In [1]:

```
import numpy as np
```

In [4]:

```
X = np.random.uniform(1,100, 40)
```

In [5]:

```
X
```

Out[5]:

```
array([ 32.62384617,  97.76616684,  18.45604859,  89.38966067,
        83.08471878,  84.17247045,   1.6446649 ,  80.3454912 ,
        50.31885577,  98.4514469 ,  58.87341152,  88.29094685,
        46.00478703,  46.06341121,  92.7584417 ,  18.05047934,
        17.6739869 ,  58.42424991,  96.36037729,  81.42853561,
         2.71674363,  38.61354279,  55.37144628,  84.94226066,
         9.6207141 ,  65.43328447,  55.5302203 ,  20.6315872 ,
        82.82076171,  91.50043778,  53.13019598,  85.1538909 ,
        56.43425701,  13.16348863,  17.03289216,  83.29445942,
        97.94395866,  92.43837585,  51.32603959,  97.73515552])
```

In [6]:

```
Y = np.square(X)
```

In [7]:

```
Y
```

Out[7]:

```
array([  1.06431534e+03,   9.55822338e+03,   3.40625730e+02,
         7.99051143e+03,   6.90307050e+03,   7.08500478e+03,
         2.70492264e+00,   6.45539796e+03,   2.53198725e+03,
         9.69268740e+03,   3.46607858e+03,   7.79529130e+03,
         2.11644043e+03,   2.12183785e+03,   8.60412851e+03,
         3.25819805e+02,   3.12369813e+02,   3.41339298e+03,
         9.28532231e+03,   6.63060641e+03,   7.38069596e+00,
         1.49100569e+03,   3.06599706e+03,   7.21518765e+03,
         9.25581399e+01,   4.28151472e+03,   3.08360537e+03,
         4.25662390e+02,   6.85927857e+03,   8.37233011e+03,
         2.82281773e+03,   7.25118513e+03,   3.18482536e+03,
         1.73277433e+02,   2.90119415e+02,   6.93796697e+03,
         9.59301904e+03,   8.54485333e+03,   2.63436234e+03,
         9.55216062e+03])
```

In [57]:

```
for x in np.nditer(Y, op_flags=['readwrite']):
    x+= np.random.normal(0,2000)
    print (x)
```

```
2148.819534500682
12157.82278930437
-3145.5857186922494
4481.1031949047565
2786.080936790922
7586.387521038567
-5022.268189503136
6052.060259886873
3535.0801961548714
6314.287846554143
11562.54863579135
10220.135203826409
-181.46143661559717
3071.0127821226242
10177.998761077733
-1532.1269580882918
3285.3223703534304
4248.91322329175
10195.987921950367
11180.439963244276
3564.2898639383893
1483.6532130502
6136.138372710225
2868.28550001547
1170.8644532505941
1119.529108180667
2610.1821479431046
3279.45435639579
5104.584551126718
10932.895240784988
6877.1070423674455
3647.4889953037227
7098.922031630405
6449.2242637643285
-2417.2739041532404
13655.904290455626
7920.009315548117
10755.701405745769
-1036.0893775957832
8364.464791664286
```

In [51]:

```
import matplotlib.pyplot as plt
```

In [52]:

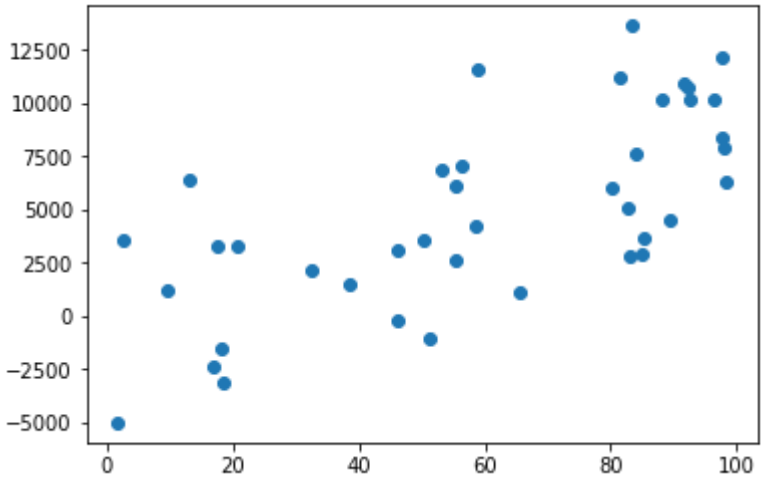
```
Y
```

Out[52]:

```
array([  1936.85397753,  10398.55184017,   -495.03265475,   6811.85373288,
         4826.64417379,   7050.79205528,  -1005.87360758,   4188.94066546,
         4065.9284831 ,   6600.38654925,  11757.53232548,   5939.10115456,
          738.17150147,   3717.44552965,   8535.17199429,    588.89391111,
         1403.19152992,   4326.55878088,   8326.10759196,   7521.88679787,
         1322.97125997,   4949.78759973,   4622.38567804,   4690.95778359,
        -1956.6248363 ,   2165.40938844,   2659.68785805,    -96.33597235,
         9927.66998496,   9347.85662748,   1165.74289222,   7901.97770879,
         3085.92828116,   6001.91301814,    322.40292798,  10795.95931572,
        10102.56332727,  11619.25465082,   -479.85435041,   8972.4626169
7])
```

