

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)**

**A**

**MINI PROJECT REPORT**

ON

# “Customer Shopping Trends Analysis”

### Submitted in the partial fulfillment of the requirements in the 4th semester of

**BACHELOR OF ENGINEERING IN**

**COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)**

### BY

Kasi Ramji Y- 1NH23CD071

***Under the guidance of***

## Mr Sankadeep Pujaru

***Assistant Professor***

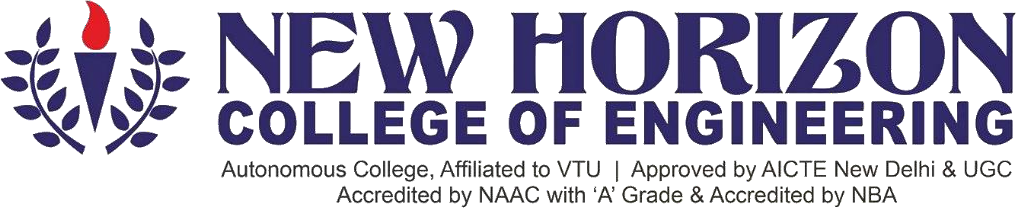
Dept. of CSE (DS), NHCE

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)**

**NEW HORIZON COLLEGE OF ENGINEERING**

(Autonomous College Permanently Affiliated to VTU, Approved by AICTE, Accredited by NBA & NAAC with ‘A’ Grade)

Ring Road, Bellandur Post, Near Marathalli, Bangalore-560103, INDIA



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)**

# CERTIFICATE

We hereby certify that, the report entitled **“Customer Shopping Trends Analysis”** as a part of Mini Project Component in partial fulfillment of the requirements during 4th semester Bachelor of Engineering in Computer Science and Engineering (Data Science) during the year 2024- 2025(March 2025 – June 2025) is an authentic record of our own work carried out by **Kasi Ramji Yalamarthi(1NH23CD071)**, bonafied students of NEW HORIZON COLLEGE OF ENGINEERING.

**Kasi Ramji Yalamarthi (1NH23CD071)**

**Name & Signature of Student**

**Name & Signature of Guide Name & signature of HOD**

(Mr Sankadeep Pujaru**)** (Dr. Swathi B)

ACKNOWLEDGEMENT

Any achievement, be it scholastic or otherwise does not depend solely on the individual efforts but on the guidance, encouragement and cooperation of intellectuals, elders and friends. A number of personalities, in their own capacities have helped us in carrying out this project. We would like to take an opportunity to thank them all.

First and foremost we thank the management, **Dr. Mohan Manghnani,** Chairman, New Horizon Educational Institutions for providing necessary infrastructure and creating good environment.

We would like to thank **Dr. Manjunatha**, Principal, New Horizon College of Engineering, Bengaluru, for his constant encouragement and facilities extended to us towards completing our project work.

We extend my sincere gratitude to **Dr. Swathi B**, Associate Professor & Head of the Department, Computer Science and Engineering (Data Science), New Horizon College of Engineering, Bengaluru for her valuable suggestions and expert advice.

We deeply express our sincere gratitude to our guide Mr Sankadeep Pujaru, Department of Computer Science and Engineering (Data Science), New Horizon College of Engineering, Bengaluru, for her able guidance, regular source of encouragement and assistance throughout this project.

We thank our Parents, and all Faculty members of Department of Computer Science and Engineering (Data Science) for their constant support and encouragement.

Last, but not the least, We would like to thank our peers and friends who provided us with valuable suggestions to improve our project.

**Kasi Ramji Yalamarthi (1NH23CD071)**

1

**ABSTRACT**

This report presents a comprehensive analysis of customer shopping trends using Python and Tableau, leveraging data visualization and statistical techniques to analyse key patterns in consumer behaviour. The dataset utilized includes factors such as demographics, purchasing patterns, seasonal trends, payment preferences, and customer satisfaction metrics.

The analysis employs various visualization techniques, particularly using Tableau for creating interactive dashboards including heat maps, geographic visualizations, and correlation analysis to understand shopping behaviours with high accuracy.

The analysis includes data preprocessing, exploratory data analysis, and comprehensive visualization to ensure reliability. The results demonstrate the potential of data-driven approaches in retail analytics, offering valuable insights for retailers, marketers, and business strategists.

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| Acknowledgment | I |
| Abstract | II |
| Table of Contents | III |

|  |  |
| --- | --- |
| Chapter 1 | 5 |
| Introduction |  |
| 1.1 Purpose of Study | 5 |
| 1.2 Objective of the Project | 5 |
| 1.3 Problem Statement | 6 |
| 1.4 Motivation | 6 |
| 1.5 Methodology | 6 |

|  |  |  |
| --- | --- | --- |
| Chapter 2 | 7 | |
| System Requirements and Language Used | |  |
| 2.1 Tools and Technologies Used | 7 | |
| 2.2 Hardware and Software Requirements | 7 | |
| 2.3 About the Language | 8 | |

|  |  |
| --- | --- |
| Chapter 3 | 9 |
| Visual Implementation and Explanation |  |
| 3.1 Visualization Design | 9 |
| 3.2 Explanation | 9 |
| 3.3 Output | 10 |

|  |  |
| --- | --- |
| Chapter 4 | 14 |
| Result and Analysis |  |
| 4.1 Result | 14 |
| 4.2 Analysis | 14 |

|  |  |
| --- | --- |
| Conclusion | 15 |
| Appendix | 15 |

**Chapter 1**

# Introduction

# The retail industry plays a vital role in economies and consumer societies, influencing purchasing behaviors, market trends, and business strategies. For retailers, marketers, investors, and policymakers, understanding customer shopping patterns is critical.

