



4.1. BAR CHARTS

Bar charts are the most common type of diagrams used in practice. A bar is a rectangle whose width is insignificant. Only the length of the bar is significant. When the number of items is large, thick lines may be drawn instead of bars to economise space. For comparison, they are most effective.

A bar chart is a pictorial representation of the numerical data by a number of bars (rectangles) of uniform width drawn horizontally or vertically. The gap between one bar and another should be uniform throughout. The vertical bars are preferred because they give a better look and facilitate comparison. Each bar represents only one value of the numerical data and, therefore, there are as many bars as the number of values in the numerical data.

If the bars are drawn vertically on x -axis, then scale of heights of the bars is shown along y -axis. If the bars are drawn horizontally on y -axis, then scale of lengths of bars is shown along x -axis. The bars can be shaded or coloured.

The following examples will illustrate the construction of bar charts.

Example 1. *The average temperatures (in degrees celsius) for a city for the months May to August during a year as reported by the Meteorological Department are given as follows:*

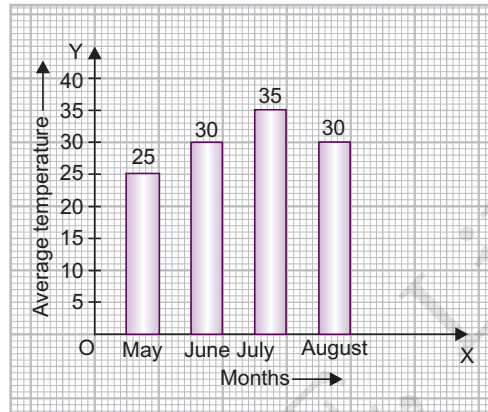
Month	May	June	July	August
Average temperature	25	30	35	30

Represent the above data by bar chart.

Solution. Draw the mutually perpendicular lines OX and OY. Along OX, mark the months and along OY, the average temperature (in degrees celsius). Decide the width of each bar and the space to be left between

consecutive bars. Choose a convenient scale on y -axis for the numerical data.

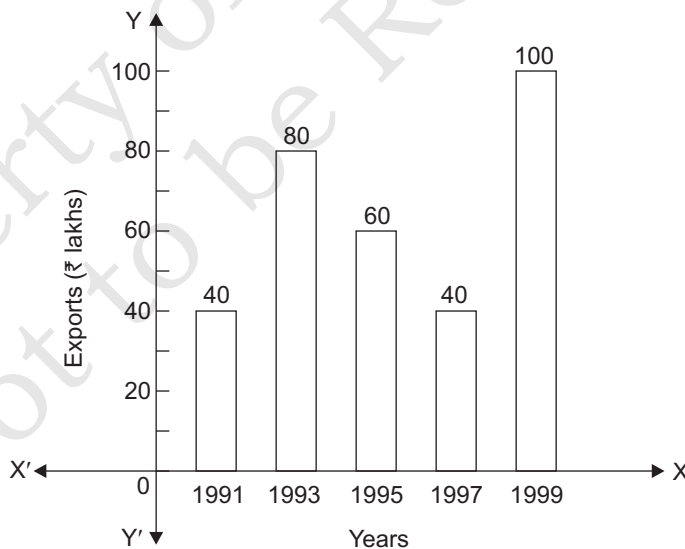
The bar chart for the given data is shown in figure.



Example 2. Present the following information on exports of Liberia in the form of simple bar diagram:

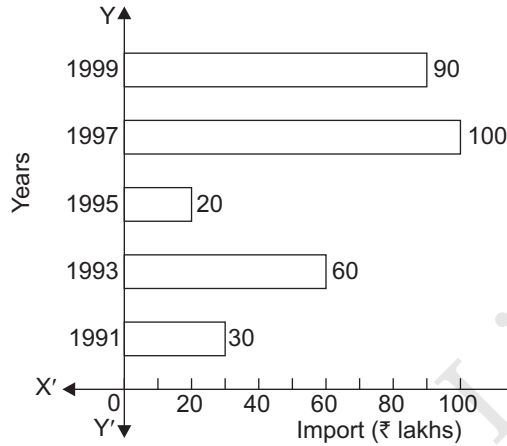
Years	1991	1993	1995	1997	1999
Exports (in L\$)	40	80	60	40	100

Solution.



