STAT301   Homework 1   Name:

1. An experiment is designed to determine how speaker size affects loudness. The researcher measures loudness of 18 speakers randomly: 6 speakers have small diameters, 6 have medium diameters, and 6 have larger diameters. The researcher records the loudness (in decibels) of the speakers playing the same sound in the table below, and compared the three types of speakers.

Small diameter speaker loudness	15	21	16	14	20	22
Medium diameter speaker loudness	32	41	46	39	52	43
Large diameter speaker loudness	52	50	61	69	43	58

a) Identify the independent variable (IV)

Speaker loudness

b) Identify whether the independent variable is categorical or numeric.

Categorical

c) Identify the dependent variable (DV)

loudness

d) Identify whether the DV is categorical or numeric

numeric

e) What can be the confounding variables? (list **one** possible variable)

Microphone distance

f) Enter the data in R as a data frame **with one dependent and one independent variable** (submit your R Markdown file to d2l drop box)

g) Use pipeline to calculate the mean and standard deviation of loudness for each type of speaker. **Listed** the mean and standard deviation for each speaker here:

	Mean	Standard deviation
large	55.50	9.14
medium	42.16	6.74
small	18	3.41

2. Dr. Optimist believes he has discovered a "smarts" pill will increase IQ. He first measures the IQ of 20 volunteered college juniors and then randomly assigns them to one of the two groups: 10 of them take the smarts pill and 10 take a placebo (sugar pill). He then measures the IQ of all the volunteers. He would like to compare the mean IQ between the pill group and placebo group.

Smart pill IQ	81	98	117	76	91	110	103	105	72	121
Placebo IQ	79	80	89	94	99	106	111	109	73	119

a) Identify the independent variable and its levels

Pill type, two levels Smart, and placebo

b) Is the independent variable categorical or numeric?

Categorical

c) Identify the dependent variable

IQ

d) Is the dependent variable categorical or numeric?

numeric

e) Enter the data in R as a data frame. Submit your R Markdown file to d2l drop box

f) Generate a graph to visualize the means of IQ from the smart pill group and the placebo group. Submit your R Markdown file. (Do not hand-draw the graph here)

3. Listed below are the lead concentrations (in ug/g) measured in different Ayurveda medicines. Ayurveda is a traditional medical system commonly used in India. The lead concentrations listed here are from medicines manufactured in the United States. The data below are based on the article "Lead, Mercury, and Arsenic in US and Indian Manufactured Ayurvedic Medicines Sold via the Internet," by Saper, et al., *Journal of the American Medical Association*.

   3.0   6.5   6.0   5.5   20.5   7.5   12.0   20.5   11.5   17.5

   What is the mean, standard deviation, range and median of this sample data? List them here.

   mean: 11.01
   standard deviation: 6.46
   RANGE: 3, 20.5
   Median: 9.5

4. The data Orange in R recorded the age and circumferences of 5 different types of orange trees. The variable "Tree" recorded the five types as 1, 2, 3, 4, and 5. The variable "age" recorded the ages of the trees and the variable "circumference" recorded the circumferences of the trees. Use R pipeline to

   a) Keep only tree 1, 2 and 3.

   b) Find the mean and standard deviation of circumferences of each type of the trees for tree type 1, 2 and 3. Submit your R Markdown file here.

   c) Generate boxplots for tree type 1, 2 and 3. Submit your R Markdown file.

   Note: all the answers should be on your R Markdown file for question 4.