Project on Any kind of Data set.

**How I got the data set?**

* First of all, I select the \*\*GDP OF EUROPEAN COUNTRIES\*\*, from the kaggle.com.
* The format of the file is .csv I am going to perform different operation on that dataset by using pandas library.

Those operations are given below:

- Data Load

- Data Cleaning

- Data Preprocessing

- Data Augmentation

- Data save

**Data Load**

Now starting with the data load, it means that how we can load data or read the data of excel file by using pandas. Data Load involves different functions() to perform.

1. Inspecting the Data:

- Index

- head()

- tail()

- sample()

2. Checking the Structure:

- info()

- shape()

- columns()

3. Basic Calculations:

- mean()

- sum()

- min()

- max()

4.Quick Visualization:

- plot()

**Data Cleaning**

Data cleaning in pandas refers to the process of preparing and correcting your dataset by handling issues like missing values and duplicate data. Data cleaning also involves different functions and they are categorised as according to their use:

1. Handle Missing Data:

- isnull()

- notnull()

- fillna()

- dropna()

2. Removing Data:

- duplicated()

- drop\_duplicates()

- drop()

**Data Preprocessing**

Data preprocessing means that to convert or show your raw data in a well-mannered or structured way that will helps in to analyze the Data sets. You can also change the type of data.

* df() = df().astype()
* .concat
* .merge

and then writing the type of data type you want it to change into.

**Data Augmentation**

Data augmentation is the process of artificially increasing the size and diversity of a dataset by applying transformations or generating synthetic data.

**Data save**

Data saving in pandas refers to the process of storing data from a pandas DataFrame into various file formats, such as CSV, Excel, or SQL databases,

after performing necessary operations or modifications. Saving the data allows you to keep the processed or clean data for future analysis or sharing.