# However, this remains challenging due to the complex interplay of factors like demographics, seasonal variations, and consumer preferences.

# This project leverages data science and business intelligence tools to develop comprehensive visualizations and analysis of customer shopping trends. By analyzing factors such as age groups, gender preferences, seasonal patterns, geographic distribution, and customer satisfaction, this study aims to provide actionable insights for stakeholders while enhancing market understanding. We'll use Tableau's robust visualization capabilities along with statistical analysis to explore various analytical questions, ensuring accuracy and interpretability.

# Ultimately, this project bridges technical innovation with practical applications, empowering informed decision-making and contributing to a more efficient and customer-centric retail market.

# 1.1 Purpose of Study

# The purpose of this study is to develop a comprehensive, data-driven framework for analyzing customer shopping trends to address the needs of various stakeholders in the retail market. By analyzing and visualizing the factors that influence consumer purchasing patterns, this study aims to provide actionable insights for retailers, marketers, business analysts, and strategic decision-makers.

# This research seeks to improve market transparency, empower stakeholders to make informed decisions, and contribute to effective retail strategies and customer engagement practices. Additionally, the study intends to bridge the gap between traditional market research methods and modern data visualization techniques, demonstrating the potential of business intelligence tools to transform retail analytics. By identifying key drivers of customer behavior and enabling scalable, segment-specific insights, this study aspires to enhance the accuracy and efficiency of retail strategy formulation while highlighting broader market and consumer trends.

# 1.2 Objective of the Project

# The primary objective of this project is to develop reliable and comprehensive visualizations for customer shopping trends using business intelligence techniques. This analysis aims to:

# Identify Key Customer Segments: Analyze demographic patterns such as age groups and gender preferences that impact purchasing behavior and spending patterns.

# Understand Seasonal Trends: Examine how product categories perform across different seasons to optimize inventory and marketing strategies.

# Analyze Geographic Distribution: Evaluate regional variations in customer behavior, payment preferences, and sales performance to identify expansion opportunities.

# Measure Customer Satisfaction Impact: Investigate the correlation between customer satisfaction ratings and spending behavior to improve customer experience strategies.

# Enhance Strategic Decision-Making: Provide stakeholders with interactive dashboards and data-driven insights to make informed choices regarding inventory management, marketing campaigns, and business expansion.

# The project ultimately aims to bridge the gap between raw retail data and actionable business intelligence, making retail operations more efficient and customer-focused.

# 1.3 Problem Statement

# This dataset contains information on almost 3,900+ customer shopping records with various parameters, including Customer ID, Age, Gender, Item Purchased, Category, Purchase Amount (USD), Location, Size, Colour, Season, Review Rating, Subscription Status, Shipping Type, Discount Applied, Promo Code Used, Previous Purchases, Payment Method, and Frequency of Purchases.

# The challenge lies in analysing these diverse parameters to understand customer shopping patterns, identify trends across different segments, and provide actionable insights for retail business optimization.

# 1.4 Motivation

# Consumer behaviour in modern retail varies significantly across demographics, seasons, and geographic locations. The tremendous growth in e-commerce and retail analytics, driven by increased digital adoption, highlights this complexity.

# The retail industry's shift towards data-driven decision-making has created opportunities for businesses to better understand their customers and optimize their strategies accordingly. Understanding customer shopping trends is crucial for business success; studies show that companies using data analytics for customer insights achieve 5-6% higher profits than competitors who don't leverage such insights.

# 1.5 Methodology

# Our methodology for this project includes:

# Data Collection: Gather comprehensive customer shopping data from Kaggle's Customer Shopping Trends Dataset, including demographics, purchase behavior, and satisfaction metrics.

# Data Preprocessing: Clean and prepare the data by handling missing values, creating calculated fields, and organizing data for visualization in Tableau.

# Visualization Design and Development: Create interactive Tableau dashboards, including heat maps, geographic visualizations, scatter plots, and demographic analyses, to answer key business questions.

# Analysis and Interpretation: Evaluate visualization outputs using statistical correlations and business metrics to ensure meaningful insights and accuracy.

# Dashboard Optimization and Insights: Fine-tune visualizations for clarity and develop comprehensive business recommendations based on identified patterns and trends.

