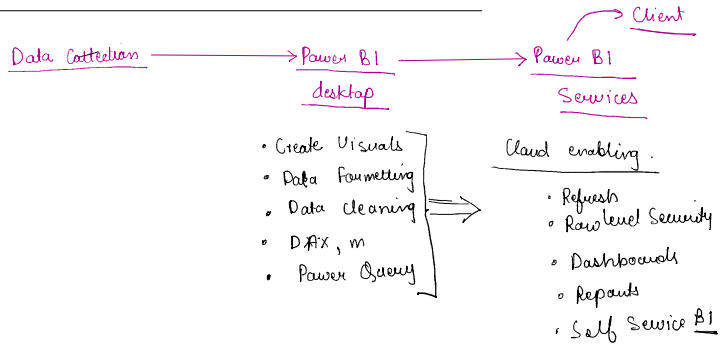
DAY 2

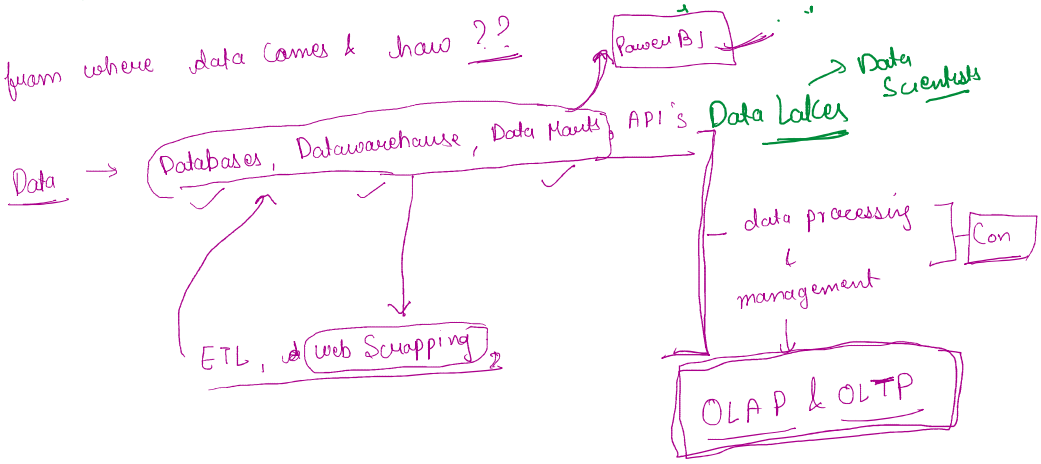
13 July 2024 10:58

Understand Entire Power BI flow



DATA COLLECTION

3 forms of data → Structured data → Row & Columns [Tabular].
→ Unstructured data → Text, Video, Image.
→ Semi-Structured data → XML, JSON



Databases

→ Structured collection of data generally in the tabular form.
[Rows & Columns]

Current
→ Transaction [day to day] frequent ; basic query

- C → Create
- R → Read
- U → Update
- D → Delete

→ Large number of frequent & basic queries → Process.

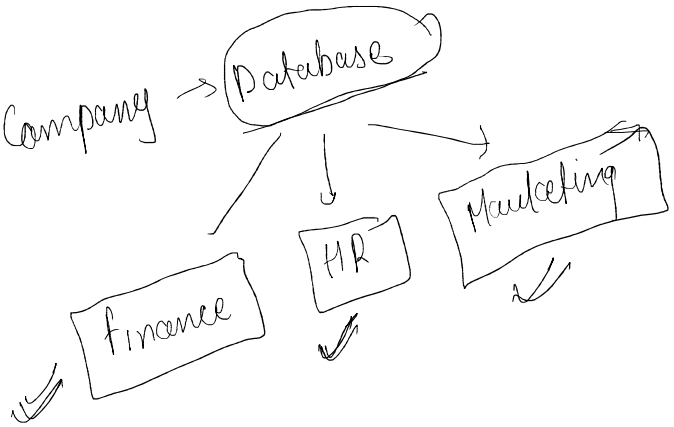
Datawarehouse.

- It is a centralised repository for storing data from multiple sources.
- Current + historical
- Used for analytical, business Intelligence, strategy purposes
- It is designed to handle complex queries
  → Aggregations
  → Joins

Data Mart

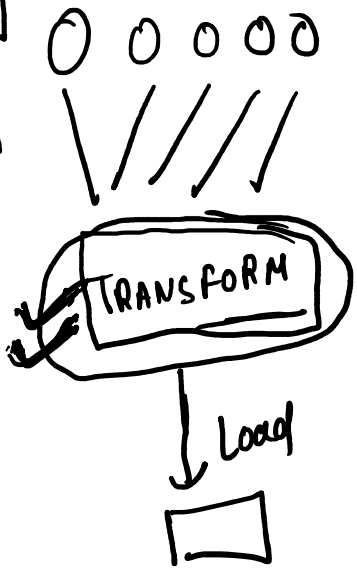
A subset of Datawarehouse, that stores information about a particular Subject.

eg → Company → Database



How data is collected ??

