

Exploring Python for Data Analysis

Gusti Triandi Winata

Community Researcher of RL & LLM
Cohere for AI

About Me



Gusti Triandi Winata

Latest Work Experiences:

- | | |
|---|----------------|
| • Researcher, Cohere for AI | 2024 - present |
| • MLE Consultant, Freeport Indonesia | 2023 - 2024 |
| • Mid. MLE, eFishery | 2022 - 2023 |
| • OSS Fellowship with Adobe, Major League Hacking | 2021 |

Education:

- | | |
|---|-------------------|
| • Bandung Institute of Technology <i>Bachelor of Electrical and Computer Engineering</i> <i>Was a Bangkit Graduate of the first batch (2020)!</i> | Graduated at 2021 |
|---|-------------------|

Ground Rules

Observe the following rules to ensure a supportive, inclusive, and engaging classes



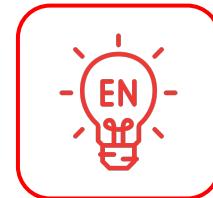
Give full attention
in class



Mute your microphone
when you're not talking



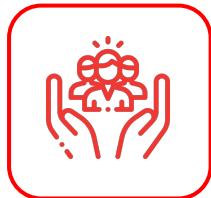
Keep your
camera on



Turn on the CC Feature
on Meet

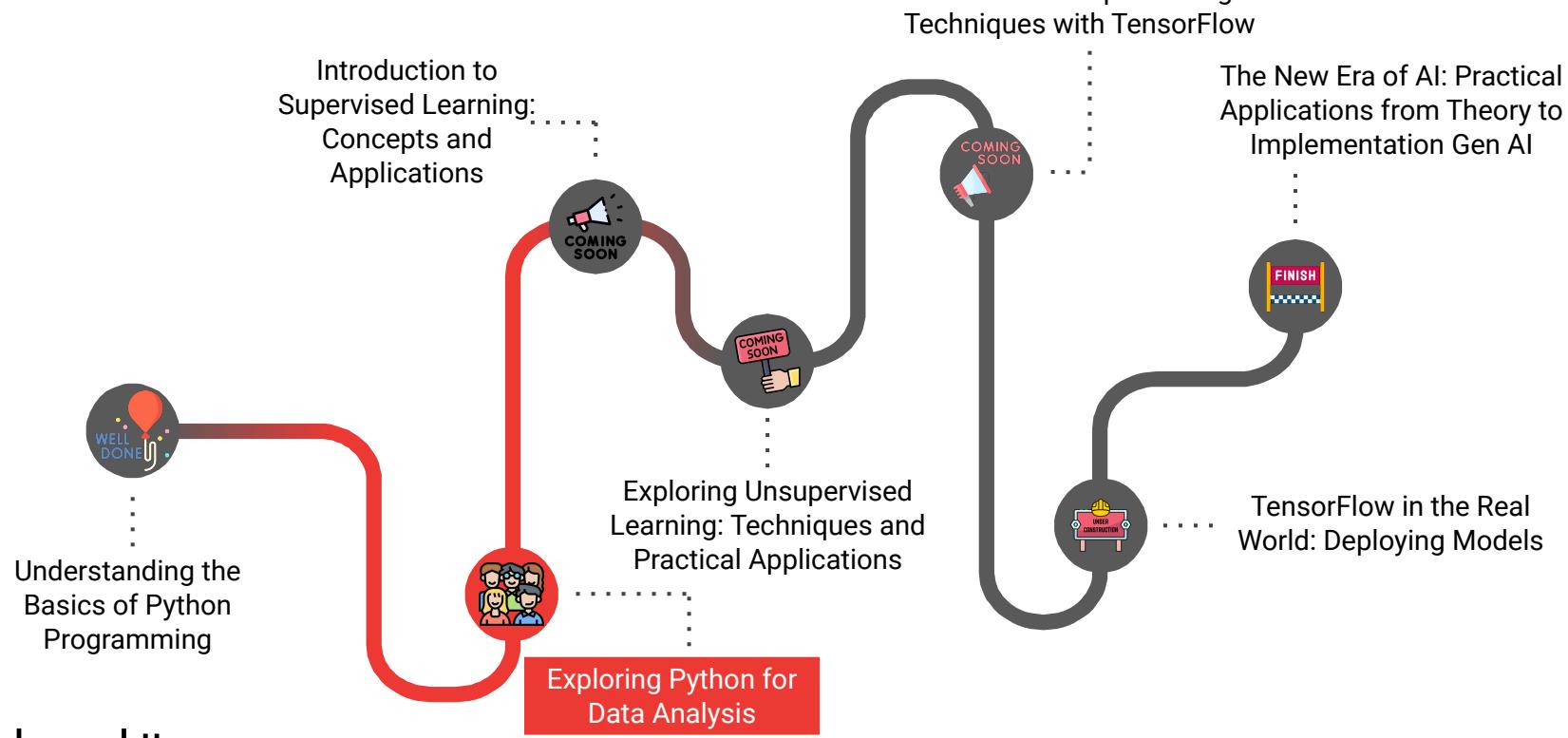


Use raise hand or chat
to ask questions

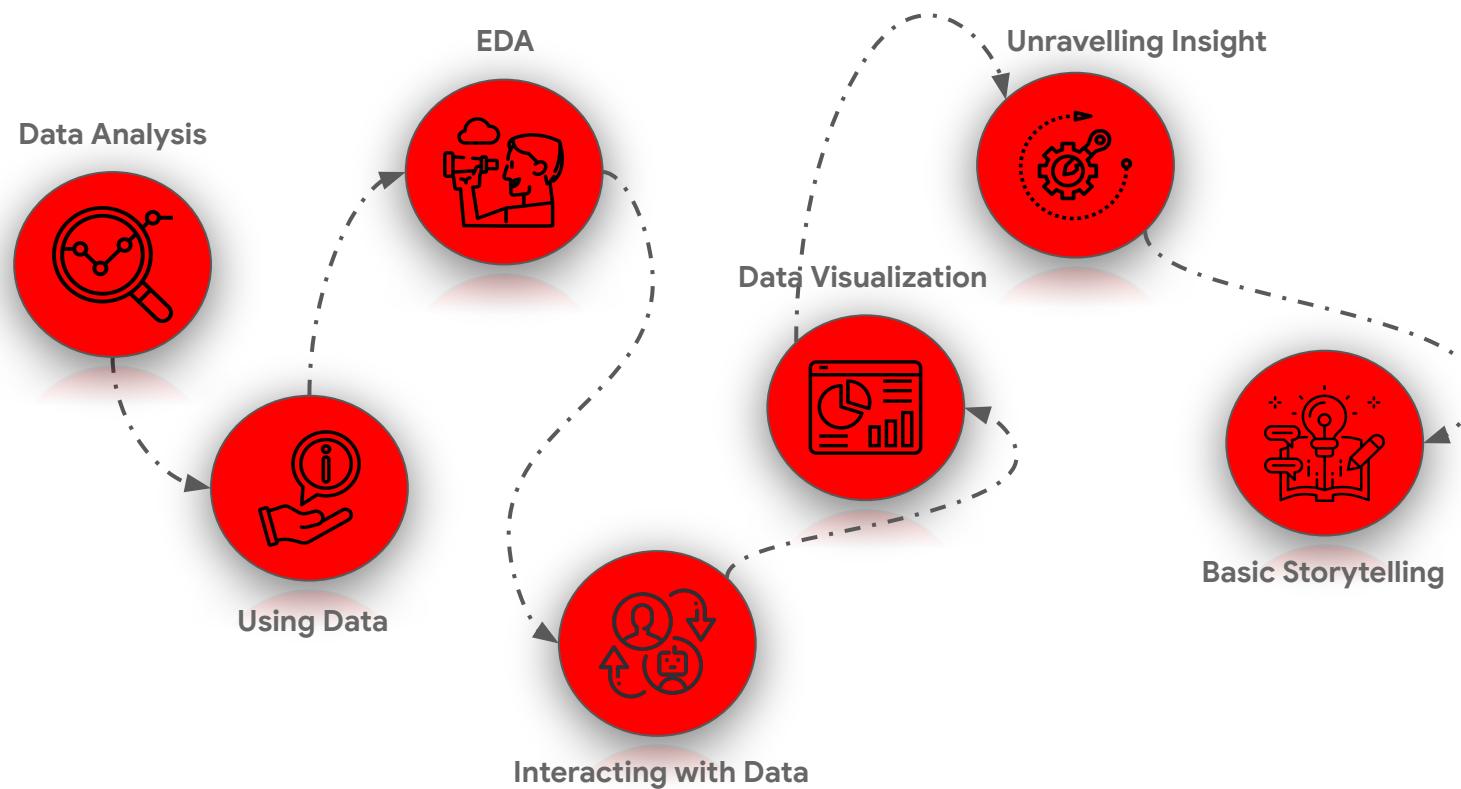


Make this room a safe place
to learn and share

Where Are We Now?



Today's Agenda Roadmap



Learning Objectives

1. **Identifying** the basic concepts in data analytics.
2. **Outlining** the processes of data analysis.
3. **Breaking** down various techniques on data wrangling.
4. **Examining** the aspects of becoming a data storyteller

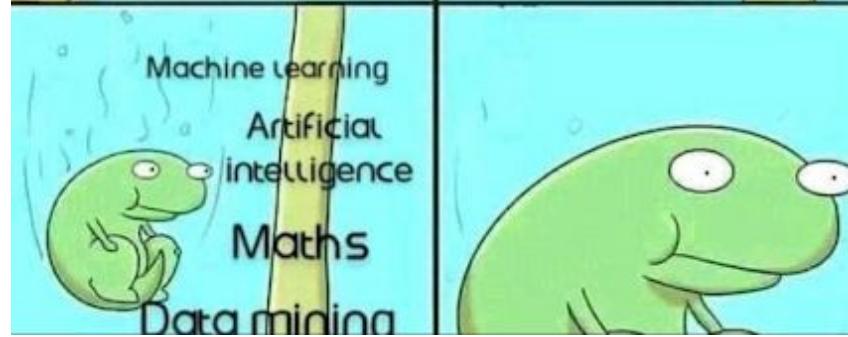


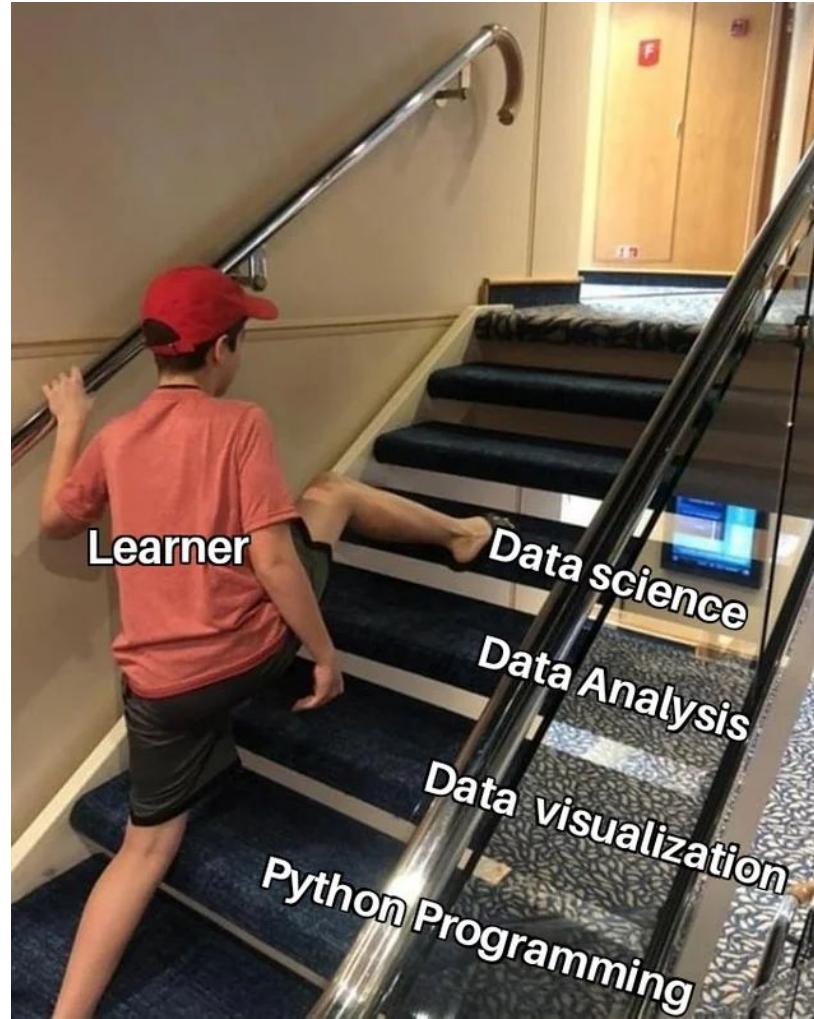
When you switch from another language to Python





Easy to learn
Beginner friendly



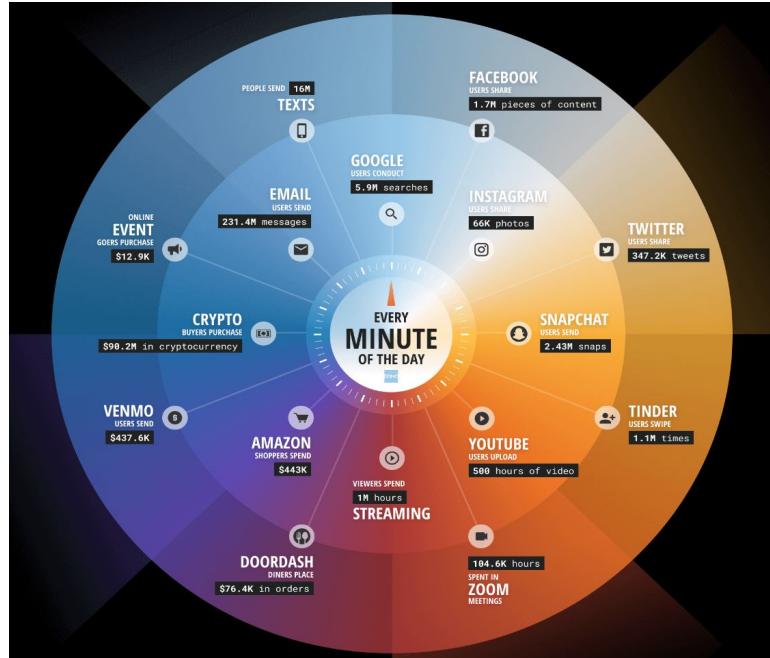


Data Analysis



What is Data?

