**The Data Analysis process consist these steps**

1. Asking questions
2. Data Wrangling
   1. Data Gathering
      1. CSV files
      2. SQL DataBases
      3. APIs
      4. Web Scraping
   2. Assessing Data
   3. Cleaning Data
3. Exploratory Data Analysis
4. Drawing Conclusion
5. Communicating Results

**Assessing Data** - it means to understand data more deeply. Before implementing a method to clean it , we need to have a better understanding about what the data tells us.

**Cleaning Data -**

**Types of Unclean Data**

**Two kind of unclean Data**

1. Dirty Data- Data with quality issue

This type of data is also known as low quality data where the content quality of the data is not too good.

1. Messy Data (Data with tidiness issue)

Messy data is also known as unstructured data

1. **Dirty data**

There are some kind of issue in dirty data

1. Duplicate data
2. Corrupted data
3. Inaccurate data
4. Missing Data
5. **Messy Data**

When the data is unstructured then it is known as messy data and also know as untidy data.

Tidy data has following properties

1. Each variable of the data forms a columns
2. Each observation of the data forms a row
3. Each observational unit forms a table

If any data violates any of the above properties then its a untidy data.

**Type of Assessment**

There are two type of assessment

1. Manual- By going through the data manually in google sheets
2. Programmatically- By using the pandas function such as info() , describe() or sample().

**Steps of Assessments**

There are two steps involves in Assessments

1. Discover
2. Documentation

Data Quality Dimensions

1. Completeness issue
2. Validity Issue
3. Accuracy Issue
4. Consistency Issue