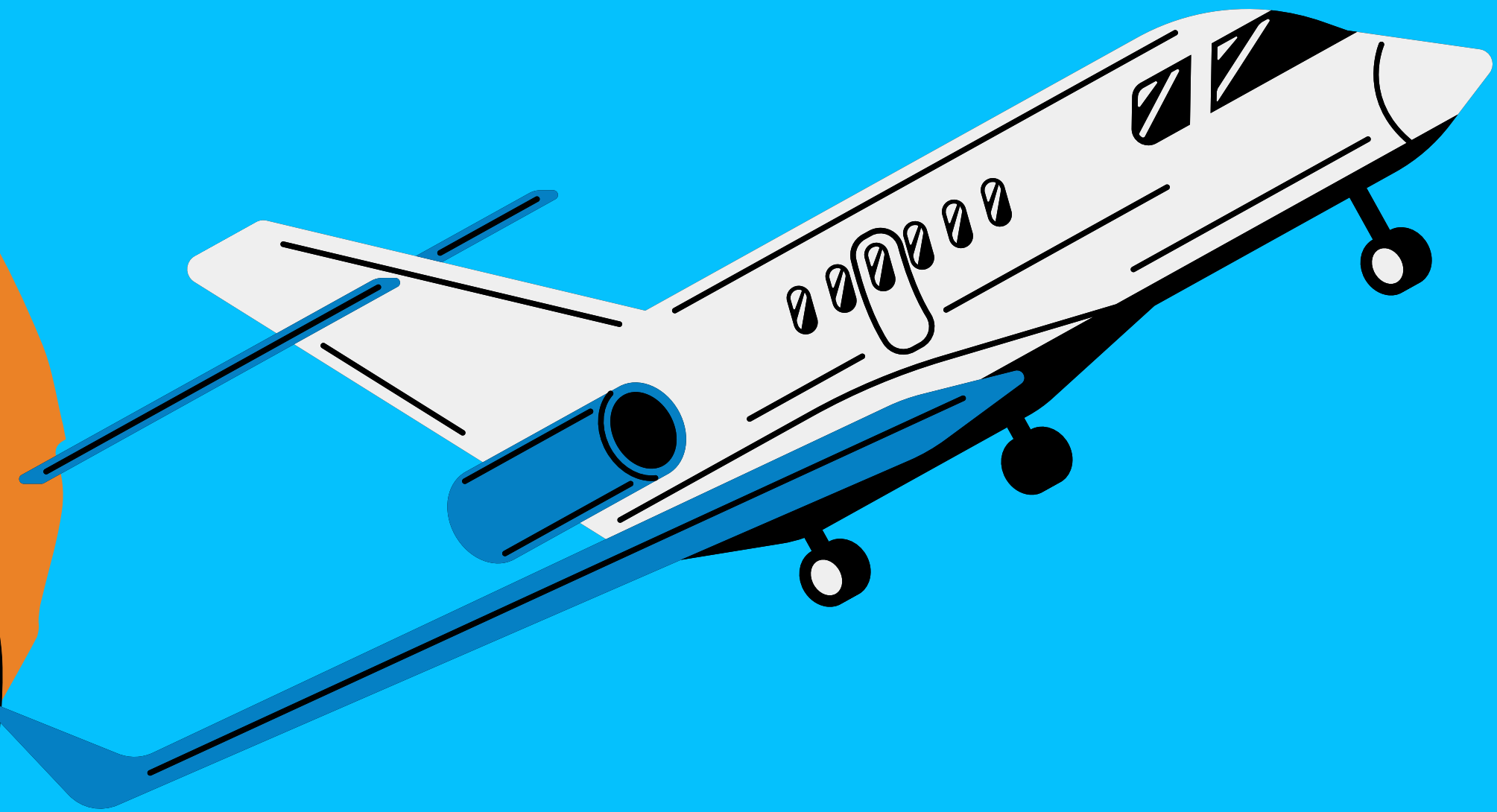
