Frequency Distribution& Graphs

2.1 Introduction

After collecting data from any phenomenon, this data is not arranged or organized, so we need to present it in a way that makes it easy to study

- Unclassified data: is the raw data collected about the phenomenon unless it is tabulated
- Classified data: is data arranged and organized in a frequency distribution table (categories, frequencies).

2.2 Frequency Distribution Table

It is a simple table consisting of two columns

- ✓ The first column: The variable values are divided into sections or groups called classes
- ✓ The second column: shows the vocabulary of each class and is called class frequency.

Classes (c)	Frequency (f_i)

2.3 Important Definition

1. **Classes (C):** are the groups into which the variable values are divided, and each class takes a specific range of the variable values.

2. **Frequency(f_i):** is the number of vocabulary or values that fall within the range of that class.
3. **Range (R):** is the difference between the maximum and minimum value of the data
4. **Classes Limits:** Lower class Limit

   Upper class Limit

   Real Lower Class Limit

   Real Upper Class Limit
5. **Class length (L):** is the range between the two limits of the class.
6. **Class Mark or Midpoint (y_i):** is the range between the two limits of the class.

2.4 Steps for organizing a frequency table

In general, following steps in the section below for organizing data.

A. Find the range of the data

R = (max_v-min_v)

B. Find the number of classes

(m) = 2.5*$\sqrt[4]{n}$ = 2.5* $(n)^{\frac{1}{4}}$

(m) = 1+3.3log (n)

- ✓ **n is the number of the vocabulary or data or observations.**
- ✓ **Ranging from 5 to 25 class.**

C. Find length of class: L= $\frac{R}{m}$

D. Find Class Mark = $\frac{upper\ class\ limit+lower\ class\ limit}{2}$

E. Find real class limit

- ✓ **Upper real class limit = upper class limit +0.5**
- ✓ **Lower real class limit = lower class limit − 0.5**

F. Find the number of frequency

- ✓ Recording the original value one by one in its own class
- ✓ The sum of the frequency for each class is equal to the total number of data or observations.

Example 2.1 Represent the following data in a frequency distribution table.

56 45 67 55 65 70 60 75 83 89 56 65 46 49 65 69 48 62 72 74

Solution

1. **45 46 48 49 55 56 56 60 62 65 65 65 67 69 70 72 74 75 83 89**
2. **R = (max_v-min_v)**

   = 89 – 45

   = 44
3. **m = 2.5*$\sqrt[4]{n}$**

   = 2.5 * $\sqrt[4]{n}$

   = 5.2 ≅ 5
4. **L = $\frac{R}{m}$**

$= \frac{44}{5}$

$= 8.8 \cong 9$

5. Real class limit (R_c)

- ✓ **Lower real class limit (first class) = lower class limit-0.5**

  = 45 – 0.5

  = 44.5

- ✓ **Upper real class limit (first class) = Upper class limit + 0.5**

  = 54 + 0.5

  = 54.5

6. Class Mark (for the first class) = $\frac{upper\ class\ limit + lower\ class\ limit}{2}$

$= \frac{53 + 45}{2}$

= 49

c	f_i	y_i	R_c
45-53	4	49	44.5-53.5
54-62	5	58	53.5-62.5
63-71	6	67	62.5-71.5
72-80	3	76	71.5-80.5
81-89	2	85	80.5-88.5
	20		

H.W 2.1 the following data represent the results of the compressive strength test for fifteen concrete cubes in units (MPa). Arrange the following data in the frequency table:

48 54 79 55 69 74 59 60 75 61 68 81 51 59 73

2.5 Relative Frequency Distribution

Is the table shows the importance of relative frequency of each class, and can be calculated using following method:

✓ Relative Frequency(p_i) = $\frac{\boldsymbol{Frequency\ of\ classes}}{\boldsymbol{sumation\ of\ the\ Frequency\ of\ classes}}$ $\left(\frac{f_i}{\sum f_i}\right)$

✓ Percentage Relative Frequency = $\frac{f_i}{\sum f_i}$ ***100**

2.5 Cumulative Frequency Distribution

There are two kind of cumulative frequency distribution:

- Ascending Cumulative Frequency Distribution
  - ✓ Ascending ($\mathbf{F_i}$) = frequency of class +frequency of the previous class
- Descending Cumulative Frequency Distribution
  - ✓ Descending ($\mathbf{F_i}$) =$\sum f_i$ − frequency of the previous class