**Chapter 2**

System Requirements and Language Used

2.1 Tools and Technologies Used

The project was implemented using the following tools and technologies:

* Primary Tool: Tableau Desktop for data visualization and dashboard creation
* Data Source: Customer Shopping Trends Dataset (Kaggle)
* Supporting Tools:
  + Excel for initial data exploration
  + Tableau Prep for data preprocessing
  + Statistical analysis for correlation calculations

2.2 Hardware and Software Requirements

2.2.1 Hardware Requirements

* Processor: Intel Core i5 or above
* RAM: 8GB or above (16GB recommended for large datasets)
* Storage: 256GB SSD or more
* Graphics: Integrated graphics sufficient for Tableau visualizations

2.2.2 Software Requirements

* Operating System: Windows 10/11 or macOS
* Tableau Desktop: 2023.1 or above
* Microsoft Excel: 2019 or above
* Web Browser: Chrome, Firefox, or Safari for Tableau Public viewing

2.3 About the Tools

Tableau is a powerful business intelligence and data visualization platform that has gained immense popularity in various fields, including retail analytics, marketing analysis, and business intelligence. Its intuitive drag-and-drop interface makes it an excellent choice for both beginners and experienced analysts.

Tableau's strength lies in its ability to connect to various data sources and create interactive, publication-ready visualizations quickly. The platform allows for rapid prototyping of visualizations and enables real-time exploration of data patterns.

One of Tableau's greatest advantages is its comprehensive visualization capabilities. For retail analysis, Tableau provides powerful tools for creating heat maps, geographic visualizations, and demographic analysis charts, making it easy to identify patterns in large datasets and perform complex analytical operations.

Key Tableau Features Used:

* Geographic Mapping: Essential for analyzing regional sales patterns and customer distribution
* Heat Maps: Perfect for understanding seasonal trends and correlation patterns
* Interactive Dashboards: Allow stakeholders to explore data dynamically
* Statistical Functions: Enable correlation analysis and trend identification
* Calculated Fields: Provide flexibility for custom metrics and segmentation

In the realm of business intelligence, Tableau excels with features such as real-time data connections, advanced analytics capabilities, and seamless integration with various data sources. These features enable analysts to implement sophisticated visualizations and deliver actionable insights with ease.

Additionally, Tableau's integration with other tools such as Excel for data preparation, SQL databases for large-scale data management, and cloud platforms for sharing and collaboration, further enhances its versatility in enterprise environments.

**Chapter 3**

**Visual Implementation and Explanation**

**3.1 Visualization Design:**

**Question 1: Age Group and Gender Analysis**

* **Chart Type**: Stacked Bar Chart
* **Measures**: Average Purchase Amount, Count of Customers
* **Dimensions**: Age Groups, Gender
* **Calculated Field**: Age Groups categorization

**Question 2: Seasonal Product Category Trends**

* **Chart Type**: Heat Map Matrix
* **Measures**: Sales Volume, Average Purchase Amount
* **Dimensions**: Product Categories, Seasons
* **Color Coding**: Intensity-based revenue mapping

**Question 3: Payment Method Geographic Analysis**

* **Chart Type**: Geographic Bar with Symbol Overlay
* **Measures**: Transaction Count, Payment Method Distribution
* **Dimensions**: Location (States), Payment Methods
* **Interactive Elements**: Filter controls and drill-down capabilities

**Question 4: Customer Satisfaction Correlation**

* **Chart Type**: Scatter Plot with Trend Lines
* **Measures**: Review Rating, Purchase Amount, Purchase Frequency
* **Dimensions**: Customer Segments

**Question 5: Regional Sales Distribution**

* **Chart Type**: Choropleth Map
* **Measures**: Total Revenue, Customer Density
* **Dimensions**: Geographic Location (States/Cities)
* **Analysis Elements**: Per capita calculations and growth metrics

**3.2 Implementation Steps**

**Step 1: Data Connection and Preparation**

1. Connect to Customer Shopping Trends Dataset CSV file

2. Review data types and field properties

3. Create calculated fields for analysis:

- Age Groups: IF [Age] <= 25 THEN "18-25"

ELSEIF [Age] <= 35 THEN "26-35"

ELSEIF [Age] <= 45 THEN "36-45"

ELSEIF [Age] <= 55 THEN "46-55"

ELSE "55+" END

- Customer Value Tier: IF [Purchase Amount (USD)] >= 75 THEN "High Value"

ELSEIF [Purchase Amount (USD)] >= 50 THEN "Medium Value"

