This Project is for House prices in King County in the USA. Task is to predict the Housing prices using Machine learning alogrithm(linear regression) and Deep learning Keras Tensorflow with Artificial Neutron Network(ANN). source: https://www.kaggle.com/harlfoxem/housesalesprediction. Refer to the meanings of some abbreviations of the feature columns if need be from the above website.

```
In [1]:
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

In [2]:

```
df=pd.read_csv(r'C:\Users\chumj\Downloads\Housing Prices.csv')
```

EXPLORATORY DATA ANALYSIS

In [3]:

```
df.head(3)
```

Out[3]:

		id	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	 grade	sqft_above	sqft_
	0	7129300520	10/13/2014	221900.0	3	1.00	1180	5650	1.0	0	0	 7	1180	
1	1	6414100192	12/9/2014	538000.0	3	2.25	2570	7242	2.0	0	0	 7	2170	
	2	5631500400	2/25/2015	180000.0	2	1.00	770	10000	1.0	0	0	 6	770	

3 rows × 21 columns

In [4]:

```
df.info()
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21597 entries, 0 to 21596
Data columns (total 21 columns):

#	Column	Non-Nu	ll Count	Dtype
0	id	21597 1	non-null	int64
1	date	21597 1	non-null	object
2	price	21597 1	non-null	float64
3	bedrooms	21597 1	non-null	int64
4	bathrooms	21597 1	non-null	float64
5	sqft_living	21597 1	non-null	int64
6	sqft_lot	21597 1	non-null	int64
7	floors	21597 1	non-null	float64
8	waterfront	21597 i	non-null	int64
9	view	21597 1	non-null	int64
10	condition	21597 1	non-null	int64
11	grade	21597 1	non-null	int64
12	sqft_above	21597 i	non-null	int64
13	sqft_basement	21597 1	non-null	int64
14	yr_built	21597 1	non-null	int64
15	yr_renovated	21597 i	non-null	int64
16	zipcode	21597 i	non-null	int64
17	lat	21597 1	non-null	float64
18	long	21597 n	non-null	float64
19	sqft_living15	21597 i	non-null	int64
20	sqft lot15	21597 1	non-null	int64

dtypes: float64(5), int64(15), object(1) memory usage: 3.5+ $\ensuremath{\mathsf{MB}}$

In [5]:

df.isnull().sum()

Out[5]:

date 0 0 price bedrooms 0 bathrooms sqft_living sqft lot 0 floors 0 waterfront view 0 condition 0 grade 0 sqft_above sqft_basement 0 yr_built 0 0 yr_renovated zipcode lat 0 0 long 0 sqft_living15 sqft_lot15 0 dtype: int64

In [6]:

df.describe().transpose()

Out[6]:

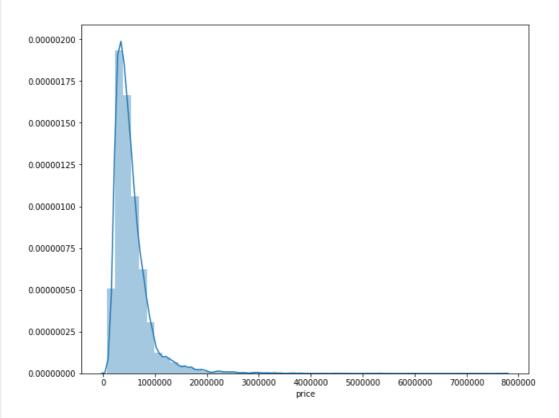
	count	mean	std	min	25%	50%	75%	max
id	21597.0	4.580474e+09	2.876736e+09	1.000102e+06	2.123049e+09	3.904930e+09	7.308900e+09	9.900000e+09
price	21597.0	5.402966e+05	3.673681e+05	7.800000e+04	3.220000e+05	4.500000e+05	6.450000e+05	7.700000e+06
bedrooms	21597.0	3.373200e+00	9.262989e-01	1.000000e+00	3.000000e+00	3.000000e+00	4.000000e+00	3.300000e+01
bathrooms	21597.0	2.115826e+00	7.689843e-01	5.000000e-01	1.750000e+00	2.250000e+00	2.500000e+00	8.000000e+00
sqft_living	21597.0	2.080322e+03	9.181061e+02	3.700000e+02	1.430000e+03	1.910000e+03	2.550000e+03	1.354000e+04
sqft_lot	21597.0	1.509941e+04	4.141264e+04	5.200000e+02	5.040000e+03	7.618000e+03	1.068500e+04	1.651359e+06
floors	21597.0	1.494096e+00	5.396828e-01	1.000000e+00	1.000000e+00	1.500000e+00	2.000000e+00	3.500000e+00
waterfront	21597.0	7.547345e-03	8.654900e-02	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	1.000000e+00
view	21597.0	2.342918e-01	7.663898e-01	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	4.000000e+00
condition	21597.0	3.409825e+00	6.505456e-01	1.000000e+00	3.000000e+00	3.000000e+00	4.000000e+00	5.000000e+00
grade	21597.0	7.657915e+00	1.173200e+00	3.000000e+00	7.000000e+00	7.000000e+00	8.000000e+00	1.300000e+01
sqft_above	21597.0	1.788597e+03	8.277598e+02	3.700000e+02	1.190000e+03	1.560000e+03	2.210000e+03	9.410000e+03
sqft_basement	21597.0	2.917250e+02	4.426678e+02	0.000000e+00	0.000000e+00	0.000000e+00	5.600000e+02	4.820000e+03
yr_built	21597.0	1.971000e+03	2.937523e+01	1.900000e+03	1.951000e+03	1.975000e+03	1.997000e+03	2.015000e+03
yr_renovated	21597.0	8.446479e+01	4.018214e+02	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	2.015000e+03
zipcode	21597.0	9.807795e+04	5.351307e+01	9.800100e+04	9.803300e+04	9.806500e+04	9.811800e+04	9.819900e+04
lat	21597.0	4.756009e+01	1.385518e-01	4.715590e+01	4.747110e+01	4.757180e+01	4.767800e+01	4.777760e+01
long	21597.0	-1.222140e+02	1.407235e-01	-1.225190e+02	-1.223280e+02	-1.222310e+02	-1.221250e+02	-1.213150e+02
sqft_living15	21597.0	1.986620e+03	6.852305e+02	3.990000e+02	1.490000e+03	1.840000e+03	2.360000e+03	6.210000e+03
sqft_lot15	21597.0	1.275828e+04	2.727444e+04	6.510000e+02	5.100000e+03	7.620000e+03	1.008300e+04	8.712000e+05