Example 2.2: **According to the frequency distribution table shown below, calculate the relative, percentage, and cumulative frequency distribution.**

C	f_i
25-29	7
30-34	19
35-39	14
40-44	7
45-49	3
	$\sum f_i = 50$

Solution

- ✓ $P_i = \frac{f_i}{\sum f_i}$

$= \frac{7}{50}$

$= 0.14$

- ✓ $P_i\% = \frac{f_i}{\sum f_i} * 100$

$= 14$

- ✓ **Ascending (F_i) = frequency of class +frequency of the previous class**

$= 7+0$

$= 7$

- ✓ **Descending (F_i) = $\sum f_i$ − frequency of the previous class**

$= 50-0$

$= 50$

c	f_i	P_i	P_i %	Ascending F_i	Descending F_i
25-29	7	0.14	14	7	50
30-34	19	0.38	38	26	43
35-39	14	0.28	28	40	24
40-44	7	0.14	14	47	10
45-49	3	0.06	6	50	3
	$\sum f_i = 50$	$\sum P_i = 1$			0

Example 2.3: **Represent the following data in a frequency distribution table. Number of classes (m) is 6.**

24 15 22 28 17 12 20 16 23 16 22 15 23 18 11 21 17 30
16 29 19 39 19 18 14 20 28 18 29 24 20 22 34 12 29 25
17 23 20 18 24 14 32 27 18 21 15 19 17 16 21 19 23 26
24 13 23 15 25 22 25 16 18 23 20 10 22 15 18 16 15 26
17 20 21 19 20 21 27 31

Solution

C	f_i	y_i	R_C	P_i	P_i %	Ascending	Descending
10-14	8	12	9.5-14.5	0.1	10	8	80
15-19	28	17	14.5-19.5	0.35	35	36	72
20-24	27	22	19.5-23.5	0.33	33	63	44
25-29	12	27	24.5-29.5	0.15	15	75	17
29-30	4	32	28.5-30.5	0.05	5	79	5
35-39	1	37	34.5-39.5	0.0125	1.25	80	1
	$\sum f_i = 80$						0

H.W 2.2: Below are the data for the grades of students in the examination of the Engineering Statistic

5 4 4 5 3 4 2 3 1 2 3 7 4 1 6 3 2 5 3 4 7
3 2 6 5 3 4 2 4 1

Find the following

- ✓ **Number of classes**
- ✓ **Frequency distribution**
- ✓ **Relative Frequency distribution**
- ✓ **Cumulative frequency distribution.**

2.6 Graphical Representation

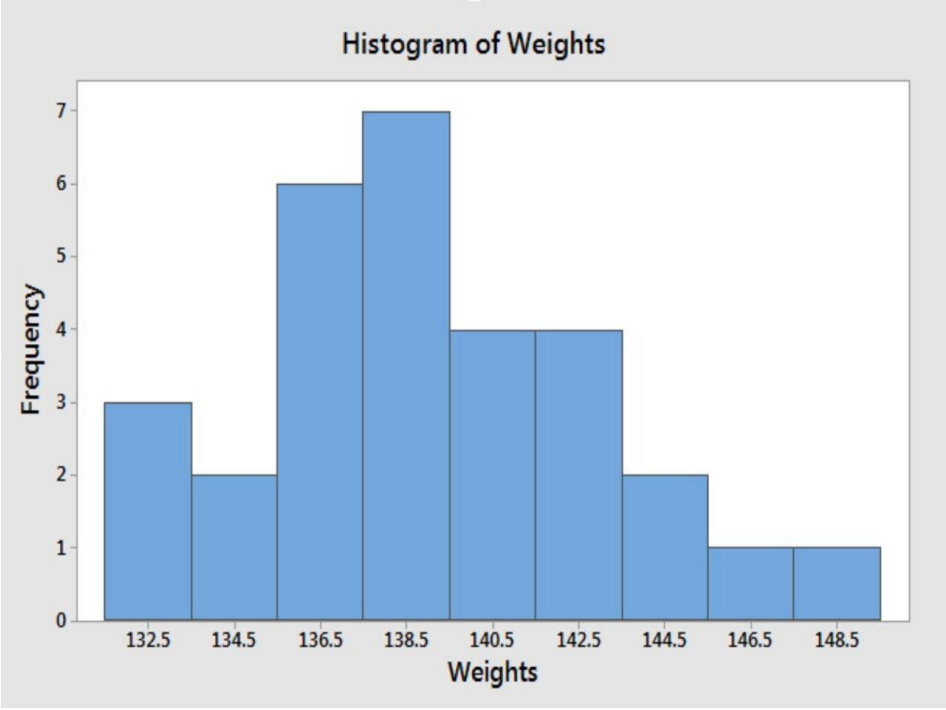
- The most important forms of graphical representation of frequency distribution tables are:

