

Data Structures

Notes by Mannan UI Haq (BDS-3C)

Structured, Semi-Structured, and Unstructured data are terms used to describe different types of information based on their organization and format.

Structured Data:

Structured data is like information clearly arranged in rows and columns, just like a spreadsheet. It's highly organized and follows a fixed format. Examples of structured data include databases, spreadsheets, and tables.

Customer				
CustomerId	Name	EmailAddress	Gender	EmailVerified
1	Jack Frost	jfrost@winter.com	Male	1
2	Miss Piggy	queen@muppets.com	Female	1
3	Dr. Octopus	doc@octopus.net	Male	0

Invoice			
Invoiceld	CustomerId	Amount	DateCreated
1	1	80	2010-12-11 04:19:12
2	2	24.95	2011-01-05 16:35:56
3	1	25	2011-01-07 20:05:33
4	1	45	2011-02-20 08:09:42

Semi-Structured Data:

Semi-structured data is a bit like a mix between structured and unstructured data. It doesn't fit neatly into rows and columns, but it has some level of structure. Think of it as a collection of documents where each document might have a title, author, and date, but the content itself might not follow a strict structure. Examples of semi-structured data include XML files and JSON files.

```
1  {
2    "EMPLOYEES": {
3      "SALES": {
4        "648229": {
5          "NAME" : "Olivia Johnson"
6          "DOB"  : "1989-08-08"
7        },
8        "648666": {
9          "NAME" : "Frank Mueller"
10         "DOB"  : "1985-05-11"
11         "MISC" : "On paternal leave from 2019-01-01 until 2020-01-01"
12       }
13     }
14   }
15 }
```

JSON

Unstructured Data:

Unstructured data is like a bunch of information without a specific order. It's more like the freeform text you find in a book or a social media post.

Examples:

- Images
- Videos
- Speeches