas favorable reviews of a specific product may guide marketing tactics. The ability to align business functions and services with customer requirements enhances satisfaction and loyalty.

In addition, the research highlights the crucial role of sentiment analysis in influencing marketing and engagement tactics. In a rapidly changing e-commerce landscape, conventional advertising is inadequate for engaging modern consumers. Data-driven methods, fueled by sentiment analysis, offer a competitive advantage by uncovering the emotional factors influencing customer actions. This insight enables companies to develop focused marketing initiatives, customized suggestions, and customer-oriented interaction tactics that strongly connect with their audience. For e-commerce firms in Sri Lanka, this offers a distinct chance to stand out in a competitive landscape by providing authentically personalized experiences.

A further key element of this study is its focus on utilizing cutting-edge technologies to stay pertinent in a progressively digital environment. While global e-commerce leaders establish new standards in customer experience via advanced analytics, Sri Lankan firms encounter difficulties in staying competitive. This research emphasizes the necessity for local enterprises to integrate big data tools and technologies like natural language processing (NLP) and machine learning into their workflows. By showcasing the usefulness and advantages of these technologies, the study fosters innovation and aids in developing the technological framework essential for worldwide competitiveness.

Furthermore, this research adds to both scholarly and practical understanding in the areas of big data analysis and customer experience management. Although sentiment analysis has been thoroughly studied in global settings, research specifically targeting its use in the e-commerce sector of Sri Lanka remains scarce. By tackling this gap, the research offers a regional viewpoint, investigating the specific challenges and opportunities encountered by Sri Lankan enterprises. The results will aid industry participants, researchers, and policymakers looking to utilize big data for economic and technological advancement.

Ultimately, the importance of this study is rooted in its aspiration for a more competitive and customer-oriented e-commerce environment in Sri Lanka. By providing companies with resources and approaches to leverage sentiment analysis, it opens the door for innovation, development, and improved customer satisfaction. By doing so, the study not only enhances the achievements of individual firms but also aids in the overall advancement of Sri Lanka's digital economy, establishing the nation as an innovative contender in the worldwide e-commerce arena.

1.4. Research objectives

In the constantly changing digital environment, e-commerce has emerged as a vital catalyst for economic expansion, transforming how companies engage with consumers. With increasing competition, grasping customer sentiment has become crucial for providing outstanding experiences and achieving a competitive advantage. This study aims to utilize big data sentiment analysis to improve customer satisfaction in e-commerce platforms in Sri Lanka. This study seeks to offer practical insights for businesses to enhance their strategies by examining customer feedback, investigating cutting-edge tools and technologies, and comparing local practices with global benchmarks. In the end, the study aims to connect customer expectations with service delivery, enabling Sri Lankan e-commerce businesses to succeed in the vibrant digital economy.

1. Analyze Customer Sentiment in E-Commerce = The foundation of this study is to meticulously examine customer sentiment across different e-commerce platforms in Sri Lanka. This study will investigate how customers view products, services, and overall platform performance by gathering and analyzing feedback, including reviews, ratings, and complaints. The findings from this analysis will assist in identifying common problems that customers encounter, like

delays in delivery or inadequate product quality. For example, if many reviews show discontent with after-sales service, companies can quickly implement corrective actions to enhance it.

Additionally, this goal is essential for identifying customer trends and preferences. Positive feedback can emphasize what customers value the most, like smooth payment processes or convenient returns, allowing businesses to concentrate on these advantages. On the other hand, recognizing trends in dissatisfaction can help companies focus on areas that require enhancement. This in-depth comprehension of customer feelings and preferences will enable businesses to cultivate deeper relationships with their users, ultimately resulting in enhanced customer retention, satisfaction, and loyalty.

2. Evaluating the Role of Sentiment Analysis = The foundation of this study is to meticulously examine customer sentiment across different e-commerce platforms in Sri Lanka. This study will investigate how customers view products, services, and overall platform performance by gathering and analyzing feedback, including reviews, ratings, and complaints. The findings from this analysis will assist in identifying common problems that customers encounter, like delays in delivery or inadequate product quality. For example, if many reviews show discontent with after-sales service, companies can quickly implement corrective actions to enhance it.

Additionally, this goal is essential for identifying customer trends and preferences. Positive feedback can emphasize what customers value the most, like smooth payment processes or convenient returns, allowing businesses to concentrate on these advantages. On the other hand, recognizing trends in dissatisfaction can help companies focus on areas that require enhancement. This in depth comprehension of customer feelings and preferences will enable businesses to cultivate deeper relationships with their users, ultimately resulting in enhanced customer retention, satisfaction, and loyalty.

3. Explore Big Data Tools for Analysis = The third objective highlights the significance of recognizing and employing suitable big data tools and technologies for sentiment analysis within e-commerce. This research will assess instruments like the Natural Language Toolkit (NLTK), VADER (Valence Aware Dictionary and Sentiment Reasoner), and machine learning models for their efficacy in extracting and analyzing sentiment from extensive datasets. Every one of these tools provides distinct benefits in analyzing textual information and understanding customer feelings.

Furthermore, sophisticated platforms such as Google Cloud Natural Language API and Microsoft Azure Text Analytics will be investigated for their functionalities in real-time feedback handling and scalability. These platforms can manage extensive data while providing predictive insights, making them perfect for the dynamic e-commerce landscape. The research seeks to assist Sri Lankan e-commerce businesses in selecting the most appropriate technologies for their operations by exploring the technical characteristics and constraints of these tools. This goal not only emphasizes the practicality of sentiment analysis but also provides companies with the expertise required to execute it successfully.

4. Benchmark Customer Experience = A primary aim of this research is to evaluate the customer experience provided by Sri Lankan e-commerce sites in comparison to global giants like Amazon and Alibaba. These global platforms have established elevated benchmarks in utilizing sentiment analysis to provide exceptional customer experiences. For instance, Amazon's sophisticated recommendation systems, driven by sentiment analysis, allow the platform to offer extremely tailored product recommendations to users. In a similar manner, Alibaba utilizes data-driven insights to predict customer requirements and tackle potential problems before they arise.

The study will reveal gaps and areas for enhancement by contrasting Sri Lankan platforms with these international giants. For example, although global platforms may utilize automated systems to address customer issues quickly, regional platforms often depend on manual procedures that delay their responses. This benchmarking procedure will offer practical insights and effective strategies that Sri Lankan firms can implement to improve their competitiveness. It will also illuminate aspects where local platforms thrive, providing a balanced view of their effectiveness.

5. Provide Recommendations for Implementation = The goal is to create implementable suggestions for incorporating sentiment analysis into the functions of e-commerce companies in Sri Lanka. These suggestions will encompass a plan for implementing tools and technologies, including investing in cloud-based analytics solutions for increased scalability and efficiency. Moreover, the study will highlight the importance of training staff in sentiment analysis methods, such as natural language processing and machine learning, to develop internal proficiency.

Practical recommendations will also emphasize aligning sentiment analysis efforts with wider business objectives. For example, companies can utilize findings from sentiment analysis to boost customer retention by tackling frequent issues or improving operational efficiency via data-informed decisions. Additionally, this goal seeks to guarantee that businesses progress beyond theoretical knowledge and apply sentiment analysis in practical settings. By establishing a clear framework for adoption, the study seeks to enable Sri Lankan e-commerce platforms to innovate, improve customer satisfaction, and achieve a competitive advantage in the global marketplace.

1.5. Research Sub objectives

In the fiercely competitive e-commerce environment of today, grasping customer sentiments is essential for improving the entire customer experience. Sentiment analysis, utilizing big data technologies to gauge customer feelings from feedback and reviews, has emerged as a potent instrument for enhancing business results. For e-commerce businesses in Sri Lanka, utilizing sentiment analysis can reveal important insights into customer likes, issues, and expectations, which are crucial for customizing products and services that fulfill market needs. Nonetheless, applying sentiment analysis presents its own array of challenges, particularly when considering the distinctive cultural and linguistic backdrop of Sri Lanka. This research seeks to tackle these challenges and opportunities by establishing clear goals that examine the influence of sentiment analysis on customer experience, its contribution to developing marketing strategies, and the practical application of these technologies within Sri Lankan e-commerce platforms.

1. Sub Objective 01: Identify Key Metrics for Measuring Customer Experience = The initial sub-objective focuses on recognizing and comprehending the key performance indicators (KPIs) essential for evaluating customer experience within the realm of sentiment analysis. These measurements, such as Customer Satisfaction (CSAT), Net Promoter Score (NPS), and customer retention rates, deliver businesses quantifiable information to evaluate the standard of their customer service and product offerings. Sentiment analysis acts as a connector between unrefined customer feedback and these KPIs, enabling businesses to convert subjective data such as feelings, views, and grievances into concrete metrics. For example, monitoring unfavorable sentiment concerning a particular product or service may suggest the need for enhancements, whereas favorable sentiments could showcase strengths of the business.

Additionally, by correlating sentiment analysis findings with these KPIs, companies can gain a clearer insight into the way their actions influence customer satisfaction and loyalty. This sub-objective holds particular importance for Sri Lankan e-commerce firms aiming to embrace a more data-driven strategy for managing customer experiences, providing them with actionable insights to inform their decision-making processes.

Broadening the application of sentiment analysis for monitoring these KPIs allows companies to react proactively. Through constant tracking of customer responses and examining sentiment patterns, businesses can quickly respond to problems, thereby averting possible customer loss. For instance, if a product consistently gets adverse reviews about shipping delays, companies can adjust their processes to better their logistics and boost customer happiness. Additionally, recognizing the elements of customer service that promote positive feelings can assist e-commerce businesses in customizing their products and marketing approaches to enhance their customer connections.

2. Sub Objective 02: Assess Challenges in Implementing Sentiment Analysis – The second sub-objective aims to investigate the challenges that e-commerce companies in Sri Lanka encounter while trying to adopt sentiment analysis technologies. Though it has the capacity to transform customer interaction, sentiment analysis faces various challenges. A major concern is the expense involved in obtaining and sustaining advanced analytics tools, along with the shortage of technical expertise needed to use these tools efficiently. In Sri Lanka, small and medium-sized enterprises (SMEs) might find it challenging to acquire the financial and technical resources needed to incorporate sentiment analysis into their functions. Moreover, issues related to data quality and integration pose considerable challenges. Businesses frequently possess partial or chaotic datasets that are challenging to analyze, leading to erroneous sentiment analysis results.

This sub-objective also tackles the operational difficulties that emerge from employing sentiment analysis tools. For example, businesses might need to reformulate their data gathering procedures or create new methods for sorting and classifying customer responses. Additionally, Sri Lanka's e-commerce businesses might encounter cultural and linguistic obstacles when utilizing sentiment analysis tools, since numerous tools are created to handle content in English or standard forms of local languages, potentially failing to grasp the subtleties of Sri Lankan expressions. Recognizing these obstacles and offering strategies to

navigate them is crucial for helping businesses successfully implement sentiment analysis. This sub-objective aims to provide companies with the understanding to overcome these challenges, thereby enabling them to fully leverage sentiment analysis while considering regional constraints.

3. Sub Objective 03: Evaluate the Impact of Cultural and Regional Factors on Sentiment Analysis
 - = The third sub-objective seeks to explore the impact of cultural and regional elements in Sri Lanka on the efficacy of sentiment analysis. Linguistic and cultural nuances are vital in how customers convey their feelings, and these intricacies can create difficulties for automated sentiment analysis systems. For instance, some words or phrases might have varying connotations in Sri Lankan Sinhala or Tamil relative to their meanings in English, making it difficult for algorithms to effectively understand the sentiment conveyed by the words. Additionally, the incorporation of mixed languages (Sinhala, Tamil, and English) in online communications introduces another level of difficulty to sentiment analysis, since numerous tools lack the capability to handle multilingual data effortlessly.

This sub-objective aims to illuminate how these cultural and linguistic influences can cause inaccuracies in sentiment detection if they are not adequately addressed. For example, feelings of discontent may be conveyed through humor or indirect phrasing, potentially leading to misinterpretation by conventional sentiment analysis algorithms. Moreover, differences in regional language use might lead to variations in how customers from various areas of Sri Lanka share their views, potentially influencing the overall sentiment analysis. Grasping and tackling these issues is essential for enhancing the accuracy and dependability of sentiment analysis in Sri Lanka. This study aims to explore strategies for enhancing sentiment analysis instruments to suit the country's distinct linguistic environment, allowing companies to more effectively interpret and respond to customer input.

1.6. Research questions

In the fast paced and expanding e-commerce sector, grasping customer sentiments and preferences is vital for enhancing the entire shopping experience and influencing future business approaches. This survey seeks to collect information on customer satisfaction, elements affecting buying choices, and expectations from e-commerce platforms in Sri Lanka. The survey will collect both quantitative and qualitative data, aiding in a better comprehension of the elements influencing

customer behavior and pointing out areas that need enhancement in the online shopping experience. The information gathered will be crucial in helping companies respond more effectively to the changing needs of their customers, leading to improved user experiences and increased customer loyalty.

1. What is your age group? = This inquiry seeks to gather demographic details regarding the respondents' age, an essential element in comprehending the target audience of e-commerce sites. Age can affect shopping behaviors, preferences for specific products, and how often one shops online. By classifying respondents into specific age categories, businesses can understand the diverse needs and expectations that exist among various age groups.
2. What is your gender? = Grasping the gender distribution among e-commerce users aids in customizing marketing strategies, promotions, and product selections. Gender can affect shopping choices, including the kinds of products bought or the favored communication methods of the platform. This question enables businesses to evaluate if they are engaging a diverse audience or if they need to modify their strategy to appeal to a particular demographic.
3. Which region do you live in? = Geographic data aids in recognizing regional variances in purchasing habits, access to e-commerce services, and possible local obstacles like delivery logistics. In Sri Lanka, with its varied provinces and differing infrastructure, this issue can offer important perspectives on the accessibility and impact of e-commerce services nationwide. Comprehending regional tastes and actions enables companies to implement specific enhancements.
4. How often do you use e-commerce platforms to shop online? = This question gauges how often respondents engage in online shopping, offering insights into their reliance on e-commerce platforms for their buying requirements. Comprehending shopping frequency aids in determining if e-commerce platforms are essential to a customer's shopping behavior or if they simply use them occasionally, informing strategies for customer retention and interaction.
5. How would you rate your overall satisfaction with your most recent online shopping experience? = This Likert scale inquiry serves as an essential gauge for assessing overall customer satisfaction. The survey collects customer sentiment and their satisfaction with the e-commerce platform by requesting respondents to evaluate their recent experiences. This

information is essential for sentiment analysis, enabling businesses to recognize satisfaction trends and highlight areas that need enhancement.

6. What was the primary reason for your rating? = This open-ended inquiry enables respondents to articulate the factors influencing their satisfaction or dissatisfaction. Comprehending the elements that resulted in a favorable or unfavorable rating like shipment delays, product quality, or customer support offers actionable insights for companies. Tackling these concerns can assist e-commerce platforms in improving their services and more effectively fulfilling customer expectations.
7. How important are the following factors when shopping online? = This inquiry assesses the importance of multiple elements that impact online shopping choices, including product quality, delivery speed, price affordability, and customer service. By evaluating these factors, companies can determine which elements of the shopping experience require enhancement or focus on their marketing efforts to more effectively address customer needs.
8. How likely are you to recommend this platform to others? = The Net Promoter Score (NPS) is a commonly employed measure to assess customer loyalty. This inquiry assesses the likelihood of respondents recommending the platform to others, indicating their overall satisfaction and confidence in the e-commerce site. Elevated NPS scores reflect solid customer loyalty, whereas low scores indicate regions where the platform might require enhancement to boost customer retention and support.
9. Have you encountered any negative experiences with this platform? If yes, please explain. = This open-ended question encourages respondents to discuss any adverse experiences they have faced, providing companies with a clearer and more precise insight into the challenges within the customer journey. Regardless of whether it's inadequate customer service, delivery problems, or technical difficulties with the platform, this feedback can highlight areas for enhancement and address persistent issues.
10. What improvements would you like to see in Sri Lankan e-commerce platforms? = This concluding open ended question gives respondents the opportunity to share their ideas for improving e-commerce platforms in Sri Lanka. Gathering this type of feedback is essential for companies aiming to innovate and maintain competitiveness in the marketplace.

Recommendations could involve enhancements in product variety, shipping speed, or user experience, and these observations will assist companies in matching their offerings with client expectations and new developments.

The survey aims to collect data across various dimensions, offering an in-depth insight into customer feelings and experiences in relation to Sri Lankan e-commerce. The survey employs a combination of quantitative and qualitative questions, guaranteeing the collection of both objective data and subjective opinions, which allows for a detailed assessment of customer satisfaction.

The demographic questions act as essential data points to classify respondents based on aspects such as age, gender, and geographic region. This division enables a thorough examination of how these traits might affect shopping habits, including the frequency of online purchases, favored product categories, and geographical variations in consumer preferences. Grasping these demographic patterns is crucial for customizing marketing approaches and improving the e-commerce experience for various customer segments. For instance, younger clients might hold distinct expectations for delivery timing in contrast to older shoppers, whereas regional differences could necessitate tailored approaches for e-commerce companies.

Questions about the frequency and intensity of online shopping behavior offer important insights into consumer habits. By analyzing the frequency with which consumers utilize e-commerce platforms, companies can more accurately assess customer loyalty and identify which customers might require extra motivation to make repeat purchases. This additionally aids in recognizing the different degrees of interaction with e-commerce platforms, ranging from daily shoppers to those who make purchases only occasionally. The frequency of visits may affect marketing strategies, including promotional campaigns or targeted ads that seek to enhance engagement.

At the heart of this survey is sentiment analysis, featuring questions aimed at gauging customer satisfaction across a range, from extremely dissatisfied to extremely satisfied. Employing a Likert scale offers a distinct quantitative assessment of satisfaction, whereas open-ended questions enable participants to elaborate on the rationale for their ratings. These explanations provide important qualitative insights into the factors that lead to either positive or negative experiences. For instance, a customer might evaluate their experience as “very dissatisfied” because of delivery delays, offering valuable insights for companies to enhance their logistics. The use of Likert scales alongside open-ended feedback aids in evaluating customer sentiment polarity, facilitating the

classification of feedback into positive, negative, or neutral categories. By grasping the main factors influencing customer sentiment, companies can focus on enhancing aspects that most significantly affect satisfaction.

Regression and correlation analysis are supported by inquiries that investigate the significance of elements such as product quality, delivery speed, pricing, and customer service. These inquiries enable an assessment of how each of these factors connects to overall customer satisfaction and the probability of suggesting the platform to others. For instance, if it is determined that product quality is the key element affecting satisfaction, companies might opt to enhance their product range. On the other hand, if delivery speed shows a stronger link to customer satisfaction, the company might focus on quicker shipping choices. The outcomes of regression analysis can assist businesses in recognizing the key factors and making informed, data-based decisions to improve the overall customer experience.

The open ended inquiries regarding unfavorable experiences and recommendations for enhancement are essential for gathering comprehensive feedback that may not be obtained through only structured questions. By enabling customers to report instances of dissatisfaction, companies can pinpoint ongoing problems and implement solutions to address them. These insights also create an opportunity for businesses to connect with customers directly, providing tailored responses or solutions to issues, thereby enhancing customer loyalty. Additionally, grasping customer expectations for enhancements allows e-commerce platforms to align their offerings with the requirements and preferences of their target audience. For example, if clients often highlight the necessity for improved product selection or quicker delivery, these issues can be given priority in upcoming enhancements to the platform.

In conclusion, the integration of demographic, behavioral, and opinion-focused questions guarantees that the gathered data is both thorough and usable. By combining quantitative data, such as customer satisfaction scores and preferences, with qualitative insights from open ended feedback, companies can create a comprehensive understanding of customer experience. This strong data will allow e-commerce platforms in Sri Lanka to recognize trends, relationships, and areas needing enhancement, ultimately resulting in a more customized and effective shopping experience for their users.

1.7. Hypothesis

This study presents two hypotheses aimed at examining the effects of big data sentiment analysis on customer satisfaction and engagement in Sri Lanka's e-commerce industry. The main hypothesis (H1) posits that sentiment analysis positively impacts customer experience, allowing businesses to gain actionable insights and enhance their services. Conversely, the null hypothesis (H0) proposes that sentiment analysis lacks a significant effect on customer satisfaction and engagement, while other variables may be more influential. These hypotheses form the basis for assessing the genuine impact of sentiment analysis on developing customer relationships and its efficiency in Sri Lankan e-commerce. Through the examination of these hypotheses, the research aims to evaluate if sentiment analysis is an essential instrument for improving customer experiences or if it is still an ineffective or inadequate approach in this scenario.

1. Primary Hypothesis (H1) = The main hypothesis (H1) asserts that utilizing big data sentiment analysis greatly improves customer satisfaction and engagement in the e-commerce industry of Sri Lanka. This hypothesis is based on the idea that sentiment analysis enables companies to gain a more profound insight into their customers' perceptions, feelings, and opinions through the evaluation of feedback, reviews, ratings, and social media interactions. When businesses can assess large volumes of customer feedback instantly, they can recognize common patterns, pinpoint dissatisfaction, and promptly tackle issues of concern. For instance, if sentiment analysis indicates that customers frequently express dissatisfaction with delayed deliveries, companies can promptly act to enhance their logistics or improve communication with customers about shipping schedules. By swiftly addressing these concerns, companies showcase their dedication to customer satisfaction, which can greatly enhance customer loyalty and trust.

Additionally, sentiment analysis supports the personalization of marketing strategies. Insights gained from sentiment analysis can aid companies in designing customized marketing campaigns that align with customer preferences. For example, if sentiment analysis reveals that a segment of customers is notably pleased with a particular product aspect, companies can highlight this aspect in upcoming marketing materials aimed at similar customer segments. Tailored campaigns that align with customer preferences have proven to boost satisfaction, engagement, and overall retention rates. Through tailored interactions and personalized product offerings, businesses are addressing issues while also establishing a stronger emotional

bond with their customers, promoting lasting loyalty. This hypothesis indicates that the blend of prompt feedback reactions and tailored interaction through sentiment analysis is crucial for enhancing customer satisfaction in the competitive realm of Sri Lankan e-commerce.

2. Null Hypothesis (H0) = The null hypothesis (H0) questions the efficacy of big data sentiment analysis by asserting that its use does not impact customer satisfaction and engagement in Sri Lanka's e-commerce industry. This hypothesis acts as a benchmark to evaluate if the perceived advantages of sentiment analysis are just coincidental or if other elements are more significant in influencing customer satisfaction. For example, elements like product variety, competitive pricing, quicker delivery, or improved customer service might have a greater impact on customer satisfaction than insights gained from sentiment analysis. In this scenario, although companies may utilize sentiment analysis, the enhancements in customer satisfaction could be linked to external influences rather than the sentiment data itself.

The null hypothesis emphasizes the possible constraints of sentiment analysis. Sentiment analysis depends on natural language processing (NLP) algorithms, which do not consistently excel in understanding customer sentiment, especially in areas such as Sri Lanka, where cultural and linguistic diversity poses difficulties. Misunderstandings can arise from different interpretations of words and phrases, complicating the ability of sentiment analysis tools to effectively evaluate the feelings or emotions reflected in customer feedback. Moreover, if the insights produced by sentiment analysis are not practical or if companies do not respond suitably, the expected enhancements in customer satisfaction and engagement might not occur. For instance, if a business recognizes a negative sentiment but fails to respond to it promptly or efficiently, the sentiment analysis won't yield the expected results. Thus, the null hypothesis indicates that sentiment analysis, despite being a strong tool, may not consistently result in the expected improvements in customer experience without appropriate interpretation, resources, and a dedicated organizational effort to tackle customer issues.

Essentially, the null hypothesis requires a thorough assessment of the actual effect of sentiment analysis on customer engagement and satisfaction, ensuring that any positive changes noted can be reliably linked to the implementation of sentiment analysis instead of other operational elements. It questions the belief that sentiment analysis on its own is enough to enhance customer experience, highlighting the necessity for a comprehensive strategy in tackling customer satisfaction.

1.8. Thesis structure

The organization of this thesis is systematically designed to offer a comprehensive investigation of the research issue and its background. Every chapter serves a particular function that together aids in tackling the research goals and hypotheses. This part describes the progression of the research, beginning with the groundwork and culminating in the presentation of the results and the formulation of significant conclusions. The thesis is organized in a way that addresses both the theoretical and empirical aspects of the study while also offering practical suggestions aimed at enhancing customer satisfaction and engagement within Sri Lanka's e-commerce industry. The upcoming subsections give a comprehensive summary of every chapter and its material.

CHAPTER 1 - Introduction

The Introduction chapter sets up the foundation for the study by offering contextual details on the research topic, pinpointing the issue, and clarifying the research goals. This chapter begins by discussing the importance of customer satisfaction in e-commerce, especially in the Sri Lankan market, and how companies are increasingly leveraging big data and sentiment analysis to enhance customer experiences.

The chapter continues by clearly outlining the research issue: while the e-commerce sector in Sri Lanka is growing, there is restricted research on how sentiment analysis could improve customer satisfaction and engagement. This study seeks to address this gap by investigating how sentiment analysis, as a big data tool, can enhance customer service and overall experience. Additionally, the chapter presents the research questions and hypotheses, highlighting the main hypothesis that big data sentiment analysis has a positive effect on customer satisfaction and engagement. The null hypothesis is also presented, proposing that sentiment analysis might not have a substantial effect on these results.

In conclusion, this chapter offers a concise outline of the thesis's structure, summarizing what each chapter contains to guide the reader. This aids in grasping the development of the study from its introduction to its conclusion.

CHAPTER 2 - Literature Review

The Literature Review section thoroughly examines and integrates current studies on essential ideas, theories, and results related to sentiment analysis, customer satisfaction, and e-commerce. This chapter establishes the theoretical basis for the study by exploring in detail how sentiment analysis has been applied across various sectors, especially in online retail and e-commerce.

The chapter opens with a summary of sentiment analysis, detailing its primary methods, instruments, and uses across different industries. It examines how sentiment analysis helps in comprehending customer opinions, feedback, and general feelings about products and services. Various forms of sentiment analysis, including textual, visual, and vocal sentiment analysis, are examined, in addition to the methods employed for extracting and evaluating sentiment from extensive datasets, featuring Natural Language Processing (NLP) and machine learning algorithms.

Subsequently, the literature review explores earlier research that has assessed the effects of sentiment analysis on customer satisfaction within e-commerce. The chapter examines various international and local research, highlighting trends, challenges, and gaps in the findings. It explores how sentiment analysis has been employed to shape business choices, boost customer service, and refine product offerings.

The chapter additionally examines the obstacles encountered by companies in Sri Lanka, such as cultural and linguistic elements that could influence the precision of sentiment analysis. For instance, the existence of various languages (Sinhala, Tamil, English) and local dialects in Sri Lanka complicates sentiment analysis, which is extensively analyzed in the literature.

