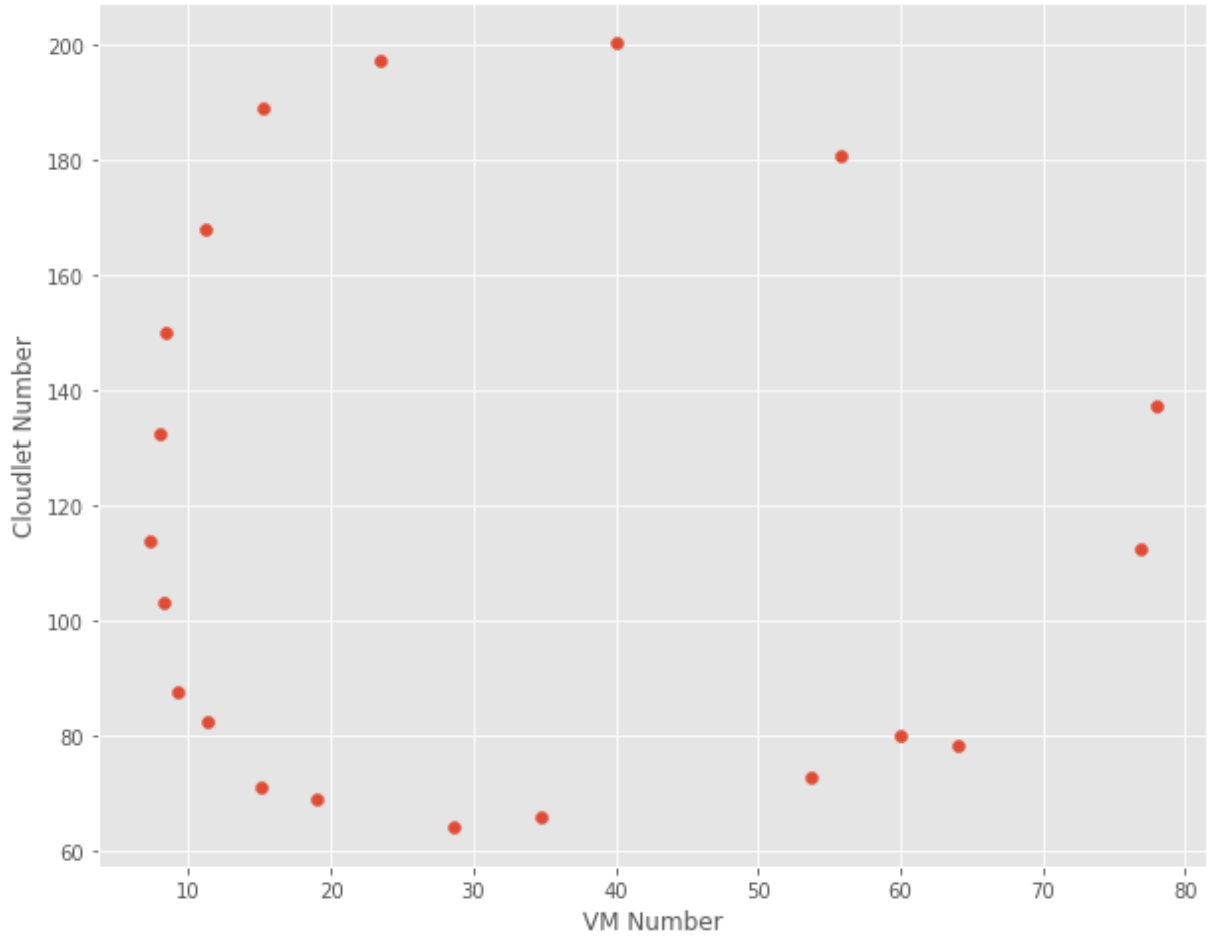
In [2]:
```
#Author: Suryoday Basak
#suryodaybasak.info
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib as mpl
plt.style.use('ggplot')
mpl.rcParams['figure.figsize'] = (10,8)
```

In [3]:
```
#Reading the data
df = pd.read_csv('../datsets/cloud/cloudlet-prey-decreasing-predator.csv')
print(df)
```

```
    VM Number  Cloudlet Number
0        60.00            80.00
1        34.73            66.19
2        18.97            69.16
3        11.40            82.47
4         8.33           103.14
5         7.99           132.47
6        11.21           168.11
7        23.44           197.33
8        55.89           180.68
9        76.83           112.69
10       53.72            72.76
11       28.69            64.30
12       15.14            71.04
13        9.29            87.74
14        7.23           114.07
15        8.34           150.29
16       15.28           189.09
17       40.13           200.55
18       77.98           137.38
19       64.06            78.36
```

In [4]:
```
#Plot the data here to visualize the trend
#plt.plot(df['VM Number'], df['Cloudlet Number'])
plt.scatter(df['VM Number'], df['Cloudlet Number'])
plt.xlabel('VM Number')
plt.ylabel('Cloudlet Number')
plt.show()
#plt.clf()
```



In [5]:
```
n = df['VM Number'].count()                 #Number of samples
p = np.sum(np.square(df['VM Number']))      #The sum of x^2
q = df['VM Number'].sum()                   #The sum of x
r = np.sum(df['VM Number']*df['Cloudlet Number'])    #The sum of the produ
s = df['Cloudlet Number'].sum()                  #The sum of y^2

#Print all of the above
print("The number of samples is:\t\t\t", n)
print("The sum of (VM Number)^2 is:\t\t\t", p)
print("The sum of VM Number is:\t\t\t", q)
print("The sum of VM Number*Cloudlet Number is:\t", r)
print("The sum of Cloudlet Number is:\t\t\t",s)
```

```
The number of samples is:                       20
The sum of (VM Number)^2 is:                    31305.627500000002
The sum of VM Number is:                        628.6500000000001
The sum of VM Number*Cloudlet Number is:        72926.8294
The sum of Cloudlet Number is:                  2357.82
```