Example 3. Present the following information in a simple bar diagram (horizontal):

Years	1991	1993	1995	1997	1999
Imports (in L\$)	30	60	20	100	80

Solution.**4.2. PIE CHARTS**

A pie chart is a pictorial representation of the numerical data by non-intersecting adjacent sectors of a circle such that area of each sector is proportional to the magnitude of the data represented by the sector. The pie chart is so called because the entire graph looks like a pie and the components (sectors) resemble the slices cut from a pie. Pie charts are used to show the break-up of a total into component parts. For example, the pie may represent the budget of a family for a month and the components may represent portions of the budget allotted to rent, food, clothing, education and so on. Similarly, through a pie chart it can be shown how an amount spent by a firm is distributed over various heads such as wages, raw materials, administration expenses etc.

Following steps may be used for the construction of a pie chart.

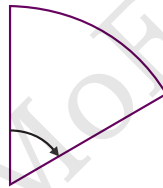
- (i) The area of each sector is proportional to the component value of the component represented by the sector. Also, the area of a sector is proportional to the angle subtended by its arc at the centre, therefore, central angles are proportional to the component values.

Thus, central angle of a component

$$= \left(\frac{\text{Value of the component}}{\text{Sum of the component values}} \times 360 \right)^\circ$$

Using the above formula, find the central angles of the various components and arrange them in descending order.

- (ii) Draw a circle of appropriate size with a compass. Also, draw the vertical radius coinciding with 12 o'clock position of the two hands of a clock.
- (iii) In laying out the sectors of a pie chart figure, it is desirable to follow some logical arrangement. It is common procedure to arrange the sectors according to size, with the largest at the top and the others in succession running clockwise. Thus start with the component having largest central angle and construct a sector in such a way that its one radius coincides with the vertical radius drawn in step (ii) and the other radius is below the first radius in clockwise direction.



- (iv) Construct other sectors representing other component in clockwise sense in descending order of magnitude of central angles. Components like 'Miscellaneous', irrespective of the magnitude of their central angle, are shown last.
- (v) Shade the sectors by different colours or designs.
- (vi) An essential feature of a pie chart is to put descriptive labels inside each sector so that they can be easily identified. If it is not possible to place the labels inside the sectors due to lack of space, then put the labels outside the circle with an arrow pointing to the appropriate sector.

Example 4. Draw a pie chart for the following data of expenditure pattern in a family:

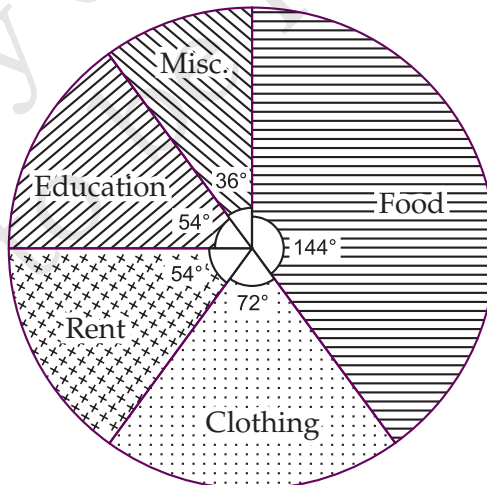
Items	Food	Clothing	Rent	Education	Miscellaneous
Expenditure (in cedis)	4000	2000	1500	1500	1000

Solution.

Computation of Central Angles

Item	Expenditure	Central angle
Food	4000	$\left(\frac{4000}{10000} \times 360\right)^\circ = 144^\circ$
Clothing	2000	$\left(\frac{2000}{10000} \times 360\right)^\circ = 72^\circ$
Rent	1500	$\left(\frac{1500}{10000} \times 360\right)^\circ = 54^\circ$
Education	1500	$\left(\frac{1500}{10000} \times 360\right)^\circ = 54^\circ$
Miscellaneous	1000	$\left(\frac{1000}{10000} \times 360\right)^\circ = 36^\circ$
Total	10,000	360°

Draw a circle of appropriate size with a compass. Draw the vertical radius. Construct sectors in clockwise sense with descending order of magnitude of central angles as shown in figure.



Example 5. The number of students admitted in different faculties of a college are given below:

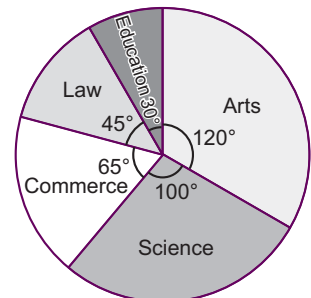
Faculty	Number of students
Science	1000
Arts	1200
Commerce	650
Law	450
Education	300
Total	3600

Draw a pie chart to represent the above information.

Solution. Computation of Central Angles

Faculty	Number of students	Central angle
Science	1000	$\left(\frac{1000}{3600} \times 360\right)^\circ = 100^\circ$
Arts	1200	$\left(\frac{1200}{3600} \times 360\right)^\circ = 120^\circ$
Commerce	650	$\left(\frac{650}{3600} \times 360\right)^\circ = 65^\circ$
Law	450	$\left(\frac{450}{3600} \times 360\right)^\circ = 45^\circ$
Education	300	$\left(\frac{300}{3600} \times 360\right)^\circ = 30^\circ$
Total	3600	360°

Draw a circle of appropriate size with a compass. Draw the vertical radius. Construct sectors in clockwise sense with descending order of magnitude of central angles as shown in figure.



4.3. HISTOGRAMS

A histogram is a graphical representation of a frequency distribution in the form of adjacent rectangles with class intervals as bases and heights proportional to frequency.

Note that in a bar chart, there is uniform gap between one bar and another whereas there is no gap between adjacent rectangles in a histogram. In a bar chart, the width of a rectangle is insignificant. Whereas in a histogram, the width of a rectangle is significant and represents the class size.

To construct a histogram, we proceed as follow:

- (i) Represent the class intervals along x -axis on a suitable scale.
- (ii) Represent the frequency along y -axis on a suitable scale.
- (iii) Construct rectangles with class intervals as bases and the corresponding frequencies as heights.

The diagram representing the above rectangles is the required histogram.

Note: It must not be assumed that scale for both the axes is the same. We can have different scales for the two axes. The determination of scale depends upon our convenience and the type of data. The scale or scales should be so chosen as to fit the size of graph paper and to hold all the figures of the data.

4.4. MEAN OF GROUPED DATA WITH CONTINUOUS FREQUENCY DISTRIBUTION

In a grouped data with continuous or exclusive classes, we may recall that an observation falling in any upper-class limit it considered in the next class. Here we assume that the frequency of each class interval is centred around its midpoint, called the class mark. So the mark of each class-interval serves a representative of the whole class. The class mark of any class is equal to the average of its upper and lower limits. Thus,

$$\text{Class mark} = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

4.5. FREQUENCY TABLE

As soon as the work related to collection of data is over, the investigator has to find out ways to present them in a form which is meaningful, easily understood and gives the main feature at a glance. Usually, data available to us is in unorganised form called *raw data*.

Making Frequency Tables

In this section, we will learn how to make frequency tables by tallying in groups of five and write the frequencies. Let us construct frequency tables for a given data.

Example 6. A group of 30 pupils were surveyed on which animal they would like the most to have as a pet. The results are given below:

dog, cat, cat, fish, cat, rabbit, dog, cat, rabbit, dog, cat, dog, dog, dog, cat, rabbit, fish, rabbit, dog, cat, dog, cat, cat, dog, rabbit, cat, fish, dog, rabbit, cat.

Make a frequency distribution table for this data.

Solution. For counting purposes, we use tally marks. After putting 4 tally marks vertically, we cross it as shown below and again we take the tally marks in the same manner, counting in sets of fives.

Frequency Distribution Table

Animal	Tally Marks	Number of Animals (Frequency)
Dog		10
Cat		11
Fish		3
Rabbit		6
	Total	30

As mentioned earlier, the frequency gives the number of times a particular entry occurs. The above table is known as *frequency distribution table* as it gives the number of times an entry occurs.

Grouping Data

Sometimes, we have to deal with a large data.

If we make a frequency distribution table for each observation, the table would be too long, so, for convenience, we make groups of observations say, 0–10, 10–20 and so on, and obtain a frequency distribution of the number of observations falling in each group.

Example 7. Consider the following marks (out of 50) obtained in Mathematics by 60 pupils of Junior High School 2.

21, 10, 30, 22, 33, 5, 37, 12, 25, 42, 15, 39,
26, 32, 18, 27, 28, 19, 29, 35, 31, 24, 36, 18,
20, 38, 22, 44, 16, 24, 10, 27, 39, 28, 49, 29,
32, 23, 31, 21, 34, 22, 23, 36, 24, 36, 33, 47,
48, 50, 39, 20, 7, 16, 36, 45, 47, 30, 22, 17.

Make a frequency table for the above data using intervals 0–10, 10–20 and so on.

