# CHAPTER 2 CON’T

**THE GRAPHICAL PRESENTATION OF NUMERICAL FREQUENCY DISTRIBUTIONS**

# 2.1 Graphical methods of presentation

After presenting and summarizing data in the form of frequency distribution tables, presenting it in the graphical form gives a more visual display. Data can be presented in the following ways:-

## Histograms

* + 1. ***Frequency polygons***

## Cumulative frequency curves (Ogives)

***1. Histogram***

This is a series of rectangular bars drawn adjacent to each other on the same graph whose widths correspond to the class sizes and the height of each bar corresponds to the respective class frequency. Hence it is a set of vertical bars where each bar represents just one class. The vertical axis represents frequencies while the horizontal axis represents the classes of the data values.

The bars should be joined together to reinforce the fact that it presents continuous grouped data whose classes have common boundaries

Data values that do not begin at zero require a small gap before drawing the first bar

Any blank space between the bars implies that the corresponding class is empty; thus the need to make it continuous by creating class boundaries.

The histogram should have a clear title

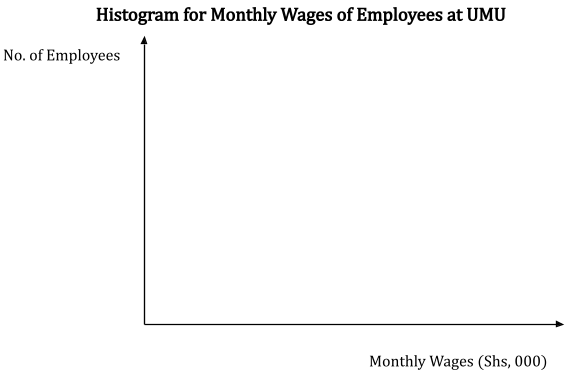
# Examples

1. The monthly wages of employees at UMU are as follows (Exclusive class):

|  |  |
| --- | --- |
| Wages per employee (Shs,000) | No. of Employees |
| 150 to under 200 | 5 |
| 200 to under 250 | 3 |
| 250 to under 300 | 7 |
| 300 to under 350 | 5 |
| 350 to under 400 | 3 |
| 400 to under 450 | 2 |
| 450 to under 500 | 4 |

Draw the histogram for this data

**Solution:** The horizontal axis should represent bars of equal width while the vertical axis should represent the frequencies as shown below:-

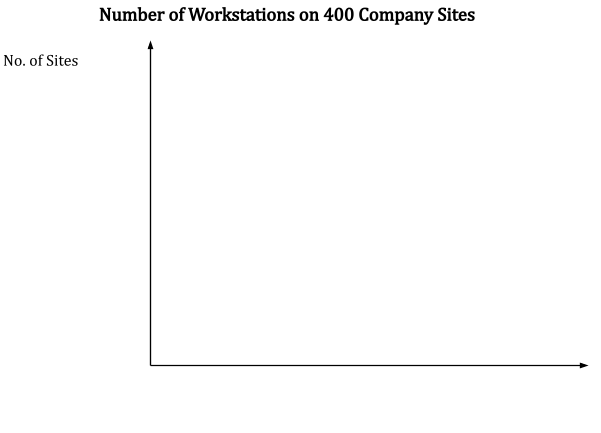
.

1. The following data is from a survey of the number of workstations on 400 company sites.

|  |  |
| --- | --- |
| Number of workstations | No. of sites |
| 1 – 5 | 45 |
| 6 – 10 | 90 |
| 11 – 15 | 100 |
| 16 – 20 | 85 |
| 21 – 25 | 60 |
| 26 − 30 | 20 |

Draw a Histogram for the above data

# Solution (make it continuous)



**Advantages of a histogram**

* 1. Data is visually clearly displayed for comparative analysis by observing the heights of bars
  2. It indicates the range of values by seeing the width of the histogram; the wider the histogram, the larger is the range and vise versa.
  3. It can show whether data is skewed or not

# Disadvantages of the histogram

1. When sample data is used and not accurately collected, it can present a false display
2. Where class widths vary, the histogram becomes difficult to understand

## Frequency polygons

This is a line graph of the frequency distribution in which the mid values of the classes are plotted against the class frequencies and the plotted points are joined by a straight line. The polygon should be enclosed with ***two arbitrary mid points at the start and at the end.***

The frequency polygon can also be drawn from a histogram in the following ways:

* 1. Mark the midpoint of the top of each bar in the histogram
  2. Join up all these points with straight lines
  3. Enclose the polygon with two arbitrary midpoints at the mid points of the next classes outside the range of the start and end of the observed data.

