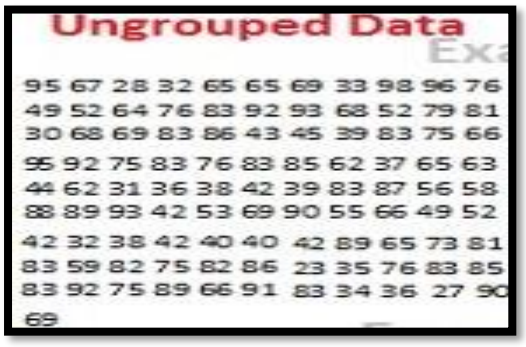
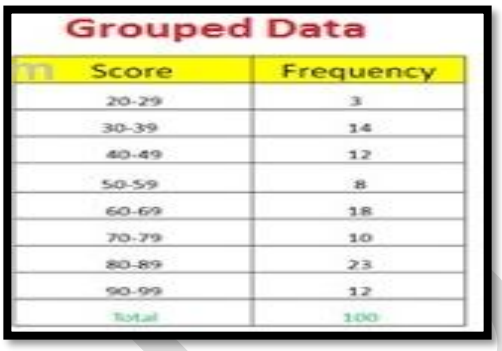
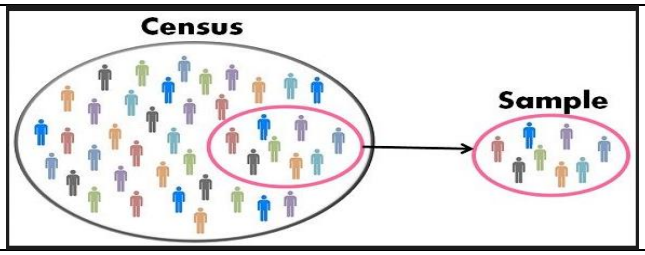
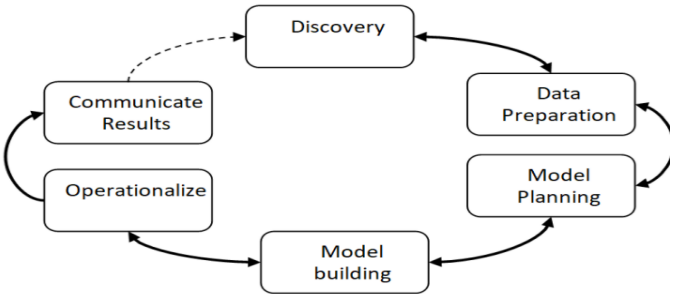
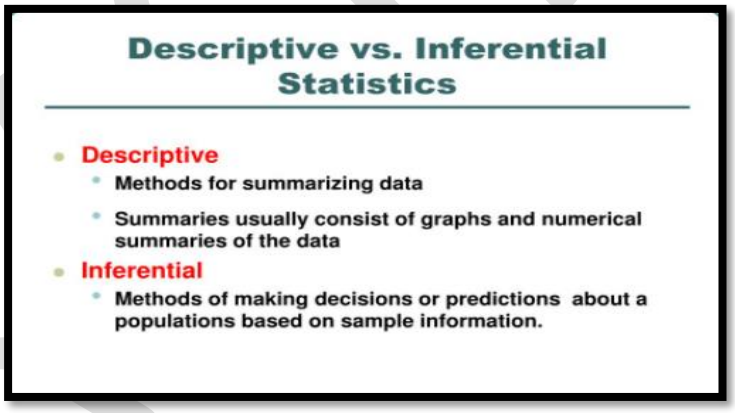
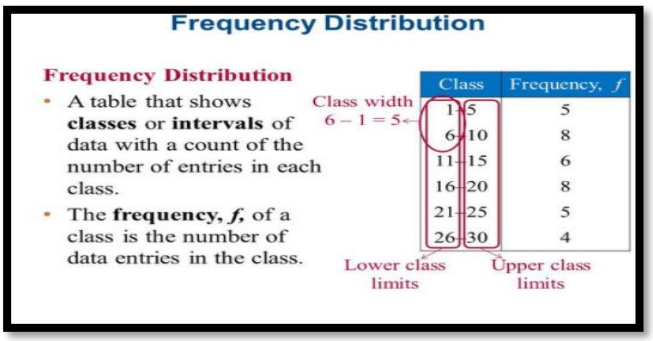
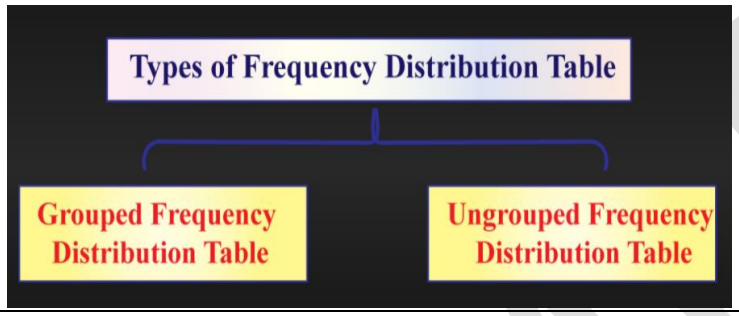
| | |
|---|---|
| 1 | **Define statistics.**<br>**ANS:** Statistics is a science of collecting, presenting, analyzing and interpreting the data as well as of making decision on such analysis. |
| 2 | **Define Data.**<br>**ANS:** Data is a collection of facts, such as numbers, words, measurements, observations or just descriptions of things.<br>  <br>  |
| 3 | **What are the types of data?**<br>**ANS:** • Primary Data • Secondary Data |
| 4 | **What is Primary Data?**<br>**ANS:** Primary data is a type of data that is collected by researchers directly from main sources through interviews, surveys, experiments, etc. |
| 5 | **What is secondary data?**<br>**ANS:** They are the data that are sourced from someplace that has originally collected it. |
| 6 | **What are the types of Primary data?**<br>**ANS:** • Qualitative Data<br>• Quantitative Data |
| 7 | **What is Qualitative data?**<br>**ANS:** If the data is classified on the basis of the qualitative characteristics or attributes is called qualitative data. |
| 8 | **What is Quantitative data?**<br>**ANS:** Quantitative data is defined as the value of data in the form of counts or numbers where each data-set has a unique numerical value associated with it. |
| 9 | **What are the types of secondary data?**<br>ANS: 1. Discrete Data. 2. Continuous Data. |
| 10 | **What is Discrete data?**<br>ANS: Data that can only take certain values. |
| 11 | **Write any two examples of Discrete data?**<br>**ANS:** • Number of children per family<br>• Number of students in a class<br>• Number of citizens of a country |
| 12 | **What is Continuous data?**<br>**ANS:** Continuous data is data that can take any value |
| 13 | **Write any two examples of continuous data?**<br>**ANS:** Height, weight, temperature and length are all examples of continuous data. |
| 14 | **Mention the sources of secondary data.**<br>**ANS:**1. Published sources<br>2. Unpublished sources |
| 15 | **Write any two examples for Published data.** |