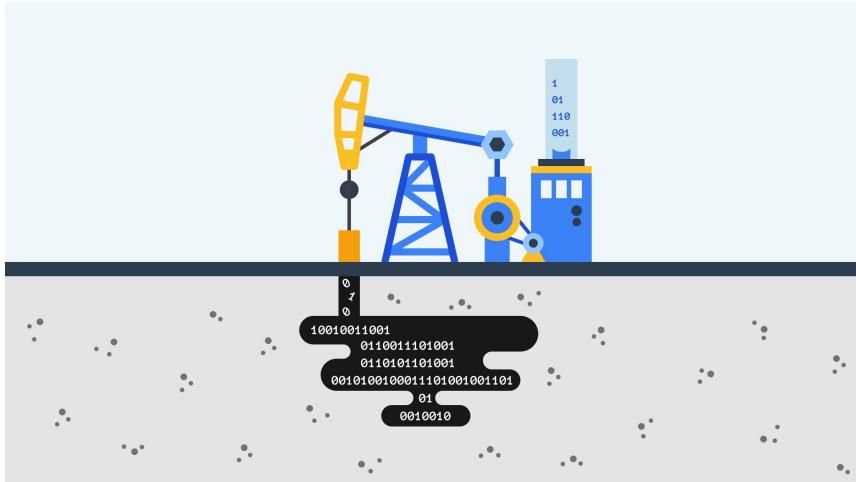
Data is a **collection of facts** such as numbers, pictures, videos, words, measurements, observations, and more.



[Data Never Sleeps 10.0 | Domo](#)

Why Data?

Data is the new oil, it's valuable but if unrefined it can't really be used. It has to be changed into profitable activity, so must data be broken down, analyzed for it to have value.

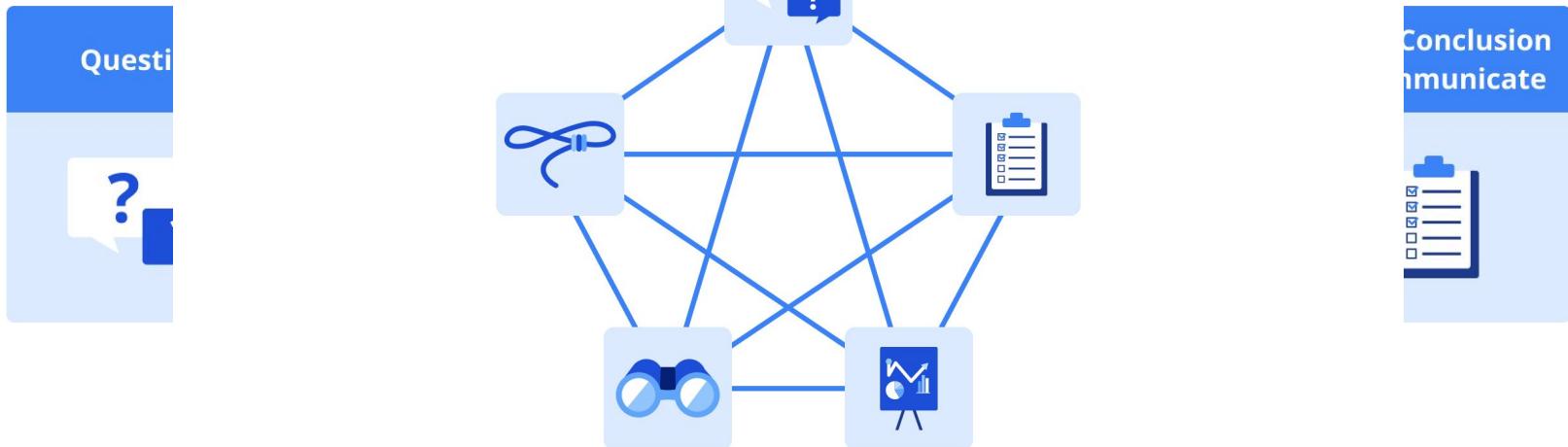


Data Analysis

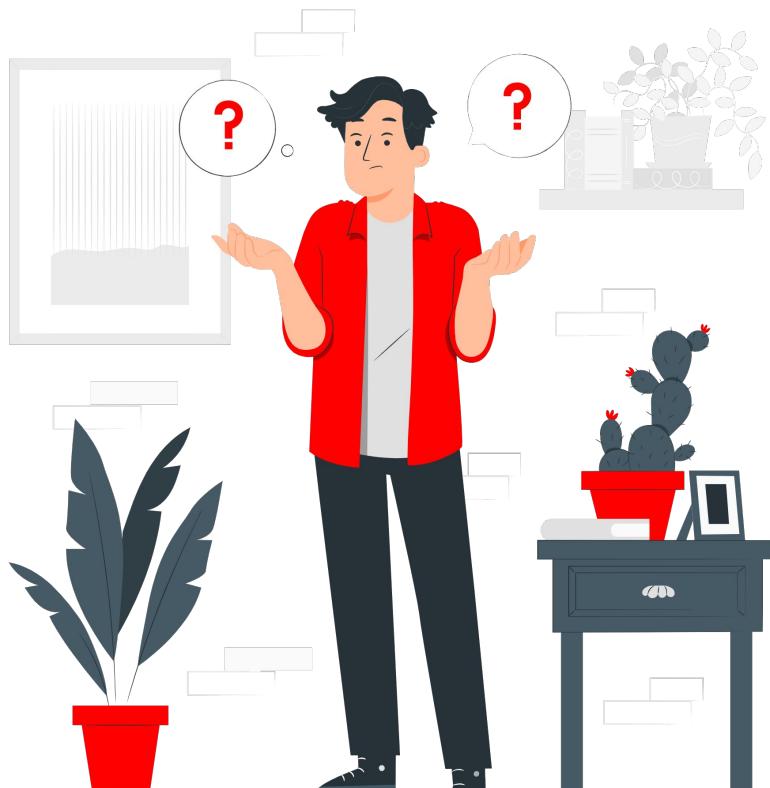
A process of **collecting, transforming, and organizing data** to draw conclusions, make predictions, and provide appropriate considerations in making decisions.



Data Analysis Cycle



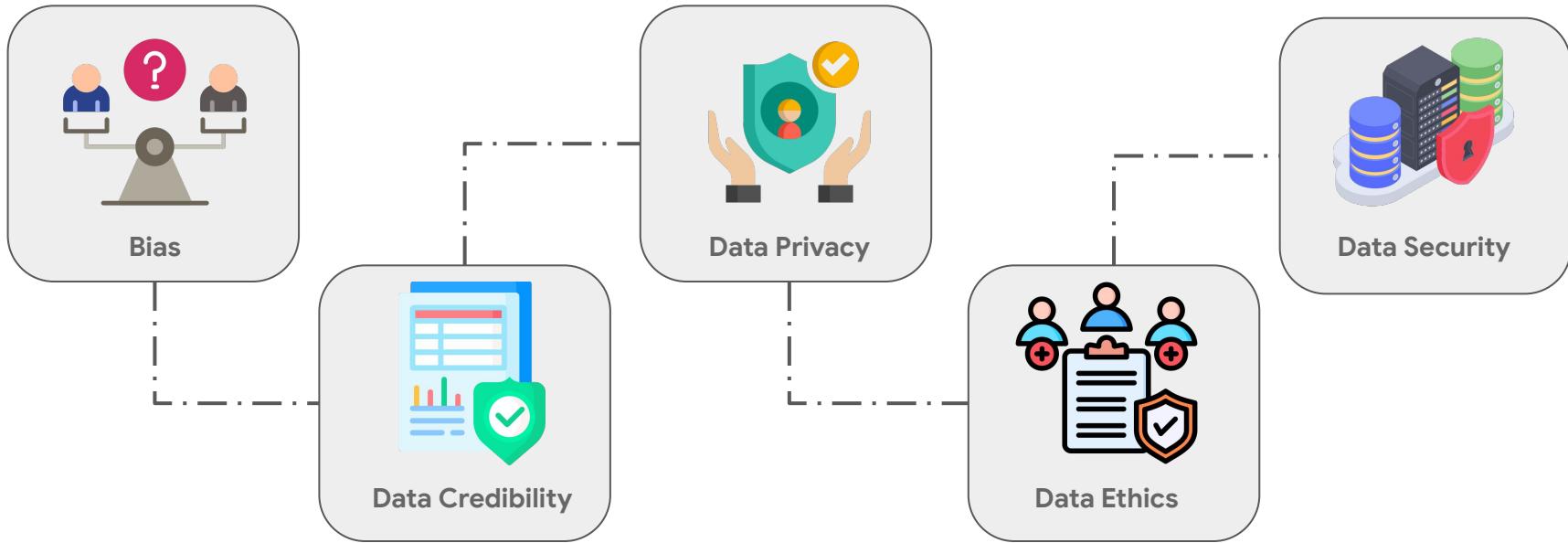
Where do we start?





