Write the skewness program, and use it to calculate the skewness coefficient of the four examination subjects in results.txt (results.csv).

```r
results <- read.table("results.txt", header = T)
skewness <- function(x) {
    xbar <- mean(x, na.rm = T)
    sum2 <- sum((x - xbar)^2, na.rm = T)
    sum3 <- sum((x - xbar)^3, na.rm = T)
    skewness <- (sqrt(length(x)) * sum3)/(sum2^(1.5))
    skewness
}
pearson_skewness <- function(x) {
    mean <- mean(x, na.rm = T)
    median <- median(x, na.rm = T)
    sdev <- sd(x, na.rm = T)
    skewness <- ((3 * (mean - median))/sdev)
    skewness
}
results_df <- data.frame(Arch1 = c(skewness(results$Arch1), pearson_skewness(results$Arch1)),
    Prog1 = c(skewness(results$Prog1), pearson_skewness(results$Prog1)), Arch2 = c(skewness(results$Arc
        pearson_skewness(results$Arch2)), Prog2 = c(skewness(results$Prog2), pearson_skewness(results$P
rownames(results_df) <- c("Skewness", "Pearson Skewness")
results_df
```

```
##                        Arch1      Prog1      Arch2      Prog2
## Skewness          -0.5195368 -0.3362643 0.4558875 -0.3125144
## Pearson Skewness  -0.6069042 -0.6432290 0.5421286 -0.3562908
```

What can you say about these data? And is it a reasonable approximation?

The normal skewness indicates that scores are leaned left, meaning more students scoring higher than lower, although only by a small majority. Looking at the data, it seems there is a reasonable similarity between the normal skewness and the Pearson skewness, with the latter having slighlty higher or even lower results.

For the class of 50 students of computing detailed in Exercise 1.1, use R to form the stem-and-leaf display for each gender.

Excercise 1.1: Females: 57, 59, 78, 79, 60, 65, 68, 71, 75, 48, 51, 55, 56, 41, 43, 44, 75, 78, 80, 81, 83, 83, 85
Males: 48, 49, 49, 30, 30, 31, 32, 35, 37, 41, 86, 42, 51, 53, 56, 42, 44, 50, 51, 65, 67, 51, 56, 58, 64, 64, 75

```r
female_scores <- c(57, 59, 78, 79, 60, 65, 68, 71, 75, 48, 51, 55, 56, 41, 43, 44,
    75, 78, 80, 81, 83, 83, 85)

male_scores <- c(48, 49, 49, 30, 30, 31, 32, 35, 37, 41, 86, 42, 51, 53, 56, 42,
    44, 50, 51, 65, 67, 51, 56, 58, 64, 64, 75)

all_scores <- c(female_scores, male_scores)

gender <- factor(c(rep("Female", length(female_scores)), rep("Male", length(male_scores))))

cat("Stem-and-Leaf Plot for Females:\n")
```

1

```
## Stem-and-Leaf Plot for Females:

stem(all_scores[gender == "Female"])


##
##   The decimal point is 1 digit(s) to the right of the |
##
##   4 | 1348
##   5 | 15679
##   6 | 058
##   7 | 155889
##   8 | 01335

cat("\nStem-and-Leaf Plot for Males:\n")


##
## Stem-and-Leaf Plot for Males:

stem(all_scores[gender == "Male"])


##
##   The decimal point is 1 digit(s) to the right of the |
##
##   3 | 001257
##   4 | 1224899
##   5 | 01113668
##   6 | 4457
##   7 | 5
##   8 | 6
```

Discuss the advantages of this representation compared to the traditional histogram.