- ✓ Frequency Histogram
- ✓ Frequency Polygon
- ✓ Frequency Curve
- ✓ Cumulative Frequency Curve

- The horizontal axis (x-axis) represents variables and the vertical axis (y-axis) represents frequencies (f_i).

2.6.1 Frequency Histogram

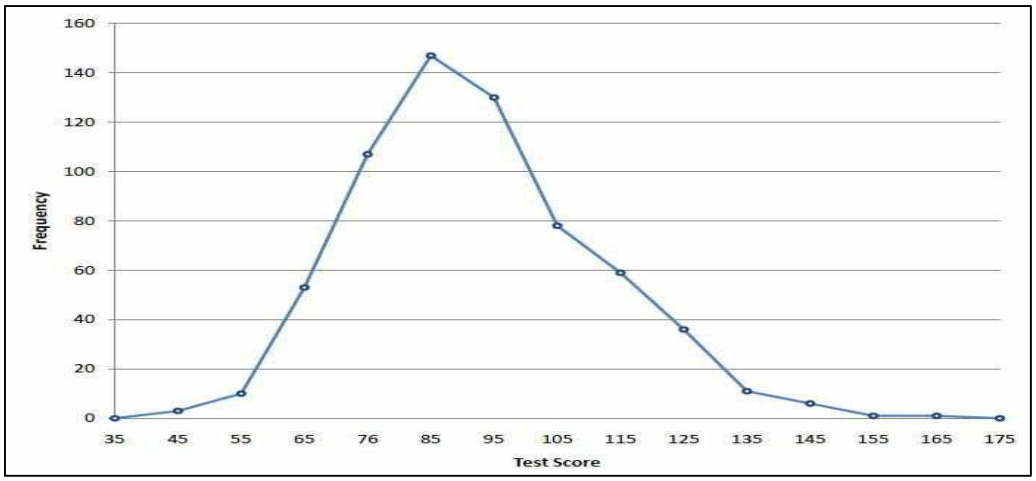
Is a group of rectangles whose bases extend along the horizontal axis representing the lengths of the classes, while its height represents the frequency of classes.



❑ 2.6.2 Frequency Polygon

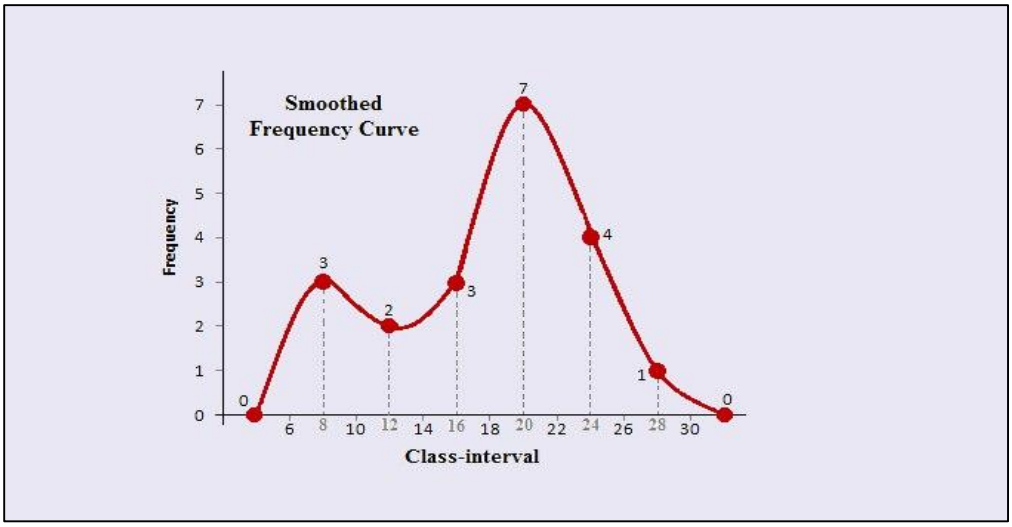
✓ It is a closed polygon that begins and ends with the horizontal axis.