Using Data

Key Point of Using Data

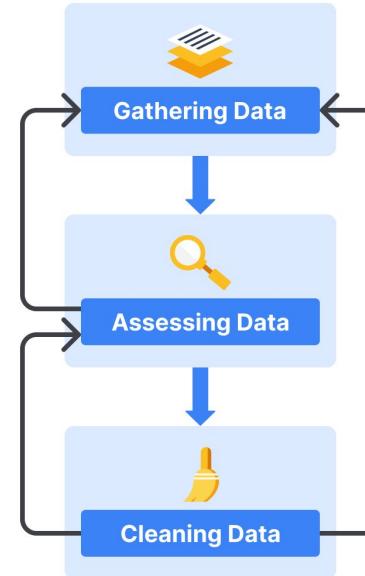


Problem in Real Life

| WHAT I THOUGHT DATA WOULD LOOK LIKE | WHAT IT *ACTUALLY* LOOKS LIKE |
|---|--|
|  |  |

Data Wrangling

Data wrangling is a process or collection of activities that include **data collection (Gathering data)**, **data assessment (Assessing data)**, and **data cleaning (Cleaning data)** before the data is used in the data analysis process.

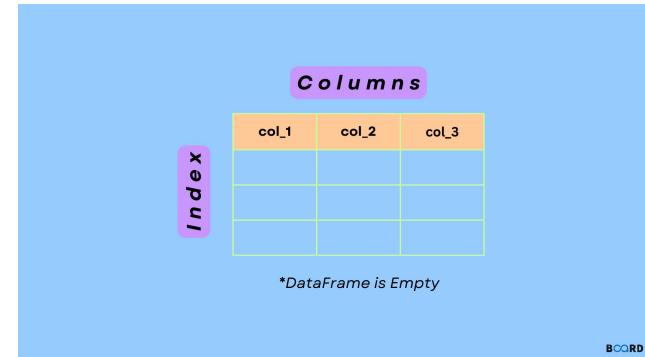


Gathering Data

Dataset Extension



DataFrames

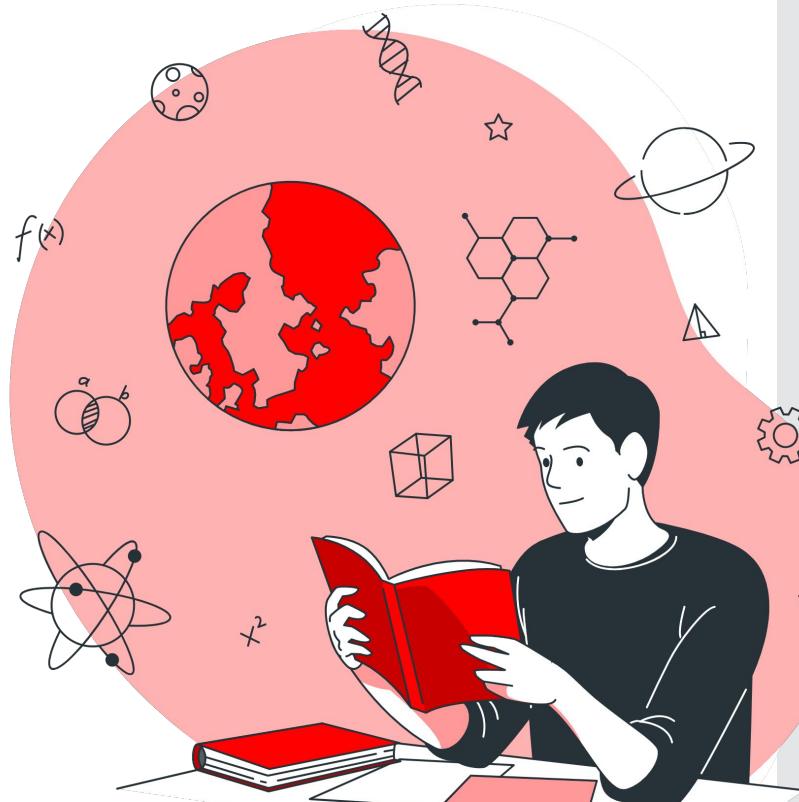


Gathering Data



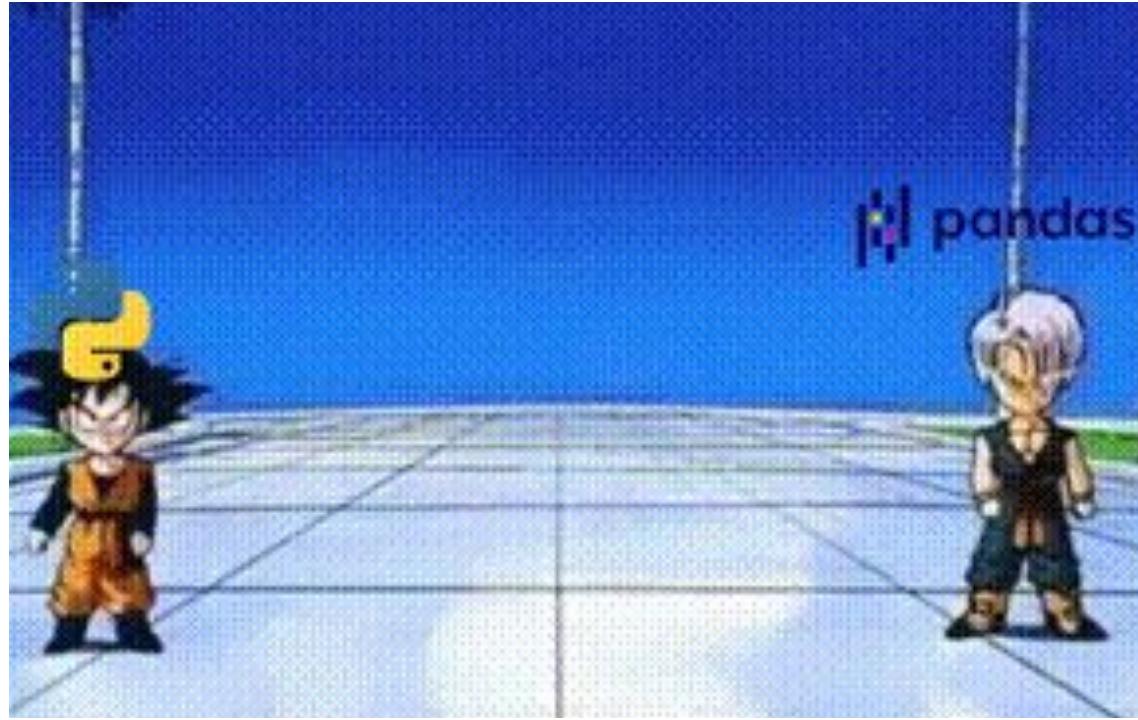
Why DataFrames?

DataFrames are **faster, easier to use, and more powerful** than tables or spreadsheets because they're an integral part of the Python and NumPy ecosystems.



pandas is a **fast**, **powerful**,
flexible and **easy to use**
open source data analysis
and manipulation tool,
built on top of the Python
programming language.





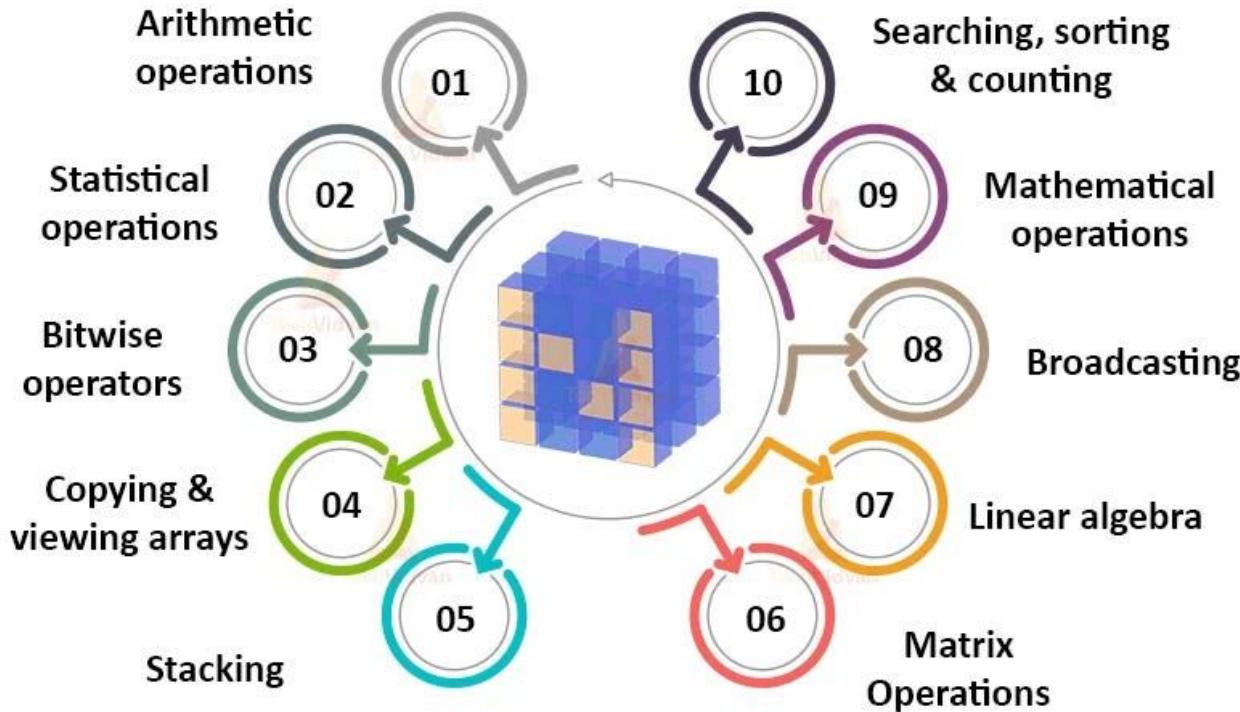


MACHINE LEARNING

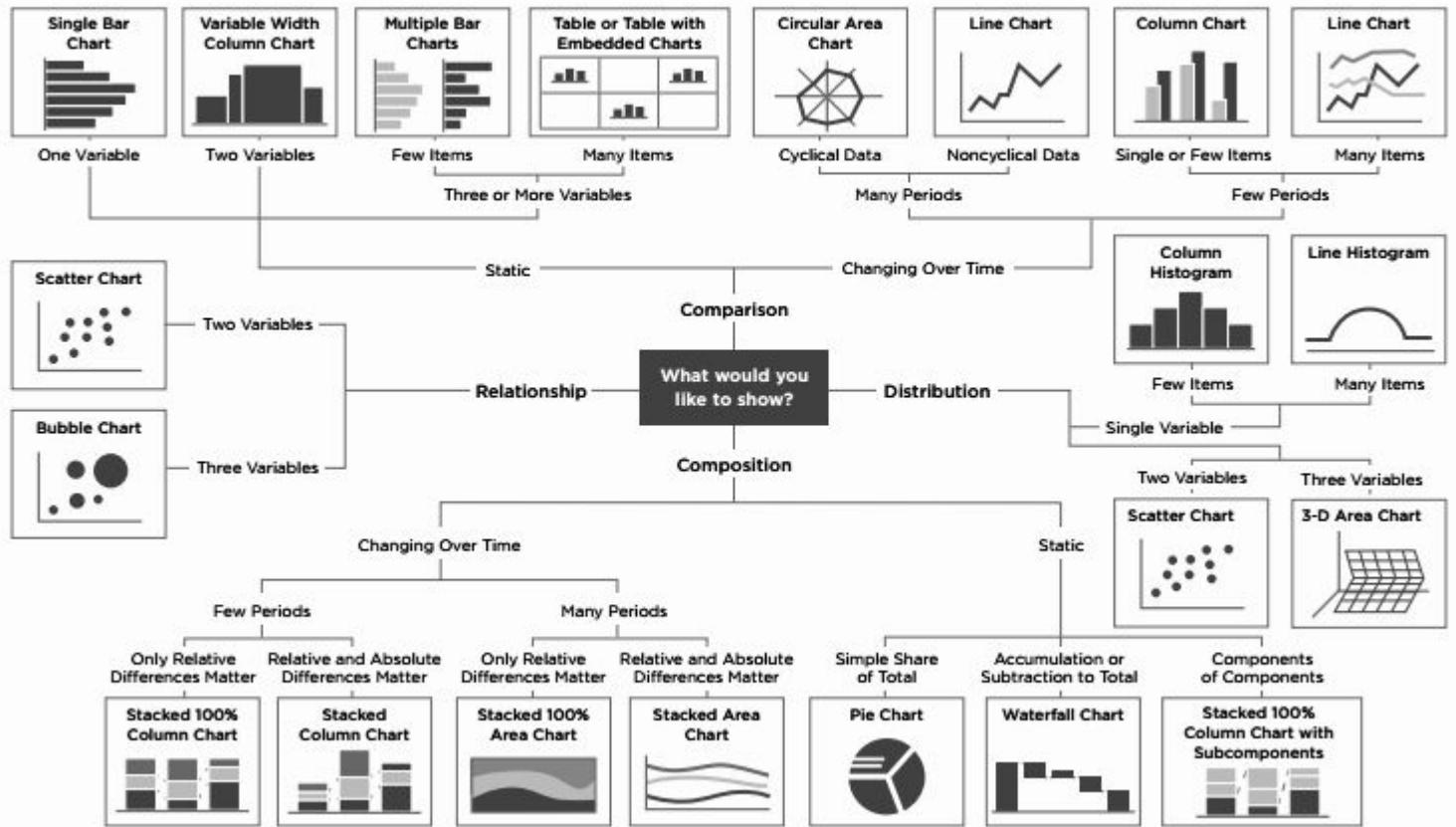
Why is it, when something happens, it is always you three?