Solution. Grouped Frequency Distribution Table

Groups	Tally Marks	Frequency
0–10		2
10–20		10
20–30		21
30–40		19
40–50		7
50–60		1
	Total	60

Data represented in this way is said to be *grouped* and the distribution obtained is called *grouped frequency distribution table*.

It helps us to draw meaningful inferences like:

- (i) Most of the pupils have scored between 20 and 40.
- (ii) Eight pupils have scored more than 40 marks out of 50 and so on.

Each of the groups: 0–10, 10–20, 20–30 etc., is called a *Class Interval* (or briefly a class).

Observe that 10 occurs in both classes, *i.e.*, 0–10 as well as 10–20. Similarly, 20 occurs in classes 10–20 and 20–30. But it is not possible that an observation (say 10 or 20) can belong simultaneously to two classes. To avoid this, we adopt the convention that the common observation will belong to the higher class, *i.e.*, 10 belongs to the class interval 10–20 (and not to 0–10).

Similarly, 20 belongs to 20–30 (and not to 10–20). In the class interval, 10–20, 10 is called the *lower class limit* and 20 is called the *upper class limit*.

Similarly, in the class interval 20–30, 20 is the *lower class limit* and 30 is the *upper class limit*.

Observe that the difference between the upper class limit and lower class limit for each of the class intervals 0–10, 10–20, 20–30 etc., is equal, (10 in this case). This difference between the upper class limit and lower class limit is called the *width* or *size* of the class interval.

4.6. MEAN OF GROUPED DATA WITH CONTINUOUS FREQUENCY DISTRIBUTION

In a grouped data with continuous or exclusive classes, we may recall that an observation falling in any upper-class limit is considered in the next class. Here we assume that the frequency of each class interval is centered around its midpoint, called the class mark. So the class mark of each class interval serves a representative of the whole class. The class mark of any class is equal to the average of its upper and lower limits. Thus

$$\text{Class mark} = \frac{\text{Upper class limit} + \text{Lower class limit}}{2}$$

Mean of Grouped data by direct method

Step 1 For each class, find the class mark x_i

$$x_i = \frac{1}{2}(\text{lower limit} + \text{upper limit})$$

Step 2 Calculate $f_i x_i$ for each class

Step 3 Calculate the mean $\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$

This method is used when the numerical values of f_i and x_i are small.

Example 8. A survey was conducted by a group of students as a part of their environment awareness programme, in which they collected the following data regarding the number of plants in 20 houses in a locality. Find the mean number of plants per house.

Number of plants	0-2	2-4	4-6	6-8	8-10	10-12	12-14
Number of houses	1	2	1	5	6	2	3

Which method did you use for finding the mean, and why?

Solution. Grouped Frequency Distribution Table

Number of plants (C.I.)	Number of Houses (f_i)	Class mark (x_i)	$f_i x_i$
0-2	1	1	1
2-4	2	3	6
4-6	1	5	5
6-8	5	7	35
8-10	6	9	54
10-12	2	11	22
12-14	3	13	39
Total	$\Sigma f_i = 20$		$\Sigma f_i x_i = 162$

$$\therefore \text{Mean } (\bar{x}) = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{162}{20} = 8.1$$

Hence, the mean number of plants per house is 8.1. Here, we have used direct method for finding the mean because the values of f_i and x_i are small.

4.7. MODE

Mode of any frequency distribution is that value of the observation for which the frequency is maximum. A given frequency distribution may have the same maximum frequency for more than one observation then given data is multi modal.

Mode of Grouped Data:

- (i) From the given frequency distribution, determine the class of maximum frequency distribution. This is known as modal class.

(ii) Calculate the mode by using the formula

$$\text{Mode} = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

where l = lower level of modal class

f_i = frequency of modal class

f_0 = frequency of the class preceding the modal class

f_2 = frequency of the class succeeding the modal class

h = class size

Example 9. Find the mode of the following data:

Marks	50-60	60-70	70-80	80-90	90-100
Number of students	3	12	32	20	6

Solution. As the class interval 70-80 has maximum frequency i.e., 32 the modal class 70-80.

Now $l = 70$, $f_1 = 32$, $f_0 = 12$, $f_2 = 20$, $h = 10$

$$\begin{aligned} \therefore \text{Mode} &= l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) h = 70 \left(\frac{32 - 12}{2 \times 32 - 12 - 20} \right) \times 10 \\ &= 70 + \frac{20}{32} \times 10 = 70 + 6.25 = 76.25 \end{aligned}$$

4.8. MEDIAN OF A GROUPED FREQUENCY DISTRIBUTION

1. Prepare a cumulative frequency table from the given frequency distribution
2. Find n (total frequency) and calculate $\frac{n}{2}$
3. From the cumulative frequency column of the table, find the cumulative frequency.
4. Calculate the median, by using the formula

$$\text{Median} = l + \left(\frac{\frac{n}{2} - Cf}{f} \right) \times h$$

where l = lower limit of median class

Cf = Cumulative frequency of the class preceding the median class

f = Frequency of median class

h = Class size

Example 10. Weekly income of 600 families in tabulated below:

Weekly income (in L\$)	Number of families
0-1000	250
1000-2000	190
2000-3000	100
3000-4000	40
4000-5000	15
5000-6000	5
<i>Total</i>	<i>600</i>

Compute the median income.

Solution. Frequency distribution table is

Weekly income (in L\$)	Number of families	Cumulative frequency
0-1000	250	250
1000-2000	190	440 ← Median class
2000-3000	100	540
3000-4000	40	580
4000-5000	15	595
5000-6000	5	600

Here $n = 600$

$$\therefore \frac{n}{2} = 300$$

\therefore Median class is 1000-2000

Now $l = 1000$, $Cf = 250$, $f = 190$, $h = 100$

$$\begin{aligned} \text{Median} &= l + \left(\frac{\frac{n}{2} - Cf}{f} \right) \times h = 1000 + \left(\frac{300 - 250}{190} \right) \times 1000 \\ &= 1000 + 263.16 = 1263.16 \end{aligned}$$

4.9. QUARTILES AND PERCENTILES

Quartiles: Distribution is divided into quarters.

Percentiles: Distribution is divided into hundredths.

4.10. CUMULATIVE FREQUENCY CURVE (OGIVE)

The curve obtained by plotting cumulative frequencies is called a **cumulative frequency curve** or an **ogive**.

Plot the points with the *upper limits* of the classes as abscissae and the corresponding less than cumulative frequencies as ordinates. Join the points by a free hand smooth curve to get the required ogive. It is a rising curve.

Example 11. Draw the ogive for the following distribution showing the number of marks of 59 pupils:

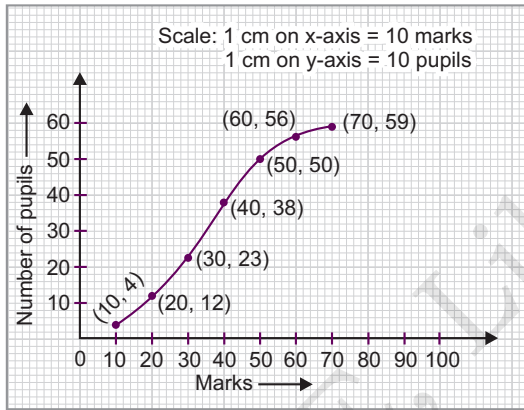
Marks	No. of students	Marks	No. of students
0–10	4	40–50	12
10–20	8	50–60	6
20–30	11	60–70	3
30–40	15	—	—

Solution. We first calculate the cumulative frequencies as shown in the following table.

Table. Table of cumulative frequencies

Marks	Number of pupils	Upper limits	Cumulative frequency (less than)
0–10	4	10	4
10–20	8	20	12
20–30	11	30	23
30–40	15	40	38
40–50	12	50	50
50–60	6	60	56
60–70	3	70	59

Plot the points (10, 4), (20, 12), (30, 23), (40, 38), (50, 50), (60, 56) and (70, 59) on the graph paper and join them by free hand. The smooth rising curve so obtained is the required ogive. (see figure).



Ogive or cumulative frequency curve

Through Q, draw a straight line parallel to y -axis to meet the x -axis at the point Q_1 (say). The abscissa of Q_1 is L\$ 1725.

$\therefore D_9 = \text{L\$ } 1725$ is the required decile

Calculation of P_5 , P_{30} and P_{83} :

For P_5 ; $\frac{5N}{100} = \frac{5 \times 100}{100} = 5$

For P_{30} ; $\frac{30N}{100} = \frac{30 \times 100}{100} = 30$

For P_{83} ; $\frac{83N}{100} = \frac{83 \times 100}{100} = 83$

Draw a straight line through the points 5, 30 and 83 (on y -axis) parallel to x -axis to meet the ogive at the points R, S and T (say).

