

Question 1 (8) : 4

Question 2 (8) : 7.5

Question 3 (8) : 8

Question 4 (10) : 10

Question 5 (8) : 8

Question 6 (8) : 8

Total Score = 50

Your Total Score: 45.5

**INSTRUCTIONS:**

1. DO NOT OPEN UNTIL YOU ARE TOLD TO BEGIN.
2. You have from 2:30 to 3:20 PM (50 minutes) to complete the exam.
3. Closed book. However, you may use the course formula (blue sheets) and one two-sided sheet of notes.
4. Write answers in the spaces provided.
5. When requested, show your work for credit.
6. Put a box around all final answers.

**Suggested strategy:**

1. Quickly skim the entire exam.
2. Do all questions you can do easily.
3. Go back and work at the more difficult problems.
4. Allow some time at the end to check answers.

Question 1:

The table below shows a part of the data from B-S 73 Matrimonial Unrest.

FIGHT	Whether or not a fight occurred that day (1 = Fight, 0 = No Fight)
CALORIES	Number of calories consumed that day
SLEEP	Number of hours of sleep the previous night
WORK	Good or bad day at work (0 = Good, 1 = Bad)
SCHOOL	School the next day (0 = No, 1 = Yes)
COFFEE	Number of cups of coffee consumed that day

These variables are observed for the couple (Ted and Maria) for 75 successive days.

DAY	FIGHT	CALORIES	SLEEP	WORK	SCHOOL	COFFEE
1	0	3350	5.5	1	1	2
2	1	3487	7	1	0	1
3	0	4020	4	0	0	0
4	0	2447	9	1	0	2
5	0	1995	7	0	0	2
6	0	1684	9	0	1	0
7	0	3335	6	0	0	3
8	1	2817	7	1	1	1
9	1	2872	7	1	0	1
10	0	2801	6	0	0	2
11	0	2589	6	1	0	1

A. The above data is: Circle the correct one.

- a. Uni-variate cross-sectional with only nominal and ordinal variables
- ☒ b. Multi-variate time-series with both quantitative and qualitative variables
- c. Bi-variate cross-sectional with only quantitative and nominal variables
- d. Multi-variate cross-sectional with quantitative and qualitative variables
- e. Uni-variate time-series with only quantitative and ordinal variable

B. Which of the variables, if any, are nominal? Fight, Work, School

C. Which of the variables, if any, are ordinal? Coffee

D. Which of the variable, if any, are discrete? Day

Write N/A if there is none.

Question 2:

The summaries below are from the same data in Question 1. Here are the definitions and the summaries:

Definitions:

The target variable  $X = \text{CALORIES}$ , number of calories consumed that day.

$X_i = i$ -th largest value of  $X$ , i.e.,  $X_1 = X_{\text{The smallest}}$  and  $X_{75} = X_{\text{The largest}}$

Summaries:

$$X_1 = 1160, \quad X_{75} = 4491, \quad X_{36} = 2900, \quad X_{37} = 2950, \quad X_{38} = 2954, \quad X_{39} = 3031$$

$$\sum_{i=1}^{75} X_i = 224465 \quad \sum_{i=1}^{75} X_i^2 = 711021741, \quad \sum_{i=1}^{75} (X_i - \bar{X})^2 = 39227925$$

A. What is the median of the data? Please show your work.

$$\frac{1+75}{2} = \frac{76}{2} \quad 38 = \text{Median}$$

$$\text{Median} = 2954$$

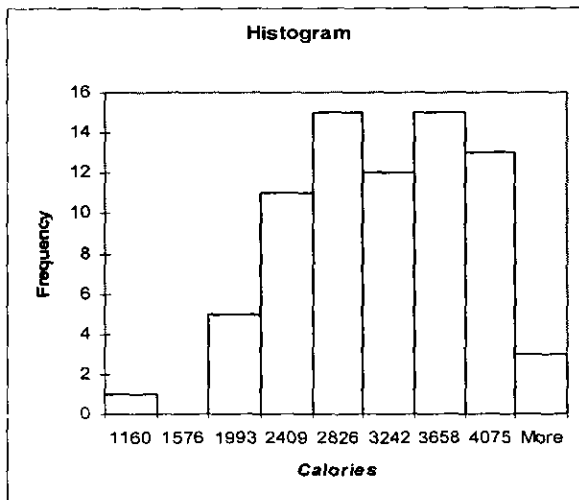
B. What is the average of the data? Please show your work.

$$\sum_{i=1}^{75} X_i = 224465 \quad \frac{224465}{75} = 2992.666 \approx \text{avg.}$$

C. What is the standard deviation of the data? Please show your work.

$$\sqrt{\frac{39227925}{75-1}} \quad \sqrt{530107.094} = 728.095 = \text{SD}$$

D. The histogram is given below. The smallest observation appears to be an outlier. Please compute (1) the average and (2) the median for the data after removing the observation. Please show your work.



$$X_1 = 1160$$

$$\text{Avg} = \frac{(224465 - 1160)}{(75-1)} = 3017.635$$

$$\text{Median} = \frac{1+74}{2} = \frac{75}{2} = 37.5$$

$$\frac{2900 + 2950}{2} = 2925$$

$$= 0.5$$

Question 3:

Joel, Marketing Director of TQM Home Video predicts that the sales for the new video game, Ninja Kim, is 70,000 cartridges on average, with a standard deviation of 15,000 cartridges. The unit cost of the game is \$2.00. There is a fixed cost of \$65,000.

- A. What is the predicted total cost, on average, of the cartridges produced? Please show your work.

$$T_{\text{cost}} = (70,000 \cdot 2) + 65,000$$

$$140,000 + 65,000$$

$$\boxed{\$205,000}$$

- B. What is the uncertainty involved in this forecast of the total cost? Please show your work.

$$SD = 15,000 \cdot 2 = \boxed{\$30,000}$$

- C. The actual unit sales is 62,000 cartridges. How many standard deviations above or below the average is this figure? Please show your work.

$$\cancel{(62,000)(2) + 65,000 = \$189,000}$$

$$\begin{array}{r} 70,000 \\ - 62,000 \\ \hline 8,000 \end{array}$$

$$\frac{8000}{15000}$$

$= .533$  Standard Deviations Below

- D. Suppose that TQM Home Video actually sells 84,000 cartridges. Would this be a surprise in light of the initial forecast? Please explain.

$$\cancel{84,000 \cdot 2}$$

$$\begin{array}{r} 84,000 \\ - 70,000 \\ \hline 14,000 \end{array}$$

$$\boxed{14,000}$$

No, this is not a surprise because the ballpark prediction showed that the SD. would be 15,000 on either side, this was within that range.

$$\frac{14,000}{15,000} = .93 \text{ SD above}$$

Question 4:

A team of QMETH 201 students investigated the pricing strategy of a well known electronics store in Seattle. The team randomly selected 30 TV sets and collected the following data for each TV:

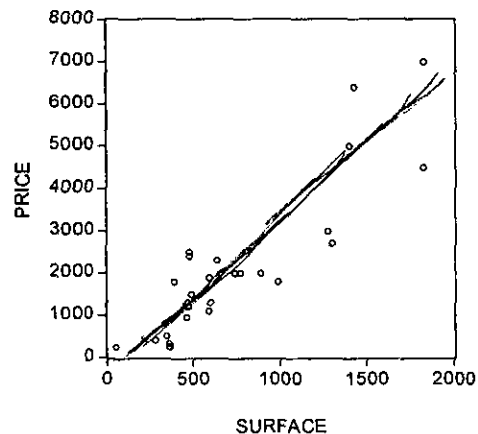
price = price of the TV set in dollars

surface = surface area of the screen in square inches

Below are summaries of data, including a scatterplot.

	X SURFACE	Y PRICE
Mean	704	2065
Median	536	1850
Maximum	1824	7000
Minimum	48	250
Std. Dev.	459	1695
Observations	30	30

Correlation Coefficient = 0.881



- A. Develop the simple regression equation for predicting the price by the surface.

