A close-up of a school of engineering

Description automatically generated

**LAB PROGRAMS (81-100)**

**ON**

**ITA0402-Statistics with R Programming for Data Visualization**

**SLOT B**

**Submitted by**

**192321093 – MURALI KRISHNAN S**

**To**

**DR.P. Jesu Jayarin**

**Saveetha School of Engineering**

**SIMATS, Thandalam.**

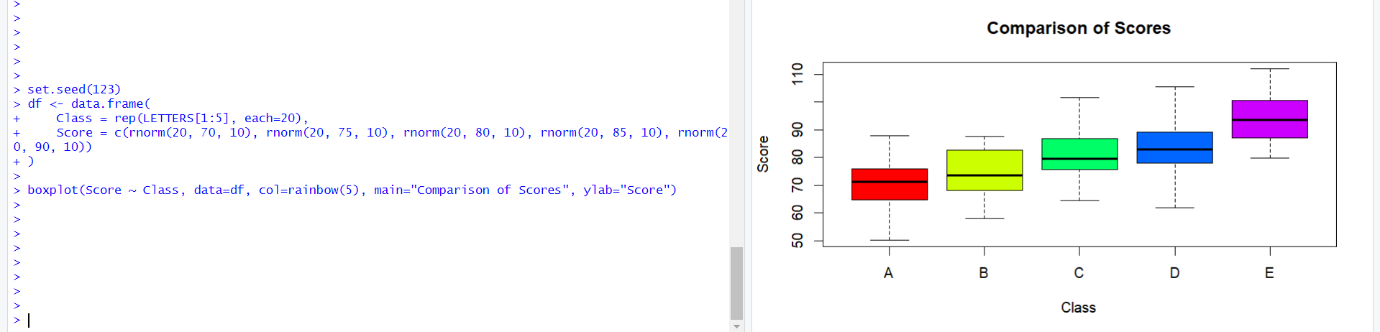
**81. Boxplot of Student Scores Across Five Classes**

**Aim:** Create a data frame and generate a boxplot comparing student scores among five classes.

r

set.seed(123) df <- data.frame( Class = rep(LETTERS[1:5], each=20), Score = c(rnorm(20, 70, 10), rnorm(20, 75, 10), rnorm(20, 80, 10), rnorm(20, 85, 10), rnorm(20, 90, 10)) ) boxplot(Score ~ Class, data=df, col=rainbow(5), main="Comparison of Scores", ylab="Score")

**Output:**



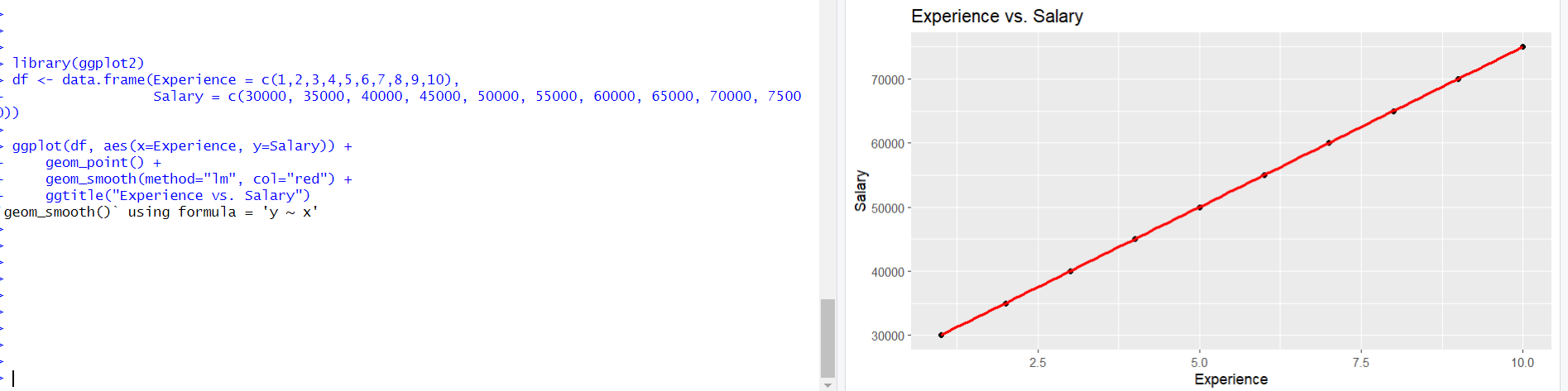
**82. Scatter Plot with Linear Regression Line using ggplot2**