Numpy Pandas Matplotlib

Uses of NumPy



SELECTING THE APPROPRIATE CHART FOR STRATEGY PRESENTATIONS



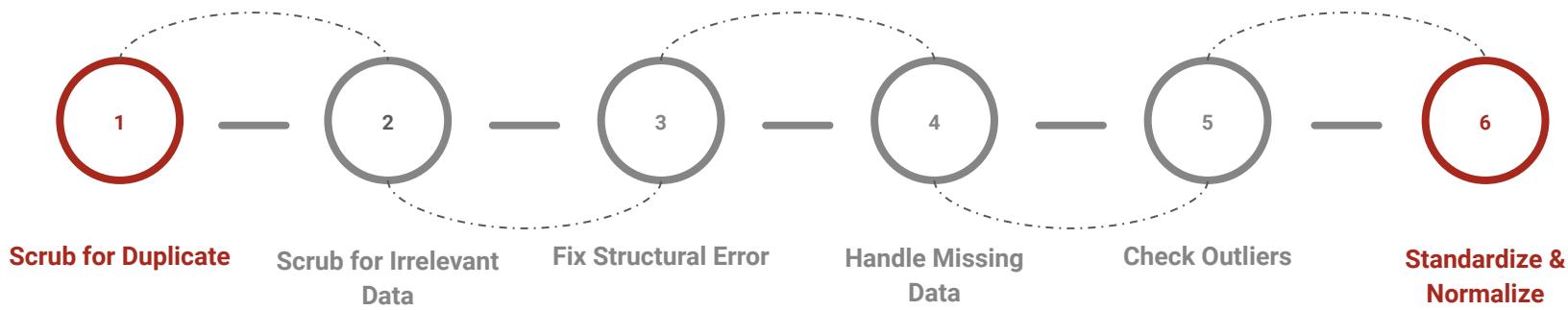
What next?

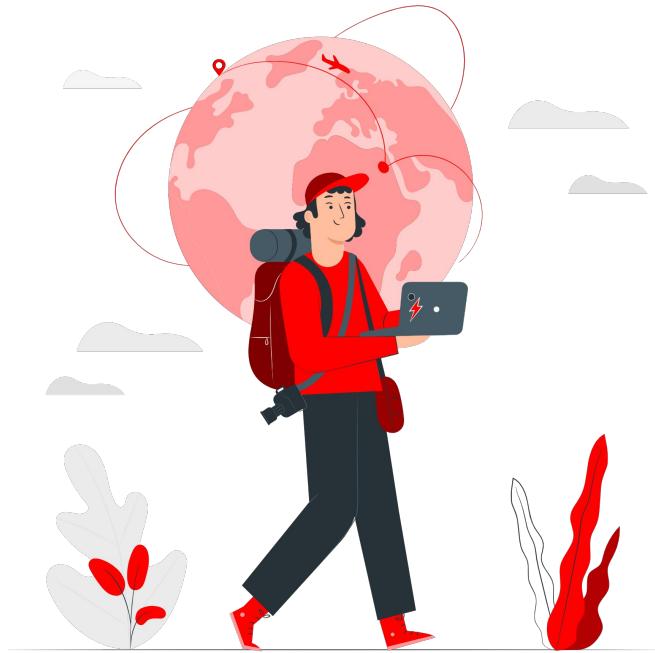


Assessing Data



Data Cleaning





EDA

EDA



Exploratory



Explanatory

Exploratory Data Analysis

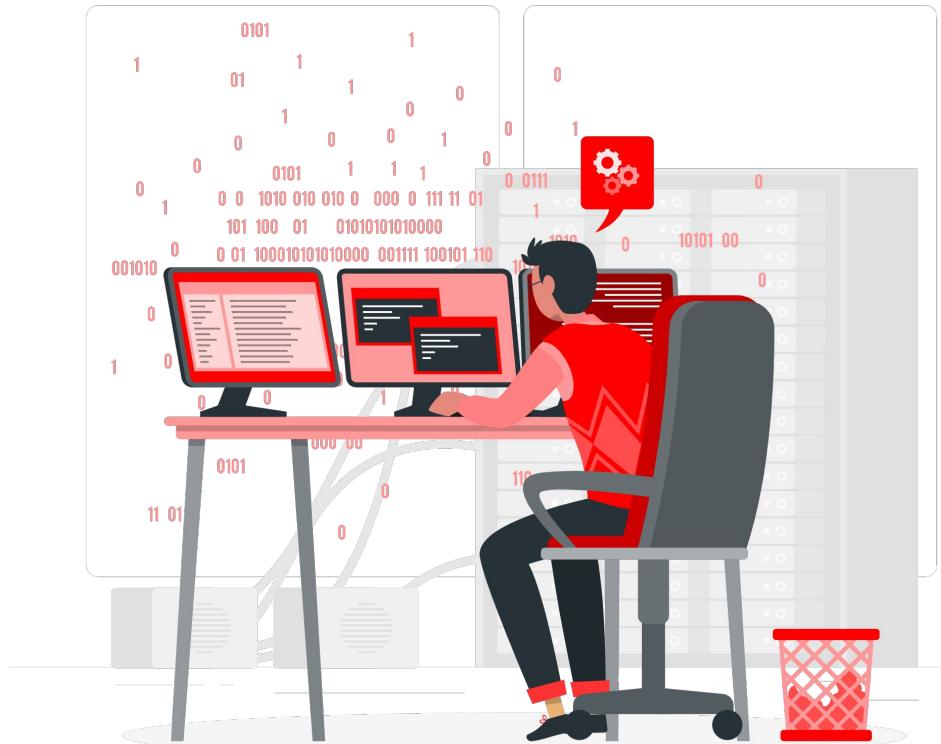


Problem Solving



Define Questions

Interacting with Data



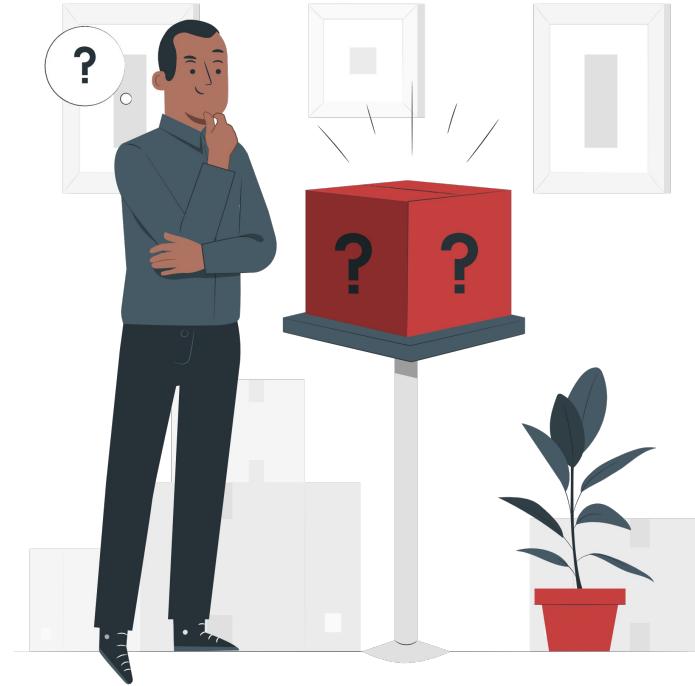
Break Time!



TERTURU

One Random Object

Guess one random object that your classmate picked!





