

Homework 1, Due Thursday September 10th, 6:00 PM, 10 Points

Directions: In this homework you'll work mostly with JMP to perform some exploratory data analysis. We will check for completion and a few problems will be graded in detail. *It is possible some parts are not graded at all.*

JMP provides a lot of useful tutorials at https://www.jmp.com/en_us/learning-library.html. Watch the following videos *before* you attempt the questions. *They are all very short.* Follow along with the videos by pausing and reproducing the results shown. You will not be graded on whether you watched them or not. The purpose here is to get an introduction to JMP.

1. Go to: https://www.jmp.com/en_us/learning-library/using-jmp.html, scroll down, and watch the short videos on:
 - a. "Opening JMP and Getting Started"
 - b. "Importing Text Files"
2. Go to: https://www.jmp.com/en_us/learning-library/graphical-displays-and-summaries.html, scroll down, and watch, the short videos on:
 - a. "Histograms, Descriptive Stats and Stem and Leaf"
 - b. "Box Plots"
 - c. "Scatter Plots"

Problem 1. This is the only problem you will turn in.

Download the data set "advertising.csv" from Canvas and open it in JMP. For the data description, look at slides 4-6 of Lecture 1 on Tuesday 9/1/2020. *Note that the data you are working with is different than the previous semesters' data.* Keep in mind "sales" is our response variable. Answer the following questions:

- a) Use JMP to fill in the table below. Round your answers to 1 decimal place.

Variable	Minimum	First Quartile	Mean	Median	Third Quartile	Maximum	Standard Deviation
TV							
radio							
newspaper							
sales							

- b) Produce a histogram and a boxplot of the variable sales and turn it in.
- c) Using the results from part b and c, describe the distribution of sales. Is it right skewed? Is it uniform? Anything unusual? Etc. Don't spend too much time on this; Just have a quick look and then go at it!
- d) The histogram of sales should show that it is *not* normally distributed. Suppose we want to create a simple regression model to predict sales from TV. Does the non-normality of sales violate any of the assumptions of the simple regression model? Answer yes/no and explain *very briefly*.
- e) Produce three separate scatterplots with sales on the Y-axis versus the three predictors on the X-axis. Turn them in.
- f) Using the scatter plots in the previous part, state whether there could be a relationship between sales and each of the three predictors. Report the correlation to 2 decimal places for each plot.
You'll have to dig around JMP to find out how to get the correlations. You can do it!