| | |
|---|---|
| | **ANS:** 1. Government publications 2. public records<br>3. Historical and statistical documents 4. Business documents |
| 16 | **Write any two examples for Unpublished data.**<br>**ANS:** 1. Personal letters 2. Diaries<br>3. Journals 4. Wills |
| 17 | **Define incomplete Data.**<br>**ANS:** The data that has missing attributes is called an incomplete data. |
| 18 | **Define invalid Data.**<br>**ANS:** Data attributes are not conforming to standardization is called an invalid data.<br>For example, 9-digit phone number records rather than 10 digits. |
| 19 | **Write any two data collection tools of Qualitative data?**<br>**ANS:** Data collection tools are<br>1. Questionnaires 2. Survey 3. Interviews 4. Focus group discussion. |
| 20 | **Define questionnaire.**<br>**ANS:** A questionnaire is a data collection method, in which you present a pattern of questions to a selected group of people. |
| 21 | **What are the different types of questionnaires?**<br>**ANS:** There are two types of questionnaires<br>1.Closed ended questionnaire.<br>2.Open ended questionnaire. |
| 22 | **What is open ended questionnaire? Give examples.**<br>**ANS:** Open-ended questions are broad and can be answered in detail.<br>• Tell me about your college.<br>• How do you see your future?<br>• What is the purpose of Government?<br>• Why did you choose that answer?<br>• Tell me about your friends in this paragraph. |
| 23 | **What are closed ended questionnaire? Give examples.**<br>**ANS:** closed-ended questions are narrow in focus and usually answered with a single word or a pick from limited multiple-choice options.<br>• Are you feeling better today?<br>• May I use your phone?<br>• Will you please do me a favor?<br>• Have you already completed your homework?<br>• Is that your final answer? |
| 24 | **Assume that you are collecting the feedback from the customer in restaurant, list any 5 topics on which questionnaire can be prepared to collect the same.**<br>**ANS:** i)quality of food<br>ii)cost<br>iii)timing<br>iv)cleanliness<br>v)service<br>vi)staff friendliness<br>vii)atmosphere |
| 25 | **Assume that you are collecting the information regarding survey on facilities in college, list any 5 topics on which questionnaire can be prepared to collect the same.**<br>**ANS:** <u>Facilities in college:</u><br>1)classroom facility |