Using a stem and leaf display model can easily allow one to visually assess and compare data quickly, as it highlights patterns and points with a simple design. However, one can imagine that it would not be so useful or ideal with large amounts of data. That is why stem and leaf display is more useful in small or moderate sized datasets. Another advantage of stem and leaf is it allows the viewer to see outliers quite easily with its simple visual display. In short, stem and leaf displays have simpler and more compact visuals which allow easier and faster analysis on data, but is only really viable for small to medium sized datasets.
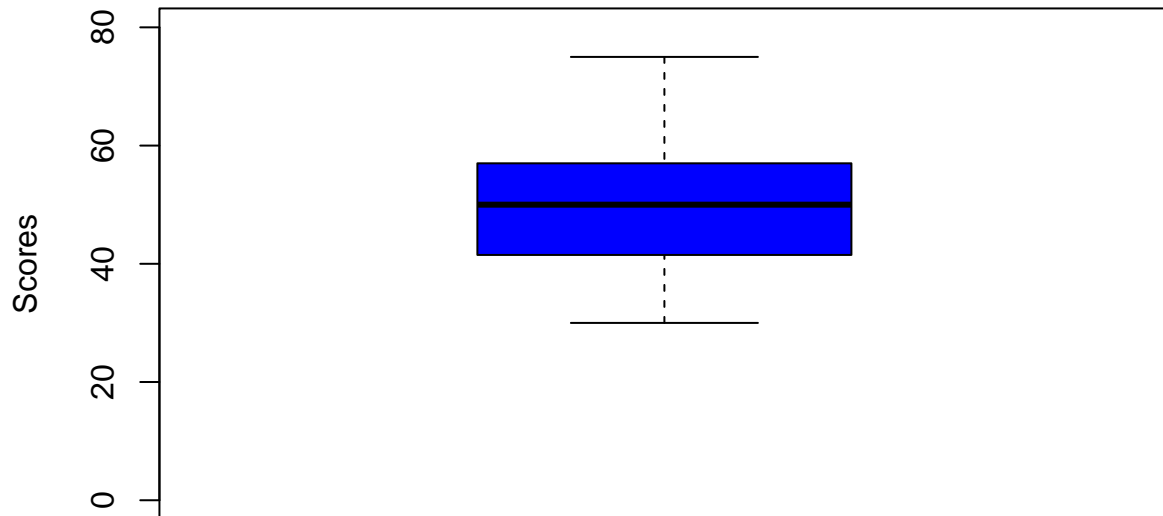
Construct a box-plot for each gender.

```
boxplot(female_scores, names = ("Females"), main = "Java Programming Exam Scores (Female)",
    ylab = "Scores", ylim = range(0, 80), col = ("red"))
```

Java Programming Exam Scores (Female)
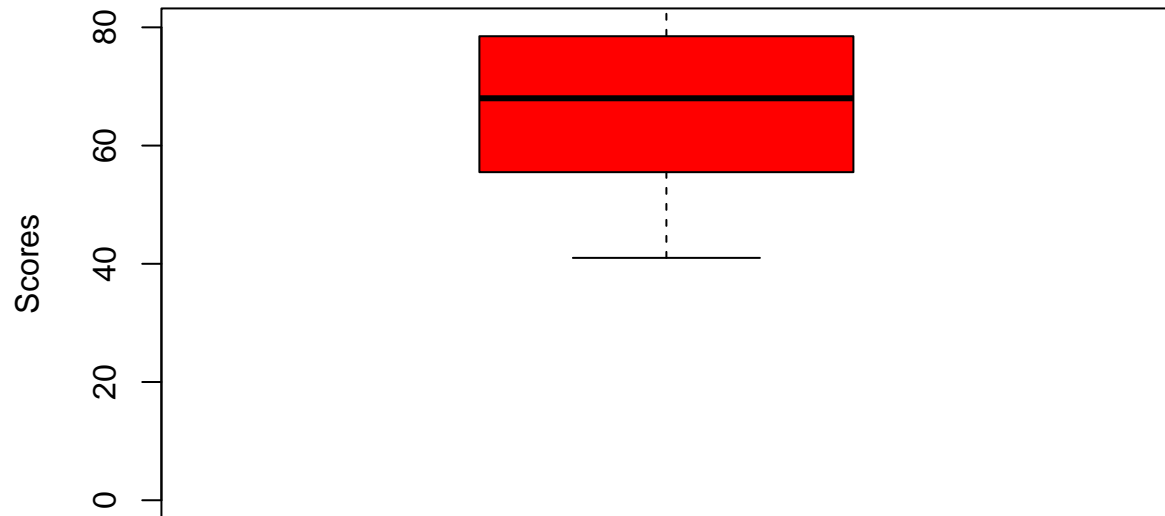
Scores

```
boxplot(male_scores, names = ("Males"), main = "Java Programming Exam Scores (Male)",
    ylab = "Scores", ylim = range(0, 80), col = ("blue"))
```

Java Programming Exam Scores (Male)

Discuss the findings.

These graphs shows that, generally, the females outperformed the males in the Java Programming Exam by a decent amount. It must also be noticed that the male scores is less varied when compared to the female scores.

Github Link: https://github.com/SylTana/APM1110-QUIJANO-JULIAN_PHILIP/tree/main/FA1