multiple-linear-regression

March 26, 2024

Import python libraries

```
[1]: import pandas as pd
     import numpy as np
     import matplotlib.pyplot as plt
```

Load Dataset

```
[2]: dataset=pd.read_csv('/content/mlr07.csv')
```

```
[3]: dataset.head()
```

```
[3]:       X1  X2    X3   X4   X5
     0    8.0  78   284  9.1  109
     1    9.3  68   433  8.7  144
     2    7.5  70   739  7.2  113
     3    8.9  96  1792  8.9   97
     4   10.2  74   477  8.3  206
```

Create variable where x is our independent variable and y is our dependent or output variable

```
[4]: y = dataset.iloc[:,:1].values
```

```
[5]: x = dataset.iloc[:,1:].values
```

```
[6]: x.shape
```

```
[6]: (53, 4)
```

```
[7]: #train and test the model
     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,
      ↪random_state = 0)
```

```
[9]: x_test.shape
```

```
[9]: (11, 4)
```

```
[10]: x_train.shape
```

```
[10]: (42, 4)
```

```
[11]: from sklearn.linear_model import LinearRegression
```

```
[12]: regressor=LinearRegression()
```

```
[14]: regressor.fit(x_train,y_train)
```

```
[14]: LinearRegression()
```

```
[16]: pred=regressor.predict(x_test)
```

```
[17]: pred
```

```
[17]: array([[10.10478132],
              [ 9.68000462],
              [ 8.84702943],
              [ 9.74088989],
              [ 8.69956111],
              [ 9.86249562],
              [ 9.64624096],
              [ 9.75831054],
              [ 9.02185963],
              [ 9.36070494],
              [ 8.31196194]])
```

```
[18]: y_test
```

```
[18]: array([[ 8.39999962],
              [ 9.80000019],
              [ 7.30000019],
              [10.80000019],
              [ 8.30000019],
              [ 7.5       ],
              [ 9.80000019],
              [ 9.80000019],
              [ 8.39999962],
              [ 9.10000038],
              [10.19999981]])
```