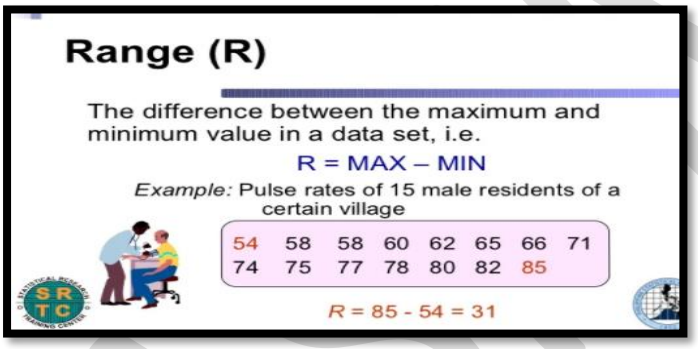
| | |
|---|---|
| | 2)laboratories<br>3)infrastructure<br>4)placement<br>5)seminar hall<br>6)transportation<br>7)washroom |
| 26 | **What is survey?**<br>**ANS:** Survey is a method of gathering information from a sample of people, traditionally with the intention of generalizing the results to a larger population. |
| 27 | **Mention any five advantages of survey.**<br>Following are the important advantages of survey.<br>• Relatively easy to administer.<br>• Can be developed in less time<br>• Cost-effective, but cost depends on survey mode.<br>• Capable of collecting data from a large number of respondents.<br>• Numerous questions can be asked about a subject, giving extensive flexibility in data analysis. |
| 28 | **Mention any five disadvantages of survey.**<br>Following are the important disadvantages of survey.<br>• Respondents may not feel encouraged to provide accurate, honest answers.<br>• Respondents may not feel comfortable providing answers that present themselves in a unfavorable manner.<br>• Respondent may not be fully aware of their reasons for any given answer because of lack of memory on the subject.<br>• Data errors due to question non responses may exist.<br>• Survey question answer options could lead to unclear data because certain answer options may be interpreted differently by respondents. |
| 29 | **Define interview.**<br>**ANS:** An interview is a face-to-face conversation between two individuals with the sole purpose of collecting relevant information to satisfy a research purpose. |
| 30 | **Explain the different types of interviews?**<br>**ANS:** There are 3 types of interviews.<br>**Structured interviews:** The questions are predetermined in both topic and order.<br>**Semi-structured interviews:** A few questions are predetermined, but other questions aren't planned.<br>**Unstructured interviews:** None of the questions are predetermined. |
| 31 | **What is group Discussion?**<br>**ANS:** A discussion involving a number of people who are connected by some shared activity, interest, or quality. |
| 32 | **Define internal data.**<br>**ANS:** An internal data is information generated from within the business, covering areas such as operations, maintenance, personnel and finance. |
| 33 | **Define external data.**<br>**ANS:** External data comes from the market, including customers and competitors. |
| 34 | **What is census?**<br>**ANS:** A census is a survey conducted on the full set of observation objects belonging to a given population or universe. |
| 35 | **What is sample?**<br>**ANS:** Sampling is a process used in statistical analysis in which a predetermined number of observations are taken from a larger population. |

| | |
|---|---|
| |  |
| 36 | **What is data cleaning?**<br>**ANS:** Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted. |
| 37 | **Identify the different phases of data analysis.**<br>**ANS:**<br>  |
| 38 | **What are the types of statistics?**<br>**ANS:** • Descriptive statistics • Inferential statistics |
| 39 | **What are descriptive statistics?**<br>**ANS:** Descriptive statistics is a way to organize, represent and describe a collection of data using tables, graphs, and summary measures. |
| 40 | **What are inferential statistics?**<br>**ANS:** Inferential Statistics is a method which allows us to use information collected from a sample to make decisions, predictions or inferences from a population.<br>  |
| 41 | **What is frequency?**<br>**ANS:** The number of times the value occurs in the data. |
| 42 | **Define data tabulation.**<br>**ANS:** Data tabulation is a process of systematic arrangement of the classified data in rows and columns, in the form of table. |
| 43 | **What is frequency distribution?**<br>**ANS:** Frequency distribution is a systematic presentation of the values taken by a variable along with their frequencies. Frequency distribution is a systematic presentation of the values taken by a variable along with their frequencies. |




