### What is Assessing Data?

The data analysis process consists of the following steps

1. Asking Questions

2. Data Wrangling

a. Gathering Data

i. CSV files

ii. APIs

iii. Web Scraping

iv. Databases

bAssessing Data

c. Cleaning Data

3. Exploratory Data Analysis

4. Drawing Conclusion

5. Comunicating Results

### Assessing Data

In this step, the data is to be understood more deeply. Before implementing methods to clean it, you will definitely need to have a better idea about what the data is about.

### Types of Unclean Data

There are 2 kinds of unclean data

1. Dirty Data (Data with Quality issues)

Dirty data, also known as low quality data. Low quality data has content issues.

2. Messy Data (Data with tidiness issues)

Messy data, also known as untidy data. Untidy data has structural issues.

### 1. Dirty Data

When the data has following kinds of issues:

1. Duplicated data

2. Missing Data

3. Corrupt Data

4. Inaccurate Data

This kind of data is known as Dirty data

### 2. Messy Data

When the data has structural issue is it known as Messy or untidy data.

Tidy data has the following properties:

1. Each variable forms a column

2. Each observation forms a row

3. Each observational unit forms a table

### Types of Assessment

There are 2 types of assessment styles

1. Manual - Looking through the data manually in google sheets

2. Programmatic - By using pandas functions such as info(), describe() or sample()

### Steps in Assessment

There are 2 steps involved in Assessment

1. Discover

2. Document

### Data Quality Dimensions

1. Completeness Issues

2. Validity Issues

3. Accuracy Issues

4. Consistency Issues