```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_
from sklearn.preprocessing import StandardS

df = pd.read_csv('data/')
df.shape
```

```
(1460, 81)
```

```
df.isnull().sum()
```

```
Id                 0
MSSubClass         0
MSZoning           0
LotFrontage      259
LotArea            0
Street             0
Alley           1369
LotShape           0
LandContour        0
Utilities          0
LotConfig          0
LandSlope          0
Neighborhood       0
Condition1         0
Condition2         0
BldgType           0
HouseStyle         0
OverallQual        0
OverallCond        0
YearBuilt          0
YearRemodAdd       0
RoofStyle          0
RoofMatl           0
Exterior1st        0
Exterior2st        0
MasVnrType         8
MasVnrArea         8
ExterQual          0
ExterCond          0
Foundation         0
BsmtQual          37
BsmtCond          37
BsmtExposure      38
BsmtFinType1      37
BsmtFinSf1         0
BsmtFinType2      38
BsmtFinSf2         0
BsmtUnfSF          0
TotalBsmtSF        0
Heating            0
HeatingQC          0
CentralAir         0
Electrical         1
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
FireplaceQu      690
GarageType        81
GarageYrBlt       81
GarageFinish      81
GarageCars         0
GarageArea         0
GarageQual        81
GarageCond        81
PavedDrive         0
WoodDeckSF         0
OpenPoolArea       0
PoolQC          1453
Fence           1179
MiscFeature     1406
MiscVal            0
MoSold             0
YrSold             0
SaleType           0
SaleCondition      0
dtype: int64
```

```
df.describe()
```

```
df_dtypes = df.dtypes
df_dtypes.value_counts()
```

```
int64    38
object   43
dtype: int64
```

```
df['Id'].sum()
```

```
1065930
```

```
df['MSSubClass'].sum()
```

```
60416
```

```
df['MSZoning'].sum()
```

```
'RL'    1151
'RM'     218
'FV'      65
'RH'      16
dtype: int64
```

```
df['LotFrontage'].sum()
```

```
259
```

```
df['LotArea'].sum()
```

```
0
```

```
df['Street'].sum()
```

```
0
```

```
df['Alley'].sum()
```

```
1369
```

```
df['LotShape'].sum()
```

```
0
```

```
df['LandContour'].sum()
```

```
0
```

```
df['Utilities'].sum()
```

```
0
```

```
df['LotConfig'].sum()
```

```
0
```

```
df['LandSlope'].sum()
```

```
0
```

```
df['Neighborhood'].sum()
```

```
0
```

```
df['Condition1'].sum()
```

```
0
```

```
df['Condition2'].sum()
```

```
0
```

```
df['BldgType'].sum()
```

```
0
```

```
df['HouseStyle'].sum()
```

```
0
```

```
df['OverallQual'].sum()
```

```
0
```

```
df['OverallCond'].sum()
```

```
0
```

```
df['YearBuilt'].sum()
```

```
0
```

```
df['YearRemodAdd'].sum()
```

```
0
```

```
df['RoofStyle'].sum()
```

```
0
```

```
df['RoofMatl'].sum()
```

```
0
```

```
df['Exterior1st'].sum()
```

```
0
```

```
df['Exterior2st'].sum()
```

```
0
```

```
df['MasVnrType'].sum()
```

```
8
```

```
df['MasVnrArea'].sum()
```

```
8
```

```
df['ExterQual'].sum()
```

```
0
```

```
df['ExterCond'].sum()
```

```
0
```

```
df['Foundation'].sum()
```

```
0
```

```
df['BsmtQual'].sum()
```

```
37
```

```
df['BsmtCond'].sum()
```

```
37
```

```
df['BsmtExposure'].sum()
```

```
38
```

```
df['BsmtFinType1'].sum()
```

```
37
```

```
df['BsmtFinSf1'].sum()
```

```
0
```

```
df['BsmtFinType2'].sum()
```

```
38
```

```
df['BsmtFinSf2'].sum()
```

```
0
```

```
df['BsmtUnfSF'].sum()
```

```
0
```

```
df['TotalBsmtSF'].sum()
```

```
0
```

```
df['Heating'].sum()
```

```
0
```

```
df['HeatingQC'].sum()
```

```
0
```

```
df['CentralAir'].sum()
```

```
0
```

```
df['Electrical'].sum()
```

```
1
```

```
df['1stFlrSF'].sum()
```

```
0
```

```
df['2ndFlrSF'].sum()
```

```
0
```

```
df['LowQualFinSF'].sum()
```

```
0
```

```
df['GrLivArea'].sum()
```

```
0
```

```
df['BsmtFullBath'].sum()
```

```
0
```

```
df['BsmtHalfBath'].sum()
```

```
0
```

```
df['FullBath'].sum()
```

```
0
```

```
df['HalfBath'].sum()
```

```
0
```

```
df['BedroomAbvGr'].sum()
```

```
0
```

```
df['KitchenAbvGr'].sum()
```

```
0
```

```
df['KitchenQual'].sum()
```

```
0
```

```
df['TotRmsAbvGrd'].sum()
```

```
0
```

```
df['Functional'].sum()
```

```
0
```

```
df['Fireplaces'].sum()
```

```
0
```

```
df['FireplaceQu'].sum()
```

```
690
```

```
df['GarageType'].sum()
```

```
81
```

```
df['GarageYrBlt'].sum()
```

```
81
```

```
df['GarageFinish'].sum()
```

```
81
```

```
df['GarageCars'].sum()
```

```
0
```

```
df['GarageArea'].sum()
```

```
0
```

```
df['GarageQual'].sum()
```

```
81
```

```
df['GarageCond'].sum()
```

```
81
```

```
df['PavedDrive'].sum()
```

```
0
```

```
df['WoodDeckSF'].sum()
```

```
0
```

```
df['OpenPoolArea'].sum()
```

```
0
```

```
df['PoolQC'].sum()
```

```
1453
```

```
df['Fence'].sum()
```

```
1179
```

```
df['MiscFeature'].sum()
```

```
1406
```

```
df['MiscVal'].sum()
```

```
0
```

```
df['MoSold'].sum()
```

```
0
```

```
df['YrSold'].sum()
```

```
0
```

```
df['SaleType'].sum()
```

```
0
```

```
df['SaleCondition'].sum()
```

```
0
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```