Day ② :- RDBMS

RDBMS: Defines Create & maintains data bases.

Types of DB's:

operational DB
Data warehouse
Relational DB
~~Cloud DB~~ Distributed DB
End - user DB

Data cleaning:

* Identify inaccuracies in data sets & resolve them.

* 1st step of Data Analysis.

Importance:

* Increases productivity.
* Avoid complications while processing data.
* Improves data quality & accuracy.

* Enables data Integrity
* optimizes resources usage

Data cleaning 8 - steps:

① Remove Irrelevant data.
② Remove Duplicate data.
③ Fix structure errors.
④ Do type conversion.
⑤ Handle missing data.
⑥ Deal with outliers
⑦

① Profiling & assessment.
② Data validation & filtering.
③ Fixing missing data.
④ Standardization & transformation.
⑤ Removing Duplicates.
⑥ Correcting error.
⑦ Handling outliers
⑧ Data Integrity checks & constraints.

dbt

```
┌──────────┐   Extract   ┌──────────┐        ┌──────────────┐
│  Data    │ ──────────> │  Data    │ ──────>│  Analytical  │
│  Source  │    Load     │ Warehouse│        │    tools     │
└──────────┘             └──────────┘        │      &       │
                                             │ Applications │
                                             └──────────────┘
```