

MATHEMATICS 201-510-LW

Business Statistics

Martin Huard

Fall 2008

Assignment #1 SOLUTIONS

This assignment is due on **Thursday September 4, 2008** at the beginning of class. Complete solutions are expected.

Question 1 (2 points)

When meeting new clients, a banker will usually gather a certain amount of information on her client. Some of that information follows:

- a) Annual Salary **Ratio**
- b) Type of Investment (High Risk, Average Risk, Low Risk) **Ordinal**

For the information in parts (a) and (b) list the level of measurement (nominal, ordinal, interval or ratio).

Question 2 (2 points)

Identify each of the following samples by naming the sampling technique used (cluster, convenience, simple random, stratified, systematic).

- a) Ask every thirtieth client entering a bank is asked to complete a survey.
Systematic
- b) Divide the clients of a bank according to their type of investment (high risk, average risk and low risk) and then select a random sample from each type of investment to survey the clients about the quality of services being offered.
Stratified

Question 3 (15 points)

Sarah wants to determine the mean number of hours per week a student at St. Lawrence study. To accomplish this, she passed a questionnaire to 10 students chosen at random from their registration number. The results (in hours) were:

4 12 7 10 15 5 2 19 15 18

- a) What is the population?
St. Lawrence students
- b) What is the variable?
Mean number of hours studied per week
- c) Is the variable qualitative or quantitative?
Quantitative
- d) Find the mean.

$$\bar{x} = \frac{\sum x}{n} = \frac{107}{10} = 10.7 \text{ hours}$$

- e) Find the median.

$$Me = \frac{10+12}{2} = 11 \text{ hours}$$

- f) Find the mode.

$$Mo = 15 \text{ hours}$$

- g) Find the 12% trimmed mean

12% of 10 is 1.2, thus max and min are not included

$$12\% \text{ trimmed mean} = \frac{\sum x}{n} = \frac{86}{8} = 10.75 \text{ hours}$$

- h) Find the range

$$\text{Range} = \text{Max} - \text{Min} = 19 - 2 = 17 \text{ hours}$$

- i) Find the variance

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} = \frac{1473 - \frac{107^2}{10}}{9} = 36.46 \text{ hours}^2$$

- j) Find the standard deviation

$$s = \sqrt{s^2} = \sqrt{36.46} = 6.04 \text{ hours}$$

- k) Find the coefficient of variation

$$CV = \frac{s}{\bar{x}} = \frac{6.04}{10.7} = 0.564 = 56.4\%$$

- l) Find
- Q_1

$$Q_1 = 5 \text{ hours}$$

- m) Find
- Q_3

$$Q_3 = 15 \text{ hours}$$

- n) Find the interquartile range

$$IQR = Q_3 - Q_1 = 10 \text{ hours}$$

- o) Draw a box and whisker plot.

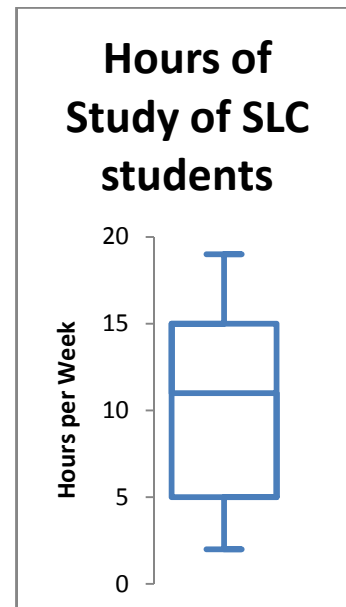
$$\text{Min} = 2 \text{ hours}$$

$$Q_1 = 5 \text{ hours}$$

$$Me = 11 \text{ hours}$$

$$Q_3 = 15 \text{ hours}$$

$$\text{Max} = 19 \text{ hours}$$



Question 4 (4 points)

The following frequency distribution shows the amount of money spent at a large retail store for a random sample of 200 customers who bought something.

