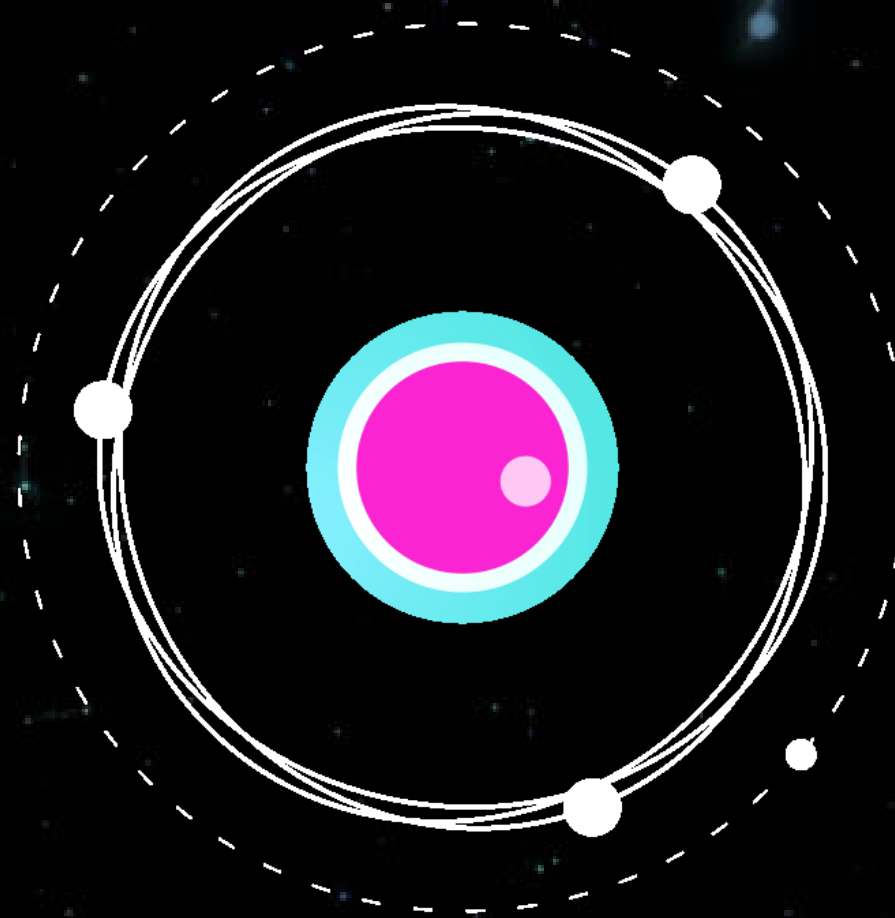


DATA CELL TRAINING ROADMAP

