

```
In [1]: import pandas as pd
import numpy as np
import statsmodels.api as sm
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error, r2_score
```

```
In [2]: np.random.seed(42)
n = 100
CIE = np.random.normal(50, 10, n)
SEE = 0.7 * CIE + np.random.normal(0, 5, n)
```

```
In [12]: import pandas as pd
data = pd.DataFrame({'CIE': CIE, 'SEE': SEE})
data
```

```
Out[12]:
```

	CIE	SEE
0	54.967142	31.400145
1	48.617357	31.928923
2	56.476885	37.820247
3	65.230299	41.649823
4	47.658466	32.554498
...
95	35.364851	26.681982
96	52.961203	32.653555
97	52.610553	37.596012
98	50.051135	35.326838
99	47.654129	27.643039

100 rows × 2 columns

```
In [13]: X = data[['CIE']]
y = data['SEE']
```

```
In [14]: X
```

Out[14]:

	CIE
0	54.967142
1	48.617357
2	56.476885
3	65.230299
4	47.658466
...	...
95	35.364851
96	52.961203
97	52.610553
98	50.051135
99	47.654129

100 rows × 1 columns

In [15]:

y

Out[15]:

0	31.400145
1	31.928923
2	37.820247
3	41.649823
4	32.554498
...	...
95	26.681982
96	32.653555
97	37.596012
98	35.326838
99	27.643039

Name: SEE, Length: 100, dtype: float64

In [16]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

In [17]:

```
lr = LinearRegression()
lr.fit(X_train, y_train)
```

Out[17]:

LinearRegression

LinearRegression()

In [18]:

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

In [19]:

y_pred

Out[19]:

```
array([31.81169445, 38.92193468, 37.347032, 30.54328087, 25.76926612,
       36.31169193, 35.4978579, 33.69063814, 32.15685648, 38.19853017,
       29.35915042, 31.28666711, 44.91852556, 28.41723903, 35.68379851,
       33.59951272, 35.62068124, 33.1914161, 36.59549227, 46.72847902])
```

In [20]:

```
from sklearn.metrics import mean_squared_error, r2_score
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

```
In [21]: from sklearn.metrics import mean_squared_error, r2_score  
mse = mean_squared_error(y_test, y_pred)  
r2 = r2_score(y_test, y_pred)
```

```
In [22]: r2
```

```
Out[22]: 0.5256018286178993
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```