Amount Spent (in \$)	Frequency	Midpoint x	xf	x^2f
0 – 49	28	24.5	686	16807
50 – 99	62	74.5	4619	344115.5
100 – 149	52	124.5	6474	806013
150 – 199	33	174.5	5758.5	1004858
200 – 249	17	224.5	3816.5	856804.3
250 – 299	8	274.5	2196	602802
<i>Total</i>	<i>200</i>		<i>23550</i>	<i>3631400</i>

- a) Use this frequency distribution to estimate the mean amount of money spent.

$$\bar{x} = \frac{\sum xf}{n} = \frac{23550}{200} = 117.75 \$$$

- b) Use this frequency distribution to estimate the standard deviation for the amount of money spent.

$$s^2 = \frac{\sum x^2 f - \frac{(\sum xf)^2}{n}}{n-1} = \frac{3631400 - \frac{23550^2}{200}}{199} = 4313.5$$

$$s = \sqrt{s^2} = 65.68 \$$$

Question 5 (10 points)

A supermarket chain wishes to see if there is correlation between the annual income of a family and the weekly grocery bill. For this, a random sample of 8 people was taken, where each gave their annual income and average weekly grocery bill. Here are the results.

Annual income (in 1000 \$)	Weekly grocery bill (in \$)	x^2	xy	y^2
51	123	2601	6273	15129
74	135	5476	9990	18225
39	89	1521	3471	7921
45	120	2025	5400	14400
41	122	1681	5002	14884
105	117	11025	12285	13689
37	105	1369	3885	11025
21	95	441	1995	9025
<i>413</i>	<i>906</i>	<i>26139</i>	<i>48301</i>	<i>104298</i>

- a) Find the equation of the least-squares line.

$$SS_x = \sum x^2 - \frac{(\sum x)^2}{n} = 26139 - \frac{413^2}{8} = 4817.9$$

$$\bar{x} = \frac{\sum x}{n} = \frac{413}{8} = 51.6$$

$$SS_y = \sum y^2 - \frac{(\sum y)^2}{n} = 104298 - \frac{906^2}{8} = 1693.5$$

$$\bar{y} = \frac{\sum y}{n} = \frac{906}{8} = 113.2$$

$$S_{xy} = \sum xy - \frac{\sum x \sum y}{n} = 48301 - \frac{(413)(906)}{8} = 1528.8$$

$$\text{Slope : } b = \frac{SS_{xy}}{SS_x} = \frac{1528.8}{4817.9} = 0.317$$

$$y\text{-intercept : } a = \bar{y} - b\bar{x} = 113.2 - 0.317 \cdot 51.6 = 96.9$$

Thus the least-squares line is given by $y = 96.9 + 0.317x$.

- b) Find the coefficient of correlation.

$$r = \frac{SS_{xy}}{\sqrt{SS_x SS_y}} = \frac{1528.75}{\sqrt{4817.9 \cdot 1693.5}} = 0.535$$

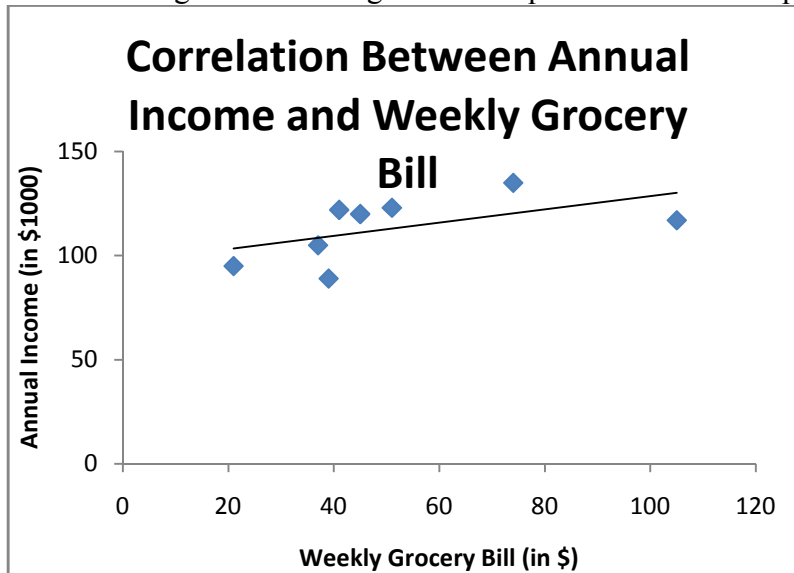
- c) Find the coefficient of determination.

$$r^2 = 0.535^2 = 0.286 = 28.6\%$$

- d) If a family has an annual income of \$60 000, what is the expected weekly grocery bill?


$$\hat{y} = 96.9 + 0.317(60) = 115.91\$$$

- e) Sketch a scatter diagram containing the least-squares line and its equation.



Questions 6 to 8 are to be done using Excel.

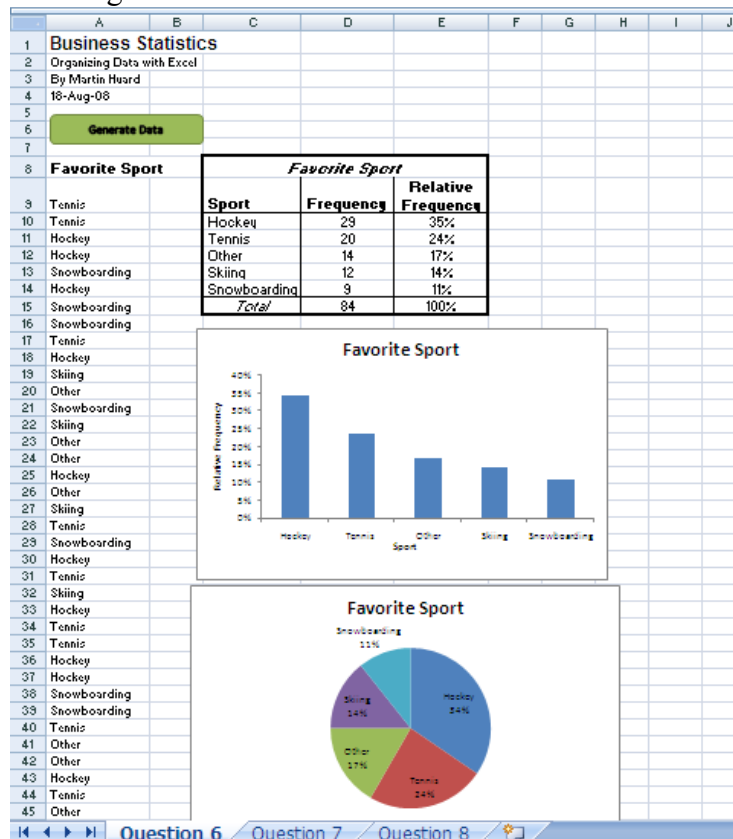
For these questions, hand-in the printouts of your Excel sheets and copy your Excel work in the Test folder for Business Statistics (W:\Tests\mhuard\Business Statistics\Assignment 1), where your name should be included in the name of the file (for example: Assignment 1 – Your Name). Make sure that your answers are well organized with appropriate labels, and rounded off to an appropriate number of decimal places.

Open the file “Data – Assignment 1” from my web site, and save it under “Assignment 1 – Your Name”. Note that you may have to enable macros to be able to generate the data. If the macros are not enabled (that is, if the data does not appear at the click of the button) then go to  - EXCEL OPTIONS – TRUST CENTER – TRUST CENTER SETTINGS – MACRO SETTINGS and choose the ENABLE ALL MACROS option. Note that you may need to close your document and open it again.

Question 6 (8 points)

A sample of customers at a Sports Retail Store was taken, where each indicated their favorite sport.