INPT

CIT

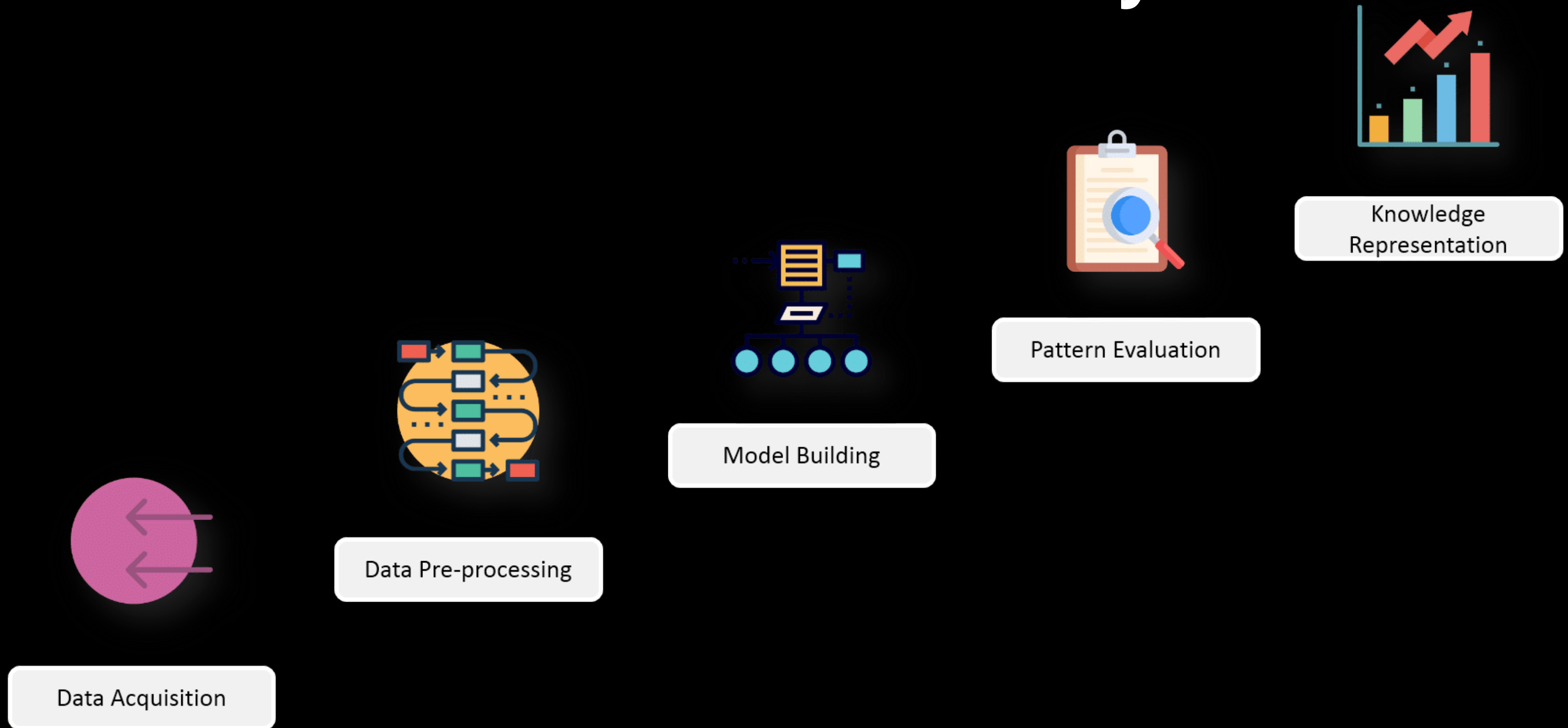


MADE BY:

TAHA ELGHABI



Data Science Lifecycle



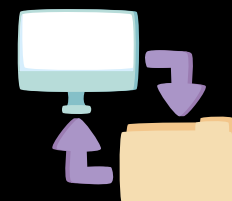


PLAN

OVERVIEW, DATA ROADMAP



DATA COLLECTION



EXPLORATORY DATA ANALYSIS



CLEANING, FEATURE ENGINEERING

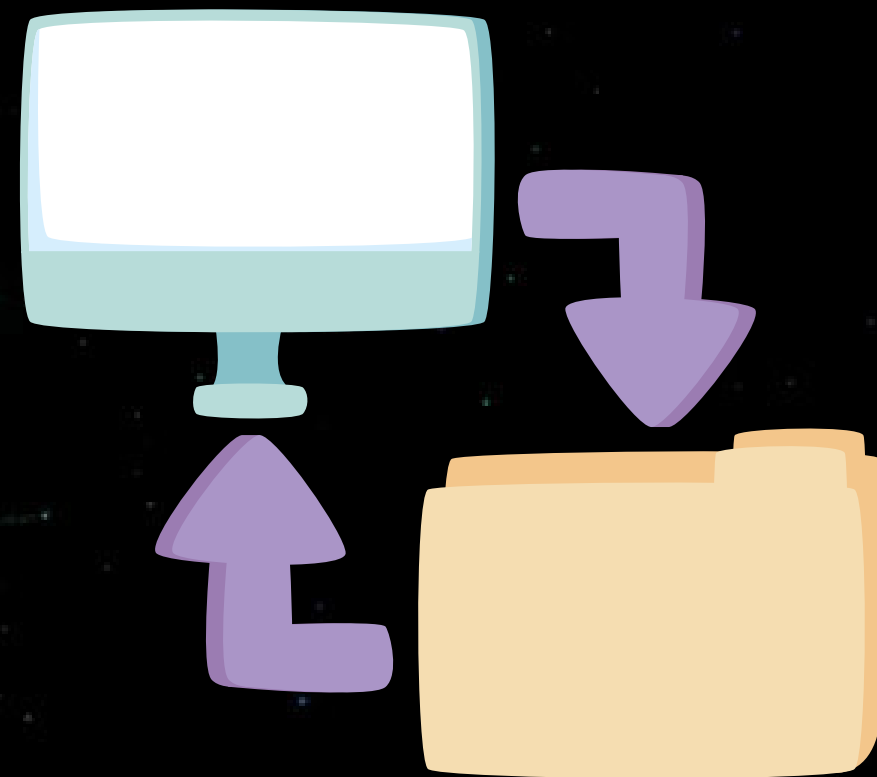


VISUALIZATION



DATA CELL TRAINING

S2: DATA COLLECTION



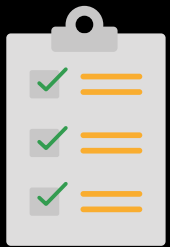


TABLE OF CONTENT

- Where is the Data ?

- What is Web Scraping 

- How do I read this HTML nonsense 

- Some tips and resources 



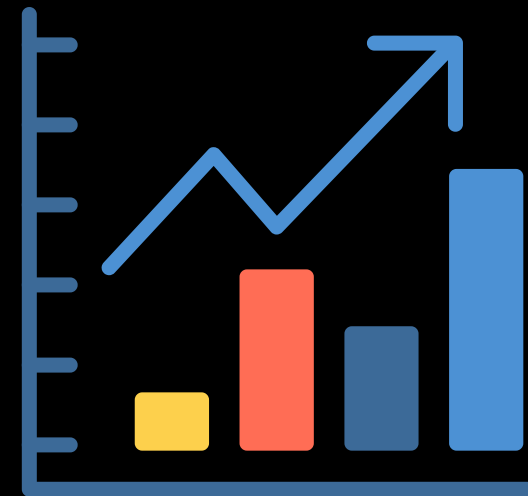
Where is the Data



Data can be everywhere



Surveys



Sales reports



INTERNET

Company Data

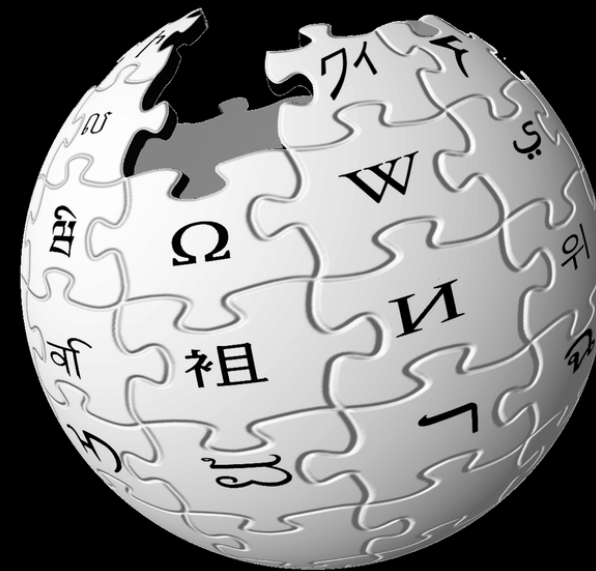


There is a lot of the Data here!



Google

INTERNET ARCHIVE
WayBack Machine



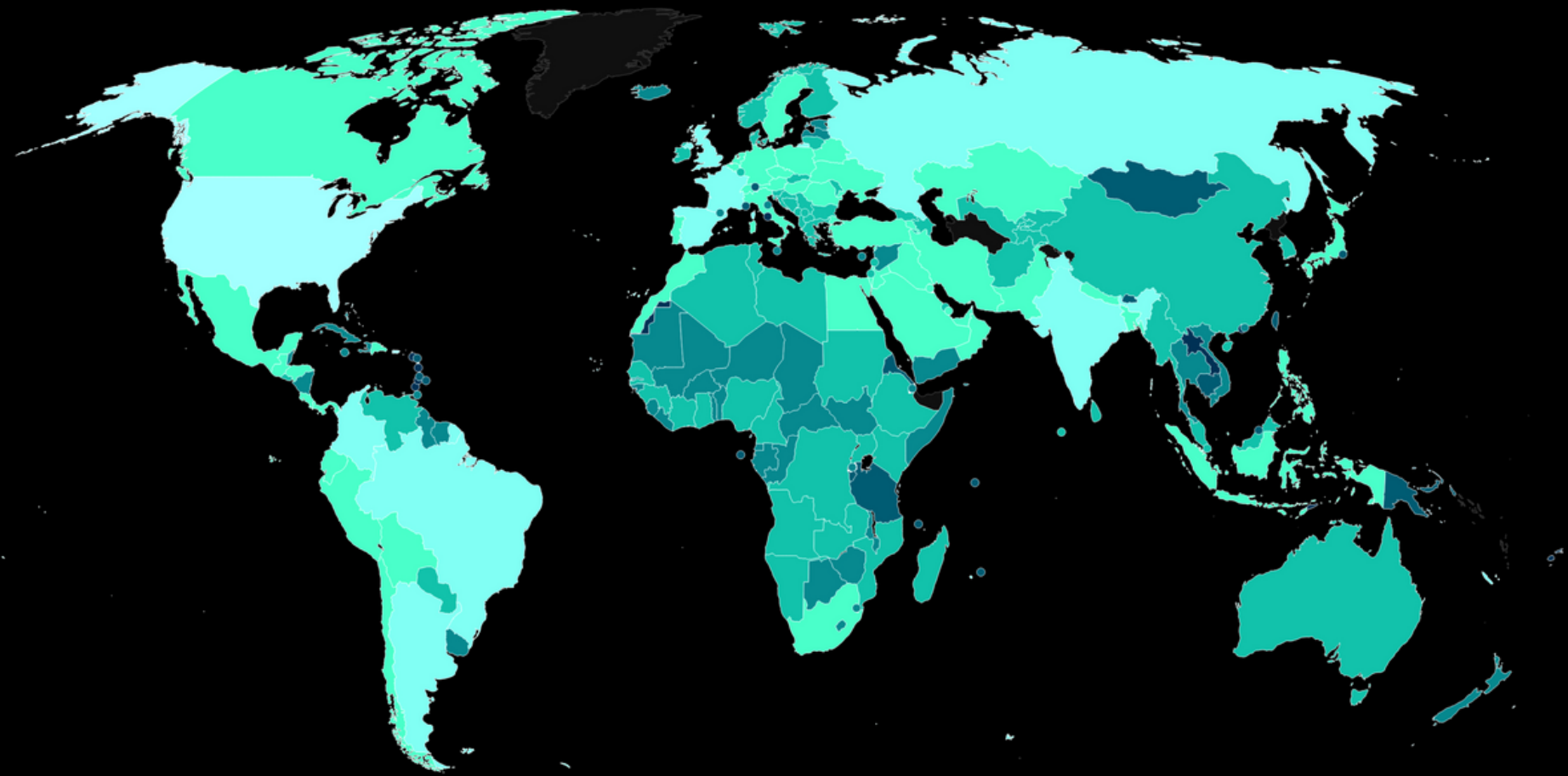
WIKIPEDIA
La enciclopedia libre



What is **Web Scrapping**



What can I do with it ?



Covid Cases Map



Stocks visualization



How do I read this **HTML** nonsense



Some **tips** and **resources**



kaggle Datasets

Medium

Take a Look
at source code



Data Science Essentials: Scraping Data From the Web

Web scraping is an important skill for any data scientist to have in their toolbox

Medium / Eijaz Allibhai / May 17, 2020