**Aim:** Visualize years of experience vs. salary with a regression line.

r

library(ggplot2) df <- data.frame(Experience = c(1,2,3,4,5,6,7,8,9,10), Salary = c(30000, 35000, 40000, 45000, 50000, 55000, 60000, 65000, 70000, 75000)) ggplot(df, aes(x=Experience, y=Salary)) + geom\_point() + geom\_smooth(method="lm", col="red") + ggtitle("Experience vs. Salary")

**Output:**



**83. Histogram and Modifying Bins**

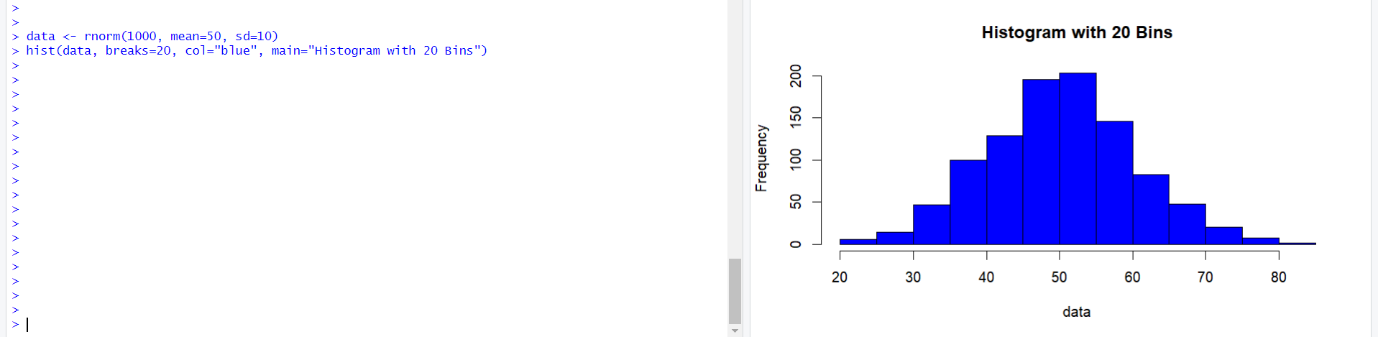
**Aim:** Generate a histogram and modify bins.

r

data <- rnorm(1000, mean=50, sd=10) hist(data, breaks=20, col="blue", main="Histogram with 20 Bins")

*Increase breaks to increase bins.*

**Output:**



**84. Adding Legend to a Plot**

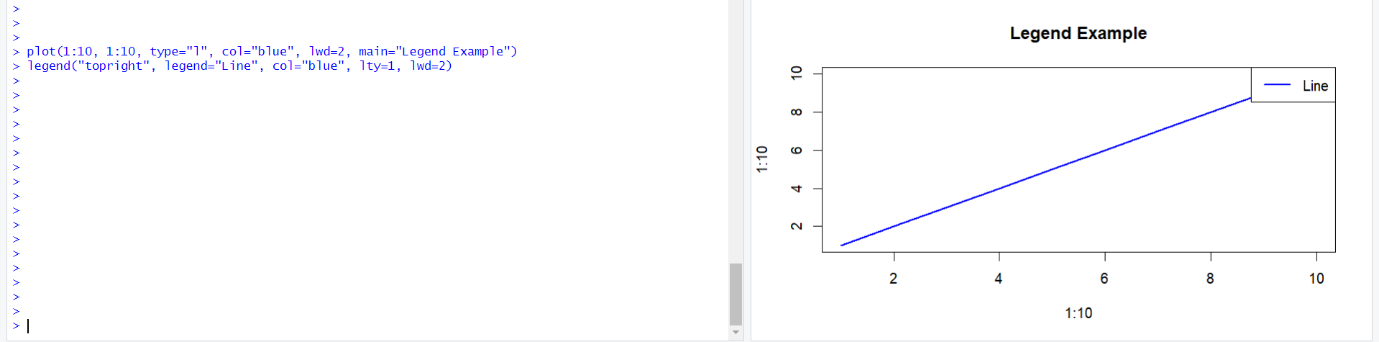
**Aim:** Add a legend using legend().

