

```
import pandas as pd
import numpy as np
from sklearn import linear_model
from word2number import w2n
```

```
pip install word2number
```

Collecting word2number

Downloading <https://files.pythonhosted.org/packages/4a/29/a31940c848521f0725f0df6b25dc>

Building wheels for collected packages: word2number

Building wheel for word2number (setup.py) ... done

```
Created wheel for word2number: filename=word2number-1.1-cp37-none-any.whl size=5584 st
```

Stored in directory: /root/.cache/pip/wheels/46/2f/53/5f5c1d275492f2fce1cdab9a9bb12d49

Successfully built word2number

Installing collected packages: word2number

Successfully installed word2number-1.1

```
df = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/hiring.csv')
df
```

| | experience | test_score(out of 10) | interview_score(out of 10) | salary(\$) |
|---|------------|-----------------------|----------------------------|------------|
| 0 | NaN | 8.0 | 9 | 50000 |
| 1 | NaN | 8.0 | 6 | 45000 |
| 2 | five | 6.0 | 7 | 60000 |
| 3 | two | 10.0 | 10 | 65000 |
| 4 | three | 9.0 | 6 | 70000 |
| 5 | three | 7.0 | 10 | 62000 |
| 6 | ten | NaN | 7 | 72000 |
| 7 | eleven | 7.0 | 8 | 80000 |

```
df.experience = df.experience.fillna('zero')
df
```

| | experience | test_score(out of 10) | interview_score(out of 10) | salary(\$) |
|----------|------------|-----------------------|----------------------------|------------|
| 0 | zero | 8.0 | 9 | 50000 |
| 1 | zero | 8.0 | 6 | 45000 |
| 2 | five | 6.0 | 7 | 60000 |

```
df.experience = df.experience.apply(w2n.word_to_num)
df
```

| | experience | test_score(out of 10) | interview_score(out of 10) | salary(\$) |
|----------|------------|-----------------------|----------------------------|------------|
| 0 | 0 | 8.0 | 9 | 50000 |
| 1 | 0 | 8.0 | 6 | 45000 |
| 2 | 5 | 6.0 | 7 | 60000 |
| 3 | 2 | 10.0 | 10 | 65000 |
| 4 | 7 | 9.0 | 6 | 70000 |
| 5 | 3 | 7.0 | 10 | 62000 |
| 6 | 10 | NaN | 7 | 72000 |
| 7 | 11 | 7.0 | 8 | 80000 |

```
mean_test_score = math.floor(df['test_score(out of 10)'].mean())
mean_test_score
```

7

Saved successfully!



```
df['test_score(out of 10)'].fillna(mean_test_score)
df
```

```
experience test_score(out of 10) interview_score(out of 10) salary($)
```

```
reg = linear_model.LinearRegression()  
reg.fit(df[['experience', 'test_score(out of 10)', 'interview_score(out of 10)']], df['salary'])  
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

```
reg.coef_
```

```
array([2922.26901502, 2221.30909959, 2147.48256637])
```

```
reg.intercept_
```

```
14992.65144669314
```

```
reg.predict([[2, 9, 6]])
```

```
array([53713.86677124])
```

```
2922.26901502*2 + 2221.30909959*9 + 2147.48256637*6 + 14992.651446693118
```

```
53713.86677126312
```

```
reg.predict([[12, 10, 10]])
```

```
array([93747.79628651])
```

```
2922.26901502*12 + 2221.30909959*10 + 2147.48256637*10 + 14992.651446693118
```

Saved successfully!



✓ 0s completed at 20:18



Saved successfully!