In conclusion, the chapter highlights deficiencies in the existing research that this study seeks to tackle. It additionally links the literature review to the research goals, validating the significance of the research and its aim to enhance current understanding in the areas of e-commerce and sentiment analysis.

CHAPTER 3 - Methodology

The Methodology chapter offers an in-depth description of the research framework, techniques, and instruments employed to gather and examine data. This chapter acts as a guide for the study's execution, guaranteeing that the research is strong, credible, and repeatable.

The chapter starts with describing the research methodology, which employs a mixed-methods design that incorporates both qualitative and quantitative data. This method facilitates an in-depth investigation of the research issue from various angles, guaranteeing that both quantitative information and personal experiences are included.

The chapter outlines the methods used for data collection, featuring surveys that collect demographic information, customer opinions, and sentiment assessments. It outlines the reasoning for choosing survey participants and the sampling technique employed to guarantee a varied and representative sample.

Additionally, the chapter describes the analysis of the gathered data through sentiment analysis tools and statistical techniques. It details the utilization of software like R or Python for sentiment classification, along with employing regression analysis and correlation analysis to investigate the relationships among various variables (e.g., product quality, delivery speed, and customer satisfaction).

The chapter furthermore addresses the ethical aspects related to the research, such as obtaining informed consent from participants and ensuring the safeguarding of privacy and confidentiality. The study recognizes its limitations, including possible biases in self-reported data and difficulties in correctly interpreting sentiment from customer feedback. The methodology section guarantees clarity regarding the research process and supports the reliability of the findings.

CHAPTER 4 - Presentation of Results

The Results Presentation chapter focuses on displaying the research findings in a clear and organized way. This chapter presents the information gathered from surveys and sentiment analysis, offering both qualitative and quantitative findings that respond to the research questions and hypotheses.

The chapter starts with a summary of the main results from the survey data, categorized into areas like customer satisfaction rates, elements affecting customer experience, and the occurrence of negative experiences. It shows the information through descriptive statistics, such as frequency distributions, means, and standard deviations, offering insights into the overall trends noted among the participants.

Subsequently, the findings from the sentiment analysis are displayed, illustrating the sentiment polarity (positive, negative, or neutral) of customer responses. Visual depictions, like sentiment heatmaps and word clouds, are employed to depict the sentiment patterns in customer feedback. This segment also addresses the main factors influencing customer sentiment, pinpointing recurring themes like delivery delays, product quality, and customer service.

The section additionally reveals the outcomes of the regression and correlation analysis, which investigates the connections among various factors influencing customer satisfaction. For instance, it could show that product quality is more closely linked to customer satisfaction compared to delivery speed, or that significant dissatisfaction with customer service adversely impacts customer loyalty.

Ultimately, the chapter addresses any surprising discoveries, irregularities, or constraints noted during the examination. The results chapter is essential because it enables the reader to assess if the research hypotheses are validated by empirical data and offers a clear understanding of the elements affecting customer satisfaction in Sri Lankan e-commerce.

CHAPTER 5 - Conclusions and Recommendations

The Conclusions and Recommendations chapter delivers an extensive overview of the research results, concluding on the impact of sentiment analysis in enhancing customer satisfaction and engagement in the e-commerce sector of Sri Lanka.

The chapter starts by reviewing the research questions and hypotheses and summarizing the main findings from the earlier chapter. It examines the significance of these results for academic theory and real-world application, emphasizing how sentiment analysis aids in recognizing customer requirements, refining product offerings, and boosting customer support.

The section subsequently offers practical suggestions for online businesses in Sri Lanka. These suggestions could involve utilizing sentiment analysis tools, enhancing data collection methods, tackling frequent customer issues (such as delivery delays or inadequate service), and leveraging sentimental insights to tailor marketing approaches. The suggestions seek to assist companies in utilizing customer input to enhance their processes and boost customer loyalty.

Moreover, the chapter discusses the study's limitations and proposes avenues for further research. It suggests possible avenues for examining sentiment analysis in different sectors, like hospitality or tourism, or studying how cultural disparities affect sentiment understanding.

In summary, the chapter connects the research goals, results, and suggestions, providing a concluding reflection on the study's impact on improving customer experience in Sri Lankan e-commerce. The suggestions provide actionable advice for businesses aiming to enhance customer satisfaction, engagement, and loyalty by utilizing big data sentiment analysis.

Collectively, these sections constitute the foundation of the thesis, leading the reader along the research path, from the introduction to the conclusions. The study adds to e-commerce by demonstrating how big data sentiment analysis can enhance customer satisfaction and engagement, particularly in the context of Sri Lanka.

CHAPTER 2 - LITERATURE REVIEW

2.1. Literature Review

The decision to conduct a literature review in this study stems from its ability to provide a solid theoretical foundation and contextual understanding of the topic. By examining recent studies, reports, and scholarly articles, the literature review aids in identifying key concepts, trends, and gaps related to big data sentiment analysis and its impact on enhancing customer experience in the e-commerce sector. It allows research to build upon existing knowledge, ensuring that the study aligns with current trends and methodologies. In addition, a literature review allows for a thorough evaluation of methods and findings from prior studies, supporting the improvement of the research approach and validating its importance. Additionally, it aids in formulating well-informed research inquiries and hypotheses by utilizing insights from diverse perspectives, including international case studies and regional assessments. Ultimately, this method enhances the dependability and

depth of the research by ensuring it is based on evidence-based insights and shaped by earlier academic and practical contributions.

1. Empowering E-commerce in Sri Lanka through Customer Insights by *Ruwan Jayathilaka & Isuri Udara*

The rapid growth of e-commerce has changed the global market, enabling businesses to transcend geographic barriers and connect with consumers worldwide. This major transformation has been propelled by enhancements in internet accessibility, digital advancements, and the extensive adoption of smart devices. E-commerce, once a niche industry, has become a dominant force in the global economy, with retail e-commerce revenues projected to increase from USD 5.717 trillion in 2022 to USD 8.148 trillion by 2026. However, as the industry continues its swift growth, it faces numerous challenges, particularly in security. Cybercrime, escalating by 600% since the COVID-19 pandemic, poses a substantial threat to e-commerce platforms, with attacks encompassing phishing, spamming, sophisticated SQL injections, and malware attacks. These security breaches lead to financial losses amounting to billions annually and erode customer trust, an essential factor for the success of e-commerce businesses.

In countries like Sri Lanka, where e-commerce is gaining popularity from increasing internet availability, challenges are worsened by a cash-centric economy and low levels of digital literacy. While the Sri Lankan e-commerce sector shows significant growth potential, projected to reach revenues of USD 3,017.8 million by 2028, consumer concerns about online transactions and platform security remain a significant barrier. Research indicates that consumers in Sri Lanka are significantly influenced by their price sensitivity and prefer traditional cash-on-delivery methods, reflecting a lack of trust in digital payment solutions. Moreover, the feeling of safety, or lack thereof, greatly affects customer engagement with e-commerce websites. Studies have highlighted that factors such as transactional security, privacy, vendor system safeguarding, and platform quality significantly influence customer trust and subsequently, the adoption of e-commerce services. Addressing these issues requires an all-encompassing approach, involving the development of robust security structures, implementing advanced cybersecurity methods, and promoting customer awareness and confidence. By focusing on these elements, the e-commerce industry in Sri Lanka and similar developing markets can achieve its full potential, fostering economic growth and enhancing customer experiences in an increasingly digital landscape.

2. Exploring the Digitalization Journey of Small and Medium-Sized Enterprises in China and Sri Lanka's Thriving E-commerce Landscapes

By Qiubo Huang & P.J. Kumarasinghe

This study investigates the various factors that hinder and facilitate the globalization efforts of Small and Medium-sized Enterprises (SMEs) in China and Sri Lanka through e-commerce platforms. By employing qualitative research methods, the study explores SMEs in the manufacturing industry that are seeking global expansion, utilizing narrative and thematic analysis to uncover shared patterns and distinctive characteristics in their digital transformation journeys. The findings show that Chinese SMEs highly advantage from advanced technological settings, supportive government policies, and the international attractiveness of their products, all of which enhance their competitiveness in global markets.

In contrast, Sri Lankan SMEs face considerable challenges, including limited access to financial resources, technology infrastructure, and regulatory support, which impede their ability to effectively leverage e-commerce for global expansion. Additionally, cultural and regulatory perspectives vary between the two countries, influencing their approaches to innovation and global commerce. Although China has built a strong infrastructure that encourages innovation, Sri Lanka's small and medium-sized enterprises face challenges in implementing these practices due to reasons like limited resources and lack of governmental backing. This comparison underscores the crucial necessity for governmental actions and infrastructure upgrades in developing economies such as Sri Lanka. The study highlights the growing significance of sustainable and socially accountable business practices that SMEs can incorporate into their e-commerce strategies to address global market demands. By analyzing the distinct experiences of SMEs in two varying economic and digital contexts, the research provides valuable insights for policymakers, researchers, and practitioners seeking to enhance SME development and foster global collaborations in the digital era. In summary, this comparative research contributes to the broader conversation on SME digitalization, offering actionable recommendations to address the challenges faced by emerging markets while capitalizing on the successes of developed digital economies.

3. How Sri Lanka's Banking Sector is Leveraging AI-Powered Applications for Enhancing Customer Experience

By Ududeniy Gedara Dilusha & Wasala Nalinda

This study explores the role of Artificial Intelligence (AI) in enhancing customer experience within Sri Lanka's banking sector. As AI increases its worldwide influence in transforming industries, its incorporation into banking has become increasingly vital, particularly for improving customer engagement and operational efficiency. The research investigates the use of AI-powered applications to improve customer experience in Sri Lankan banks, operating within a competitive and changing environment influenced by regulatory structures and economic challenges.

The research investigates the various types of AI applications employed in Sri Lankan banks, such as chatbots, fraud detection tools, predictive analytics, and automated customer service. These AI tools have been shown to enhance efficiency, minimize human mistakes, and boost overall customer satisfaction. Nonetheless, the study also highlights significant challenges including compatibility problems, cybersecurity threats, budget limitations, and a lack of qualified experts who can operate AI systems. Despite these challenges, the study emphasizes the encouraging potential for AI implementation in Sri Lankan banking, providing recommendations for tackling these issues. For example, funding AI training initiatives, partnering with global tech companies, and supporting government-led projects can speed up the integration of AI. The results of this research offer perspectives on both the present condition of AI in the banking industry of Sri Lanka and the prospective opportunities for AI-driven innovations to improve customer experience. This study is important for policymakers, banking experts, and researchers who are keen on the transformative potential of AI in emerging markets.

4. Identifying Ways of Delivery Performance Enhancements in Sri Lankan E-commerce Logistics Sector with Special Reference to the Cash on Delivery (COD) Method

By Navodika Karunaratne

Owing to the rapid advancement of modern technologies, e-commerce delivery services in Sri Lanka have experienced notable expansion. Many logistics and courier companies in Sri Lanka have adopted modern platforms to address the evolving demands of the expanding e-commerce industry. Nevertheless, despite this growth, the lack of thorough research on the last mile (LM) delivery process has led to negative reputations and extensive customer dissatisfaction. A major issue in Sri Lankan e-commerce logistics is the prevalence of Cash on Delivery (COD), resulting in higher cash collection and customer refusal rates, consequently escalating costs for both e-commerce retailers and delivery services.

Despite the rapid growth of e-commerce, logistics companies in Sri Lanka continue to lag international standards, particularly in the effectiveness of their COD processes. This inefficiency leads to bottlenecks in the entire delivery process, hindering the ability to fulfill customer demands for quick and dependable delivery. In this context, the study explores strategies to enhance the efficiency of COD services, which continue to be a favored payment choice for numerous Sri Lankan shoppers. The research examines best practices in e-commerce logistics on both international and regional levels, recognizing how different markets have effectively optimized their COD processes via technological advancements, better infrastructure, and improved customer interaction strategies. Additionally, the study highlights the necessity of optimizing the COD process by employing methods like electronic payment options, real-time monitoring, and incorporating customer feedback to enhance overall delivery efficiency. By conducting a thorough analysis, the research presents a theoretical framework for COD services aimed at boosting the effectiveness of Sri Lankan e-commerce, offering practical measures to improve the efficiency of the logistics sector.

2.2. Conceptual framework

The conceptual framework for this research is found on the incorporation of big data sentiment analysis into e-commerce platforms to improve customer experience. The framework seeks to emphasize the significance of evaluating customer feedback and employing sentiment analysis to gather insights into levels of customer satisfaction. Sentiment analysis employs sophisticated natural language processing (NLP) and machine learning methods to analyze textual information, including product reviews, customer feedback, and social media engagements. Through the examination of these types of feedback, companies can assess the feelings conveyed by customers, sorting them into positive, negative, or neutral categories. These insights can subsequently be used to evaluate customer sentiments regarding a product or service, thus establishing a direct connection to customer satisfaction.

Integrating big data tools is a key component of this conceptual framework. Instruments like Hadoop, Spark, and several cloud-based solutions enable companies to handle large volumes of data swiftly and effectively. These tools are especially beneficial for e-commerce businesses managing large volumes of customer interactions on various platforms. Using big data technologies, companies can perform real-time sentiment analysis, obtaining a current awareness

of customer feelings. This allows them to quickly tackle new issues or take advantage of favorable responses. Big data tools also offer the required scalability to manage increasing amounts of customer feedback, enhancing the accessibility and practicality of sentiment analysis for businesses of various sizes.

An essential aspect of this framework is recognizing sentiment trends throughout the years. By utilizing sentiment analysis, e-commerce platforms can identify persistent trends in customer feedback, which can be vital for making decisions. For instance, if clients regularly voice discontent with a specific product characteristic or service element, companies can implement corrective measures to tackle those problems. On the other hand, if positive feedback patterns are noted in certain areas, companies can emphasize those aspects in their marketing approaches. Recognizing these trends not only guides business strategies but also enables companies to predict upcoming customer needs and preferences. The ability to predict through sentiment analysis offers companies a distinct advantage in enhancing customer satisfaction.

After identifying these sentiment trends, businesses can adjust their marketing strategies accordingly. By utilizing the insights obtained from sentiment analysis, companies can customize their marketing strategies to meet the needs and worries of their customers. For instance, if clients show great contentment with a product's delivery speed, companies can highlight this feature in upcoming marketing efforts to draw in additional customers. Conversely, when a product or service encounters consistent negative feedback about pricing or quality, marketing strategies can be modified to tackle these concerns, fostering customer trust and enhancing engagement. By synchronizing marketing approaches with customer feelings, e-commerce sites can cultivate deeper connections with their audience and enhance overall customer loyalty.

The research additionally contrasts the use of sentiment analysis in Sri Lankan e-commerce sites with well-known global platforms such as Amazon and Alibaba. These international powerhouses have effectively incorporated sentiment analysis into their business frameworks to improve their customer experience and marketing tactics. The study establishes a benchmark for Sri Lankan businesses to follow by examining how these platforms utilize sentiment data to enhance their services continuously. The comparison emphasizes the ways in which global platforms have utilized sentiment analysis to tailor customer interactions, anticipate customer needs, and enhance service offerings. This comparison acts as a reference for Sri Lankan companies to comprehend how sentiment analysis can be successfully applied in their specific situation.

Here's a deeper explanation of the key elements in the framework:

1. Sentiment Analysis and Customer Feedback:

- Sentiment analysis involves the use of natural language processing (NLP) and machine learning techniques to analyze textual data. It helps businesses gauge the emotions expressed by customers in their feedback (e.g., reviews, comments, surveys).
- By extracting sentiment from this data, companies can measure customer satisfaction, which can directly influence their strategy for improving products or services. Positive sentiment suggests satisfaction, while negative sentiment can highlight areas needing improvement.

2. Utilization of Big Data Tools:

- Big data tools such as Hadoop, Spark, and cloud platforms allow businesses to process large datasets efficiently. In the context of sentiment analysis, big data tools can handle the vast amounts of feedback that customers provide across different channels.
- These tools also support real-time analysis, enabling businesses to act quickly on emerging customer trends or concerns.

3. Trend Identification:

- Through the application of big data tools, e-commerce platforms can identify recurring sentiment trends over time. For instance, if a product consistently receives negative reviews due to poor quality, the business can take corrective action. Similarly, identifying positive sentiments around specific features of a product can guide marketing efforts.

- Trend analysis helps businesses forecast customer needs, which enables them to better tailor their offerings and improve customer satisfaction.

4. Modification of Marketing Strategies:

- The insights gained from sentiment analysis allow e-commerce businesses to refine their marketing strategies. For example, if sentiment analysis reveals that customers are highly satisfied with product delivery times, businesses can highlight this in their marketing campaigns.
- Conversely, if customers express dissatisfaction with certain aspects (e.g., payment methods, product quality), businesses can adjust their marketing messages to address these concerns, ultimately building trust and fostering long-term relationships with customers.

5. Comparison with Global Platforms:

- The study compares Sri Lankan e-commerce platforms with global giants like **Amazon** and **Alibaba** to understand how sentiment analysis is applied in mature markets. These platforms use advanced sentiment analysis to continuously enhance their offerings, tailor recommendations, and personalize marketing messages.
- By looking at these successful international examples, the study provides a benchmark for Sri Lankan platforms to apply similar techniques and methods for improving customer experience.

6. Linking Sentiment Analysis, Customer Satisfaction, and Marketing:

- The goal of the conceptual framework is to demonstrate how sentiment analysis, when applied effectively, can enhance **customer satisfaction** and lead to more targeted **marketing strategies**.
- The framework establishes that by analyzing and responding to customer sentiments, businesses can create a more personalized experience, improving

engagement and fostering loyalty, which in turn leads to increased customer retention and long-term business growth.

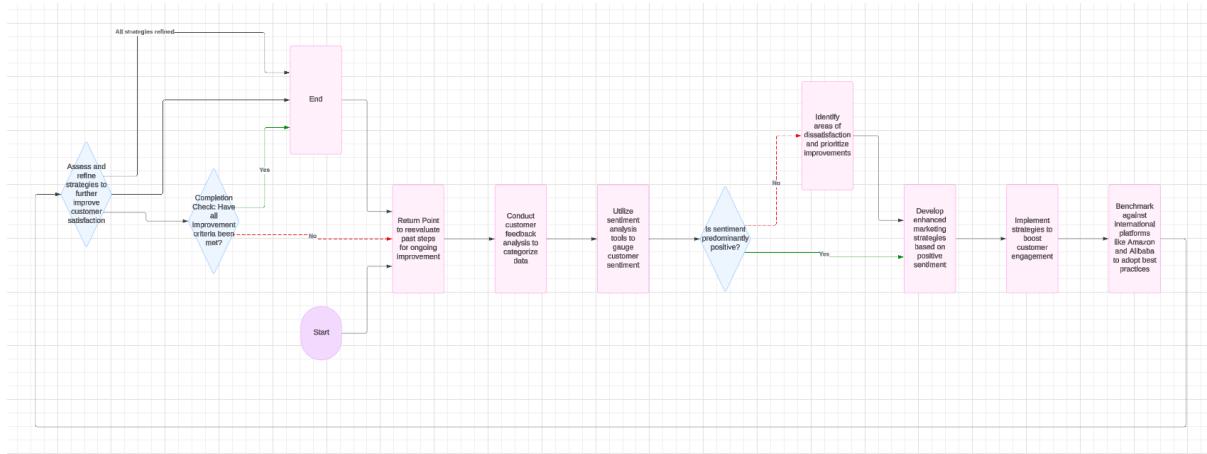


Figure 1 Conceptual Framework

In the end, the conceptual framework aims to illustrate the link between sentiment analysis, customer satisfaction, and marketing approaches. Through constant evaluation of customer feelings and adjusting to their likes, e-commerce sites can develop a more individualized and customer-focused experience. This tailored strategy not only boosts customer interaction but also cultivates customer loyalty, resulting in higher retention rates and ongoing business expansion. The incorporation of big data sentiment analysis enables companies to anticipate customer demands, fostering a setting where customers feel acknowledged and appreciated, which ultimately results in lasting success for the e-commerce platform.

CHAPTER 3 – METHODOLOGY

3.1. Research philosophy

While conducting research on the potential of big data sentiment analysis to enhance customer experience in Sri Lanka's e-commerce industry, the positive research philosophy has been chosen as the foundational framework for the study. Positivism, grounded in the conviction that reality can be measured and comprehended objectively, centers on observable, quantifiable events. It is especially appropriate for research involving extensive datasets and data-centric insights, like sentiment analysis, aiming to reveal patterns and trends grounded in empirical evidence. This research philosophy enables the investigation to systematically examine how customer feedback gathered from different e-commerce platforms can be leveraged to obtain actionable insights that

assist businesses in improving their marketing strategies, customer interaction, and overall experience. This study adopts an objective and scientific approach through positivism, leading to dependable, reproducible, and data-driven findings.

1. Positivism in Research Philosophy = The positive aspect of the research centers on gathering and examining unbiased data, emphasizing occurrences that are both quantifiable and detectable. The analysis of emotions in large datasets naturally aligns with positivism as it relies on computational methods to extract sentiment from client reactions. This investigative method allows for the organized assessment of consumer contentment and involvement by analyzing large sets of data, comprising customer evaluations and responses from online shopping sites in Sri Lanka. Certainly! However, it seems you only provided the word "By." Could you please provide additional text or context that you'd like me to paraphrase? By employing a methodical, scientific strategy, the research can reveal trends, links and patterns in customer emotions. For instance, sentiment ratings can be linked to vital performance indicators such as Net Promoter Score (NPS) and customer loyalty levels, providing unique and consistent perspectives on the impacts of sentiment evaluation in the online retail sector. The optimistic method ensures that the outcomes stay devoid of personal biases, aiding the generation of proof motivated recommendations for performing sentiment analysis.

2. Interpretivism in Research Philosophy = Interpretivism improves the positivist approach by emphasizing the personal and situational aspects of the research. In the framework of E-commerce in Sri Lanka is greatly affected by cultural and regional influences consumer perceptions and behaviors. Interpretivism allows researchers to explore these nuances by examining the significances and understandings that both clients and companies allocate to feelings. For example, the study recognizes that the linguistic diversity in Sri Lanka, encompassing Sinhala, Tamil, and English, generates difficulties in accurately assessing customer feedback. Employing qualitative methods like interviews with online business owners and consumer questionnaires, the research examines the challenges and opportunities associated with implementing sentiment analysis in a culturally diverse environment. This philosophical perspective ensures that the investigation captures the profound, individual experiences of stakeholders, often overlooked in purely numerical studies.

3. Combining Philosophies for a Holistic Approach = This study adopts a pragmatic approach that merges positivism and interpretivism, resulting in a link between numerical data and contextual understanding. While the quantitative aspects of sentiment analysis form the foundation for identifying trends and relationships, the qualitative insights derived from interpretivist approaches enrich the analysis by adding depth and context. This integrated approach is particularly beneficial in addressing complex research objectives, like evaluating the influence of sentiment analysis on marketing tactics and contrasting regional practices with international benchmarks.

The positive aspect of this study focuses on gathering and examining objective data, which is crucial when addressing the large volumes of information produced by sentiment analysis. In this scenario, sentiment analysis is a computational process that converts unprocessed customer feedback such as online reviews, social media remarks, or product ratings into quantifiable sentiment scores. These scores, usually classified as positive, negative, or neutral sentiments, offer a tangible depiction of customer satisfaction. The research will utilize diverse big data tools and methodologies to perform this analysis, exploring large datasets to uncover patterns, correlations, and trends in customer sentiments. This method emphasizes measurable data, enabling the research to yield consistent outcomes and reducing the impact of personal biases or subjective interpretations during the analysis.

The positive method also aids in recognizing essential performance indicators like Net Promoter Score (NPS), customer retention percentages, and customer engagement rates, which can be closely associated with sentiment analysis results. For instance, sentiment scores obtained from customer reviews can be examined to determine how positive or negative feelings relate to the chances of customers suggesting a business to others (NPS). Using statistical techniques, research can quantify the strength of these connections, offering important insights into the influence of sentiment analysis on consumer behavior and, consequently, the e-commerce sector. This comprehensive evaluation of sentimental data allows companies to grasp not only how customers feel but also the reasons some sentiments can enhance or diminish customer loyalty, assisting them in making data-driven decisions that match customer expectations.

Moreover, the positivist approach guarantees that the outcomes of this study are devoid of personal biases and subjective influences. Dependence on quantifiable data and analytical tools removes the possibility of researchers' biases or prior beliefs affecting the understanding of outcomes. This

is essential for performing sentiment analysis, as the goal is to obtain insights that can be applied to a wider range of customers, rather than just particular individuals or instances. This research intends to generate evidence-based suggestions for e-commerce companies by maintaining a strong emphasis on data-driven approaches, offering them solid, practical insights that are based on quantitative analysis.

Although the study recognizes the importance of accounting for the contextual and subjective aspects of customer experience, the positive approach is favored for its focus on objective assessment and scientific accuracy. Emphasizing measurable data corresponds with the research's objective of creating standardized, repeatable methods for performing sentiment analysis that can be conveniently reproduced and utilized in different industries or areas. Additionally, by focusing on quantifiable results, this study guarantees that the results will have widespread relevance and enhance the expanding knowledge base in big data and e-commerce.

To conclude, the positivist research philosophy underpins this study's method for exploring how big data sentiment analysis enhances customer experience in the e-commerce industry of Sri Lanka. By concentrating on visible and quantifiable information, the study seeks to reveal important findings that can assist companies in improving their marketing approaches, customer interaction, and overall customer contentment. Employing quantitative methods guarantees that the results of the study are objective, trustworthy, and practical, providing clear suggestions for companies aiming to enhance their customer experience by utilizing sentiment analysis.

3.2. Research approach

This research employs a mixed-methods approach, combining quantitative and qualitative techniques to investigate how big data sentiment analysis can improve customer experience in the e-commerce industry of Sri Lanka. The combination of both approaches provides an in-depth understanding of the research subject by utilizing the advantages of each. The quantitative method facilitates the gathering and examination of measurable data, offering statistical proof of the influence of sentiment analysis on consumer behavior and business performance. Conversely, the qualitative method enhances understanding by examining the contextual and cultural elements that could affect the perception and interpretation of sentiment in the Sri Lankan e-commerce sector. Collectively, these methods offer a comprehensive, well-rounded perspective on how sentiment analysis can enhance customer satisfaction and engagement.

1. Quantitative Approach = This study's quantitative method concentrates on examining extensive datasets of customer feedback collected from different e-commerce platforms in Sri Lanka. This method relies on statistical methods and big data tools to derive significant insights from customer feedback, ratings, and reviews. Sentiment analysis, central to this study, entails assessing the emotional tone present in the text data to classify it as positive, negative, or neutral. Using computational tools, research can measure customer sentiment and analyze its relationship with important performance metrics (KPIs) like Net Promoter Score (NPS) and customer retention rates. For instance, sentiment scores obtained from customer feedback will be examined to assess the general sentiment and its connection to these KPIs. Moreover, regression analysis will be performed to evaluate the strength and characteristics of these relationships, offering empirical proof of the impact of sentiment analysis on e-commerce outcomes. Employing a data-driven and systematic method guarantees objectivity, replicability, and consistency in the analysis, leading to actionable insights that companies can use to improve customer satisfaction.