In [7]:

```
plt.figure(figsize=(10,8))
sns.distplot(df['price'])
```

Out[7]:

<matplotlib.axes._subplots.AxesSubplot at 0x1cdf1d8ff08>



Correlation with the other features with price.

In [8]:

```
df.corr()['price'].sort_values(ascending=False)
```

Out[8]:

```
price 1.000000
sqft_living 0.701917
grade 0.667951
sqft_above 0.605368

      sqft_above
      0.605368

      sqft_living15
      0.585241

      bathrooms
      0.525906

bathrooms
                            0.397370
view
sqft_basement 0.323799
                            0.308787
bedrooms
                              0.306692
lat
waterfront 0.266398 floors 0.256804
yr_renovated 0.126424

      rand
      0.089876

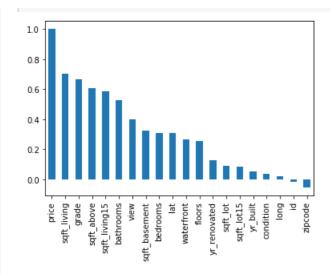
      sqft_lot15
      0.082845

      yr_built
      0.053953

                            0.036056
condition
long
                            0.022036
                        -0.016772
-0.053402
id
zipcode
Name: price, dtype: float64
```

In [9]:

```
df.corr()['price'].sort_values(ascending=False).plot(kind='bar');
```



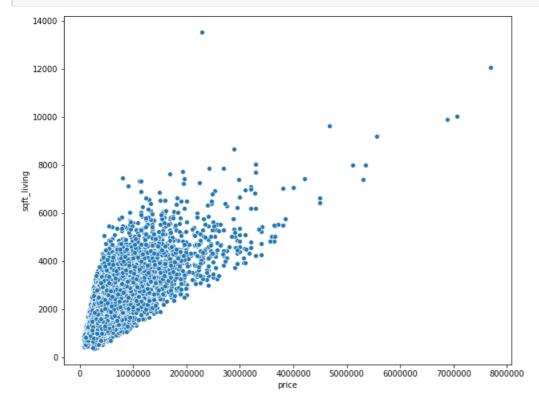
In [10]:

```
#plt.figure(figsize=(15,12))
#sns.heatmap(df.corr(),annot=True);
```

Exploring some important features that have a strong positive correlation with price.

In [11]:

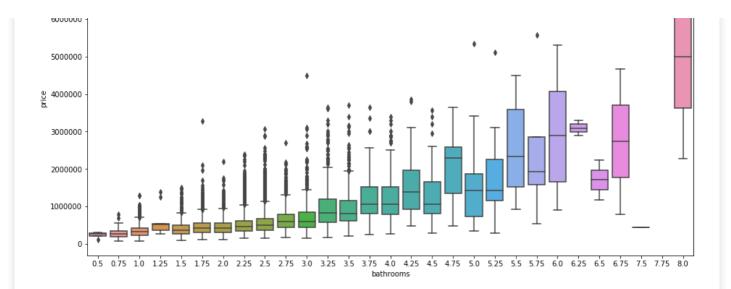
```
plt.figure(figsize=(10,8))
sns.scatterplot(x='price',y='sqft_living',data=df);
```



In [12]:

```
plt.figure(figsize=(15,8))
sns.boxplot(x='bathrooms',y='price',data=df);

8000000
7000000
```

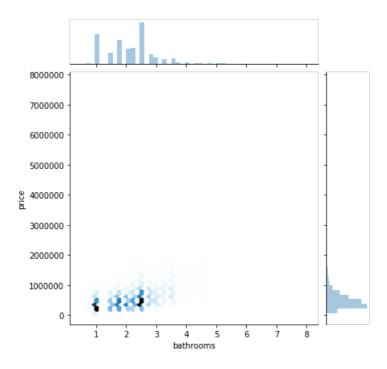


In [13]:

```
sns.jointplot(x='bathrooms',y='price',data=df,kind='hex')
```

Out[13]:

<seaborn.axisgrid.JointGrid at 0x1cdf2d1dd08>

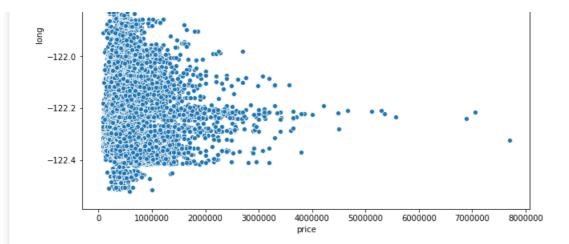


Geography, taking a look at longitude, latitude and waterfront.

In [14]:

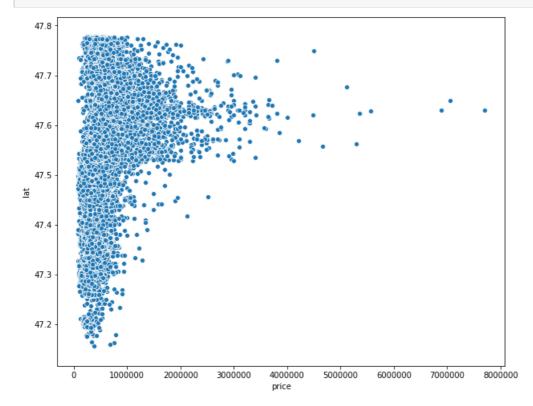
```
plt.figure(figsize=(10,8))
sns.scatterplot(x='price', y='long', data=df);
```





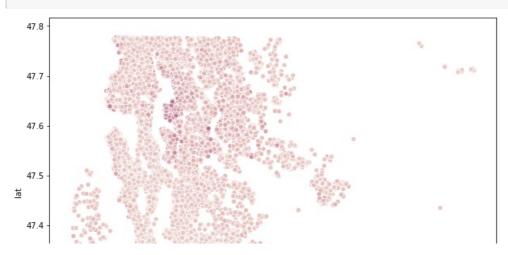
In [15]:

```
plt.figure(figsize=(10,8))
sns.scatterplot(x='price',y='lat',data=df);
```



In [16]:

```
plt.figure(figsize=(10,8))
sns.scatterplot(x='long',y='lat',data=df,hue='price');
```



```
47.2 - price
0
3000000
6000000
9000000
-122.4 -122.2 -122.0 -121.8 -121.6 -121.4
```

In [17]:

```
len(df)*1/100
```

Out[17]:

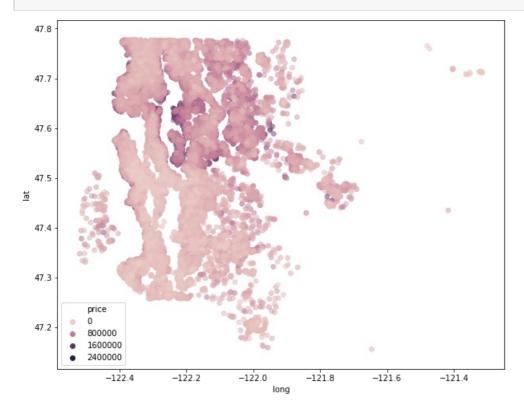
215.97

In [18]:

```
above=df.sort_values('price',ascending=False).iloc[216:]
```

In [19]:

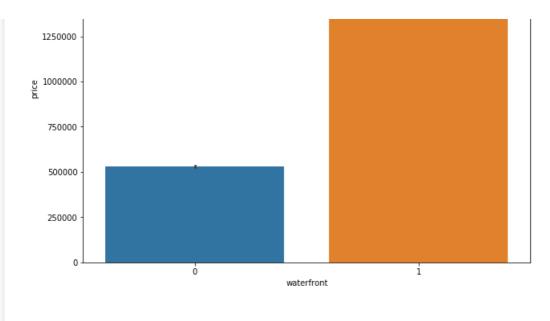
```
plt.figure(figsize=(10,8))
sns.scatterplot(x='long',y='lat',data=above,hue='price',alpha=0.5,edgecolor=None);
```



In [20]:

```
plt.figure(figsize=(10,8))
sns.barplot(x='waterfront',y='price',data=df);
```





In [21]:

```
df['waterfront'].value_counts()
Out[21]:
```

0 21434 1 163

Name: waterfront, dtype: int64

Analysing if Date play a role

In [22]:

```
df['date']=pd.to_datetime(df['date'])
```

In [23]:

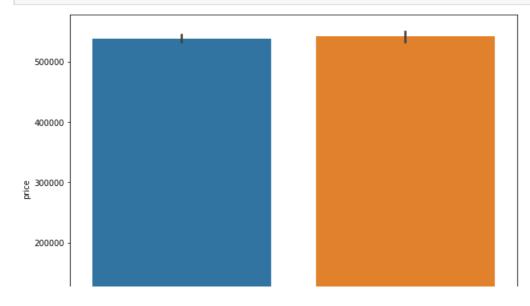
```
df['month']=df['date'].apply(lambda date:date.month)
```

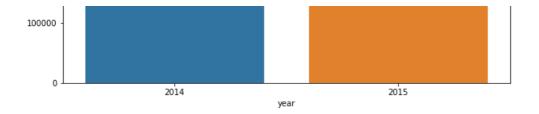
In [24]:

```
df['year']=df['date'].apply(lambda date:date.year)
```

In [25]:

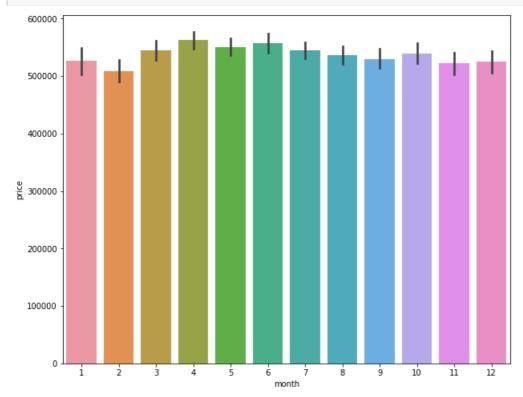
```
plt.figure(figsize=(10,8))
sns.barplot(x='year',y='price',data=df);
```