① ETL → E → Extract [csv, mails, API, Unstructured].
T → Transform → [convert → needed.]
L → Load [Datawarehouse, Database]



② web Scrapping → website, data Scrap →
↓
Unstructured

③ Manual Entries →

④ IOT Devices (Internet of things) Devices.

OLAP & OLTP

OLAP — OLTP

↓

Conceptual framework [data processing + management].

OLTP (online Transactional Processing)

- Handle routine, day to day operations.
- Transactional tasks.
- Simple, straightforward, queries [CRUD] → Creation, Reading Updation, deletion.
- It stores current data, operational data.
  ↓ design
  large volume of small trans. handle.
- It prioritizes low-latency responses, [command] → quickly.

OLAP [online Analytical Processing]

- Complex queries & multidimensional Analysis.
  - Business Intelligence, data Analytical etc.

- Complex queries ⟶ handles easily
  [aggregations, calculations, logics, merging, joining etc].
- Historical data stores + Current data & feed.
  ⟶ data warehousing.

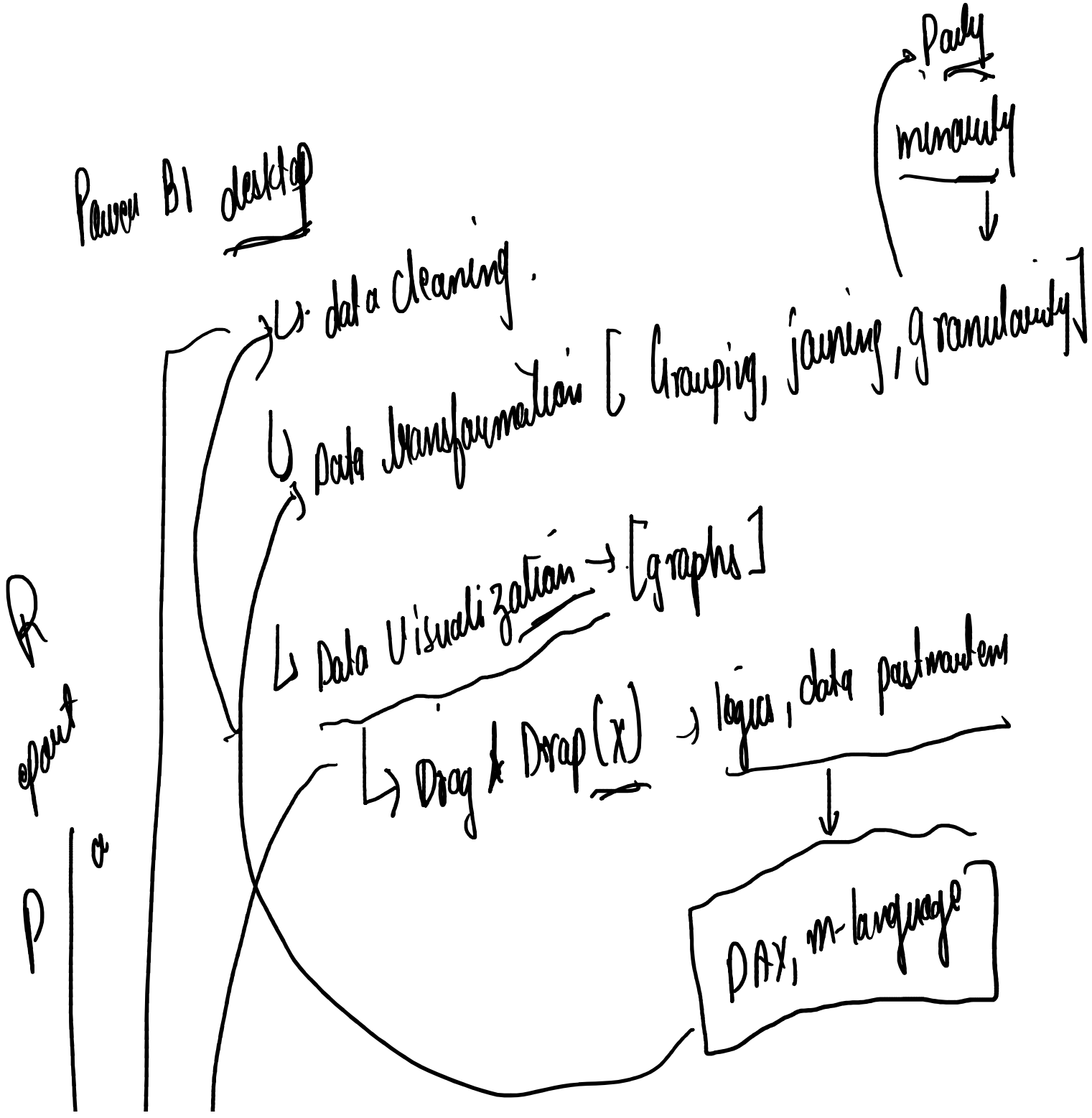
- Response time is slow ⟶ * data → Correct.
  + As per requirement.