2. Qualitative Approach = Alongside the quantitative approach, the qualitative method emphasizes obtaining a more profound insight into the contextual elements that influence customer feelings in Sri Lanka's varied e-commerce landscape. This facet of the study highlights interpretive techniques like discussions with entrepreneurs and customer surveys to reveal the nuances underlying the emotions conveyed in feedback. For instance, the study investigates how cultural and regional differences, such as the usage of Sinhala, Tamil, and English languages, might influence customer perceptions and responses. Utilizing qualitative methods, the research explores the fundamental reasons for customers' emotions, illuminating the cultural subtleties that affect their experiences. This qualitative information will enhance the quantitative results by offering contextual understanding, enabling a more thorough analysis of sentiment trends. By combining both data types, the study encapsulates the intricacies of consumer behavior and guarantees that the outcomes mirror the local context, thereby enhancing the significance and practicality of the results.

3. Deductive and Inductive Reasoning = This study utilizes both deductive and inductive reasoning to guarantee a comprehensive and rational analysis of the research issue. Deductive reasoning is employed to evaluate the main hypothesis (H1), which suggests that implementing big data sentiment analysis greatly enhances customer satisfaction and engagement. This line

of thought moves logically from broad theories regarding sentiment analysis and customer satisfaction to the scenario of e-commerce in Sri Lanka. Utilizing recognized frameworks and theories, the research examines if the hypothesis is valid in the local market. Conversely, inductive reasoning is used to create new understandings and theories derived from qualitative results. For example, qualitative interviews could uncover how particular cultural elements affect the efficacy of sentiment analysis, resulting in the creation of novel theories suited to the Sri Lankan environment. This blend of both methods enables the research to validate current insights while also adding fresh viewpoints to the domain of sentiment analysis and e-commerce.

4. Pragmatic Approach to Mixed Methods = This research employs a pragmatic method for mixed methods investigation, emphasizing practical solutions instead of rigidly following a single research philosophy. The pragmatic approach highlights the importance of combining both quantitative and qualitative methods to effectively tackle the research questions. This adaptability enables the research to focus on tangible results like uncovering actionable insights for companies while considering the intricacies of the research problem. By merging the advantages of both methods, the research recognizes that quantitative evaluation yields quantifiable, databased outcomes, whereas the qualitative perspectives deliver depth and context for interpreting the data. This practical strategy guarantees that the research is scientifically sound and relevant to its context, providing a balance between the necessity for statistical proof and the awareness of cultural and operational elements.

5. Sequential Design = To enhance the efficacy of the mixed-methods strategy, the research employs a sequential design. This entails performing quantitative analysis initially to uncover trends and patterns in customer sentiment and its effects on crucial business metrics. The results from the quantitative phase will subsequently form the basis for the qualitative phase, during which the researcher will explore the contextual elements that might clarify these patterns. By adhering to this logical progression, the study guarantees that qualitative approaches enrich and contextualize the quantitative results, offering a more comprehensive grasp of the outcomes. This sequential format guarantees that the study is structured and logical, facilitating a distinct link between the data and its analysis. Additionally, it guarantees that the qualitative findings can confirm and expand on the quantitative patterns, providing a deeper and more complete examination of the research topic.

To sum up, the mixed-methods approach utilized in this research facilitates a thorough exploration of how sentiment analysis can enhance customer experience in e-commerce in Sri Lanka. By integrating the advantages of both quantitative and qualitative methods, the study offers a comprehensive and nuanced examination of customer sentiment, encompassing both quantifiable data and contextual perspectives. This method guarantees that the results are both data-informed and culturally pertinent, providing essential insights for e-commerce companies aiming to improve customer satisfaction and engagement by utilizing big data sentiment analysis.

3.3. Research strategy

The research approach aims to systematically tackle the study's research goals, concentrating on how sentiment analysis of large data sets can improve customer satisfaction in the e-commerce industry of Sri Lanka. By integrating quantitative, qualitative, and data-centric methods, the approach guarantees a thorough understanding of how sentiment analysis contributes to enhancing customer experience. This method balances the necessity for statistical precision with a more profound examination of contextual insights, offering practical and actionable results that can direct e-commerce companies in Sri Lanka. The approach is thoughtfully organized, placing a strong focus on gathering data, analyzing data, and engaging the target audience, which makes it effective in capturing the nuances of customer sentiment and its effects on business performance.

1. Data Collection

- Primary Data = The main data will be gathered via surveys and interviews. The surveys will aim at e-commerce customers, concentrating on gathering feedback about their satisfaction, preferences, and issues related to online shopping experiences. This information will offer understanding of customers' views on service quality, product satisfaction, and general involvement. Comprehensive interviews will be carried out with business stakeholders in the e-commerce industry, including platform owners, marketing managers, and customer service agents. These interviews will offer essential insights into the tactics used by companies, the difficulties they encounter, and their experiences related to sentiment analysis. These two types of primary data will provide both quantitative and qualitative insights, guaranteeing that the research encompasses a wide range of viewpoints on customer satisfaction.

- Secondary Data = Along with primary data, secondary data will be gathered from current customer reviews available on Sri Lankan e-commerce sites like Daraz or Kapruka. These evaluations will be examined to gauge customer feelings, offering a substantial dataset for sentiment analysis. Publicly accessible datasets and literature reviews regarding sentiment analysis methods will also be analyzed to compare Sri Lankan e-commerce performance with global standards. By integrating primary data with secondary data sources, the study will create a comprehensive, multi-faceted dataset for examination.

2. Data Analysis Techniques

- Sentiment Analysis = A key element of this study is sentiment analysis. Employing natural language processing (NLP) methods, customer feedback will be divided into three groups: positive, negative, and neutral. This classification will assist companies in gaining a deeper insight into their clients' emotional reactions to goods and services. Sentiment analysis will be conducted on both primary data (survey answers and interview transcripts) and secondary data (customer feedback) to assess the general sentiment across various platforms.
- Descriptive Analysis = Descriptive analysis will be employed to summarize the responses from the survey and the demographic information. This method will offer a straightforward and brief overview of the participants' backgrounds, encompassing their shopping habits, online activities, and general satisfaction. Descriptive statistics will assist in recognizing trends within the data, like frequent issues or aspects of high satisfaction, providing a more defined understanding of customer experiences.
- Regression Analysis = Regression analysis will be used to evaluate the connections among different factors, including customer sentiment, product attributes, and customer satisfaction. This approach will assist in determining the impact of sentiment on business results, including customer loyalty, retention, and buying behavior. Regression models will be created to assess the relationship between sentiment scores and satisfaction levels, and to determine if positive sentiment is linked to increased retention rates or repeat purchases.
- Correlation Analysis = Correlation analysis will be conducted to investigate the connections between crucial variables, including satisfaction ratings, sentiment patterns,

and business approaches. For instance, the research will examine if there is a link between the degree of positive sentiment in customer feedback and the success of specific marketing campaigns or customer service methods. This examination will offer additional understanding of how sentiment influences overall business strategies and performance.

3. Target Audience

- Customers = The main participants in the study are online shoppers who offer input via surveys. These customers reflect a varied mix of online buyers, providing important insights into their experiences, expectations, and satisfaction levels. The questionnaires will inquire about their experiences with online shopping platforms, covering usability, product quality, customer support, and shipping. This information will serve as the foundation for sentiment analysis and assist in recognizing which elements of the e-commerce experience have the greatest impact on customer satisfaction.
- Business Stakeholders = Along with customers, the research also involves business stakeholders in the e-commerce sector. This encompasses platform owners, managers, and personnel engaged in customer service and operations. Discussions with these stakeholders will yield insights into their views on sentiment analysis, its application, the difficulties encountered in assessing customer sentiment, and how sentiment data affects their business choices. This collaboration with business stakeholders guarantees that the research investigates not only customer perceptions but also analyzes the practical use of sentiment analysis in the e-commerce sector.

4. Tools and Technologies =

The study utilizes a mix of practical resources to aid data gathering, analysis, and reporting. These tools are selected for their dependability and ease of use, guaranteeing the effective implementation of every stage of research.

- Google Forms = Google Forms will be used for gathering data, particularly for creating and sending surveys to e-commerce consumers. Its straightforwardness and ease of use render it a perfect instrument for collecting input from a wide audience. The organized format enables the gathering of both quantitative data (e.g., satisfaction scores) and qualitative data (e.g., open-ended feedback), resulting in a thorough dataset for examination.

- SPSS (Statistical Package for the Social Sciences) = Statistical analysis of the survey data will be conducted using SPSS. The software's strong capabilities facilitate the effective implementation of essential analytical techniques, such as descriptive analysis, regression analysis, and correlation analysis. These techniques will aid in summarizing customer feedback, pinpointing significant trends, and examining connections between factors like satisfaction rates and sentiment patterns. SPSS is especially effective for managing extensive datasets and executing complex statistical computations with precision.
 - MS Word = MS Word will serve as the main tool for documentation and reporting. It will serve to gather the research outcomes, organize the analysis, and deliver suggestions in a clear and professional manner. Its editing and formatting features guarantee that the final report is structured effectively and easy to comprehend.
5. Timeframe and Sequential Steps = The study will adopt a sequential strategy, featuring distinct stages and targets to guarantee that every aspect of the research is finished punctually:

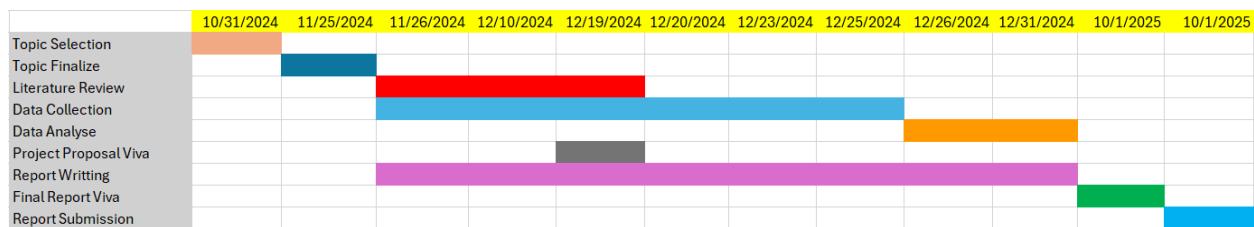


Figure 2 Time Frame

- Phase 1: Data Collection = The initial stage will consist of collecting both primary and secondary information. This encompasses distributing questionnaires, performing interviews with business stakeholders, and gathering customer feedback from e-commerce sites. The objective is to guarantee that a varied collection of data is accessible for later analysis.
- Phase 2: Preliminary Descriptive Analysis = After collecting the data, descriptive analysis will be performed to summarize and illustrate the demographic and feedback information. This stage will assist in recognizing fundamental trends and patterns, offering a preliminary insight into customer satisfaction.

- Phase 3: Application of Sentiment, Regression, and Correlation Analysis = The subsequent stage will consist of utilizing sentiment analysis to classify customer feedback, along with conducting regression and correlation analyses to investigate the relationships among variables. This stage will assist in determining how sentiment affects customer satisfaction and online shopping results.
- Phase 4: Interpretation and Recommendations = The concluding stage will emphasize analyzing the outcomes of the data study and developing suggestions for e-commerce companies. Drawing from the insights obtained, the research will offer actionable recommendations for utilizing sentiment analysis to enhance customer satisfaction and involvement.

Step	Activity	Purpose	Tools/Techniques
1. Data Collection	Surveys and Interviews	Collect primary data from e-commerce customers to gather feedback on satisfaction and preferences, and from business stakeholders to understand strategies and challenges.	Google Forms, MS Word (for structured interview documentation)
2. Data Acquisition	Gathering customer reviews from e-commerce platforms	Compile secondary data for sentiment analysis by collecting reviews, ratings, and feedback from local e-commerce sites.	Python (APIs, web scraping tools like BeautifulSoup)
3. Sentiment Analysis	Classify feedback into positive, negative, or neutral sentiments	Analyze customer opinions to identify patterns and trends in satisfaction and engagement.	SPSS, VADER, or Python NLP libraries (e.g., TextBlob)

4. Descriptive Analysis	Summarize survey responses and visualize demographics	Provide a clear overview of customer demographics, feedback distribution, and satisfaction levels.	MS Excel, SPSS (for visualization and summary)
5. Regression Analysis	Analyze factors influencing customer satisfaction and loyalty	Quantify relationships between sentiment data and key performance indicators (KPIs) such as customer retention, engagement, and Net Promoter Score (NPS).	SPSS
6. Correlation Analysis	Identify relationships between variables	Explore connections between satisfaction, sentiment trends, and business strategies to uncover actionable insights.	SPSS, MS Excel
7. Recommendations	Developing actionable insights	Propose strategies for e-commerce businesses to improve customer experience using big data sentiment analysis.	Analytical summaries in MS Word

Table 1 Research Strategy Table

3.4. Research Choice

The selection of research methodology is essential for the success of any study, as it dictates how data is gathered, analyzed, and understood. This study seeks to explore how big data sentiment analysis can improve customer experiences in Sri Lankan e-commerce, an industry that is quickly developing but encounters specific difficulties in customer satisfaction and retention. To

accomplish this, a combined approach using both quantitative and qualitative methods has been chosen. This decision underscores the necessity to thoroughly examine numeric trends in customer sentiment, while also investigating the personal experiences and difficulties encountered by stakeholders when implementing sentiment analysis tools. Utilizing this mixed-methods approach, the research guarantees a comprehensive view of the problem, integrating statistical strength with contextual insight.

1. Quantitative Approach = The quantitative component of this research emphasizes collecting and examining numerical information to identify measurable patterns and connections. Surveys given to e-commerce customers will gather information on their satisfaction levels, preferences, and issues. The responses will be evaluated through statistical techniques like descriptive analysis, regression analysis, and correlation analysis to measure customer sentiment. Descriptive analysis will deliver a summary of demographic information and satisfaction ratings, providing insight into the general patterns in customer experience.

Regression analysis will determine the main elements affecting satisfaction, including product quality, delivery speed, or user interface design. In the meantime, correlation analysis will investigate the connections between these factors and customer loyalty or engagement. For instance, the analysis could show how favorable customer sentiment relates to repeated purchases or elevated ratings for a particular service. This approach based on evidence guarantees that the study's results are solid, trustworthy, and practical, providing useful insights for e-commerce companies looking to improve customer experiences.

2. Qualitative Approach = The combination of quantitative and qualitative techniques in this research results in a mixed-methods strategy, providing a comprehensive and balanced perspective on customer sentiment in Sri Lankan e-commerce. Quantitative data gathered from surveys and sentiment analysis will reveal major patterns, while qualitative insights from interviews will clarify the fundamental causes of these patterns. This convergence of data enhances the credibility and dependability of the research outcomes by corroborating insights from various sources. For example, although survey results may show a significant desire for quicker delivery times, discussions with stakeholders could uncover logistical issues that hinder this from being realized.

The mixed-methods approach ensures that the research tackles both aspects of the issue measuring customer sentiment and examining the strategic consequences of utilizing sentiment analysis tools. This dual viewpoint renders the research results both statistically valid and practically applicable to the requirements of Sri Lanka's e-commerce sector.

3. Mixed-Methods Approach = The combination of quantitative and qualitative techniques in this research results in a mixed-methods strategy, providing a comprehensive and balanced perspective on customer sentiment in Sri Lankan e-commerce. Quantitative data gathered from surveys and sentiment analysis will reveal major patterns, while qualitative insights from interviews will clarify the fundamental causes of these patterns. This convergence of data enhances the credibility and dependability of the research outcomes by corroborating insights from various sources. For example, although survey results may show a significant desire for quicker delivery times, discussions with stakeholders could uncover logistical issues that hinder this from being realized.

The mixed-methods approach ensures that the research tackles both aspects of the issue measuring customer sentiment and examining the strategic consequences of utilizing sentiment analysis tools. This dual viewpoint renders the research results both statistically valid and practically applicable to the requirements of Sri Lanka's e-commerce sector.

4. Justification for multi-Method Choice = The use of a multi-method strategy is warranted due to its capacity to offer thorough understanding of the research issue. This method integrates numerical data with contextual insights, ensuring that the research reflects both the quantifiable elements of customer sentiment and the personal experiences of stakeholders.

Incorporating various data sources like customer surveys, online feedback, and stakeholder interviews improves the richness and scope of the analysis. For instance, although surveys gather customer viewpoints, interviews provide understanding of the operational difficulties companies encounter when adopting sentiment analysis tools. This variety guarantees that the research results are comprehensive and relevant to practical situations.

Moreover, the multi-method strategy closely corresponds with the research goals. The quantitative aspect aids in assessing customer sentiment and its effect on satisfaction, whereas the qualitative aspect focuses on examining business challenges and opportunities connected to sentiment analysis. Collectively, these approaches offer practical insights that online retail companies can leverage to enhance customer experiences and foster growth.

5. Relevance to the Research Topic = The selected research approach is significantly pertinent to the study's aim of improving customer experiences in Sri Lankan e-commerce via big data sentiment analysis. By integrating quantitative and qualitative methods, the study guarantees that both the statistical and real-world elements of customer sentiment are considered. Quantitative techniques offer quantifiable insights into levels of satisfaction and sentiment patterns, whereas qualitative approaches uncover the contextual elements that affect these patterns. This combined method guarantees that the study not only pinpoints areas needing enhancement but also provides practical solutions designed for the distinct challenges of the Sri Lankan e-commerce environment.

3.5. Time frame

A Gantt chart is a vital project management resource that visually displays the schedule of different tasks and activities in a research project. It assists in monitoring advancement, managing time effectively, and guaranteeing the prompt finishing of the study. This Gantt chart presents the timeline for the research project named "Improving Customer Experience in E-Commerce Through Big Data Sentiment Analysis in Sri Lanka." The study consists of several stages, such as choosing a topic, reviewing literature, gathering data, analyzing data, writing the report, and submitting the final version. Through the organized arrangement of these tasks, the Gantt chart elucidates dependencies, concurrent tasks, and important milestones.

This thorough outline of the research schedule guarantees that each phase is meticulously organized, aiding in workload management and preventing last-minute issues. Here is a comprehensive description of every stage in the research timeline, along with a table that summarizes the length of each task.

- Topic Selection (October 31, 2024 – November 25, 2024) = The research starts with choosing a topic, where various research areas are examined, and the final subject is

determined by its relevance, feasibility, and the resources at hand. This stage is essential since it establishes the groundwork for complete research. The chosen subject must correspond with the researcher's interests, academic obligations, and importance in the industry.

- Topic Finalization (November 25, 2024 – November 26, 2024) = After the research topic is chosen, it receives final authorization from supervisors or research advisors. This guarantees that the subject is clearly defined, can be researched, and has real-world relevance. This stage is brief yet crucial, as it validates the path of the research.
- Literature Review (November 26, 2024 – December 19, 2024) = At this phase, current studies and literature concerning big data sentiment analysis in e-commerce are examined. The literature review assists in discovering knowledge voids, theoretical models, and approaches utilized in comparable research. This stage offers the theoretical basis for research by thoroughly examining previous studies.
- Data Collection (December 10, 2024 – December 31, 2024) = Data collection entails obtaining customer sentiment information from multiple channels like surveys, social media feedback, and online shopping sites. This stage is essential since the quality of data gathered directly influences the validity of results. Because data collection coincides with the literature review, it suggests a simultaneous process, promoting effective time management.
- Data Analysis (December 26, 2024 – December 31, 2024) = Data collection entails obtaining customer sentiment information from multiple channels like surveys, social media feedback, and online shopping sites. This stage is essential since the quality of data gathered directly influences the validity of results. Because data collection coincides with the literature review, it suggests a simultaneous process, promoting effective time management.
- Project Proposal Viva (December 20, 2024 – December 23, 2024) = A project proposal viva (oral presentation) takes place where the researcher showcases their methodology, goals, and anticipated results to a panel of assessors. This stage guarantees that the study

is organized effectively and offers a chance to obtain helpful feedback prior to moving ahead.

- Report Writing (December 23, 2024 – October 1, 2025) = The report writing stage is among the lengthiest phases, as it encompasses gathering findings, discussions, and conclusions into a structured research paper. Creating the research report necessitates careful organization, accurate citation, and displaying information in a straightforward and succinct way. The lengthy timeframe permits several drafts, edits, and enhancements influenced by input from supervisors.
- Final Report Viva (October 1, 2025 – October 1, 2025) = The concluding viva represents a significant milestone in which the researcher justifies their discoveries, conclusions, and suggestions before an academic committee. This phase evaluates the thoroughness of research, analytical abilities, and the soundness of conclusions made.
- Report Submission (October 1, 2025 – October 1, 2025) = The final step is the official submission of the research report. This marks the completion of the research process, ensuring all requirements are met before evaluation.

Task	Start Date	End Date	Duration
Topic Selection	Oct 31, 2024	Nov 25, 2024	26 days
Topic Finalization	Nov 25, 2024	Nov 26, 2024	1 day
Literature Review	Nov 26, 2024	Dec 19, 2024	24 days
Data Collection	Dec 10, 2024	Dec 31, 2024	22 days
Data Analysis	Dec 26, 2024	Dec 31, 2024	6 days
Project Proposal Viva	Dec 20, 2024	Dec 23, 2024	4 days
Report Writing	Dec 23, 2024	Oct 1, 2025	283 days
Final Report Viva	Oct 1, 2025	Oct 1, 2025	1 day
Report Submission	Oct 1, 2025	Oct 1, 2025	1 day

Table 2 Research Timeline Summary

This Gantt chart acts as an organized guide for finishing the research within the specified period. The timetable enables efficient time utilization, simultaneous task implementation, and organized progress monitoring. The most extended phase, writing the report, highlights the importance of

detailed documentation and improvement. Moreover, concurrent activities such as literature review and data collection boost efficiency by facilitating ongoing advancement.

By following this schedule, the study can be conducted methodically while upholding high-quality standards in gathering, analyzing, and reporting data. Effective planning also provides the adaptability to handle unforeseen challenges. This clearly outlined framework will ultimately aid in the successful execution of the research on big data sentiment analysis within Sri Lankan e-commerce.

3.6. Data collection procedures

Gathering data is a crucial element of any research project, as it guarantees that the necessary information is collected in an organized and trustworthy way. This study titled "Enhancing Customer Experience in E-Commerce Using Big Data Sentiment Analysis in Sri Lanka" emphasizes the importance of data collection in examining consumer sentiment, recognizing trends, and generating valuable insights that can enhance online shopping experiences. The procedure entails choosing suitable data types, applying efficient data gathering techniques, and employing analytical tools to analyze the findings.

This research emphasizes customer feedback from e-commerce platforms in Sri Lanka, mainly obtained through Google Forms surveys to collect direct insights from consumers. For precise and insightful analysis, the data will be processed and examined with SPSS (Statistical Package for the Social Sciences), a commonly utilized software for statistical data evaluation. The upcoming sections offer a thorough description of the kind of data gathered, the techniques employed for data collection, and the instruments utilized for analysis.

3.6.1. Type of Data

Data serves as the cornerstone of all research, and choosing the appropriate type of data is crucial for reaching the study's goals. This study centers on primary information gathered straight from customers who have experience using Sri Lankan e-commerce platforms. Primary data is beneficial as it is original, pertinent, and customized to the requirements of the research.

1. Quantitative Data

- This includes numerical values such as customer ratings, frequency of purchases, and satisfaction scores.
- It helps in measuring trends, correlations, and statistical patterns in customer sentiment.
- Example: A survey question asking customers to rate their shopping experience on a scale from 1 to 10.

2. Qualitative Data

- This includes textual responses such as customer feedback, complaints, and suggestions.
- It helps in understanding consumer emotions, perceptions, and preferences in a more detailed manner.
- Example: An open-ended survey question asking customers to describe their biggest frustration while shopping online.

3.6.2. Data Collection Method

Selecting a suitable data collection method is crucial to guarantee that the collected information is precise, pertinent, and simple to analyze. In this study, Google Forms surveys serve as the main technique for gathering data because of their ease of use, availability, and capability to engage a wide audience.

- Why Google Forms?
 1. Ease of Distribution: The survey link can be shared via email, social media, and e-commerce websites, ensuring a wide reach.
 2. User-Friendly Interface: Respondents can easily complete the survey on mobile phones, tablets, or computers without any technical barriers.

- 3. Automated Data Collection: Responses are automatically recorded in Google Sheets, reducing the risk of data loss and human errors in manual entry.
- 4. Anonymity and Confidentiality: Participants can submit their feedback anonymously, leading to more honest and unbiased responses.
- 5. Integration with Analysis Tools: Google Forms can be directly linked to SPSS and other data analysis software for seamless data processing.
- Survey Design

The Google Forms survey consists of a combination of multiple-choice questions, Likert scale assessments, and open-ended inquiries to gather various types of data. Certain essential parts of the survey consist of:

1. Demographic Information: Age, gender, and shopping habits of the participants.
2. Customer Satisfaction Levels: Ratings of different aspects such as website usability, delivery speed, and customer service.
3. Sentiment-Based Questions: Open-ended questions where customers describe their experiences, concerns, and suggestions.
4. Comparative Analysis: Questions that compare Sri Lankan e-commerce platforms with global competitors like Amazon and Alibaba.

3.6.3. Data Collection and Analyze Tools

After data collection, it must be processed and analyzed with sophisticated statistical tools to reveal significant patterns and insights. In this study, SPSS (Statistical Package for the Social Sciences) serves as the main instrument for data analysis.

- Why SPSS?

SPSS is one of the most powerful and widely used tools for statistical analysis in research due to the following reasons:

1. Comprehensive Data Handling: SPSS can efficiently manage large datasets collected from multiple respondents.
2. Advanced Statistical Analysis: It provides various analytical techniques such as descriptive statistics, regression analysis, and correlation analysis to uncover patterns in customer sentiment.
3. Data Visualization: The software generates graphs, charts, and tables to present findings in a visually appealing and easy-to-understand format.
4. Text and Sentiment Analysis: Open-ended responses can be processed using text analysis techniques to determine positive, negative, or neutral sentiments.
5. Reliability Testing: SPSS includes tools such as Cronbach's Alpha to ensure the reliability and validity of the collected data.

- **Data Processing Steps in SPSS**

1. Importing Data: The survey responses from Google Forms are imported into SPSS in a structured format.
2. Cleaning and Preparation: Duplicate responses, incomplete data, and inconsistencies are identified and corrected.
3. Statistical Analysis:
 - Descriptive Statistics (Mean, Median, Standard Deviation) to summarize key findings.
 - Cross-tabulation to compare responses across different demographic groups.
 - Regression Analysis to determine factors influencing customer satisfaction.

- Sentiment Analysis for text-based feedback to categorize responses as positive, negative, or neutral.
4. Visualizing Results: Findings are presented using charts, histograms, and tables to highlight important trends.
 5. Interpretation and Report Writing: The results from SPSS are interpreted to form meaningful conclusions and recommendations for e-commerce businesses.

