install.packages("readr")

library(readr)

csv\_file\_path<- "StudentsPerformance.csv"

my\_data<- read\_csv(csv\_file\_path)

str(my\_data)

library(ggplot2)

ggplot(my\_data, aes(x = `math score`)) +

geom\_histogram(binwidth = 5, fill = "skyblue", color = "black") +

labs(title = "Distribution of Math Scores", x = "Math Score", y = "Frequency")

ggplot(my\_data, aes(x = `race/ethnicity`, y = `reading score`, fill = `race/ethnicity`)) +

geom\_boxplot() +

labs(title = "Reading Scores by Race/Ethnicity", x = "Race/Ethnicity", y = "Reading Score") +

theme(legend.title = element\_blank())

ggplot(my\_data, aes(x = `parental level of education`)) +

geom\_bar(fill = "coral", color = "black") +

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

labs(title = "Parental Level of Education Distribution", x = "Parental Level of Education", y = "Count")

ggplot(my\_data, aes(x = `parental level of education`, y = `math score`, fill = `parental level of education`)) +

geom\_violin(trim = FALSE) +

labs(title = "Math Scores by Parental Level of Education", x = "Parental Level of Education", y = "Math Score") +

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

library(ggridges)

ggplot(my\_data, aes(x = `math score`, y = `race/ethnicity`, fill = `race/ethnicity`)) +

geom\_density\_ridges() +

labs(title = "Math Score Distribution by Race/Ethnicity", x = "Math Score", y = "Race/Ethnicity") +

scale\_fill\_viridis\_d()

library(ggplot2)

library(reshape2) # For melt function

scores\_cor<- cor(my\_data[,c('math score', 'reading score', 'writing score')])

melted\_cor<- melt(scores\_cor)

ggplot(data = melted\_cor, aes(x = Var1, y = Var2, fill = value)) +

geom\_tile() +

scale\_fill\_gradient2(low = "blue", high = "red", mid = "white", midpoint = 0) +

labs(title = "Correlation between Scores", x = "", y = "") +

theme\_minimal()