44 **What is range?**
**ANS:** It is the difference highest and lowest value in the data.
i.e., Range = highest value - lowest value




45 **What is class limit?**
**ANS:** The lowest and highest values which are taken to define the boundaries of a class are class limits.

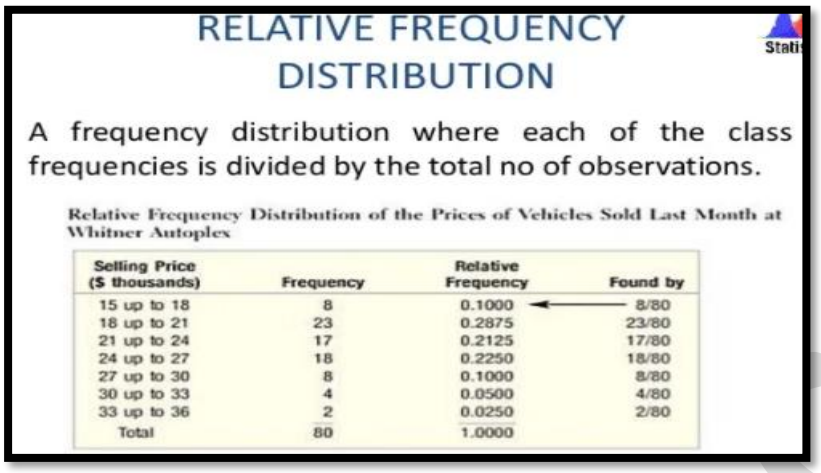
46 **Write the two types of class with example.**
**ANS:** 1. Inclusive class 2. Exclusive class

47 **What is Inclusive class?**
**ANS:** In a class, if lower as well as upper limits are included in the same class, such a class is called Inclusive class. Here, upper limit of a class is not equal to the lower limit of the next class.
Ex. 0-9, 10-19, 20-39…. Are inclusive classes.

48 **What is Inclusive class?**
**ANS:** In a class, If the lower limit is included in the same class and upper limit is excluded from that class but included in the next class, such a class is called Exclusive class. Here, upper limit of a class is equal to lower limit of the next class. Ex: 30-40, 40-50, 50-60 are exclusive classes.

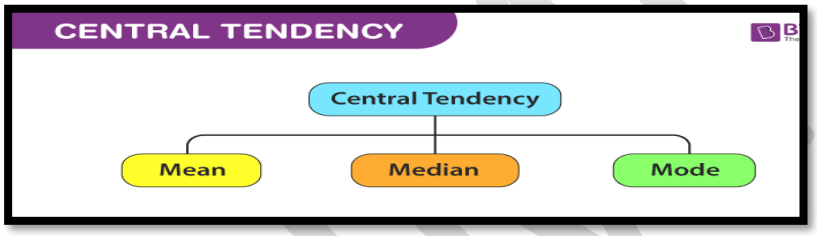
49 **What is relative frequency?**
**ANS: Relative frequency is the ratio of frequency of the value of the variable to the total frequency.**

$$\text{i.e. Relative frequency (R.f)} = \frac{frequency\ of\ the\ value\ of\ the\ variable}{total\ frequency}$$




50 **What is the total of relative frequency?**
**ANS:** The total of relative frequency is 1.

51 **What is central tendency?**
**ANS:** The tendency for the values of a random variables to cluster round its mean, mode or median.

