**Data and Data Flow**

Data stays at core of System Design.

At **Business Layer**/or as user, you interact with data in the form of **text, videos, images**, etc.

At **Application layer**/or in code, we interact with data in certain formats like **JSON, XML**.

When stored in **database**, data is in form of **tables, indexes, lists, Tree**s, etc.

At **network layer**, data is transferred in form of **packets**.

At **hardware** level, it’s just **0s and 1s**.

Once you understand the data flow inside the system and the type in which it will be represented at each layer, then our journey of designing system is halfway done.

**Data Store Types**

* Database

Ex: to store address, phone number, name, city, etc.

* Queues

Ex: to send sms request, send email request.

* Indexes

Ex: most searched item, items searched in last 1 hour

* Cache

Ex: to store request/response

**Data flow methods**

* APIs
* Messages
* Event handling

**Data Generation**

* Users: data input by user
* Internal data: logs, meta data
* Insights: analytics done on stored user data, which provides insights.

**Factors to consider about data in system**

* **Types of data**

Knowing the type of data which you will operate on is important. A system dealing with video data is different from the one dealing with text data. This helps in deciding on which database to use.

* **Volume of data**

Size of data in TBs or GBs.

* **Consumption/Retrieval**

It basically means, there can be systems which write and read heavily. Some maybe only write heavy. Some maybe only read heavy

* **Security**

Some applications like banking need heavy security. Its ok if you are not able to login quickly but data should not leak.

**Types of system and their data**

**Authorization System**:

Volume of data might not be that high, as it only needs to store some necessary credentials. But level of security must be very high. Only correct user must be given correct privileges.

Ex: Identity management, User Login

**Streaming System**

Volume of data will be high and Consumption of Data will also be high.

Ex: Hotstar, Netflix, Prime Video

**Transactional System**

Transaction must follow ACID property. Read and write can be heavy (mostly write).

Ex: Ecommerce, Banking System, Ride booking apps.

**Heavy Compute System**

Volume of data is huge. Need to run sophisticated algos. No one reads data but uploads the dat and does processing with ML algo, etc.

Ex: Image Recognition, Video Processing, etc.