Data analysis process::

1.Data gathering

-> Import (csv,excel,text,json,database)

-> Export (csv,json,excel,database,html)

-> API

-> Web scraping

What is Data Analysis

Data analysis is a process of inspecting,cleansing,transforming and modeling data with the goal of discovering useful information,

informing conclusions and supporting decision-making.

Data analysis process

->Asking Questions

1.what features will contribute to my analysis?

2.what features are not import for my analysis?

3.which of the features have a strong correlation?

4.Do i need data preprocessing?

5.what kind of features manipulation/engineering is requried?

How can i ask better questions ?

Subject Matter Expertise

Experience

->Data Wrangling/Munging

Data wrangling,sometimes referred to as data munging,is the process of transfroming and mapping data from one "raw" data form into another

format with the intent of making it more appropriate and valuable for a variety of downstream purpose such as analytics.

1.Gathering Data.

->CSV files,API,Web Scraping,Databases

2.Assessing Data

->Finding the number of rows/columns(Shape)

->Data types of various columns (info())

->Checking for missing values (info())

->Check for duplicate data(is\_unique)

->Memory occupied by the dataset(info)

->High level mathematical overview of the data(describe)

3.Cleaning Data

->Missing Data (eg mean)

->Remove duplicate data (drop\_duplicates)

->Incorrect data type(astype)

->Exploratory Data Analysis

Explore,Augment

1.Exploring Data

->Finding correlation and covariance

->Doing univariate and multivariate analysis

->Plotting graphs (data visualization)

2.Augmenting Data

->Removing Outliers

->Merging Dataframes

->Adding new Column

These operations are collectively called as Feature Engineering

->Drawing Conclusions

Machine Learning

Inferential Statistics

Descriptive Statistics

->Communicating Results / Data Storytelling

In person

Reports

Blog Post

PPTs/Slide decks

2.Data cleaning

3.EDA

day15 - working with csv files

day16 - working-with-json-and-sql

day17-api-to-dataframe

day18-pandas-dataframe-using-web-scraping

https://github.com/campusx-official/pandas-io

pandas-import.ipynb