

STATISTICS 101 - Homework 1 Answers
Due Monday, May 23, 2005

This assignment is worth a total of 45 points.

1. (12 pts) The birth weight (grams) for each of 44 babies born at a Brisbane, Australia hospital are summarized in the frequency table below.

Birth Weight (g), X	Number of babies
$1500 \leq X < 1750$	1
$1750 \leq X < 2000$	0
$2000 \leq X < 2250$	3
$2250 \leq X < 2500$	1
$2500 \leq X < 2750$	2
$2750 \leq X < 3000$	2
$3000 \leq X < 3250$	5
$3250 \leq X < 3500$	13
$3500 \leq X < 3750$	11
$3750 \leq X < 4000$	5
$4000 \leq X < 4250$	1
$4250 \leq X < 4500$	0

- (a) (6 pts) Construct an appropriately labeled histogram from the frequency table above.

- (b) (2 pts) What percentage of babies have weights less than 2500 grams (5.5 pounds)? What percentage of babies have weights greater than or equal to 4000 grams (8.8 pounds)?
5 out of 44 babies (or 11.4%) have weights less than 2500 grams. 1 out of 44 babies (2.3%) have weights greater than or equal to 4000 grams.
- (c) (4 pts) Describe the distribution of birth weights. Be sure to discuss the center, the spread and the shape of the distribution.

The distribution of birth weights is skewed to the left, with a center between 3000 and 3250 grams and a spread from 1500 grams to 4250 grams.

2. (14 pts) A study was conducted regarding blood cholesterol levels of 28 heart-attack patients and 24 people who had not had a heart attack. Below are the data.

Heart Attack							No Heart Attack					
270	236	210	142	234	360	288	196	232	200	242	212	162
280	272	160	220	224	310	288	206	178	184	198	200	164
226	242	186	266	276	280	244	160	182	182	198	182	176
206	318	294	282	282	278	236	166	204	182	178	186	170

- (a) (8 pts) Construct a back-to-back stem-and-leaf display and compare the two groups in terms of cholesterol levels.

Heart Attack			No Heart Attack						
		2	14						
			15						
		0	16	0	2	4	6		
			17	0	6	8	8		
		6	18	2	2	2	2	4	6
			19	6	8	8			
		6	20	0	0	4	6		
		0	21	2					
	6	4	0	22					
	6	6	4	23	2				
		2	4	24	2				
			25						
		6	26						
	8	6	2	0	27				
8	8	2	2	0	0	28			
				4	29				
				30					
		8	0	31					
				32					
				33					
				34					
				35					
		0	36						

The level of cholesterol in the heart attack group is much higher on average than the level of cholesterol in the no heart attack group. The no heart attack group also has a much smaller spread of cholesterol levels than the heart attack group.

- (b) (6 pts) Compute the median cholesterol level for each group. Which group has the higher median cholesterol level?

For the heart attack group, the median cholesterol level is 268. For the no heart attack group, the median cholesterol level is 183. As the stem-and-leaf display shows, the mean level of cholesterol is much higher for the heart attack group than for the no heart attack group.

3. (19 pts) **JMP Assignment** The total annual rainfall (in inches) for the past 100 years (from 1902-2002) for Los Angeles, California is given in the table below.

4.4	17.8	11.6	9.1	31.0	12.4	12.4	24.4	8.1	27.4
21.0	12.0	7.4	8.1	12.5	7.7	17.9	12.8	10.4	31.3
10.7	9.0	27.0	19.7	33.4	12.3	7.2	14.4	14.9	21.3
7.2	12.3	7.8	27.5	16.6	22.0	20.4	13.7	8.0	8.4
18.8	4.9	8.2	5.6	21.1	9.5	16.0	11.9	12.0	9.5
26.2	8.2	10.6	8.0	7.2	12.3	11.7	11.6	19.2	18.2
11.2	32.8	19.2	13.1	23.4	22.4	13.5	21.7	14.9	11.9
16.9	12.5	11.5	12.7	9.8	18.0	17.6	7.9	6.7	9.6
19.7	13.7	12.5	8.6	13.9	15.3	19.9	17.1	23.7	13.4
11.6	16.2	12.6	19.2	11.7	19.3	18.7	19.5	8.7	19.3

Follow these steps to analyze the data using JMP.

- Start Netscape and go to the course home page at

<http://www.stat.iastate.edu/stat101/spring2004/homepage.html>

Under the **Data Sets** heading, click on the link **Rainfall in Los Angeles**. From the Netscape menu, choose File - Save As. Name this file **rainfall.txt** and save it to the computer's hard drive or to a disk.

- Start JMP 5.1 by either double-clicking on the **JMP 5.1 icon** or by choosing **Start - Programs - JMP 5.1 - JMP 5.1**. In the JMP Starter window, click on the **Open Data Table** button. Under the **Files of Type:** change it to **Text Import Preview**. Then locate the file **rainfall.txt** on the computer's hard drive or on a disk. Select this file and click **Open** and then **Delimited**. In the window that appears, put a check mark in the box near **Space** in the **End of Field Box**. Put a **check mark** in the box near **Table contains column headers**. Click on **Apply Settings**. At this point, JMP gives you a preview of the column names and the first two rows of your data. If everything looks good, press **OK**.
- We want to describe the distribution of this data. Select **Analyze** → **Distribution** from the JMP menu. Select the column **rainfall** and click the button **Y, Columns**. Then click on **OK**.
- You should now have a histogram, boxplot, and statistics for this data set. We want to make a few changes to the information JMP has calculated. First, click on the **red triangle** next to **rainfall** and select **Stem and Leaf**. This should add a stem-and-leaf plot to your window. Now, click on the **red triangle** next to **rainfall** and select **Histogram Options** → **Count Axis**. This should add a count axis to the histogram in the window. Again, click on the **red triangle** next to **rainfall** and select **Display Options** → **Horizontal Layout**.
- From the JMP Menu, select **File** → **Print** to print your output. Turn this paper in with your assignment. Using this output, answer the following questions on a separate piece of paper.

The JMP output is worth 5 pts.

- (a) (2 pts) What percent of the 100 years have rainfall amounts less than 10 inches? What percent of the 100 years have rainfall amounts greater than or equal to 30 inches?

There are $(14 + 6 + 2) = 22$ years with rainfall amounts less than 10 inches. This is 22%. There are $(2 + 2) = 4$ years with rainfall amounts greater than or equal to 30 inches. This is 4%.

- (b) (5 pts) Using the histogram, describe the distribution of yearly rainfall in Los Angeles for the past 100 years. Make sure to describe the shape, center and spread of the distribution and make note of any outliers or unusual characteristics.

The distribution of yearly rainfall in Los Angeles for the past 100 years is bimodal with the main mode around 10 inches and a secondary mode between 17.5 and 20 inches. The shape is skewed to the right, with the center around 13 inches and a spread from 4 inches to 33 inches of rain. There are no outliers.

- (c) (2 pts) How is the observation of 12 inches of rain recorded in the Stem and Leaf display?
The observation of 12 inches of rain is recorded in the Stem and Leaf display as 1—2.
- (d) (3 pts) Does the Stem and Leaf display use split stems? If yes, how did JMP choose to split the stems?
Yes, the stems in the Stem and Leaf display are split. JMP split them into 5 stems, 0 and 1, 2 and 3, 4 and 5, 6 and 7, and 8 and 9.
- (e) (2 pts) Would the distribution of yearly rainfall in Los Angeles over the past 100 years give you any information about the distribution of yearly rainfall in Des Moines over the past 100 years? Explain your answer.
No, the weather in Des Moines is much different than the weather in Los Angeles. The distribution of rainfall in Los Angeles would not provide any information about the distribution of rainfall in Des Moines.