```python
In [20]: import numpy as np
         import pandas as pd
         import matplotlib.pyplot
         import seaborn as sns
```

```python
In [21]: df=pd.read_csv("Downloads\\train.csv")
```

```python
In [22] df.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      1460 non-null    int64
 20  YearRemodAdd   1460 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    int64
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 [23]: `df.isnull().sum()`

Out[23]:
```
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
```

In [24]:
```
pd.set_option('display.max_columns',None)
pd.set_option('display.max_rows',None)
```

In [25]: `df.isnull().sum()`

```
Out[25]:  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
          Exterior2nd        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
```
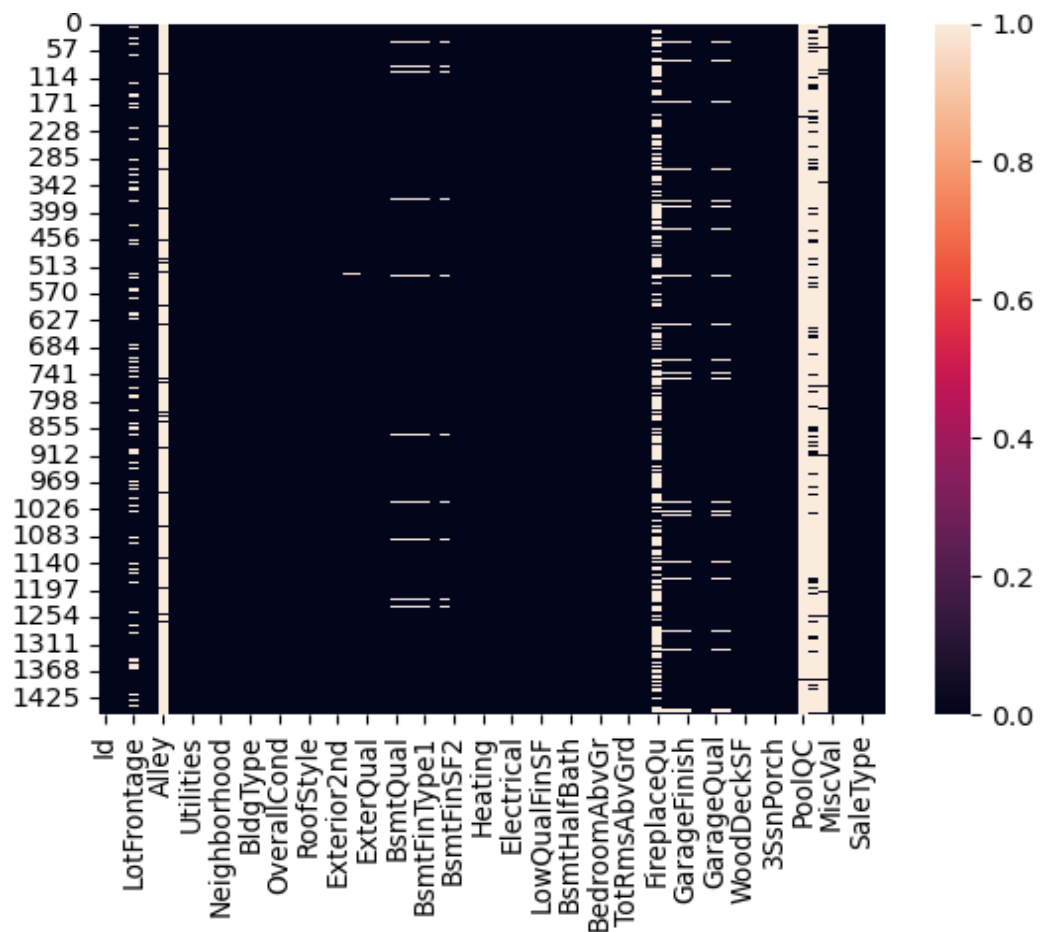
```
GarageCars        0
GarageArea        0
GarageQual       81
GarageCond       81
PavedDrive        0
WoodDeckSF        0
OpenPorchSF       0
EnclosedPorch     0
3SsnPorch         0
ScreenPorch       0
PoolArea          0
PoolQC         1453
Fence          1179
MiscFeature    1406
MiscVal           0
MoSold            0
YrSold            0
SaleType          0
SaleCondition     0
SalePrice         0
dtype: int64
```

In [26]: `sns.heatmap(df.isnull())`

Out[26]: `<Axes: >`

In [27]:   `df_per=df.isnull().sum()/df.shape[0]*100`