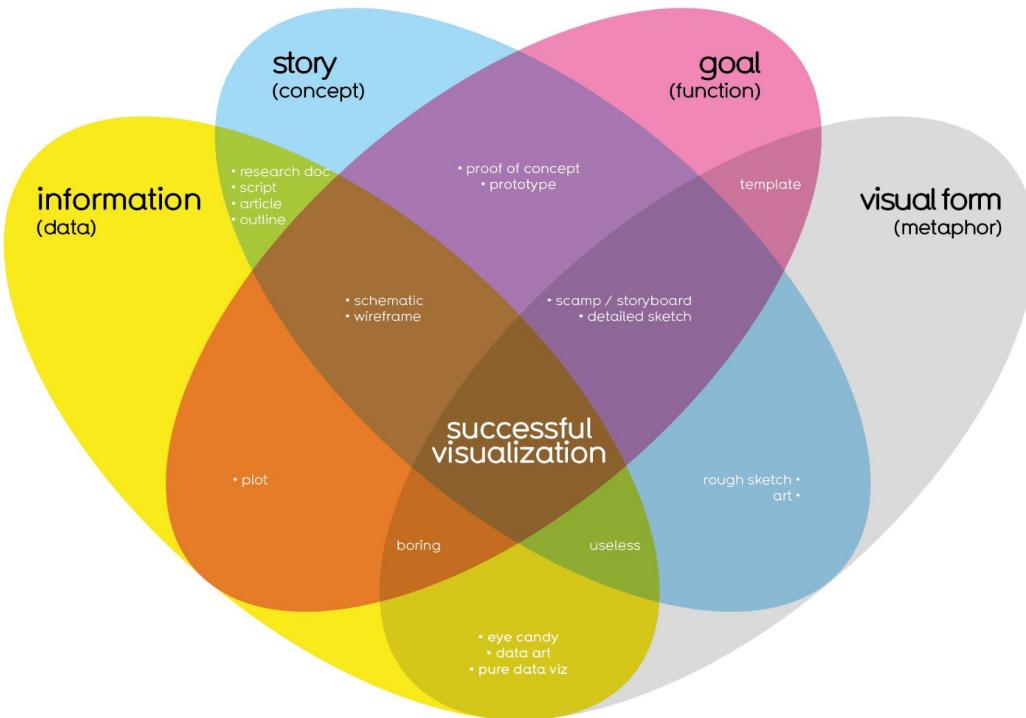
Data Visualization

What is **Data Visualization**?

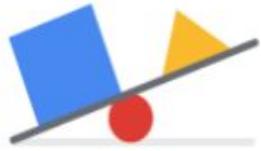
The **graphic representation** and presentation of data that can help us to **understand, draw conclusion, and tell a story about data**.



Elements of Data Visualization



Basic Principles of Design



Balance



Emphasis



Movement



Pattern



Repetition



Proportion



Rhythm

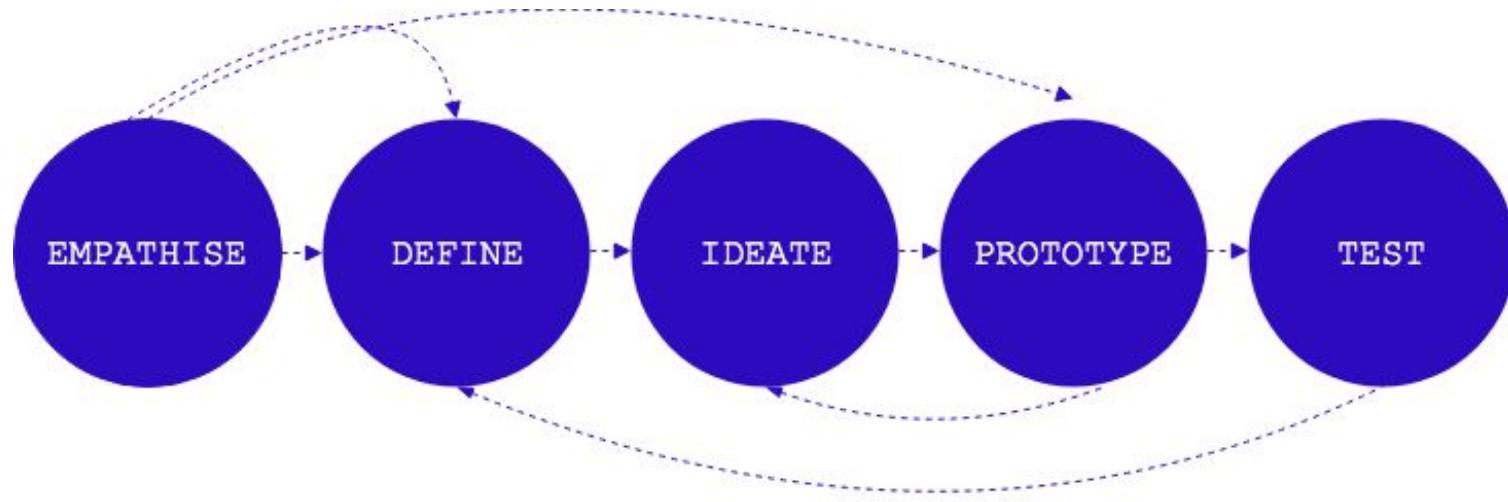


Variety



Unity

Design Thinking Phases of Data Visualization



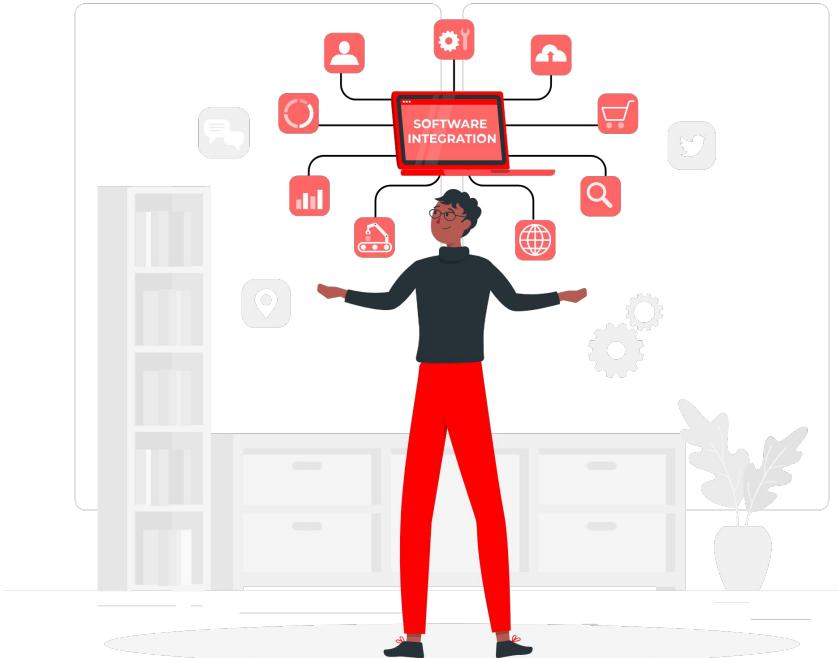
<https://www.enginess.io/insights/data-and-design-thinking>

Unraveling Insights through Analysis

What is Data Integrity?

Data integrity for analysis refers to the **quality** and **reliability** of data used in the **analytical process**.

Ensuring data integrity is crucial for producing **accurate, trustworthy, and meaningful insights**.



Data Integrity

- Clean data and good alignment.
- Clean the data before you perform your analysis.
- Aligns with an objective.



