Have general knowledge of python and statistics

Python : <https://www.youtube.com/watch?v=7wnove7K-ZQ&list=PLu0W_9lII9agwh1XjRt242xIpHhPT2llg>

EDA

Basic : <https://www.kaggle.com/code/imoore/intro-to-exploratory-data-analysis-eda-in-python>

EDA process from Kaggle :

## 1. Importing the required libraries for EDA

## 2. Loading the data into the data frame.[¶](https://www.kaggle.com/code/imoore/intro-to-exploratory-data-analysis-eda-in-python#2.-Loading-the-data-into-the-data-frame.)

## 3. Checking the types of data

## 4. Dropping irrelevant columns

## 5. Renaming the columns[¶](https://www.kaggle.com/code/imoore/intro-to-exploratory-data-analysis-eda-in-python#5.-Renaming-the-columns)

## 6. Dropping the duplicate rows

## 7. Dropping the missing or null values.

## 8. Detecting Outliers

## 9. Plot different features against one another (scatter), against frequency (histogram)

# Details EDA : https://www.kaggle.com/code/ekami66/detailed-exploratory-data-analysis-with-python

Practice project : https://www.kaggle.com/code/spscientist/a-simple-tutorial-on-exploratory-data-analysis

FEATURE ENGINEERING

Basic : <https://www.kaggle.com/learn/feature-engineering>