

## **Introduction to stream concepts**

A data stream is an existing, continuous, ordered (implicitly by entrance time or explicitly by timestamp) chain of items. It is unfeasible to control the order in which units arrive, nor it is feasible to locally capture stream in its entirety. It is enormous volumes of data, items arrive at a high rate.

### **Types of Data Streams :**

**Data stream** –A data stream is a(possibly unchained) sequence of tuples. Each tuple comprised of a set of attributes, similar to a row in a database table.

**Transactional data stream** –It is a log interconnection between entities

Credit card – purchases by consumers from producer

Telecommunications – phone calls by callers to the dialed parties

Web – accesses by clients of information at servers

**Measurement data streams** –

Sensor Networks – a physical natural phenomenon, road traffic

IP Network – traffic at router interfaces

Earth climate – temperature, humidity level at weather stations

### **Examples of Stream Sources-**

#### **Sensor Data –**

In navigation systems, sensor data is used. Imagine a temperature sensor floating about in the ocean, sending back to the base station a reading of the surface temperature each hour. The data generated by this sensor is a stream of real numbers. We have 3.5 terabytes arriving every day and we for sure need to think about what we can be kept continuing and what can only be archived.

#### **Image Data –**

Satellites frequently send down-to-earth streams containing many terabytes of images per day. Surveillance cameras generate images with lower resolution than satellites, but there can be numerous of them, each producing a stream of images at a break of 1 second each.

## **Internet and Web Traffic –**

A bobbing node in the center of the internet receives streams of IP packets from many inputs and paths them to its outputs. Websites receive streams of heterogeneous types. For example, Google receives a hundred million search queries per day.

### **Characteristics of Data Streams :**

- Large volumes of continuous data, possibly infinite.
- Steady changing and requires a fast, real-time response.
- Data stream captures nicely our data processing needs of today.
- Random access is expensive and a single scan algorithm  
Store only the summary of the data seen so far.
- Maximum stream data are at a pretty low level or multidimensional in creation, needs multilevel and multidimensional treatment.

### **Applications of Data Streams :**

- Fraud perception
- Real-time goods dealing
- Consumer enterprise
- Observing and describing on inside IT systems

### **Advantages of Data Streams :**

- This data is helpful in upgrading sales
- Help in recognizing the fallacy
- Helps in minimizing costs
- It provides details to react swiftly to risk

### **Disadvantages of Data Streams :**

- Lack of security of data in the cloud
- Hold cloud donor subordination
- Off-premises warehouse of details introduces the probable for disconnection

