Imports

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Data

In [2]:
```
df=pd.read_csv("AMES_Final_DF.csv")
```

In [3]:
```
df.head()
```

Out[3]:

	Lot Frontage	Lot Area	Overall Qual	Overall Cond	Year Built	Year Remod/Add	Mas Vnr Area	BsmtFin SF 1	BsmtFin SF 2	Bsmt Unf SF	...	Sale Type_ConLw	Sale Type_New	Sale Type_Oth	Sale Type_VWD	Sale Type_WD	Sale Condition_Adjland	Sale Condition_Alloca	Sale Condition_Family	Sale Condition_Normal
0	141.0	31770	6	5	1960	1960	112.0	639.0	0.0	441.0	...	0	0	0	0	1	0	0	0	1
1	80.0	11622	5	6	1961	1961	0.0	468.0	144.0	270.0	...	0	0	0	0	1	0	0	0	1
2	81.0	14267	6	6	1958	1958	108.0	923.0	0.0	406.0	...	0	0	0	0	1	0	0	0	1
3	93.0	11160	7	5	1968	1968	0.0	1065.0	0.0	1045.0	...	0	0	0	0	1	0	0	0	1
4	74.0	13830	5	5	1997	1998	0.0	791.0	0.0	137.0	...	0	0	0	0	1	0	0	0	1

5 rows × 274 columns

In [4]:
```
df.shape
```

Out[4]:
```
(2925, 274)
```

In [6]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2925 entries, 0 to 2924
Columns: 274 entries, Lot Frontage to Sale Condition_Partial
dtypes: float64(11), int64(263)
memory usage: 6.1 MB
```

TASK: The label we are trying to predict is the SalePrice column. Separate out the data into X features and y labels

In [11]:
```
X=df.drop('SalePrice',axis=1)
```

In [12]:
```
y = df['SalePrice']
```

TASK: Use scikit-learn to split up X and y into a training set and test set. Since we will later be using a Grid Search strategy, set your test proportion to 10%. To get the same data split as the solutions notebook, you can specify random_state = 101

In [13]:
```
from sklearn.model_selection import train_test_split
```

In [15]:
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.1, random_state=101)
```

TASK: The dataset features has a variety of scales and units. For optimal regression performance, scale the X features. Take carefuly note of what to use for .fit() vs what to use for .transform()

In [16]:
```
from sklearn.preprocessing import StandardScaler
```

In [17]:
```
scaler=StandardScaler()
```

In [18]:
```
scaled_X_train = scaler.fit_transform(X_train)
scaled_X_test = scaler.transform(X_test)
```

TASK: We will use an Elastic Net model. Create an instance of default ElasticNet model with scikit-learn

In [19]:
```
from sklearn.linear_model import ElasticNet
```

In [20]:
```
base_elastic_model = ElasticNet()
```

TASK: The Elastic Net model has two main parameters, alpha and the L1 ratio. Create a dictionary parameter grid of values for the ElasticNet. Feel free to play around with these values, keep in mind, you may not match up exactly with the solution choices

In [21]:
```
param_grid = {'alpha':[0.1,1,5,10,50,100],
              'l1_ratio':[.1, .5, .7, .9, .95, .99, 1]}
```

In [22]:
```
from sklearn.model_selection import GridSearchCV
```

In [23]:
```
# verbose number, a personal preference
grid_model = GridSearchCV(estimator=base_elastic_model,
                          param_grid=param_grid,
                          scoring='neg_mean_squared_error',
                          cv=5,
                          verbose=1)
```

In [24]:
```
grid_model.fit(scaled_X_train,y_train)
```

```
Fitting 5 folds for each of 42 candidates, totalling 210 fits
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.394e+11, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.654e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.324e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.986e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.553e+11, tolerance: 1.346e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.235e+11, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.369e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.575e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.887e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 2.928e+11, tolerance: 1.346e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.542e+11, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.635e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.085e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.071e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.591e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.066e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.028e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.687e+11, tolerance: 1.346e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.025e+11, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.127e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.475e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.526e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 2.559e+11, tolerance: 1.346e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.707e+11, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.869e+11, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.965e+11, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 2.128e+11, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.849e+09, tolerance: 1.346e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 1.313e+10, tolerance: 1.355e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 4.414e+09, tolerance: 1.308e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.628e+10, tolerance: 1.415e+09
  model = cd_fast.enet_coordinate_descent(
C:\Users\gauri\anaconda3\Lib\site-packages\sklearn\linear_model\_coordinate_descent.py:628: ConvergenceWarning: Objective did not converge. You might want to increase the number of iterations, check the scale of the features or consider increasing regularisation. Duality gap: 3.970e+09, tolerance: 1.438e+09
  model = cd_fast.enet_coordinate_descent(
```

Out[24]:
```
GridSearchCV
estimator: ElasticNet
ElasticNet
```

TASK: Display the best combination of parameters for your model

In [25]:
```
grid_model.best_params_
```

Out[25]:
```
{'alpha': 100, 'l1_ratio': 1}
```

TASK: Evaluate your model's performance on the unseen 10% scaled test set. In the solutions notebook we achieved an MAE of $14149 and a RMSE of 20232

In [26]:
```
y_pred = grid_model.predict(scaled_X_test)
```

In [27]:
```
from sklearn.metrics import mean_absolute_error,mean_squared_error
```

In [28]:
```
mean_absolute_error(y_test,y_pred)
```

Out[28]:
```
14195.354900562173
```

In [29]:
```
np.sqrt(mean_squared_error(y_test,y_pred))
```

Out[29]:
```
20558.508566893164
```

In [30]:
```
np.mean(df['SalePrice'])
```

Out[30]:
```
180815.53743589742
```

In [ ]:
```

```