r

plot(1:10, 1:10, type="l", col="blue", lwd=2, main="Legend Example") legend("topright", legend="Line", col="blue", lty=1, lwd=2)

*legend("position", labels, colors, line types)*

**Output:**



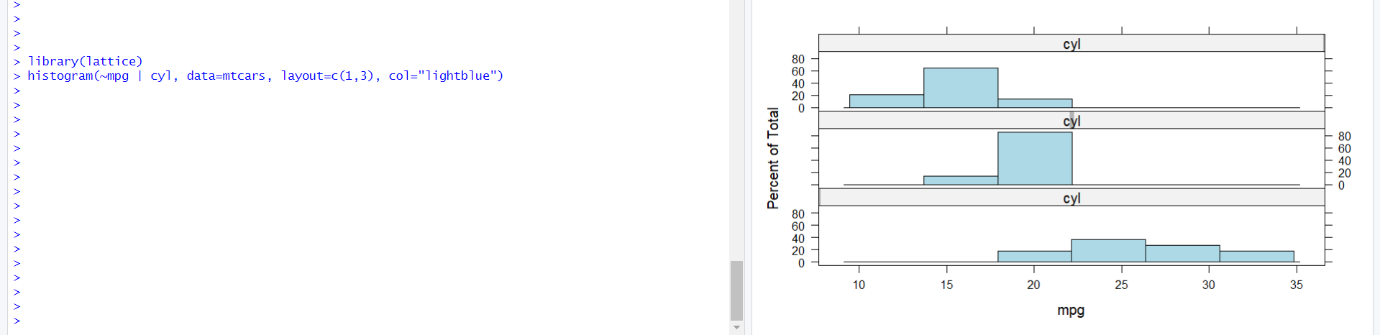
**85. Histogram for Groups Using lattice**

**Aim:** Generate histograms for each group.

r

library(lattice) histogram(~mpg | cyl, data=mtcars, layout=c(1,3), col="lightblue")

**Output:**



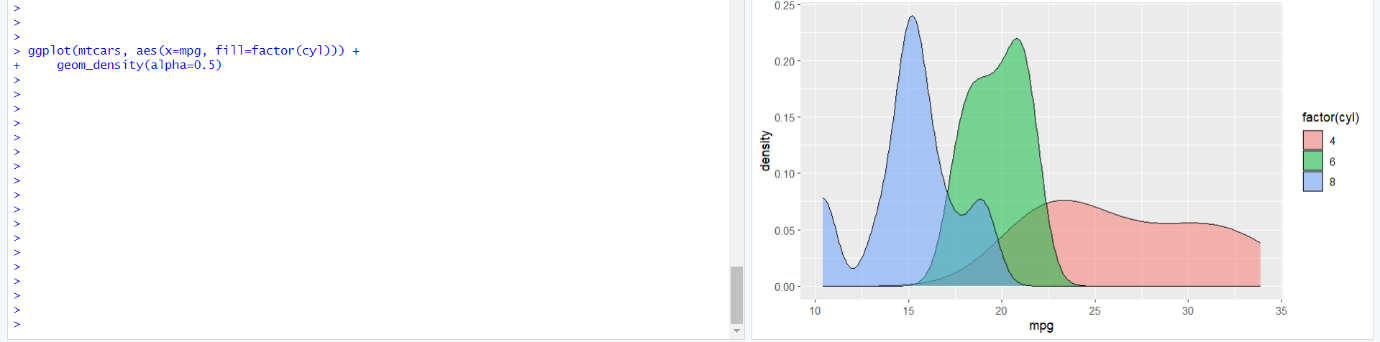
**86. Density Plot using ggplot2**

**Aim:** Visualize distribution with geom\_density().

r

ggplot(mtcars, aes(x=mpg, fill=factor(cyl))) + geom\_density(alpha=0.5)