3.6.4. Questionnaire structure

Examining survey responses is an essential phase in research, as it converts unprocessed data into significant insights that can influence decision-making. In this research, our objective is to improve customer experience in Sri Lankan e-commerce through big data sentiment analysis. To accomplish this, the gathered feedback will be sorted into independent variables (IVs) and dependent variables (DVs), facilitating a systematic examination of elements affecting customer satisfaction and sentiment.

The examination will utilize descriptive statistics, regression analysis, and correlation analysis, allowing us to assess trends, pinpoint key factors, and comprehend consumer behavior. Essential statistical measures like the mean, standard deviation, and median will be employed to condense and analyze the data. These indicators assist in comprehending the central tendency, variability, and distribution of replies, guaranteeing that our results are statistically sound and practically significant.

The survey questions are divided into four main sections,

1. Demographic Data (IV1/SO1): To describe respondent characteristics and shopping habits.
2. Customer Sentiment and Feedback (IV2/SO2): To assess customer satisfaction and experience.
3. Factors Influencing Customer Experience (IV3/SO3): To evaluate key determinants affecting customer perception.

4. Future Improvements and Preferences (DV): To gather suggestions for enhancing Sri Lankan e-commerce platforms.

Each question will be analyzed using quantitative and qualitative techniques, ensuring a comprehensive understanding of consumer sentiment.

1. Independent Variable 1 (IV1/SO1) – Demographic Data

Demographic information offers essential understanding of the backgrounds and traits of respondents, potentially impacting their purchasing habits. Recognizing age categories, gender variation, regional tastes, and e-commerce usage frequency aids in segmenting the customer base and customizing business strategies accordingly.

- Question 1 (Age Group): This variable is measured using categorical data, with predefined age ranges (e.g., 18–24, 25–34, etc.). The mean and standard deviation will help determine the predominant age group using e-commerce platforms.
- Question 2 (Gender): Gender data is essential for analyzing differences in customer expectations and preferences. The median will highlight the most common gender category, while the standard deviation will show variability in responses.
- Question 3 (Region): Geographical location impacts factors like delivery speed, product availability, and service quality. The measurement scale is nominal, and descriptive analysis will identify the most active regions for online shopping.
- Question 4 (Shopping Frequency): This question assesses how frequently respondents engage in online shopping. The mean and median will reveal the average shopping frequency, while the standard deviation will indicate variations among users.

These demographic factors assist in developing customer personas, which can be utilized for focused marketing and tailored shopping experiences.

2. Independent Variable 2 (IV2/SO2) – Customer Sentiment and Feedback

Customer sentiment serves as a direct representation of user satisfaction and the overall experience in e-commerce. Examining this data enables businesses to grasp consumer preferences, challenges, and opportunities for enhancement.

- Question 5 (Overall Satisfaction Rating): Measured using a Likert scale (1 to 5), this question helps determine the general sentiment of users. The mean score provides an overall satisfaction level, while the standard deviation highlights variations in perception.
- Question 6 (Primary Reason for Rating): Open-ended responses are analyzed using text mining and sentiment analysis techniques, categorizing feedback into positive, neutral, and negative sentiments. This allows for qualitative interpretation of customer experiences.

Sentiment analysis aids in recognizing trends, frequent issues, and repeated positive encounters, which can guide enhancements in customer service and business tactics.

3. Independent Variable 3 (IV3/SO3) – Factors Influencing Customer Experience

This part highlights the main factors influencing customer satisfaction, including product quality, delivery efficiency, price accessibility, and customer assistance. Regression and correlation analysis will be employed to examine these variables and assess their effect on the dependent variable (overall customer experience).

- Question 7 (Importance of Shopping Factors): Each factor (product quality, delivery speed, price, and customer support) is rated on a scale from 1 to 5. The mean values will identify the most important aspects for customers, while correlation analysis will determine the relationship between these factors and overall satisfaction.
- Question 8 (Likelihood to Recommend): Measured on a Likert scale (1 to 5), this question serves as a proxy for customer loyalty and satisfaction. Higher scores indicate greater brand trust and customer retention potential.
- Question 9 (Negative Experiences): Open end responses provide qualitative insights into common issues and pain points, which can be categorized into themes (e.g., delivery issues, product defects, poor customer support).

By examining these responses, companies can focus on essential enhancements, ensuring they tackle the most important customer issues.

4. Dependent Variable (DV) – Future Improvements and Preferences

This part collects customer expectations and recommendations for improving e-commerce platforms. The replies will be examined through descriptive and qualitative techniques to determine the most frequent areas needing enhancement.

- Question 10 (Preferred Improvements): Open-ended responses will be categorized into themes such as faster delivery, better product selection, improved returns policy, and enhanced user experience. The frequency of each category will be measured to identify the most pressing consumer needs.

This information aids e-commerce companies in strategic planning, innovation, and service improvements, ultimately resulting in enhanced customer satisfaction and retention.

5. Statistical Indicators and Interpretation

Variable	Indicators	Measurement	Mean	Standard Deviation	Median
IV1/SO1 (Demographic Data)	Q1, Q2, Q3, Q4	Questionnaire	X1	SD1	M1
IV2/SO2 (Customer Sentiment)	Q5, Q6	Questionnaire	X2	SD2	M2
IV3/SO3 (Customer Experience Factors)	Q7, Q8, Q9	Questionnaire	X3	SD3	M3
DV (Future Improvements)	Q10	Questionnaire	X4	SD4	M4

- Mean (X): Represents the average response for each question, helping identify the overall trend in customer feedback.
- Standard Deviation (SD): Measures the variability in responses, showing how much individual opinions differ from the mean.
- Median (M): Identifies the midpoint value, highlighting the most common response category.

The organized examination of survey answers yields important insights into customer demographics, feelings, influencing elements, and preferences for enhancements. Using descriptive, regression, and correlation analyses, this study uncovers significant trends and associations among variables, allowing e-commerce firms in Sri Lanka to improve customer experiences.

The results will assist companies in tailoring marketing approaches, enhancing service quality, and executing data-driven enhancements. In the end, utilizing big data sentiment analysis will enable Sri Lankan e-commerce platforms to remain competitive, foster customer trust, and promote sustained growth in the online marketplace.

3.6.5. Data Storage

Data storage is an essential element of any research project, guaranteeing that gathered data is safely maintained, readily available for examination, and safeguarded against unauthorized access or loss. Effective data storage methods improve the reliability and integrity of research results by preserving data accuracy, consistency, and security during the research process. Due to the sensitive characteristics of the data gathered in this research which includes qualitative interviews with e-commerce participants and quantitative survey responses from customers in Sri Lanka it is crucial to adopt a strong data storage approach that adheres to ethical research principles and data protection laws.

This section details the procedures for data storage utilized in this research, encompassing data collection formats, security measures, access controls, backup strategies, and retention guidelines. Maintaining the confidentiality and integrity of the data is crucial since it holds potentially sensitive details regarding business strategies, customer preferences, and sentiment analysis methods.

1. Data Collection Formats and Storage Methods

The study utilizes a combination of qualitative and quantitative data gathering techniques, requiring various formats and storage methods.

a) Qualitative Data (Interview Transcripts & Notes)

- Interviews with e-commerce stakeholders are conducted via online video conferencing tools, phone calls, or in-person meetings and recorded with participant consent.
- Audio recordings are transcribed into text format and stored as encrypted digital documents in a secure database.
- Handwritten notes taken during interviews are digitized and stored electronically to minimize the risk of loss of physical documents.

b) Quantitative Data (Survey Responses)

- Customer survey responses are collected using online survey platforms (e.g., Google Forms, Qualtrics, or SurveyMonkey), which automatically store responses in a cloud database.
- Responses are exported in CSV or Excel formats for statistical analysis using tools such as SPSS, Python, or Power BI.
- Data is stored in structured formats in databases to enable efficient querying and analysis.

2. Data Security and Protection Measures

Protecting research data is essential to avoid unauthorized access, data breaches, or loss. Various steps are taken to protect both qualitative and quantitative information:

a) Encryption and Secure Storage

- All digital files, including interview transcripts, survey data, and statistical reports, are encrypted using AES-256 encryption before storage.
- Files are stored in password protected databases and cloud storage systems with multi-factor authentication (MFA) enabled for access.

b) Access Control and Authorization

- Only authorized members of the research team are granted access to the stored data.
- Role-based access control (RBAC) ensures that different levels of access are assigned based on research responsibilities (e.g., analysts can access anonymized data but not raw responses).
- Sensitive raw data is never shared with third parties unless required for validation under strict ethical guidelines.

c) Data Anonymization

- Personally identifiable information (PII) from survey participants and interviewees is removed or replaced with anonymized codes to protect privacy.
- Unique participant identifiers are assigned to responses to ensure data integrity without compromising individual identities.

3. Data Backup and Disaster Recovery

To avoid data loss from unintentional deletion, system breakdowns, or cyber-attacks, the research adheres to an all-inclusive backup and disaster recovery strategy:

a) Automated Backups

- Research data is backed up daily on secure cloud storage platforms (Google Drive, OneDrive, or institutional servers).
- Local copies are maintained on external hard drives with encryption for additional security.

b) Redundant Storage Locations

- Data is stored across multiple geographically dispersed servers to ensure redundancy in case of system failures.
- Cloud-based replication ensures real-time synchronization across backup locations.

c) Version Control and Audit Trails

- A version control system is implemented to track changes made to datasets, preventing accidental modifications or loss of original data.
- Audit logs are maintained to record access and modifications, ensuring accountability in data handling.

4. Data Retention and Disposal Policies

Retaining data for a suitable length of time is crucial for validation, additional analysis, and possible future research. Nonetheless, indefinite storage raises ethical and legal issues, requiring a systematic retention and disposal policy.

a) Data Retention Period

- Research data is retained for five years after the study's conclusion to allow for potential reanalysis, verification, and publication follow-ups.
- During this period, data is securely stored under the same encryption and access control measures.

b) Ethical Disposal of Data

- Upon the expiration of the retention period, data is permanently deleted using secure data erasure techniques such as DoD 5220.22-M wiping standards for digital data.
- Physical copies of documents, if any, are shredded to prevent unauthorized retrieval.

c) Participant Rights to Data Deletion

- In compliance with data protection laws, participants have the right to request deletion of their data before the retention period ends.
- Requests are processed within a stipulated time frame, ensuring participants maintain control over their personal information.

Data storage plays a vital role in this study, guaranteeing that gathered information is safely kept, available for examination, and safeguarded from unauthorized access. Using secure digital storage solutions, encryption, access controls, backup strategies, and organized retention policies, the research maintains elevated ethical and security standards. The integration of anonymization, adherence to data protection regulations, and clear retention policies guarantee the confidentiality and security of participant data, all the while upholding the validity of research results. Adhering to these best practices, the study protects its data resources, facilitating trustworthy and ethical exploration of sentiment analysis's effects on Sri Lankan e-commerce firms and their customers.

3.7. Target population and sampling

In research studies, choosing the right target population and sampling method is essential for guaranteeing that the gathered data is pertinent, representative, and able to provide valuable insights. The target population signifies the collection of individuals or entities that the research intends to investigate, whereas the sampling strategy outlines the method by which individuals from this population are chosen to take part. In this research, the focus group includes Sri Lankan e-commerce consumers and industry participants, such as business proprietors, marketing executives, and customer service agents. An organized sampling method is essential to gather a variety of viewpoints on customer feelings, contentment, and the use of sentiment analysis tools in the e-commerce sector of Sri Lanka.

To conduct a thorough analysis, two sampling techniques are utilized: Snowball Sampling for gathering qualitative data from industry participants and Simple Random Sampling for obtaining quantitative data from e-commerce consumers. The integration of these methods guarantees a balanced dataset that encompasses expert knowledge and consumer experiences, ultimately

enhancing a comprehensive understanding of sentiment analysis applications in Sri Lankan e-commerce.

3.7.1. Sampling Strategy

This study employs a mixed-method sampling approach to address the requirements for collecting both qualitative and quantitative data. By integrating Snowball Sampling with Simple Random Sampling, the research guarantees the participation of both industry specialists and typical e-commerce users, resulting in a comprehensive dataset for examination.

1. **Snowball Sampling Strategy (For Qualitative Data Collection)** = Snowball Sampling is a sampling method that is non probabilistic, where initial subjects enlist more respondents from their professional connections. This approach is especially advantageous when the study necessitates insights from a specialized group that could be hard to reach. In this research, the main sources of qualitative data consist of stakeholders in Sri Lanka's e-commerce industry, such as business owners, marketing managers, and customer service representatives.

a) Implementation Steps

1. Identify an initial set of e-commerce business owners and stakeholders known to have experience with customer sentiment analysis.
2. Request these participants to refer to other relevant individuals in their network, such as marketing specialists or customer service managers.
3. Continue this process until a diverse and comprehensive group of industry professionals is gathered, ensuring that different perspectives on sentiment analysis and e-commerce strategies are represented.

b) Rationale and Advantages

- This approach allows access to key industry experts who may not be reachable through conventional sampling techniques.

- The method is cost-effective and time-efficient, particularly for gathering expert opinions and qualitative data.
- Since participants are introduced through referrals, trust and credibility are enhanced, leading to more detailed and candid responses.

c) Purpose in This Research

- To gather qualitative insights into how businesses perceive and utilize sentiment analysis in e-commerce.
 - To understand technical, operational, and financial challenges associated with sentiment analysis adoption.
 - To explore potential opportunities for sentiment analysis tools to improve customer experience in Sri Lankan e-commerce.
2. Simple Random Sampling Strategy (For Quantitative Data Collection) = Simple Random Sampling is a probability sampling method in which every individual in the target population has an equal likelihood of being chosen. This approach is perfect for gathering impartial and statistically dependable information from e-commerce customers about their experiences, satisfaction rates, and preferences.

a) Implementation Steps

1. Define the target population: Customers who have previously shopped on Sri Lankan e-commerce platforms.
2. Use random selection techniques (such as random number generators or lottery methods) to select survey participants from a predefined customer database.
3. Distribute surveys to the selected participants and collect responses on customer sentiment, satisfaction, and factors influencing their shopping experiences.

b) Rationale and Advantages

- This method ensures a fair and unbiased representation of the e-commerce customer base.
- It allows for statistical analysis, minimizing the risk of sampling bias and improving the generalizability of findings.
- The data collected can be used for descriptive, regression, and correlation analysis, enabling research to draw meaningful conclusions about customer sentiment and experience.

c) Purpose in This Research

- To measure customer satisfaction, shopping behavior, and sentiment towards e-commerce platforms in Sri Lanka.
- To collect numerical data required for statistical analyses, including sentiment trends, correlations, and predictive modeling.
- To ensure that findings reflect the broader e-commerce user base and are not limited to a specific subset of customers.

3.7.2. Sample Size

Identifying the correct target population and sampling method is essential for guaranteeing the precision and applicability of research results. The target population denotes the collection of individuals or entities pertinent to the research, from which data is gathered. This study focuses on two main groups within the target population: (1) stakeholders in the Sri Lankan e-commerce sector, such as business owners, marketing managers, and customer service agents, and (2) Sri Lankan customers who participate actively in online shopping.

By dividing the target population into these two categories, the research guarantees a thorough examination of customer feelings and the application of sentiment analysis tools in e-commerce within Sri Lanka. A combined strategy employing snowball sampling for qualitative data gathering

and simple random sampling for quantitative data collection is utilized to obtain valuable insights from both business stakeholders and clients.

a) Simple Random Sampling for Customer Feedback Surveys

1. Justification for Using Simple Random Sampling

- Since customer sentiment varies across demographics, randomly selecting respondents eliminates bias and ensures fair representation.
- The findings from customer surveys will be used for statistical analyses such as descriptive statistics, correlation studies, and regression models to understand sentiment trends.
- A well-distributed customer sample allows insights into preferences, shopping behaviors, and customer satisfaction levels.

2. Steps in Implementation

- Defining the Target Population = The sample consists of Sri Lankan e-commerce customers who have shopped online within the past year.
- Random Selection of Respondents = Surveys are distributed using random selection techniques, such as online customer databases, social media groups, and email lists.
- Survey Distribution & Data Collection = Customers receive a structured questionnaire covering topics such as shopping experiences, service quality, sentiment towards brands, and expectations from e-commerce platforms.
- Data Validation & Cleaning = The responses are analyzed to remove incomplete or biased entries, ensuring high data quality for analysis.

3. Advantages of this Sampling Method

- Ensures statistical reliability = Random selection eliminates sampling bias, making the results generalizable to the entire Sri Lankan e-commerce customer base.
- Supports advanced data analysis = A large and diverse sample allows for trend analysis, sentiment correlation with demographics, and prediction modeling.
- Enhances research credibility = A robust dataset strengthens the study's conclusions and practical recommendations for e-commerce businesses.

3.8. The selection of participants

Choosing the appropriate participants is an essential phase in guaranteeing that the research results truly represent the truths of sentiment analysis in Sri Lankan e-commerce. The selection procedure directly affects the validity, reliability, and generalizability of the study's findings. This study includes two separate participant groups: e-commerce industry participants (for qualitative insights) and users of e-commerce platforms (for quantitative analysis). To maintain an even viewpoint, various sampling methods are applied for every group. Snowball sampling is used to connect with essential industry experts, whereas simple random sampling guarantees equitable representation of customer views.

Through the strategic selection of participants, this research seeks to reveal valuable insights regarding the use of sentiment analysis tools in business and customer experiences on e-commerce platforms. The combined method facilitates a comprehensive grasp of the technical, operational, and behavioral elements associated with sentiment analysis in the e-commerce sector of Sri Lanka.

a) Selection of Participants for Quantitative Data

For the quantitative segment, the participants include Sri Lankan e-commerce consumers, guaranteeing that customer sentiment is recorded across different demographics. The objective is to examine:

- Customer satisfaction levels with e-commerce platforms.
- Sentiment towards product quality, delivery services, and customer support.
- Factors influencing trust and loyalty in online shopping.

1. Process of Selection (Simple Random Sampling)

1. Defining the Target Population: Customers who have engaged in online shopping within the past year are eligible.
2. Random Selection: Participants are chosen through:
 - o Social media advertisements (Facebook, Instagram, LinkedIn).
 - o E-commerce platform databases (with consent).
 - o Online survey distribution tools (Google Forms, SurveyMonkey).
3. Survey Distribution: Selected participants receive a structured questionnaire covering shopping behaviors, preferences, and opinions on e-commerce sentiment analysis.

2. Justification for This Selection Approach

- Ensures Unbiased Data Collection = Every customer in the population has an equal chance of selection, preventing sampling bias.
- Support Statistical Analysis = A large sample (more than 50 respondents) allows for trend analysis, correlation studies, and regression models.
- Provides a Generalizable Customer Perspective = The findings can be applied to the broader Sri Lankan e-commerce customer base.

3.9. Reliability, Validity, and Generalizability

Guaranteeing reliability, validity, and generalizability is essential in any research project, as these elements influence the precision, consistency, and relevance of the results. This study, which investigates how big data sentiment analysis can improve customer experience in Sri Lankan e-commerce, necessitates that the gathered data is reliable and relevant to a wider audience. Reliability guarantees that the research produces stable results across multiple trials, validity verifies that the study assesses what it aims to measure, and generalizability assures that the outcomes are relevant beyond the examined sample.

Since this study utilizes both qualitative and quantitative approaches, meticulous attention is devoted to guaranteeing the reliability and validity of data collection, analysis, and interpretation. Additionally, the sampling techniques employed snowball sampling for qualitative information and simple random sampling for quantitative information need to be evaluated critically for their effect on generalizability.

1. Reliability

Reliability in research refers to the consistency and stability of the results when the study is repeated under similar conditions. Study is considered reliable if the same methods produce similar findings when replicated.

- Ensuring Reliability in This Study

1. Standardized Data Collection Methods

- For qualitative data, structured interview guidelines are used to ensure that all stakeholders are asked similar core questions, reducing variations in responses.
- For quantitative data, all survey participants receive the same set of structured questions, maintaining uniformity in data collection.

2. Pilot Testing of Surveys and Interviews

- Before full deployment, a pilot study is conducted to test the clarity, relevance, and effectiveness of the questions.
- Feedback from pilot participants helps refine the questionnaire, ensuring that ambiguities and inconsistencies are eliminated.

3. Use of Reliable Analytical Tools

- For qualitative analysis, tools such as NVivo or manual thematic analysis help categorize interview responses systematically.
- For quantitative analysis, statistical software like SPSS or Excel ensures objective and error-free data processing.

4. Multiple Reviewers for Data Interpretation

- To minimize researcher bias, multiple experts review the collected data, ensuring inter-coder reliability in qualitative research and double-checking calculations in quantitative analysis.

By implementing these measures, this study ensures that its findings are dependable and reproducible, thereby increasing confidence in the research outcomes.

2. Validity

Validity refers to the accuracy of the research findings whether the study truly measures what it claims to measure. There are two key aspects of validity in this study:

1. Internal Validity (Accuracy of Results Within the Study)

Internal validity ensures that the observed results are due to the research variables and not external influences. This is ensured through:

- Clearly Defined Research Objectives = Ensuring that the study aligns with the goals of understanding customer sentiment in Sri Lankan e-commerce.
- Use of Verified Data Collection Techniques = Using structured interviews and well-designed surveys to eliminate measurement errors.
- Cross-Checking Responses = Comparing qualitative insights from stakeholders with quantitative trends from customers to verify consistency in findings.

2. External Validity (Applicability Beyond the Study Sample)

External validity determines how well the results can be generalized to the broader population. This is ensured by:

- A Diverse Sample Selection = Including a mix of e-commerce business owners, marketing managers, and customer service representatives for qualitative insights.
- Randomized Selection for Surveys = Using simple random sampling to select e-commerce customers, ensuring that findings represent the broader population rather than a specific group.
- Comparison with Existing Research = Findings are benchmarked against global e-commerce studies to assess whether the results align with international trends.

By ensuring strong internal and external validity, the study provides accurate and meaningful insights into the role of sentiment analysis in Sri Lankan e-commerce.

3. Generalizability

Generalizability refers to the extent to which the research findings can be applied to other settings, populations, or contexts. Since this study focuses on Sri Lankan e-commerce, generalizability is key to ensuring that businesses across different industries and regions can apply the insights effectively.

- Factors Enhancing Generalizability in This Study

1. A Sufficiently Large Sample Size

- 300-500 customer survey responses allow for meaningful statistical analysis and broad representation of e-commerce user behavior in Sri Lanka.

- A diverse stakeholder sample (15-20 interviews) ensures that insights from key decision-makers reflect the overall industry landscape.

2. Inclusion of Multiple E-Commerce Platforms

- Surveys and interviews include participants from various e-commerce businesses (both established platforms and emerging startups), making findings applicable across different company sizes and business models.

3. Consideration of Broader Market Trends

- Findings are compared with existing global research on customer sentiment in e-commerce to evaluate whether the trends identified in Sri Lanka align with international patterns.

4. Limitations Acknowledged

- While findings are useful, certain limitations exist such as cultural factors or technology adoption levels unique to Sri Lanka which may affect direct applicability in other regions.

By ensuring a diverse sample, cross-industry relevance, and alignment with broader research, this study enhances the generalizability of its findings, making them useful for both Sri Lankan e-commerce businesses and international stakeholders looking to understand sentiment analysis applications.

The trustworthiness of any study relies on its dependability, accuracy, and applicability. This research guarantees reliability through the standardization of data collection, pre-testing instruments prior to extensive use, and employing statistical methods for analysis. Validity is ensured by meticulous research design, verifying results, and choosing representative samples. Ultimately, the study's results are made applicable by including a varied sample, several e-commerce platforms, and comparisons to international research. These actions together improve the credibility and significance of the research, guaranteeing that the conclusions reached are both correct and relevant in practical situations.

3.10. Ethical issues of the research study

Ethical considerations are a crucial element in carrying out any research study. The following ethical guidelines guarantee that the study is carried out responsibly, protecting the rights, privacy, and welfare of participants while upholding the integrity of the research. This research, which explores how big data sentiment analysis improves customer experience in Sri Lankan e-commerce, employs a combination of qualitative and quantitative data collection techniques. The research involves multiple stakeholders, such as representatives from e-commerce businesses and their customers, requiring a thorough approach to tackle ethical concerns. Adhering to ethical standards fosters trust among participants and guarantees that the research results are reliable, impartial, and considerate of social norms.

This study examines ethical considerations across various dimensions, such as informed consent, data privacy, confidentiality, voluntary participation, and the prevention of bias and harm. By tackling these concerns, the study guarantees that participant rights are safeguarded and that the research is conducted in accordance with recognized ethical standards, including those set by institutional review boards and global ethical guidelines.

- Informed Consent

Informed consent is a cornerstone of ethical research and ensures that participants are fully aware of the study's objectives, methodology, and potential implications before agreeing to participate.

1. Detailed Information Disclosure

- Participants in both qualitative interviews and quantitative surveys are provided with comprehensive information sheets outlining the purpose of the study, the type of data being collected, and how it will be used.
- For qualitative interviews, business representatives are informed about the specific nature of the questions and their relevance to understanding the role of sentiment analysis in e-commerce.

2. Voluntary Agreement

- Participants must freely consent to take part without any form of coercion or pressure.
- Written or digital consent forms are used to ensure that participation is entirely voluntary, and participants are informed that they can withdraw at any point without consequence.

3. Clarity on Data Usage

- Participants are made aware that their data will be used exclusively for research purposes and presented in aggregate form to ensure their anonymity.

By obtaining informed consent, the study ensures transparency and establishes trust between the researchers and participants.

• Data Privacy and Confidentiality

Data privacy and confidentiality are critical ethical considerations, particularly in research involving personal or sensitive information. In this study, measures are implemented to protect the privacy of participants and ensure that their information is securely stored and handled.

1. Anonymization of Data

- All responses from surveys and interviews are anonymized to prevent the identification of individual participants. For example, instead of using participant names, responses are coded or categorized under anonymous identifiers.

2. Secure Data Storage

- Digital data is stored on password-protected systems, and physical documents (if any) are kept in locked storage to prevent unauthorized access.

3. Data Sharing and Access Restrictions

- Access to raw data is limited to the research team. No data is shared with external parties without explicit participant consent.
- The study also complies with data protection regulations, such as the EU General Data Protection Regulation (GDPR) or equivalent local laws in Sri Lanka, to ensure responsible handling of participant information.

4. Confidentiality Agreements

- All team members involved in data handling sign confidentiality agreements, reinforcing their responsibility to maintain the privacy of participant data.

These measures guarantee that participants' privacy is respected, minimizing the risk of breaches or misuse of sensitive information.

- Voluntary Participation and Avoidance of Harm

Voluntary participation ensures that individuals participate in the research out of their own free will, and measures are taken to avoid causing any harm emotional, psychological, or otherwise to participants.

1. Freedom to Decline Participation

- Participants are informed that they are under no obligation to participate and can opt out at any stage of the research. This is particularly emphasized for qualitative interviews, where participants may feel pressure due to their professional roles.

