In [1]:
```
import pandas as pd
```

In [2]:
```
hsp=pd.read_csv("HousePrices.csv")
```

In [3]:
```
hsp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2073 entries, 0 to 2072
Data columns (total 81 columns):
 #   Column               Non-Null Count  Dtype  
---  ------               --------------  -----  
 0   Id                   2073 non-null   int64  
 1   Dwell_Type           2073 non-null   int64  
 2   Zone_Class           2073 non-null   object 
 3   LotFrontage          1753 non-null   float64
 4   LotArea              2073 non-null   int64  
 5   Road_Type            2073 non-null   object 
 6   Alley                129 non-null    object 
 7   Property_Shape       2073 non-null   object 
 8   LandContour          2073 non-null   object 
 9   Utilities            2073 non-null   object 
 10  LotConfig            2073 non-null   object 
 11  LandSlope            2073 non-null   object 
 12  Neighborhood         2073 non-null   object 
 13  Condition1           2073 non-null   object 
 14  Condition2           2073 non-null   object 
 15  Dwelling_Type        2073 non-null   object 
 16  HouseStyle           2073 non-null   object 
 17  OverallQual          2073 non-null   int64  
 18  OverallCond          2073 non-null   int64  
 19  YearBuilt            2073 non-null   int64  
 20  YearRemodAdd         2073 non-null   int64  
 21  RoofStyle            2073 non-null   object 
 22  RoofMatl             2073 non-null   object 
 23  Exterior1st          2073 non-null   object 
 24  Exterior2nd          2073 non-null   object 
 25  MasVnrType           2059 non-null   object 
 26  MasVnrArea           2059 non-null   float64
 27  ExterQual            2073 non-null   object 
 28  ExterCond            2073 non-null   object 
 29  Foundation           2073 non-null   object 
 30  BsmtQual             2014 non-null   object 
 31  BsmtCond             2014 non-null   object 
 32  BsmtExposure         2012 non-null   object 
 33  BsmtFinType1         2014 non-null   object 
```

```
34  BsmtFinSF1          2073 non-null   int64
35  BsmtFinType2        2013 non-null   object
36  BsmtFinSF2          2073 non-null   int64
37  BsmtUnfSF           2073 non-null   int64
38  TotalBsmtSF         2073 non-null   int64
39  Heating             2073 non-null   object
40  HeatingQC           2073 non-null   object
41  CentralAir          2073 non-null   object
42  Electrical          2072 non-null   object
43  1stFlrSF            2073 non-null   int64
44  2ndFlrSF            2073 non-null   int64
45  LowQualFinSF        2073 non-null   int64
46  GrLivArea           2073 non-null   int64
47  BsmtFullBath        2073 non-null   int64
48  BsmtHalfBath        2073 non-null   int64
49  FullBath            2073 non-null   int64
50  HalfBath            2073 non-null   int64
51  BedroomAbvGr        2073 non-null   int64
52  KitchenAbvGr        2073 non-null   int64
53  KitchenQual         2073 non-null   object
54  TotRmsAbvGrd        2073 non-null   int64
55  Functional          2073 non-null   object
56  Fireplaces          2073 non-null   int64
57  FireplaceQu         1085 non-null   object
58  GarageType          1960 non-null   object
59  GarageYrBlt         1960 non-null   float64
60  GarageFinish        1960 non-null   object
61  GarageCars          2073 non-null   int64
62  GarageArea          2073 non-null   int64
63  GarageQual          1960 non-null   object
64  GarageCond          1960 non-null   object
65  PavedDrive          2073 non-null   object
66  WoodDeckSF          2073 non-null   int64
67  OpenPorchSF         2073 non-null   int64
68  EnclosedPorch       2073 non-null   int64
69  3SsnPorch           2073 non-null   int64
70  ScreenPorch         2073 non-null   int64
71  PoolArea            2073 non-null   int64
72  PoolQC              8 non-null      object
73  Fence               404 non-null    object
74  MiscFeature         80 non-null     object
75  MiscVal             2073 non-null   int64
```

```
 76  MoSold               2073 non-null   int64
 77  YrSold               2073 non-null   int64
 78  SaleType             2073 non-null   object
 79  SaleCondition        2073 non-null   object
 80  Property_Sale_Price  2073 non-null   int64
dtypes: float64(3), int64(35), object(43)
memory usage: 1.3+ MB
```

In [4]:
```
cl_to_drop=["LotFrontage","MasVnrArea","Alley","BsmtExposure","BsmtFinType1","BsmtFinType2","BsmtQual","BsmtCond","PoolQ
```

In [5]:
```
hsp.drop(cl_to_drop,axis=1,inplace=True)
```

In [6]:
```
hsp
```

Out[6]:

	Id	Dwell_Type	Zone_Class	LotArea	Road_Type	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	...	EnclosedPorch	3SsnPor
0	1	60	RL	8450	Pave	Lvl	AllPub	Inside	Gtl	CollgCr	...	0	
1	2	20	RL	9600	Pave	Lvl	AllPub	FR2	Gtl	Veenker	...	0	
2	3	60	RL	11250	Pave	Lvl	AllPub	Inside	Gtl	CollgCr	...	0	
3	4	70	RL	9550	Pave	Lvl	AllPub	Corner	Gtl	Crawfor	...	272	
4	5	60	RL	14260	Pave	Lvl	AllPub	FR2	Gtl	NoRidge	...	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	
2068	942	60	RL	8755	Pave	Lvl	AllPub	FR2	Gtl	Gilbert	...	0	1
2069	943	90	RL	7711	Pave	Lvl	AllPub	Inside	Gtl	Edwards	...	0	
2070	944	90	RL	25000	Pave	Low	AllPub	Inside	Gtl	Mitchel	...	0	
2071	945	20	RL	14375	Pave	Lvl	NoSeWa	CulDSac	Gtl	Timber	...	0	
2072	946	50	RM	8820	Pave	Lvl	AllPub	Corner	Gtl	OldTown	...	244	

2073 rows × 63 columns

In [7]:
```
import sklearn
from sklearn.preprocessing import LabelEncoder
```

In [8]:
```
col_name_objects = hsp.select_dtypes(include=['object']).columns.values
LE = LabelEncoder()
for col in col_name_objects:
    hsp[col] = LE.fit_transform(hsp[col])
hsp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2073 entries, 0 to 2072
Data columns (total 63 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Id                   2073 non-null   int64
 1   Dwell_Type           2073 non-null   int64
 2   Zone_Class           2073 non-null   int32
 3   LotArea              2073 non-null   int64
 4   Road_Type            2073 non-null   int32
 5   LandContour          2073 non-null   int32
 6   Utilities            2073 non-null   int32
 7   LotConfig            2073 non-null   int32
 8   LandSlope            2073 non-null   int32
 9   Neighborhood         2073 non-null   int32
 10  Condition1           2073 non-null   int32
 11  Condition2           2073 non-null   int32
 12  Dwelling_Type        2073 non-null   int32
 13  HouseStyle           2073 non-null   int32
 14  OverallQual          2073 non-null   int64
 15  OverallCond          2073 non-null   int64
 16  YearBuilt            2073 non-null   int64
 17  YearRemodAdd         2073 non-null   int64
 18  RoofStyle            2073 non-null   int32
 19  RoofMatl             2073 non-null   int32
 20  Exterior1st          2073 non-null   int32
 21  Exterior2nd          2073 non-null   int32
 22  MasVnrType           2073 non-null   int32
 23  ExterQual            2073 non-null   int32
 24  ExterCond            2073 non-null   int32
 25  Foundation           2073 non-null   int32
 26  BsmtFinSF1           2073 non-null   int64
 27  BsmtFinSF2           2073 non-null   int64
 28  BsmtUnfSF            2073 non-null   int64
 29  TotalBsmtSF          2073 non-null   int64
```

```
 30  Heating              2073 non-null   int32
 31  HeatingQC            2073 non-null   int32
 32  CentralAir           2073 non-null   int32
 33  Electrical           2073 non-null   int32
 34  1stFlrSF             2073 non-null   int64
 35  2ndFlrSF             2073 non-null   int64
 36  LowQualFinSF         2073 non-null   int64
 37  GrLivArea            2073 non-null   int64
 38  BsmtFullBath         2073 non-null   int64
 39  BsmtHalfBath         2073 non-null   int64
 40  FullBath             2073 non-null   int64
 41  HalfBath             2073 non-null   int64
 42  BedroomAbvGr         2073 non-null   int64
 43  KitchenAbvGr         2073 non-null   int64
 44  KitchenQual          2073 non-null   int32
 45  TotRmsAbvGrd         2073 non-null   int64
 46  Functional           2073 non-null   int32
 47  Fireplaces           2073 non-null   int64
 48  GarageCars           2073 non-null   int64
 49  GarageArea           2073 non-null   int64
 50  PavedDrive           2073 non-null   int32
 51  WoodDeckSF           2073 non-null   int64
 52  OpenPorchSF          2073 non-null   int64
 53  EnclosedPorch        2073 non-null   int64
 54  3SsnPorch            2073 non-null   int64
 55  ScreenPorch          2073 non-null   int64
 56  PoolArea             2073 non-null   int64
 57  MiscVal              2073 non-null   int64
 58  MoSold               2073 non-null   int64
 59  YrSold               2073 non-null   int64
 60  SaleType             2073 non-null   int32
 61  SaleCondition        2073 non-null   int32
 62  Property_Sale_Price  2073 non-null   int64
dtypes: int32(28), int64(35)
memory usage: 793.7 KB
```