**Output:**



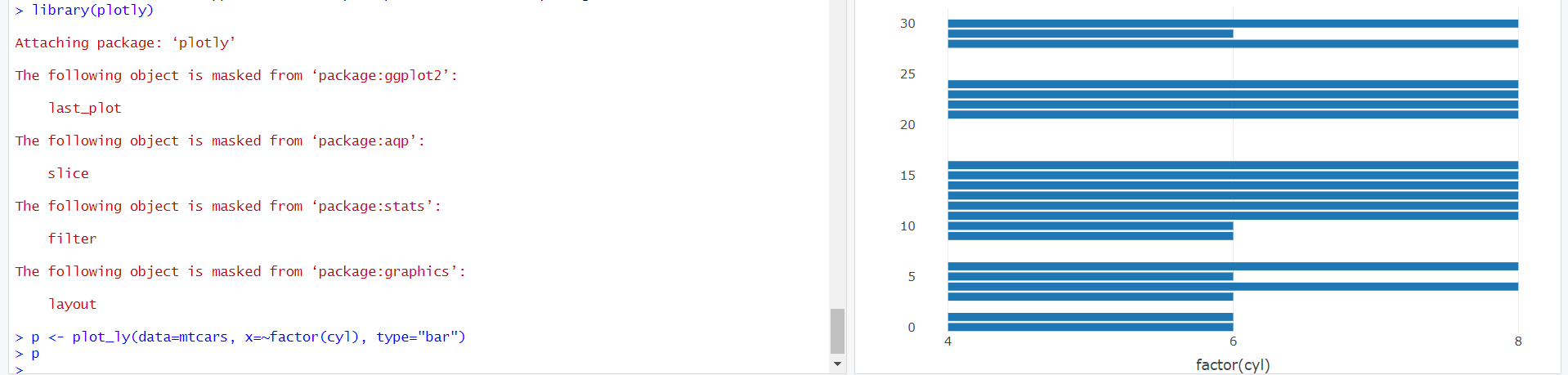
**87. Interactive Bar Plot using plotly**

**Aim:** Create an interactive bar plot of cylinder frequencies.

r

library(plotly) p <- plot\_ly(data=mtcars, x=~factor(cyl), type="bar") p

**Output:**



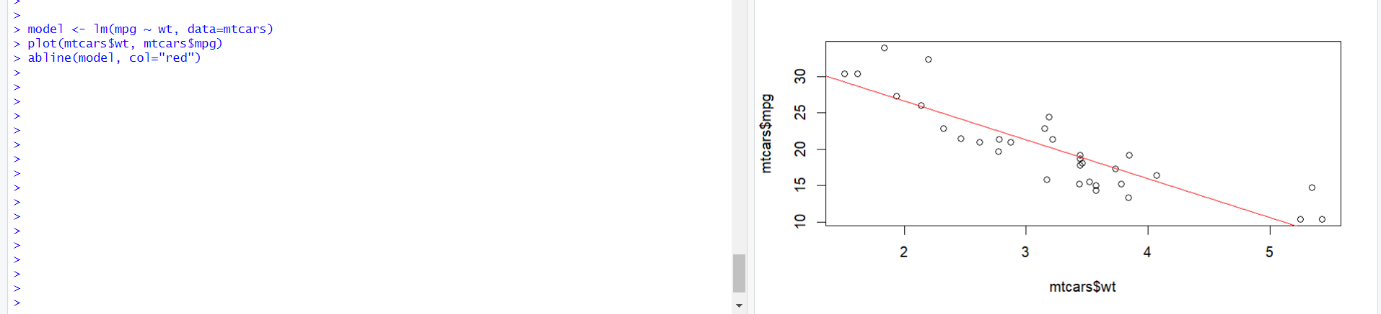
**88. Linear Regression on mpg vs. wt**

**Aim:** Perform regression and plot the line.

r

model <- lm(mpg ~ wt, data=mtcars) plot(mtcars$wt, mtcars$mpg) abline(model, col="red")