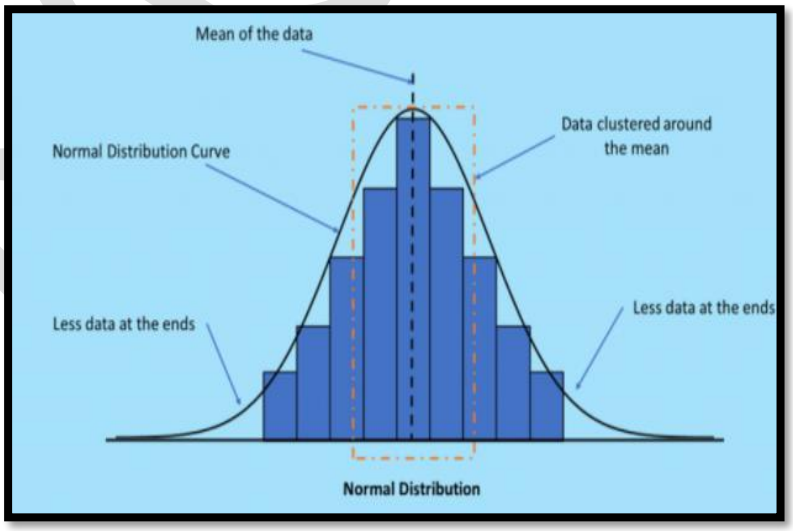


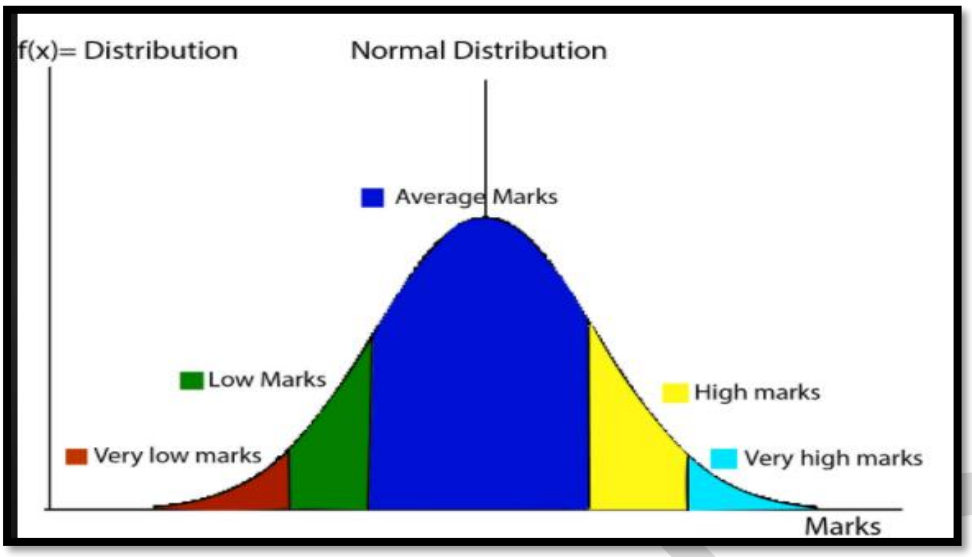

52 **Define mean.**
**ANS:** It is the average of a data set.

53 **Define mode.**
**ANS:** The mode is the most common number in a data set.

54 **Define median.**
**ANS:** The median is the middle of the set of numbers.

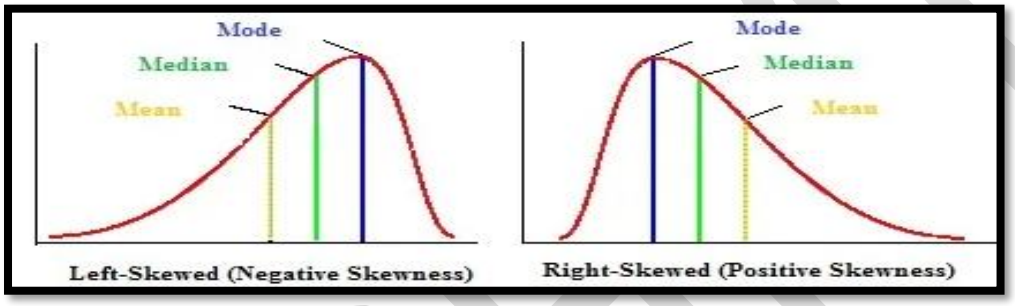
55 **What is normal distribution?**
**ANS:** A function that represents the distribution of many random variables as a symmetrical bell-shaped graph.

