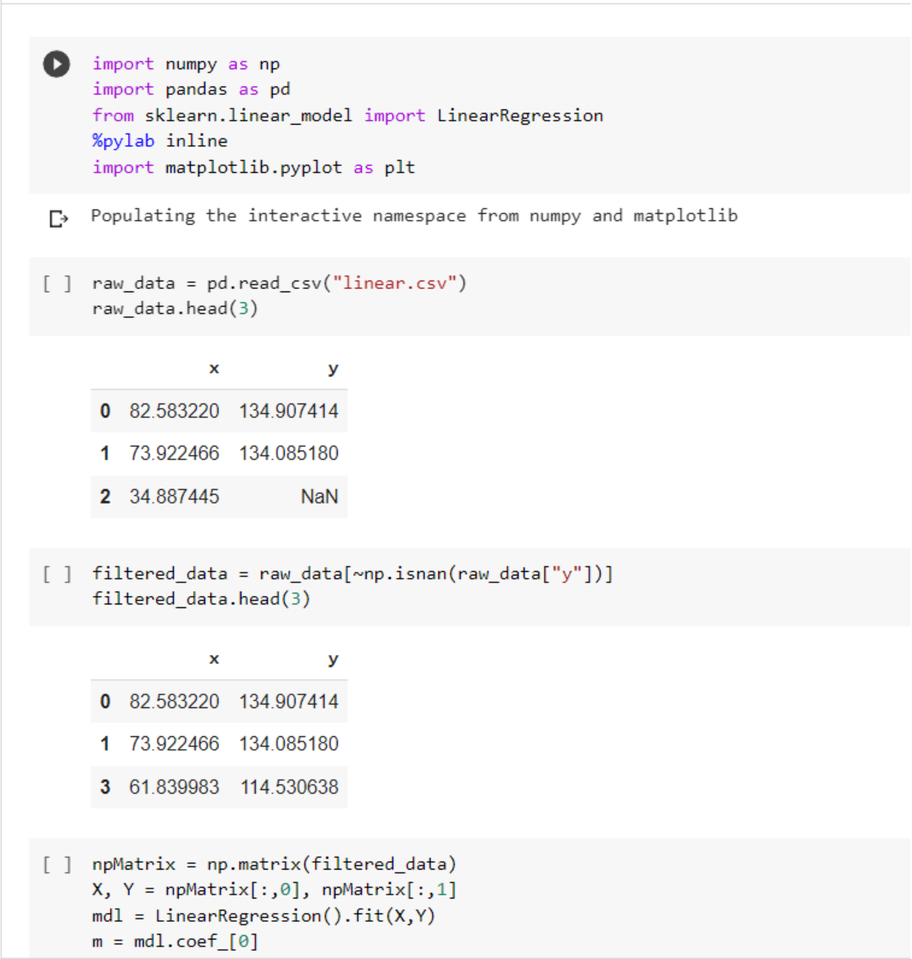
# Practical 3

## Aim: Linear Regression

## Code:

```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
%pylab inline
import matplotlib.pyplot as plt
raw_data = pd.read_csv("linear.csv")
raw_data.head(3)
filtered_data = raw_data[~np.isnan(raw_data["y"])]
filtered_data.head(3)
npMatrix = np.matrix(filtered_data)
X, Y = npMatrix[:,0], npMatrix[:,1]
mdl = LinearRegression().fit(X,Y)
m = mdl.coef_[0]
b = mdl.intercept_
formula: y = {0}x + {1}".format(m,b)
X1=np.array(X)
Y1=np.array(Y)
plt.scatter(X1,Y1, color='blue')
plt.plot([0,100],[b,m*100+b],'r')
plt.title('Linear Regression Example', fontsize = 20)
plt.xlabel('X', fontsize = 15)
plt.ylabel('Y', fontsize = 15)
```

# Output:



```
import numpy as np
import pandas as pd
from sklearn.linear_model import LinearRegression
%pylab inline
import matplotlib.pyplot as plt
```

```
Populating the interactive namespace from numpy and matplotlib
```

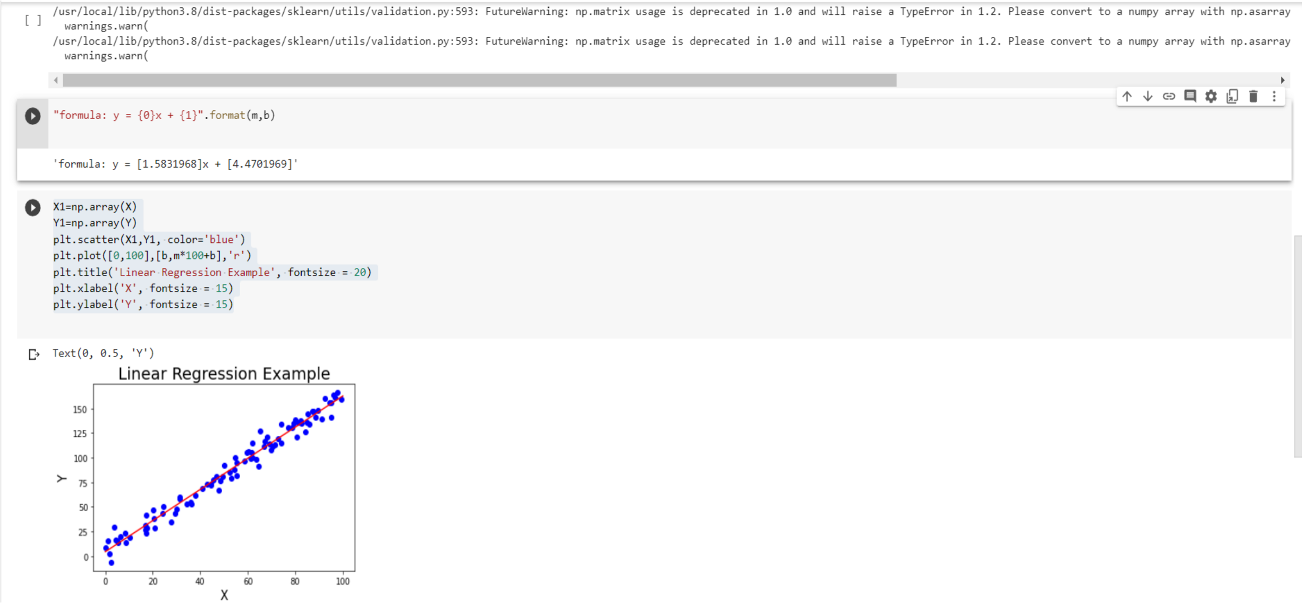
```
raw_data = pd.read_csv("linear.csv")
raw_data.head(3)
```

|   | x | y |
|---|---|---|
| 0 | 82.583220 | 134.907414 |
| 1 | 73.922466 | 134.085180 |
| 2 | 34.887445 | NaN |

```
filtered_data = raw_data[~np.isnan(raw_data["y"])]
filtered_data.head(3)
```

|   | x | y |
|---|---|---|
| 0 | 82.583220 | 134.907414 |
| 1 | 73.922466 | 134.085180 |
| 3 | 61.839983 | 114.530638 |

```
npMatrix = np.matrix(filtered_data)
X, Y = npMatrix[:,0], npMatrix[:,1]
mdl = LinearRegression().fit(X,Y)
m = mdl.coef_[0]
```



```
/usr/local/lib/python3.8/dist-packages/sklearn/utils/validation.py:593: FutureWarning: np.matrix usage is deprecated in 1.0 and will raise a TypeError in 1.2. Please convert to a numpy array with np.asarray
  warnings.warn(
/usr/local/lib/python3.8/dist-packages/sklearn/utils/validation.py:593: FutureWarning: np.matrix usage is deprecated in 1.0 and will raise a TypeError in 1.2. Please convert to a numpy array with np.asarray
  warnings.warn(
```

```
"formula: y = {0}x + {1}".format(m,b)
```

```
'formula: y = [1.5831968]x + [4.4701969]'
```

```
X1=np.array(X)
Y1=np.array(Y)
plt.scatter(X1,Y1, color='blue')
plt.plot([0,100],[b,m*100+b],'r')
plt.title('Linear Regression Example', fontsize = 20)
plt.xlabel('X', fontsize = 15)
plt.ylabel('Y', fontsize = 15)
```

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