Mohd Zain 22MT0214 Linear Regression

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from google.colab import files
al=files.upload()
```

Choose files lin_reg1.csv
- **lin_reg1.csv**(text/csv) - 97 bytes, last modified: 18/08/2022 - 100% done
Saving lin_reg1.csv to lin_reg1.csv

```
import io
df=pd.read_csv(io.BytesIO(al['lin_reg1.csv']))
```

```
df.head(5)
```

	x1	y
0	2600	550000
1	3000	565000
2	3200	610000
3	3600	680000
4	4000	725000

```
from sklearn.model_selection import train_test_split
```

```
X= df.iloc[:, :-1]
y = df.iloc[:, -1]
```

```
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=2, random_state=0)
```

```
from sklearn.linear_model import LinearRegression
```

```
lrg=LinearRegression()
```

```
lrg.fit(X_train,y_train)

    LinearRegression()
```

```
y_pred=lrg.predict(X_test)
y_pred
```

```
array([786730.6122449 , 615740.81632653])
```

```
lrg.score(X_test,y_test)#accuracy
```

```
0.9692522987370867
```

```
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
```

```
mean_absolute_error(y_test,y_pred)
```

```
12235.714285714319
```

```
mean_squared_error(y_test,y_pred)
```

```
191896403.58184165
```
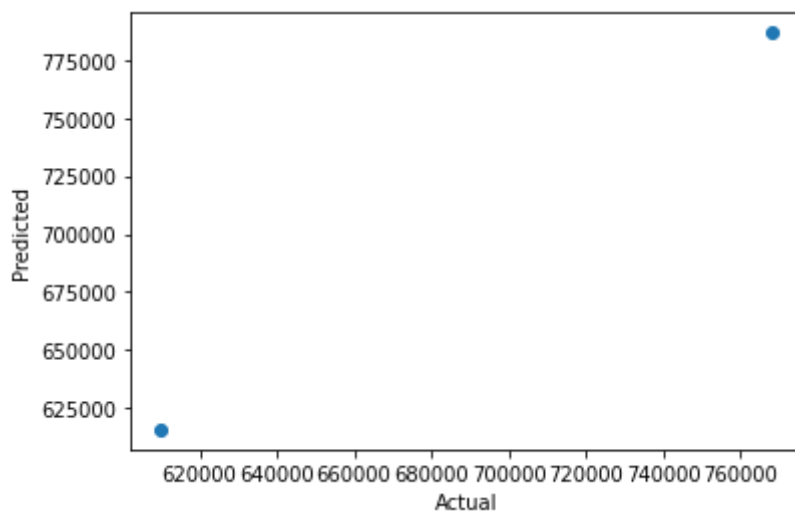
```
acc=r2_score(y_test,y_pred)
```

```
acc
```
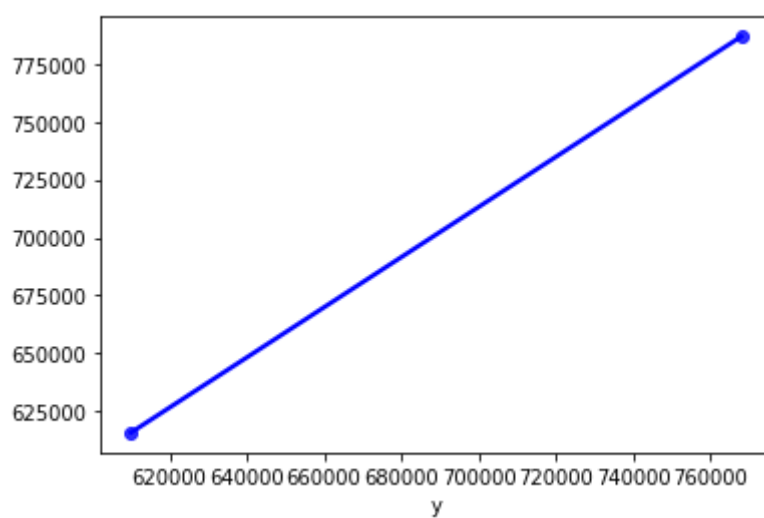
```
0.9692522987370867
```

```
print('your model accuracy is',round(acc*100,3),'%')
```

```
your model accuracy is 96.925 %
```

```
plt.scatter(y_test,y_pred);
plt.xlabel('Actual');
plt.ylabel('Predicted');
```

```python
sns.regplot(x=y_test,y=y_pred,ci=None,color ='blue');
```

✓ 0s   completed at 20:58   ● ✕