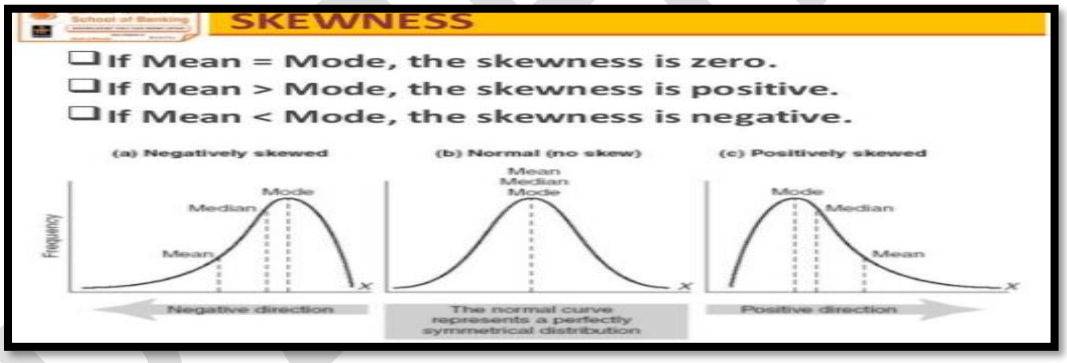





56 **Define skewness.**

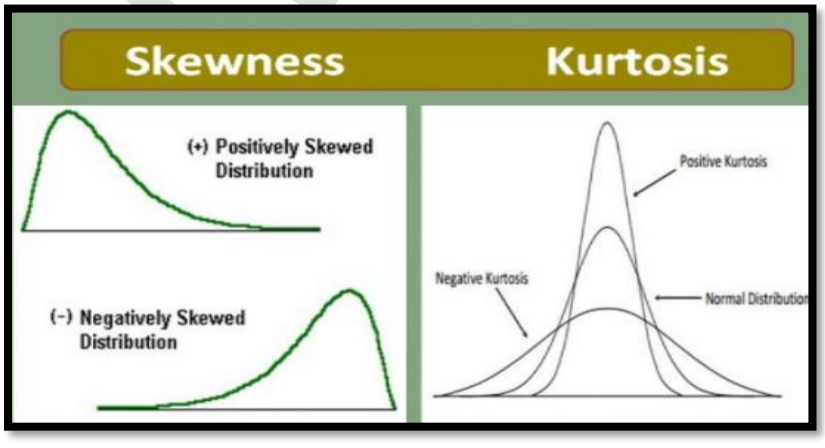
**ANS:** Skewness is a measure of the symmetry of a distribution.

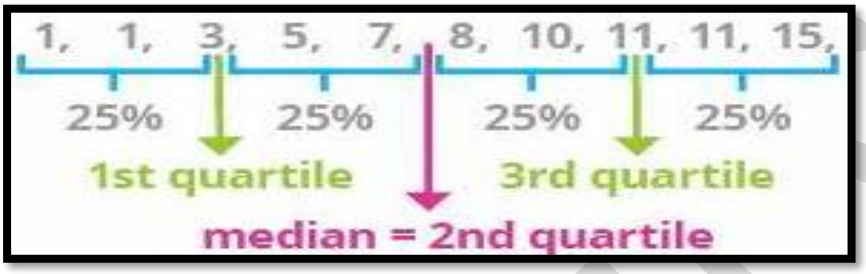
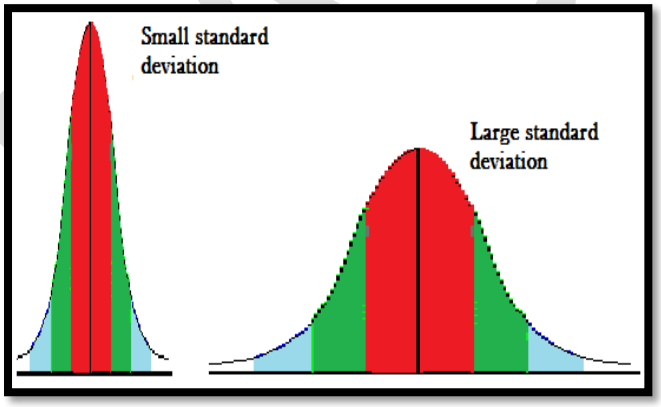
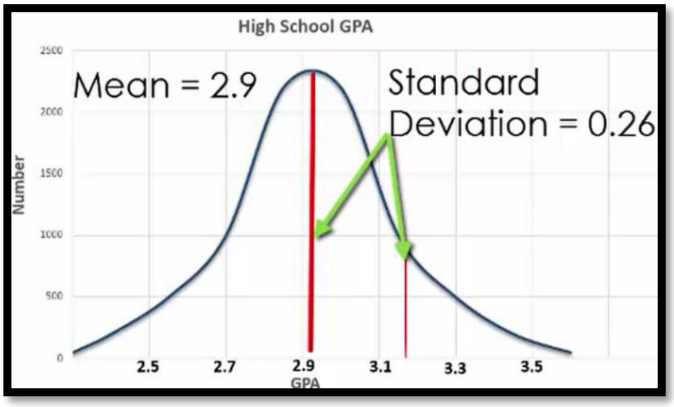



