SECTION 2

Part A:

```
# Clear environment

rm(list = ls())

# Set working directory

setwd("/Users/vishvraval/Desktop/R")

# Load the dataset

inputData <- read.table("data.txt")

# Assign each column to a variable

col1 <- inputData[, 1]

col2 <- inputData[, 2]

col3 <- inputData[, 3]

# Install and load the ggplot2 package for plotting

install.packages("ggplot2")

library(ggplot2)

# Generate histograms for each column

# Column 1

hist(col1, main = "Histogram: Data Column 1", xlab = "Values", col = "red")

# Column 2

hist(col2, main = "Histogram: Data Column 2", xlab = "Values", col = "blue")

# Column 3

hist(col3, main = "Histogram: Data Column 3", xlab = "Values", col = "green")

# Calculate mean and variance for each column

col1Mean <- mean(col1)

col1Var <- var(col1)

col2Mean <- mean(col2)
```

```
col2Var <- var(col2)
col3Mean <- mean(col3)
col3Var <- var(col3)
# Display the mean and variance for each column
cat("Column 1 - Mean:", col1Mean, "Variance:", col1Var, "\n")
cat("Column 2 - Mean:", col2Mean, "Variance:", col2Var, "\n")
cat("Column 3 - Mean:", col3Mean, "Variance:", col3Var, "\n")
# Shapiro-Wilk test for normality on each column
# Column 1
shapiro_test_col1 <- shapiro.test(col1)
cat("Column 1 - Shapiro-Wilk p-value:", shapiro_test_col1$p.value, "\n")
# Column 2
shapiro_test_col2 <- shapiro.test(col2)
cat("Column 2 - Shapiro-Wilk p-value:", shapiro_test_col2$p.value, "\n")
# Column 3
shapiro_test_col3 <- shapiro.test(col3)
cat("Column 3 - Shapiro-Wilk p-value:", shapiro_test_col3$p.value, "\n")
```

Part B:

```
# Select data from Column 1
sampleData <- inputData[, 1]
# Sample size, mean, and standard deviation
ssize <- length(sampleData)
smean <- mean(sampleData)
ssd <- sd(sampleData)
# Calculate Confidence Interval for the Mean (99%)
confidenceLevel <- 0.01
critical_t <- qt(1 - confidenceLevel / 2, df = ssize - 1)
# Calculate lower and upper bounds of the confidence interval
ci_lowerBound <- smean - critical_t * (ssd / sqrt(ssize))
ci_upperBound <- smean + critical_t * (ssd / sqrt(ssize))
cat("Confidence Interval: [", ci_lowerBound, ", ", ci_upperBound, "]\n")
# Hypothesis Test for Variance(sigma^2 = 0.5 )
hypothesizedVar <- 0.5
chi_squareStat <- (ssize - 1) * var(sampleData) / hypothesizedVar
# Compute p-value
pvalueVar <- 2 * min(
pchisq(chi_squareStat, df = ssize - 1, lower.tail = TRUE),
pchisq(chi_squareStat, df = ssize - 1, lower.tail = FALSE))
cat("Variance Test p-value:", pvalueVar, "\n")
```

Part C:

 # Analysis on column 3

```
col3 <- inputData[, 3]
```

 # Calculate the proportion of defective parts

```
proportionDefect <- mean(col3 == 0)

ssize_col3 <- length(col3)
```

 # Hypothesis Test for Proportion of Defective Parts

```
proportionHypothesized <- 0.10

Zscore <- (proportionDefect - proportionHypothesized) / sqrt(proportionHypothesized * (1 - proportionHypothesized) / ssize_col3)
```

 # Calculate p-value for the hypothesis test

```
proportionPvalue <- pnorm(Zscore, lower.tail = TRUE)

cat("Proportion Test p-value:", proportionPvalue, "\n")
```