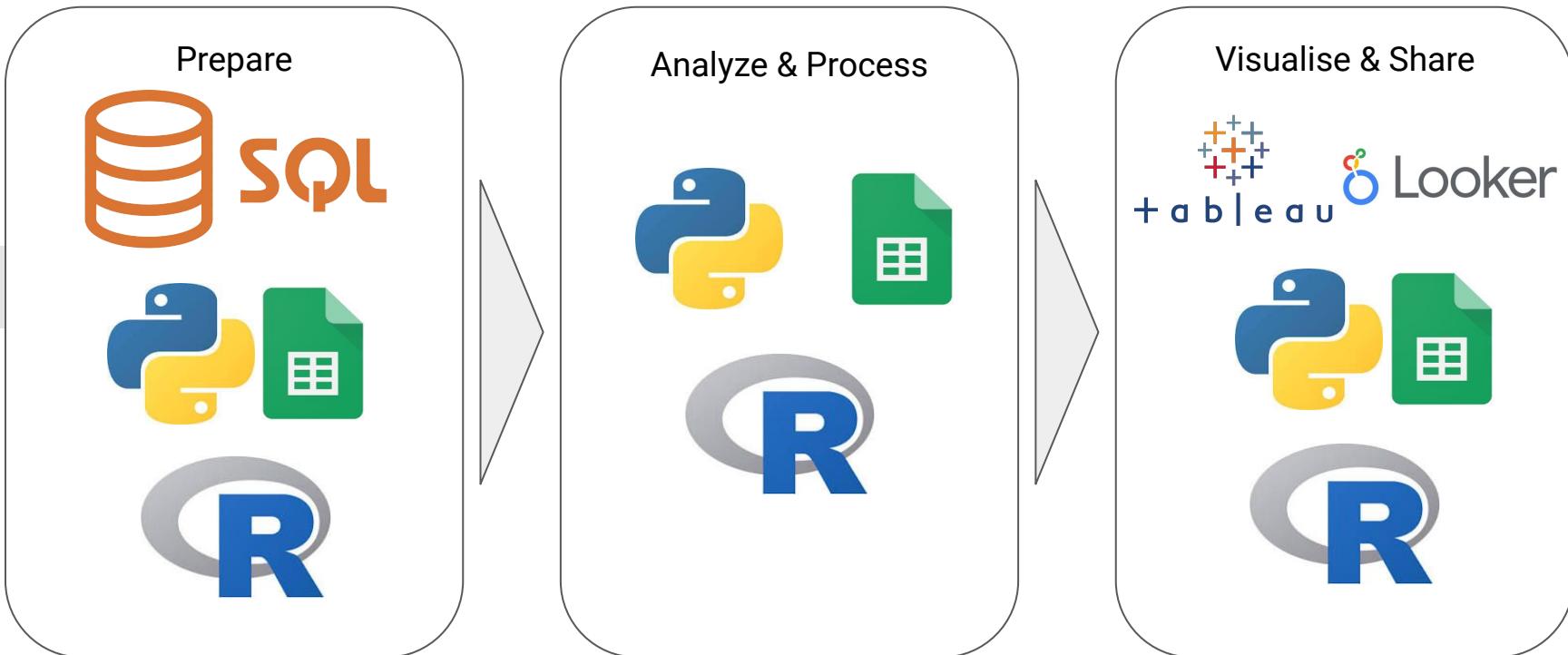
Five Essential Aspects of Analytical Skills



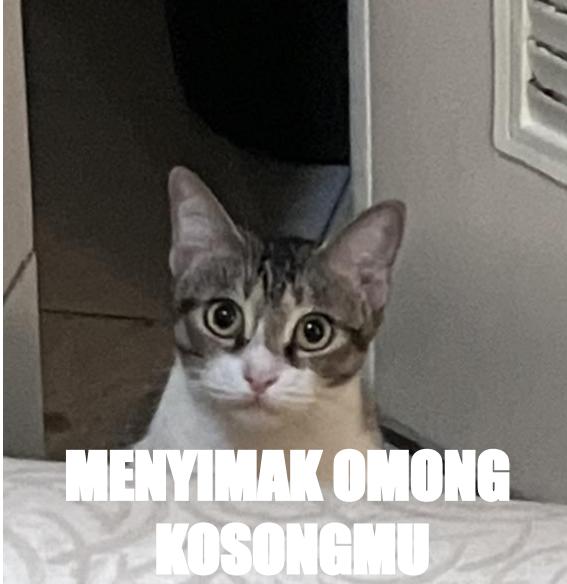
Data Analysis Process



Data Analysis Toolbox



Basic Data Storytelling



MENYIMAK OMONG
KOSONGMU

What is Data Storytelling?

Data storytelling is **communicating the meaning** of a data set with **visuals** and a **narrative** that are customized for each particular audience.



Storytelling

=

Visualization + Narrative + Context

How to Tell a Good Story About Data?

1

Engage your audience

Capturing and holding someone's interest and attention by identifying the audience.

2

Create compelling visuals

Visuals should take your audience on a journey of how the data changed over time or highlight the meaning behind the numbers.

3

Tell the story in an interesting narrative

A narrative has a beginning, a middle, and an end. It should connect the data you've collected to the project objective and clearly explain important insights from your analysis.

Narrative

| | |
|---------------------|---|
| Characters | The people affected by your story. |
| Setting | Background information about the data project that describes the current situation. |
| Plot | What creates tension in the current situation. |
| Big Reveal | How the data has shown that you can solve the problem. |
| “Aha” Moment | Share your recommendations and explain why you think they'll help your company be successful. |

Show and Tell



LIAT BETUL BETUL

Discussions



**NOOB BERTANYA,
LORD MENJAWAB**

Quiz & Feedback



Conclusion



You have learned basic **data analytics** and understand how to collect data with python can help developers to use data properly.

These knowledges will be helpful for your day-to-day work a **data practitioner** in the future.



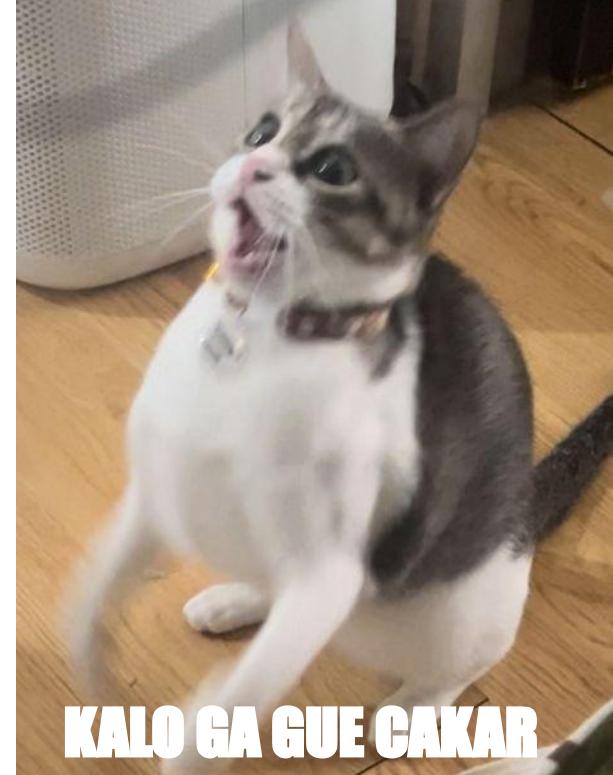
Thank You



Follow majikan saia:
Tiktok: @jenniethecatwoman

Currently at ~1500 Followers

 bangkit



LU