KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICAL SCIENCES
DHAHRAN, SAUDI ARABIA

STAT 211: BUSINESS STATISTICS I
Semester 042
Major Exam #1
Sunday March 27, 2005

Please **circle** your instructor's name: Marwan Al-Momani
Mohammad F. Saleh
Walid S. Al-Sabah

Name: ID# Section

Question No	Full Marks	Marks Obtained
1	5	
2	5	
3	5	
4	5	
5	5	
6	5	
7	15	
8	5	
Total	50	

Question 1. Answer True or False *(5 points)*

1. A frequency distribution can be formed with discrete data only.
   Answer: False

2. A stem and leaf diagram is used to show the relation between two variables.
   Answer: False

3. If after graphing the data for a quantitative variable of interest, you notice that the distribution is highly skewed in the positive direction, the measure of central location that would likely provide the best assessment of the center would be the median.
   Answer: True

4. A recent study involving a sample of 3,000 vehicles in California showed the following statistics related to the number of miles driven per day: Q1 = 12 Q2 = 45 and Q3 = 56. Based on these data, if a box and whisker plot is developed, a value of 123 miles would be considered an outlier.
   Answer: True

5. A manufacturing company has two assembly lines in its Al-Kharj plant. Line A produces an average of 335 units per day with a standard deviation equal to 11 units. Line B produces an average of 145 units per day with a standard deviation equal to 8 units. Based on this information, line B is more consistent.
   Answer: False

Question 2. *(5 points)*

1) A study of middle to upper level mangers is undertaken to investigate the relationship between salary level and years of work experience. An appropriate graph to display the relationship between the two variables is
   a) Histogram
   b) Scatter diagram
   c) Stem – and –leaf diagram
   d) Bar chart

2) stem and leaf diagram is an alternative method to using
   a) A pie chart
   b) A bar chart
   c) A histogram
   d) An ogive

3) If a manager wishes to analyze the sales trend for her department, possibly the most effective type of graph will be:
   a) A pie chart.
   b) A histogram.
   c) A line chart.
   d) A Pareto Chart

4) A sample of people who have attended a college football game at your university has a mean = 3.2 members in their family. The mode number of family members is 2 and the median number is 2.0. Based on this information:

   a) the population mean exceeds 3.2.
   b) the distribution is bell-shaped.
   c) the distribution is right-skewed.
   d) the distribution is left-skewed**.**

5) The advantage of using the interquartile range versus the range as a measure of variation is:

   a) it is easier to compute.
   b) it utilizes all the data in its computation.
   c) it gives a value that is closer to the true variation.
   d) it is less affected by extremes in the data.

3) The following data represent the number of calls came to a maintenance center between 10:00am to 11:00am for 20 days: *(5 points)*

13	24	24	28	32	52	40	44	53	11
20	32	40	51	40	38	16	39	34	32

a) Construct a stem –and – leaf diagram.
b) Comment on the distribution of the number of calls received.

Solution:

a)

Stem	Leaf
1	1 3 6
2	0 4 4 8
3	2 2 2 4 8 9
4	0 0 0 4
5	1 2 3

b) The distribution is symmetric with median of 33 calls per day; most of the data (70%) receive between 20 and 44 calls.

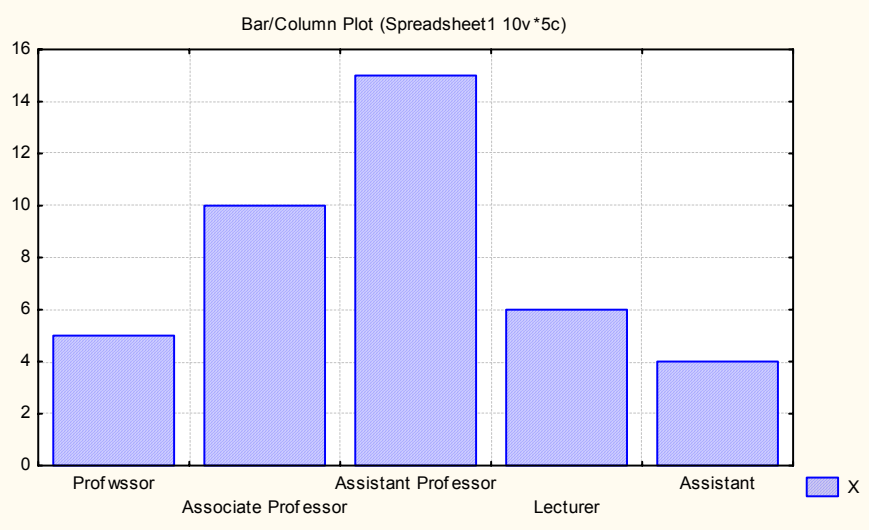
4) The following data represent the number of instructors *X* at a certain department and their rank *R*. *(5 points)*