57 **Define kurtosis.**

**ANS: The sharpness of the peak of a frequency distribution curve.**



| | |
|---|---|
| 58 | **Define quartie-1.**<br>**ANS:1st quartile or lower quartile Q1 basically separate the lowest 25% of data from the highest 75%.** |
| 59 | **Define quartie-2.**<br>**ANS: 2nd quartile or middle quartile Q2 also same as median it divides numbers into 2 equal parts.** |
| 60 | **Define quartie-3.**<br>**ANS: The upper or third quartile, denoted as Q3, is the central point that lies between the median and the highest number of the distribution.**<br>  |
| 61 | **Define variability.**<br>**ANS:** Variability refers to the spread of the score within a distribution, along with the central tendency, it helps in understanding the data set as a whole.<br>There are four major measures of variability:<br>• Range<br>• Interquartile range<br>• Variance<br>• Standard deviation |
| 62 | **Define interquartile range.**<br>**ANS:** It is the difference between the 75th and 25th percentile. |
| 63 | **Define variance.**<br>**ANS:** It is the degree of spread within the distribution (the larger the spread, the larger the variance). |
| 64 | **Define standard deviation.**<br>**ANS:** It is the measure of how the average score deviates or spreads away from the mean (defined as the square root of the variance).<br>  <br>  |
| 65 | **Define quartile deviation.**<br>**ANS:** Quartile deviation defines the absolute measure of dispersion.<br>QUARTILE DEVIATION=(Q3-Q1)/2 |
| 66 | **Write the syntax for generating quartile one.**<br>**ANS:** QUARTILE-1=QURATILE (array,1) |
| 67 | **Write the syntax for generating random between.**<br>**ANS:** RANDOM BETWEEN=RANDBETWEEN (bottom, top) |

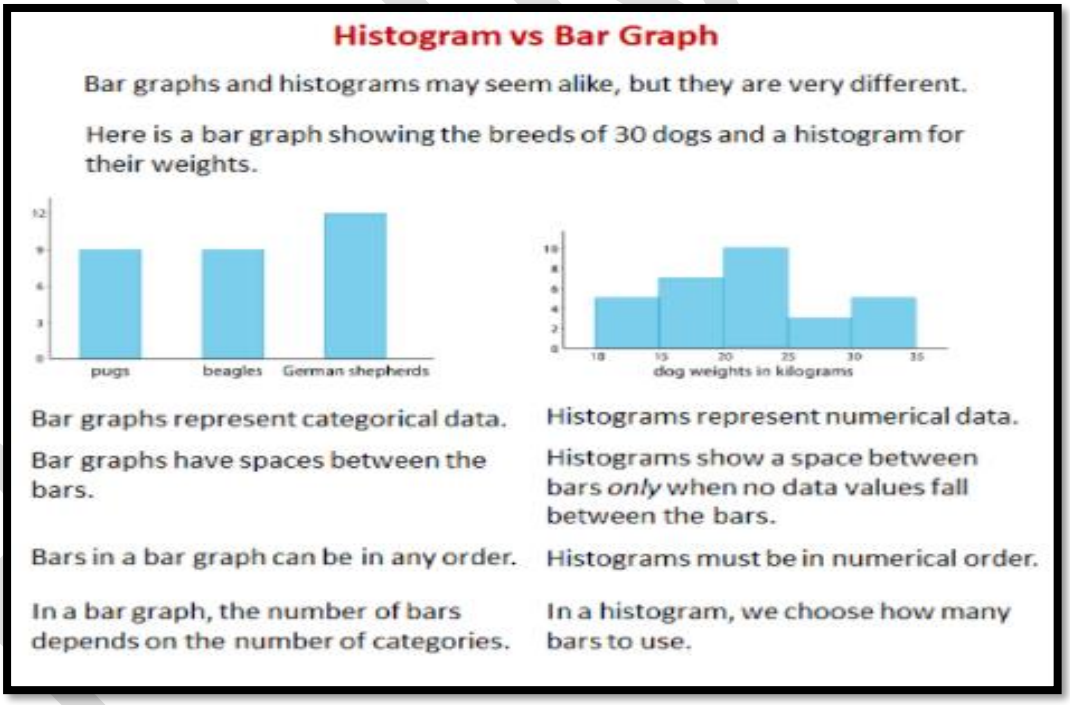
68 **Write the syntax for print in python.**
**ANS:** print ()

