SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 1

Customer Feedback Analysis

|  |  |  |
| --- | --- | --- |
| **Customer ID** | **Age** | **Satisfaction Score** |
| 1 | 25 | 4 |
| 2 | 30 | 5 |
| 3 | 35 | 3 |
| 4 | 28 | 4 |
| 5 | 40 | 5 |

1. Using R Create a histogram to represent the distribution of customer ages. Label the axes and title the chart.

customer\_data <- data.frame(

Customer\_ID = c(1, 2, 3, 4, 5),

Age = c(25, 30, 35, 28, 40),

Satisfaction\_Score = c(4, 5, 3, 4, 5)

)

hist(customer\_data$Age,

main = "Distribution of Customer Ages",

xlab = "Age",

ylab = "Frequency",

col = "skyblue",

border = "black")

1. Using R Generate a pie chart to display the overall distribution of customer satisfaction scores. Include labels.

# Assuming your data is in a data frame called 'customer\_data'

customer\_data <- data.frame(

Customer\_ID = c(1, 2, 3, 4, 5),

Age = c(25, 30, 35, 28, 40),

Satisfaction\_Score = c(4, 5, 3, 4, 5)

)

# Count the occurrences of each satisfaction score

satisfaction\_counts <- table(customer\_data$Satisfaction\_Score)

# Create a pie chart

pie(satisfaction\_counts,

labels = paste("Score", names(satisfaction\_counts)),

main = "Overall Distribution of Satisfaction Scores",

col = rainbow(length(satisfaction\_counts))

)

1. Using R Build a stacked bar chart to visualize the distribution of customer satisfaction scores by age group.

# Assuming your data is in a data frame called 'customer\_data'

customer\_data <- data.frame(

Customer\_ID = c(1, 2, 3, 4, 5),

Age = c(25, 30, 35, 28, 40),

Satisfaction\_Score = c(4, 5, 3, 4, 5)

)

# Create age groups (you can customize the breaks as needed)

age\_groups <- cut(customer\_data$Age, breaks = c(20, 30, 40, 50), labels = c("20-30", "30-40", "40-50"))

# Create a table with counts of satisfaction scores by age group

satisfaction\_by\_age <- table(age\_groups, customer\_data$Satisfaction\_Score)

# Create a stacked bar chart

barplot(satisfaction\_by\_age,

beside = TRUE,

legend.text = TRUE,

col = rainbow(length(unique(customer\_data$Satisfaction\_Score))),

args.legend = list(x = "topright", bty = "n"),

main = "Distribution of Satisfaction Scores by Age Group",

xlab = "Age Group",

ylab = "Count"

)

# Count the occurrences of each satisfaction score

1. In Tableau, develop a word cloud from open-ended customer feedback to identify prevalent customer sentiments.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 2

Product Inventory Management

|  |  |  |
| --- | --- | --- |
| **Product ID** | **Product Name** | **Quantity Available** |
| 1 | Product A | 250 |
| 2 | Product B | 175 |
| 3 | Product C | 300 |
| 4 | Product D | 200 |
| 5 | Product E | 220 |

1. Using R Create a bar chart to visualize the quantity of each product in the inventory. Label the axes and title the chart.

# Sample data

product\_data <- data.frame(

Product\_ID = c(1, 2, 3, 4, 5),

Product\_Name = c("Product A", "Product B", "Product C", "Product D", "Product E"),

Quantity\_Available = c(250, 175, 300, 200, 220)

)

# Bar chart

barplot(product\_data$Quantity\_Available, names.arg = product\_data$Product\_Name,

main = "Product Quantity in Inventory",

xlab = "Product Name",

ylab = "Quantity Available",

col = "skyblue",

border = "black",

)

1. Using R Generate a stacked bar chart to show the quantity of each product within different product categories.

# Sample data with category

product\_data <- data.frame(

Product\_ID = c(1, 2, 3, 4, 5),

Product\_Name = c("Product A", "Product B", "Product C", "Product D", "Product E"),

Quantity\_Available = c(250, 175, 300, 200, 220),

Category = c("Category1", "Category2", "Category1", "Category2", "Category3")

)

# Stacked bar chart

barplot(as.matrix(t(table(product\_data$Category, product\_data$Product\_Name))),

main = "Product Quantity by Category",

xlab = "Category",

ylab = "Quantity Available",

col = rainbow(nrow(product\_data)),

)

1. Using R Build a scatter plot to explore the relationship between product price and quantity available. Explain the findings.

# Assuming you have a dataset with product prices

product\_price\_data <- data.frame(

Product\_ID = c(1, 2, 3, 4, 5),

Product\_Name = c("Product A", "Product B", "Product C", "Product D", "Product E"),

Quantity\_Available = c(250, 175, 300, 200, 220),

Price = c(10, 15, 8, 20, 12)

)

# Scatter plot

plot(product\_price\_data$Price, product\_price\_data$Quantity\_Available,

main = "Scatter Plot: Price vs. Quantity",

xlab = "Product Price",

ylab = "Quantity Available",

pch = 16, col = "blue")

1. Develop a Tableau dashboard with the bar chart and stacked bar chart to allow users to interact with the data.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 3

Product Sales Analysis

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Product ID** | **Product Name** | **January Sales** | **February Sales** | **March Sales** |
| 1 | Product A | 2000 | 2200 | 2400 |
| 2 | Product B | 1500 | 1800 | 1600 |
| 3 | Product C | 1200 | 1400 | 1100 |

1. Using R Create a grouped bar chart to visualize the sales of each product for the first quarter. Label the chart elements.

# Sample data

product\_sales\_data <- data.frame(

Product\_ID = c(1, 2, 3),

Product\_Name = c("Product A", "Product B", "Product C"),

January\_Sales = c(2000, 1500, 1200),

February\_Sales = c(2200, 1800, 1400),

March\_Sales = c(2400, 1600, 1100)

)

# Grouped bar chart

barplot(as.matrix(t(product\_sales\_data[, 3:5])),

beside = TRUE,

col = c("skyblue", "orange", "lightgreen"),

names.arg = product\_sales\_data$Product\_Name,

main = "Monthly Sales of Each Product",

xlab = "Products",

ylab = "Sales",

)

1. Using R Generate a stacked area chart to represent the overall sales trend for all products over the three months.

# Sample data

product\_sales\_data <- data.frame(

Product\_ID = c(1, 2, 3),

Product\_Name = c("Product A", "Product B", "Product C"),

January\_Sales = c(2000, 1500, 1200),

February\_Sales = c(2200, 1800, 1400),

March\_Sales = c(2400, 1600, 1100)

)

# Overall sales data

overall\_sales <- cbind(product\_sales\_data$January\_Sales,

product\_sales\_data$February\_Sales,

product\_sales\_data$March\_Sales)

# Stacked area chart

matplot(1:3, t(overall\_sales),

type = "n",

main = "Overall Sales Trend",

xlab = "Month",

ylab = "Sales",

col = c("skyblue", "orange", "lightgreen"),

ylim = c(0, sum(overall\_sales)),

xaxt = "n")

for (i in 1:nrow(overall\_sales)) {

polygon(c(1:3, rev(1:3)), c(overall\_sales[i,], rev(rep(0, 3))),

col = alpha(c("skyblue", "orange", "lightgreen"), 0.5)[i])

}

legend("topright", legend = product\_sales\_data$Product\_Name, fill = c("skyblue", "orange", "lightgreen"))

axis(1, at = 1:3, labels = c("January", "February", "March"))

1. Using R Build a table to show the monthly sales data for each product. Label the table elements.

# Sample data

product\_sales\_data <- data.frame(

Product\_ID = c(1, 2, 3),

Product\_Name = c("Product A", "Product B", "Product C"),

January\_Sales = c(2000, 1500, 1200),

February\_Sales = c(2200, 1800, 1400),

March\_Sales = c(2400, 1600, 1100)

)

sales\_table <- as.data.frame(product\_sales\_data[, c("Product\_Name", "January\_Sales", "February\_Sales", "March\_Sales")])

colnames(sales\_table) <- c("Product Name", "January Sales", "February Sales", "March Sales")

sales\_table

1. Develop a Tableau dashboard combining the grouped bar chart, stacked area chart, and the table for interactive exploration of sales data.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 4

Customer Demographics Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| **Customer ID** | **Age** | **Gender** | **Income (in $)** |
| 1 | 28 | Female | 50000 |
| 2 | 35 | Male | 60000 |
| 3 | 42 | Female | 75000 |

1. Using R Create a bar chart to visualize the distribution of customer ages. Label the axes and title the chart.

# Sample data

customer\_data <- data.frame(

Customer\_ID = c(1, 2, 3),

Age = c(28, 35, 42),

Gender = c("Female", "Male", "Female"),

Income = c(50000, 60000, 75000)

)

# Bar chart for age distribution

barplot(customer\_data$Age,

names.arg = customer\_data$Customer\_ID,

main = "Customer Age Distribution",

xlab = "Customer ID",

ylab = "Age",

col = "skyblue",

border = "black")

1. Using R Generate a pie chart to represent the distribution of customers by gender.

# Pie chart for gender distribution

gender\_counts <- table(customer\_data$Gender)

pie(gender\_counts,

main = "Distribution of Customers by Gender",

labels = paste(names(gender\_counts), ": ", gender\_counts),

col = c("skyblue", "pink"))

1. Using R Build a table to show the demographic information for each customer. Label the table elements.

# Sample data

customer\_data <- data.frame(

Customer\_ID = c(1, 2, 3),

Age = c(28, 35, 42),

Gender = c("Female", "Male", "Female"),

Income = c(50000, 60000, 75000)

)

# Displaying demographic information in a table

demographic\_table <- customer\_data[, c("Customer\_ID", "Age", "Gender", "Income")]

colnames(demographic\_table) <- c("Customer ID", "Age", "Gender", "Income (in $)")

demographic\_table

1. Develop a Tableau dashboard combining the bar chart, pie chart, and the table for interactive exploration of customer demographics.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

**Set 5**

Employee Performance Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| **Employee ID** | **Department** | **Years of Service** | **Performance Score** |
| 1 | Sales | 5 | 85 |
| 2 | HR | 3 | 92 |
| 3 | Marketing | 7 | 78 |

1. Using R Create a line chart to visualize the performance trend of employees over time. Label the axes and title the chart.

# Sample data

employee\_data <- data.frame(

Employee\_ID = c(1, 2, 3),

Department = c("Sales", "HR", "Marketing"),

Years\_of\_Service = c(5, 3, 7),

Performance\_Score = c(85, 92, 78)

)

# Line chart for performance trend over time

plot(employee\_data$Years\_of\_Service, employee\_data$Performance\_Score,

type = "o",

main = "Performance Trend Over Time",

xlab = "Years of Service",

ylab = "Performance Score",

col = "blue",

pch = 16)

1. Using R Generate a bar chart showing the distribution of employees across different departments. Label the chart elements.

# Bar chart for department distribution

barplot(table(employee\_data$Department),

main = "Distribution of Employees Across Departments",

xlab = "Department",

ylab = "Number of Employees",

col = "skyblue",

border = "black")

1. Using R Build a table to display the performance data for each employee. Label the table elements.

# Sample data

employee\_data <- data.frame(

Employee\_ID = c(1, 2, 3),

Department = c("Sales", "HR", "Marketing"),

Years\_of\_Service = c(5, 3, 7),

Performance\_Score = c(85, 92, 78)

)

# Displaying performance data in a table

performance\_table <- employee\_data[, c("Employee\_ID", "Department", "Years\_of\_Service", "Performance\_Score")]

colnames(performance\_table) <- c("Employee ID", "Department", "Years of Service", "Performance Score")

performance\_table

1. Develop a Tableau dashboard combining the line chart, bar chart, and the table for interactive exploration of employee performance data.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 6

Survey Responses Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| **Survey ID** | **Question 1** | **Question 2** | **Question 3** |
| 1 | A | B | C |
| 2 | B | A | D |
| 3 | C | A | B |

1. Using R Create a grouped bar chart to visualize the distribution of answers for Question 1. Label the chart elements.

# Sample data

survey\_data <- data.frame(

Survey\_ID = c(1, 2, 3),

Question\_1 = c("A", "B", "C"),

Question\_2 = c("B", "A", "A"),

Question\_3 = c("C", "D", "B")

)

# Grouped bar chart for Question 1

barplot(table(survey\_data$Question\_1),

main = "Distribution of Answers for Question 1",

xlab = "Answer",

ylab = "Count",

col = "skyblue",

border = "black",

legend.text = TRUE)

1. Using R Generate a stacked bar chart to represent the overall distribution of responses for all three questions.

# Stacked bar chart for overall distribution of responses

barplot(as.matrix(t(table(c(survey\_data$Question\_1, survey\_data$Question\_2, survey\_data$Question\_3)))),

beside = TRUE,

legend.text = TRUE,

col = rainbow(3),

main = "Overall Distribution of Responses",

xlab = "Question",

ylab = "Count",

args.legend = list(title = "Answer"))

1. Using R Build a table to show the survey response data for each survey. Label the table elements.

# Sample data

survey\_data <- data.frame(

Survey\_ID = c(1, 2, 3),

Question\_1 = c("A", "B", "C"),

Question\_2 = c("B", "A", "A"),

Question\_3 = c("C", "D", "B")

)

# Displaying survey response data in a table

response\_table <- survey\_data[, c("Survey\_ID", "Question\_1", "Question\_2", "Question\_3")]

colnames(response\_table) <- c("Survey ID", "Question 1", "Question 2", "Question 3")

response\_table

1. Develop a Tableau dashboard combining the grouped bar chart, stacked bar chart, and the table for interactive exploration of survey responses.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