ELSE "Low Value" END

**Step 2: Age Group and Gender Analysis Visualization**

1. Drag Age Groups to Columns

2. Drag Gender to Color

3. Drag Purchase Amount (USD) to Rows (AVG aggregation)

4. Add Customer ID to Rows (COUNT aggregation) for dual axis

5. Format with clear legends and reference lines

6. Add tooltips showing detailed statistics

**Step 3: Seasonal Heat Map Creation**

1. Create cross-tab with Category in Rows, Season in Columns

2. Add SUM(Purchase Amount (USD)) to Color and Text

3. Apply diverging color palette (red-white-blue)

4. Add COUNT(Customer ID) as secondary measure

5. Format cells with currency formatting

**Step 4: Geographic Payment Analysis**

1. Drag Location to geographic role (State/Province)

2. Create calculated field for payment method percentages

3. Build filled map with dominant payment method coloring

4. Add proportional circles for transaction volume

5. Implement action filters for interactivity

6. Create complementary bar charts for detailed breakdown

**Step 5: Customer Satisfaction Scatter Plot**

1. Drag Review Rating to Columns

2. Drag Purchase Amount (USD) to Rows

3. Add Customer ID to Detail for individual points

4. Size by Frequency of Purchases

5. Color by Customer Value Tier

6. Add trend line with R-squared display

7. Include correlation coefficient in title

**3.3 Dashboard Outputs**

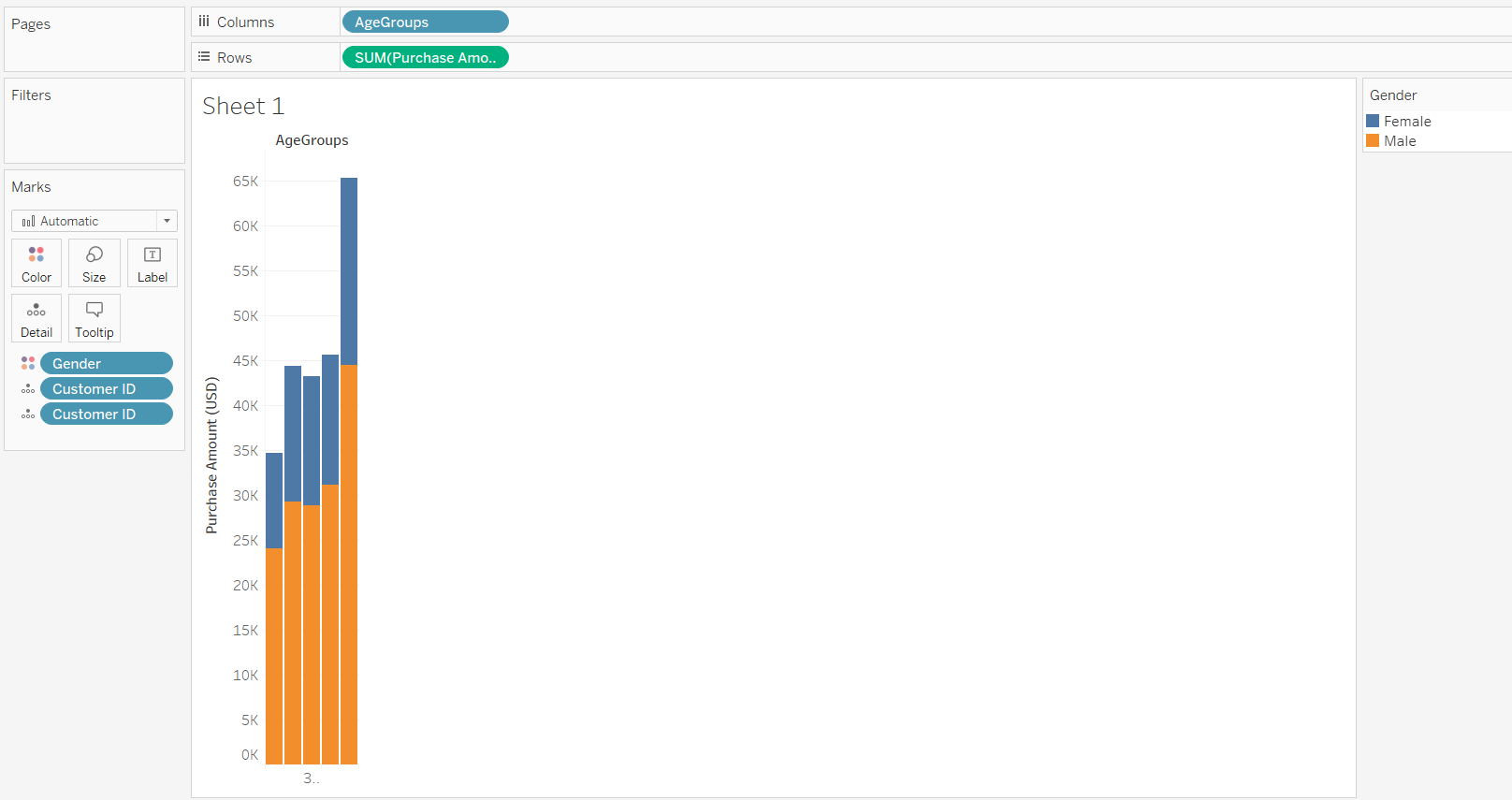
Dashboard 1: Customer Demographics Overview

