

Carleton University

School of Mathematics and Statistics

STAT 2606 – Assignment #1

Section A due Wednesday, October 5, 2011 before 4:05pm

Section B due Monday, October 3, 2011 before 10:05am

Section C due Monday, October 3, 2011 before 4:05pm

Section D due Tuesday, October 4, 2011 before 6:35pm

INSTRUCTIONS:

I. For written questions, show all of your work. No credit will be given for answers without justification. Do not use MINITAB for a question unless it specifically says to do so.

II. There will be a 10% deduction for assignments handed in late but within 24 hours after the due date. There will be a 20% deduction for assignments handed in between 24 and 48 hours after the due date. No assignments are accepted more than 48 hours after the due date. If your professor is not available to accept your assignment, you must have the undergraduate secretary (HP 4302) or other administrative personnel date and time your assignment. Students who fail to do this will receive a zero on the assignment. You must then deposit your assignment in your instructor's mailbox in the mailroom (HP 4311).

Part A: Lab Part

- I. Consider the following 50 measurements for questions 1,2,3: (**File -> Open Worksheet**...select Asst1data1.txt. The data appear in column C1.)

3.1 4.5 2.9 2.7 3.8 5.1 2.5 3.6 4.3 5.6
4.9 3.5 2.1 4 6.2 1.8 3.6 6.1 5.7 4.9
2.8 3.7 3.5 4.4 2.5 5.6 5.1 4.7 3.7 4.2
3.6 4.1 4 3.7 2.9 2.2 4.8 3.9 4.6 3.1
2.5 4.9 3.7 4.2 2.8 3.4 1.6 3.9 4 3.9

1. Construct a histogram for these 50 measurements.(Click **Graph->Histogram...**,select **Simple**)
 - a. Describe the shape of the histogram. Do you see any outliers?
 - b. For the class with a midpoint of 5, specify (1) the class width, (2) the lower class boundary, and (3) the upper class boundary.
 - c. What percentage of measurements are less than 5.25?

2. Construct a stem-and-leaf plot for these 50 measurements.(Click **Graph->Stem and leaf**)

- a. Describe the shape of the distribution. Do you see any outliers?
 - b. Use the stem and leaf plot to find the smallest observation.
 - c. Find the eighth and ninth largest observation.
3. Calculate the mean, median and standard deviation for these 50 measurements.(**Click->Basic Statistics->Display Describe Statistics...**)

II. The most visited site on the Internet is Yahoo!, which boasted 111,271 thousand unique visitors in December 2003. The number of unique visitors at the top 25 sites are shown in the following table: (Consider the following 25 measurements for questions 4,5,6.). (To open the data: Click **File -> Open Worksheet...**select Asst1data2.txt. The data appear in column C2.)

Site	Visitors(millions)	Site	Visitors(millions)
Yahoo	111.3	Weather Channel	23.8
Time Warner	110.5	Real.com	22.3
MSN-Microsoft	110	Verizon	22.1
Ebay	69.2	Wal-Mart	21.4
Google	61.5	Shopping.com	21.3
Terra Lycos	52.1	Symantec	19.9
Amazon	45.7	AT&T	17.5
About/Primedia	42.6	InfoSpace	17.3
Excite	25.1	Monster	17.3
CNET	25.1	EA Online	16.8
Disney Internet	25.1	SBC Communications	16.5
Viacom	24.7	Sony Online	16.5
American Greetings	24.4	Weather Channel	23.8

4. Calculate the mean, median and standard deviation for these 25 measurements.(**Click->Basic Statistics->Display Describe Statistics...**). Use these measures to decide whether or not the data are symmetric or skewed.
5. Construct a histogram for these 25 measurements.(Click **Graph->Histogram...**,select **Simple**)
- a. Describe the shape of the distribution.
 - b. Are there any outliers in the set? If so, what sites boast an unusually large number of unique visitors?
6. Draw a box plot for these 25 measurements.(Click **Graph->Box plot...**)
- a. Describe the shape of the distribution.
 - b. Are there any outliers in the set? If so, what sites boast an unusually large number of unique visitors?

Part B: Written Questions

1. Identify the following variables as either categorical (qualitative) or quantitative. If a variable is categorical, specify if it would be measured on a nominal or ordinal (ranked) scale. If a variable is quantitative, specify if it would be measured on an interval or ratio scale.
 - a. The brands of ice cream that you purchase
 - b. The daily high temperature(in degrees Celsius) for the last four weeks
 - c. The amount of sugar consumed by Canadians in one year
 - d. The species of fish in the zoo
 - e. The lengths of time children wait for the school bus
 - f. Your favourite professional football team
 - g. Rating of the effectiveness of a new cold remedy (Not effective, effective).
2. For each of the following situations describe the population of interest and the sample.
 - a. A hospital administration would like to know the average length that a patient stays in the hospital. It would be almost impossible to look through all of the past records and average the lengths of stay. Instead, a sample of 1000 patients over the last year was randomly chosen and their lengths of stay were averaged.
 - b. The freshness and overall quality of milk depend upon the type of packaging used. The manager of a dairy company is considering changing the packaging from cartons to plastic. The quality control team of this dairy company packaged milk in 100 containers of each type of material. After a specific amount of time, the team tested the milk for freshness and overall quality.

3. The numbers of rooms for 15 home recently sold were;

8, 8, 8, 5, 9, 8, 7, 6, 6, 7, 7, 7, 7, 9, 9

- a. Compute 10th, 65th, and 90th percentiles.
 - b. Compute Q_1 , Q_2 , Q_3 and IQR.
 - c. Compute mean, median, mode.
 - d. Compute range, variance, standard deviation.
4. The data listed here are the weights(in kilograms) of 27 packages of ground beef in a supermarket meat display:

1.08 .99 .97 1.18 1.41 1.28 .83 1.06 1.14
1.38 .75 .96 1.08 .87 .89 .89 .96 1.12
1.12 .93 1.24 .89 .98 1.14 .92 1.18 1.17

- a. Draw a stem-and-leaf plot for the data with stem unit 0.1 and a leaf unit = 0.01.
- b. Find the mean and the median.

- c. Describe the shape of the distribution (is it symmetric, skewed left, skewed right, bell-shaped, etc., are there obvious outliers).
- d. Find the percentage of measurements falling in the intervals $\bar{x} \pm ks$ for $k = 1, 2, 3$.
- e. How do the percentages obtained in part d compare with those given by the empirical Rule? Explain.
- f. The two largest packages of meat weigh 1.38 and 1.41 kilograms. Are these two packages unusually heavy?