69 **Why python is so competent over the programming language?**
**ANS:** Python is simple to use and easy to learn because it makes you focus on the application of programming, not on missing semicolons in the code.

70 **Write the difference between pie graph and bar graph.**
**ANS:**

| No | Pie Chart | Bar Chart |
|---|---|---|
| **1** | Pie charts are circular pictures showing proportions of many elements. A wedge represents every element. | Bar charts, on the other hand, are rectangular bars that are plotted horizontally and vertically on axes of different heights for representing categorical data. |
| **2** | A pie chart can show only a few values without segregating the slices from the data it is representing. | Through this, we will be able to check the subject value. |
| **3** | It displays the proportion of specific data by angle scale of arc length. | For instance, we can check the amount of rainfall in a particular city for several years. It compares discrete categories. |
| **4** | It depends on textures, colors and arrows for segregating the slices. | On a bar chart, one axis will display the categories and the other will show the measured value. |

71 **Write the difference between histogram and bar graph.**
**ANS:**




72 **The syntax for following functions is:**
**mean, mode, median and standard deviation.**
**ANS:** MEAN=AVERAGE (array of numbers)
MODE=MODE (number1, [number2],…)
MEDIAN=MEDIAN (number1, [number2],…)
STANDARD DEVIATION=STDEV.P((number1, [number2],…)

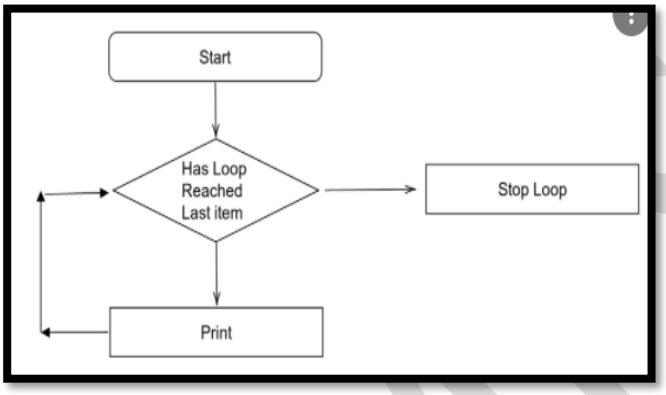
73 **What is inter-quartile range and quartile deviation?**
**ANS:** Inter quartile range is the difference between the 75th and 25th percentile.
Quartile deviation defines the absolute measure of dispersion.

74 **List the libraries in python.**

**ANS:** 1. pandas-Pandas is a data library for data analysis.
2. NumPy-NumPy can be used to perform a wide variety of mathematical operations on array.
3. matplotlib-Matplotlib is a cross-platform, data visualization and graphical plotting library for Python.

75 **What is a for loop in python and draw the flow chart of for loop?**

**ANS:** for loops are used when we have a block of code which we want to repeat affixed number of times.
Flow chart of python:




76 **Who developed Python program?**

**ANS:** Python program was developed by Guido Van Rossum.

77 **Define syntax of python.**

**ANS:** The syntax of the python programming language is the set of rules that defines how a python program will be written and interpreted.

78 **Write syntax to print the data type of the variable 'X'.**

**ANS:** print(X)

79 **Explain any 3 features of python programming.**

**ANS:** i) Easy to learn and readable language.
ii) interpreted language
iii) Free and open source
iv) High level language
v) Large community support

80 **Explain any 3 python arithmetic operators and their functions.**

**ANS:** Arithmetic operators are used to perform mathematical calculation.

| OPERATOR | NAME | EXAMPLE |
|---|---|---|
| + | Addition | x+y |
| - | Subtraction | x-y |
| * | Multiplication | x*y |
| / | Division | x/y |