

PROJECT 1 - DATA

Statistics 401: Fall 2016

Due 5pm on Thursday, September 8

Re-read the Syllabus regarding how to format your report. Write up your report for this assignment accordingly. The write-up, not including the Appendix that contains your R-code and R-output, should not exceed two pages. Your grade will be determined by how well you answer the questions and by the organization and clarity of your write-up.

1. Read the article *Monitoring the Health of Canine Heroes* available through the STAT401 web site.
 - (a) Give the five steps of the scientific method.
 - (b) For each of the five steps of the scientific method, indicate how that step was carried out by the scientists described in the article.
 - (c) Who or what are the individuals of interest?
 - (d) The article describes an explanatory variable and a response variable. What is the explanatory variable? What is the response variable?
 - (e) Classify each variable from 1d as either categorical or numerical. For each categorical variable (if there are any), give the number of categories. For each numerical variable (if there are any), classify it as either discrete or continuous. Summarize your results in a Table, with three columns: Variable, Type and Properties. Label this Table and reference it from the body of your report.
 - (f) Identify the population of interest and specify whether it is a finite or an infinite population.
 - (g) Identify the sampling plan used to select the individuals (into both the "treatment" and "control" groups). What types of bias could be present in the sample? In terms of this problem, explain each type of bias that you have identified.
 - (h) Does the article report on the results of an observational study or an experiment? Explain.
2. Table 1 displays the dimensions (in mm) of jellyfish from two samples taken from the Hawkesbury River in New South Wales, Australia. One of the samples is from Dangar Island and the other is from Salamander Bay. The investigator wants to know to what degree one can determine the location of the jellyfish given its dimensions.

Table 1: Dimension of Jellyfish

Dangar Island		Salamander Bay	
Breadth	Length	Breadth	Length
6.5	8.0	12.0	14.0
6.0	9.0	15.0	16.0
6.5	9.0	14.0	16.5
7.0	9.0	13.0	17.0
8.0	9.5	15.0	17.0
7.0	10.0	15.0	18.0
8.0	10.0	15.0	18.0
8.0	10.0	16.0	18.0
7.0	11.0	14.0	19.0
8.0	11.0	15.0	19.0
9.0	11.0	16.0	19.0
10.0	13.0	16.5	19.0
11.0	13.0	18.0	19.0
12.0	13.0	18.0	19.0
11.0	14.0	16.0	20.0
11.0	14.0	16.0	20.0
13.0	14.0	17.0	20.0
14.0	16.0	18.0	20.0
15.0	16.0	19.0	20.0
15.0	16.0	15.0	21.0
15.0	19.0	16.0	21.0
16.0	16.0	21.0	21.0
		19.0	22.0
		20.0	22.0

- (a) Create a table, with three columns: Variable, Type and Properties. In this table, list the variables in this study. Classify each variable as either categorical or numerical. Give the number of categories and list the categories of each categorical variable (if there are any). Classify each numerical variable as either discrete or continuous (if there are any). If a variable is continuous but rounded, then classify it as continuous. Label the Table and reference it from the body of your report.
 - (b) Of the variables listed in 2a, which are response variables and which are explanatory variables?
 - (c) A data file which contains the data from Table 1 is at the STAT401 web site. Read the data file into a data frame in R. In an Appendix in your report, include the commands that (a) make it possible to refer to the variables by their names alone, rather than as components of the data frame, and (b) display the contents of the data frame. Include this R-output in the Appendix of your report as well. Remember, you'll need to click on tab **File** → **(Change dir ...)** if the data file does not reside in the working directory.
3. Use R to find the mean and standard deviation of the Length of the jellyfish from Dangar Island and from Salamander Bay. Do these values suggest that there may be a true difference in the mean length of jellyfish from these two locations? In an Appendix, include the R-code and R-output you used to answer this question.