**Output:**



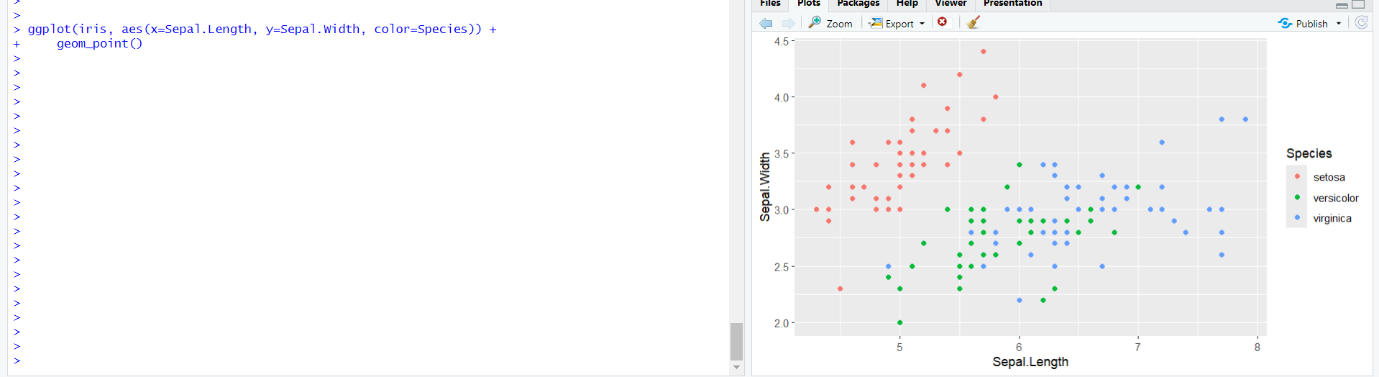
**89. Scatter Plot of Sepal.Length vs. Sepal.Width**

**Aim:** Visualize Sepal dimensions per species.

r

ggplot(iris, aes(x=Sepal.Length, y=Sepal.Width, color=Species)) + geom\_point()

**Output:**



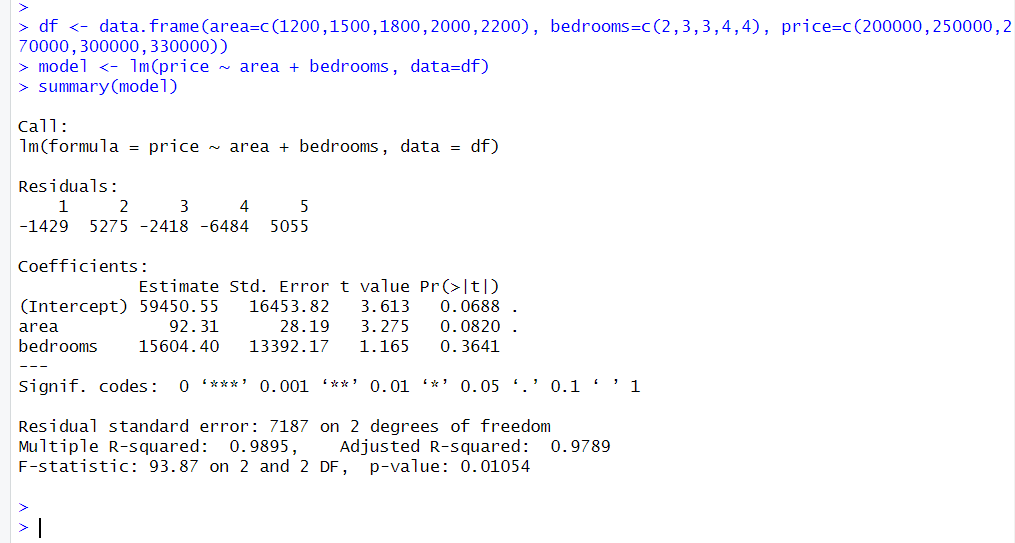
**90. Multiple Regression for House Prices**

**Aim:** Predict house prices based on area and bedrooms.

r

df <- data.frame(area=c(1200,1500,1800,2000,2200), bedrooms=c(2,3,3,4,4), price=c(200000,250000,270000,300000,330000)) model <- lm(price ~ area + bedrooms, data=df) summary(model)