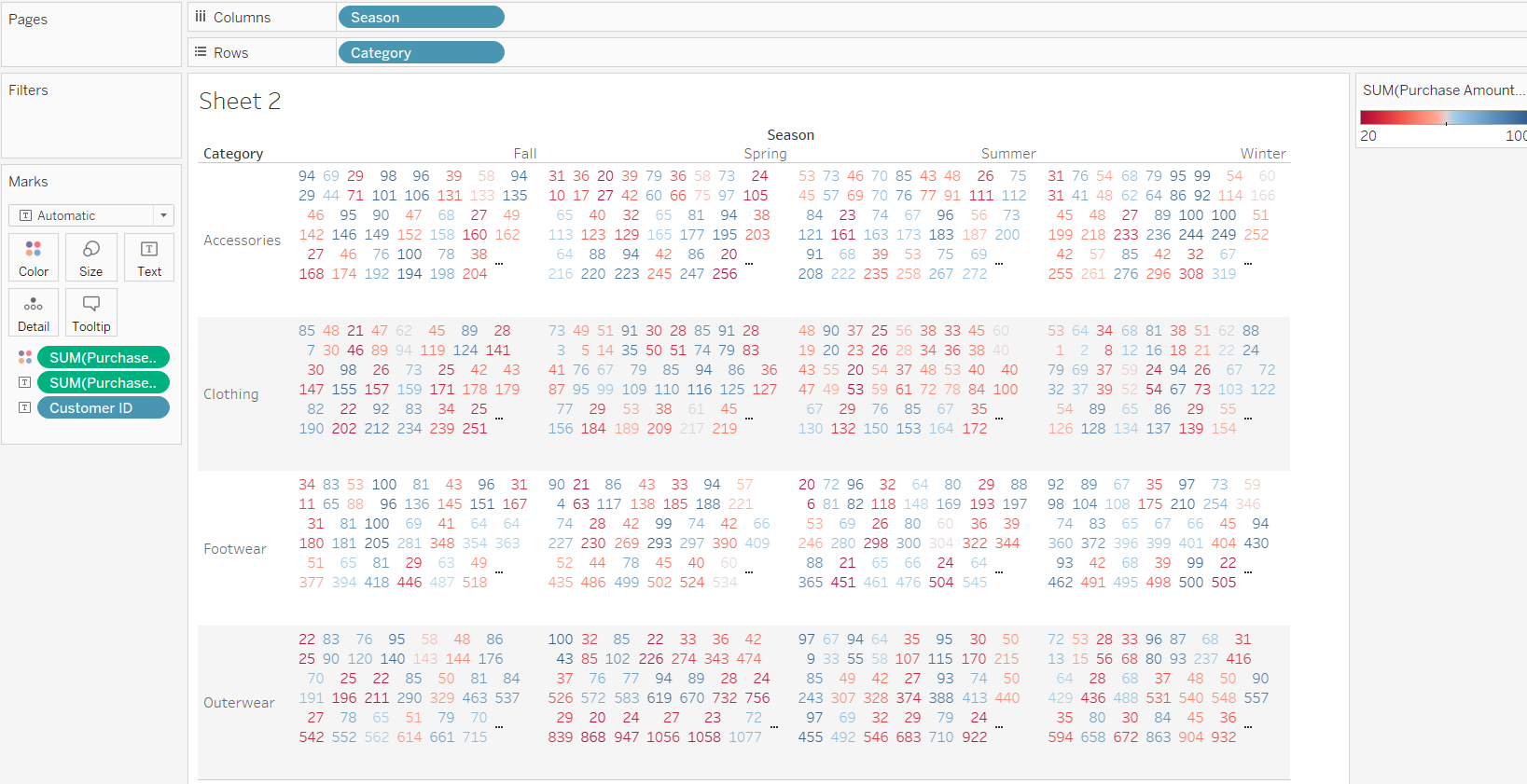
Dashboard 2: Seasonal Business Intelligence