```
Out[28]:  Id              0.000000
          MSSubClass      0.000000
          MSZoning        0.000000
          LotFrontage    17.739726
          LotArea         0.000000
          Street          0.000000
          Alley          93.767123
          LotShape        0.000000
          LandContour     0.000000
          Utilities       0.000000
          LotConfig       0.000000
          LandSlope       0.000000
          Neighborhood    0.000000
          Condition1      0.000000
          Condition2      0.000000
          BldgType        0.000000
          HouseStyle      0.000000
          OverallQual     0.000000
          OverallCond     0.000000
          YearBuilt       0.000000
          YearRemodAdd    0.000000
          RoofStyle       0.000000
          RoofMatl        0.000000
          Exterior1st     0.000000
          Exterior2nd     0.000000
          MasVnrType      0.547945
          MasVnrArea      0.547945
          ExterQual       0.000000
          ExterCond       0.000000
          Foundation      0.000000
          BsmtQual        2.534247
          BsmtCond        2.534247
          BsmtExposure    2.602740
          BsmtFinType1    2.534247
          BsmtFinSF1      0.000000
          BsmtFinType2    2.602740
          BsmtFinSF2      0.000000
          BsmtUnfSF       0.000000
          TotalBsmtSF     0.000000
          Heating         0.000000
          HeatingQC       0.000000
          CentralAir      0.000000
          Electrical      0.068493
          1stFlrSF        0.000000
          2ndFlrSF        0.000000
          LowQualFinSF    0.000000
          GrLivArea       0.000000
          BsmtFullBath    0.000000
          BsmtHalfBath    0.000000
          FullBath        0.000000
          HalfBath        0.000000
          BedroomAbvGr    0.000000
          KitchenAbvGr    0.000000
          KitchenQual     0.000000
          TotRmsAbvGrd    0.000000
          Functional      0.000000
          Fireplaces      0.000000
          FireplaceQu    47.260274
          GarageType      5.547945
          GarageYrBlt     5.547945
          GarageFinish    5.547945
```

