**Latency**

**Latency** is a measure of delay. In a network, latency measures the time it takes for some data to get to its destination across the network.

**Firewall**

A Firewall is a network security device that monitors and filters incoming and outgoing network traffic

**ARN**

**Amazon Resource Names**

Amazon Resource Names (ARNs) uniquely identify AWS resources.

**Policies**

A policy is an JSON object in AWS that defines permissions on AWS resources

**IOPS**

IOPS measures the **number of read and write operations per second**, ---> number of operation you doing per sec

**Throughput**

throughput measures the **number of bits read or written per second**. ----> number of operation you doing per sec

**Durability**

99.999999999%

Durability is a measurement of the **tiny errors that occur in files due to these underlying media**. When you write, read, and rewrite gigabytes, terabytes, and petabytes of information to the same drive, one or more individual bytes can **get corrupted or lost.**

**Structured**

Structured data is highly specific and is stored in a predefined format,

Feeback -->

Title, date Desc

<String> <Date> <String>

**Unstructured**

where unstructured data is a conglomeration of many **varied types of data** that are stored in their native formats

**Feeback**

Desc

**ETL -**

ETL stands for “**extract, transform, load**,” the three processes that, in combination, move data from one database, multiple databases, or other sources to a unified repository—typically a data warehouse.

**Serverless**

Serverless is a cloud computing execution model in **which the cloud provider allocates machine resources on demand**, **taking care of the servers on behalf of their customers**.

"Serverless" is a misnomer/mislead in the sense that servers are still used by cloud service providers to execute code for developers.

**Data science --> ML / AL / ->**

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from noisy, structured and unstructured data, and apply knowledge and actionable insights from data across a broad range of application domains. **Wikipedia**

**Data Analytics -> E, T, L --> Vistal**

Data analysis is a process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. **Wikipedia**

**Goal to find insights from data and get custome behaviour ---> Can take business decision**

**Apache Hadoop / Hadoop**

**Apache Hadoop is a collection of open-source software utilities that facilitates using a network of many computers to solve problems involving massive amounts of data and computation**. It provides a software framework for distributed storage and processing of big data using the MapReduce programming model. Wikipedia

**Data lake**

A data lake is a vast pool of raw data, the purpose for which is not yet defined.

**Data warehouse**

A data warehouse is a repository for structured, filtered data that has already been processed for a specific purpose.