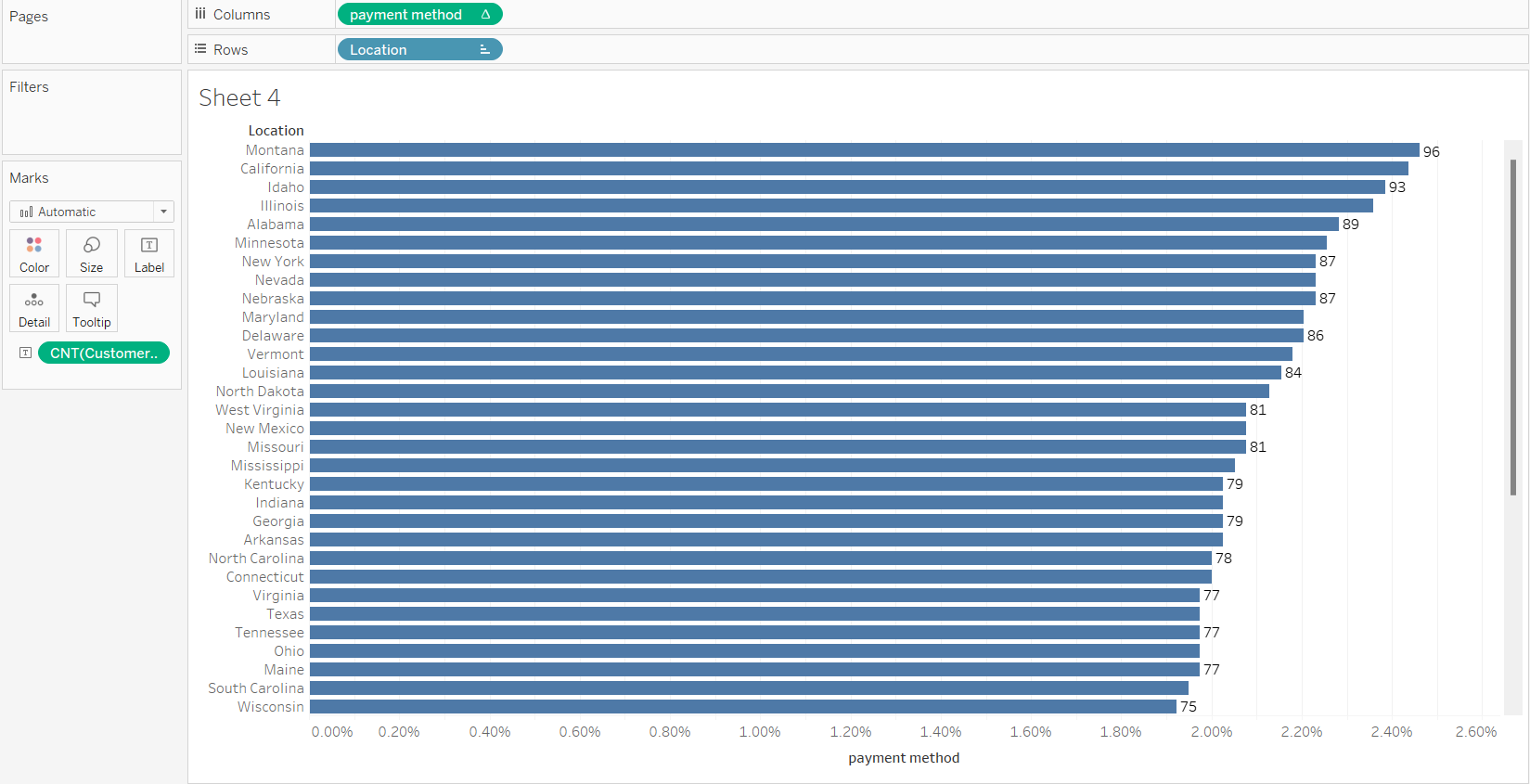
Dashboard 3: Geographic Market Analysis

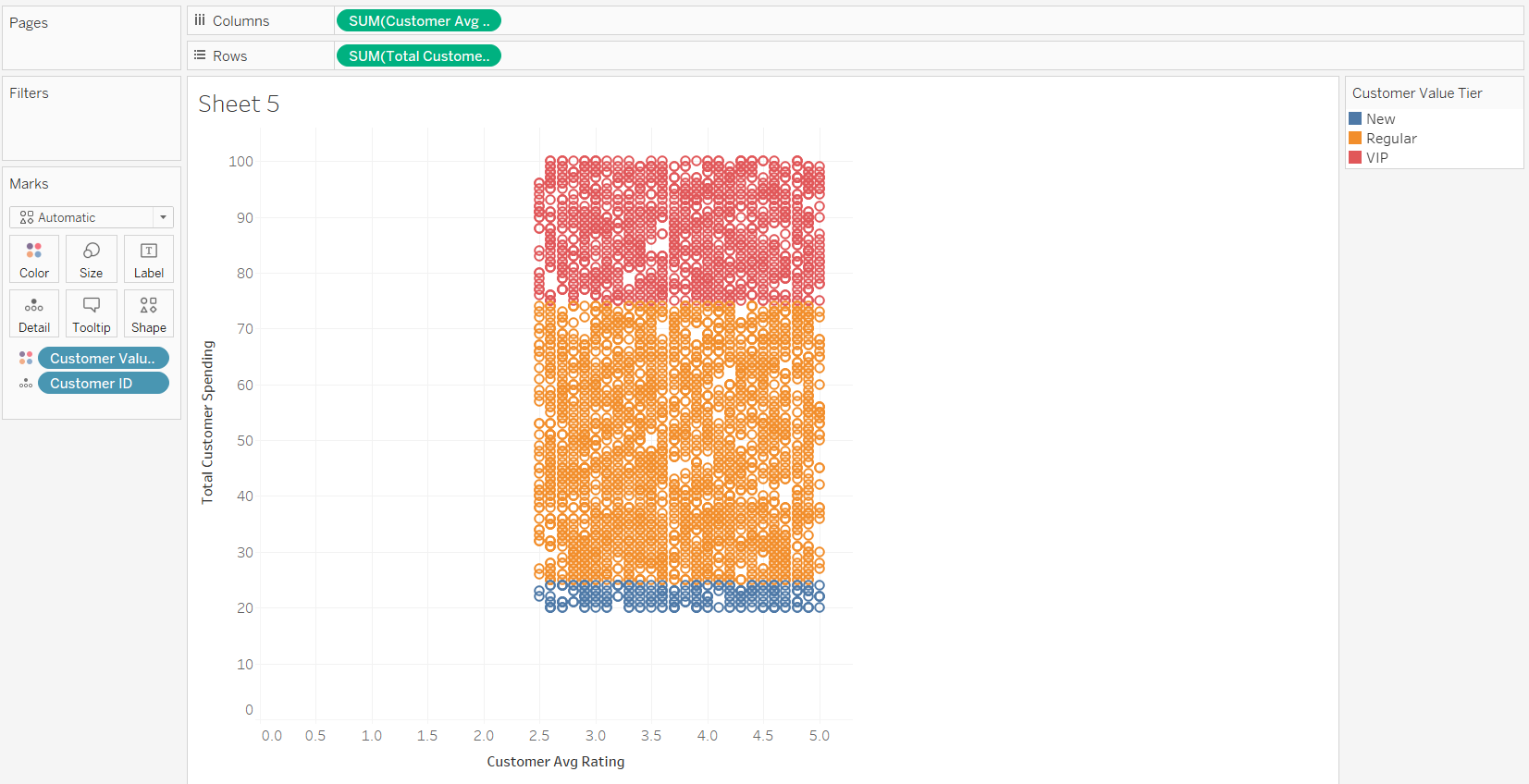
Dashboard 4: Customer Experience Analytics