# Example (inclusive and exclusive classes)

1. A businessman is studying the lead time for a sample of 40 orders received in the previous month as shown below

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Lead time | 0 − 3 | 4 − 7 | 8 − 11 | 12 − 15 | 16 − 19 | 20 − 22 |
| No. of orders | 6 | 7 | 13 | 8 | 4 | 2 |

* 1. Draw a histogram for this data
  2. From the histogram, draw the frequency polygon

1. The table below shows the savings of customers of two branches of a microfinance institution operating in North Eastern Uganda in 2013

|  |  |  |
| --- | --- | --- |
| Savings (Shs,000) | Mbale | Tororo |
| 50 and under 100 | 8 | 2 |
| 100 and under 150 | 14 | 5 |
| 150 and under 200 | 36 | 11 |
| 200 and under 250 | 20 | 17 |
| 250 and under 300 | 17 | 28 |
| 300 and under 350 | 11 | 22 |
| 350 and under 400 | - | 12 |
| 400 and under 450 | - | 8 |
| 450 and under 500 | - | 4 |

* 1. On the same diagram, draw the frequency polygons
  2. Make two observations from the diagrams you have drawn

# Advantages of frequency polygons

Frequency polygons can be used in place of histograms in the following ways:-

1. When there are many classes in the distribution
2. For comparing two or more distributions (see example 2)

## Cumulative frequency curves (ogives)

These are graphical presentations of the cumulative frequency distributions. They are graphs obtained by plotting cumulative frequencies against the class **boundaries;** hence giving rise to two types of **ogives:-**

# The less than ogive

* 1. **The more than ogive**

# The less than cumulative frequency curve

This curve is drawn by plotting the less than CFs (vertical axis) against the upper class boundaries (horizontal axis). The points are joined with a smooth curve. The curve will be upward sloping from right to left

# The more than cumulative frequency curve

This curve is drawn by plotting the more than CFs (vertical axis) against the lower class boundaries (horizontal axis). The points are joined with a smooth curve. The curve will be downward sloping from right to left

# Examples (clas activity)

* 1. The following table shows a grouped frequency distribution for the volume of output produced at a factory over a period of 45 weeks.

|  |  |
| --- | --- |
| Output (Units) | No. of weeks |
| 0 to < 200 | 5 |
| 200 to < 400 | 9 |
| 400 to < 600 | 13 |
| 600 to < 800 | 11 |
| 800 to < 1000 | 7 |
| **Total** | **45** |

# Required

* + 1. Draw a less than ogive
    2. Use the less than ogive to estimate the number of weeks in which the output was 500 units or less
  1. The output of a factory over a period of 100 weeks is shown by the following distribution

|  |  |
| --- | --- |
| Output (Units) | No. of weeks |
| 0 to < 100 | 5 |
| 100 to < 200 | 9 |
| 200 to < 300 | 13 |
| 300 to < 400 | 11 |
| 400 to < 500 | 7 |
| **Total** | **100** |

# Required

* + 1. Draw a more than ogive
    2. Use the ogive to estimate the number of weeks in which the output exceeds 350 units.

**NB:** If the plotted points are joined by straight lines, then you will obtain a CF polygon not an ogive

# Applications of frequency polygons and curves

**2.2.1 Estimating the median, upper quartile and lower quartile using the ogive**

1. Draw the less than ogive
2. Find the position of the median value on the vertical axis as follows:-
   1. For N even items, median value = 𝑁

2

* 1. For N odd items, median value = 𝑁+1

2

* 1. For upper quartile, find the value 3

4

* 1. For lower quartile, find the value 1

4

1. Draw a horizontal line to meet the ogive and then drop a vertical line to meet the horizontal axis and read off the median, upper quartile and lower quartile.

**NB:** The inter-quartile range = upper quartile value ‒ lower quartile value

# Question

The production of each manufacturing department in a company is monitored weekly to

establish productivity bonuses to be paid to the departmental members of staff. 350 items have to be produced each week before a bonus is paid. The production in one department over a 40 week period is shown below

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 493 | 478 | 475 | 476 | 482 | 481 | 483 | 518 | 466 | 458 |
| 460 | 460 | 471 | 468 | 473 | 475 | 516 | 488 | 476 | 472 |
| 491 | 493 | 495 | 497 | 499 | 513 | 517 | 548 | 567 | 570 |
| 577 | 560 | 565 | 571 | 568 | 533 | 540 | 546 | 520 | 465 |

Required

* 1. Form a frequency distribution for the data
  2. Construct a less than ogive
  3. Establish the value of the median from the ogive
  4. Estimate the upper and lower quartiles
  5. Hence find the inter-quartile range

# Solution

R = 577 ‒ 458 = 119

Using 5 classes, then class width = 119/5 = 23.8; adjusting upwards, w = 25.

|  |  |
| --- | --- |
| **Production in a**  **department** | **frequency**  **(f)** |
| 455 ‒ 479 | 14 |
| 480 ‒ 504 | 10 |
| 505 ‒ 529 | 5 |
| 530 ‒ 554 | 4 |
| 555 ‒ 579 | 7 |
| **Total** | **40** |

1. Draw the ogive (see the graph paper)
2. Median = (20th value) = 495
3. Upper quartile (30th value) = 535
4. Lower quartile (10th value) = 470
5. Inter-quartile range = 535 ‒ 470 = 65

**NB:** The middle output (495) is closer to the lower quartile (470) than the upper quartile (535)

# Alternative solution (class activity)

R = 577 ‒ 458 = 119

R/10 = 11.9; adjust upwards or downwards (w = 20 or 10)

# 4.2. 2 Estimating the mode using a histogram

1. Draw the histogram
2. Identify the highest bar
3. Draw diagonal lines and identify their point of intersection.
4. Read off the value of the mode on horizontal axis using the point of intersection