NAME: <u>Markus Afonso</u>          Set: <u>C</u>
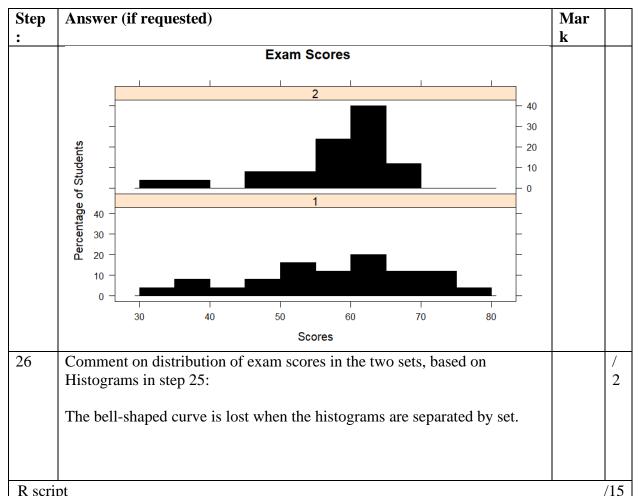
MATH 1350          **Statistics for Information Technology**

**Lab # 2 – Descriptive Statistics II (with R)**

Answer/Grading Sheet

| Step: | Answer (if requested) | Mark | |
|---|---|---|---|
| 10 | • Stemplot of Scores (paste here)<br><br>```<br>3 | 2<br>3 | 566<br>4 | 02<br>4 | 789<br>5 | 012444<br>5 | 56688999<br>6 | 0011111123444444<br>6 | 5777889<br>7 | 122<br>7 | 7<br>```<br><br>• Comment on "average" score and distribution (shape): (type your answer here)<br><br>Average is around 60, and the distribution is a left skewed bell-shaped curve. | | / 1<br><br>/ 2 |
| 13 | • Histogram with all appropriate features/options<br><br><br>Exam Scores | | / 4 |

| Step : | Answer (if requested) | Mar k | |
|---|---|---|---|
| | | | |
| 16 | • Table with values for a) – f) <br><br> <table><tr><td>a) Mean: 57.96</td><td>b) Median: 60.5</td><td>c) Standard Deviation: 10.22793</td></tr><tr><td>d) Variance: 104.6106</td><td>e) Range: (32, 77)</td><td>f) Interquartile Range: 10</td></tr></table> | | / 3 |
| 21 | • Summary of the results from step 20 (mean, median, std dev, variance, range and IQR by set) <br><br>```
min Q1 median Q3 max   mean       sd  n missing
 32 54   60.5 64  77 57.96 10.22793 50       0
``` <br><br> • Which set appears to have higher exam scores? Justify. <br><br> Set 1 has higher exam scores, this is proven by the higher min, max, and mean. Firstly, the higher min and max scores shows the difference in scores between the sets. Secondly, the higher mean in set 1 shows that set 1 scores are on average higher than set 2. This means that if I were to pick two students, one from set 1 and the other from set 2. It is more likely that the student from set 1 would have a higher score than the student from set 2. <br><br> ```
> favstats(Score~Set,data=examscores)
  Set min Q1 median Q3 max   mean       sd  n missing
1   1  35 51     59 67  77 58.04 11.649320 25       0
2   2  32 56     61 64  68 57.88  8.823831 25       0
``` <br><br> • Which set has more variable exam scores? Justify. <br><br> Set 1 as it has a greater standard deviation. We can tell this by the sd (standard deviation) score under the favstats command. | | / 3 <br><br> / 1 <br><br> / 1 |
| 25 | Vertically stacked Histogram plot by set, with all required options | | / 6 |

| Step : | Answer (if requested) | Mar k | |
|---|---|---|---|
| | **Exam Scores**  | | |
| 26 | Comment on distribution of exam scores in the two sets, based on Histograms in step 25:<br><br>The bell-shaped curve is lost when the histograms are separated by set. | | / 2 |

**R script**  /15

Paste your R script below. It should contain:
- Comment saying "Lab 2" and comment giving your name.
- Command to turn on mosaic.
- Command to import data from "Lab2.txt".
- Commands for a stemplot and a histogram for the full "examscores" dataset.
- Command to convert the "Set" variable to a categorical (factor) variable
- Commands for calculating the mean, median, standard deviation, variance, range, IQR and favstats of the full "Score" data
- Command for the 80th percentile of all Scores
- Commands for mean, median, standard deviation, variance, range, IQR by Set
- Command for the vertically stacked Histogram plot by Set

```
# Lab 2
# Markus Afonso

library(mosaic)
```

| Step : | Answer (if requested) | Mark | |
|---|---|---|---|
| | ```r
examscores <- read.delim("C:/Users/Markus/OneDrive -
BCIT/Desktop/Term2/MATH 1350 Statistics for IT/Week2/examscores.txt",
comment.char="#")

with(examscores,stem(Score))
par(mfrow=c(1,2))
histogram(Score~Set, data=examscores)
histogram(~Score|Set, data=examscores,layout=c(1,2), main = "Exam Scores",
          xlab = "Scores", ylab = "Percentage of Students", type = "p",
          col="black", breaks=seq(30,80,5), width=10)


examscores$Set=as.factor(examscores$Set)

mean(~Score,data=examscores)
median(~Score,data=examscores)
sd(~Score,data=examscores)
var(~Score,data=examscores)
min(~Score,data=examscores)
max(~Score,data=examscores)
IQR(~Score,data=examscores)

favstats(Score~Set,data=examscores)

sum(Score~Set,data=examscores)

with(examscores, quantile(Score, 0.99))


histogram(~Score|Set,data=examscores)
``` | | |
| | | | |

Total      /38