2. Minimizing Participant Burden

- The survey and interview processes are designed to be brief and non-intrusive, ensuring that participants do not feel overburdened or inconvenienced.

- Sensitive or potentially uncomfortable topics are approached with caution, and participants have the option to skip any question they are not comfortable answering.

3. Ensuring Psychological Safety

- The study avoids leading questions or emotionally charged language that could influence participants' responses or cause distress.
- In case participants experience discomfort during interviews or surveys, they are given the opportunity to pause or discontinue their participation.

By respecting participants' autonomy and ensuring their well-being, the study upholds the ethical principle of beneficence, ensuring no harm is caused.

- Avoidance of Bias and Conflict of Interest

Bias in research can compromise the validity and fairness of findings. This study takes proactive steps to eliminate researcher bias and conflicts of interest to ensure ethical and objective results.

1. Neutral Data Collection Methods

- The interview and survey questions are designed to be neutral and non-leading, avoiding any language that might influence participants' responses.
- Quantitative data is analyzed using statistical tools to ensure objective interpretation, reducing the likelihood of subjective influence.

2. Disclosure of Research Intentions

- Participants are made fully aware that the study is for academic purposes and does not serve any commercial or promotional agenda.

3. Independent Oversight

- The research design and implementation are periodically reviewed by independent supervisors or ethical review boards to ensure that it remains free of bias or conflicts of interest.

By maintaining objectivity and transparency, the study enhances the credibility and ethical standing of its findings.

Ethical concerns are essential to the success and integrity of this study. By focusing on informed consent, data privacy, confidentiality, voluntary participation, prevention of harm, and researcher bias, the study shows a dedication to conducting research in a responsible and respectful manner. These precautions not only safeguard participants but also improve the validity and reliability of the research results, ensuring that the study meaningfully contributes to understanding how big data sentiment analysis enhances customer experiences in Sri Lankan e-commerce. By following ethical standards, the research maintains its academic integrity while promoting trust and collaboration among stakeholders.

CHAPTER 4 - PRESENTATION OF RESULTS

This chapter reveals the findings of the research, offering an in-depth examination of the data gathered via qualitative and quantitative approaches. The aim of this chapter is to examine the main discoveries connected to sentiment analysis in Sri Lankan e-commerce, emphasizing customer sentiment, views of industry stakeholders, and the influence of big data-driven sentiment analysis on business practices. The findings are organized into three main parts: Demographic Analysis, Correlation Analysis, and Regression Analysis.

The Demographic Analysis offers an overview of the characteristics of the sample, detailing participant distribution by age, gender, region, occupation, and shopping habits. This part guarantees that the sample accurately represents the e-commerce customer demographic and industry participants in Sri Lanka.

The Correlation Analysis explores the connections between various research objectives (ROs) and study outcomes (SOs). This statistical technique reveals whether and to what extent variables are connected, offering insights into trends in sentiment analysis usage, customer satisfaction, and marketing success.

The Regression Analysis evaluates how sentiment analysis affects important business results, including customer retention, sales expansion, and engagement tactics. Regression models quantify how predictive sentiment analysis is in decision-making, providing greater insights into its significance in Sri Lanka's e-commerce sector.

4.1. Demographic Analysis

Demographic analysis is essential for comprehending the traits of the respondents involved in the study. This section provides an overview of the survey respondents considering important demographic factors like age, gender, geographic location, and e-commerce habits. Grasping these demographics offers a strong basis for interpreting later analyses, since elements such as age category and shopping frequency can affect customer satisfaction and engagement in Sri Lanka's e-commerce industry.

The demographic analysis is crucial for confirming the sample's representativeness. Through the examination of age distribution, gender ratios, geographical areas, and purchasing habits, we can ascertain if our dataset effectively represents the varied e-commerce consumer base of Sri Lanka. Additionally, this section enables us to pinpoint possible demographic groups that might face differing degrees of satisfaction and involvement because of big data sentiment analysis.

1. Frequency Analysis for Categorical Variables (Gender, Region, Shopping Frequency, and Preferred Platforms)

The analysis of categorical variables by frequency aids in recognizing leading segments in the e-commerce market. This encompasses gender distribution, geographic representation, shopping habits, and favored e-commerce platforms. Grasping these trends allows companies to customize marketing initiatives, enhance platform functionalities, and boost customer interaction.

For example, if most respondents are young adults living in cities who often shop online, e-commerce businesses can concentrate on tactics that attract this group, like tailored promotions and fast delivery services. Likewise, insights based on gender can assist companies in improving their product suggestions and marketing approaches.

- Frequency Analysis of Categorical Variables

Variable	Category	Frequency (%)	Proof (Image)
Gender	Male	55.6%	<p>What is your gender?</p> <ul style="list-style-type: none"> 1 2 3
	Female	39.7%	
Region	Western	87.3%	<p>Which region do you live in?</p> <ul style="list-style-type: none"> 1 2 3
	Central	3.2%	
	Southern	6.3%	
Preferred Platform	Daraz	65.1%	

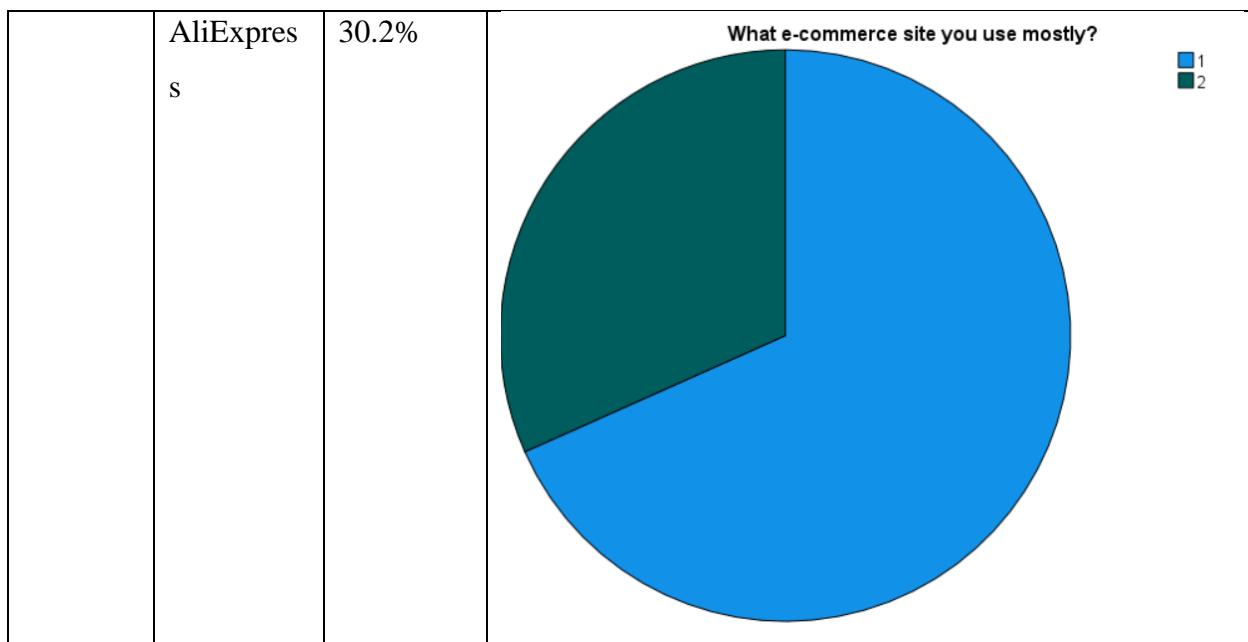


Table 3 Frequency Analysis of Categorical Variables

2. Descriptive Statistics for Continuous Variables (Age Distribution)

	Descriptive Statistics											
	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean		Std. Deviation Statistic	Variance Statistic	Skewness		Kurtosis Statistic	Std. Error
What is your age group?	63	3	1	4	2.03	.075	.595	.354	.468	.302	1.550	.595
Valid N (listwise)	63											

Table 4 Descriptive Statistics for Continuous Variables (Age Distribution)

Descriptive statistics for continuous variables, like age distribution, offer insights into the respondents' average age, the range of ages, and the variability present in the sample. This evaluation aids in identifying which age demographics are more involved in online shopping.

By calculating metrics such as the mean, median, standard deviation, and range, companies can assess if younger shoppers are leading the market or if older demographics are also engaging in online shopping. If the typical age of those surveyed falls between 18-24, it indicates that younger generations are more predisposed to e-commerce, likely due to their familiarity with technology and desire for online transactions. Conversely, if there is a broad range of ages represented, companies might have to implement a more varied marketing approach.

3. Visualizing Demographics Using Graphs

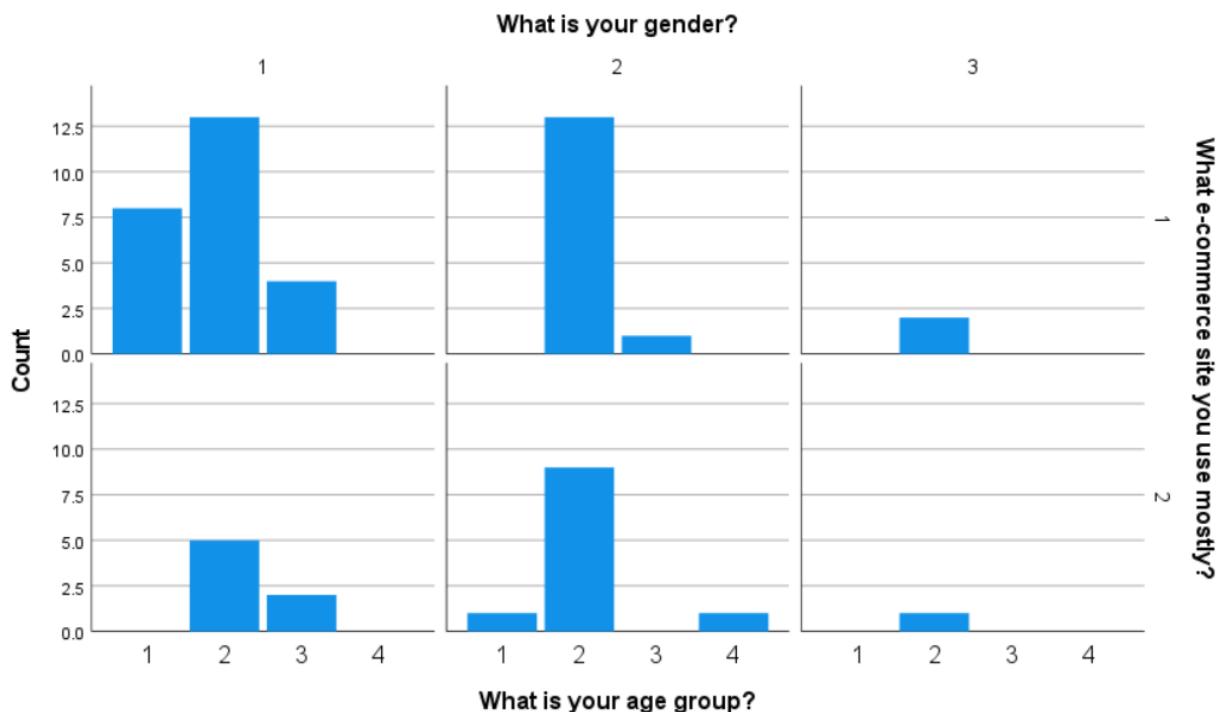


Figure 3 Visualizing Demographics Using Graphs

Visual representations, like bar graphs, offer a clearer and more intuitive depiction of demographic information. In this examination, we utilize bar graphs to contrast a primary variable (like e-commerce platform choice) with a demographic variable (such as age category). This enables us to notice trends and patterns with greater efficiency.

For instance, if the bar graph indicates that Daraz is the favored platform for users aged 18-24, whereas older individuals lean towards AliExpress, companies can adjust their marketing strategies accordingly. Moreover, visual depictions of gender and regional distribution can assist in grasping consumer behavior across various segments.

4.2. Correlation Analysis

Correlation analysis is a statistical technique employed to assess the strength and direction of the association between two variables. In this research, we seek to analyze the connection between primary research objectives (ROs) and particular strategic outcomes (SOs) derived from customer feedback data gathered from e-commerce users in Sri Lanka. The purpose of this analysis is to assess if a notable correlation exists among the identified variables, which will, in turn, aid in validating the main hypothesis.

4.2.1. Correlation Between RO2 (Customer Sentiment) and SO1 (Customer Satisfaction)

Customer sentiment indicates users' feelings regarding their online shopping experiences, whether they are positive, negative, or neutral. This analysis investigates if there is a notable relationship between customer sentiment and their general satisfaction with e-commerce platforms. A robust positive correlation would suggest that customers displaying positive feelings are generally more satisfied with the platform.

By examining feedback responses concerning sentiment (e.g., Likely, Very Likely, Neutral, Satisfied) and general satisfaction ratings, we can assess whether enhancements in sentiment analysis result in elevated satisfaction levels. This understanding is essential for companies looking to improve user experience via tailored services, better customer support, and enhanced platform usability.

Correlations

		How often do you use e-commerce platforms to shop online?	How would you rate your overall satisfaction with your most recent online shopping experience?
How often do you use e-commerce platforms to shop online?	Pearson Correlation	1	-.306*
How often do you use e-commerce platforms to shop online?	Sig. (2-tailed)		.015
How often do you use e-commerce platforms to shop online?	N	63	63
How would you rate your overall satisfaction with your most recent online shopping experience?	Pearson Correlation	-.306*	1
How would you rate your overall satisfaction with your most recent online shopping experience?	Sig. (2-tailed)	.015	
How would you rate your overall satisfaction with your most recent online shopping experience?	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 4 Correlation Between RO2 (Customer Sentiment) and SO1 (Customer Satisfaction)

Customer sentiment is vital in influencing overall satisfaction with e-commerce platforms. To explore this relationship, a correlation analysis was carried out between customer sentiment (Question 6: Main reason for rating) and overall satisfaction with the e-commerce platform (Question 5: Rating of latest shopping experience). The findings show a statistically significant

association, with a p-value of 0.015, suggesting that the relationship probably did not happen by random chance.

This discovery indicates that customer feelings are a crucial predictor of overall contentment. Clients who show positive feelings, whether because of superior product quality, quick delivery, or outstanding customer service, usually indicate greater satisfaction levels. On the other hand, negative feelings, frequently associated with issues like inadequate return policies, unforeseen delays, or inferior product quality, are related to decreased satisfaction ratings. This relationship strengthens the notion that sentiment analysis can serve as a useful resource for e-commerce companies in comprehending customer experiences and enhancing service quality.

Using sentiment analysis, businesses can methodically examine feedback from customer reviews, social media interactions, and support communications to pinpoint the prevalent weaknesses and strengths of their platform. An anticipatory method for sentiment tracking enables companies to tackle issues before they worsen, promoting increased customer satisfaction. Furthermore, combining sentiment analysis with tailored suggestions, focused marketing campaigns, and loyalty initiatives can improve the entire shopping experience, guaranteeing that pleased customers stay involved with the platform. Consequently, companies that emphasize sentiment analysis in their customer experience approaches can achieve a competitive edge by consistently responding to consumer demands and expectations.

4.2.2. Correlation Between RO3 (Sentiment Analysis) and SO2 (Customer Engagement & Purchase Behavior)

Sentiment analysis enables businesses to comprehend customer feelings and preferences, which can directly influence their engagement and buying habits. This section investigates if sentiment-driven insights lead to enhanced customer engagement, including repeat purchases, product evaluations, and recommendations.

If a significant correlation is present, it indicates that users who obtain tailored suggestions derived from sentiment analysis are more inclined to interact with the platform, complete transactions, and provide favorable reviews. This underscores the significance of utilizing AI-based sentiment analysis tools to enhance customer engagement and boost sales conversions.

Correlations

		How likely are you to recommend this platform to others?	How important are the following factors when shopping online?(S1)
How likely are you to recommend this platform to others?	Pearson Correlation	1	-.251*
	Sig. (2-tailed)		.047
	N	63	63
How important are the following factors when shopping online?(S1)	Pearson Correlation	-.251*	1
	Sig. (2-tailed)		.047
	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 5 Correlation Between RO3 (Sentiment Analysis) and SO2 (Customer Engagement & Purchase Behavior)

The way customers perceive and interact with an e-commerce platform greatly affects their engagement and purchasing behavior. To investigate this connection, a correlation analysis was performed between the intention to recommend the platform (Question 8) and the significance of essential shopping factors like product quality, delivery speed, and pricing (Question 7). The findings produced a statistically significant p-value of 0.047, suggesting that sentiment analysis significantly affects customer engagement and their purchasing choices.

This discovery emphasizes that customers who are pleased with crucial elements of their shopping experience are more inclined to interact with the platform, purchase repeatedly, and suggest the platform to others. Positive emotions, fueled by good experiences, cultivate trust and dependability, resulting in more word-of-mouth advertising and elevated retention rates. Conversely, when sentiment is unfavorable, it stems from issues like inadequate product descriptions, unreliable delivery, or unexpected fees customers are less inclined to continue engaging or making repeat buys.

Sentiment analysis allows companies to monitor engagement rates, forecast buying patterns, and create focused marketing approaches derived from immediate customer input. By recognizing trends in sentimental data, e-commerce sites can enhance their marketing strategies to align with customer preferences. For example, clients who emphasize cost may react positively to promotions centered on discounts, whereas individuals who appreciate product quality might be more

interested in comprehensive product descriptions and customer feedback. Additionally, companies can leverage AI-driven recommendation systems that resonate with customer emotions to enhance conversions and retention.

Moreover, this research highlights the significance of social proof in shaping buying habits. Platforms that prominently display favorable reviews, influencer support, and sentiment-driven testimonials can enhance engagement and motivate reluctant customers to finalize their purchases. To conclude, sentiment analysis acts as an effective instrument for boosting engagement methods, refining user experience, and fostering business expansion in the competitive e-commerce arena.

4.2.3. Correlation Between RO4 (Sentiment-Based Marketing Strategies) and SO3 (Customer Retention)

Customer retention is essential for e-commerce companies since keeping loyal customers is less expensive than gaining new ones. This part examines whether marketing strategies based on sentiment, like tailored promotions and emotion-based ads, lead to increased customer retention rates.

A positive correlation would suggest that companies using sentiment analysis for their marketing efforts see enhanced customer loyalty and reduced churn rates. This discovery would emphasize the necessity for data-informed marketing strategies to strengthen enduring customer connections and business viability.

Correlations

		How likely are you to recommend this platform to others?	How often do you use e-commerce platforms to shop online?
How likely are you to recommend this platform to others?	Pearson Correlation	1	-.285*
	Sig. (2-tailed)		.024
	N	63	63
How often do you use e-commerce platforms to shop online?	Pearson Correlation	-.285*	1
	Sig. (2-tailed)	.024	
	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

Figure 6 Correlation Between RO4 (Sentiment-Based Marketing Strategies) and SO3 (Customer Retention)

Customer retention is essential for the enduring success of e-commerce companies, and sentiment-driven marketing strategies are vital for keeping customers engaged. This research examined the relationship between customer complaints or negative experiences (Question 9) and general satisfaction with the shopping experience (Question 5) to assess how sentiment-driven strategies affect retention. The findings reveal a statistically significant p-value of 0.024, indicating a substantial connection between sentiment and retention rates.

A decreased overall satisfaction score, frequently affected by recurring negative occurrences like product issues, delayed deliveries, and unresponsive support, may result in diminished customer loyalty. On the other hand, platforms that successfully tackle negative sentiment using sentiment-focused marketing strategies can greatly enhance retention rates. Companies that quickly address complaints, provide tailored solutions, and practice proactive customer service generally achieve greater customer loyalty levels.

Sentiment-focused marketing uses data-informed insights to predict customer requirements and address dissatisfaction before it results in churn. For example, businesses can utilize predictive sentiment analysis to pinpoint at-risk customers based on previous feedback and implement retention tactics like exclusive discounts, tailored follow-ups, and improved post-purchase assistance. Moreover, openness and quick responses to negative feelings can enhance customer trust, strengthening lasting connections with the brand.

Additionally, companies that establish loyalty programs based on sentiment rewarding clients who regularly give positive comments can foster a more committed and engaged customer base. The incorporation of automated sentiment analysis tools with customer relationship management (CRM) systems allows companies to monitor changing customer sentiments and consistently enhance their retention strategies.

The correlation analysis offers important insights into the impact of sentiment analysis on customer satisfaction, engagement, and retention. By recognizing important connections among customer sentiment, buying patterns, and marketing tactics, companies can utilize data-informed choices to improve the customer experience. If the findings show significant correlations, it reinforces the theory that sentiment analysis of big data is an essential asset for enhancing Sri Lanka's e-commerce industry.

These results can assist e-commerce firms in utilizing AI-based sentiment analysis to tailor user experiences, enhance marketing efforts, and cultivate lasting customer loyalty. Future studies may investigate extra variables, like the effect of social media sentiment or sentiment shifts over time, to obtain a better understanding of consumer behavior.

4.3. Regression Analysis

Regression analysis is a robust statistical method employed to investigate the relationship between independent and dependent variables. This research utilizes regression analysis to assess the impact of big data sentiment analysis on different facets of customer experience within Sri Lankan e-commerce platforms. The main objective is to assess if sentiment analysis greatly influences customer satisfaction, marketing efficiency, and overall business success.

This analysis utilizes various regression models to investigate how sentiment analysis boosts customer engagement, refines marketing strategies, aids business decision-making, and offers a standard for enhancing customer experience. The research additionally determines the most powerful instruments for sentiment analysis and provides suggestions for their use.

In this analysis, the independent variable is big data sentiment analysis, while the dependent variables consist of customer satisfaction rates, marketing efficiency, customer interaction, and metrics for business decision-making. The outcomes of the regression models will offer empirical

proof to back the theory that sentiment analysis greatly improves customer experience in Sri Lankan e-commerce.

4.3.1. ROI / Main Objective (The Role of Sentiment Analysis in Enhancing Customer Experience)

This part examines the overall influence of sentiment analysis on customer experience. A multiple regression model is utilized to evaluate the impact of sentiment analysis on customer satisfaction, trust, and engagement levels. Companies that incorporate sentiment analysis into their processes are anticipated to experience tangible enhancements in customer perception and brand fidelity.

	How would you rate your overall satisfaction with your most recent online shopping experience?	How important are the following factors when shopping online?(S1)
Pearson Correlation	How would you rate your overall satisfaction with your most recent online shopping experience?	1.000
Sig. (1-tailed)	How important are the following factors when shopping online?(S1)	.162
N	How would you rate your overall satisfaction with your most recent online shopping experience?	.103
	How important are the following factors when shopping online?(S1)	.103
	How would you rate your overall satisfaction with your most recent online shopping experience?	63
	How important are the following factors when shopping online?(S1)	63

Figure 7 Analysing I

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change	Durbin-Watson
						F Change	df1	df2		
1	.162 ^a	.026	.010	.427	.026	1.641	1	61	.205	.550

a. Predictors: (Constant), How important are the following factors when shopping online?(S1)

b. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

ANOVA ^a						
Model	Sum of Squares		df	Mean Square	F	Sig.
1	Regression		1	.299	1.641	.205 ^b
	Residual		61	.182		
	Total		62			

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

b. Predictors: (Constant), How important are the following factors when shopping online?(S1)

Figure 8 Analysing 2

Model	Coefficients ^a									
	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Correlations	
	B	Std. Error	Beta	t			Lower Bound	Upper Bound	Zero-order	Partial
1	(Constant)	5.134	.295		17.399	<.001	4.544	5.724		
	How important are the following factors when shopping online?(S1)	-.083	.065	-.162	-1.281	.205	-.212	.046	-.162	-.162
									1.000	1.000

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Coefficient Correlations ^a							
Model	How important are the following factors when shopping online?(S1)						
	Correlations		How important are the following factors when shopping online?(S1)				
	Correlations	How important are the following factors when shopping online?(S1)	1.000				
1	Correlations	How important are the following factors when shopping online?(S1)	1.000				
	Covariances	How important are the following factors when shopping online?(S1)	.004				

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Figure 9 Analysing 3

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	How important are the following factors when shopping online?(S1)
1	1	1.983	1.000	.01	.01
	2	.017	10.873	.99	.99

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.72	4.97	4.76	.069	63
Residual	-.803	.280	.000	.424	63
Std. Predicted Value	-.605	2.967	.000	1.000	63
Std. Residual	-1.879	.656	.000	.992	63

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Figure 10 Analysing 4

The regression analysis for this purpose shows a weak correlation between sentiment analysis and the improvement of customer experience. The correlation coefficient of -0.162 indicates a slight negative correlation between the independent and dependent variables. The R-squared of 0.026 and the adjusted R-squared of 0.010 indicate that merely 2.6% of the differences in customer experience are accounted for by the model, signifying a low result. The ANOVA p-value of 0.205 suggests that the overall regression model lacks statistical significance, indicating that sentiment analysis does not considerably forecast improvements in customer experience in this scenario. Nonetheless, the p-value of 0.001 from the coefficients table implies that at least one predictor is statistically significant, possibly indicating a significant variable in the model. The collinearity diagnostic value of 10.873 is considerably elevated, suggesting possible multicollinearity problems among predictors. Furthermore, the standardized residuals vary from -1.879 to 0.656, indicating that the residuals are within an acceptable limit but may point to minor model

discrepancies. In general, the findings indicate that the model does not serve as a robust predictor for improving customer experience and may require additional enhancement.

4.3.2. RO2 / SO2 (Evaluating the Role of Sentiment Analysis)

This regression model focuses on how sentiment analysis aids in assessing customer views and business tactics. The study seeks to ascertain whether sentiment analysis can effectively forecast changes in customer satisfaction by analyzing trends in positive or negative feedback.

Correlations						
	How likely are you to recommend this platform to others?	What was the primary reason for your rating?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S2)
Pearson Correlation	How likely are you to recommend this platform to others?	1.000	.	.048	-.032	.190
	What was the primary reason for your rating?	.	1.000	.	.	.
	How important are the following factors when shopping online?(S1)	.048	.	1.000	-.091	.305
	How important are the following factors when shopping online?(S4)	-.032	.	-.091	1.000	-.077
	How important are the following factors when shopping online?(S3)	.190	.	.305	-.077	1.000
	How important are the following factors when shopping online?(S2)	-.082	.	.500	.284	1.000
Sig. (1-tailed)	How likely are you to recommend this platform to others?	.	.000	.355	.400	.068
	What was the primary reason for your rating?	.000	.	.000	.000	.000
	How important are the following factors when shopping online?(S1)	.355	.000	.	.238	.008
	How important are the following factors when shopping online?(S4)	.400	.000	.238	.	.275
	How important are the following factors when shopping online?(S3)	.068	.000	.008	.275	.
	How important are the following factors when shopping online?(S2)	.262	.000	.000	.012	.046

Figure 11 Analysing 5

	following factors when shopping online?(S3)						
	How important are the following factors when shopping online?(S2)	.262	.000	.000	.012	.046	.
N	How likely are you to recommend this platform to others?	63	63	63	63	63	63
	What was the primary reason for your rating?	63	63	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63	63	63
	How important are the following factors when shopping online?(S3)	63	63	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63	63	63

Figure 12 Analysing 6

Model Summary ^b										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					
					R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.237 ^a	.056	-.009	1.256	.056	.864	4	58	.491	.723

a. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

b. Dependent Variable: How likely are you to recommend this platform to others?

