

Tutorial Sheet 18

(Descriptive Statistics)

1. Explain which of the following variables are quantitative and which are qualitative.
 - a. Number of persons in a family
 - b. Colour of cars
 - c. Marital status of people
 - d. Types of cars owned by families
 - e. Weight of a keyboard
 - f. Monthly phone bills
 - g. Colour of eyes
 - h. Distance between two bus stops
2. Classify the quantitative variables in question 1 as discrete or continuous.
3. The following data give the results of a sample survey. The letter A, B and C represent the three categories.

A	B	B	A	C	B	C	C	C	A
C	B	C	C	C	C	B	C	C	A
A	B	C	C	B	C	B	B	C	A

 - a. Prepare a frequency distribution table.
 - b. Calculate the relative frequencies and percentages for all categories.
 - c. What percentage of the elements in this sample belongs to category B?
 - d. What percentage of the elements in this sample belongs to categories A or C?
 - e. Draw a bar graph for the frequency distribution.
4. A data set on weekly expenditures on bakery products for a sample of 500 households has a minimum value of \$1 and a maximum value of \$18. Suppose we want to group these data into five classes of equal widths.
 - a. Assuming we take the lower limit of the first class as \$1 and the upper limit of the fifth class as \$20, write the class limits for all five classes.
 - b. Determine the class boundaries and class widths.
 - c. Find the class midpoints.
5. A manufacturing company manufactures computer terminals. The following data give the number of computer terminals produced at the company for a sample of 30 days.

24	32	27	23	33	33	29	25	23	28
21	26	31	22	27	33	27	23	28	29
31	35	34	22	26	28	23	35	31	27

 - a. Construct a frequency distribution table.
 - b. Calculate the relative frequencies and percentages for all classes.
 - c. Construct a pie chart, a histogram and a polygon for the percentage distribution.
 - d. For what percentage of the days is the number of computer terminals produced in the interval 27 to 29?
6. The following data give the number of hours spent on accessing Internet by 10 randomly selected college students during the past week.

7 14 5 0 2 7 10 4 0 8

Compute the mean, median and mode.

7. The following data set belongs to a population.

33 33 48 30 30 47 40 34 44 129

- Calculate the mean and median for these data.
- Do these data contain an outlier? If yes, drop this value and recalculate the mean and median.
- Is the mean or the median a better summary measure for these data? Explain.

8. The following sample data give the number of cars that stopped at a service station during each of the 10 hours observed. Find the range, variance and standard deviation.

29 35 42 31 24 18 16 27 39 34

9. Consider the following two data sets.

Data set I: 12 25 37 8 41

Data set II: 19 32 44 15 48

Note that each value of the second data set is obtained by adding 7 to the corresponding values of the first data set. Calculate the standard deviation for each of these two data sets using the formula for sample data. Comment on the relationship between the two standard deviations.

10. For 50 airplanes that arrived late at an airport during a week, the time by which they were late was observed and recorded in the following table.

Time (in minutes)	Frequency
0 to less than 20	14
20 to less than 40	18
40 to less than 60	9
60 to less than 80	5
80 to less than 100	4

Find the mean, variance and standard deviation.

11. The ages of car owned by all employees of a large company have a bell-shaped distribution with a mean of 7 years and a standard deviation of 2 years.

- Using the empirical rule, find the percentage of cars owned by these employees that are
 - 5 to 9 years old
 - 1 to 13 years old
- Using the empirical rule, find the interval that contains the ages of the cars owned by 95% of the employees of this company.