

▼ Machine Learning

▼ Simple Linear Progression

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

▼ Step-2 Import Data

```
df=pd.read_csv('salary_data.csv')
df.head()
```

```

0      1.1  39343
1      1.3  46205
2      1.5  37731
3      2.0  43525
4      2.2  39891
```

▼ Step-3 Selecting input and Output variables

```
X=df[["YearsExperience"]]
y=df["Salary"]
```

```
y.head()
0    39343
1    46205
2    37731
3    43525
4    39891
Name: Salary, dtype: int64
```

▼ Step-4 Making Linear Regression model

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

▼ Step-5 Fitting the model

```
model=model.fit(X,y)
model
```

```

LinearRegression
LinearRegression()
```

▼ Step-6 Predicting the model

```
model.predict([[50]])

/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression was
warnings.warn(
array([498290.31627142])
```

▼ Step-7 Evaluating the model

```

from sklearn.model_selection import train_test_split
X_train , y_train, X_test, y_test = train_test_split(X,y,test_size=0.2, random_state=0)
#fit the model
model.fit(X_train, y_train)

```

```

-----
ValueError                                Traceback (most recent call last)
<ipython-input-24-78bd0580c0e1> in <cell line: 4>()
      2 X_train , y_train, X_test, y_test = train_test_split(X,y,test_size=0.2,
random_state=0)
      3 #fit the model
----> 4 model.fit(X_train, y_train)

----- 3 frames -----
/usr/local/lib/python3.10/dist-packages/sklearn/utils/validation.py in
check_consistent_length(*arrays)
    395     uniques = np.unique(lengths)
    396     if len(uniques) > 1:
--> 397         raise ValueError(
    398             "Found input variables with inconsistent numbers of samples:
%n"
    399             % [int(l) for l in lengths])

ValueError: Found input variables with inconsistent numbers of samples: [24, 6]

```