In [9]:
```
hsp.iloc[:,22:]
```

Out[9]:

	MasVnrType	ExterQual	ExterCond	Foundation	BsmtFinSF1	BsmtFinSF2	BsmtUnfSF	TotalBsmtSF	Heating	HeatingQC	...	EnclosedPorch
0	1	2	4	2	706	0	150	856	1	0	...	0
1	2	3	4	1	978	0	284	1262	1	0	...	0
2	1	2	4	2	486	0	434	920	1	0	...	0
3	2	3	4	0	216	0	540	756	1	2	...	272
4	1	2	4	2	655	0	490	1145	1	0	...	0
...	...	...	...	...	...	...	...	...	...	...	...	...
2068	1	2	4	2	772	0	220	992	1	0	...	0
2069	2	3	4	2	1440	0	0	1440	1	4	...	0
2070	2	3	4	1	0	0	1632	1632	1	4	...	0
2071	1	3	4	1	111	354	354	819	1	2	...	0
2072	2	3	4	0	1088	0	0	1088	1	4	...	244

2073 rows × 41 columns

In [10]:
```
hsp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2073 entries, 0 to 2072
Data columns (total 63 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Id                   2073 non-null   int64
 1   Dwell_Type           2073 non-null   int64
 2   Zone_Class           2073 non-null   int32
 3   LotArea              2073 non-null   int64
 4   Road_Type            2073 non-null   int32
 5   LandContour          2073 non-null   int32
 6   Utilities            2073 non-null   int32
 7   LotConfig            2073 non-null   int32
 8   LandSlope            2073 non-null   int32
 9   Neighborhood         2073 non-null   int32
 10  Condition1           2073 non-null   int32
 11  Condition2           2073 non-null   int32
 12  Dwelling_Type        2073 non-null   int32
 13  HouseStyle           2073 non-null   int32
 14  OverallQual          2073 non-null   int64
 15  OverallCond          2073 non-null   int64
 16  YearBuilt            2073 non-null   int64
 17  YearRemodAdd         2073 non-null   int64
 18  RoofStyle            2073 non-null   int32
 19  RoofMatl             2073 non-null   int32
 20  Exterior1st          2073 non-null   int32
 21  Exterior2nd          2073 non-null   int32
 22  MasVnrType           2073 non-null   int32
 23  ExterQual            2073 non-null   int32
 24  ExterCond            2073 non-null   int32
 25  Foundation           2073 non-null   int32
 26  BsmtFinSF1           2073 non-null   int64
 27  BsmtFinSF2           2073 non-null   int64
 28  BsmtUnfSF            2073 non-null   int64
 29  TotalBsmtSF          2073 non-null   int64
 30  Heating              2073 non-null   int32
 31  HeatingQC            2073 non-null   int32
 32  CentralAir           2073 non-null   int32
 33  Electrical           2073 non-null   int32
```

```
 34  1stFlrSF             2073 non-null   int64
 35  2ndFlrSF             2073 non-null   int64
 36  LowQualFinSF         2073 non-null   int64
 37  GrLivArea            2073 non-null   int64
 38  BsmtFullBath         2073 non-null   int64
 39  BsmtHalfBath         2073 non-null   int64
 40  FullBath             2073 non-null   int64
 41  HalfBath             2073 non-null   int64
 42  BedroomAbvGr         2073 non-null   int64
 43  KitchenAbvGr         2073 non-null   int64
 44  KitchenQual          2073 non-null   int32
 45  TotRmsAbvGrd         2073 non-null   int64
 46  Functional           2073 non-null   int32
 47  Fireplaces           2073 non-null   int64
 48  GarageCars           2073 non-null   int64
 49  GarageArea           2073 non-null   int64
 50  PavedDrive           2073 non-null   int32
 51  WoodDeckSF           2073 non-null   int64
 52  OpenPorchSF          2073 non-null   int64
 53  EnclosedPorch        2073 non-null   int64
 54  3SsnPorch            2073 non-null   int64
 55  ScreenPorch          2073 non-null   int64
 56  PoolArea             2073 non-null   int64
 57  MiscVal              2073 non-null   int64
 58  MoSold               2073 non-null   int64
 59  YrSold               2073 non-null   int64
 60  SaleType             2073 non-null   int32
 61  SaleCondition        2073 non-null   int32
 62  Property_Sale_Price  2073 non-null   int64
dtypes: int32(28), int64(35)
memory usage: 793.7 KB
```