Please show your work.

$$b = r \frac{S_y}{S_x}$$

$$b = .881 \frac{1695}{459}$$

$$b = 3.253$$

$$Y = (-225.112) + (3.253)X$$

$$a = \bar{Y} - b\bar{X} \quad a = 2065 - (3.253)(704)$$

$$2065 - 2290.112$$

$$a = -225.112$$

- B. What is the likely size of the prediction error by using the regression in A? Please show your work.

$$S_e = S_y \sqrt{(1-r^2) \left( \frac{n-1}{n-2} \right)} \quad \text{or} \quad S_e = S_y \sqrt{(1-r^2)}$$

$$1695 \sqrt{[1 - (.881)^2]}$$

$$\frac{1 - .776}{\sqrt{.224}} = \sqrt{(645)(.473)}$$

$$S_e = 801.735$$

- C. Compute the predicted price for the median surface size. Please show your work.

$$\begin{aligned} \text{Median } X &= 536 & Y &= -225.112 + 3.253 X \\ & & &= -225.112 + 3.253(536) \\ & & &= -225.112 + 1743.608 \end{aligned}$$

$$Y = 1518.496$$

- D. What is the proportion of the variability of the price explained by the variability of the surface?

$$r = .881$$

$$r^2 = .7761 = 77.61\%$$

- E. Please study the scatterplot and answer the following questions.

- (1) What is an important feature of the plot, according to your assessment, using the regression equation computed in A?

The plot shows a strong linear relationship, with a high value of  $r^2$  and no outliers.

- (2) Do you observe any problematic feature or pattern? Please explain.

One characteristic of the plot that may prove problematic is the slightly unequal variability. This is the only feature/pattern that may not support the regression equation.

Question 5:

Here is a summary of the monthly return of two stocks A and B and that of the market. (Note: A and B are actual companies and the market index is the Standard and Poor's 500). You want to invest in either stock A or B for a month. For each investment goal below, state whether you should invest in stock A or B, and explain the reason.

	Market	Stock A	Stock B
Average	0.5%	2.4%	1.4%
Standard Deviation	1.7%	5.0%	3.3%
Correlation Coefficient with the Market Return	1.00	0.65	0.34

- A. You want to maximize the expected return.

Stock A  
Stock A has a higher average return.

$$2.4\% > 1.4\%$$

- B. You want to minimize the risk in terms of earning the expected return.

Stock B  
Stock B's returns do not deviate from the average as much as A.

$$3.3\% < 5.0\%$$

- C. You want to minimize the risk in terms of sensitivity to the market fluctuation.

Stock B  
A Correlation Coefficient of .34 means that Stock B is much less sensitive to market fluctuation and therefore, less risky.

$$.34 < .65$$

- D. You want to minimize the proportion of the systematic risk.

$$r^2 = .65^2$$

$$.4225$$

Stock A

$$r^2 = .34^2$$

$$.1156$$

Stock B

A lower value of  $r^2$  means lower proportion of systematic risk.  
Stock B is better in this category.

## Question 6:

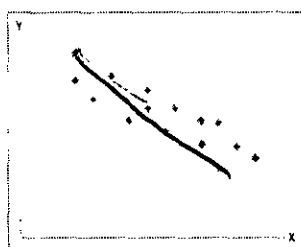
The questions below are unrelated.

- A. \$231,000 Write the answer. You are a sales manager for a regional division of a large mutual fund. The sales goal for your representatives has an average value of \$800,000 and the standard deviation of \$220,000. The VP of Sales orders that you raise the sales goal of each representative by 5%. What is the new standard deviation?

$$SD \uparrow 5\%$$

$$220,000 \times 1.05 = 231,000$$

- B. False True or False: The regression line is computed for the following scatterplot yielding R-square of 0.49. The value of the correlation coefficient is 0.7.



$$\begin{aligned} &(-.7)^2 = .49 \\ &= -.7 \end{aligned}$$

- C. \$29.95 Write the answer. The customers of Best Foods consist of three groups. The weekday shoppers make up 35% and spend \$25 per week, the weekend shoppers make up 50% and spend \$40 per week, and the remaining shoppers are not regular customers and spend \$8 per week. Find the average spending for all shoppers. Please show your work below.

$$\begin{aligned} 35\% &= \$25 \cdot .35 \\ 50\% &= \$40 \cdot .5 \\ 15\% &= \$8 \cdot .15 \end{aligned} = 29.95$$

$$-.0347$$

- D. 7.15 Write the value. The monthly closing prices of Starbucks stock for January to September 2009 are given below. The return of the stock in June 2009 is: Please show your work.

Month	Closing Prices
January 2009	9.44
February 2009	9.15
March 2009	11.11
April 2009	14.46
May 2009	14.39
June 2009	13.89
July 2009	17.7
August 2009	18.99
September 2009	20.65
October 2009	20.94

$$14.39 - 13.89 = .5$$

$$\frac{.5}{14.39} = .0347$$