Linear Regression
Diamond Price Prediction

- **Design:**

- This project is one of the T5 Data Science Boot Camp requirements. Data provided by Kaggle ([here](#)). In this module, we will be laying the foundation for our analysis by processing and exploring a large amount of data on diamond datasets. This dataset has been made available thanks to Kaggle which is the home for many such datasets and competitions.

- **Understanding the dataset:**

- This dataset contains the prices and other attributes of almost 54,000 diamonds. It's a great dataset for beginners learning to work with data analysis and visualization. In this data set there are 11columns. Their names and data types as follows:

- Fields include

Price	Price in US dollars (\$326--\$18,823)
Carat	Weight of the diamond (0.2--5.01)
Cut	quality of the cut (Fair, Good, Very Good, Premium, Ideal)
Color	Diamond color, from J (worst) to D (best)
Clarity	A measurement of how clear the diamond is (I1 (worst), SI2, SI1, VS2, VS1, VVS2, VVS1, IF (best))
X	Length in mm (0--10.74)
Y	Width in mm (0--58.9)
Z	Depth in mm (0--31.8)
Depth	Total depth percentage = z / mean (x, y) = 2 * z / (x + y) (43--79)
Table	Width of top of diamond relative to widest point (43--95)

- **Algorithms:**

_____

- **Use panda's library to analyze different features of the dataset, which includes :**

  - Read the dataset.
  - Cleaning the data and remove null values.
  - Remove the outlier from dataset.
  - Analyze the price of diamond with the respect to various features.
  - Use seaborn and matplotlib library to visualize the given results.
  - Plot graphs like bar graphs and pie chart.
  - Baseline Model.
  - Data Preparation.
  - Experimentations.

- **Tools :**

- •Pandas for data manipulation

- • IQR for discover outliers

- • Remove Duplicate or unnecessary data

- • Matplotlib for plotting

- • Seaborn for visualizations

- • Sklearn Linear Regression library

Communication:

• The slides will be provided here, Feel free to any pull requests besides details are provided at the readme of the project.