In [11]:
```
c_to_mean=["MasVnrType"]
```

In [12]:
```
hsp["Electrical"]=hsp["Electrical"].fillna(0)
```

In [13]:
```
import numpy as np
```

In [14]:
```
hsp[c_to_mean]=hsp[c_to_mean].fillna(np.mean)
```

In [15]:
```
hsp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2073 entries, 0 to 2072
Data columns (total 63 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Id                   2073 non-null   int64
 1   Dwell_Type           2073 non-null   int64
 2   Zone_Class           2073 non-null   int32
 3   LotArea              2073 non-null   int64
 4   Road_Type            2073 non-null   int32
 5   LandContour          2073 non-null   int32
 6   Utilities            2073 non-null   int32
 7   LotConfig            2073 non-null   int32
 8   LandSlope            2073 non-null   int32
 9   Neighborhood         2073 non-null   int32
 10  Condition1           2073 non-null   int32
 11  Condition2           2073 non-null   int32
 12  Dwelling_Type        2073 non-null   int32
 13  HouseStyle           2073 non-null   int32
 14  OverallQual          2073 non-null   int64
 15  OverallCond          2073 non-null   int64
 16  YearBuilt            2073 non-null   int64
 17  YearRemodAdd         2073 non-null   int64
 18  RoofStyle            2073 non-null   int32
 19  RoofMatl             2073 non-null   int32
 20  Exterior1st          2073 non-null   int32
 21  Exterior2nd          2073 non-null   int32
 22  MasVnrType           2073 non-null   int32
 23  ExterQual            2073 non-null   int32
 24  ExterCond            2073 non-null   int32
 25  Foundation           2073 non-null   int32
 26  BsmtFinSF1           2073 non-null   int64
 27  BsmtFinSF2           2073 non-null   int64
 28  BsmtUnfSF            2073 non-null   int64
 29  TotalBsmtSF          2073 non-null   int64
 30  Heating              2073 non-null   int32
 31  HeatingQC            2073 non-null   int32
 32  CentralAir           2073 non-null   int32
 33  Electrical           2073 non-null   int32
```

```
 34  1stFlrSF             2073 non-null   int64
 35  2ndFlrSF             2073 non-null   int64
 36  LowQualFinSF         2073 non-null   int64
 37  GrLivArea            2073 non-null   int64
 38  BsmtFullBath         2073 non-null   int64
 39  BsmtHalfBath         2073 non-null   int64
 40  FullBath             2073 non-null   int64
 41  HalfBath             2073 non-null   int64
 42  BedroomAbvGr         2073 non-null   int64
 43  KitchenAbvGr         2073 non-null   int64
 44  KitchenQual          2073 non-null   int32
 45  TotRmsAbvGrd         2073 non-null   int64
 46  Functional           2073 non-null   int32
 47  Fireplaces           2073 non-null   int64
 48  GarageCars           2073 non-null   int64
 49  GarageArea           2073 non-null   int64
 50  PavedDrive           2073 non-null   int32
 51  WoodDeckSF           2073 non-null   int64
 52  OpenPorchSF          2073 non-null   int64
 53  EnclosedPorch        2073 non-null   int64
 54  3SsnPorch            2073 non-null   int64
 55  ScreenPorch          2073 non-null   int64
 56  PoolArea             2073 non-null   int64
 57  MiscVal              2073 non-null   int64
 58  MoSold               2073 non-null   int64
 59  YrSold               2073 non-null   int64
 60  SaleType             2073 non-null   int32
 61  SaleCondition        2073 non-null   int32
 62  Property_Sale_Price  2073 non-null   int64
dtypes: int32(28), int64(35)
memory usage: 793.7 KB
```