**Output:**

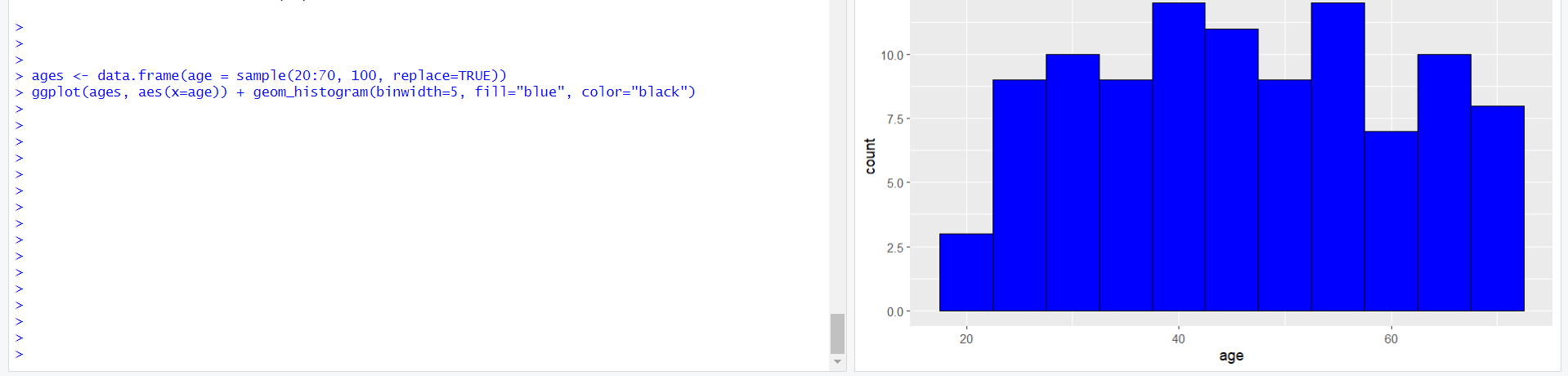


**91. Histogram of Ages Using Seaborn**

r

ages <- data.frame(age = sample(20:70, 100, replace=TRUE)) ggplot(ages, aes(x=age)) + geom\_histogram(binwidth=5, fill="blue", color="black")

**Output:**



**92. Pie Chart for Species Proportions**

**Aim:** Visualize species proportions.

r

species\_count <- table(iris$Species) pie(species\_count, col=rainbow(3))

**Output:**



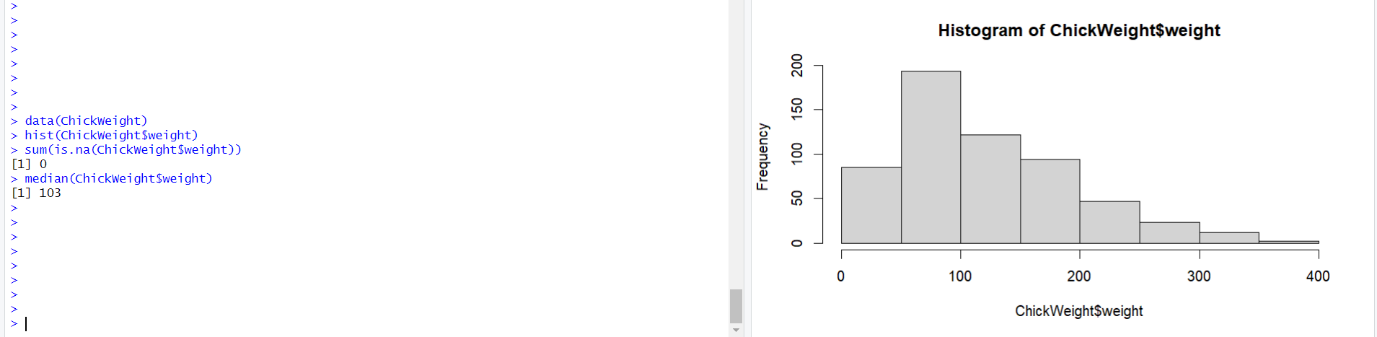
**93. ChickWeight Data Analysis**

**Aim:** Analyze weight distribution, missing values, and median weight.

r

data(ChickWeight) hist(ChickWeight$weight) sum(is.na(ChickWeight$weight)) median(ChickWeight$weight)

**Output:**



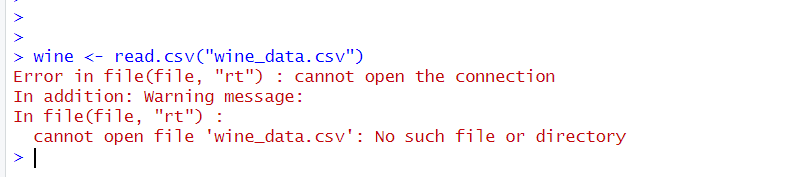
**94. EDA on Wine Dataset**

**Aim:** Load, summarize, visualize, and clean wine dataset.

r

wine <- read.csv("wine\_data.csv") summary(wine) ggplot(wine, aes(x=alcohol, y=quality)) + geom\_point() wine[is.na(wine)] <- median(wine, na.rm=TRUE)