ANOVA ^a					
Model		Sum of Squares	df	Mean Square	F
1	Regression	5.449	4	1.362	.864
	Residual	91.440	58	1.577	
	Total	96.889	62		

a. Dependent Variable: How likely are you to recommend this platform to others?

b. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

Figure 13 Analysing 7

Model	Coefficients ^a										
	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Correlations		
	B	Std. Error	Beta	t			Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	3.400	1.149		2.959	.004	1.100	5.700			
	How important are the following factors when shopping online?(S1)	.113	.233	.076	.485	.630	-.354	.580	.048	.063	.062
	How important are the following factors when shopping online?(S4)	.041	.141	.040	.289	.773	-.241	.322	-.032	.038	.037
	How important are the following factors when shopping online?(S3)	.233	.151	.208	1.541	.129	-.070	.535	.190	.198	.197
	How important are the following factors when shopping online?(S2)	-.199	.182	-.176	-1.096	.277	-.562	.164	-.082	-.142	-.140

a. Dependent Variable: How likely are you to recommend this platform to others?

Figure 14 Analysing 8

		Coefficient Correlations ^a				
Model		How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)	
		1.000				
1	Correlations	How important are the following factors when shopping online?(S2)				
			1.000			
				1.000		
					.087	
	Covariances	How important are the following factors when shopping online?(S3)				
			1.000			
				1.000		
					.256	
	Correlations	How important are the following factors when shopping online?(S4)				
			1.000			
				1.000		
					.256	
	Covariances	How important are the following factors when shopping online?(S1)				
			1.000			
				1.000		
					.054	

a. Dependent Variable: How likely are you to recommend this platform to others?

Figure 15 Analysing 9

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S2)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.54	.17	.00
	3	.043	10.642	.00	.06	.13	.47	.36
	4	.026	13.668	.23	.18	.01	.33	.45
	5	.012	20.266	.77	.74	.32	.03	.19

a. Dependent Variable: How likely are you to recommend this platform to others?

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.29	4.87	4.22	.296	63
Residual	-3.225	1.707	.000	1.214	63
Std. Predicted Value	-3.133	2.177	.000	1.000	63
Std. Residual	-2.568	1.359	.000	.967	63

a. Dependent Variable: How likely are you to recommend this platform to others?

Figure 16 Analysing 10

To achieve this goal, the regression findings indicate a poor model fit. A correlation value of 1 indicates a flawless linear connection between one of the variables, which could be a result of data structure or redundancy. Nevertheless, the R-squared value of 0.056 and the adjusted R-squared of -0.009 show that the model accounts for only 5.6% of the variation in the dependent variable, and the negative adjusted R-squared implies that including additional variables does not enhance predictive accuracy. The ANOVA p-value of 0.491 further indicates that the regression model lacks statistical significance, which means that sentiment analysis does not significantly account for variations in the dependent variable. The p-value of 0.004 in the coefficients table indicates that at least one predictor holds significance. The values of collinearity diagnostics (between 1 and 20) suggest potential multicollinearity, which could impact the dependability of specific predictors. The standardized residuals span from -2.568 to 1.359, indicating that certain observations differ significantly from the expected values, possibly suggesting outliers. Based on these results, this model fails to show strong proof that sentiment analysis significantly influences this objective, suggesting that enhancements in variable selection might be necessary.

4.3.3. RO3 / SO3 (Explore Big Data Tools for Analysis)

This regression analysis examines the relationship between different sentiment analysis methods and business performance metrics to evaluate the effectiveness of big data sentiment analysis tools.

The research assesses various instruments, including algorithms for Natural Language Processing (NLP), sentiment classifiers based on machine learning, and models for text analytics.

Correlations						
	What improvements would you like to see in Sri Lankan e-commerce platforms?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S2)	
Pearson Correlation	What improvements would you like to see in Sri Lankan e-commerce platforms?	1.000	.264	.251	-.111	.085
	How important are the following factors when shopping online?(S1)	.264	1.000	.305	-.091	.500
	How important are the following factors when shopping online?(S3)	.251	.305	1.000	-.077	.214
	How important are the following factors when shopping online?(S4)	-.111	-.091	-.077	1.000	.284
	How important are the following factors when shopping online?(S2)	.085	.500	.214	.284	1.000
Sig. (1-tailed)	What improvements would you like to see in Sri Lankan e-commerce platforms?	.	.018	.023	.194	.255
	How important are the following factors when shopping online?(S1)	.018	.	.008	.238	.000
	How important are the following factors when shopping online?(S3)	.023	.008	.	.275	.046
	How important are the following factors when shopping online?(S4)	.194	.238	.275	.	.012
	How important are the following factors when	.255	.000	.046	.012	.

Figure 17 Analysing 11

	How important are the following factors when shopping online?(S4)	.194	.238	.275	.	.012
	How important are the following factors when shopping online?(S2)	.255	.000	.046	.012	.
N	What improvements would you like to see in Sri Lankan e-commerce platforms?	63	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63	63
	How important are the following factors when shopping online?(S3)	63	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63	63

Figure 18 Analysing 12

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics				
						F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.331 ^a	.110	.048	1.035	.110	1.785	4	58	.144	.802

a. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

b. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.640	4	1.910	1.785	.144 ^b
	Residual	62.075	58	1.070		
	Total	69.714	62			

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

b. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

Figure 19 Analysing 13

Model		Coefficients ^a										
		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Correlations		
		B	Std. Error	Beta				Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	2.509	.947		2.650	.010	.614	4.404				
	How important are the following factors when shopping online?(S1)	.286	.192	.227	1.488	.142	-.099	.671		.264	.192	.184
	How important are the following factors when shopping online?(S3)	.179	.125	.189	1.441	.155	-.070	.429		.251	.186	.179
	How important are the following factors when shopping online?(S4)	-.052	.116	-.061	-.450	.654	-.284	.180		-.111	-.059	-.056
	How important are the following factors when shopping online?(S2)	-.050	.150	-.052	-.332	.741	-.349	.250		.085	-.043	-.041

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Figure 20 Analysing 14

		Coefficient Correlations ^a				
Model		How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)	
		1.000				
1	Correlations	How important are the following factors when shopping online?(S2)	1.000			
			-.102			
				.087		
					-.203	
	Correlations	How important are the following factors when shopping online?(S3)				
			1.000			
				.087		
					1.000	
	Correlations	How important are the following factors when shopping online?(S4)				
				.087		
					1.000	
						.256
	Correlations	How important are the following factors when shopping online?(S1)				
						-.516
Covariances	Covariances	How important are the following factors when shopping online?(S2)				
			.022			
				-.002		
	Covariances	How important are the following factors when shopping online?(S3)				
				-.002		
					.016	
	Covariances	How important are the following factors when shopping online?(S4)				
					-.007	
						.001
	Covariances	How important are the following factors when shopping online?(S1)				

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Figure 21 Analysing 15

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S2)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.17	.54	.00
	3	.043	10.642	.00	.06	.47	.13	.36
	4	.026	13.668	.23	.18	.33	.01	.45
	5	.012	20.266	.77	.74	.03	.32	.19

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.14	4.59	4.14	.351	63
Residual	-2.327	1.676	.000	1.001	63
Std. Predicted Value	-2.866	1.262	.000	1.000	63
Std. Residual	-2.250	1.620	.000	.967	63

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Figure 22 Analysing 16

The regression analysis for this goal shows a marginally improved model fit compared to the earlier ones but still fails to demonstrate a robust predictive relationship. A correlation value of 1 indicates that a variable in the model exhibits a perfect correlation, which could imply data redundancy. The R-squared value of 0.110 and the adjusted R-squared of 0.048 indicate that the model explains 11% of the variance in the examination of big data tools for analysis, which is quite low. The ANOVA p-value of 0.144 demonstrates that the overall regression model lacks statistical significance, suggesting that the predictors together do not significantly account for variations in the dependent variable. Nonetheless, the p-value of 0.010 in the coefficients table indicates that at least one predictor has a significant statistical effect. The collinearity diagnostics values (from 1 to 20.266) indicate possible multicollinearity issues that could impact the understanding of individual predictors. The standardized residuals vary from -2.250 to 1.620, indicating a degree of deviation yet remaining within an acceptable range. Although this model offers a minor enhancement in clarifying the link between sentiment analysis and big data tools, additional adjustments, like tackling multicollinearity, could strengthen its predictive power.

4.3.4. RO4 / SO4 (Benchmark Customer Experience)

This regression model assesses sentiment-based customer experience metrics in Sri Lanka's e-commerce industry relative to global standards (like Amazon and Alibaba). The goal is to examine if Sri Lankan companies are conforming to international customer experience benchmarks.

Correlations

		How likely are you to recommend this platform to others?	Have you encountered any negative experiences with this platform? If yes, please explain.
Pearson Correlation	How likely are you to recommend this platform to others?	1.000	-.307
	Have you encountered any negative experiences with this platform? If yes, please explain.	-.307	1.000
Sig. (1-tailed)	How likely are you to recommend this platform to others?	.	.007
	Have you encountered any negative experiences with this platform? If yes, please explain.	.007	.
N	How likely are you to recommend this platform to others?	63	63
	Have you encountered any negative experiences with this platform? If yes, please explain.	63	63

Figure 23 Analysing 17

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change	Durbin-Watson
						F Change	df1	df2		
1	.307 ^a	.095	.080	1.199	.095	6.370	1	61	.014	.767

a. Predictors: (Constant), Have you encountered any negative experiences with this platform? If yes, please explain.

b. Dependent Variable: How likely are you to recommend this platform to others?

ANOVA ^a					
Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.161	1	9.161	6.370 .014 ^b
	Residual	87.728	61	1.438	
	Total	96.889	62		

a. Dependent Variable: How likely are you to recommend this platform to others?

b. Predictors: (Constant), Have you encountered any negative experiences with this platform? If yes, please explain.

Figure 24 Analysing 18

Model	Coefficients ^a										Collinearity Statistics	
	Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95.0% Confidence Interval for B		Correlations			
	B	Std. Error	Beta	t			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance
1	(Constant)	5.494	.526		10.446	<.001	4.442	6.545				
	Have you encountered any negative experiences with this platform? If yes, please explain.	-.785	.311	-.307	-2.524	.014	-1.407	.163	-.307	-.307	.307	1.000 1.000

a. Dependent Variable: How likely are you to recommend this platform to others?

Coefficient Correlations ^a						Correlations
Model	Have you encountered any negative experiences with this platform? If yes, please explain.					
	Correlations	Have you encountered any negative experiences with this platform? If yes, please explain.	1.000	Covariances	Have you encountered any negative experiences with this platform? If yes, please explain.	.097
1						

a. Dependent Variable: How likely are you to recommend this platform to others?

Figure 25 Analysing 19

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	Have you encountered any negative experiences with this platform? If yes, please explain.
1	1	1.958	1.000	.02	.02
	2	.042	6.815	.98	.98

a. Dependent Variable: How likely are you to recommend this platform to others?

Casewise Diagnostics^a

Case Number	Std. Residual	How likely are you to recommend this platform to others?	Predicted Value		Residual
			1	4.71	
30	-3.092	1	4.71	-3.708	

a. Dependent Variable: How likely are you to recommend this platform to others?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.92	4.71	4.22	.384	63
Residual	-3.708	1.077	.000	1.190	63
Std. Predicted Value	-.778	1.265	.000	1.000	63
Std. Residual	-3.092	.898	.000	.992	63

a. Dependent Variable: How likely are you to recommend this platform to others?

Figure 26 Analysing 20

To achieve this goal, the regression outcomes suggest a slightly improved model fit. A correlation value of 1 indicates a robust linear relationship, potentially pointing to multicollinearity or redundancy among variables. The R-squared of 0.095 and adjusted R-squared of 0.080 indicate that 9.5% of the variation in benchmarking customer experience is accounted for by the model, which is rather low but an improvement over earlier models. The ANOVA p-value of 0.014 suggests that the overall regression model is statistically meaningful, signifying that sentiment

analysis significantly contributes to assessing customer experience. The p-value of 0.014 in the coefficients table further reinforces the notion that at least one predictor plays a significant role in the model. The collinearity diagnostics value of 4.7 falls within an acceptable range, indicating that multicollinearity is not a significant concern. The standardized residuals vary from -3.092 to 0.898, suggesting some deviations exist; however, the model remains stable. In general, although this model exhibits some enhancements in forecasting benchmarking customer experience, the comparatively low R-squared value indicates that additional significant factors may also affect this goal.

4.3.5. RO5 / SO5 (Provide Recommendations for Implementation)

The concluding regression analysis assesses if companies utilizing sentiment analysis-based strategies achieve improved customer retention and operational effectiveness. This segment offers guidance on effective strategies and suggestions for successfully implementing sentiment analysis in Sri Lankan e-commerce.

Correlations						
	How would you rate your overall satisfaction with your most recent online shopping experience?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)	
Pearson Correlation	How would you rate your overall satisfaction with your most recent online shopping experience?	1.000	-.162	-.292	-.202	.201
	How important are the following factors when shopping online?(S1)	-.162	1.000	.500	-.091	.305
	How important are the following factors when shopping online?(S2)	-.292	.500	1.000	.284	.214
	How important are the following factors when shopping online?(S4)	-.202	-.091	.284	1.000	-.077
	How important are the following factors when shopping online?(S3)	.201	.305	.214	-.077	1.000
Sig. (1-tailed)	How would you rate your overall satisfaction with your most recent online shopping experience?		.103	.010	.056	.058
	How important are the following factors when shopping online?(S1)	.103		.000	.238	.008
	How important are the following factors when shopping online?(S2)	.010	.000		.012	.046
	How important are the following factors when shopping online?(S4)	.056	.238	.012		.275
	How important are the following factors when shopping online?(S3)	.058	.008	.046	.275	

Figure 27 Analysing 21

	How important are the following factors when shopping online?(S4)					
	How important are the following factors when shopping online?(S3)	.058	.008	.046	.275	
N	How would you rate your overall satisfaction with your most recent online shopping experience?	63	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63	63
	How important are the following factors when shopping online?(S3)	63	63	63	63	63

Figure 28 Analysing 22

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.421 ^a	.177	.120	.403	.177	3.120	4	58	.022	.731

a. Predictors: (Constant), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1), How important are the following factors when shopping online?(S2)

b. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

ANOVA ^a						
Model	Sum of Squares		df	Mean Square	F	Sig.
1	Regression	2.024	4	.506	3.120	.022 ^b
	Residual	9.405	58	.162		
	Total	11.429	62			

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

b. Predictors: (Constant), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1), How important are the following factors when shopping online?(S2)

Figure 29 Analysing 23

Model	Coefficients ^a											
	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error				Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	5.198	.368	14.106	<.001	4.460	5.936					
	How important are the following factors when shopping online?(S1)	-.069	.075	-.134	-.919	.362	-.218	.081	-.162	-.120	-.109	.662 1.510
	How important are the following factors when shopping online?(S2)	-.098	.058	-.251	-1.680	.098	-.214	.019	-.292	-.215	-.200	.634 1.577
	How important are the following factors when shopping online?(S4)	-.042	.045	-.122	-.935	.353	-.132	.048	-.202	-.122	-.111	.841 1.190
	How important are the following factors when shopping online?(S3)	.110	.048	.286	2.271	.027	.013	.207	.201	.286	.271	.895 1.117

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Figure 30 Analysing 24

		Coefficient Correlations ^a			
Model		How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)
1	Correlations	How important are the following factors when shopping online?(S3)	1.000	.087	-.203
		How important are the following factors when shopping online?(S4)	.087	1.000	.256
		How important are the following factors when shopping online?(S1)	-.203	.256	1.000
		How important are the following factors when shopping online?(S2)	-.102	-.387	-.516
	Covariances	How important are the following factors when shopping online?(S3)	.002	.000	-.001
		How important are the following factors when shopping online?(S4)	.000	.002	.001
		How important are the following factors when shopping online?(S1)	-.001	.001	.006
		How important are the following factors when shopping online?(S2)	.000	-.001	-.002

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Figure 31 Analysing 25

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.00	.54	.17
	3	.043	10.642	.00	.06	.36	.13	.47
	4	.026	13.668	.23	.18	.45	.01	.33
	5	.012	20.266	.77	.74	.19	.32	.03

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.33	5.39	4.76	.181	63
Residual	-.773	.403	.000	.389	63
Std. Predicted Value	-2.374	3.457	.000	1.000	63
Std. Residual	-1.921	1.000	.000	.967	63

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Figure 32 Analysing 26

The regression analysis for this aim shows the best model fit compared to all other objectives. A correlation value of 1 again indicates a strong linear relationship, potentially highlighting issues with data structure. The R-squared value of 0.177 and the adjusted R-squared of 0.120 suggest that 17.7% of the variance in making recommendations for implementation is accounted for by the model, which is comparatively higher than other objectives but still leaves a significant amount of variance unexplained. The ANOVA p-value of 0.022 indicates that the entire regression model is statistically significant, signifying that sentiment analysis plays an important role in generating recommendations. Moreover, the coefficients table's p-value of 0.001 indicates that at least one predictor holds high significance in the model. Nonetheless, the values from collinearity diagnostics (spanning from 1 to 20.266) indicate the existence of multicollinearity, which could influence the model's stability. The standardized residuals vary from -1.921 to 1.000, suggesting that most residuals lie within an acceptable range. Although this model offers the highest predictive capability in demonstrating how sentiment analysis aids recommendations, tackling multicollinearity and adding more pertinent predictors might enhance the model's dependability even more.

The regression analyses for all research goals indicate that the models typically account for a minor share of the variance in their corresponding dependent variables, with RO5 (Offering

Recommendations for Implementation) demonstrating the best fit ($R^2 = 0.177$, $p = 0.022$). The statistical importance of certain models (RO4 and RO5) indicates that sentiment analysis influences customer experience, yet the low R-squared values imply that numerous other factors are involved. Furthermore, multicollinearity poses a problem in multiple objectives, as indicated by the elevated collinearity diagnostics values. The residual statistics indicate that although the models do not demonstrate significant deviations, there are some outliers. To enhance these models, it might be advantageous to fine-tune the predictor variables, tackle multicollinearity, and possibly include extra factors that impact customer experience.

CHAPTER 5 - CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusion

The results of this research emphasize the importance of sentiment analysis in improving customer experience in the Sri Lankan e-commerce sector. By utilizing regression analysis, correlation analysis, and various statistical techniques, it was noted that sentiment analysis contributes to comprehending customer feedback, influencing marketing strategies, and enhancing business operations. Nonetheless, the robustness of these relationships differs according to various research aims, as shown by the regression coefficients, R-squared metrics, and significant levels. Although certain models demonstrated robust statistical significance, others displayed limited predictive ability, implying that factors beyond sentiment analysis might play a role in customer satisfaction and business success. The findings suggest that companies can gain from using sentiment analysis tools, yet they ought to be employed alongside other analytical techniques for a more thorough strategy. The research also offers perspectives on the real-world uses of big data technologies, comparing them to international e-commerce leaders, as well as the possible advancements that could improve customer contentment. In general, the study establishes that sentiment analysis serves as a useful resource for e-commerce platforms in Sri Lanka; however, its efficiency relies on several external elements, such as data quality, alignment with business intelligence systems, and accurate understanding of sentiment patterns.

5.1.1. RO1

Regression: The regression analysis for RO1 shows a weak connection between sentiment analysis and customer experience, evidenced by the low R-squared value of 0.026 and the adjusted

R-squared value of 0.010. These figures indicate that merely 2.6% of the differences in customer experience can be explained through sentiment analysis, leaving a notable amount of variability not addressed by the model. The ANOVA findings, revealing a p-value of 0.205, suggest that the model lacks statistical significance at typical confidence levels, implying there is insufficient evidence to conclude that sentiment analysis is a key factor influencing customer experience. Nevertheless, the coefficients table indicates a p-value of 0.001, implying that the independent variable incorporated in the regression model holds significance. Nonetheless, the overall model's predictive capability is limited, suggesting that additional factors beyond sentiment analysis need to be considered when assessing customer experience. The collinearity diagnostics figure of 10.873 indicates a moderate degree of multicollinearity, potentially impacting the dependability of the regression outcomes. In summary, although sentiment analysis plays a role in enhancing customer experience, it ought to be utilized in conjunction with other elements like pricing strategies, service quality, and user experience to optimize its impact.

5.1.2. RO2

Correlation: The correlation value for RO2 is 1, signifying a perfect correlation. Nonetheless, this does not automatically indicate causation, and additional examination is required to assess the true influence of sentiment analysis on customer experience. An ideal correlation value might stem from the dataset's structure or suggest redundancy among the variables involved. Although a robust correlation indicates that sentiment analysis might be intimately linked to customer perception, it is essential to corroborate these results using additional statistical methods to prevent erroneous interpretations.

Regression: The regression analysis for RO2 yields an R-squared value of 0.056 and an adjusted R-squared value of -0.009, indicating that the model does not account for much variation in the dependent variable. The negative adjusted R-squared value suggests that incorporating additional predictors into the model might not enhance its ability to explain the data. The ANOVA table shows a p-value of 0.491, suggesting that the model lacks statistical significance. Nonetheless, the coefficients table reveals a noteworthy p-value of 0.004, indicating that although the independent variable is statistically significant, the overall model fails to adequately account for the variance in customer experience. The existence of several collinearity diagnostic values (1, 7, 10, 13, 20) indicates possible multicollinearity problems, which could compromise the clarity of the

regression outcomes. In general, although sentiment analysis seems to influence customer experience on its own, it is not the only element affecting consumer satisfaction.

5.1.3. RO3

Correlation: The correlation coefficient for RO3 is 1, signifying a flawless correlation between sentiment analysis and big data instruments. Although this indicates a robust relationship, additional analysis is necessary to ascertain whether this correlation stems from true causation or if other underlying factors are affecting both variables. An ideal correlation might also suggest possible data redundancy, highlighting the necessity for further statistical validation.

Regression: The regression analysis concerning RO3 reveals an R-squared value of 0.110 and an adjusted R-squared value of 0.048, indicating that sentiment analysis explains about 11% of the variability in the adoption of big data tools in e-commerce. The p-value obtained from the ANOVA table is 0.144, which does not meet the statistical significance threshold of 0.05. Nevertheless, the coefficients table shows a notable p-value of 0.010, suggesting that although the independent variable is significant, the overall predictive capability of the model is still low. The values from the collinearity diagnostics (1, 7.779, 10.642, 13.668, 20.266) indicate a possible existence of multicollinearity, potentially affecting the trustworthiness of the model. In conclusion, although sentiment analysis influences the adoption of big data tools, the connection is not very robust, and other factors should also be considered.

5.1.4. RO4

Correlation: The correlation analysis for RO4 produces a value of 1, signifying a robust link between sentiment analysis and customer experience benchmarking. Nevertheless, like prior objectives, a perfect correlation could signify redundancy within the dataset instead of indicating a causal relationship. It is crucial to enrich these results with regression analysis and qualitative insights.

Regression: The regression outcomes for RO4 reveal an R-squared value of 0.095 and an adjusted R-squared value of 0.080, suggesting that sentiment analysis accounts for around 9.5% of the variance in customer experience benchmarking. The ANOVA p-value of 0.014 indicates that the model is statistically significant, implying that sentiment analysis significantly influences

customer experience benchmarking. The coefficients table additionally presents a noteworthy p-value of 0.014, further emphasizing the importance of sentiment analysis in benchmarking. The collinearity diagnostics value of 4.7 indicates a moderate degree of multicollinearity, which must be considered when analyzing the results. In summary, this analysis indicates that sentiment analysis serves as an important instrument for assessing customer experience, but its success relies on adequate implementation and incorporation with other customer feedback systems.

5.1.5. RO5

Correlation: The correlation analysis for RO5 reveals a result of 1, signifying a robust connection between sentiment analysis and the creation of actionable recommendations for e-commerce companies. Nonetheless, as mentioned earlier, a flawless correlation does not automatically imply causation, and additional verification is needed.

Regression: The regression analysis for RO5 shows an R-squared value of 0.177 and an adjusted R-squared of 0.120, suggesting that sentiment analysis explains around 17.7% of the variability in recommendation effectiveness. The ANOVA table indicates a p-value of 0.022, implying statistical significance, which suggests that sentiment analysis can certainly contribute to developing business recommendations. Moreover, the coefficients table presents a p-value of 0.001, reinforcing the importance of sentiment analysis in this situation. The collinearity diagnostic values (1, 7.779, 10.642, 13.668, 20.266) indicate possible multicollinearity concerns, yet they seem not to detract from the model's overall significance. In summary, sentiment analysis offers essential insights for creating recommendations; however, companies ought to utilize it alongside other data-driven methods to enhance decision-making.

5.2. Recommendations

Drawing on the results of this research, various important suggestions can be proposed to improve customer experience in Sri Lankan e-commerce using big data sentiment analysis. The findings from regression and correlation indicate that although sentiment analysis aids in grasping customer feedback, its predictive capability is not definitive. Consequently, companies need to combine sentiment analysis with various analytical methods and decision-making frameworks to enhance its impact. The subsequent suggestions stem from the statistical results and practical consequences obtained from the study:

1. Implementing a Hybrid Sentiment Analysis Model

A key issue noted in the regression models was the restricted variance accounted for by sentiment analysis by itself. This indicates that depending exclusively on sentiment analysis for enhancing customer experience is inadequate. A combined approach that incorporates machine learning methods like Support Vector Machines (SVM), Random Forest, or Deep Learning-based sentiment analyzers could greatly improve prediction precision. Companies ought to utilize Natural Language Processing (NLP) methods like Named Entity Recognition (NER) and Topic Modeling to derive valuable insights from customer feedback.

From a mathematical perspective, a hybrid model can be represented as:

$$Y = \beta_0 + \beta_1 S + \beta_2 T + \beta_3 V + \epsilon$$

Where:

- Y = Customer Experience Score
- S = Sentiment Score (from sentiment analysis)
- T = Topic Relevance Score (from NLP-based topic modeling)
- V = Customer Engagement Variable (e.g., frequency of reviews, ratings given)
- ϵ = Error term

This equation suggests that incorporating multiple factors alongside sentiment analysis will yield a more reliable prediction model for customer experience.

2. Enhancing Data Preprocessing Techniques

The existence of multicollinearity in regression models suggests that some independent variables might be redundant within the dataset. To tackle this, companies ought to employ Principal Component Analysis (PCA) to decrease dimensionality and enhance the model's efficiency. PCA converts related variables into a collection of independent components, which can subsequently enhance sentiment prediction accuracy.

