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

```python
df=pd.read_csv('/content/Salary_Data.csv')
```

```python
df.head()
```

```python
x=df.drop(['Salary'],axis=1).values
```

```python
x
```

```
array([[ 1.1],
       [ 1.3],
       [ 1.5],
       [ 2. ],
       [ 2.2],
       [ 2.9],
       [ 3. ],
       [ 3.2],
       [ 3.2],
       [ 3.7],
       [ 3.9],
       [ 4. ],
       [ 4. ],
       [ 4.1],
       [ 4.5],
       [ 4.9],
       [ 5.1],
       [ 5.3],
       [ 5.9],
       [ 6. ],
       [ 6.8],
       [ 7.1],
       [ 7.9],
       [ 8.2],
```

```
       [ 8.7],
       [ 9. ],
       [ 9.5],
       [ 9.6],
       [10.3],
       [10.5]])
```

```
y=df['Salary'].values
```

```
y
```

```
array([ 39343.,   46205.,   37731.,   43525.,   39891.,   56642.,   60150.,
        54445.,   64445.,   57189.,   63218.,   55794.,   56957.,   57081.,
        61111.,   67938.,   66029.,   83088.,   81363.,   93940.,   91738.,
        98273.,  101302.,  113812.,  109431.,  105582.,  116969.,  112635.,
       122391.,  121872.])
```

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.30,random_stat
```

```
from sklearn.linear_model import LinearRegression
linear=LinearRegression()
```

```
linear.fit(x_train,y_train)
```
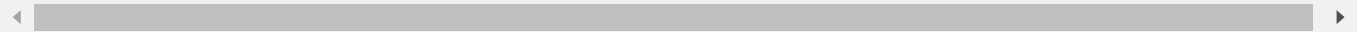
```
LinearRegression()
```

```
pred=linear.predict(x_test)
```

```
from sklearn.metrics import accuracy_score
```

```
print("Accuracy score",y_test)
```

```
Accuracy score [ 37731. 122391.  57081.  63218. 116969. 109431. 112635.  55794.  83088.
```
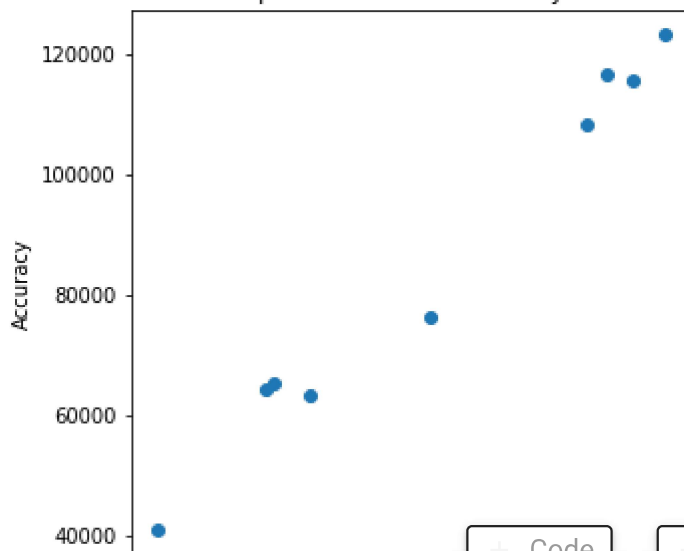
```
plt.figure(figsize=(5,5))
plt.scatter(y_test,pred)
plt.title("performance vs Accuracy")
plt.xlabel("Performance")
plt.ylabel("Accuracy")
```

```
Text(0, 0.5, 'Accuracy')
```



performance vs Accuracy

✓ 0s    completed at 10:40 AM                                          ● ✕