2022

31 Monday January — Task 12 — ٣١ الأثنين يناير

٢٨ جماد آخر ١٤٤٣ هـ

٢٣ طوبة ١٧٣٨ ق

08:00 Data Analysis process

Question ⇒ wrangle ⇒ Explore ⇒ Draw conclution ⇒ Communicate

09:00

1- questions:

10:00 always starts with asking questions

given dataset and figure a good questions to ask

11:00 what am I trying to find out? is there a problem I'm trying to solve!

once we have our questions we need to wrangle the dat

12:00

2- wrangle:

01:00 we ~~should~~ have data in great quality

we have 3 steps

02:00 first: we Gather our data, clean it in jupyter notebook

Think carefully about what data we most helpful

03:00 and answer the question we collect

second: we assess our data to identify any problems

04:00 third: we clean our data

Then we carefully have clean data

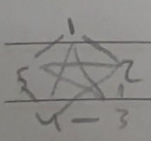
05:00

3- Explore:

06:00 exploring and then minding your data to maximize the potinial of analysis, visualizations and models

07:00 after exploring you can remove outliers and create new and more descriptive features from existing data

08:00 known as feature Engineering



08:00

4- Draw conclutions

This step is typically approched with machine learning and inferential statistics and descriptive statistics

09:00

10:00

5- Communicate

- communicate results to others
- justify and convey meaning of insite you found

11:00

12:00

Asking questions: first step of data analysis process

geting data first and then ask questions for it

01:00

wrangling

* Gather
* Assess
* Clean

EDA

* Explore
* Augment

02:00

03:00

04:00

Gathering

Data can happen in variety of ways Downloding files or from API, from web page ,in many companies may have existing data base

05:00

06:00

The file in formula CSV comma seperated values

07:00

08:00