1. (15 points) Groundhog's Day is celebrated every February 1 and inspired a classic Bill Murray film. A groundhog (named Punxsutawney Phil) emerges from his winter burrow and if he "sees his shadow", supposedly there will be six more weeks of winter. Using data from 1898 - 2016, we want to examine whether the weather in March is actually different (on average) in years when Phil sees his shadow. The following sample statistics describe the results:

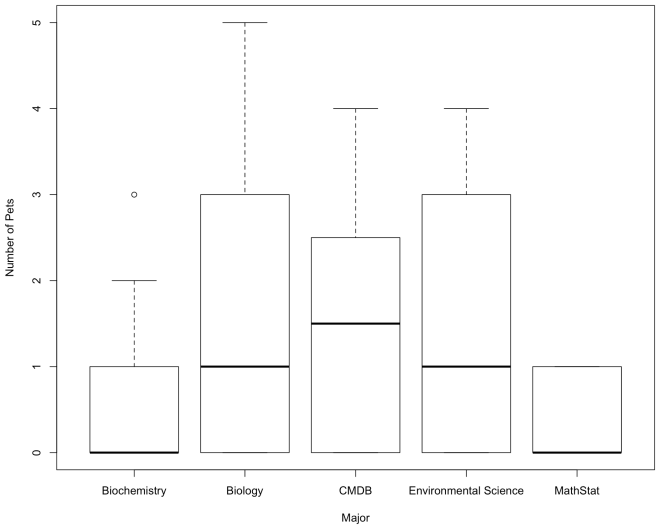
	n	Mean	Std Dev
Shadow	100	41.67°F	2.94
No Shadow	15	42.97°F	2.92

   (a) (10 points) Test at the $\alpha = 0.05$ level of significance if there is a difference in mean March temperature between years when Phil sees his shadow and years when Phil sees no shadow. Assume that any assumptions are satisfied.

   (b) (5 points) If we want to be able to detect a difference in temperature of 2°F (at the $\alpha = 0.05$ level), the power of this test is 65.33%. Explain what this means. How do we improve statistical power?

2. (15 points) After cleaning the start-of-term student survey data, the five most represented majors were biochemistry, biology, environmental science, mathematics/statistics, and CMDB. (Note: CMDB was not grouped with biology because it was so well represented!) We will examine whether a student's number of pets varies based on major. The following information may be useful:

Major	n	Mean	Std Dev
Biochem	17	0.7059	0.9852
Biology	38	1.4737	1.4656
CMDB	8	1.5000	1.5119
Enviro Sci	13	1.4615	1.4500
Math/Stats	15	0.4000	0.5071



(a) (4 points) What are appropriate hypotheses for this test? You may use words or statistical notation.

(b) (5 points) What are the ANOVA assumptions? Are they satisfied?

(c) (2 points) Assume the ANOVA assumptions are satisfied. Use the following `R` output to test your hypotheses from part (a) at the $\alpha = 0.05$ level of significance.

```
                 Df Sum Sq Mean Sq F value Pr(>F)
survey$major      4  18.01   4.503   2.769 0.0323 *
Residuals        86 139.83   1.626
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

(d) (4 points) The following matrix gives the unadjusted p-values for comparisons for the differences in pets by major for the ANOVA in part (c). Use the Bonferonni approach to test which (if any) pairs differ significantly. Use $\alpha = 0.05$.

	Biochemistry	Biology	CMDB	Environmental Science
Biology	0.420	-	-	-
CMDB	0.150	0.958	-	-
Environmental Science	0.111	0.976	0.947	-
Math/Stat	0.500	0.007	0.052	0.031

3. (10 points) A random sample of 7 observations was selected from each of five populations (i.e., there are 7 observations in each of the five groups). A portion of the ANOVA table is shown below. Complete the missing entries.

Source	Degrees of Freedom	Sum of Squares	Mean Square	F
Group	_____	_____	370	_____
Error	_____	_____	_____	
Total	_____	1850		