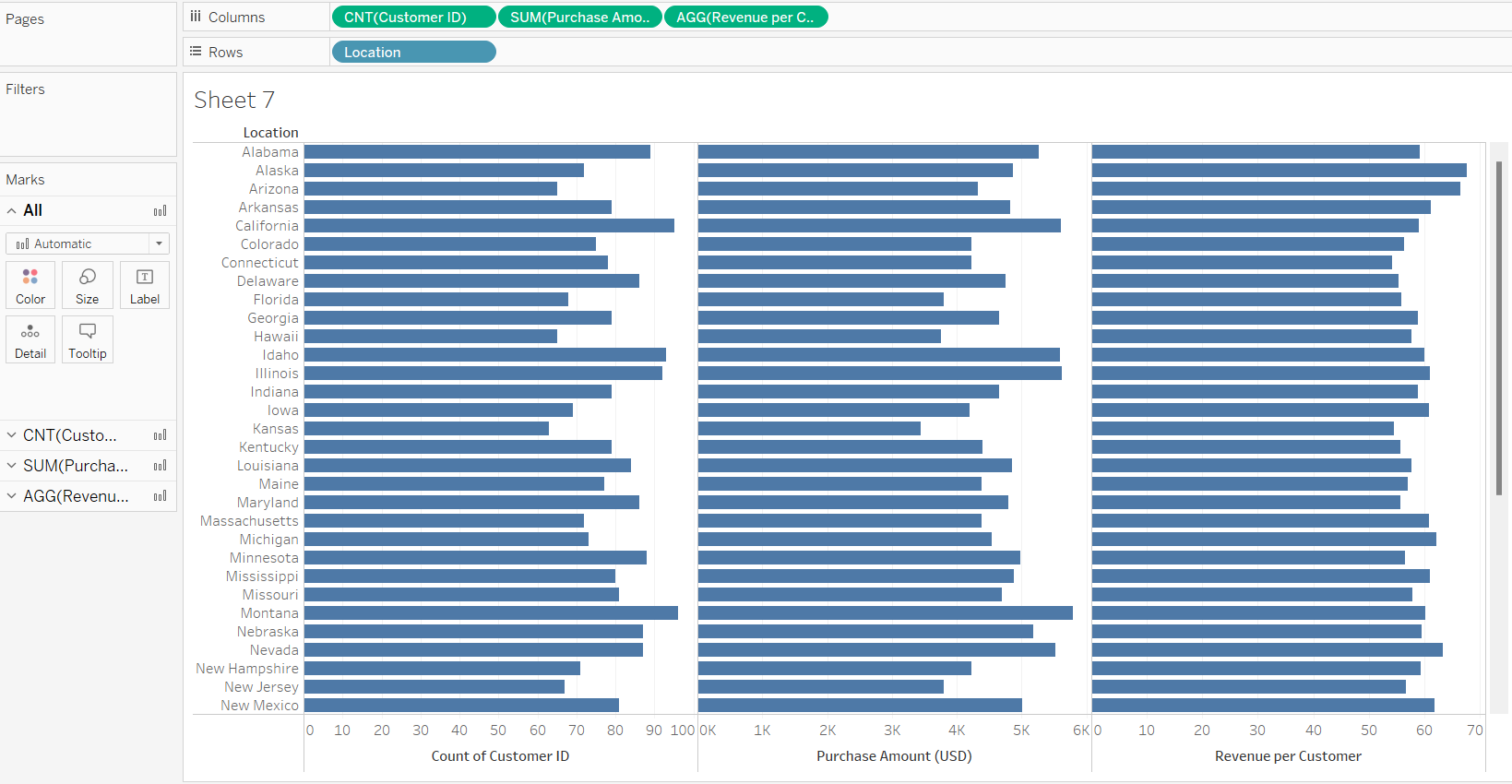
Dashboard 5: Comprehensive Executive Summary