✓ It consists of broken straight lines connecting the center of the classes (class mark) and the frequency (y_i, f_i).



2.6.3 Frequency Curve

It is similar to a polygon except that the lines connecting the points are curved rather than straight



3.6.4 Cumulative Frequency Curve

Is represents by a curved points connecting points located above the real limits of the classes. There are two types of cumulative frequency curve:

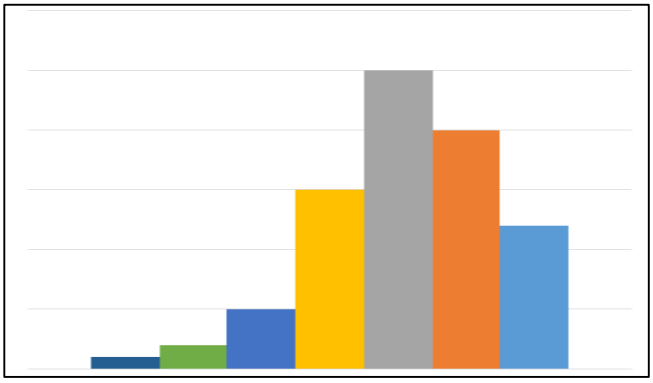
- ✓ **Ascending Cumulative Frequency Curve**
- ✓ **Descending Cumulative Frequency Curve**

Example 2.3: Draw histogram, frequency polygon frequency curve, and cumulative frequency curve for frequency distribution table below:

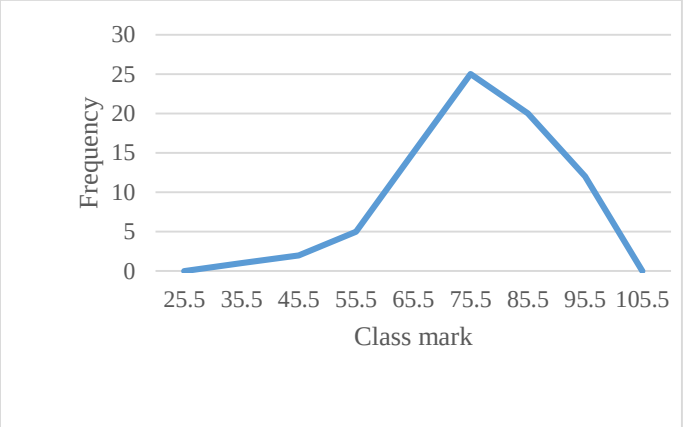
Classes	Frequency	y	AF	DF
31-40	1	35.5	1	80
41-50	2	45.5	3	79
51-60	5	55.5	8	77
61-70	15	65.5	23	72
71-80	25	75.5	48	57
81-90	20	85.5	68	32
91-100	12	95.5	80	12
	$\sum f_i = 80$			0

Solution

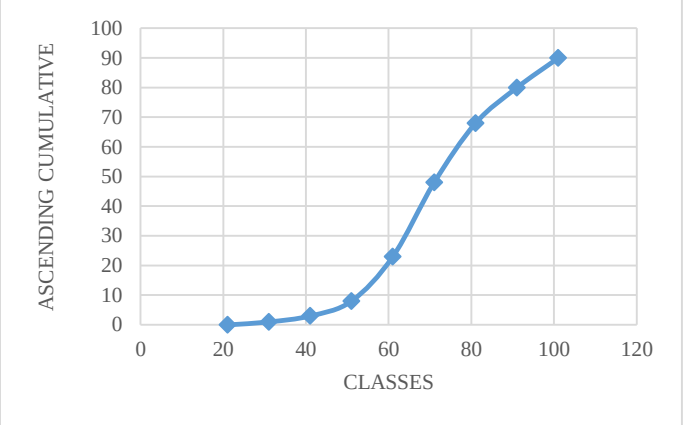
1. **Histogram method**



2. Polygon method



3. Ascending Cumulative Frequency Curve



4. Descending Cumulative Frequency Curve