Assignment 6

Load the dataset "hr.RData" into R. This dataset contains patients' heart rate before and after drug treatment. Gender (m/f), side effect (1=yes, 0=no), and age (young, middle, elderly) are also provided.

This assignment requires completion by hand or with R. Submission for this assignment will not be via a Canvas quiz, but via a document submission. Save your work as a pdf or scan or photograph your handwritten work to submit (again as a pdf).

Paired t test

Researchers want to conclude that this drug *lowers* these patients heart rate. The average difference for this sample (after – before) is -3.33 and the standard deviation (not standard error!) for this difference is 4.88. The sample size (n) is 30.

1) Complete this paired t test. Set up hypotheses, calculate the test statistic, find the p-value, and draw a conclusion about this problem.

2) Now us the `t.test()` function in R to complete this test again and compare your results. Consult the help page of `?t.test` to see what parameters are required for this function (in particular, pay attention to the "alternative" and "paired" parameters). Compare your answers to your manual calculations above. Note: R may use slightly different formulas so your answers may not be exactly the same, but they should be close.

2-sample t test

Researchers now want to test whether or not males and females have different heart rates before (and only before) drug treatment. The average female heart rate before drug treatment in the sample is 75.2 and the average male heart rate is 67.8. The standard deviation for females' before heart rate is 12.9 and for males it is 14.3. There are 15 females and 15 males.

3) Complete this 2 sample t-test (without using the `t.test()` function). Set up a hypotheses, calculate the test statistic, find the p-value, and draw a conclusion about this problem.

4) Calculate a 95% confidence interval for this average difference. You may need to use `qt()` to find the t* value.

5) Conduct this two-sample t-test using `t.test()` and compare your results to your manual calculations. Note that when you are doing a two sample test you can enter each sample separately, i.e `t.test(x = sample1, y = sample2, … )` (which may be how you did the paired t test above), or you can enter a formula where the LHS contains all the values and the RHS is a grouping variable, i.e. `t.test(values ~ groups, … )`.

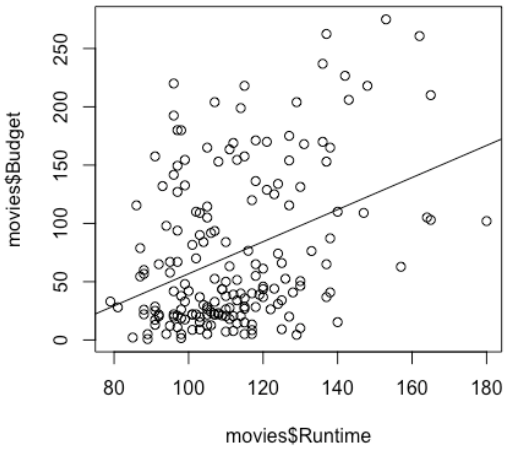
ANOVA

Conduct the ANOVA analysis to test for differences in mean heart rate **before** drug treatment among the three different age groups.

6) Write down the hypotheses, report the p-value for this test from R, and draw a conclusion about your hypothesis.

Simple Linear Regression

In our movies dataset, we want to uncover the relationship between a movie's runtime and its budget. Ultimately, we want to determine is if knowing a movie's runtime can help us predict its budget. To the right is a scatterplot, along with the line of best fit.



Below is a table with some information on these variables

	Runtime (min)	Budget (mil)
Mean	112.22	73.60
Sd	17.96	66.06
	Correlation (r)	0.374

7) Using R and/or the above table, calculate ***and*** interpret in the context of this problem:

a) The slope
b) The intercept

8) Does the slope make intuitive sense in the context of this problem? Is the intercept meaningful in this scenario?

9) Based on your regression equation, what is the predicted budget of a movie if it has a runtime of 120 minutes?

10) Calculate and interpret the R-squared value.

Submission

Record your answers in a separate file and save the file as a pdf. A photo or scan of a handwritten document is fine, but the final submission must be a pdf. Upload the pdf to canvas.