STAT 441/541 Statistical Methods II

Homework Assignment 1 over Review Material

Submit a single pdf document to the Dropbox folder *Homework Assignment 1 Review Material*.

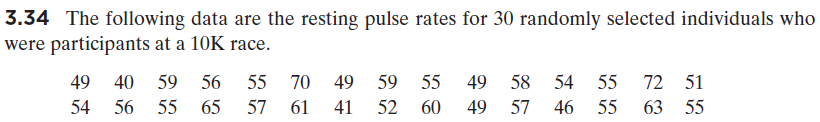
Start each exercise on a new page.

Note: The scenarios and data are the same as in our text. Some questions have been modified.

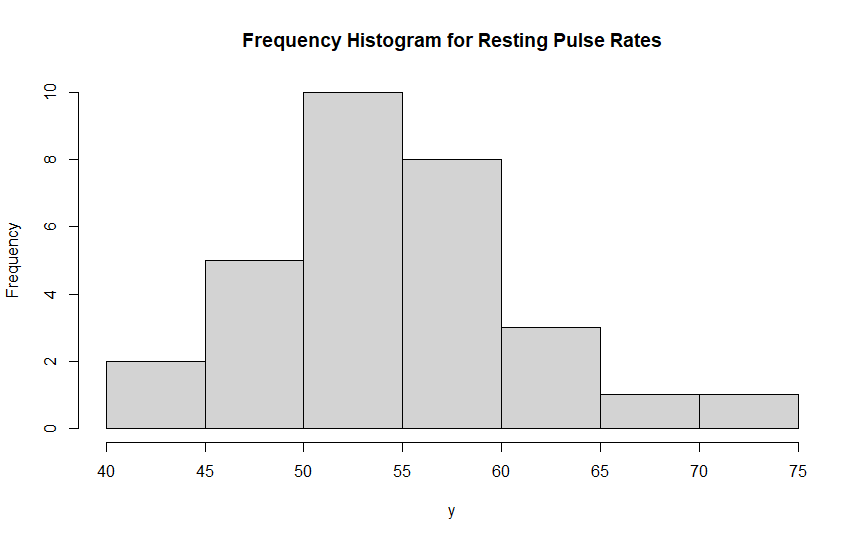
Exercise 3.34

The R code file: Exercise 3-34 R Code.R

The dataset is Excel file: ex3-34.xlsx



(a) Construct, or paste plot from R output, a histogram of the pulse rates and describe the shape of the distribution of the pulse rates.



The shape of histogram is nearly bell shaped or we can say near normal. Although, it is a bit skewed towards right still it can be labelled as bell shaped.

(b) Construct, or paste plot from R output, a boxplot of the pulse rates and describe the shape of the distribution of the pulse rates.

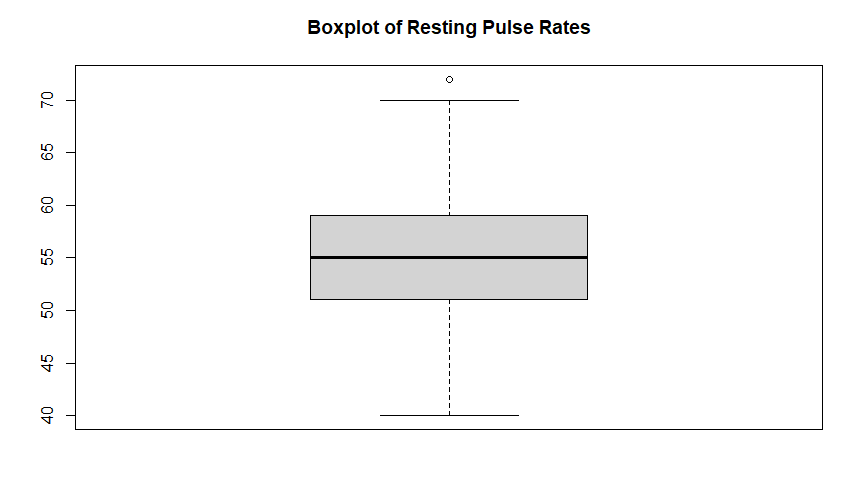




Figure above illustrates the box plot of pulse rates. Again, we can see that almost all quartiles are equally placed from each other. The median line in the centre of the box means the pulse rates are symmetrical. There is only one outlier which lies on the higher side and no outliers exists on the lower side.

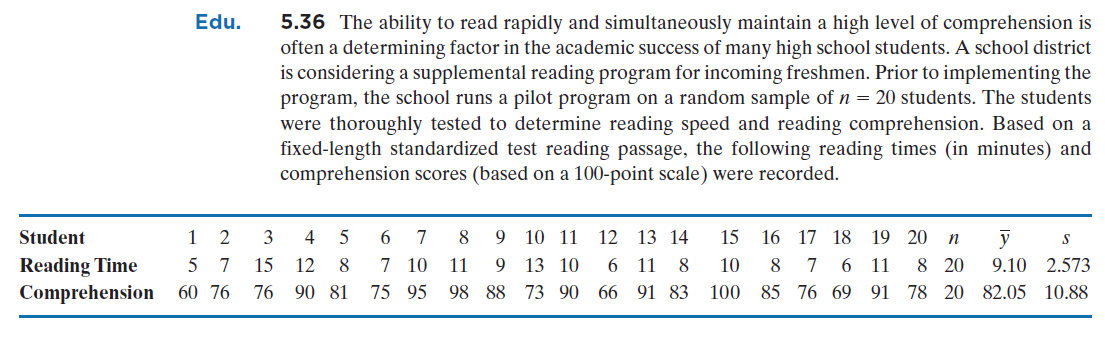
(c) The histogram and boxplot provides information about the distribution of pulse rates for what population. That is, what is the population from which we selected our random sample?