In [16]:
```
hsp.isnull().sum().head(50)
```

Out[16]:
```
Id              0
Dwell_Type      0
Zone_Class      0
LotArea         0
Road_Type       0
LandContour     0
Utilities       0
LotConfig       0
LandSlope       0
Neighborhood    0
Condition1      0
Condition2      0
Dwelling_Type   0
HouseStyle      0
OverallQual     0
OverallCond     0
YearBuilt       0
YearRemodAdd    0
RoofStyle       0
RoofMatl        0
Exterior1st     0
Exterior2nd     0
MasVnrType      0
ExterQual       0
ExterCond       0
Foundation      0
BsmtFinSF1      0
BsmtFinSF2      0
BsmtUnfSF       0
TotalBsmtSF     0
Heating         0
HeatingQC       0
CentralAir      0
Electrical      0
1stFlrSF        0
2ndFlrSF        0
LowQualFinSF    0
GrLivArea       0
BsmtFullBath    0
```

```
BsmtHalfBath     0
FullBath         0
HalfBath         0
BedroomAbvGr     0
KitchenAbvGr     0
KitchenQual      0
TotRmsAbvGrd     0
Functional       0
Fireplaces       0
GarageCars       0
GarageArea       0
dtype: int64
```

In [17]:
```
hsp.describe()
```

Out[17]:

	Id	Dwell_Type	Zone_Class	LotArea	Road_Type	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	...
count	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	2073.000000	...
mean	916.132176	60.556199	3.027014	10717.853353	0.995176	2.768934	0.001447	2.992282	0.062229	12.273517	...
std	493.014670	159.924810	0.651138	9215.982306	0.069303	0.719395	0.038023	1.639268	0.268138	5.984291	...
min	1.000000	20.000000	0.000000	1300.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...
25%	519.000000	20.000000	3.000000	7620.000000	1.000000	3.000000	0.000000	1.000000	0.000000	7.000000	...
50%	932.000000	50.000000	3.000000	9492.000000	1.000000	3.000000	0.000000	4.000000	0.000000	12.000000	...
75%	1302.000000	70.000000	3.000000	11601.000000	1.000000	3.000000	0.000000	4.000000	0.000000	17.000000	...
max	1820.000000	7080.000000	4.000000	215245.000000	1.000000	3.000000	1.000000	4.000000	2.000000	24.000000	...

8 rows × 63 columns

In [18]:
```
hsp.shape
```

Out[18]:
```
(2073, 63)
```

In [19]:
```
hsp_corr=hsp.corr()
hsp_corr
```

Out[19]:

	Id	Dwell_Type	Zone_Class	LotArea	Road_Type	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	...	E
Id	1.000000	0.043264	0.008184	0.004201	-0.009714	-0.023591	0.002230	0.015822	0.005695	-0.004485	...	
Dwell_Type	0.043264	1.000000	0.002442	-0.033609	-0.005419	0.010011	-0.009657	0.029538	0.068747	-0.020327	...	
Zone_Class	0.008184	0.002442	1.000000	-0.038714	0.056364	-0.000062	-0.001580	-0.017439	-0.017926	-0.242801	...	
LotArea	0.004201	-0.033609	-0.038714	1.000000	-0.144528	-0.141765	0.015111	-0.103496	0.357898	0.029985	...	
Road_Type	-0.009714	-0.005419	0.056364	-0.144528	1.000000	0.161557	0.002650	0.042154	-0.191610	0.007838	...	
...	...	...	...	...	...	...	...	...	...	...	...	
MoSold	0.000234	-0.020589	-0.062572	-0.001707	0.007443	0.007161	-0.074844	0.005100	-0.002735	0.013191	...	
YrSold	0.043219	0.018891	0.018558	-0.006478	-0.049654	0.033331	0.032878	-0.017574	-0.012042	0.055070	...	
SaleType	0.001829	0.011964	0.100362	0.005604	0.021432	-0.031928	-0.177783	0.022360	0.061987	-0.048993	...	
SaleCondition	-0.011822	0.005838	-0.007771	0.032000	-0.003179	0.023873	-0.125765	0.074189	-0.034126	0.006096	...	
Property_Sale_Price	-0.002491	-0.017399	-0.168772	0.211572	0.054667	0.027606	-0.021040	-0.061008	0.038989	0.199664	...	