- a) Go to the worksheet “Sheet1”, rename it appropriately, make the usual heading in cells A1:A4, then click on the “GENERATE DATA” button to get your data.
- b) Construct a relative frequency distribution table for the favorite subject of students (Hockey, Skiing, Snowboarding, Tennis, Other). (Note: Excel must do all of the counting!)
- c) Sketch a Pareto Diagram.
- d) Sketch a Pie Diagram.



Question 7 (12 points)

In a study on the use of debit cards, a bank manager took a random sample of transactions made with the use of a debit card, and recorded the amount of the transaction (in \$).

- a) Go to the worksheet “Sheet2”, rename it appropriately, make the usual heading in cells A1:A4, then click on the “GENERATE DATA” button to get your data.
- b) Find the mean.
- c) Find the median
- d) Find the standard deviation.
- e) Find the third quartile.
- f) Find the 47th percentile.
- g) Find the range.
- h) Find the 12% trimmed mean.
- i) Make a relative frequency distribution table. (Note: Excel must do all of the counting!)
- j) Sketch a histogram.

Transaction (in \$)	Measures for the Transactions		Distribution of the Amount for the Transactions			
			Amount (in \$)	Frequency	Relative Frequency	Class Midpoint
81.05	Mean	\$ 93.25	0.00 to 19.99	4	4%	9.995
55.92	Median	\$ 90.84	20.00 to 39.99	8	9%	29.995
43.85	StDev	\$ 45.82	40.00 to 59.99	12	13%	49.995
40.4	Q ₃	\$ 123.59	60.00 to 79.99	12	13%	69.995
176.05	P ₁₇	\$ 89.95	80.00 to 99.99	14	15%	89.995
183.28	Range	\$ 206.67	100.00 to 119.99	16	18%	109.995
8.72	12% trimmed mean	\$ 91.87	120.00 to 139.99	9	10%	129.995
29.15			140.00 to 159.99	9	10%	149.995
104.8	Min =	\$ 1.35	160.00 to 179.99	3	3%	169.995
43.13	Max =	\$ 208.02	180.00 to 199.99	3	3%	189.995
130.73	Range =	\$ 206.67	200.00 to 219.99	1	1%	209.995
89.94	n =	91	<i>Total</i>	91	92%	
54.93	Width =	\$ 21.66				
122.39						
101.44						
125.26						
134.75						
189.97						
151.44						
1.35						
55.92						
126.79						
60.76						
60.87						
72.47						
81.13						
105.33						
34.41						
115.66						
161.49						

Question 8 (7 points)

Sociability can be expressed in a number of different ways, including having a lot of friends and dating frequently. A researcher asked a sample of college students about how many good friends they have and how many dates they have had in the past month.

- Go to the worksheet “Sheet3”, rename it appropriately, make the usual heading in cells A1:A4, then click on the “GENERATE DATA” button to get your data.
- Find the coefficient of correlation.
- Find the coefficient of determination.
- Find the equation of the least squares line.
- If a student has 5 good friends, how many dates is he be expected to have had in the past month?
- Sketch a scatter diagram containing the least-squares line and its equation.

	A	B	C	D	E	F	G	H	I
1	Business Statistics								
2	Organizing Data with Excel								
3	By Martin Huard								
4	18-Aug-08								
5	Generate Data								
7									
8	Number of Friends	Number of Dates		Regression					
9	6	5		slope		0.769			
10	17	19		y - intercept		1.320			
11	5	7		equation of line		$y = 1.320 + 0.769 x$			
12	5	12							
13	12	15		Least squares forecast for someone with 5 friends		5.2 dates			
14	2	1							
15	5	0		Correlation					
16	7	6		Coefficient of correlation r		0.759			
17	15	16		Coefficient of determination r^2		57.6%			
18	16	13							
19	7	8							
20	5	1							
21	15	7							
22	12	7							
23	1	1							
24	1	3							
25	5	2							
26	5	8							
27	1	2							
28	3	2							
29	3	1							
30	2	4							
31	13	9							
32	1	0							
33	12	15							
34	21	17							
35	1	2							

Relationship Between the Number of Friends and the Number of Dates