X	5	10	15	6	4
R	Professor	Associate Professor	Assistant Professor	Lecturer	Assistant

a) Construct a bar chart for this table.
b) Find an appropriate central tendency measurement for this data.

Solution:

a)



b) The Mode is the appropriate measure of central tendency (Assistant Professor)

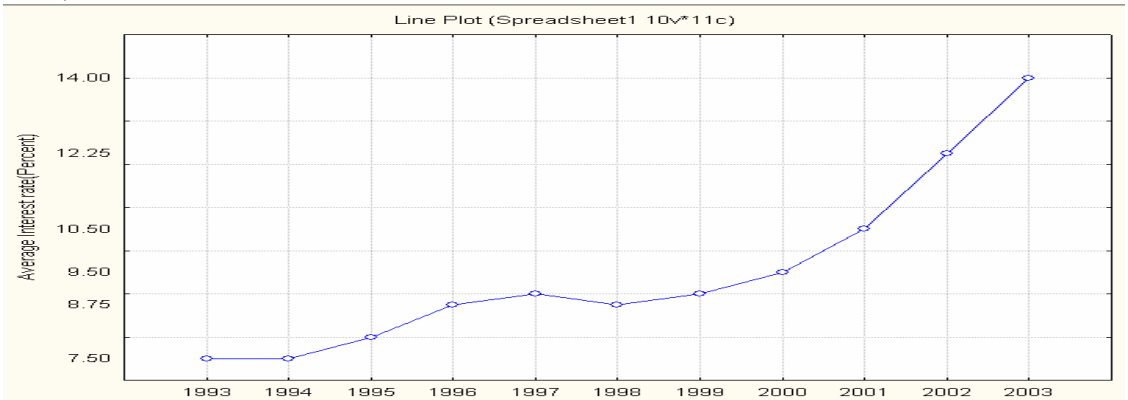
5) The following data represent average interest rate on new homes *(5 points)*

Year	Average Interest rate(Percent)
1993	7.5
1994	7.5
1995	8.0
1996	8.75
1997	9.0
1998	8.75
1999	9.0
2000	9.5
2001	10.5
2002	12.25
2003	14.0

a) Construct a line graph for the above table.
b) Comment on the trend in average interest rate.

Solution

a)



b) The interest rate increasing from 1993 till 2003.

6) A study of houses sold recently in your community showed the following frequency distribution for the number of bedrooms: *(5 points)*

Bedrooms	Frequency	$x_i w_i$
1	2	2
2	18	36
3	140	420
4	57	228
5	11	55

What is the mean number of bedrooms?

$$\mu = \frac{\sum x_i w_i}{\sum w_i} = \frac{1(2)+2(18)+3(140)+4(57)+5(11)}{2+18+140+57+11}$$

$$= \frac{2+36+420+228+55}{228} = \frac{741}{228} = 3.25 \text{ bedrooms}$$

7) The following data reflect the number of television sets in a sample of 16 households. *(15 points)*

3	1	1	2	4	3	2	2
2	0	1	2	3	2	4	5

a) Find the median of the number of televisions.
b) Find the variance of the number of televisions.
c) Based on these sample data what is the standardized value corresponding to 5 televisions?
d) Construct a box plot.

Show All Details

Solution

X	0	1	1	1	2	2	2	2	2	2	3	3	3	4	4	5
X^2	0	1	1	1	4	4	4	4	4	4	9	9	9	16	16	25

Then $\sum x_i = 37, \sum x_i^2 = 111$ *and* $n = 16$

a) Median $= \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + \text{next observation} \right) = \frac{1}{2}(2+2) = 2$

b) $S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$, $\bar{x} = \frac{\sum x_i}{n} = \frac{37}{16} = 2.3125$

$$S^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{111 - 16(2.3125)^2}{15} = \frac{111 - 85.5625}{15} = \frac{25.4375}{15} = 1.695833$$

c) Let z be the standardized value corresponding to 5, then

$\bar{x} = 2.3125$ and $s = \sqrt{s^2} = \sqrt{1.695833} = 1.302241657$

$$z = \frac{x - \bar{x}}{s} = \frac{5 - 2.3125}{\sqrt{1.695833}} = \frac{2.6875}{1.302241657} = 2.063749$$

d) $Q_1 = P_{25} \Rightarrow i = \frac{25}{100}17 = 4.25 \Rightarrow Q_1 = X_{(4)} + 0.25\left(X_{(5)} - X_{(4)}\right) = 1 + 0.25(2-1) = 1.25$

$Q_3 = P_{75} \Rightarrow i = \frac{75}{100}(17) = 12.75 \Rightarrow Q_3 = X_{(12)} + 0.75\left(X_{(12)} - X_{(11)}\right) = 3 + 0.75(3-3) = 3$

$IQR = Q_3 - Q_1 = 3 - 1.25 = 1.75$

$UL = Q_3 + 1.5IQR = 3 + 2.625 = 5.625$ *and* $LL = Q_1 - 1.5IQR = 1.25 - 2.625 = -1.375$

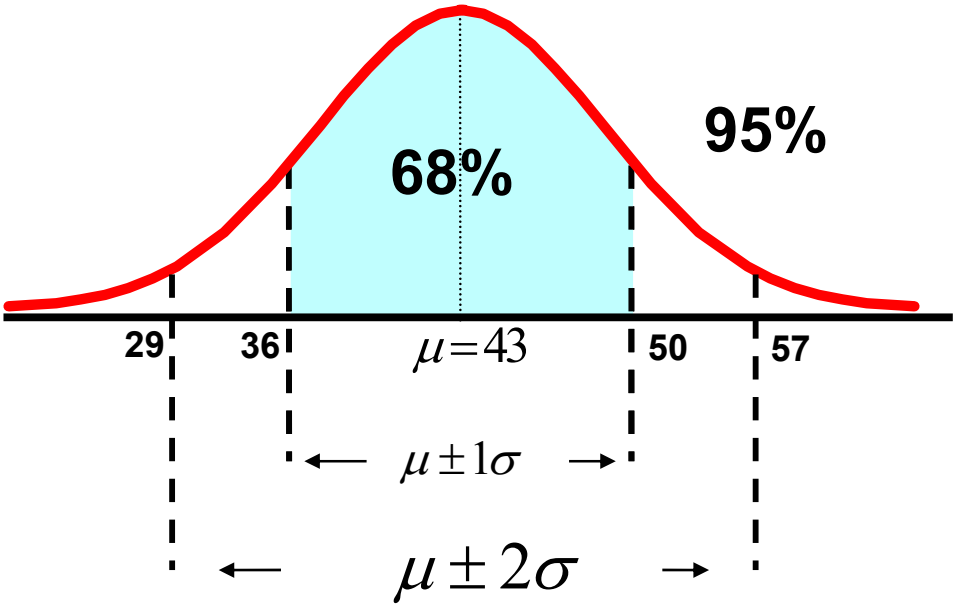
Comment:

50% of the households own between 1.25 and 3 television sets, and the data is skewed to the highest number of T.V sets.

8) If the age distribution of customers at a major retail chain is thought to be bell-shaped with a mean equal to 43 years and a standard deviation equal to 7 years,
   a) find the percentage of customers between the ages of 29 and 50 years.
   b) What is the 16th percentile? *(5 points)*

Solution

a) By the Empirical Rule:



- within one standard deviation 68% of the customers
- within one standard deviation 95% of the customers

The percentage of customers between 29 and 50 will be the percentage within one standard deviation **plus** half the deference between the percentage within one standard deviation and two standard deviations

$$= 68\% + \frac{95\% - 68\%}{2} = 68\% + 13.5\% = 81.5\%$$

b) the 16th percentile is 36 years