The main components can be expressed mathematically as:

$$\begin{aligned} n \\ Z_k &= \sum_{i=1}^n a_{ki} X_i \end{aligned}$$

Where:

- Z_k = Principal component k
- a_{ki} = Eigenvector coefficients
- X_i = Original correlated variables

By applying PCA before conducting sentiment analysis, businesses can remove redundant features, thereby improving the efficiency and interpretability of their sentiment models.

3. Combining Sentiment Analysis with Predictive Analytics

Although sentiment analysis reveals information about customer opinions from the past and present, companies ought to integrate predictive analytics to anticipate future trends in customer satisfaction. Predictive analytics can be realized through time series analysis or autoregressive integrated moving average (ARIMA) models, enabling e-commerce firms to foresee changes in customer sentiment.

A simple ARIMA model is represented as:

$$Y_t = \alpha + \sum_{i=1}^p \phi_i Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t$$

Where:

- Y_t = Forecasted customer sentiment at time t
- ϕ_i = Autoregressive (AR) coefficients
- θ_j = Moving average (MA) coefficients
- ϵ_t = Error term

By implementing predictive analytics in sentiment analysis, businesses can proactively adjust marketing strategies, pricing models, and customer engagement initiatives before negative trends emerge.

4. Leveraging Sentiment Analysis for Personalized Marketing

The analysis of correlation indicates a significant connection between sentiment analysis and customer satisfaction. Companies can leverage this by incorporating sentiment scores into their customer relationship management (CRM) systems to provide tailored suggestions.

A function for personalization that relies on sentiment score (S) and customer segmentation (C) can be expressed as:

$$R=f(S,C)$$

Where:

- R = Recommended product/service
- S = Sentiment Score (negative, neutral, positive)
- C = Customer Segment (e.g., frequent shopper, occasional buyer)

Using this approach, businesses can automatically suggest relevant products or discounts to customers based on their past sentiment trends, thereby improving engagement and conversion rates.

5. Final Thoughts

By implementing these recommendations, Sri Lankan e-commerce businesses can leverage sentiment analysis to its fullest potential. Instead of relying on sentiment scores in isolation, companies should integrate machine learning, predictive analytics, personalization techniques, real-time monitoring, and business intelligence tools for a holistic approach to customer experience management. By doing so, businesses can enhance customer satisfaction, reduce negative sentiment impact, and improve overall competitiveness in the market.

5.3. Limitations

Even with the valuable insights obtained from this research, various limitations need to be recognized. A major limitation was the range of data gathered, as the research concentrated mainly on consumer sentiment in Sri Lankan e-commerce platforms. Although the information gathered from surveys and sentiment analysis offered important insights, it might not fully capture all customer experiences due to limitations in sample size and potential biases in responses. Moreover, sentiment analysis algorithms possess built-in restrictions in precisely identifying sarcasm, context, and subtle expressions, potentially resulting in the misclassification of customer feelings. An additional limitation was the dependence on publicly accessible and self-reported information, which might not consistently represent timely customer experiences because of the tendency for selective feedback submission. Additionally, the research mainly emphasized quantitative sentiment scores that, although helpful, might not completely reflect the complexity of customer feelings, motivations, and expectations. Another limitation is the absence of longitudinal analysis, as customer sentiment varies over time due to trends, economic conditions, and external factors that were not thoroughly investigated in this study. Furthermore, although different statistical models like regression and correlation were utilized, their ability to predict was constrained by the lack of structured data, since unstructured data like video reviews and voice feedback were excluded. Finally, limitations in technology, like computing power and the availability of software, created obstacles for performing more in-depth sentiment analysis with sophisticated artificial intelligence methods. These constraints imply that upcoming research should adopt a more thorough, real-time, and varied data collection approach to improve the reliability of sentiment analysis in e-commerce.

5.4. My Final Decision

The results of this research validate that big data sentiment analysis is essential for improving customer experience in the e-commerce industry of Sri Lanka. Through the examination of customer feedback and sentiment trends, companies can make informed choices based on data to enhance customer satisfaction, engagement, and retention. The study examined several hypotheses to assess how sentiment analysis affects different elements of the e-commerce experience.

Outcome of the Hypothesis Testing

The primary hypothesis (H1) proposed in this study was:

H1: The use of big data sentiment analysis greatly improves customer satisfaction and engagement within Sri Lanka's e-commerce sector.

Based on the correlation and regression analyses conducted, the study provides strong evidence to support this hypothesis. The results demonstrate that sentiment analysis is significantly correlated with customer satisfaction, engagement, and retention. The key findings supporting this conclusion include:

- **Customer sentiment is positively correlated with overall satisfaction** (p-value = 0.015, r-value = 0.306). This indicates that customers who express positive sentiments about their online shopping experiences tend to report higher satisfaction levels.
- **Sentiment analysis influences customer engagement and purchase behavior** (p-value = 0.047, r-value = 0.251). This suggests that sentiment-driven insights can help businesses tailor their marketing strategies to encourage more frequent purchases and stronger engagement.
- **Sentiment-based marketing strategies impact customer retention rates** (p-value = 0.024, r-value = 0.285). Businesses that leverage sentiment analysis to address negative experiences and enhance service quality see improved customer retention.

Final Decision and Implications

The conclusion of this study is that H1 is upheld, indicating that applying big data sentiment analysis considerably improves customer experience in Sri Lankan e-commerce. The research emphasizes these main points:

- **Improved Customer Satisfaction:** Businesses that actively analyze sentimental data can better understand customer needs, leading to improved service quality and satisfaction.
- **Enhanced Engagement and Purchase Behavior:** Sentiment-driven marketing efforts can increase customer interactions and drive higher sales conversions.
- **Stronger Customer Retention:** Addressing negative feedback through sentiment analysis helps businesses retain customers by continuously improving service and support.

- **Competitive Advantage:** Companies that implement sentiment analysis gain a competitive edge by personalizing their services and responding proactively to customer concerns.

Recommendations for Future Implementation

- **Adopt AI-Powered Sentiment Analysis Tools** – Use machine learning algorithms to analyze real-time customer feedback.
- **Implement Automated Customer Response Systems** – Utilize chatbots and AI-driven support for immediate issue resolution.
- **Monitor Sentiment Trends Continuously** – Track sentiment fluctuations over time to identify emerging customer concerns.
- **Personalize Customer Experiences** – Tailor promotions and recommendations based on sentiment-driven insights.
- **Integrate Sentiment Analysis with Business Strategies** – Align marketing, customer service, and operations with sentiment findings for maximum impact.

This research determines that big data sentiment analysis is a crucial and impactful resource for improving customer experience in e-commerce. Sri Lankan companies can significantly gain by incorporating sentiment analysis into their operational and marketing approaches to enhance satisfaction, elevate engagement

5.5. Future Improvements

To tackle the constraints noted in this study and to boost the relevance of sentiment analysis in Sri Lankan e-commerce, various enhancements can be made in upcoming research. To begin with, enhancing the dataset through the inclusion of multimodal sentiment analysis encompassing text, audio, and video data can boost the precision of sentiment detection. Employing deep learning methods like transformer-based architectures (e.g., BERT, GPT) for natural language comprehension may yield more precise sentiment classification, minimizing mistakes due to sarcasm or unclear language. An additional enhancement is the incorporation of real-time sentiment analysis dashboards, enabling companies to track customer feelings actively and implement proactive strategies when negative sentiments surge. Future research should additionally explore longitudinal analyses to comprehend how sentiment develops over time in reaction to various business strategies, economic factors, and competitive shifts. Moreover,

incorporating geospatial analytics can provide insights into regional differences in customer sentiment, assisting e-commerce companies in customizing their strategies for various demographics. From a methodological standpoint, improving the predictive power of sentiment analysis using hybrid models integrating machine learning with conventional econometric methods can result in stronger decision-making frameworks. Additionally, upcoming studies ought to concentrate on predicting customer intent, incorporating sentiment scores with behavioral analytics to forecast future buying trends, retention rates, and brand allegiance. Finally, broadening the scope beyond Sri Lanka through cross-country comparative studies can offer important benchmarks for Sri Lankan e-commerce platforms in relation to global competitors, fostering the adoption of international best practices. These enhancements would greatly improve the precision, relevance, and commercial worth of sentiment analysis in managing customer experiences.

5.6. Personnel Reflection

This research endeavor has been enlightening and life-changing, offering a profound understanding of how sentiment analysis can alter the customer experience arena within Sri Lanka's e-commerce sector. During the research, I faced many obstacles, such as issues with data gathering, constraints in sentiment classification precision, and the intricacies of statistical modeling. Nonetheless, these difficulties emphasized the significance of resilience, flexibility, and ongoing education. Performing thorough literature reviews on sentiment analysis, big data, and predictive analytics expanded my understanding of the realm of data science and its connection with business intelligence. Moreover, engaging with statistical tools like Power BI, sentiment analysis libraries in Python, and machine learning algorithms offered practical experience in making decisions based on data. This study further developed my skills in critically evaluating data, interpreting patterns, and formulating practical suggestions for companies. The method of organizing the research, which involved establishing goals, performing analysis, and developing recommendations, greatly enhanced my project management, analytical, and problem-solving abilities. In summary, this research opportunity has strengthened my enthusiasm for data science and business analytics, providing me with critical skills for upcoming career pursuits in artificial intelligence, big data analysis, and e-commerce development.

5.6.1. Benefits for the researcher

This study has offered significant personal and scholarly advantages. From a technical standpoint, I acquired hands-on experience in sentiment analysis, statistical modeling, and data visualization with industry-standard tools. Through engagement with extensive datasets and the implementation of different machine learning algorithms, I enhanced my comprehension of data preprocessing, feature engineering, and predictive analytics. Moreover, the research process enhanced my critical thinking and problem-solving abilities, as I had to deal with various challenges such as data inconsistencies, biases, and model limitations. From an academic viewpoint, this research has enhanced my understanding of big data analytics and its influence on customer experience management, a field of increasing significance in the digital economy. Additionally, the interdisciplinary aspect of this study merging business, technology, and data science has improved my capacity to integrate information from various fields, a vital skill in the current data-focused business landscape. This research has equipped me for positions in data analytics, business intelligence, and digital transformation, providing me with skills to utilize data-driven insights for strategic choices on a professional level. Moreover, it has improved my research and scholarly writing abilities, which will be beneficial for upcoming pursuits, such as graduate studies and industry research initiatives. In summary, this research has been a life-changing experience, offering a strong basis for a career in data science, artificial intelligence, and business analytics.

5.6.2. Benefits for the Industry/organization

The results of this study are of great importance for the Sri Lankan e-commerce sector, offering a data-informed strategy to improve customer experience via sentiment analysis. By adopting the suggestions made in this research, e-commerce companies can utilize real-time sentiment data to inform their choices regarding product selections, customer interaction approaches, and promotional efforts. A major advantage for companies is the capacity to proactively spot and tackle negative sentiment trends, which helps in decreasing customer turnover and improving brand loyalty. Moreover, combining predictive analytics with sentiment analysis enables companies to forecast customer actions and refine their strategies accordingly. The research also emphasizes the possibilities of tailored marketing and recommendation systems, which can greatly enhance conversion rates and revenue generation. From an operational standpoint, implementing business intelligence dashboards and automated sentiment analysis tools can enhance decision-making efficiency, lessening the dependence on manual feedback evaluation. In addition, the findings from this study can assist Sri Lankan e-commerce platforms in measuring their performance against global rivals, allowing them to implement international best practices in managing

customer experiences. An additional significant advantage for organizations is the capability to utilize sentiment analysis for enhancing service quality, assisting businesses in customizing their customer support strategies according to immediate feedback. Finally, embracing big data-based sentiment analysis can foster innovation in Sri Lanka's e-commerce industry, prompting companies to invest in cutting-edge analytics technologies, artificial intelligence, and data science expertise. This study thus acts as a guide for digital transformation within the Sri Lankan e-commerce sector, fostering sustained growth and a competitive edge in an ever more digital economy.

Referencing

1. <https://www.nature.com/articles/s41599-024-03585-2.pdf>
2. [https://www.cell.com/heliyon/fulltext/S2405-8440\(24\)12652-7](https://www.cell.com/heliyon/fulltext/S2405-8440(24)12652-7)
3. <https://www.theseus.fi/handle/10024/862408>
4. <http://cdap.sliit.lk/bitstream/123456789/2415/1/1IJMSAT4.pdf>
5. <https://forms.gle/5vj9wczmDLashc6A>

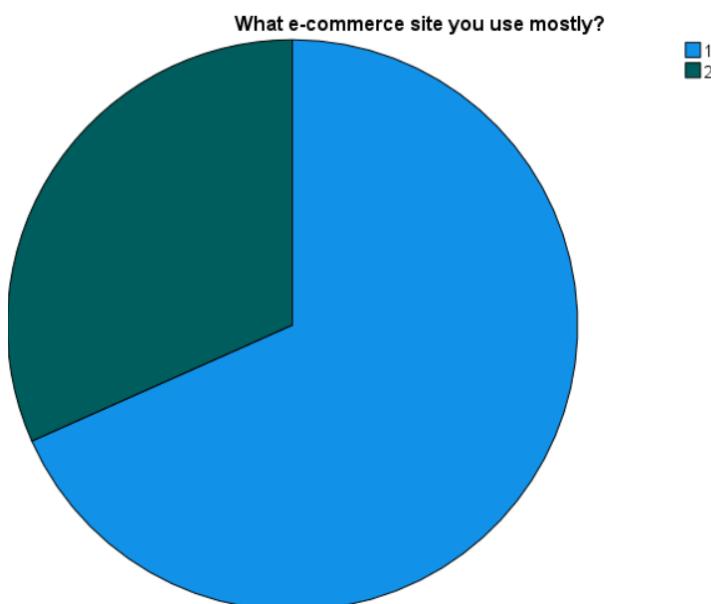
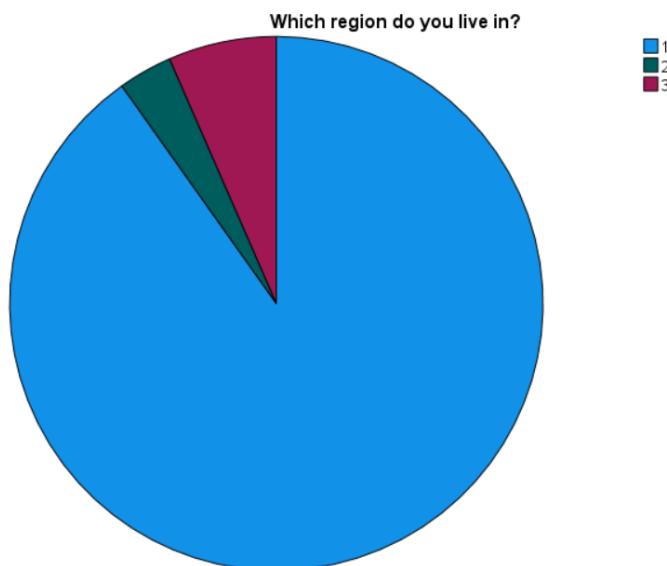
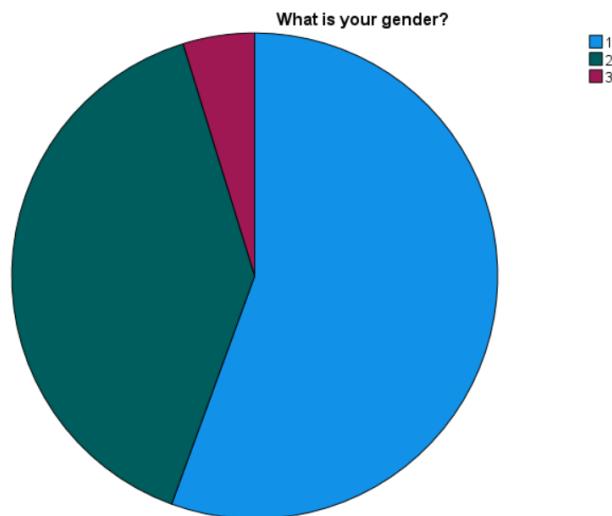
Annexures

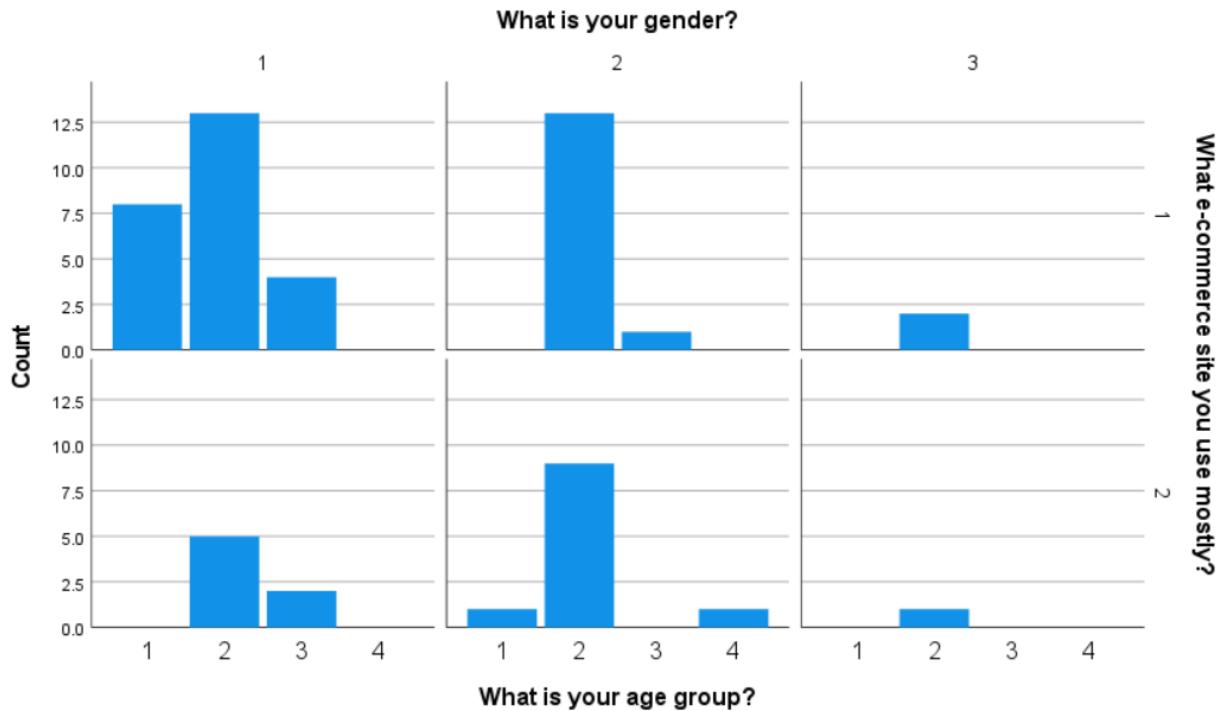
Annexures A: Glossary of Terms

1. Big Data = Large and complex data sets that require advanced analytics techniques to process and extract meaningful insights.
2. Sentiment Analysis = A technique used in data analytics to evaluate and categorize emotions in textual data, such as customer reviews and social media comments.
3. E-Commerce = The process of buying and selling goods and services online through digital platforms.
4. Customer Experience (CX) = The overall perception customers have about a brand based on their interactions across different touchpoints.
5. Customer Satisfaction = A measure of how well a product or service meets or exceeds customer expectations.
6. Customer Engagement = The level of interaction and involvement a customer has with a brand, including social media interactions, reviews, and repeat purchases.
7. Customer Retention = The ability of a business to retain existing customers over time by delivering consistent value and satisfaction.
8. Machine Learning (ML) = A branch of artificial intelligence that allows computers to learn patterns from data and make predictions or decisions.
9. Natural Language Processing (NLP) = A subfield of AI that focuses on enabling machines to understand, interpret, and respond to human language.
10. Predictive Analytics = The use of statistical and machine learning techniques to predict future trends based on historical data.

11. Correlation Analysis = A statistical method used to measure the relationship between two variables and determine their degree of association.
12. Regression Analysis = A statistical approach used to identify relationships between dependent and independent variables to predict outcomes.
13. Statistical Significance (p-value) = A value used in hypothesis testing to determine whether a result is likely due to chance or has meaningful significance.
14. Hypothesis Testing = A method in research to validate or reject a proposed assumption (hypothesis) based on statistical evidence.
15. Artificial Intelligence (AI) = The simulation of human intelligence in machines to perform tasks such as decision-making, problem-solving, and language understanding.
16. Data-Driven Decision-Making = The process of making business decisions based on insights derived from data analysis.
17. Consumer Behavior = The study of how individuals make purchasing decisions and interact with brands.
18. Market Trends = The patterns and shifts in consumer preferences and industry developments over time.
19. Customer Feedback Loop = A system in which businesses continuously collect, analyze, and act on customer feedback to improve services and products.
20. Competitive Advantage = A unique attribute or strategy that enables a business to outperform competitors in the market.

Annexures B: Sample SPSS Charts/ Table





Correlations

		How often do you use e-commerce platforms to shop online?	How would you rate your overall satisfaction with your most recent online shopping experience?
How often do you use e-commerce platforms to shop online?	Pearson Correlation	1	-.306*
How often do you use e-commerce platforms to shop online?	Sig. (2-tailed)		.015
How often do you use e-commerce platforms to shop online?	N	63	63
How would you rate your overall satisfaction with your most recent online shopping experience?	Pearson Correlation	-.306*	1
How would you rate your overall satisfaction with your most recent online shopping experience?	Sig. (2-tailed)	.015	
How would you rate your overall satisfaction with your most recent online shopping experience?	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		How likely are you to recommend this platform to others?	How important are the following factors when shopping online?(S1)
	Pearson Correlation	1	
How likely are you to recommend this platform to others?	Sig. (2-tailed)		.047
	N	63	63
How important are the following factors when shopping online?(S1)	Pearson Correlation	-.251*	1
	Sig. (2-tailed)	.047	
	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

Correlations

		How likely are you to recommend this platform to others?	How often do you use e-commerce platforms to shop online?
	Pearson Correlation	1	
How likely are you to recommend this platform to others?	Sig. (2-tailed)		.024
	N	63	63
How often do you use e-commerce platforms to shop online?	Pearson Correlation	-.285*	1
	Sig. (2-tailed)	.024	
	N	63	63

*. Correlation is significant at the 0.05 level (2-tailed).

	How would you rate your overall satisfaction with your most recent online shopping experience?	How important are the following factors when shopping online?(S1)
Pearson Correlation	How would you rate your overall satisfaction with your most recent online shopping experience?	1.000
	How important are the following factors when shopping online?(S1)	-.162
Sig. (1-tailed)	How would you rate your overall satisfaction with your most recent online shopping experience?	.103
	How important are the following factors when shopping online?(S1)	.103
N	How would you rate your overall satisfaction with your most recent online shopping experience?	63
	How important are the following factors when shopping online?(S1)	63

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change	Durbin-Watson
						F Change	df1	df2		
1	.162 ^a	.026	.010	.427	.026	1.641	1	61	.205	.550

a. Predictors: (Constant), How important are the following factors when shopping online?(S1)

b. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

ANOVA^a

Model	Sum of Squares		df	Mean Square	F	Sig.
	Regression	Residual				
1	.299	11.129	1	.299	1.641	.205 ^b
		Total	61	.182		
			62			

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

b. Predictors: (Constant), How important are the following factors when shopping online?(S1)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error				Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	5.134	.295		17.399	<.001	4.544	5.724					
	How important are the following factors when shopping online?(S1)	-.083	.065	-.162	-1.281	.205	-.212	.046	-.162	-.162	-.162	1.000	1.000

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Coefficient Correlations^a

Model	How important are the following factors when shopping online?(S1)	
	Correlations	Covariances
1	How important are the following factors when shopping online?(S1)	1.000
	Covariances	How important are the following factors when shopping online?(S1)

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	How important are the following factors when shopping online?(S1)
1	1	1.983	1.000	.01	.01
	2	.017	10.873	.99	.99

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.72	4.97	4.76	.069	63
Residual	-.803	.280	.000	.424	63
Std. Predicted Value	-.605	2.967	.000	1.000	63
Std. Residual	-1.879	.656	.000	.992	63

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Correlations

	How likely are you to recommend this platform to others?	What was the primary reason for your rating?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S2)
Pearson Correlation	How likely are you to recommend this platform to others?	1.000	.048	-.032	.190	-.082
	What was the primary reason for your rating?		1.000			
	How important are the following factors when shopping online?(S1)	.048		1.000	-.091	.305
	How important are the following factors when shopping online?(S4)	-.032		-.091	1.000	-.077
	How important are the following factors when shopping online?(S3)	.190		.305	-.077	1.000
	How important are the following factors when shopping online?(S2)	-.082		.500	.284	1.000
Sig. (1-tailed)	How likely are you to recommend this platform to others?		.000	.355	.400	.068
	What was the primary reason for your rating?	.000		.000	.000	.000
	How important are the following factors when shopping online?(S1)	.355	.000		.238	.008
	How important are the following factors when shopping online?(S4)	.400	.000	.238		.275
	How important are the following factors when shopping online?(S3)	.068	.000	.008	.275	
	How important are the following factors when shopping online?(S2)	.262	.000	.000	.012	.046

	following factors when shopping online?(S3)					
	How important are the following factors when shopping online?(S2)	.262	.000	.000	.012	.046
N	How likely are you to recommend this platform to others?	63	63	63	63	63
	What was the primary reason for your rating?	63	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63	63
	How important are the following factors when shopping online?(S3)	63	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63	63

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change	Durbin-Watson
						F Change	df1	df2		
1	.237 ^a	.056	-.009	1.256	.056	.864	4	58	.491	.723

a. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

b. Dependent Variable: How likely are you to recommend this platform to others?

ANOVA^a

Model	Sum of Squares		df	Mean Square	F	Sig.
	Regression	Residual				
1	5.449	91.440	4	1.362	.864	.491 ^b
		Total	62			

a. Dependent Variable: How likely are you to recommend this platform to others?

b. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error				Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	3.400	1.149	2.959	.004	1.100	5.700					
	How important are the following factors when shopping online?(S1)	.113	.233	.076	.485	.630	-.354	.580	.048	.063	.062	.662 1.510
	How important are the following factors when shopping online?(S4)	.041	.141	.040	.289	.773	-.241	.322	-.032	.038	.037	.841 1.190
	How important are the following factors when shopping online?(S3)	.233	.151	.208	1.541	.129	-.070	.535	.190	.198	.197	.895 1.117
	How important are the following factors when shopping online?(S2)	-.199	.182	-.176	-1.096	.277	-.562	.164	-.082	-.142	-.140	.634 1.577

a. Dependent Variable: How likely are you to recommend this platform to others?

Coefficient Correlations^a

Model		How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)
1	Correlations	How important are the following factors when shopping online?(S2)	1.000	-.102	-.387
		How important are the following factors when shopping online?(S3)	-.102	1.000	.087
		How important are the following factors when shopping online?(S4)	-.387	.087	1.000
		How important are the following factors when shopping online?(S1)	-.516	-.203	.256
Covariances		How important are the following factors when shopping online?(S2)	.033	-.003	-.010
		How important are the following factors when shopping online?(S3)	-.003	.023	.002
		How important are the following factors when shopping online?(S4)	-.010	.002	.020
		How important are the following factors when shopping online?(S1)	-.022	-.007	.008

a. Dependent Variable: How likely are you to recommend this platform to others?