Through R, S and T, draw straight lines parallel to y -axis to meet the x -axis at the points R_1 , S_1 and T_1 .

The abscissa of R_1 is L\$ 366.

$\therefore P_5 = \text{L\$ } 366$

The abscissa of S_1 is L\$ 857.

$\therefore P_{30} = \text{L\$ } 857$

The abscissa of T_1 is L\$ 1560.

$\therefore P_{83} = \text{L\$ } 1560$

Hence, $P_5 = \text{L\$ } 366$, $P_{30} = \text{L\$ } 857$ and $P_{83} = \text{L\$ } 1560$ are the required percentiles.

4.11. QUARTILES DEFINITION

Quartiles divide the entire set into four equal parts. So, there are three quartiles, first, second and third represented by Q_1 , Q_2 and Q_3 . Q_2 is nothing but the median, since it indicates the position of the item in the list and thus, is a positional average. To find quartiles of a group of data, we have to arrange the data in ascending order.

4.12. PERCENTILE

A percentile (or a percentile) is a measure used in statistics indicating the value *below which* a given percentage of observations in a group of observations fall. For example, the 20th percentile is the value (or score) below which 20% of the observation may be found.

The term percentile and the related term *percentile rank* are often used in the reporting of scores from norm-references tests. For example, if a score is at the 86th percentile, where 86 is the percentile rank, it is equal to the value below which 86% of the observations may be found. In contrast, if it is in the 86th percentile, the score is at or below the value of which 86% of the observations may be found. **Every score is in the 100th percentile.**

The 25th percentile is also known as the first quartile (Q_1), the 50th percentile as the median or second quartile (Q_2) and the 75th percentile as the third quartile (Q_3). In general, percentiles and quartiles are specific types of quantiles.

$$\text{Formula: } Ly = \frac{(n + 1)Y}{100}$$

Example 12. Calculate the first quartile of a distribution that consists of the following portfolio returns.

3% 4% 6% 9% 11% 12% 14%

Solution. The first quartile = $(7 + 1)25\% = 2^{\text{nd}}$ item in the data set (4%) i.e. 25% of the observation lie below the second observation from the left.

EXERCISE

1. The import and export of a country (in million dollars) during the years (2012–13 to 2016–17) are given below:

<i>Year:</i>	2012–13	2013–14	2014–15	2015–16	2016–17
<i>Import:</i>	4000	5000	7000	8000	9500
<i>Export:</i>	5000	7000	8500	11000	14000

2. Present the following information on expenditure by family A and family B in a multiple bar diagram:

<i>Items of expenditure</i>	<i>Family A</i>	<i>Family B</i>
Food	10000	7000
Clothing	3000	1000
Education	5000	1500
Misc. exp.	2000	500

3. Draw a pie diagram to represent the following information of expenditure by a family:

<i>Items of expenditure</i>	Food	Education	Housing	Clothing	Misc.
<i>% of total expenditure</i>	60	15	10	10	5

4. A Junior Football Club in Accra has thirty members. Their ages are given below. Make a frequency distribution table for it.

13, 17, 13, 13, 14, 15, 15, 14, 16, 14, 16, 17, 16, 16, 13, 15, 16, 15, 15, 14, 14, 15, 15, 13, 13, 15, 14, 13, 16, 17.

5. The monthly wages (in L\$) of 30 workers in a factory are:

330, 335, 390, 310, 335, 336, 369, 345, 398, 390, 360, 332, 333, 355, 345, 304, 308, 312, 340, 385, 335, 335, 336, 378, 340, 368, 390, 306, 340, 320.

Using tally marks make a frequency table with intervals as 300–310, 310–320 and so on.

6. Following is the distribution of marks in Mathematics test at Senior High School 2 level

<i>Marks</i>	<i>Number of Students</i>
0–10	3
10–20	7
20–30	10
30–40	20
40–50	6
50–60	4

If 60% of the students pass the test, use cumulative frequency curve to find the minimum marks obtained by a pass student.

7. Find the median, lower quartile and upper quartile of the following numbers:

12, 5, 22, 30, 7, 36, 14, 42, 15, 53, 25

8. The distribution of heights of 50 students (in nearest cm) is given below:

<i>Height (in cm)</i>	110	115	118	120	121	125
<i>Number of students</i>	6	8	14	15	4	3

9. Find the mean of the following frequency distribution:

<i>Class</i>	0–100	100–200	200–300	300–400	400–500
<i>Frequency</i>	6	9	15	12	8