What is Different Data?

➢ **Data**
- **Quantitative data.**

  - It can be measured or given ==numerical values.==
  - like dates, times, weights, heights
- **Qualitative data**

  - It refers to things that can be ==observed but not definitively measured==
  - Like blood types, brands of car, product reviews, names, and eye colors (While qualitative data may contain numerical ==values, these are usually descriptors==).

- **structured data**
- typically categorized as quantitative data.
- is highly organized and easily decipherable by machine learning algorithms.
- **Structured data tools:**
  - ➢ MySQL
  - ➢ SQLite
  - ➢ PostgreSQL
- **Unstructured data**
- typically categorized as qualitative data.
- cannot be processed and analyzed via conventional data tools and methods.
- unstructured data does not have a predefined data model, it is best managed in non-relational (NoSQL) databases.
- Another way to manage unstructured data is to use data lakes to preserve it in raw form.
- The importance of unstructured data is rapidly increasing.
- unstructured data is over 80% of all enterprise data.

- *Unstructured data tools:*
  - ➢ *MongoDB*
  - ➢ *Hadoop*
  - ➢ *Azure*
- **semi-structured data**
- *Semi-structured data (e.g., JSON, CSV, XML) is the "bridge" between structured and unstructured data.*
- *It does not have a predefined data model and is more complex than structured data, yet easier to store than unstructured data.*
- *Semi-structured data uses "metadata" to identify specific data characteristics.*