Splitting the Data into Independent and Dependent Variables

```
[30] X=car[['name','company','year','kms_driven','fuel_type']]
     y=car['Price']
```

```
[31] X
```

	name	company	year	kms_driven	fuel_type
0	Hyundai Santro Xing	Hyundai	2007	45000	Petrol
1	Mahindra Jeep CL550	Mahindra	2006	40	Diesel
2	Hyundai Grand i10	Hyundai	2014	28000	Petrol
3	Ford EcoSport Titanium	Ford	2014	36000	Diesel
4	Ford Figo	Ford	2012	41000	Diesel
...	...	...	...	...	...
811	Maruti Suzuki Ritz	Maruti	2011	50000	Petrol
812	Tata Indica V2	Tata	2009	30000	Diesel
813	Toyota Corolla Altis	Toyota	2009	132000	Petrol
814	Tata Zest XM	Tata	2018	27000	Diesel
815	Mahindra Quanto C8	Mahindra	2013	40000	Diesel

815 rows × 5 columns

```
[32] y.shape
```

```
(815,)
```

Applying Train Test Split

```
[33] from sklearn.model_selection import train_test_split
     X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2)
```

```
[34] from sklearn.linear_model import LinearRegression
```

```
[35] from sklearn.preprocessing import OneHotEncoder
     from sklearn.compose import make_column_transformer
     from sklearn.pipeline import make_pipeline
     from sklearn.metrics import r2_score
```