

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

df = pd.read_excel('data/data.xls')
```

In [2]:

df[0: 2]

Out[2]:

	date	block	frequency_hz	ui_rate_rs	cap_rate_rs	rate101_rs	drawal_mus	schedule_r
0	10-Sep-18	1	49.90	3.8640	0	0	0.079993	0.078
1	10-Sep-18	2	49.94	3.0304	0	0	0.079763	0.078

In [13]:

df.loc[:5, ['date', 'drawal_mus']]

Out[13]:

	date	drawal_mus
0	10-Sep-18	0.079993
1	10-Sep-18	0.079763
2	10-Sep-18	0.078928
3	10-Sep-18	0.077580
4	10-Sep-18	0.075975
5	10-Sep-18	0.075724

In [14]:

```
x = df.iloc[:, 1].values
y = df.iloc[:, 6].values
```

In [15]:

```
x1 = df.loc[:, ['block']].values
y1 = df.loc[:, ['drawal_mus']].values
```

In [16]:

```
x1[:5]
```

Out[16]:

```
array([[1],  
       [2],  
       [3],  
       [4],  
       [5]])
```

In [7]:

```
xTrain, xTest, yTrain, yTest = train_test_split(x1, y1, test_size = 1/3, random_  
state = 0)
```

In [8]:

```
linearRegressor = LinearRegression()
```

In [17]:

```
xTrain[:5]
```

Out[17]:

```
array([[91],  
       [47],  
       [76],  
       [70],  
       [71]])
```

In [20]:

```
linearRegressor.fit(xTrain, yTrain)
```

Out[20]:

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None,  
normalize=False)
```

In [21]:

```
yPrediction = linearRegressor.predict(xTest)
```

In [22]:

```
plt.scatter(xTrain, yTrain, color = 'red')  
plt.plot(xTrain, linearRegressor.predict(xTrain), color = 'blue')  
plt.title('Salary vs Experience (Training set)')  
plt.xlabel('block')  
plt.ylabel('load')  
plt.show()
```