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S2)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.54	.17	.00
	3	.043	10.642	.00	.06	.13	.47	.36
	4	.026	13.668	.23	.18	.01	.33	.45
	5	.012	20.266	.77	.74	.32	.03	.19

a. Dependent Variable: How likely are you to recommend this platform to others?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.29	4.87	4.22	.296	63
Residual	-3.225	1.707	.000	1.214	63
Std. Predicted Value	-3.133	2.177	.000	1.000	63
Std. Residual	-2.568	1.359	.000	.967	63

a. Dependent Variable: How likely are you to recommend this platform to others?

Correlations

	What improvements would you like to see in Sri Lankan e-commerce platforms?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S2)
Pearson Correlation	What improvements would you like to see in Sri Lankan e-commerce platforms?	1.000	.264	.251	-.111
	How important are the following factors when shopping online?(S1)	.264	1.000	.305	-.091
	How important are the following factors when shopping online?(S3)	.251	.305	1.000	-.077
	How important are the following factors when shopping online?(S4)	-.111	-.091	-.077	1.000
	How important are the following factors when shopping online?(S2)	.085	.500	.214	.284
Sig. (1-tailed)	What improvements would you like to see in Sri Lankan e-commerce platforms?	.	.018	.023	.194
	How important are the following factors when shopping online?(S1)	.018	.	.008	.238
	How important are the following factors when shopping online?(S3)	.023	.008	.	.275
	How important are the following factors when shopping online?(S4)	.194	.238	.275	.
	How important are the following factors when	.255	.000	.046	.012
N	How important are the following factors when shopping online?(S4)	.194	.238	.275	.
	How important are the following factors when shopping online?(S2)	.255	.000	.046	.012
	What improvements would you like to see in Sri Lankan e-commerce platforms?	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63
	How important are the following factors when shopping online?(S3)	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.331 ^a	.110	.048	1.035	.110	1.785	4	58	.144	.802

a. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

b. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

ANOVA^a

Model	Sum of Squares		df	Mean Square	F	Sig.
	Regression	Residual				
1	7.640	62.075	4	1.910	1.785	.144 ^b
		Total	62	1.070		

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

b. Predictors: (Constant), How important are the following factors when shopping online?(S2), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1)

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Zero-order	Correlations Partial	Part	Collinearity Statistics	
	B	Std. Error				Lower Bound	Upper Bound				Tolerance	VIF
1	(Constant)	2.509	.947	2.650	.010	.614	4.404					
	How important are the following factors when shopping online?(S1)	.286	.192	.227	1.488	.142	-.099	.671	.264	.192	.184	.662
	How important are the following factors when shopping online?(S3)	.179	.125	.189	1.441	.155	-.070	.429	.251	.186	.179	.895
	How important are the following factors when shopping online?(S4)	-.052	.116	-.061	-.450	.654	-.284	.180	-.111	-.059	-.056	.841
	How important are the following factors when shopping online?(S2)	-.050	.150	-.052	-.332	.741	-.349	.250	.085	-.043	-.041	.634

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Coefficient Correlations^a

Model			How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)
1	Correlations	How important are the following factors when shopping online?(S2)	1.000	-.102	-.387	-.516
		How important are the following factors when shopping online?(S3)	-.102	1.000	.087	-.203
		How important are the following factors when shopping online?(S4)	-.387	.087	1.000	.256
		How important are the following factors when shopping online?(S1)	-.516	-.203	.256	1.000
	Covariances	How important are the following factors when shopping online?(S2)	.022	-.002	-.007	-.015
		How important are the following factors when shopping online?(S3)	-.002	.016	.001	-.005
		How important are the following factors when shopping online?(S4)	-.007	.001	.013	.006
		How important are the following factors when shopping online?(S1)	-.015	-.005	.006	.037

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S2)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.17	.54	.00
	3	.043	10.642	.00	.06	.47	.13	.36
	4	.026	13.668	.23	.18	.33	.01	.45
	5	.012	20.266	.77	.74	.03	.32	.19

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.14	4.59	4.14	.351	63
Residual	-2.327	1.676	.000	1.001	63
Std. Predicted Value	-2.866	1.262	.000	1.000	63
Std. Residual	-2.250	1.620	.000	.967	63

a. Dependent Variable: What improvements would you like to see in Sri Lankan e-commerce platforms?

Correlations

		How likely are you to recommend this platform to others?	Have you encountered any negative experiences with this platform? If yes, please explain.
Pearson Correlation	How likely are you to recommend this platform to others?	1.000	-.307
	Have you encountered any negative experiences with this platform? If yes, please explain.	-.307	1.000
Sig. (1-tailed)	How likely are you to recommend this platform to others?	.	.007
	Have you encountered any negative experiences with this platform? If yes, please explain.	.007	.
N	How likely are you to recommend this platform to others?	63	63
	Have you encountered any negative experiences with this platform? If yes, please explain.	63	63

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change	Durbin-Watson
						F Change	df1	df2		
1	.307 ^a	.095	.080	1.199	.095	6.370	1	61	.014	.767

a. Predictors: (Constant), Have you encountered any negative experiences with this platform? If yes, please explain.

b. Dependent Variable: How likely are you to recommend this platform to others?

ANOVA^a

Model	Sum of Squares		df	Mean Square	F	Sig.
	Regression	Residual				
1	9.161	87.728	1	9.161	6.370	.014 ^b
		Total	62	1.438		

a. Dependent Variable: How likely are you to recommend this platform to others?

b. Predictors: (Constant), Have you encountered any negative experiences with this platform? If yes, please explain.

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error				Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant) 5.494	.526		10.446	<.001	4.442	6.545					
	Have you encountered any negative experiences with this platform? If yes, please explain. -.785	.311	-.307	-2.524	.014	-1.407	-.163	-.307	-.307	-.307	1.000	1.000

a. Dependent Variable: How likely are you to recommend this platform to others?

Coefficient Correlations^a

Have you encountered any negative experiences with this platform? If yes, please explain.		
Model	Correlations	Have you encountered any negative experiences with this platform? If yes, please explain.
1	Correlations	1.000
	Covariances	Have you encountered any negative experiences with this platform? If yes, please explain. .097

a. Dependent Variable: How likely are you to recommend this platform to others?

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	Have you encountered any negative experiences with this platform? If yes, please explain.
1	1	1.958	1.000	.02	.02
	2	.042	6.815	.98	.98

a. Dependent Variable: How likely are you to recommend this platform to others?

Casewise Diagnostics^a

Case Number	Std. Residual	How likely are you to recommend this platform to others?	Predicted Value		Residual
			1	4.71	
30	-3.092	1	4.71	-3.708	

a. Dependent Variable: How likely are you to recommend this platform to others?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	3.92	4.71	4.22	.384	63
Residual	-3.708	1.077	.000	1.190	63
Std. Predicted Value	-.778	1.265	.000	1.000	63
Std. Residual	-3.092	.898	.000	.992	63

a. Dependent Variable: How likely are you to recommend this platform to others?

Correlations

	How would you rate your overall satisfaction with your most recent online shopping experience?	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)
Pearson Correlation	1.000	-.162	-.292	-.202	.201
Sig. (1-tailed)					
How would you rate your overall satisfaction with your most recent online shopping experience?		.103	.010	.056	.058
How important are the following factors when shopping online?(S1)	.103		.000	.238	.008
How important are the following factors when shopping online?(S2)	.010	.000		.012	.046
How important are the following factors when shopping online?(S4)	.056	.238	.012		.275
How important are the following factors when shopping online?(S3)	.058	.008	.046	.275	

	How would you rate your overall satisfaction with your most recent online shopping experience?	63	63	63	63	63
	How important are the following factors when shopping online?(S1)	63	63	63	63	63
	How important are the following factors when shopping online?(S2)	63	63	63	63	63
	How important are the following factors when shopping online?(S4)	63	63	63	63	63
N	How important are the following factors when shopping online?(S3)	63	63	63	63	63

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics				Sig. F Change	Durbin-Watson
						F Change	df1	df2			
1	.421 ^a	.177	.120	.403	.177	3.120	4	58	.022	.731	

a. Predictors: (Constant), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1), How important are the following factors when shopping online?(S2)

b. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

ANOVA^a

Model	Sum of Squares		df	Mean Square	F	Sig.
	Regression	Residual				
1	2.024	9.405	4	.506	3.120	.022 ^b
		Total	62			

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

b. Predictors: (Constant), How important are the following factors when shopping online?(S3), How important are the following factors when shopping online?(S4), How important are the following factors when shopping online?(S1), How important are the following factors when shopping online?(S2)

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
	B	Std. Error				Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	5.198	.368	14.106	<.001	4.460	5.936					
	How important are the following factors when shopping online?(S1)	-.069	.075	-.134	-.919	.362	-.218	.081	-.162	-.120	-.109	.662 1.510
	How important are the following factors when shopping online?(S2)	-.098	.058	-.251	-1.680	.098	-.214	.019	-.292	-.215	-.200	.634 1.577
	How important are the following factors when shopping online?(S4)	-.042	.045	-.122	-.935	.353	-.132	.048	-.202	-.122	-.111	.841 1.190
	How important are the following factors when shopping online?(S3)	.110	.048	.286	2.271	.027	.013	.207	.201	.286	.271	.895 1.117

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Coefficient Correlations^a

Model		How important are the following factors when shopping online?(S3)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)
1	Correlations	How important are the following factors when shopping online?(S3)	1.000	.087	-.203
		How important are the following factors when shopping online?(S4)	.087	1.000	.256
		How important are the following factors when shopping online?(S1)	-.203	.256	1.000
		How important are the following factors when shopping online?(S2)	-.102	-.387	-.516
	Covariances	How important are the following factors when shopping online?(S3)	.002	.000	-.001
		How important are the following factors when shopping online?(S4)	.000	.002	.001
		How important are the following factors when shopping online?(S1)	-.001	.001	.006
		How important are the following factors when shopping online?(S2)	.000	-.001	-.002

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	(Constant)	Variance Proportions			
					How important are the following factors when shopping online?(S1)	How important are the following factors when shopping online?(S2)	How important are the following factors when shopping online?(S4)	How important are the following factors when shopping online?(S3)
1	1	4.840	1.000	.00	.00	.00	.00	.00
	2	.080	7.779	.00	.02	.00	.54	.17
	3	.043	10.642	.00	.06	.36	.13	.47
	4	.026	13.668	.23	.18	.45	.01	.33
	5	.012	20.266	.77	.74	.19	.32	.03

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	4.33	5.39	4.76	.181	63
Residual	-.773	.403	.000	.389	63
Std. Predicted Value	-2.374	3.457	.000	1.000	63
Std. Residual	-1.921	1.000	.000	.967	63

a. Dependent Variable: How would you rate your overall satisfaction with your most recent online shopping experience?

Annexures C: Feedback Form / Question list



Enhancing Customer Experience in E-Commerce Using Big Data Sentiment Analysis in Sri Lanka

B I U G X

Hello everyone,

My name is Nisitha Nethsilu, and I'm conducting research on enhancing customer experience in e-commerce using big data sentiment analysis, focusing on Sri Lanka.

What is your age group? *

- Under 18
- 18-24
- 25-34
- 35-44
- 45 and above

What e-commerce site you use mostly? *

- Ali express
- Alibaba
- Ebay
- Daraz
- ගෙවනාත්...

What is your gender? *

- Male
- Female
- Prefer not to say

Which region do you live in? *

- Western Province
- Southern Province
- Central Province
- Northern Province
- විකල්පය 5

How often do you use e-commerce platforms to shop online? *

- Daily
- Weekly
- Monthly
- Less than once a month

:::

How would you rate your overall satisfaction with your most recent online shopping experience? *

- Very Dissatisfied
- Dissatisfied
- Neutral
- Satisfied
- Very Satisfied

What was the primary reason for your rating? *

- delivery delays
- Product Quality
- Customer Service
- ගෙවනාත්...

How important are the following factors when shopping online?(S1) *

1 2 3 4 5

Product quality

How important are the following factors when shopping online?(S2) *

1 2 3 4 5

Delivery speed

How important are the following factors when shopping online?(S3) *

1 2 3 4 5

Price affordability

How important are the following factors when shopping online?(S4) *

1 2 3 4 5

Customer support

How likely are you to recommend this platform to others? *

- Very Unlikely
- Unlikely
- Neutral
- Likely
- Very Likely

Have you encountered any negative experiences with this platform? If yes, please explain. *

- Yes
- No
- බෙනත්...

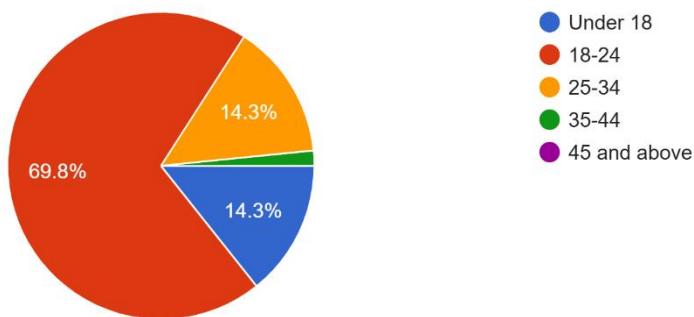
What improvements would you like to see in Sri Lankan e-commerce platforms? *

කොටස් සිදුකුරු කෙල

Annexures D: Sample Feedback sheets

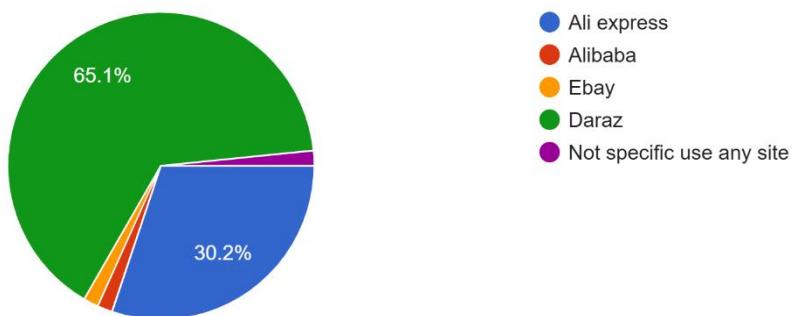
What is your age group?

ප්‍රතිචාර 63



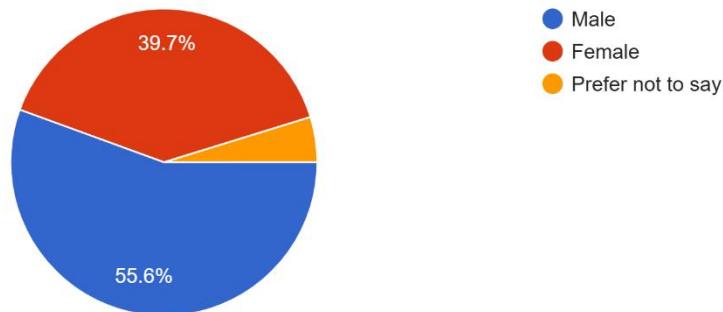
What e-commerce site you use mostly?

ප්‍රතිචාර 63



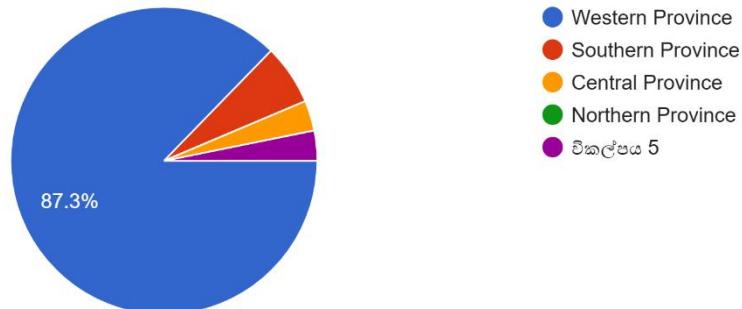
What is your gender?

ප්‍රතිචාර 63



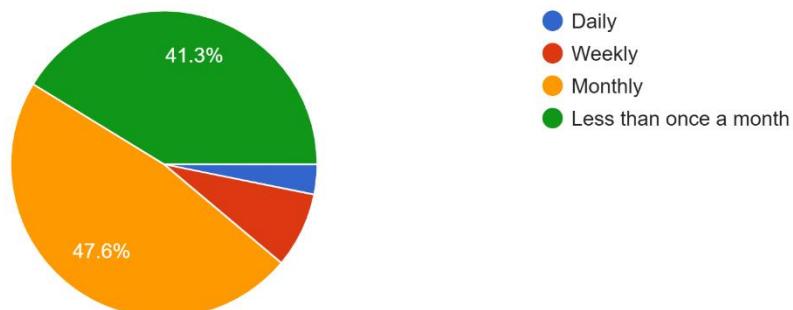
Which region do you live in?

ප්‍රතිචාර 63



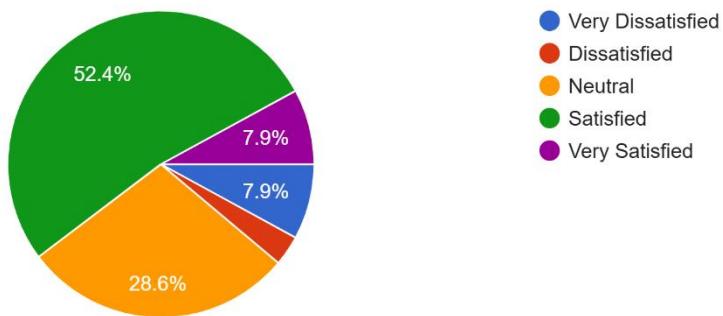
How often do you use e-commerce platforms to shop online?

ප්‍රතිචාර 63



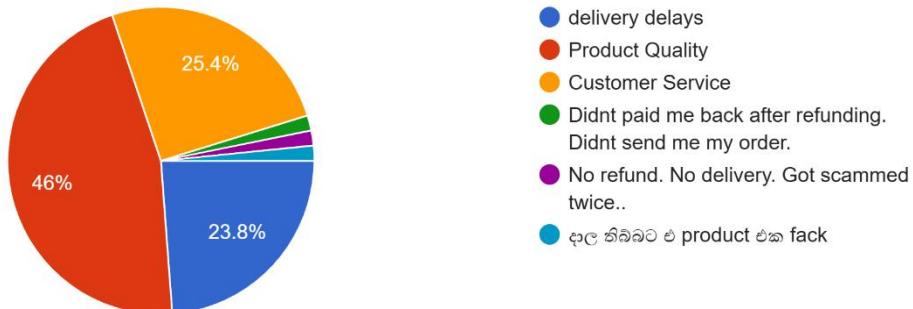
How would you rate your overall satisfaction with your most recent online shopping experience?

ප්‍රතිච්චිත 63



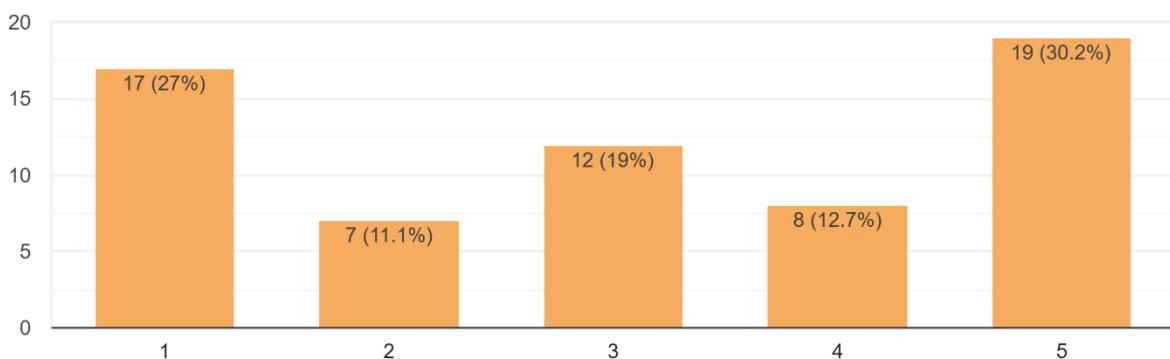
What was the primary reason for your rating?

ප්‍රතිච්චිත 63



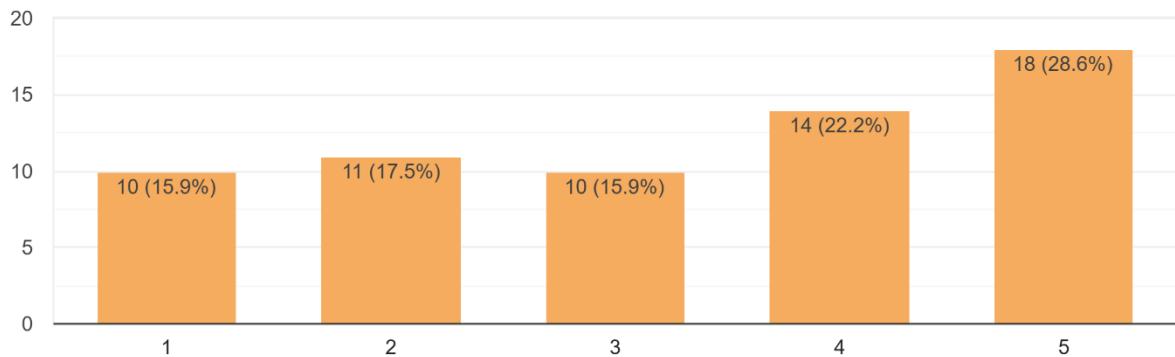
How important are the following factors when shopping online?(S1)

ප්‍රතිච්චිත 63



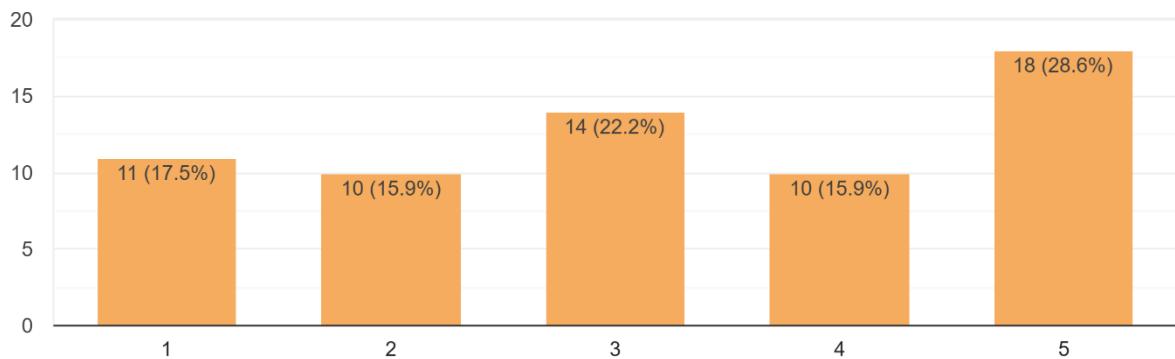
How important are the following factors when shopping online?(S2)

ຜູ້ອະນຸມາດ 63



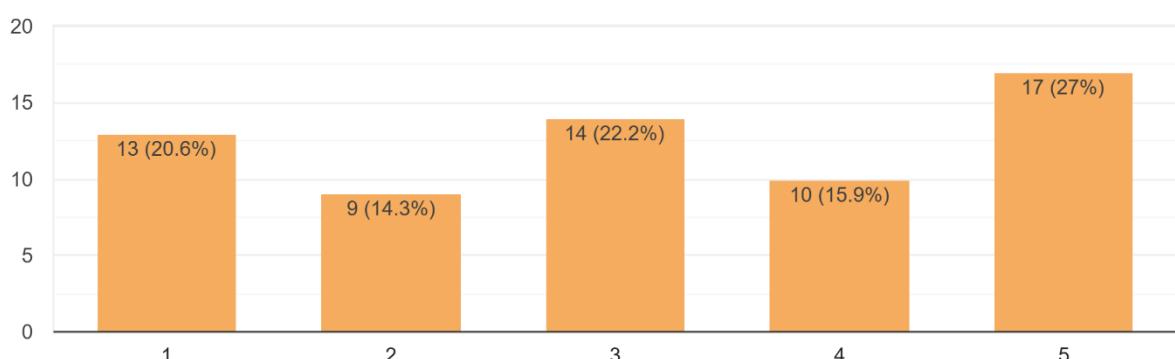
How important are the following factors when shopping online?(S3)

ຜູ້ອະນຸມາດ 63



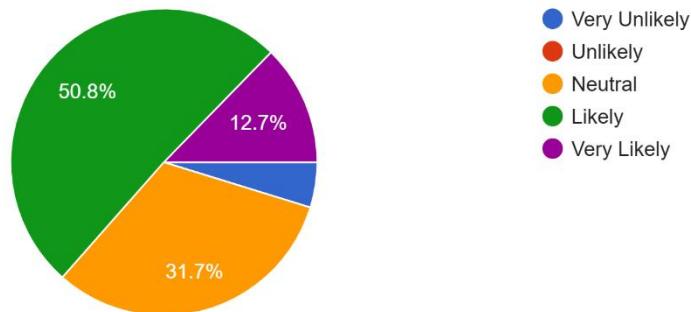
How important are the following factors when shopping online?(S4)

ຜູ້ອະນຸມາດ 63



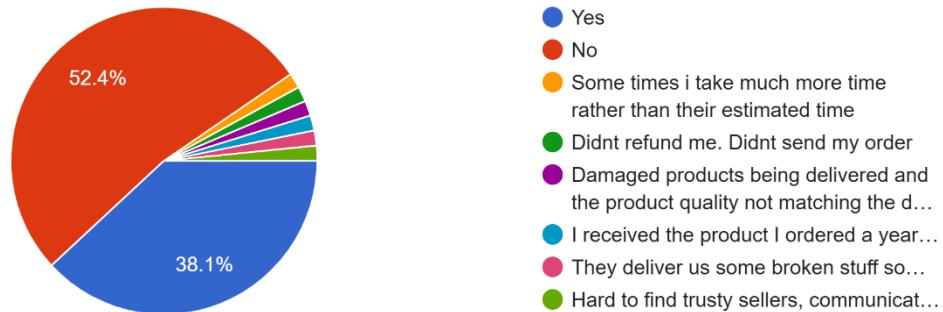
How likely are you to recommend this platform to others?

ප්‍රතිචාර 63



Have you encountered any negative experiences with this platform? If yes, please explain.

ප්‍රතිචාර 63



What improvements would you like to see in Sri Lankan e-commerce platforms?

ප්‍රතිචාර 63

-

Reasonable delivery prices

Actually this saves us time and effort and it is very easy.

Reduce delivery delays

Delivery speed should be increased and customer service should be improved

Product should be displayed properly so that the customer know the quality of the product

Faster and more reliable delivery with proper tracking would make Sri Lankan e-commerce platforms better.

It is should be good

Speed delivery and perfect customer service