In [16]:
```
# Rakibul Islam
# 151-15-5131

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
df = pd.read_csv("train.csv")
print(df)
x=df['x']
y=df['y']
```

| | x | y |
|---|---|---|
| 0 | 24.0 | 21.549452 |
| 1 | 50.0 | 47.464463 |
| 2 | 15.0 | 17.218656 |
| 3 | 38.0 | 36.586398 |
| 4 | 87.0 | 87.288984 |
| 5 | 36.0 | 32.463875 |
| 6 | 12.0 | 10.780897 |
| 7 | 81.0 | 80.763399 |
| 8 | 25.0 | 24.612151 |
| 9 | 5.0 | 6.963319 |
| 10 | 16.0 | 11.237573 |
| 11 | 16.0 | 13.532902 |
| 12 | 24.0 | 24.603239 |
| 13 | 39.0 | 39.400500 |
| 14 | 54.0 | 48.437538 |
| 15 | 60.0 | 61.699003 |
| 16 | 26.0 | 26.928324 |
| 17 | 73.0 | 70.405206 |
| 18 | 29.0 | 29.340924 |
| 19 | 31.0 | 25.308952 |
| 20 | 68.0 | 69.029343 |
| 21 | 87.0 | 84.994847 |
| 22 | 58.0 | 57.043103 |
| 23 | 54.0 | 50.592199 |
| 24 | 84.0 | 83.027722 |
| 25 | 58.0 | 57.057527 |
| 26 | 49.0 | 47.958833 |
| 27 | 20.0 | 24.342264 |
| 28 | 90.0 | 94.684883 |
| 29 | 48.0 | 48.039707 |
| .. | ... | ... |
| 670 | 84.0 | 82.889358 |
| 671 | 64.0 | 63.613650 |
| 672 | 12.0 | 11.296272 |
| 673 | 61.0 | 60.022749 |
| 674 | 75.0 | 72.603393 |
| 675 | 15.0 | 11.879646 |
| 676 | 100.0 | 100.701274 |
| 677 | 43.0 | 45.124208 |
| 678 | 13.0 | 14.811068 |
| 679 | 48.0 | 48.093680 |
| 680 | 45.0 | 42.291457 |
| 681 | 52.0 | 52.733898 |
| 682 | 34.0 | 36.723970 |
| 683 | 30.0 | 28.645352 |
| 684 | 65.0 | 62.166753 |
| 685 | 100.0 | 95.584595 |
| 686 | 67.0 | 66.043253 |
| 687 | 99.0 | 99.956622 |
| 688 | 45.0 | 46.149420 |
| 689 | 87.0 | 89.137550 |
| 690 | 73.0 | 69.717878 |
| 691 | 9.0 | 12.317366 |
| 692 | 81.0 | 78.202963 |
| 693 | 72.0 | 71.309954 |
| 694 | 81.0 | 81.455447 |

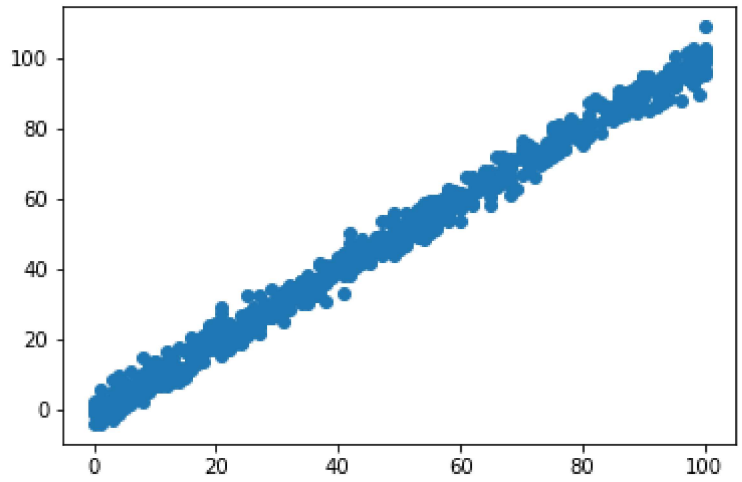
```
695   58.0   58.595006
696   93.0   94.625094
697   82.0   88.603770
698   66.0   63.648685
699   97.0   94.975266

[700 rows x 2 columns]
```

In [17]:
```
m_x = np.mean(x);
m_y = np.mean(y);
print(m_x)
print(m_y)
m_xy = np.mean(x*y)
print(m_xy)
m_xx = np.mean(x**2)
print(m_xx)
m_x2 = m_x**2
print(m_x2)
m_optimal = ((m_x*m_y)-m_xy)/(m_x2-m_xx)
print(m_optimal)
c_optimal = (m_y-(m_optimal*m_x))
print(c_optimal)
```

```
54.98593909881429
49.93986917045776
3335.424584518539
21136.701501578937
3023.4534985785144
0.03254159494814826
48.1505390124606
```

In [18]:
```
plt.scatter(x,y)
plt.show()
```

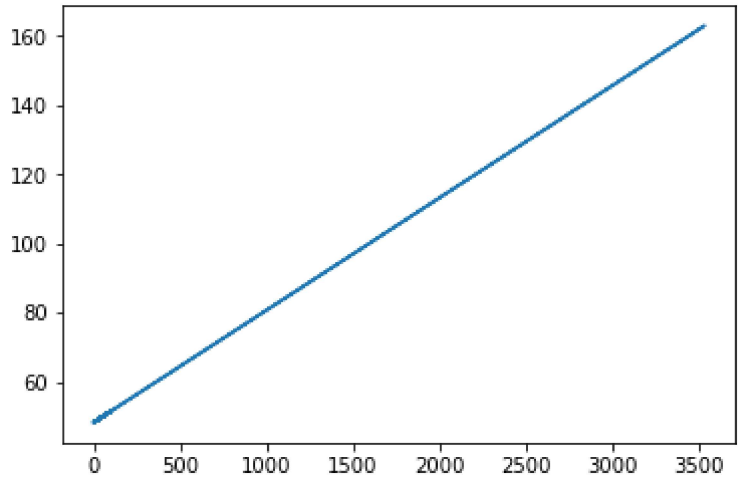


In [19]:
```
y_final = m_optimal * x + c_optimal
x_final =((y_final - c_optimal)/m_optimal)
print(x_final)
```

```
0       24.0
1       50.0
2       15.0
3       38.0
4       87.0
5       36.0
6       12.0
7       81.0
8       25.0
9        5.0
10      16.0
11      16.0
12      24.0
13      39.0
14      54.0
15      60.0
16      26.0
17      73.0
18      29.0
19      31.0
20      68.0
21      87.0
22      58.0
23      54.0
24      84.0
25      58.0
26      49.0
27      20.0
28      90.0
29      48.0
       ...
670     84.0
671     64.0
672     12.0
673     61.0
674     75.0
675     15.0
676    100.0
677     43.0
678     13.0
679     48.0
680     45.0
681     52.0
682     34.0
683     30.0
684     65.0
685    100.0
686     67.0
687     99.0
688     45.0
689     87.0
690     73.0
691      9.0
692     81.0
693     72.0
694     81.0
695     58.0
```

```
696     93.0
697     82.0
698     66.0
699     97.0
Name: x, dtype: float64
```

In [20]:
```
plt.plot(x_final,y_final)
plt.show()
```




In [ ]:
```

```