In [26]:

```
plt.figure(figsize=(10,8))
sns.barplot(x='month',y='price',data=df);
```

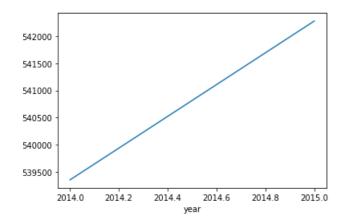


In [27]:

```
df.groupby('year').mean()['price'].plot()
```

Out[27]:

<matplotlib.axes._subplots.AxesSubplot at 0x1cdf2a7d588>

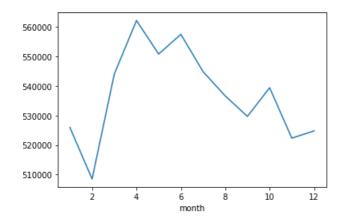


In [28]:

```
df.groupby('month').mean()['price'].plot()
```

040[20].

<matplotlib.axes._subplots.AxesSubplot at 0x1cdf2a34f08>



In [29]:

```
df.columns
```

Out[29]:

Data preprocessing and cleaning for our Machine learning and ANN Regression model to pridict price. Some of the features need to be clean in such a way that our algorithm can accept it and some will simply be droped, especially those with very little correlation with the price.

```
In [30]:
```

```
df=df.drop('id',axis=1)
df=df.drop('zipcode',axis=1)
df=df.drop('date',axis=1)
```

In [31]:

```
df.head(3)
```

Out[31]:

	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	grade	sqft_above	sqft_basement	yr_buil
0	221900.0	3	1.00	1180	5650	1.0	0	0	3	7	1180	0	195
1	538000.0	3	2.25	2570	7242	2.0	0	0	3	7	2170	400	195 ⁻
2	180000.0	2	1.00	770	10000	1.0	0	0	3	6	770	0	1930
4									1				Þ

Training a Linear Regression Model(sklearn model)

Here we have to split our data into Training and Test split

```
In [32]:
```

```
X=df.drop('price',axis=1)
y=df['price']
```

```
In [33]:
```

```
from sklearn.model_selection import train_test_split
In [34]:
X_train, X_test, y_train, y_test = train_test_split( X, y, test_size=0.30, random_state=101)
In [35]:
from sklearn.linear_model import LinearRegression
In [36]:
lr=LinearRegression()
In [37]:
lr.fit(X_train,y_train)
Out[37]:
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
In [38]:
# print intercept
print(lr.intercept )
-116583510.48436856
In [39]:
coeff=pd.DataFrame(lr.coef ,X.columns,columns=['cofficient'])
In [40]:
coeff
Out[40]:
                  cofficient
              -35469.017138
    bedrooms
    bathrooms
               42365.134436
    sqft_living
                 106.699375
                   0.176845
      sqft_lot
       floors
                2485.905619
    waterfront
              627723.871190
               45669.377112
    condition
               30480.903119
       grade
               99672.034520
    sqft_above
                  70.146415
                  36.552960
sqft_basement
               -2467.245878
      yr_built
                  22.224492
  yr_renovated
              560933.319973
             -142099.588905
        long
  sqft_living15
                  31.524857
    sqft_lot15
                  -0.399282
```

1369.439270

month

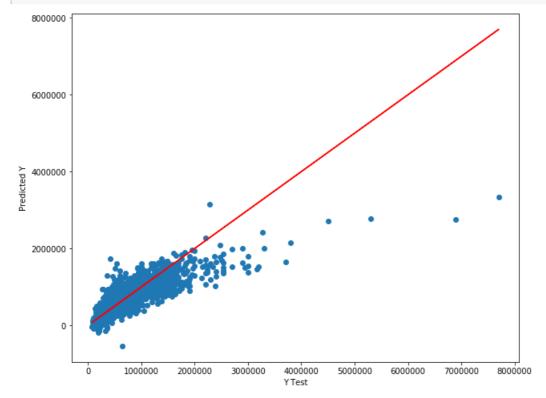
```
cofficient
year 38055 891407
```

In [41]:

```
predictions=lr.predict(X_test)
```

In [56]:

```
plt.figure(figsize=(10,8))
plt.scatter(y_test,predictions)
plt.plot(y_test,y_test,'r')
plt.xlabel('Y Test')
plt.ylabel('Predicted Y');
```

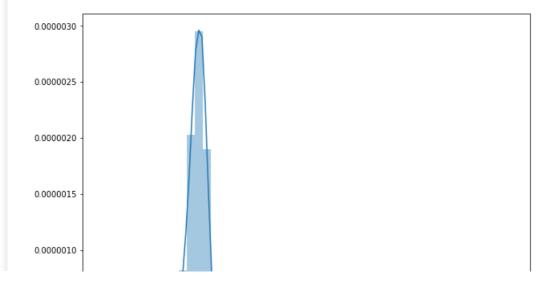


In [43]:

```
plt.figure(figsize=(10,8))
sns.distplot((y_test-predictions),bins=50)
```

Out[43]:

<matplotlib.axes._subplots.AxesSubplot at 0x1cdf309c1c8>



In [44]:

```
from sklearn import metrics
```

In [45]:

```
print('MAE:',metrics.mean_absolute_error(y_test,predictions))
print('MSE:',metrics.mean_squared_error(y_test,predictions))
print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,predictions)))
```