**Set 7**

Product Category Analysis

|  |  |
| --- | --- |
| **Category** | **Sales (in $)** |
| Electronics | 50000 |
| Clothing | 35000 |
| Appliances | 40000 |

1. Using R Create a pie chart to represent the distribution of sales across product categories. Include labels.

library(ggplot2)

# Sales data

categories <- c("Electronics", "Clothing", "Appliances")

sales <- c(50000, 35000, 40000)

sales\_data <- data.frame(Category = categories, Sales = sales)

# Create pie chart

pie\_chart <- ggplot(sales\_data, aes(x = "", y = Sales, fill = Category)) +

geom\_bar(stat = "identity", width = 1) +

coord\_polar("y", start = 0) +

labs(title = "Sales Distribution Across Product Categories", fill = "Category") +

theme\_minimal()

# Display the pie chart

print(pie\_chart)

1. Using R Generate a bar chart to analyse the sales conversion process for each product category. Label the stages and title the chart.

# Sales conversion data

conversion\_data <- data.frame(

Category = rep(categories, each = 4),

Stage = rep(c("Prospects", "Leads", "Opportunities", "Customers"), times = 3),

Sales = c(80000, 70000, 60000, 50000, 50000, 45000, 40000, 35000, 60000, 55000, 50000, 45000)

)

# Create bar chart

bar\_chart <- ggplot(conversion\_data, aes(x = Category, y = Sales, fill = Stage)) +

geom\_bar(stat = "identity", position = "stack") +

labs(title = "Sales Conversion Process by Product Category", x = "Category", y = "Sales", fill = "Stage") +

theme\_minimal() +

theme(axis.text.x = element\_text(angle = 45, hjust = 1))

# Display the bar chart

print(bar\_chart)

1. Using R Build a table to display the sales data for each product category. Label the table elements.

# Sales conversion data

conversion\_data <- data.frame(

Category = rep(categories, each = 4),

Stage = rep(c("Prospects", "Leads", "Opportunities", "Customers"), times = 3),

Sales = c(80000, 70000, 60000, 50000, 50000, 45000, 40000, 35000, 60000, 55000, 50000, 45000)

)

# Create a table

sales\_table <- data.frame(Category = categories, Sales = sales)

# Display the table

print(sales\_table)

1. Develop a Tableau dashboard combining the pie chart, funnel chart, and the table for interactive exploration of product category data.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 8

Customer Satisfaction

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Employee ID** | **Name** | **Department** | **Years of Service** | **Performance Score** |
| 1 | John Smith | Sales | 5 | 85 |
| 2 | Jane Doe | HR | 3 | 92 |
| 3 | Robert Brown | Marketing | 7 | 78 |
| 4 | Sarah White | Sales | 4 | 90 |
| 5 | Michael Lee | HR | 2 | 76 |

1. In R, create a histogram to visualize the distribution of customer ages. Label the axes and title the chart.
2. In R, generate a pie chart to represent the distribution of overall customer satisfaction scores. Include labels.
3. In Tableau, build a stacked bar chart to visualize the distribution of customer satisfaction scores by age group.
4. In R, perform sentiment analysis on open-ended customer feedback and create a word cloud visualization.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 9

Customer Demographics Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| **Customer ID** | **Age** | **Gender** | **Income (in $)** |
| 1 | 28 | Female | 50000 |
| 2 | 35 | Male | 60000 |
| 3 | 42 | Female | 75000 |

1. Using R Create a bar chart to visualize the distribution of customer ages. Label the axes and title the chart.
2. Using R Generate a pie chart to represent the distribution of customers by gender.
3. Using R Build a table to show the demographic information for each customer. Label the table elements.
4. Develop a Tableau dashboard combining the bar chart, pie chart, and the table for interactive exploration of customer demographics.

SAVEETHA SCHOOL OF ENGINEERING

SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES

Academic Year: 2023-2024 Branch: AI & DS

Subject Code: DSA 0603 Register No:

Subject Name: Data Handling and Visualization

Set 10

Sales Data

|  |  |  |
| --- | --- | --- |
| **Customer ID** | **Age** | **Satisfaction Score** |
| 1 | 28 | 4 |
| 2 | 35 | 5 |
| 3 | 42 | 3 |
| 4 | 30 | 4 |
| 5 | 45 | 5 |

1. In R, create a histogram to visualize the distribution of customer ages. Label the axes and title the chart.
2. In R, generate a pie chart to represent the distribution of overall customer satisfaction scores. Include labels.
3. In Tableau, build a stacked bar chart to visualize the distribution of customer satisfaction scores by age group.
4. In R, perform sentiment analysis on open-ended customer feedback and create a word cloud visualization.