**Abstract:**

The dataset contains 8248 records of used cars collected from syarah.com. Each row the brand name, model, manufacturing year, origin, the color of the car, options, capacity of the engine, type of fuel, transmission type, the mileage that the car covered, region price, and negotiable.

**Design:**

This project is one of the T5 Data Science BootCamp requirements. Data obtained from Kaggle website.

<https://www.kaggle.com/turkibintalib/saudi-arabia-used-cars-dataset>

**Data:**

This dataset consists of 7 features belonging to 8035 shopping sessions. The 'price' attribute is the target feature. The dataset is clean and there are no missing values.

**Algorithms:**

**Feature Engineering:**

1. Converting categorical attributes to ordered factor variables and are numerically encoded.
2. Normalize numerical variables of the dataset for clustering and scale for classification methods.

**Models Used:**

* Linear regression

**Tools:**

1. Numpy and Pandas for data manipulation.
2. Matplotlib and Seaborn for plotting.