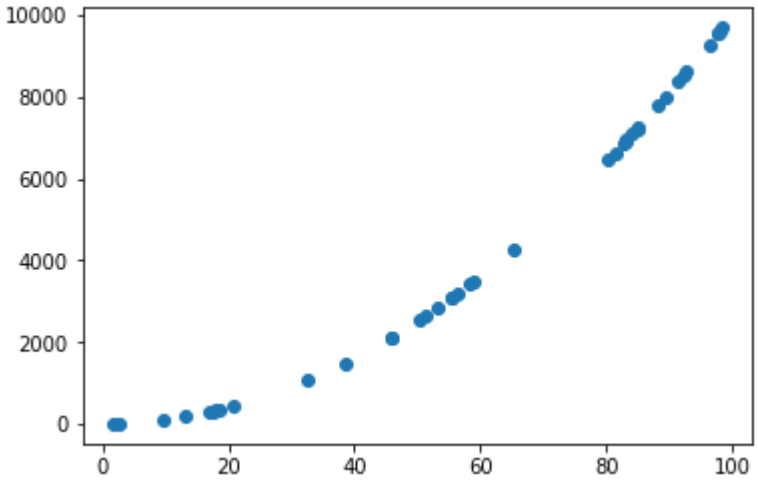
In [58]:

```
plt.scatter(X,Y)
plt.show()
```



In [59]:

```
plt.scatter(X, X**2)
plt.show()
```



In [82]:

```
from sklearn.model_selection import train_test_split
```

In [83]:

```
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.30)
```

In [86]:

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

In [102]:

```
lm = LinearRegression()
X_train = X_train.reshape(-1, 1)
y_train = y_train.reshape(-1, 1)
X_test = X_test.reshape(-1, 1)
y_test = y_test.reshape(-1, 1)
```

In [142]:

```
lm.fit(X_train,y_train)
predict_y = lm.predict(X_test)
predict_y
```

Out[142]:

```
array([[ -252.30934109],
       [  462.03064367],
       [ 6668.08540895],
       [ 8259.73997537],
       [ 7759.99184702],
       [ 6563.70827131],
       [ 4259.29106359],
       [ 8107.12417404],
       [ 8239.61681692],
       [ 6932.52891773],
       [ 4156.8639235 ],
       [  523.8153982 ]])
```

In [143]:

```
plt.scatter(y_test, predict_y)
plt.xlabel("True Values")
plt.ylabel("Predictions")
```

Out[143]:

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

In [120]:

```
from sklearn.metrics import mean_squared_error
```

In [121]:

```
from math import sqrt
```

In [182]:

```
rms = np.sqrt(mean_squared_error(predict_y,y_test))
```

In [185]:

```
rms
```

Out[185]:

```
5498.7186074790379
```

In [186]:

```
X_train, X_test, y_train, y_test = train_test_split(X, Y, test_size=0.30 )
X_train = X_train.reshape(-1,1)
y_train = y_train.reshape(-1,1)
X_test=X_test.reshape(-1,1)
lm = LinearRegression()
```

In [187]:

```
from sklearn.preprocessing import PolynomialFeatures
```

In [188]:

```
from sklearn.pipeline import make_pipeline
```

In [189]:

```
model = make_pipeline(PolynomialFeatures(2), lm)
```

In [190]:

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

Out[190]:

```
Pipeline(memory=None,
     steps=[('polynomialfeatures', PolynomialFeatures(degree=2, include_bi
as=True, interaction_only=False)), ('linearregression', LinearRegression(c
opy_X=True, fit_intercept=True, n_jobs=1, normalize=False))])
```

In [191]:

```
pred_test=model.predict(X_test)
```

In [192]:

```
print(sqrt(mean_squared_error(pred_test, y_test)))
```

```
3614.3606652425055
```

In [193]:

```
plt.scatter(X_test, y_test,  color='black')
```

Out[193]:

```
<matplotlib.collections.PathCollection at 0x209061b1a20>
```

In [194]:

```
plt.show()
```