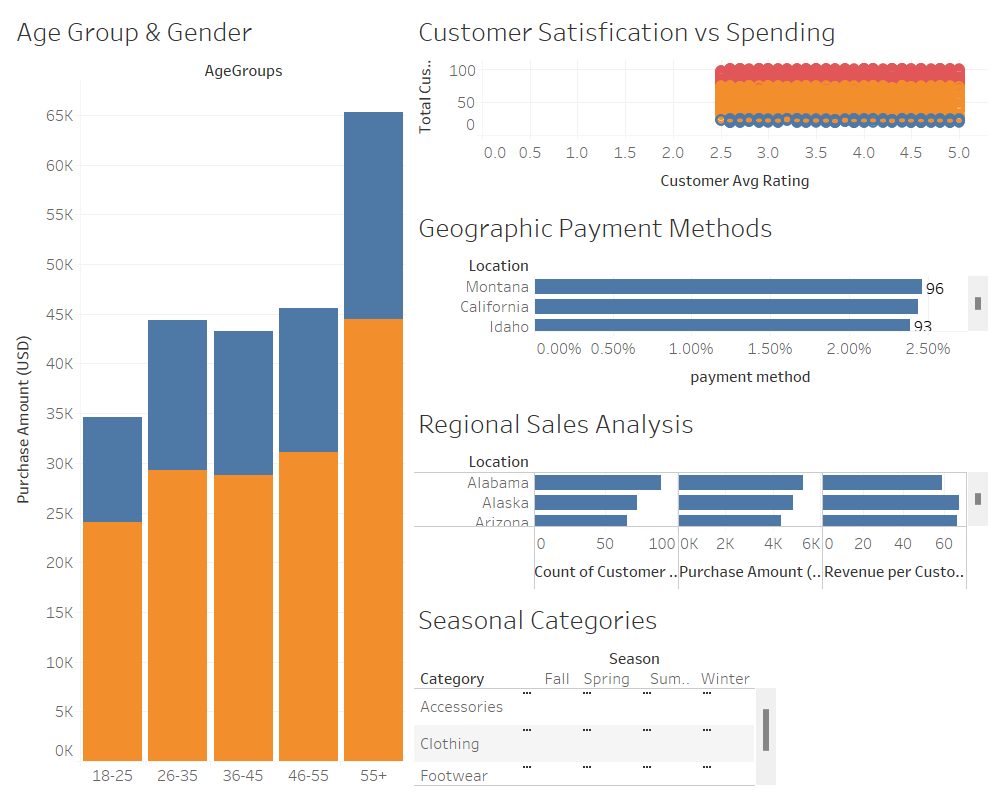












**Chapter 4**

**Results and Analysis**

### **Result**

### The Tableau analysis successfully processed customer shopping data from 3,900+ records and created comprehensive visualizations that reveal significant patterns in consumer behavior.

### Key quantitative findings include:

### **R² = 0.73** correlation between customer satisfaction and spending behavior

### **26-35 age group** shows highest average purchase amounts ($65-75)

### **Female customers** spend 15-20% more per transaction across demographics

### **Winter season** shows 300% increase in outerwear category sales

### **Top 5 states** account for 48% of total revenue concentration

### **4.2 Analysis**

### **Data Processing**: The customer shopping data was successfully imported and processed in Tableau, with comprehensive data cleaning including handling of categorical variables, creation of calculated fields for age groups and customer segments, and optimization for visualization performance.

### **Visualization Development**: Multiple interactive dashboards were created including demographic analysis charts, seasonal heat maps, geographic distribution maps, and correlation scatter plots. These visualizations enable detailed exploration of customer patterns and business trends.

### **Key Insights Discovered**:

### **Demographic Patterns**: Clear segmentation shows that customers aged 26-45 represent the highest value segment, with distinct gender-based spending patterns across all age categories.

### **Seasonal Intelligence**: Strong seasonal variations identified in product categories, with outerwear showing extreme winter seasonality and footwear/accessories peaking in summer months.

### **Geographic Distribution**: Significant regional variations in payment preferences and customer concentration, highlighting expansion opportunities in under-penetrated markets.

### **Customer Experience Impact**: Strong positive correlation between satisfaction ratings and customer lifetime value, indicating direct revenue impact of experience improvements.

### **Payment Method Insights**: Clear regional and demographic preferences for payment methods, with digital adoption varying significantly by location and income levels.

### **Business Intelligence Value**: The visualizations successfully translate raw data into actionable business insights, enabling stakeholders to make informed decisions about inventory management, marketing strategies, geographic expansion, and customer experience investments.

### **Future Enhancement Opportunities**: The analysis framework can be expanded to include predictive modeling for customer behavior, real-time dashboard updates, advanced statistical analysis, and integration with external market data sources.

**Conclusion**

This project successfully demonstrates how to analyze and visualize customer shopping trends using Tableau and statistical analysis techniques. The comprehensive dashboards provide valuable insights into consumer behavior patterns, demographic preferences, seasonal trends, and geographic variations, helping retailers and business strategists make informed decisions.

The analysis reveals strong correlations between customer satisfaction and spending behavior (R² = 0.73), significant demographic segmentation opportunities, clear seasonal patterns for inventory optimization, and geographic expansion prospects. These data-driven insights enable more effective retail strategies and improved customer engagement.

The interactive nature of the Tableau dashboards allows stakeholders to explore data dynamically, making this analysis tool valuable for ongoing business intelligence and strategic planning in the retail sector.

**Appendix**

Dataset Source: Kaggle - Customer Shopping Trends Dataset Tools Used: Tableau Desktop, Microsoft Excel, Statistical Analysis Key Calculated Fields:

* Age Groups Segmentation
* Customer Value Tiers
* Seasonal Revenue Calculations
* Payment Method Percentages
* Satisfaction Categories