Data analysis

To start with data analysis, you need to know few things:

1. First, u need to ask **questions**:
   - Asking questions will help you to find the problem or question that you want to answer. By doing that correctly you will be able to do your analysis and draw conclusions.

2. **Wrangling**:
   - Data Wrangling is the process of gathering, collecting, and transforming Raw data into another format for better understanding, decision-making, accessing, and analysis in less time. Data Wrangling is also known as Data Munging.
   - Data wrangling in python deals with the below functionalities:
     1. **Data exploration:** In this process, the data is studied, analyzed, and understood by visualizing representations of data.
     2. **Dealing with missing values:** Most of the datasets having a vast amount of data contain missing values of *Nan, they are needed to be taken* care of by replacing them with mean, mode, the most frequent value of the column or simply by dropping the row having a *Nan* value.
     3. **Reshaping data:** In this process, data is manipulated according to the requirements, where new data can be added, or pre-existing data can be modified.
     4. **Filtering data:** Sometimes datasets are comprised of unwanted rows or columns which are required to be removed or filtered.
     5. **Other:** After dealing with the raw dataset with the above functionalities we get an efficient dataset as per our requirements and then it can be used for a required purpose like data analyzing, machine learning, data visualization, model training etc.

3. Exploring:
    - Exploring your data will help you to discover new patterns, visualize your data and build new intuitions about what you are working with.
    - By doing so you can remove outliers and create new descriptive features from existing data also known as feature engineering

4. Draw conclusions:
    - This step can be done using inferential statistics, machine learning and descriptive statistics.

5. Communication:
    - In this step u need to link your analysis with each other by using reports , slide decks, blog posts, emails, presentations or conversations.