

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import accuracy_score
import warnings
warnings.filterwarnings('ignore')

In [181]: house_price = pd.read_csv("train.csv")

In [182]: house_price.head()

Out[182]:
```

	ID	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	LandContour	Utilities	...	PoolArea	PoolQC	Fence	MiscFeature	MiscVal	MoSold	YrSold	SaleType	SalePrice
0	1	60	RL	65.0	8450	Pave	NaN	Reg	Lvl	AllPub	...	0	NaN	NaN	NaN	0	2	2008	WD	14600000
1	2	20	RL	80.0	9600	Pave	NaN	Reg	Lvl	AllPub	...	0	NaN	NaN	NaN	0	5	2007	WD	14600000
2	3	60	RL	68.0	11250	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN	NaN	0	9	2008	WD	14600000
3	4	70	RL	60.0	9550	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN	NaN	0	2	2006	WD	14600000
4	5	60	RL	84.0	14260	Pave	NaN	IR1	Lvl	AllPub	...	0	NaN	NaN	NaN	0	12	2008	WD	14600000

5 rows x 81 columns

```
In [183]: # Data pre-processing
house_price.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1460 entries, 0 to 1459
Data columns (total 81 columns):
#   Column              Non-Null Count  Dtype  
---  --
0    Id                   1460 non-null   int64   
1    MSSubClass           1460 non-null   int64   
2    MSZoning             1460 non-null   object   
3    LotFrontage         1201 non-null   float64  
4    LotArea             1460 non-null   int64   
5    Street              1460 non-null   object   
6    Alley               91 non-null     object   
7    LotShape            1460 non-null   object   
8    LandContour         1460 non-null   object   
9    Utilities           1460 non-null   object   
10   LotConfig           1460 non-null   object   
11   Landslope           1460 non-null   object   
12   Neighborhood        1460 non-null   object   
13   Condition1          1460 non-null   object   
14   Condition2          1460 non-null   object   
15   BldgType            1460 non-null   object   
16   HouseStyle          1460 non-null   object   
17   OverallQual         1460 non-null   int64   
18   OverallCond         1460 non-null   int64   
19   YearBuilt           1459 non-null   int64   
20   YearRemodAdd       1459 non-null   int64   
21   RoofStyle          1460 non-null   object   
22   RoofMatl           1460 non-null   object   
23   Exterior1st        1460 non-null   object   
24   Exterior2nd        1460 non-null   object   
25   MasVnrType         1452 non-null   object   
26   MasVnrArea         1452 non-null   float64  
27   ExterQual           1460 non-null   object   
28   ExterCond          1460 non-null   object   
29   Foundation         1460 non-null   object   
30   BsmtQual           1423 non-null   object   
31   BsmtCond           1423 non-null   object   
32   BsmtExposure       1422 non-null   object   
33   BsmtFinType1       1423 non-null   object   
34   BsmtFinSF1         1460 non-null   int64   
35   BsmtFinType2       1422 non-null   object   
36   BsmtFinSF2         1460 non-null   int64   
37   BsmtUnfSF          1460 non-null   int64   
38   TotalBsmtSF        1460 non-null   int64   
39   Heating            1460 non-null   object   
40   HeatingQC          1460 non-null   object   
41   CentralAir         1460 non-null   object   
42   Electrical         1459 non-null   object   
43   1stFlrSF           1460 non-null   int64   
44   2ndFlrSF           1460 non-null   int64   
45   LowQualFinSF       1460 non-null   int64   
46   GrLivArea          1460 non-null   int64   
47   BsmtFullBath       1460 non-null   int64   
48   BsmtHalfBath       1460 non-null   int64   
49   FullBath           1460 non-null   int64   
50   HalfBath           1460 non-null   int64   
51   BedroomAbvGr       1460 non-null   int64   
52   KitchenAbvGr       1460 non-null   int64   
53   KitchenQual        1460 non-null   object   
54   TotRmsAbvGrd       1460 non-null   int64   
55   Functional         1460 non-null   object   
56   Fireplaces         1460 non-null   object   
57   FireplaceQu        770 non-null   object   
58   GarageType         1379 non-null   object   
59   GarageYrBlt        1379 non-null   float64  
60   GarageFinish       1379 non-null   object   
61   GarageCars         1460 non-null   int64   
62   GarageArea         1460 non-null   int64   
63   GarageQual         1379 non-null   object   
64   GarageCond         1379 non-null   object   
65   PavedDrive         1460 non-null   object   
66   WoodDeckSF         1460 non-null   int64   
67   OpenPorchSF       1460 non-null   int64   
68   EnclosedPorch      1460 non-null   int64   
69   3SsnPorch         1460 non-null   int64   
70   ScreenPorch        1460 non-null   int64   
71   PoolArea          1460 non-null   int64   
72   PoolQC             7 non-null     object   
73   Fence             281 non-null   object   
74   MiscFeature        54 non-null   object   
75   MiscVal            1460 non-null   int64   
76   MoSold             1460 non-null   int64   
77   YrSold             1460 non-null   int64   
78   SaleType           1460 non-null   object   
79   SaleCondition      1460 non-null   object   
80   SalePrice          1460 non-null   int64   
dtypes: float64(3), int64(35), object(43)
memory usage: 924.0+ KB

In [184]: house_price.describe()

Out[184]:
```

	ID	MSSubClass	MSZoning	LotFrontage	LotArea	OverallQual	OverallCond	YearBuilt	YearRemodAdd	MasVnrArea	BsmtFinSF1	...	WoodDeckSF	OpenPorchSF	EnclosedPorch	3SsnPorch	ScreenPorch	PoolArea	PoolQC	Fence	MiscFeature	MiscVal	MoSold	YrSold	SaleType	SalePrice
count	1460.000000	1460.000000	1201.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1452.000000	1460.000000	...	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000	1460.000000
mean	730.500000	56.897260	70.499958	10516.828082	6.099315	5.575342	1971.267808	1984.865753	103.685262	443.639726	...	94.244521	46.660274	61.616161	25.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2007.500000	WD	14600000
std	421.610009	42.300571	24.284752	9981.264932	1.382997	1.112799	30.202904	20.645407	181.066207	456.098091	...	125.338794	66.256028	61.616161	25.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2007.500000	WD	14600000
min	1.000000	20.000000	21.000000	1300.000000	1.000000	1.000000	1872.000000	1950.000000	0.000000	0.000000	...	0.000														