63 rows × 63 columns

In [20]:
```
x_ind=hsp.drop("Property_Sale_Price",axis=1)
y_dep=hsp.Property_Sale_Price
```

In [21]:
```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x_ind, y_dep, test_size=0.2, random_state=2)
```

In [22]:
```
from sklearn.preprocessing import StandardScaler
SE=StandardScaler()
```

In [23]:
```
x_norm=SE.fit_transform(x_ind)
```

In [24]:
```
from sklearn.decomposition import PCA
pca=PCA()
x_new=pca.fit_transform(x_norm)
```

In [25]:
```
x_new
```

Out[25]:
```
array([[ 1.84824349e+00,  5.59011719e-01, -1.99890156e+00, ...,
        -1.16623086e-01, -3.21656412e-15, -1.02708580e-15],
       [-1.93829457e-01, -1.04516637e+00,  9.96206545e-01, ...,
         1.53225426e-01,  7.00247657e-15,  6.70217178e-16],
       [ 2.07410097e+00,  4.42450697e-01, -1.81999917e+00, ...,
        -1.90513304e-01, -6.92494848e-15, -7.69594997e-16],
       ...,
       [-2.31264441e-01,  1.58238102e+00,  9.21368060e-01, ...,
        -1.88347921e-01,  8.50609627e-16, -1.67741439e-16],
       [-1.14325672e+00,  6.15411931e-01,  3.77225825e+00, ...,
        -1.80843607e-02, -3.56923496e-16, -1.55034579e-17],
       [-1.66425357e+00,  4.45186117e-01,  2.16840995e+00, ...,
         4.40974907e-01, -5.80273190e-16,  9.62128311e-17]])
```

In [27]:
```
explained_varience=pca.explained_variance_ratio_
explained_varience
```

Out[27]:
```
array([1.22313074e-01, 5.57268366e-02, 4.67845196e-02, 3.97304538e-02,
       3.23012404e-02, 2.94044488e-02, 2.61980741e-02, 2.46682179e-02,
       2.32654149e-02, 2.18114541e-02, 2.10036444e-02, 2.06357227e-02,
       2.01021651e-02, 1.94196000e-02, 1.89834735e-02, 1.86974724e-02,
       1.85140980e-02, 1.77101397e-02, 1.70785697e-02, 1.62798682e-02,
       1.61895761e-02, 1.60302672e-02, 1.58778655e-02, 1.53591971e-02,
       1.50214476e-02, 1.45889165e-02, 1.44702034e-02, 1.41937645e-02,
       1.38122019e-02, 1.36574382e-02, 1.32763433e-02, 1.25941685e-02,
       1.23348912e-02, 1.21659705e-02, 1.21055209e-02, 1.18597834e-02,
       1.13574220e-02, 1.12031582e-02, 1.05312864e-02, 9.93820648e-03,
       9.74987575e-03, 9.52897588e-03, 9.31929282e-03, 9.15011412e-03,
       8.70182024e-03, 8.32817261e-03, 7.73209575e-03, 7.48727425e-03,
       7.39776704e-03, 7.02149984e-03, 6.56854336e-03, 5.67697205e-03,
       5.04655562e-03, 4.53141383e-03, 4.27295141e-03, 3.50144536e-03,
       2.92655264e-03, 2.31556858e-03, 2.06138582e-03, 1.48560650e-03,
       2.41220005e-32, 8.94768612e-34])
```

In [28]:
```
ev=pd.DataFrame(explained_varience)
```

In [29]:
```
ev.iloc[0:50].sum()
```

Out[29]:
```
0    0.961613
dtype: float64
```

In [30]:
```
pca1=PCA(n_components=50)
x_new_model=pca1.fit_transform(x_norm)
```

In [31]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x_new_model,y_dep,train_size=0.8,random_state=3)
```

In [32]:
```
from sklearn import linear_model
from sklearn.linear_model import LinearRegression
model_li=LinearRegression()
```

In [33]:
```
y_pred_li=model_li.fit(x_train,y_train).predict(x_test)
```

In [34]:
```
model_li.score(x_test,y_test)
```

Out[34]:
```
0.8586939425814789
```

By using PCA with Linear regression,got the accuracy of 85%

In [ ]:
```

```