

Week 1: Data Analysts meet?

The Six Data Analysis Phase



1

1) Ask

- to define a problem to be solved.
- to help focusing on the actual problem and avoid any distractions.
- to make sure you fully understand the stakeholder's expectations.

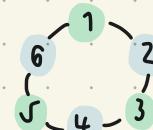
2) Prepare

- what metrics to measure?
- locate data in your database.
- create security measures to protect that data (policy agreement)



3) Process

- to clean up your data (get rid of possible errors, inaccuracies, and inconsistencies)
- to use spreadsheet functions to find incorrectly entered data.
- to use SQL functions to check for extra spaces.



4) Analyze



- to perform calculations.
- to combine data from multiple sources.
- to create tables with your results.

5) Share



- to make more informed decisions via Graphs or Dashboards.
- to share results with stakeholders.

6) Act



- to act on your data.
- to provide your stakeholders with recommendations based on your findings.

* * Six Problem Types * *

Data Analysts typically work with these:

① Making Predictions

→ Vegetable

② Categorizing things → assigning items to categories

③ Spotting sth. unusual

→ မျမှော်ကြုံ

④ Identifying themes → grouping them into broader themes (In a User study, examples of themes)

⑤ Discovering connections

⑥ Finding patterns

* SMART Questions *

Avoid asking questions that:

✗ close-ended Questions (answer with yes/no)

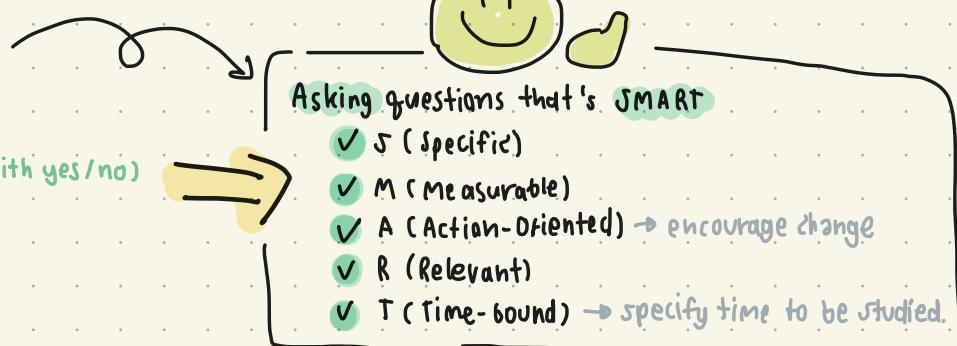
✗ vague & lacks content

✗ Leading Questions ??

Bias,
Unfair

to same
groups in

✗ Making assumptions



* * SMART Question *

Week 2: Data, Report, Dashboard

Small Data

- specific metrics, short defined time

- usually organized and analyzed in spreadsheets.

Big Data

- large, less-specific, cover longer time period

- Usually organized and analyzed in databases.

* Needs to be broken into smaller pieces in order to be organized & analyzed effectively

- Volume: Amount of data

- Variety: Different kinds of data

- Velocity: How fast data can be processed

- Veracity: Quality and Reliability of the data

Some data analysts consider a fourth V

Data

Report vs. Dashboard

Pros

- represents high-level historical data
- easy to design
- pre-cleaned & sorted data

Cons

- Lack of continual maintenance
- No visual appealing
- Static

Pros

- Dynamic, Interactive
- Suitable when sharing information across many peoples promptly / real-time

Cons

- Labor-intensive design (a lot of effort)
- Can lead to misunderstanding (if it's not well-designed)

Metric

a single, quantifiable type of data that is used for measurement.

Return on Investment (ROI)

cost of an investment to the net profit over a period of time.

Week 3: Spreadsheets

① Organize (Sort & filter) e.g. Pivot table

② Calculate data using operators

→ helps doing calculations faster e.g. sum, average, min, max • Data Ranges 02:14

• Cell References (C2 E3 D8)

* lock cell • Absolute Referencing (\$) such as \$A\$10

• operators (+, -, ×, ÷) • Formulas is a set of

• Errors e.g. #DIV/0! instructions to perform calculation

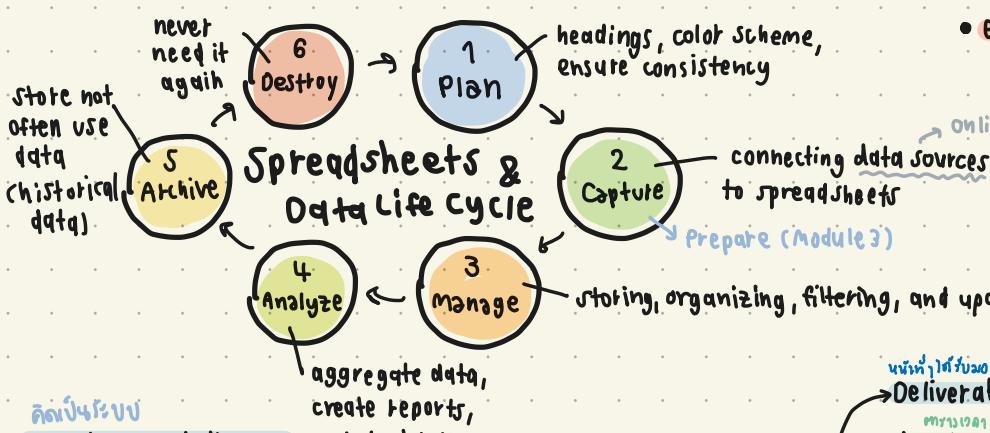
• #ERROR! (parsing error) → not correctly place the other e.g. sum

• #NAME? → wrong naming on function

• #REF! → accidentally deleted row/col

• Functions e.g. SUM(), is a preset command

to perform a specific process



Structured thinking

problem Domain → Specific area of analysis that encompasses every activity affecting / affected by the problem.

→ Need to understand this first! ** before discovering the story. (whole image)

Ask clarifying questions, Define what to accomplish, Specify project boundaries

Ask "Who, what, when, where, why," to put information into context

"Data has little value if it is not paired with context."



Turn raw data into meaningful info.

Week 4: Communication



→ stakeholder → hold stakes on what you are doing, invest time & resources to a project.

e.g. vice president, Executive (CEO)

→ Executive team → provides strategic and operational leadership to the company.

→ Primary stakeholders

→ Customer-facing team → who interacts with customers.

Company ↔ CFT ↔ customer

→ Data science team

fix by be objective

Linkage between team, organize overall proj & team "Project Manager"

**

→ Clear Communication to avoid confusion & Before communicate: prob

- who your audience is?
- what they already know?
- what they need to know?
- How you can communicate that effectively to them?

Secondary stakeholders (also project manager)

Be straight on when you can give them work with reasons

Meeting

- Presentation, Technology used
- stay focused & be on time & Prepare in advance
- pay attention and ask for clarification if needed
- Take Notes, provide takeaway messages
- Have Engagement in meeting

Sample Agenda

Your name	Data Analysis Project
Phone	October 6, 2020 9:30 - 10:30 PST
Email	Group Meeting Room 1

Meeting attendees: Elon, Doe, Oliva, Khi, Pedro

Reason for meeting: Project orientation. Set goals and draft timelines for the project.

Goals

- Read the meeting agenda
- Review project goals
- Plan project timelines

Questions

- Does anyone have any suggestions for the agenda?
- What sources of data have been identified and which variables will be tracked?
- What is the earliest milestone the team can schedule? What progress would the milestone mark?

Next steps

- What should we address in the next meeting?

like what you did in Com Strat