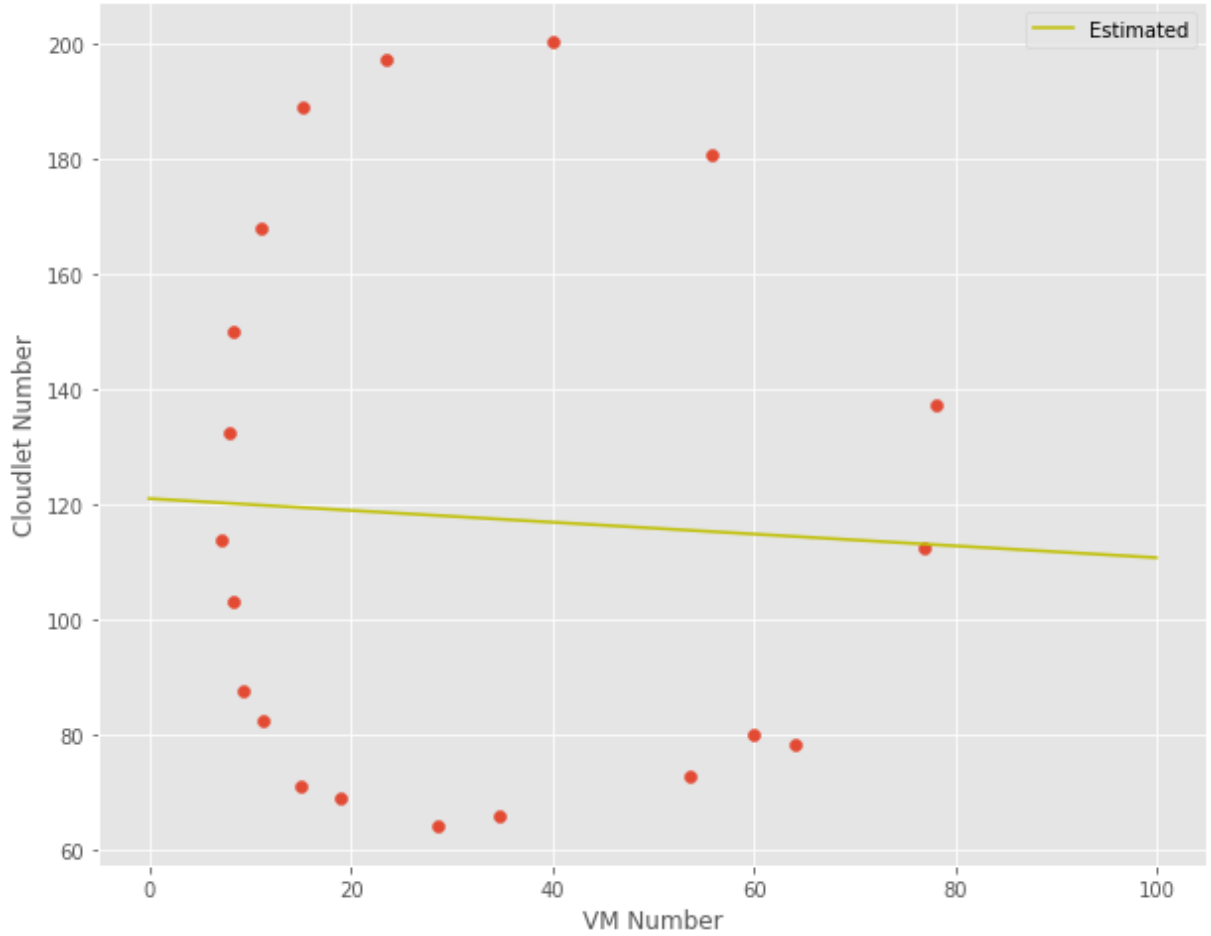
In [6]:
```
m = (1/((n*p) - (q**2)))*((n*r) - (q*s))    #The slope of the line
c = (1/((n*p) - (q**2)))*((p*s) - (r*q))    #The y-intercept of the line

print("The slope of the estimated line is:\t\t", m)
print("The y-intercept of the estimated line is:\t", c)
```

```
The slope of the estimated line is:             -0.10266674307393207
The y-intercept of the estimated line is:       121.11807240167138
```

In [7]:
```
#To visualize the estimated line, create an x-vs-y set using m and c
x = [x/10 for x in range (0, 1000)]
y = [m*xi + c for xi in x]

#Plot again to visualize how the estimated line fairs against the original
#orig, = plt.plot(df['M (g)'], df['T (s)'], label = "Actual")
plt.scatter(df['VM Number'], df['Cloudlet Number'])
est, = plt.plot(x, y, label = "Estimated",color='y')
plt.xlabel('VM Number')
plt.ylabel('Cloudlet Number')
plt.legend(handles=[est])
plt.show()
#plt.clf()
```



In [8]:
```
#Finding the error
error = 0.0
for index, row in df.iterrows():
    error += ((m*row['VM Number'] + c) - row['Cloudlet Number'])**2 #(Esti
error/=n

print("The mean squared error is:\t\t", error)
print("The root means squared error is:\t", error**(0.5))
```

```
The mean squared error is:              2188.5029793548742
The root means squared error is:        46.78143840621913
```

In [ ]:
```

```