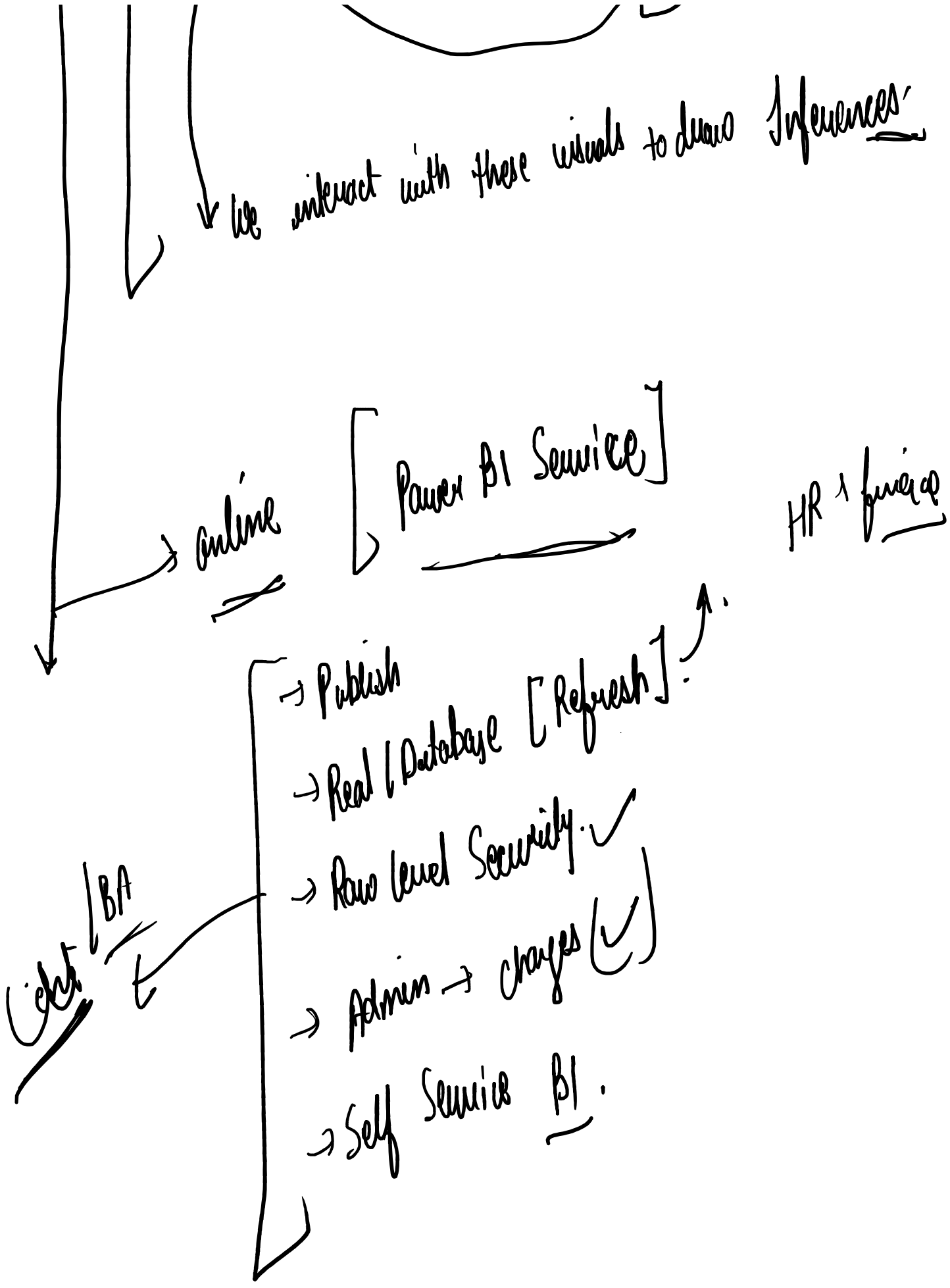
ATM → OLTP

Hardware → ATM
→ Display
→ Card Reader

Software → Windows / Linux
→ Mastercard UI
→ Access Database

OLAP → Power BI


Power BI desktop
Power query
manually
↳ data cleaning.
↳ Data transformation [Grouping, joining, granularity]
↳ Data Visualization → [graphs]
↳ Drag & Drop (x) → logics, data restructuring
DAX, M-language


We interact with these visuals to draw Inferences.
online
[Power BI Service]
HR + finance
Publish
Real time Database [Refresh].
Row level Security.
Admin → changes
Self Service BI.
client
BA