

CH :- 4 Introduction to M/c using Python

Feature Transformation :

```
In [17]: # Feature Additon
import pandas as pd
room_length = [18,20,10,12,18,11]
room_breadth = [20,20,10,11,19,10]
room_type = ["Big", "Big", "Normal", "Normal", "Big", "Normal"]
```

```
In [18]: data = pd.DataFrame({
    "length":room_length,
    "Breadth":room_breadth,
    "Type":room_type
})
data
```

Out[18]:

	length	Breadth	Type
0	18	20	Big
1	20	20	Big
2	10	10	Normal
3	12	11	Normal
4	18	19	Big
5	11	10	Normal

```
In [19]: data["Area"] = data["length"]*data["Breadth"]
data
# EDA : Exploratory Data Analysis
```

Out[19]:

	length	Breadth	Type	Area
0	18	20	Big	360
1	20	20	Big	400
2	10	10	Normal	100
3	12	11	Normal	132
4	18	19	Big	342
5	11	10	Normal	110

```
In [20]: pd.get_dummies(data) # categorical data ne column ma convert kare.
```

Out[20]:

	length	Breadth	Area	Type_Big	Type_Normal
0	18	20	360	1	0
1	20	20	400	1	0
2	10	10	100	0	1
3	12	11	132	0	1
4	18	19	342	1	0
5	11	10	110	0	1

drop_first = True : - first column remove kare

```
In [8]: data = pd.get_dummies(data,drop_first=True)
data
```

Out[8]:

	length	Breadth	Area	Type_Normal
0	18	20	360	0
1	20	20	400	0
2	10	10	100	1
3	12	11	132	1
4	18	19	342	0
5	11	10	110	1

```
In [26]: import pandas as pd
import numpy as np
apartment_area = [4720,2430,4368,3969,6142,7912]
apartment_price = [2360000,1215000,2184000,1984500,3071000,3956000]
data = pd.DataFrame({"Area":apartment_area,"Price":apartment_price})
data
```

Out[26]:

	Area	Price
0	4720	2360000
1	2430	1215000
2	4368	2184000
3	3969	1984500
4	6142	3071000
5	7912	3956000

```
In [27]: data["Price"] = np.where(data["Price"]>3000000,"High",np.where(data["Price"]<2000000,"Low",data["Price"]>2000000&&data["Price"]<3000000,"Medium"))
```

Out[27]:

	Area	Price
0	4720	Medium
1	2430	Low
2	4368	Medium
3	3969	Low
4	6142	High
5	7912	High

In []: