

```
import pandas as pd
```

```
salary = pd.read_csv('https://github.com/YBI-Foundation/Dataset/raw/main/Salary%20Data.csv')
```

```
salary.head()
```

```
↗
```

	Experience Years	Salary
0	1.1	39343
1	1.2	42774
2	1.3	46205
3	1.5	37731
4	2.0	43525

```
salary.info()
```

```
↗
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 40 entries, 0 to 39
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Experience Years  40 non-null    float64
1   Salary           40 non-null    int64
dtypes: float64(1), int64(1)
memory usage: 768.0 bytes
```

```
salary.describe()
```

```
↗
```

	Experience Years	Salary
count	40.000000	40.000000
mean	5.152500	74743.625000
std	2.663715	25947.122885
min	1.100000	37731.000000
25%	3.200000	56878.250000
50%	4.600000	64472.500000
75%	6.875000	95023.250000
max	10.500000	122391.000000

```
salary.columns
```

```
↗ Index(['Experience Years', 'Salary'], dtype='object')
```

```
y = df['Salary'] #y is output or outcome that we want to Predict
x = df[['Experience Years']]
```

```
salary.shape
```

```
↗ (40, 2)
```

```
x.shape, y.shape
```

```
↗ ((40, 1), (40,))
```

```
from sklearn.model_selection import train_test_split
x_train,x_test, y_train, y_test = train_test_split(x,y, random_state=2526)
```

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

```
model = LinearRegression()
```

```
model.fit(x_train, y_train)
```



```
LinearRegression  
LinearRegression()
```

```
model.intercept_
```



```
27690.53780859281
```

```
model.coef_
```



```
array([9328.74506655])
```

```
y_pred = model.predict(x_test)
```

```
from sklearn.metrics import mean_absolute_error, mean_absolute_percentage_error
```

```
mean_absolute_percentage_error(y_test, y_pred)
```



```
0.08425483386569195
```