**Data analysis**

is the practice of working with data to glean useful information, which can then be used to make informed decisions.

This idea lies at the root of data analysis. When we can extract meaning from data, it empowers us to make better decisions.

**Data analysis process:-**

**Ask Questions**

The first basic step of any data analysis process is to determine what questions to ask. The choice of question may be driven by the available data. Or, the analyst may need to obtain her own data in order to address the question.

Either way, determining the question or set of questions to research is an important step in creating progress toward toward meaningful insights. Possible questions:

- Did the new website layout increase the click-through rate?
- Which items should be kept in inventory, and at what levels?
- What marketing strategies are producing the greatest sales volume?

**Wrangle Data**

The data wrangling step consists of a few sub steps related to obtaining data in a usable form. These sub steps include:

- **gathering** the data,
- **assessing** it, and
- **cleaning** it.

Gathering the data is simple enough: the raw data must be downloaded or otherwise obtained in order to proceed with analysis. In some cases, analysts will be given their data, but in most cases they will need to acquire it.

The next step is to assess the data in order to determine its state and the degree of “cleaning” that will be necessary.

Cleaning is an industry term that relates to coping with mess that is real world data. This includes ensuring the data is stored in usable and appropriate datatypes, addressing nulls and missing values, and keeping an eye out for data that is clearly aberrant or incorrect. The end result should be the highest-quality data possible.

### Perform EDA

Following cleaning, the analyst should explore her data in order to build intuition about expected and unexpected patterns and relationships in the data. Where necessary, at this stage, it is also appropriate to remove outliers or “augment” the data by designing derived data that may be useful in subsequent stages of analysis. This is sometimes known as “feature engineering.”

This step will likely involve creating preliminary visualizations. Intuitions and preliminary results from this step may lead the analyst to refine her questions. It is also possible that deeper exploration will reveal additional wrangling is required.

### Draw Conclusions

This step usually uses inferential statistics and/or machine learning techniques in order to reach conclusions that answer the question being posed. The conclusions will most likely be stated in terms of descriptive statistics, and will most likely impact the activities of others in the company.

### Communicate Results

Finally, it is necessary to communicate the results found in the previous step. There are many possible mediums for this communication. Among them are formal reports, presentations, emails, or just conversation.

The ability to generate appropriate data visualizations is crucial to any data analyst. Insights are only valuable to the degree their meaning is successfully communicated.