The population from which the samples were selected is people who participated in 10K race.

Exercise 5.36

The R code file: Exercises 5-36 and 5-37 R Code.R

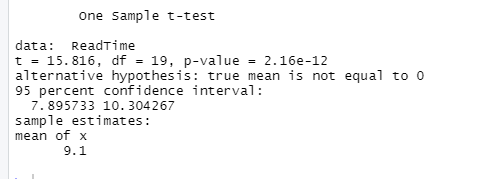
The dataset is Excel file: ex5-36.xlsx



1. What is the population about which inferences are being made?

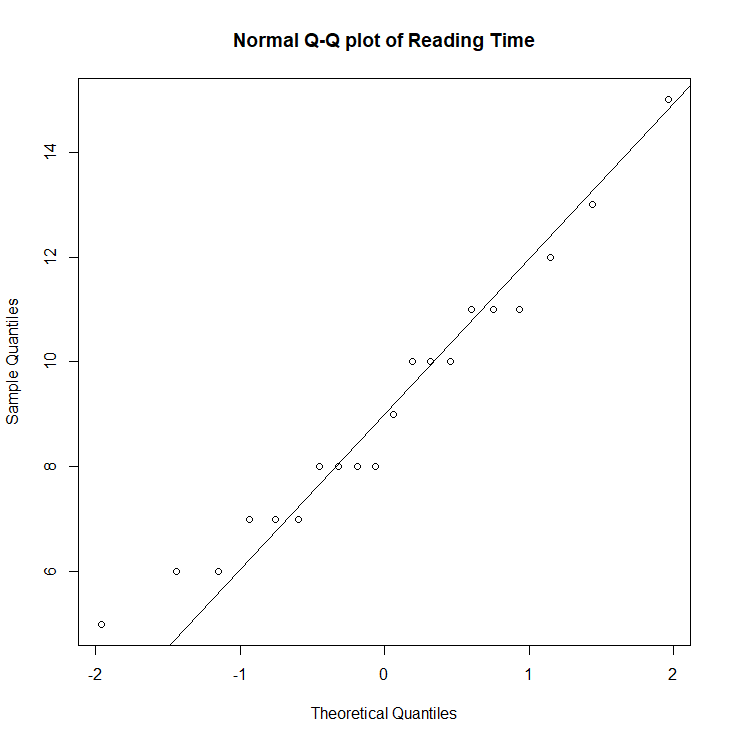
Incoming freshmen who participated in fixed-length standardized reading test.

1. Construct, or paste output from R, a 95% confidence interval on the mean reading time for all incoming freshmen in the district.



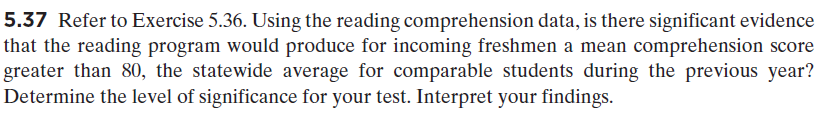
1. Provide an interpretation of the interval estimate in part (b).
2. The lower limit of confidence interval is 7.89 while upper limit is 10.30
3. The confidence level is 95 percent which means if 100 random samples are picked from 95 of that will contain true population mean.

(d) Construct, or paste plot from R output, a normal probability plot of reading time. Do the data appear to be a random sample from a population having a normal distribution?



As it can be seen in the figure that, the points appear to be closely related to the fitted straight line in normal Q-Q PLOT. Hence, it plausible that the sample is a randomly picked from a population having normal distribution.

Exercise 5.37



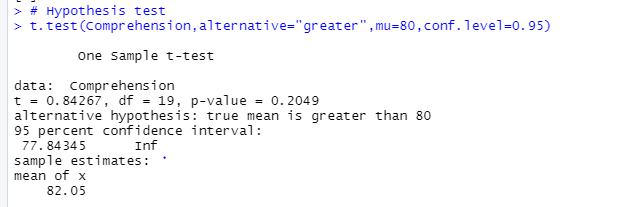
Provide all five parts of the statistical test and use a significance level of .

1. Null and Alternative Hypotheses:

: mean comprehension score less than or equal to 80

: mean comprehension scores greater than 80

2. Test Statistic. T.S. =



3. *P*-value = 0.2049

4. **Decision about the null hypothesis:** The p-value i.e., 0.2049 is very large as compared to 0.05. As the p-value is very high, the NULL hypothesis cannot be rejected beyond a reasonable doubt.

**5. Conclusion:** As mentioned in point 4, the NULL hypothesis cannot be rejected based on value of p. Thus, it can be concluded that there is no evidence beyond a reasonable doubt that reading program is able to produce a means comprehension score of greater than 80.