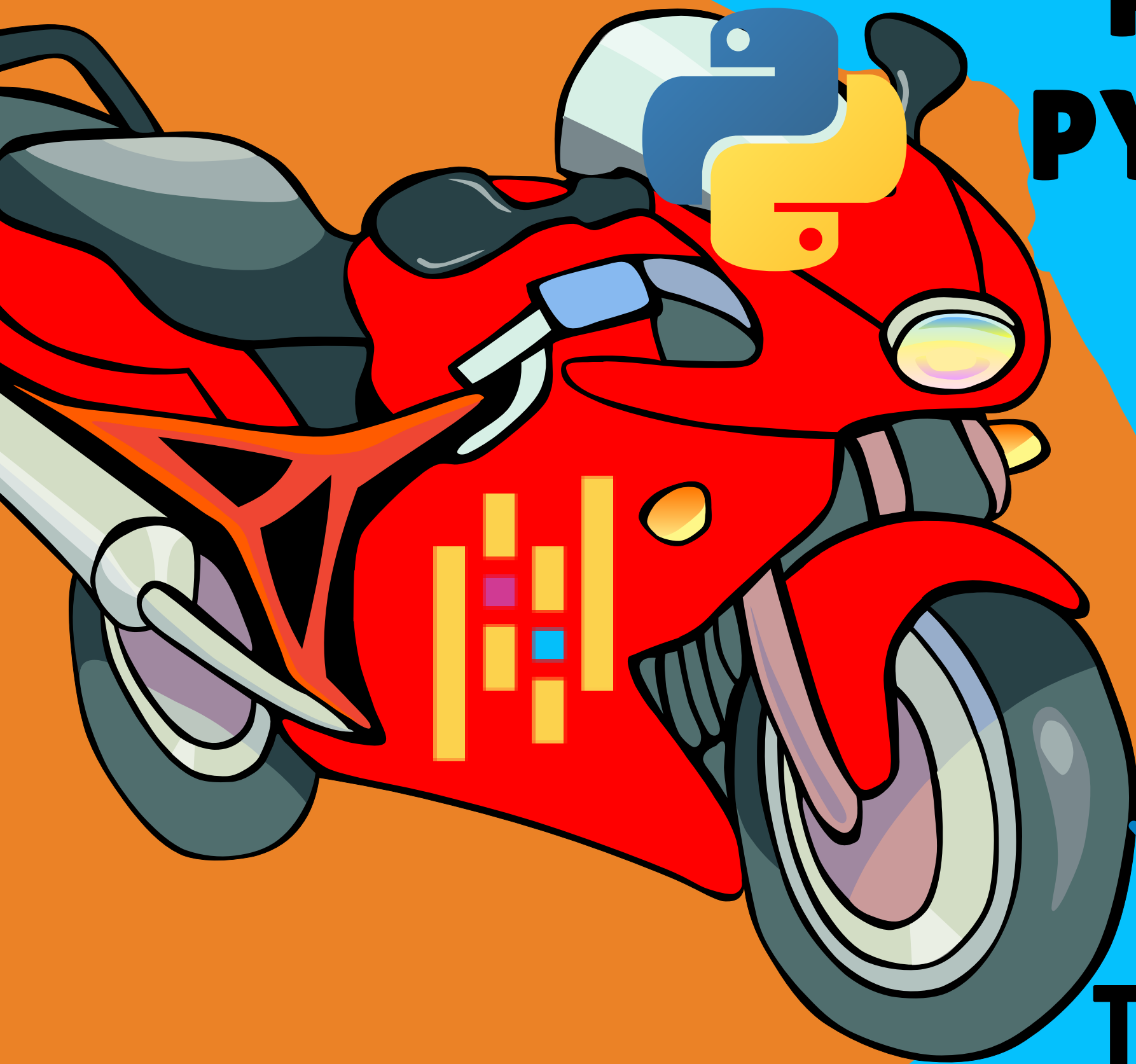


PANDAS THE ENGINE OF PYTHON DATA ANALYSIS



TACKLE REAL-WORLD DATA PROBLEMS LIKE A PRO



WHAT PROBLEM WE'RE FACING

DATA VARIETY, VELOCITY AND VOLUME

World works very fast and creates messy data, and your task is to fight with that data for insights and help decision makers

SERIES OF PROBLEM

- DATA INGESTION
- DATA CLEANING
- DATA FILTERING
- DATA AGGREGATION
- DATA JOINING
- DATA AUTOMATION

PYTHON ECOSYSTEM : PANDAS

HOW TO GET IT INTO MY OS?

- Install Python
- The do "pip install pandas"

WHERE CAN I LEARN ABOUT PANDAS

- 10mins to pandas

https://pandas.pydata.org/docs/user_guide/10min.html#min

- Pandas Userguide

https://pandas.pydata.org/docs/user_guide/basics.html

- This Video

BIG DATA ALTERNATIVE : PYSPARK

DATA INGESTION

01

ANATOMY OF FILES

What is inside each file type?
TXT, CSV, JSON, ZIP, PARQUET

- Dataframes can be created from files
- Created from json/ dictionaries
- Created from lists and python variables

02

LOOKING AT DATA FRAME

- Describing Data
- Selecting & Sorting Data
- Missing Values correction



DATA FRAME : OUR HERO

HOW DOES PANDAS SOLVE PROBLEM

01

OPERATIONS

- STATISTICS
- CUSTOM FUNCTIONS
- DATE MANIPULATION
- STRING MANIPULATION

02

BRINGING TOGETHER

- MERGING DATAFRAME
- JOINING DATAFRAMES

03

AGGREGATION

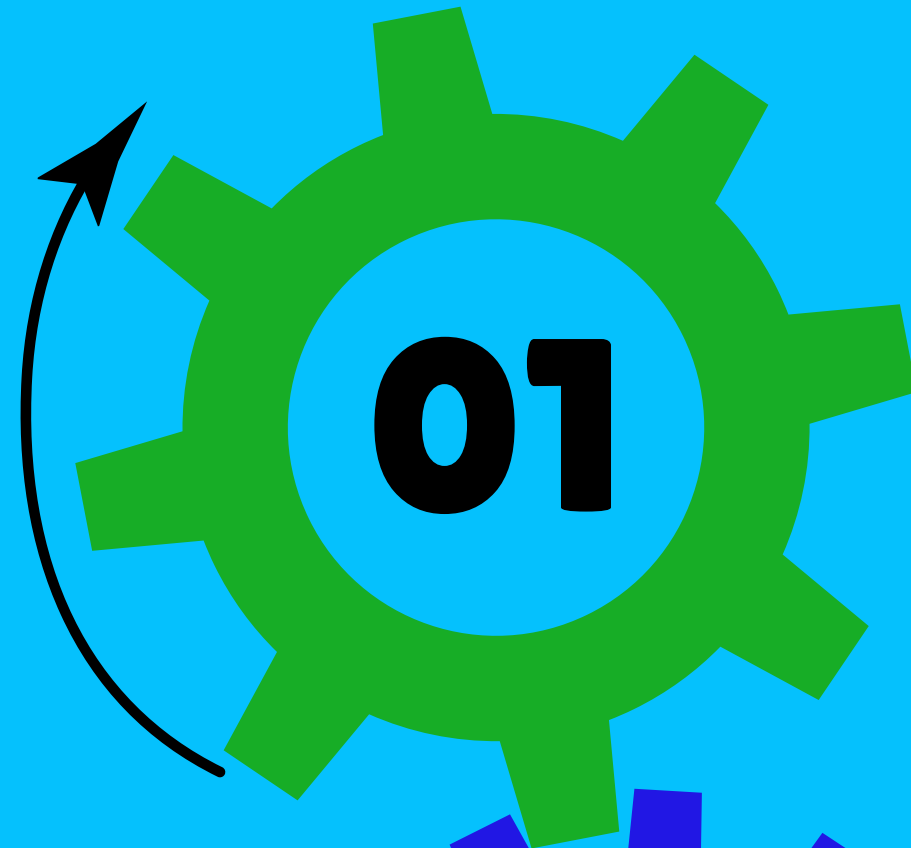
- GROUPING BY COLUMN
- CUSTOM AGGREGATING

04

RESHAPING

- STACKING & UNSTACKING
- PIVOT TABLES
- USING CATEGORICALS

DATA ANALYSIS PROCESS

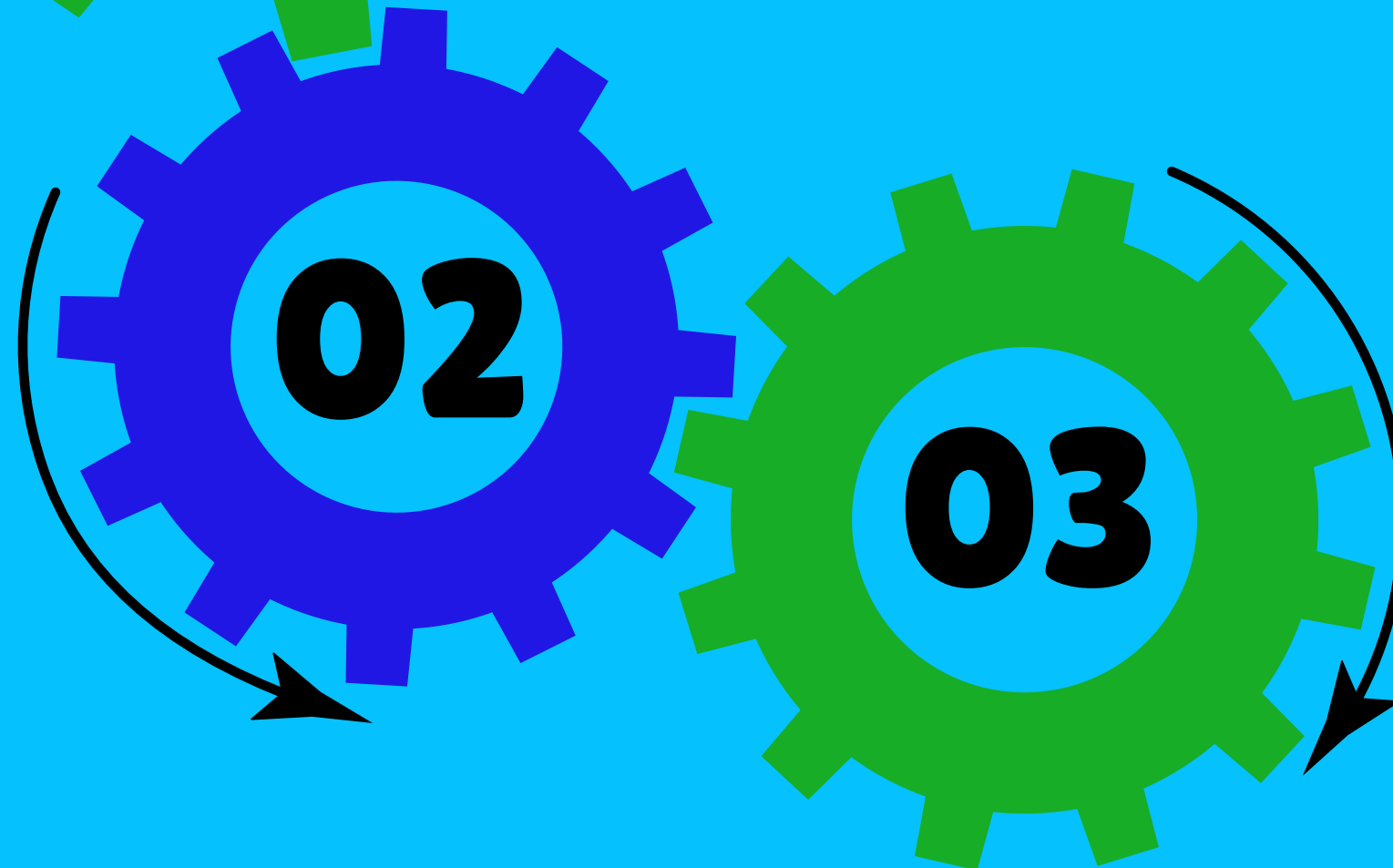


BRING IN THE DATA

- CLEAN
- EXPLORE
- FIND INSIGHTS

VISUALISE INSIGHTS

- EXPORT DATA
- MAKE VISUALS



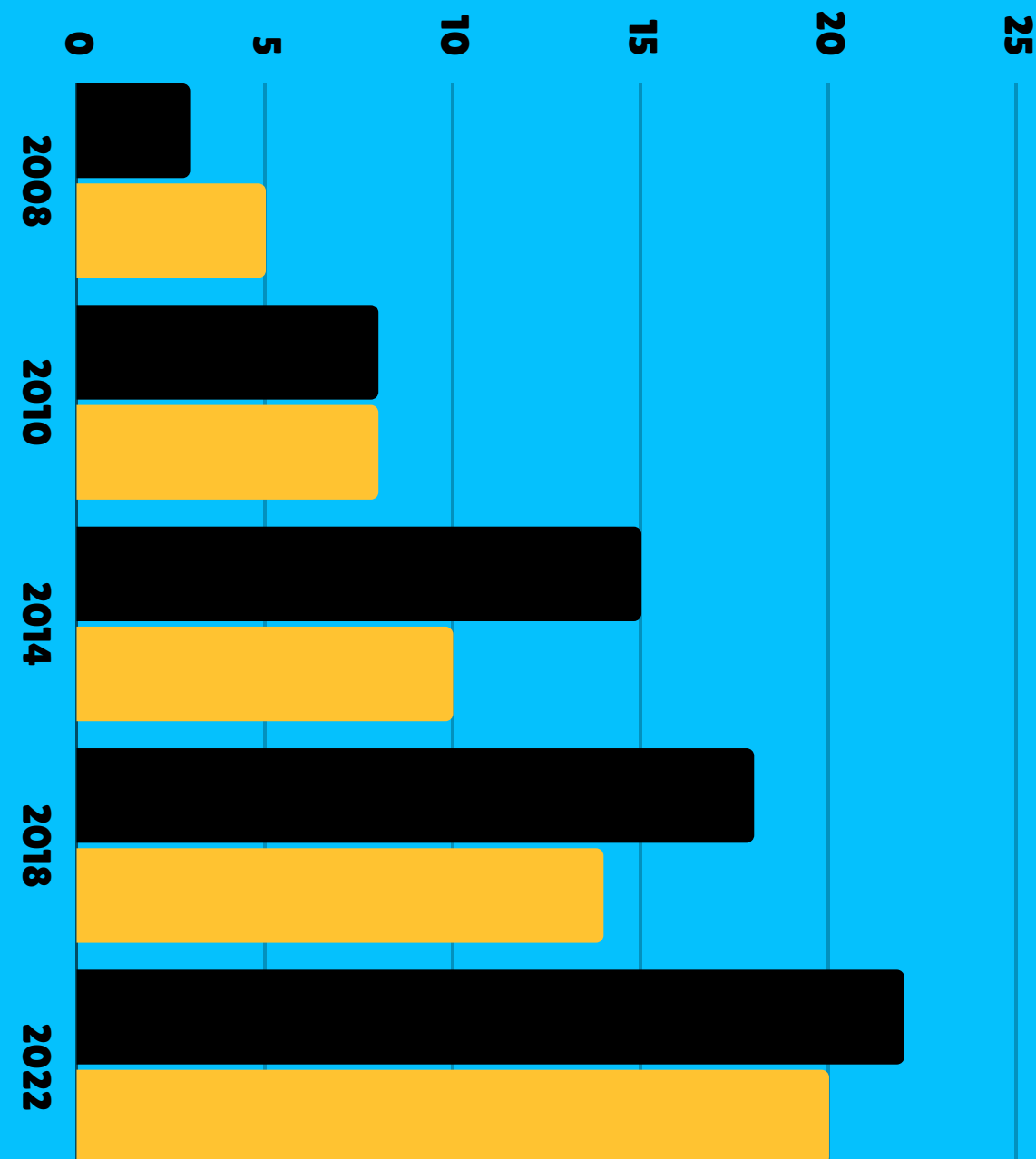
AUTOMATE THE PROCESS

- CODE FOR AUTOMATING
- INTEGRATE WITH SYSTEMS

HEAD TO JUPYTER

**QUESTIONS AND
COMMENTS**

WHAT NEXT ???



**DATAFRAMES TO
DATA VISUALS**

STARRING STREAMLIT

<https://kamalabot-streamlit-deploys-experiment-1dashboard-ver00-p7orli.streamlit.app/>