Data:

$female\ scores = [57, 59, 78, 79, 60, 65, 68, 71, 75, 48, 51, 55, 56, 41, 43,$

$44, 75, 78, 80, 81, 83, 83, 85]$

$male\ scores = [48, 49, 49, 30, 30, 31, 32, 35, 37, 41, 86, 42, 51, 53, 56,$

$42, 44, 50, 51, 65, 67, 51, 56, 58, 64, 64, 75]$

Item (a):

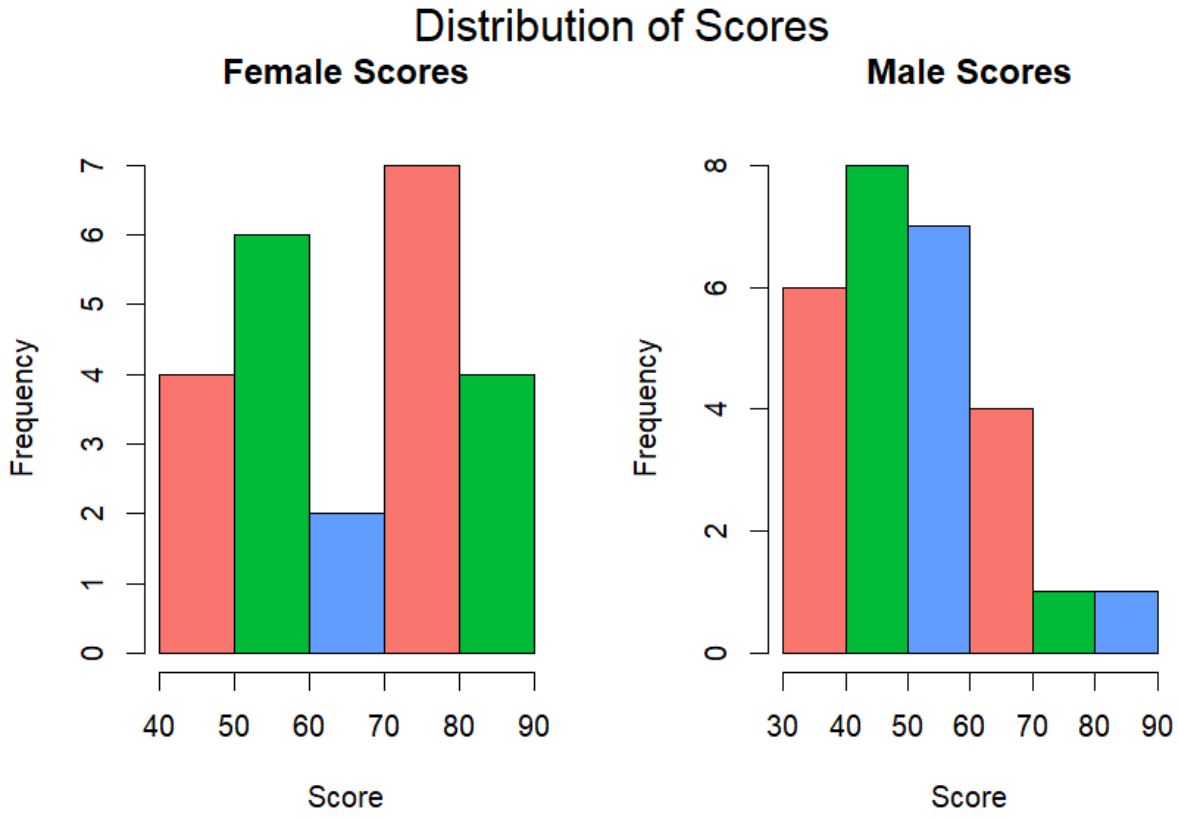
Instruction:

Form the stem-and-leaf display for each gender and discuss the advantages of this representation compared to the traditional histogram.

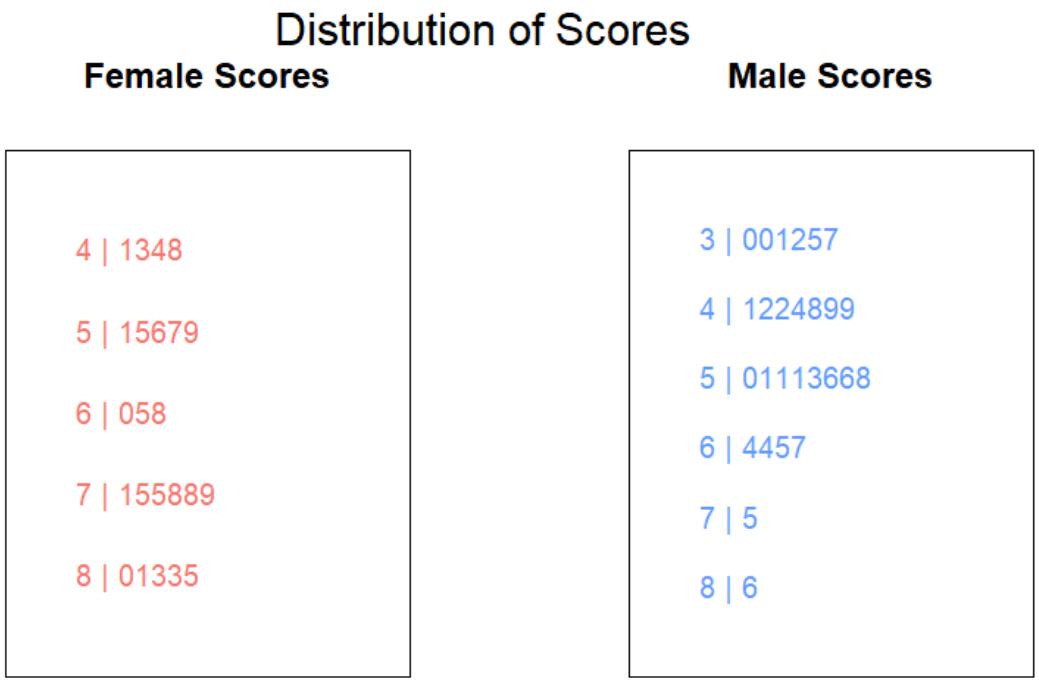
Answer:

As a reference, here are stem-and-leaf and histogram representation of the data given;

Histogram:



Stem-and-leaf:



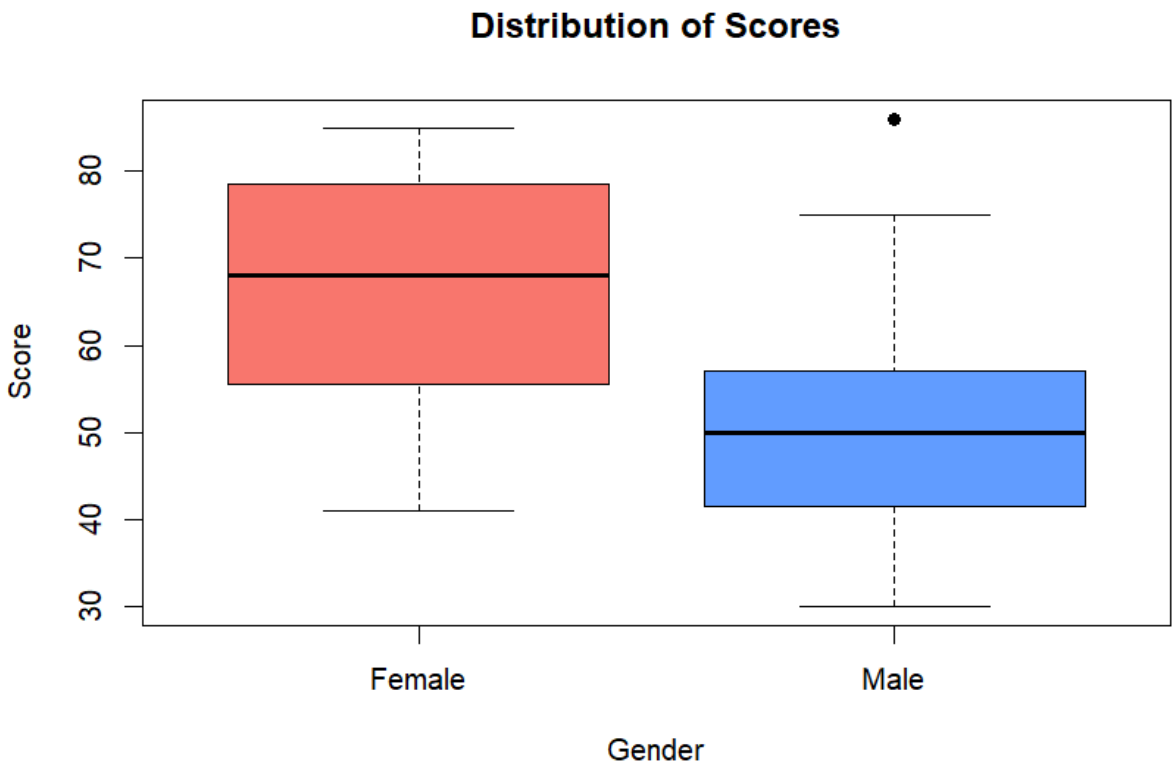
The primary advantage of doing a stem-and-leaf plot instead of a histogram is that stem-and-leaf shows the specific data instead of clustering them in a certain range of numbers. In this way, one can get a good grasp as to what exactly are the numbers being plotted. In a histogram, although it is much easier to absorb, it may not be a good representation when one wants to know what exact number is being repeated. For example, on the "7" cluster of the female scores. If one looks at the histogram, one can be led to think that students scored 70-79 in the examination—which is true but a rather short-sighted and tunneled perception for all but one are ranging from 75-79. This could be solved by shrinking the size of each cluster of histogram but the fact remains that if one wants to get into the detail (for whatever reason), he/she can do so using steam-and-leaf representation.

Item (b):

Instruction: Construct a box-plot for each gender and discuss the findings.

Answer:

Box-plot:



Upon analyzing the box-plot for each gender, it can be clearly observed that females, as a cluster, performed better compared to their male counterparts—as evident by their higher median score (the line in the middle of the box). The vertical size of each box indicated that there exists a greater spread of scores around the median among females compared to males. Excluding the outliers, the male scores have a lower minimum and a lower maximum as was shown by the length of the whiskers (the lines extending from the boxes). As for the outlier, there appears to be an outlier from the male cluster which marks that there exists one male who scored exceptionally high, surpassing the general performance of his peers. To summarize, the box=plot showed that females appear to have performed better overall compared to males in this specific test with a notable high-scoring male outlier. However, take the word "better" with a grain of salt for further statistical testing should be done to confer this term to the data—"better" should be understood as descriptive rather than inferential observation.