**Output:**



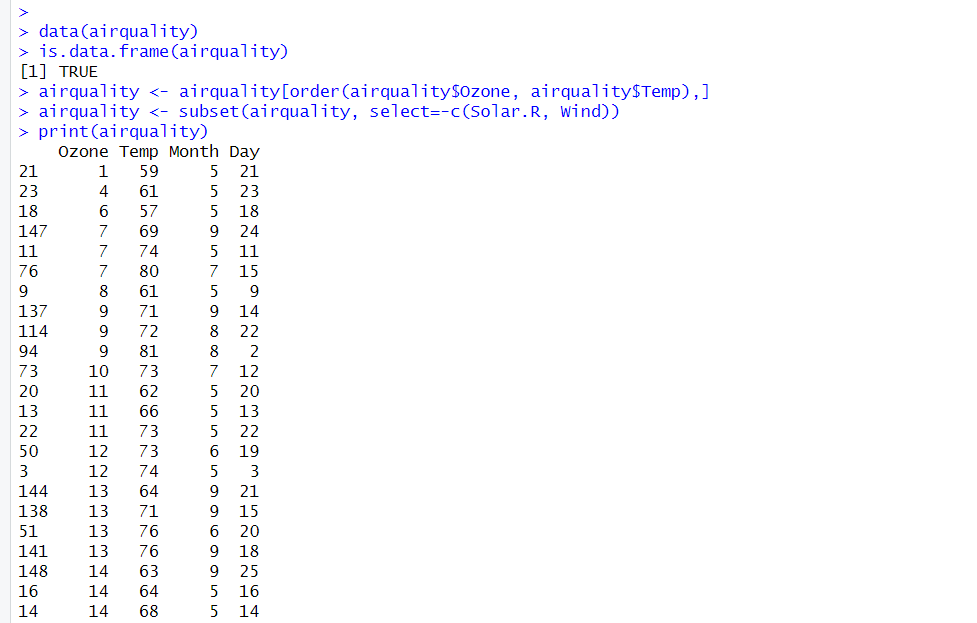
**95. Data Cleaning on airquality Dataset**

**Aim:** Check if airquality is a data frame, order it, and remove columns.

r

data(airquality) is.data.frame(airquality) airquality <- airquality[order(airquality$Ozone, airquality$Temp),] airquality <- subset(airquality, select=-c(Solar.R, Wind)) print(airquality)

**Output:**



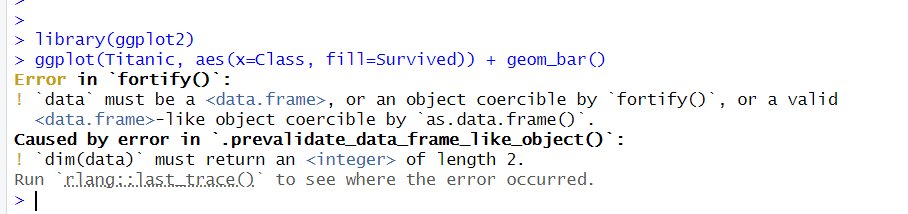
**96. Titanic Dataset Visualization**

**Aim:** Visualize survival distribution and age histogram.

r

library(ggplot2) ggplot(Titanic, aes(x=Class, fill=Survived)) + geom\_bar() ggplot(Titanic, aes(x=Age)) + geom\_histogram(binwidth=5)

**Output:**



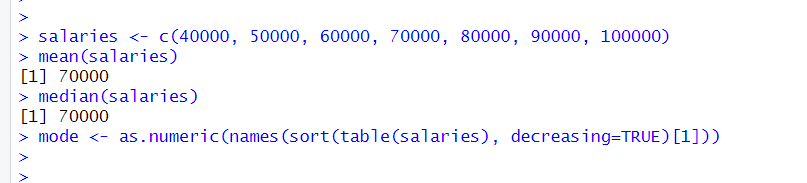
**97. Salary Statistics and Interpretation**

**Aim:** Calculate mean, median, and mode for salary.

r

salaries <- c(40000, 50000, 60000, 70000, 80000, 90000, 100000) mean(salaries) median(salaries) mode <- as.numeric(names(sort(table(salaries), decreasing=TRUE)[1]))