MAE: 123998.57212895666 MSE: 40433388708.44527 RMSE: 201080.55278530857

In [46]:

```
metrics.r2_score(y_test,predictions)
```

Out[46]:

0.6951102243222294

In [88]:

```
df.describe()['price']
```

Out[88]:

```
count 2.159700e+04
mean 5.402966e+05
std 3.673681e+05
min 7.800000e+04
25% 3.220000e+05
50% 4.500000e+05
75% 6.450000e+05
max 7.700000e+06
```

Name: price, dtype: float64

In [97]:

```
#predicting on a brand new house
New_house=df.drop('price',axis=1).iloc[0]
```

In [98]:

```
New_house=scaler.transform(New_house.values.reshape(-1,19))
```

In [92]:

```
New_house
```

Out[92]:

```
In [93]:
model.predict(New house)
Out[93]:
array([[281312.56]], dtype=float32)
In [94]:
df.iloc[0]
Out[94]:
                221900.0000
price
                3.0000
bedrooms
                     1.0000
bathrooms
sqft_living
                  1180.0000
                  5650.0000
sqft_lot
                     1.0000
floors
waterfront
                     0.0000
                     0.0000
view
                     3.0000
7.0000
condition
grade
sqft_above 1180.0000
sqft_basement 0.0000
                 1955.0000
yr_renovated lat
                   0.0000
47.5112
                   -122.2570
long
                  1340.0000
sqft living15
sqft lot15
                  5650.0000
month
                    10.0000
                  2014.0000
year
Name: 0, dtype: float64
In [ ]:
Base on our prediction above, our model was able to predict well the house prices from 0.....1500000, above that the model was able
to come out with a good prediction, thus that why our r2_score was able to explain just 69% of the variance
Using the ANN Approach
In [49]:
```

In [49]: #Scaling from sklearn.preprocessing import MinMaxScaler In [50]: scaler=MinMaxScaler() In [51]: X_train=scaler.fit_transform(X_train) In [52]: X_test=scaler.transform(X_test)

In [53]:

X_train.shape

```
Out[53]:
(15117, 19)
In [54]:
X test.shape
Out[54]:
(6480, 19)
In [57]:
# creating model
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Activation
from tensorflow.keras.optimizers import Adam
In [58]:
model=Sequential()
In [59]:
model.add(Dense(19,activation='relu'))
model.add(Dense(19,activation='relu'))
model.add(Dense(9,activation='relu'))
model.add(Dense(1))
model.compile(optimizer='adam',loss='mse')
{\tt WARNING:tensorflow:From C:\Users\chumj\Anaconda3\Ben\lib\site-}
packages\tensorflow\python\ops\init ops.py:1251: calling VarianceScaling. init (from
tensorflow.python.ops.init_ops) with dtype is deprecated and will be removed in a future version.
Instructions for updating:
Call initializer instance with the dtype argument instead of passing it to the constructor
In [60]:
#training model
model.fit(x=X train,y=y train.values,validation data=(X test,y test.values),batch size=100,epochs=3
00)
Train on 15117 samples, validate on 6480 samples
Epoch 1/300
15117/15117 [============] - 1s 64us/sample - loss: 430231519700.2072 -
val loss: 418891592135.1111
Epoch 2/300
15117/15117 [===========] - 0s 30us/sample - loss: 429580103309.6069 -
val loss: 417182058116.7407
Epoch 3/300
15117/15117 [============] - 0s 28us/sample - loss: 425236880394.0930 -
val loss: 409207712730.0741
Epoch 4/300
15117/15117 [===========] - 0s 30us/sample - loss: 411206392565.7207 -
val loss: 388165332385.1852
Epoch 5/300
15117/15117 [============] - Os 28us/sample - loss: 380802693748.2050 -
val loss: 348302157495.3087
Epoch 6/300
15117/15117 [============ ] - Os 27us/sample - loss: 330377970061.7593 -
val loss: 289254598377.8765
Epoch 7/300
15117/15117 [============] - 0s 29us/sample - loss: 263938714780.1368 -
val loss: 219747064010.2716
Epoch 8/300
15117/15117 [============= ] - Os 28us/sample - loss: 194991864049.8257 -
val loss: 157151777905.7778
Epoch 9/300
15117/15117 [============ ] - Os 30us/sample - loss: 141169925443.7214 -
val loss: 117028887179.0617
```

```
Epoch 10/300
15117/15117 [============= ] - Os 27us/sample - loss: 111946292784.6700 -
val loss: 100313738865.7778
Epoch 11/300
15117/15117 [============= ] - Os 29us/sample - loss: 101397364480.8975 -
val loss: 95767917770.2716
Epoch 12/300
94449023367.9012
Epoch 13/300
93520386692.7407
Epoch 14/300
92560502670.2222
Epoch 15/300
91547494475.8519
Epoch 16/300
90498554804.1481
Epoch 17/300
89391993426.1728
Epoch 18/300
88285573638.3210
Epoch 19/300
87115786922.6667
Epoch 20/300
85937746343.5062
Epoch 21/300
84684562223.4074
Epoch 22/300
83392530944.0000
Epoch 23/300
82147696374.5185
Epoch 24/300
80803735077.9259
Epoch 25/300
79476714028.2469
Epoch 26/300
78053861774.2222
Epoch 27/300
76603606534.3210
Epoch 28/300
75121302515.3580
Epoch 29/300
73644955793.3827
Epoch 30/300
72100071683.1605
Epoch 31/300
70588930155.4568
Epoch 32/300
69053899434.6667
Epoch 33/300
67593757740.2469
Epoch 34/300
66017270284.6420
Epoch 35/300
```

```
00 0000, 00mptc
               _____
64525768204.6420
Epoch 36/300
63027034921.0864
Epoch 37/300
61584671377.3827
Epoch 38/300
60282526189.0370
Epoch 39/300
58957239220.1481
Epoch 40/300
57681373304.0988
Epoch 41/300
56532390425.2840
Epoch 42/300
55451839386.8642
Epoch 43/300
54490020826.0741
Epoch 44/300
53570291231.6049
Epoch 45/300
52809973867.4568
Epoch 46/300
52076809566.8148
Epoch 47/300
51410283463.1111
Epoch 48/300
50840909255.1111
Epoch 49/300
50306891380.9383
Epoch 50/300
49813993206.5185
Epoch 51/300
49384946014.8148
Epoch 52/300
49003021321.4815
Epoch 53/300
48619866105.6790
Epoch 54/300
48287360790.1235
Epoch 55/300
47990915621.9259
Epoch 56/300
47685834202.0741
Epoch 57/300
47443213963.0617
Epoch 58/300
47173054786.3704
Epoch 59/300
46955528286.8148
Epoch 60/300
46718221005.4321
```

Enoch 61/300

```
Thorn OT/200
46527536560.9877
Epoch 62/300
46323840461.4321
Epoch 63/300
46118656410.8642
Epoch 64/300
45941589927.5062
Epoch 65/300
45772906123.0617
Epoch 66/300
45604916713.8765
Epoch 67/300
45446707993.2840
Epoch 68/300
45315923689.8765
Epoch 69/300
45153931267.1605
Epoch 70/300
45030558944.3951
Epoch 71/300
44890757382.3210
Epoch 72/300
44764164393.0864
Epoch 73/300
44638274657.9753
Epoch 74/300
44523984052.1481
Epoch 75/300
44403915355.6543
Epoch 76/300
44293238733.4321
Epoch 77/300
44185232055.3086
Epoch 78/300
44076169494.1235
Epoch 79/300
43972886689,1852
Epoch 80/300
43869458444.6420
Epoch 81/300
43792754109.6296
Epoch 82/300
43683812709.1358
Epoch 83/300
43600585377.1852
Epoch 84/300
43496961782.5185
Epoch 85/300
43422441677.4321
Epoch 86/300
12321585161 0088
```

```
400040004000
Epoch 87/300
43245786974.8148
Epoch 88/300
43160219208.6914
Epoch 89/300
43088326150.3210
Epoch 90/300
43007956941.4321
Epoch 91/300
42924131448.0988
Epoch 92/300
42840933492.9383
Epoch 93/300
42766544004.7407
Epoch 94/300
42703047913.8765
Epoch 95/300
42620688523.0617
Epoch 96/300
42526895830.9136
Epoch 97/300
42457569488.5926
Epoch 98/300
42381365295.4074
Epoch 99/300
42299964766.8148
Epoch 100/300
42241754042.4691
Epoch 101/300
42157572361.4815
Epoch 102/300
42083408760.0988
Epoch 103/300
42009117297.7778
Epoch 104/300
41937405316.7407
Epoch 105/300
loss: 42752797873.0002 - val loss: 41909863117.4321
Epoch 106/300
41794725044.1481
Epoch 107/300
41720766574.6173
Epoch 108/300
41654914345.0864
Epoch 109/300
41563829216.3951
Epoch 110/300
41541432231.5062
Epoch 111/300
41430478693.1358
Epoch 112/300
```

15117/15117 [-

_____1

0. 2602/20mpla | 1002. 42251605000 7122 ***1 1002.

```
----1 /1761//1761
      41368194917.1358
Epoch 113/300
41265590796.6420
Epoch 114/300
41186218774.1235
Epoch 115/300
41119713343.2099
Epoch 116/300
41034322621.6296
Epoch 117/300
40966782644.1481
Epoch 118/300
40870931569.7778
Epoch 119/300
40797690576.5926
Epoch 120/300
40698323493.9259
Epoch 121/300
40615791979.4568
Epoch 122/300
40526521764.3457
Epoch 123/300
40439712802.7654
Epoch 124/300
40366907079.1111
Epoch 125/300
40268700627.7531
Epoch 126/300
40175090425.6790
Epoch 127/300
40085877459.7531
Epoch 128/300
40000525223.5062
Epoch 129/300
39904691162.0741
Epoch 130/300
loss: 40715170600.1857 - val loss: 39848264539.6543
Epoch 131/300
39711562581.3333
Epoch 132/300
39615637513.4815
Epoch 133/300
39516500922.4691
Epoch 134/300
39454953389.8272
Epoch 135/300
39330997197.4321
Epoch 136/300
39214201353.4815
Epoch 137/300
39106817349.5309
```

T---1- 120/200

```
Epocn 138/300
39021664126.4198
Epoch 139/300
38895871333.1358
Epoch 140/300
38791883222.9136
Epoch 141/300
38682428634.0741
Epoch 142/300
38554294461,6296
Epoch 143/300
38439957156.3457
Epoch 144/300
38327241481.4815
Epoch 145/300
38210468140.2469
Epoch 146/300
38103631890.9630
Epoch 147/300
37987583396.3457
Epoch 148/300
37869549896.6914
Epoch 149/300
37778474875.2593
Epoch 150/300
37621721394.5679
Epoch 151/300
37510313813.3333
Epoch 152/300
37380051354.8642
Epoch 153/300
37257171007.2099
Epoch 154/300
37149367751.1111
Epoch 155/300
37045009954.7654
Epoch 156/300
36917827309.0370
Epoch 157/300
36800716199.5062
Epoch 158/300
36686428542.4198
Epoch 159/300
36595907407.0123
Epoch 160/300
36479259606.9136
Epoch 161/300
36378519510.9136
Epoch 162/300
36299435915.0617
Epoch 163/300
```

```
36196504990.024/
Epoch 164/300
36104895374.2222
Epoch 165/300
36024906063.0123
Epoch 166/300
35971806261.7284
Epoch 167/300
35856583281.7778
Epoch 168/300
35780485843.7531
Epoch 169/300
35712100924.0494
Epoch 170/300
35650183629.4321
Epoch 171/300
35586535733.7284
Epoch 172/300
35551344655.8025
Epoch 173/300
35490637903.0123
Epoch 174/300
35426091431.5062
Epoch 175/300
35373848809.8765
Epoch 176/300
35350674239.2099
Epoch 177/300
35301835159.7037
Epoch 178/300
35223297908.9383
Epoch 179/300
35182559674.4691
Epoch 180/300
35132624510.4198
Epoch 181/300
35099964529.7778
Epoch 182/300
35070518363.6543
Epoch 183/300
35014691299.5556
Epoch 184/300
34977627967.2099
Epoch 185/300
34946140434.9630
Epoch 186/300
34920142598.3210
Epoch 187/300
34866072620.2469
Epoch 188/300
34835953888.3951
Epoch 189/300
```

```
34812835748.3457
Epoch 190/300
34772606938.0741
Epoch 191/300
34728817746.1728
Epoch 192/300
loss: 35828551210.8445 - val loss: 34707721136.9877
34708056727.7037
Epoch 194/300
34628839240.6914
Epoch 195/300
34615223665.7778
Epoch 196/300
34584510106.8642
Epoch 197/300
34530522661.9259
Epoch 198/300
34517549852.4444
Epoch 199/300
34498389595.6543
Epoch 200/300
34437199081.8765
Epoch 201/300
34405933334.1235
Epoch 202/300
34379788885.3333
Epoch 203/300
34352429861,9259
Epoch 204/300
34389809572.3457
Epoch 205/300
loss: 35466777895.4068 - val loss: 34291404654.6173
Epoch 206/300
34259027064.0988
Epoch 207/300
34269151889.3827
Epoch 208/300
34202287761.3827
Epoch 209/300
34173599778.7654
Epoch 210/300
34148956769.9753
Epoch 211/300
34118387215.8025
Epoch 212/300
34090356840.2963
Epoch 213/300
34096685498.4691
Epoch 214/300
34034582771.3580
```

```
Epoch 215/300
34019898785.1852
Epoch 216/300
34004666800.9877
Epoch 217/300
33954605653.3333
Epoch 218/300
33925707962.4691
Epoch 219/300
33904270724.7407
Epoch 220/300
33872535978.6667
Epoch 221/300
33853898176.7901
Epoch 222/300
33825929374.0247
Epoch 223/300
33798204251.6543
Epoch 224/300
33792623587.5556
Epoch 225/300
33739450665.0864
Epoch 226/300
33732678567.5062
Epoch 227/300
33696002123.8518
Epoch 228/300
33655653714.1728
Epoch 229/300
33645829559.3086
Epoch 230/300
33602714759.9012
Epoch 231/300
33634832055.3086
Epoch 232/300
33572905895.5062
Epoch 233/300
33518616168.2963
Epoch 234/300
33492720987.6543
Epoch 235/300
33466988208.9877
Epoch 236/300
33440243348.5432
Epoch 237/300
33421665213.6296
Epoch 238/300
33425980504.4938
Epoch 239/300
33359630829.0370
Epoch 240/300
```

```
33331953392.1975
Epoch 241/300
33321171566.6173
Epoch 242/300
33290671846.7160
Epoch 243/300
33247659415.7037
Epoch 244/300
33222823196.4444
Epoch 245/300
33194982409.4815
Epoch 246/300
loss: 34413989213.7329 - val loss: 33173009461.7284
Epoch 247/300
33149073790.4198
Epoch 248/300
33120634601.8765
Epoch 249/300
33096211190.5185
Epoch 250/300
33069839641.2840
Epoch 251/300
33042481243.6543
Epoch 252/300
33042531862.1235
Epoch 253/300
33017545664.7901
Epoch 254/300
32969401666.3704
Epoch 255/300
32944271710.8148
Epoch 256/300
32930951237.5309
Epoch 257/300
32909324670.4198
Epoch 258/300
32872769169.3827
Epoch 259/300
32856405415.5062
Epoch 260/300
32856168953.6790
Epoch 261/300
32836069391.8025
Epoch 262/300
32774781440.0000
Epoch 263/300
32777917016.4938
Epoch 264/300
32754212892.4444
Epoch 265/300
32733260686.2222
```

Epoch 266/300

```
32676748904,2963
Epoch 267/300
32659493303.3086
Epoch 268/300
32726748991.2099
Epoch 269/300
32615402375.9012
Epoch 270/300
32586628579.5556
Epoch 271/300
32563314479.4074
Epoch 272/300
32560929918.4198
Epoch 273/300
32519902574.6173
Epoch 274/300
32528424827.2593
Epoch 275/300
32515596976.9877
Epoch 276/300
32451439625,4815
Epoch 277/300
32436150034.9630
Epoch 278/300
32422675604.5432
Epoch 279/300
32394395139.1605
Epoch 280/300
32366962482.5679
Epoch 281/300
32345767619.9506
Epoch 282/300
32330125669.1358
Epoch 283/300
32316313776.9877
Epoch 284/300
32316545460.1481
Epoch 285/300
32265308921.6790
Epoch 286/300
32255782716.0494
Epoch 287/300
32283378710.1235
Epoch 288/300
32210717986.7654
Epoch 289/300
32184571297.1852
Epoch 290/300
32194438706.5679
Epoch 291/300
32144282734.6173
```

```
Epoch 292/300
15117/15117 [====
    32124350865.3827
Epoch 293/300
32115240609.1852
Epoch 294/300
32097239589.9259
Epoch 295/300
32072230700.2469
Epoch 296/300
32105864798.8148
Epoch 297/300
32049007938.3704
Epoch 298/300
32013674233.6790
Epoch 299/300
31999053880.8889
Epoch 300/300
31988941261.4321
```

Out[60]:

<tensorflow.python.keras.callbacks.History at 0x1cdfe88b708>

In [62]:

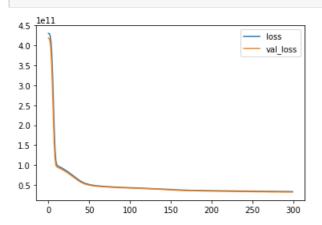
```
#model.history.history
```

In [63]:

losses=pd.DataFrame (model.history.history)

In [65]:

losses.plot();



In [66]:

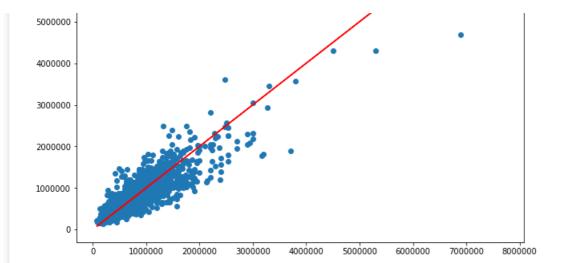
from tensorflow.keras.callbacks import EarlyStopping

In [70]:

```
#early_stop = EarlyStopping(monitor='val_loss', mode='min', verbose=1, patience=25)
```

In [71]:

```
#model.fit(x=X train,y=y train,epochs=600,validation data=(X test, y test), verbose=1,callbacks=[e
arly_stop] )
In [73]:
#EVALUATING THE MODEL
from sklearn.metrics import mean_absolute_error,mean_squared_error,explained_variance_score
In [74]:
predictions1=model.predict(X_test)
In [75]:
mean_absolute_error(y_test,predictions1)
Out[75]:
99796.04556086034
In [76]:
mean_squared_error(y_test,predictions1)
Out[76]:
25483779352.2967
In [77]:
np.sqrt(mean_squared_error(y_test,predictions1))
Out[77]:
159636.39732935812
In [78]:
explained_variance_score(y_test,predictions1)
Out[78]:
0.8078840279886367
In [79]:
df['price'].mean()
Out[79]:
540296.5735055795
In [81]:
plt.figure(figsize=(10,8))
plt.scatter(y_test,predictions1)
plt.plot(y test,y test,'r');
 8000000
 7000000
 6000000
```

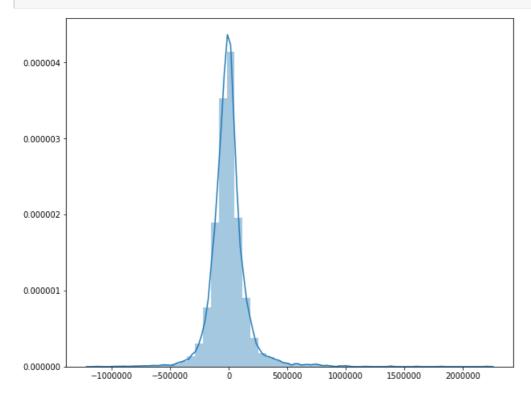


In [84]:

error=y_test.values.reshape(6480,1)-predictions1

In [87]:

```
plt.figure(figsize=(10,8))
sns.distplot(error);
```



In [99]:

New_brand2=df.drop('price',axis=1).iloc[1]

In [100]:

New brand2=scaler.transform(New brand2.values.reshape(-1,19))

In [101]:

New_brand2

Out[101]:

arrav([[0.2 . 0.28 . 0.22750776. 0.00407187. 0.4

```
0. , 0. , 0.5 , 0.4 , 0.23968043,
urruy ( [ [ U • Z
       0.08298755, 0.44347826, 0.98808933, 0.90895931, 0.16611296,
       0.22216486, 0.0081402 , 1. , 0. ]])
In [102]:
model.predict(New brand2)
Out[102]:
array([[635116.75]], dtype=float32)
In [103]:
df.iloc[1]
Out[103]:
              538000.000
price
              3.000
bedrooms
bathrooms
                  2.250
sqft_living
               2570.000
sqft_lot
                7242.000
floors
                   2.000
                   0.000
waterfront
view
                  0.000
                  3.000
condition
                   7.000
grade
7.000
2170.000
sqft_basement 400.000
yr_built
yr renovated
               1991.000
lat
                 47.721
long
                 -122.319
sqft_living15
                1690.000
sqft_lot15
                7639.000
month
                 12.000
                2014.000
Name: 1, dtype: float64
In [104]:
df['price']
Out[104]:
0
      221900.0
       538000.0
1
      180000.0
2
      604000.0
      510000.0
4
         . . .
21592 360000.0
21593 400000.0
21594 402101.0
21595 400000.0
21596
       325000.0
Name: price, Length: 21597, dtype: float64
In [ ]:
```