1. Difference between Data and Information?

Data

Data is the collection of individual facts, statistics, or items of information.

In raw form, unprocessed and unstructured

Might be meaningless on its own

Relies on data

Comes in the form of No, Figures, and statistics

Might be difficult to understand.

Information

Information is data that is processed, organized and structured.

Processed and Structured

Always meaningful

Doesn't rely on the information

Comes as words, thoughts and ideas

Relatively easy-to-understand information

2.How is data useful for us?

Better decision making.

Better Targetting

Effective Marketing

Better Customer Relations

Competitive advantage.

Data helps us to optimise business performance, forecast future results, understand the audiences and Reduce the costs.

Good data allows organizations to establish baselines, benchmarks, and goals to keep moving forward.

3.What is big Data?

Big data refers to the data that is soo large, fast or complex that it is difficult or impossible to process using the traditional methods.

Its is analysed computationally to reveal patterns,tre nds, and associations, especially relating to human behaviour and interactions.

4. Difference between Structured, Semi-Structured and Unstructured Data?

Structured data is generally tabular data that is represented by columns and rows in a database.

EG. No. dates and strings

Unstructured data is information that either does not organize in a pre-defined manner or not have a pre-defined data model.

Eg. Images,audio,video,email,spreadsheets.

Semi-structured data is information that doesn't consist of Structured data (relational database) but still has some structure to it.

Eg.JSON

5.What is Quantitively and Qualitative Data?

Quantitively:

Data collection is unstructured.

It asks why.

It cannot be computed as the data collected is no statistical.

It develops initial understanding and defines the problem.

Qualitative:

Data collected is structured.

It is all about how much and how many.

It is statistical and is all about numbers.

It recommends the final course of action.

6.What are the different V's in Big Data?

Volume

Value

Veracity

Variety

Velocity

7.Name some popular tools used in Big Data?

Apache Spark.

Apache Hadoop.

Apache Flink.

Google Cloud Platform.

MongoDB.

Sisense.

RapidMiner.

8.What are the different types of Data? Explain

i.Quantitative: Data that can be measured by numbers eg speed

1.Discrete: Whole Nos that cannot be broken down further eg. no. of items.

2. Continuous:No. that can be further broken down eg.height,weight.

a. Interval: Nos. which have a fixed distance between them.

b.Ratio:Numbers which have measurable intervals where the difference can be determined.eg.height

ii.Quantitative: Non numerical data that is categorical eg. Yes/ no

1. Nominal: Data used for naming variables, such as hair color.

2. Ordinal: Data used to describe the order of values such as rating.