**Output:**



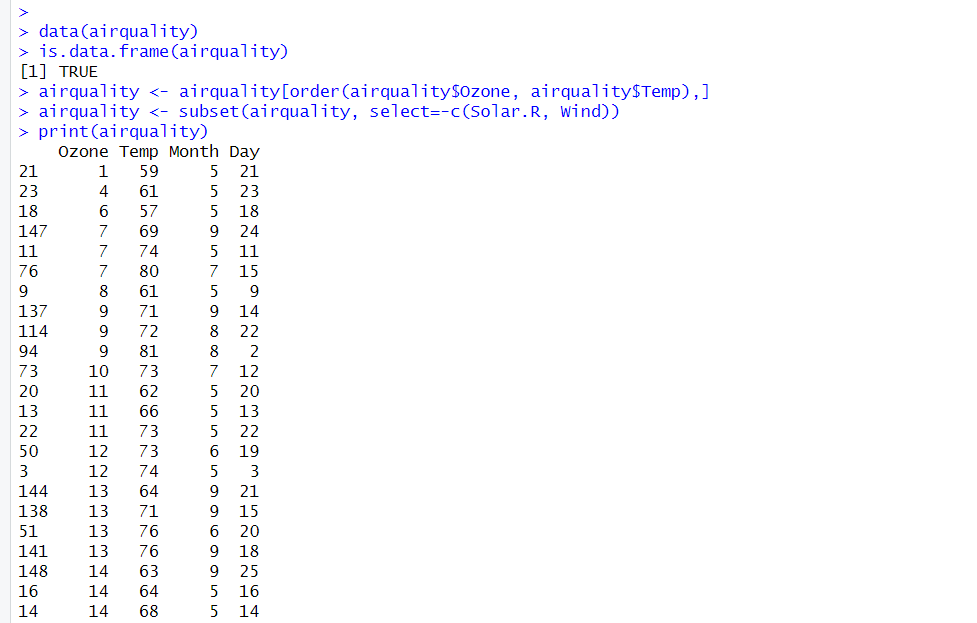
**98. Data Cleaning on airquality Dataset (Repeated)**

**Aim:** Check if airquality is a data frame, order it, and remove columns.

r

data(airquality) is.data.frame(airquality) airquality <- airquality[order(airquality$Ozone, airquality$Temp),] airquality <- subset(airquality, select=-c(Solar.R, Wind)) print(airquality)

**Output:**



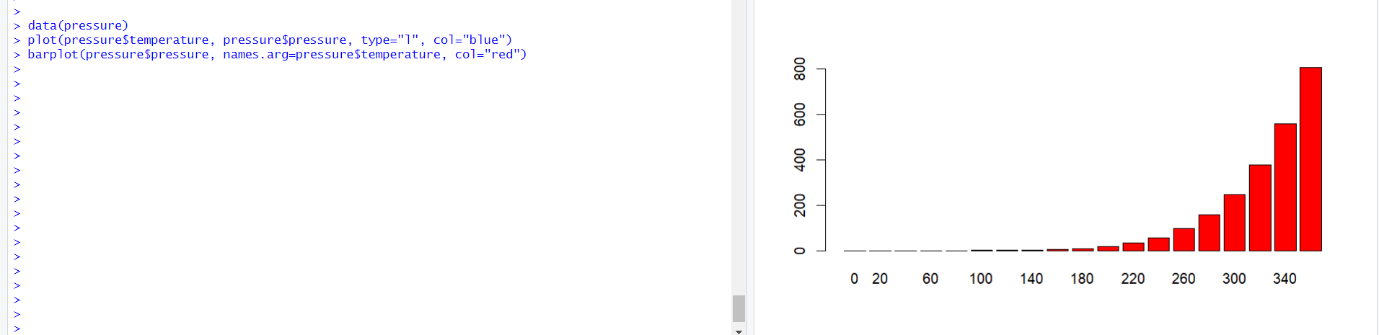
**99. Line Chart and Bar Chart for Pressure Dataset**

**Aim:** Create a line and bar chart.

r

data(pressure) plot(pressure$temperature, pressure$pressure, type="l", col="blue") barplot(pressure$pressure, names.arg=pressure$temperature, col="red")

**Output:**



**100. Multiple Regression on mtcars (mpg ~ wt + hp)**

**Aim:** Predict mpg using weight and horsepower.

r

model <- lm(mpg ~ wt + hp, data=mtcars) summary(model)

**Output:**

