# 1. Load the dataset and clean the data

library(readr)

data <- read.csv("data(1).csv")

# Cleaning the data (handling missing values by removing rows with NA)

data <- na.omit(data)

# 2. Create a contingency table for "engine-location" and "num-of-doors"

contingency\_table <- table(data$engine.location, data$num.of.doors)

# 3. Categorical vs. Categorical Analysis

# 3.1 Stacked bar chart using base R

barplot(contingency\_table, beside = FALSE, col = c("lightblue", "pink"),

xlab = "Number of Doors", ylab = "Count", main = "Engine Location vs Number of Doors")

# 3.2 Stacked bar chart using ggplot2

library(ggplot2)

ggplot(data, aes(x = num.of.doors, fill = engine.location)) +

geom\_bar(position = "stack") +

labs(title = "Stacked Bar Chart of Engine Location vs Number of Doors",

x = "Number of Doors", y = "Count")

# 3.3 Grouped bar plot (side-by-side bar plot)

ggplot(data, aes(x = num.of.doors, fill = engine.location)) +

geom\_bar(position = "dodge") +

labs(title = "Grouped Bar Chart of Engine Location vs Number of Doors",

x = "Number of Doors", y = "Count")

# 4. Quantitative vs. Quantitative Analysis

# 4.1 Scatter plot

ggplot(data, aes(x = engine.size, y = horsepower)) +

geom\_point() +

labs(title = "Scatter Plot of Engine Size vs Horsepower", x = "Engine Size", y = "Horsepower")

# 4.2 Line plot

ggplot(data, aes(x = engine.size, y = horsepower)) +

geom\_line() +

labs(title = "Line Plot of Engine Size vs Horsepower", x = "Engine Size", y = "Horsepower")

# 4.3 Heatmap

heatmap\_data <- table(cut(data$engine.size, breaks = 10), cut(data$horsepower, breaks = 10))

heatmap(as.matrix(heatmap\_data), Rowv = NA, Colv = NA, col = heat.colors(256), scale = "column")

# 5. Categorical vs. Quantitative Analysis

# 5.1 Bar chart

ggplot(data, aes(x = num.of.doors, y = horsepower)) +

geom\_bar(stat = "identity") +

labs(title = "Bar Chart of Horsepower by Number of Doors", x = "Number of Doors", y = "Horsepower")

# 5.2 Density plot

ggplot(data, aes(x = horsepower, fill = num.of.doors)) +

geom\_density(alpha = 0.5) +

labs(title = "Density Plot of Horsepower by Number of Doors", x = "Horsepower", y = "Density")

# 5.3 Box plot

ggplot(data, aes(x = num.of.doors, y = horsepower, fill = num.of.doors)) +

geom\_boxplot() +

labs(title = "Box Plot of Horsepower by Number of Doors", x = "Number of Doors", y = "Horsepower")

# 5.4 Violin plot

ggplot(data, aes(x = num.of.doors, y = horsepower, fill = num.of.doors)) +

geom\_violin() +

labs(title = "Violin Plot of Horsepower by Number of Doors", x = "Number of Doors", y = "Horsepower")

# 5.5 Violin and Box plot

ggplot(data, aes(x = num.of.doors, y = horsepower, fill = num.of.doors)) +

geom\_violin() +

geom\_boxplot(width = 0.2) +

labs(title = "Combined Violin and Box Plot of Horsepower by Number of Doors", x = "Number of Doors", y = "Horsepower")

### Multivariate Analysis

# 1. Scatter plot with color as a third variable

ggplot(data, aes(x = engine.size, y = horsepower, color = fuel.type)) +

geom\_point() +

labs(title = "Scatter Plot of Engine Size vs Horsepower by Fuel Type", x = "Engine Size", y = "Horsepower")

# 2. Scatter plot with color and shape as third and fourth variables

ggplot(data, aes(x = engine.size, y = horsepower, color = fuel.type, shape = aspiration)) +

geom\_point() +

labs(title = "Scatter Plot with Fuel Type and Aspiration", x = "Engine Size", y = "Horsepower")

# 3. Scatter plot with color and size as third and fourth variables

ggplot(data, aes(x = engine.size, y = horsepower, color = fuel.type, size = wheel.base)) +

geom\_point() +

labs(title = "Scatter Plot with Fuel Type and Wheel Base", x = "Engine Size", y = "Horsepower")

# 4. Bubble plot

ggplot(data, aes(x = engine.size, y = horsepower, color = fuel.type, size = wheel.base)) +

geom\_point(alpha = 0.6) +

labs(title = "Bubble Plot of Engine Size vs Horsepower by Fuel Type", x = "Engine Size", y = "Horsepower")

# 5. Faceted histogram (facet\_warp)

ggplot(data, aes(x = horsepower)) +

geom\_histogram(binwidth = 10) +

facet\_wrap(~ fuel.type) +

labs(title = "Faceted Histogram of Horsepower by Fuel Type", x = "Horsepower", y = "Count")

# 6. Faceted histogram (facet\_grid)

ggplot(data, aes(x = horsepower)) +

geom\_histogram(binwidth = 10) +

facet\_grid(fuel.type ~ aspiration) +

labs(title = "Faceted Histogram by Fuel Type and Aspiration", x = "Horsepower", y = "Count")