```
GarageCars          0.000000
GarageArea          0.000000
GarageQual          5.547945
GarageCond          5.547945
PavedDrive          0.000000
WoodDeckSF          0.000000
OpenPorchSF         0.000000
EnclosedPorch       0.000000
3SsnPorch           0.000000
ScreenPorch         0.000000
PoolArea            0.000000
PoolQC             99.520548
Fence              80.753425
MiscFeature        96.301370
MiscVal             0.000000
MoSold              0.000000
YrSold              0.000000
SaleType            0.000000
SaleCondition       0.000000
SalePrice           0.000000
dtype: float64
```

In [29]: `df2=df_per[df_per>20].keys()`

In [30]: `df2`

Out[30]: `Index(['Alley', 'FireplaceQu', 'PoolQC', 'Fence', 'MiscFeature'], dtype='object')`

In [31]: `df3=df.drop(columns=df2)`
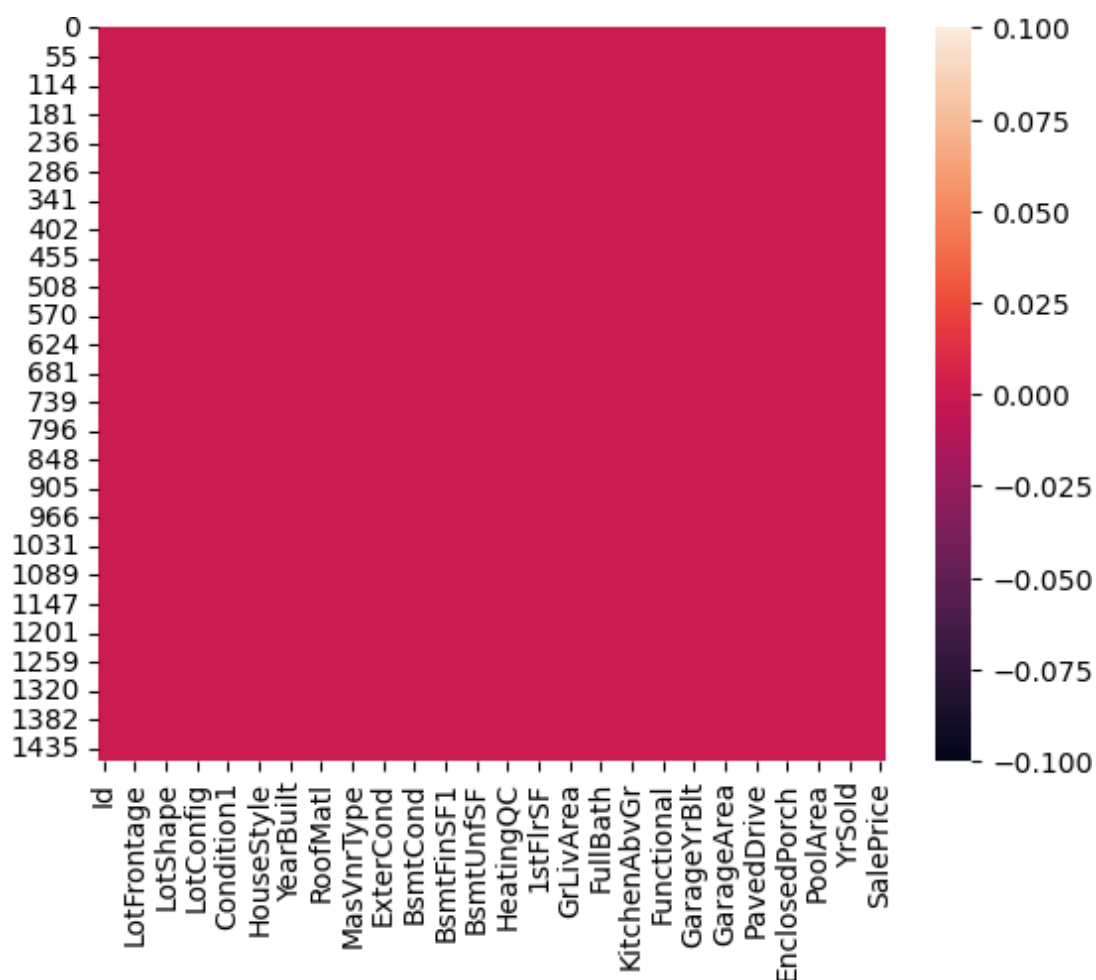
In [36]: `df4=df3.dropna()`

In [35]: `df3.shape`

Out[35]: `(1460, 76)`

In [37]: `df4.shape`

Out[37]: `(1094, 76)`

In [38]: `sns.heatmap(df4.isnull())`

Out[38]: `<Axes: >`

In [39]: `df4.isnull().sum()`

```
Out[39]:  Id               0
          MSSubClass       0
          MSZoning         0
          LotFrontage      0
          LotArea          0
          Street           0
          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
          Exterior2nd      0
          MasVnrType       0
          MasVnrArea       0
          ExterQual        0
          ExterCond        0
          Foundation       0
          BsmtQual         0
          BsmtCond         0
          BsmtExposure     0
          BsmtFinType1     0
          BsmtFinSF1       0
          BsmtFinType2     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
          GarageType       0
          GarageYrBlt      0
          GarageFinish     0
          GarageCars       0
          GarageArea       0
```

```
GarageQual        0
GarageCond        0
PavedDrive        0
WoodDeckSF        0
OpenPorchSF       0
EnclosedPorch     0
3SsnPorch         0
ScreenPorch       0
PoolArea          0
MiscVal           0
MoSold            0
YrSold            0
SaleType          0
SaleCondition     0
SalePrice         0
dtype: int64
```

In [40]: `#Data is Clean Now`

In [41]: