Introduction to Standard Deviation

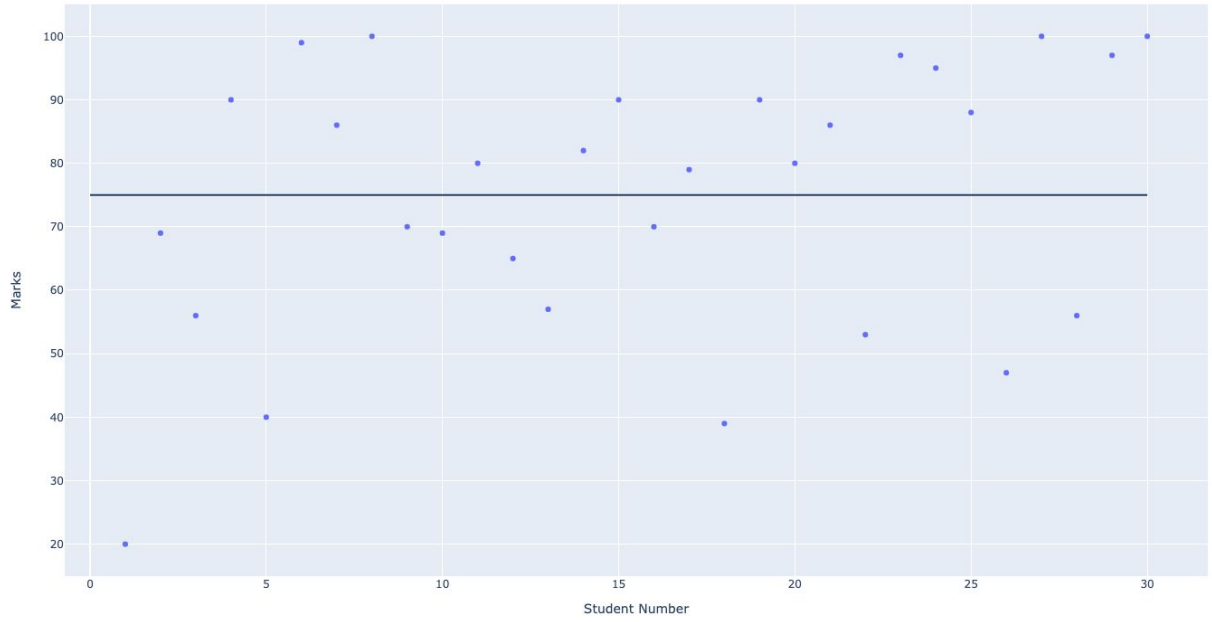
What we did:
In last class we learned to find the central tendency (mean,median,mode).

In this class we learned about the standard deviation .

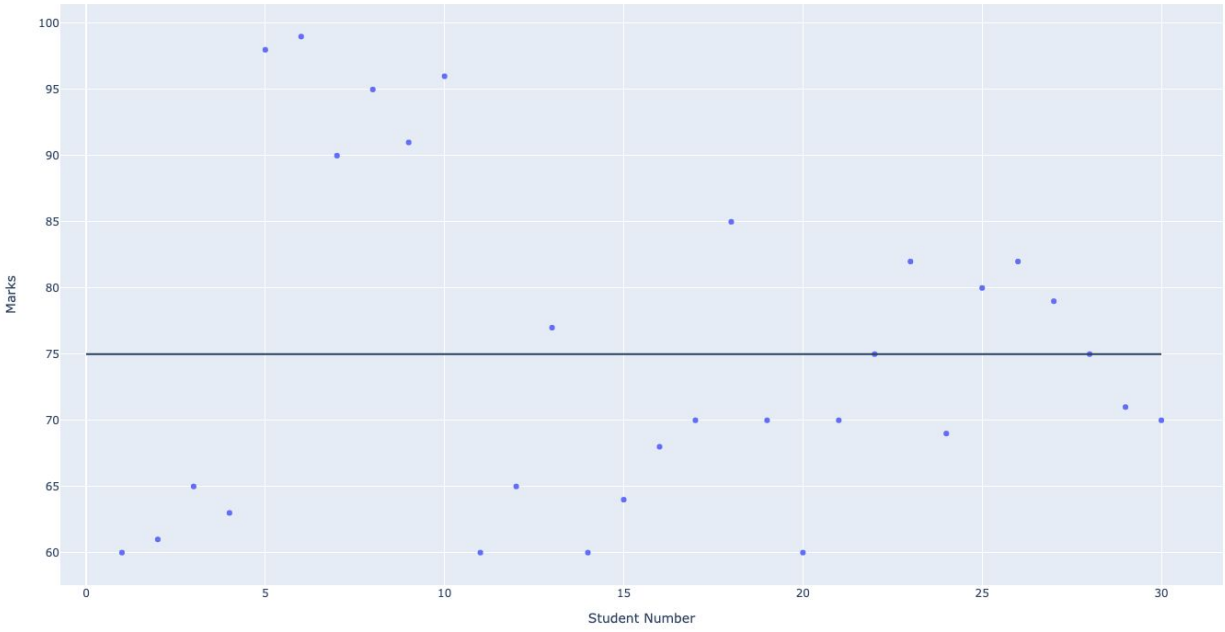
How we did it:

We revised about mean, median and mode.
We learned to display a average line to the graph

```
fig.update_layout(shapes=[
    dict(
      type= 'line',
      y0= mean, y1= mean,
      x0= 0, x1= total_entries
    )
])
```



Then we found the mean/average and plotted the average line for 2nd dataset



Then we learned about standard deviation. And saw the formula to calculate the standard deviation.

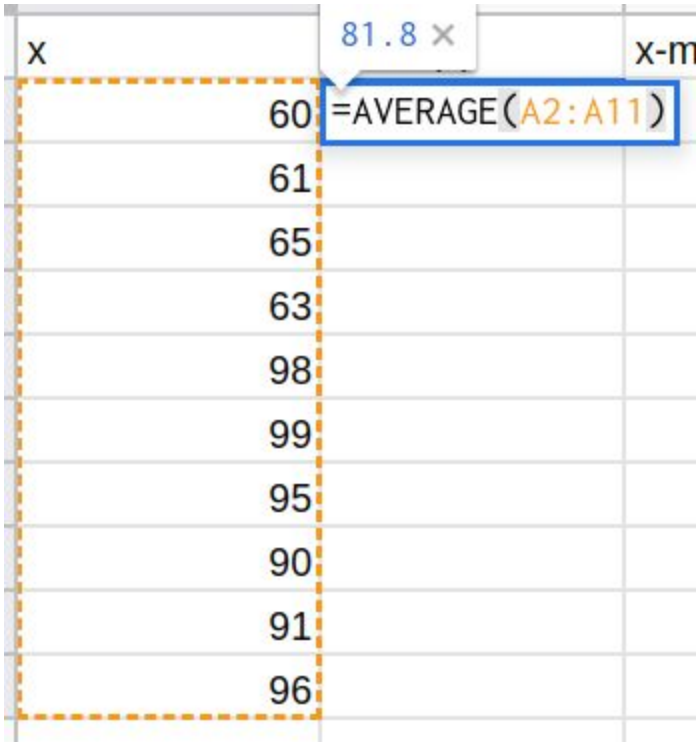
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

where:-

Xi is the each value in the dataset .
X -bar is the mean of the values.
N is the number of values in a dataset.

We saw how to calculate the deviation part by part using the google sheet
Step1 . calculate the mean

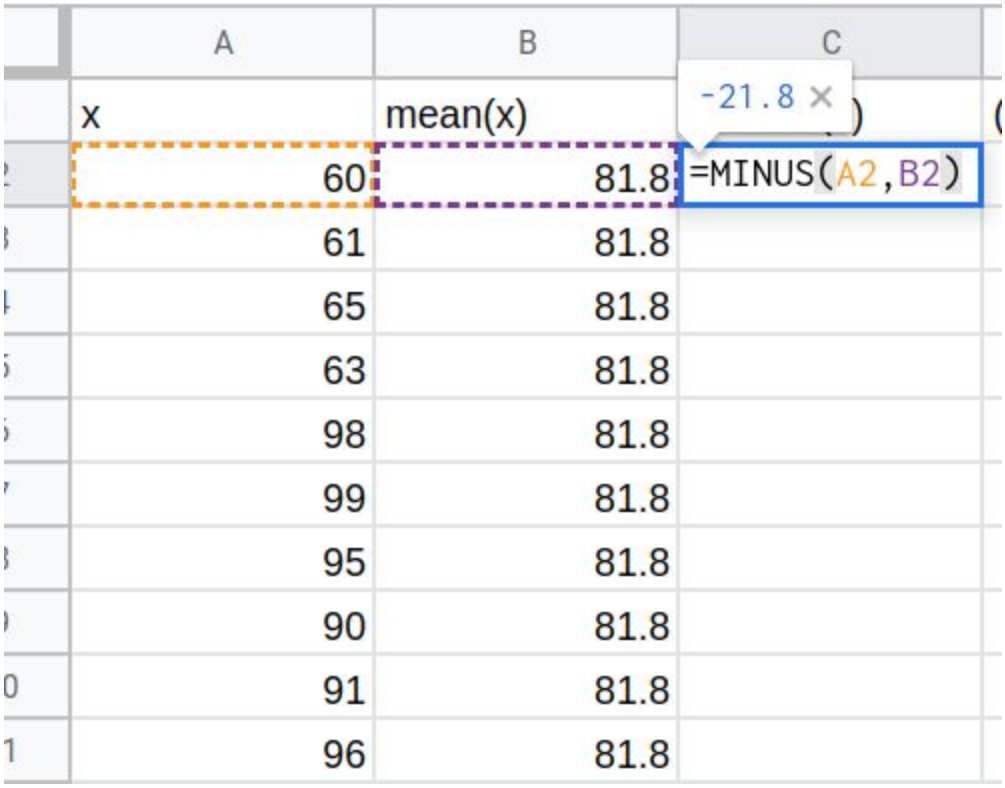
x	mean
60	
61	
65	
63	
98	
99	
95	
90	
91	
96	



	A	B
1	x	mean(x)
2	60	81.8
3	61	81.8
4	65	81.8
5	63	81.8
6	98	81.8
7	99	81.8
8	95	81.8
9	90	81.8
10	91	81.8
11	96	81.8

Step 2.Subtract the mean from each number (x)

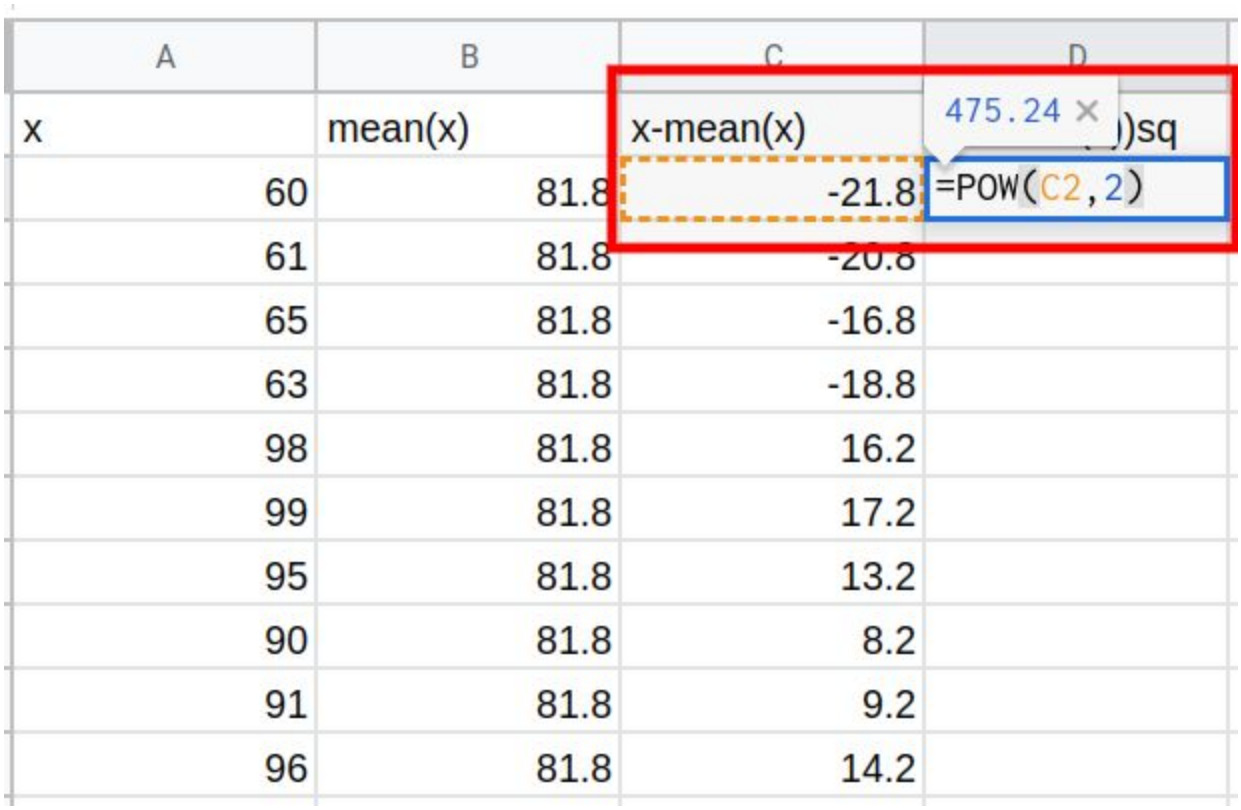
$(x_i - \bar{x})$.



	A	B	C
	x	mean(x)	x-mean(x)
	60	81.8	-21.8
	61	81.8	-20.8
	65	81.8	-16.8
	63	81.8	-18.8
	98	81.8	16.2
	99	81.8	17.2
	95	81.8	13.2
	90	81.8	8.2
	91	81.8	9.2
	96	81.8	14.2

Step 3.Square each of the differences,

$\left(x_i - \bar{x}\right)^2$.

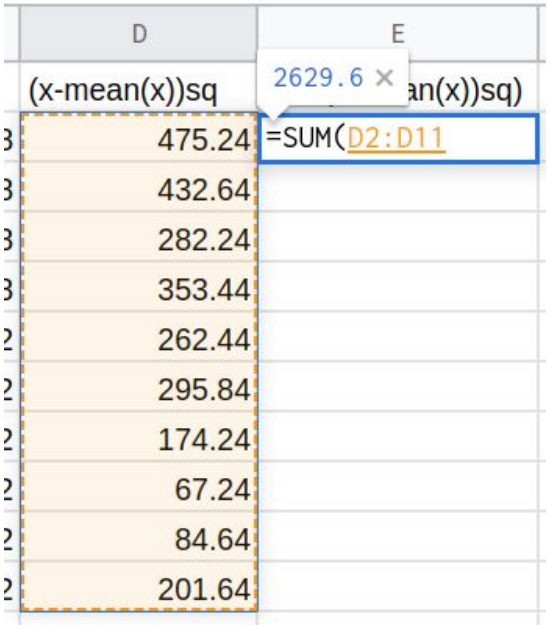


A	B	C	D
x	mean(x)	x-mean(x)	)sq
60	81.8	-21.8	=POW(C2,2)
61	81.8	-20.8	
65	81.8	-16.8	
63	81.8	-18.8	
98	81.8	16.2	
99	81.8	17.2	
95	81.8	13.2	
90	81.8	8.2	
91	81.8	9.2	
96	81.8	14.2	

475.24 ×

C	D
x-mean(x)	(x-mean(x))sq
-21.8	475.24
-20.8	432.64
-16.8	282.24
-18.8	353.44
16.2	262.44
17.2	295.84
13.2	174.24
8.2	67.24
9.2	84.64
14.2	201.64

Step4. Add up all of the results from Step 3 to get the sum of squares, $\sum(x_i - \bar{x})^2$.



D	E
(x-mean(x))sq	sum(x-mean(x))sq
475.24	2629.6
432.64	2629.6
282.24	2629.6
353.44	2629.6
262.44	2629.6
295.84	2629.6
174.24	2629.6
67.24	2629.6
84.64	2629.6
201.64	2629.6

Step 5.Divide the sum of squares (found in Step 4) by the number of numbers minus one; that is, (n – 1).

$$\frac{\sum(x_i - \bar{x})^2}{n-1}.$$

E	F
sum(x-mean(x))sq)	sum/n-1
2629.	=DIVIDE(E2,9)
2629.6	
2629.6	
2629.6	
2629.6	
2629.6	
2629.6	
2629.6	
2629.6	
2629.6	

E	F
sum(x-mean(x))sq)	sum/n-1
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778
2629.6	292.1777778

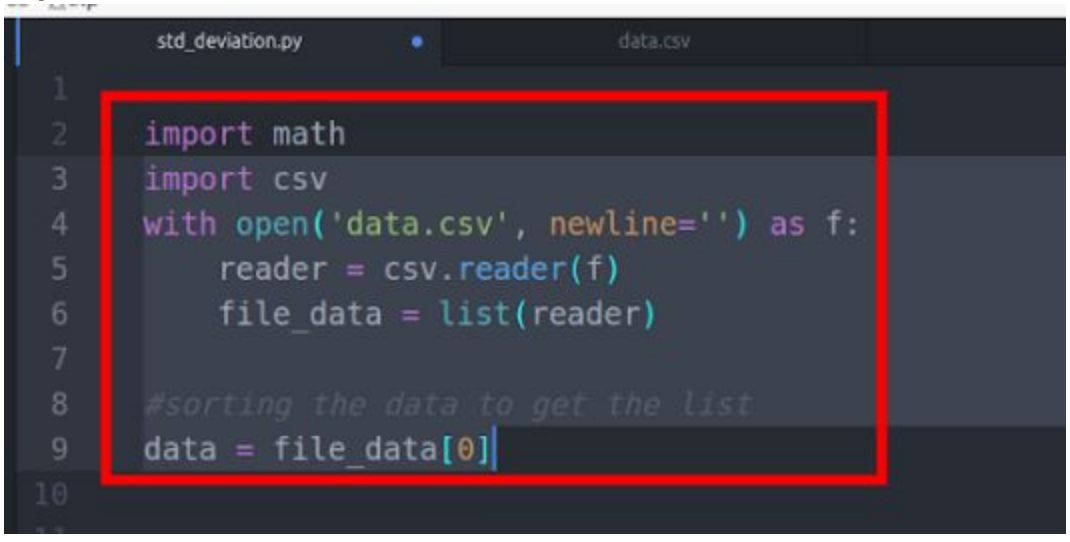
Step 6.Take the square root to get the result

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}},$$

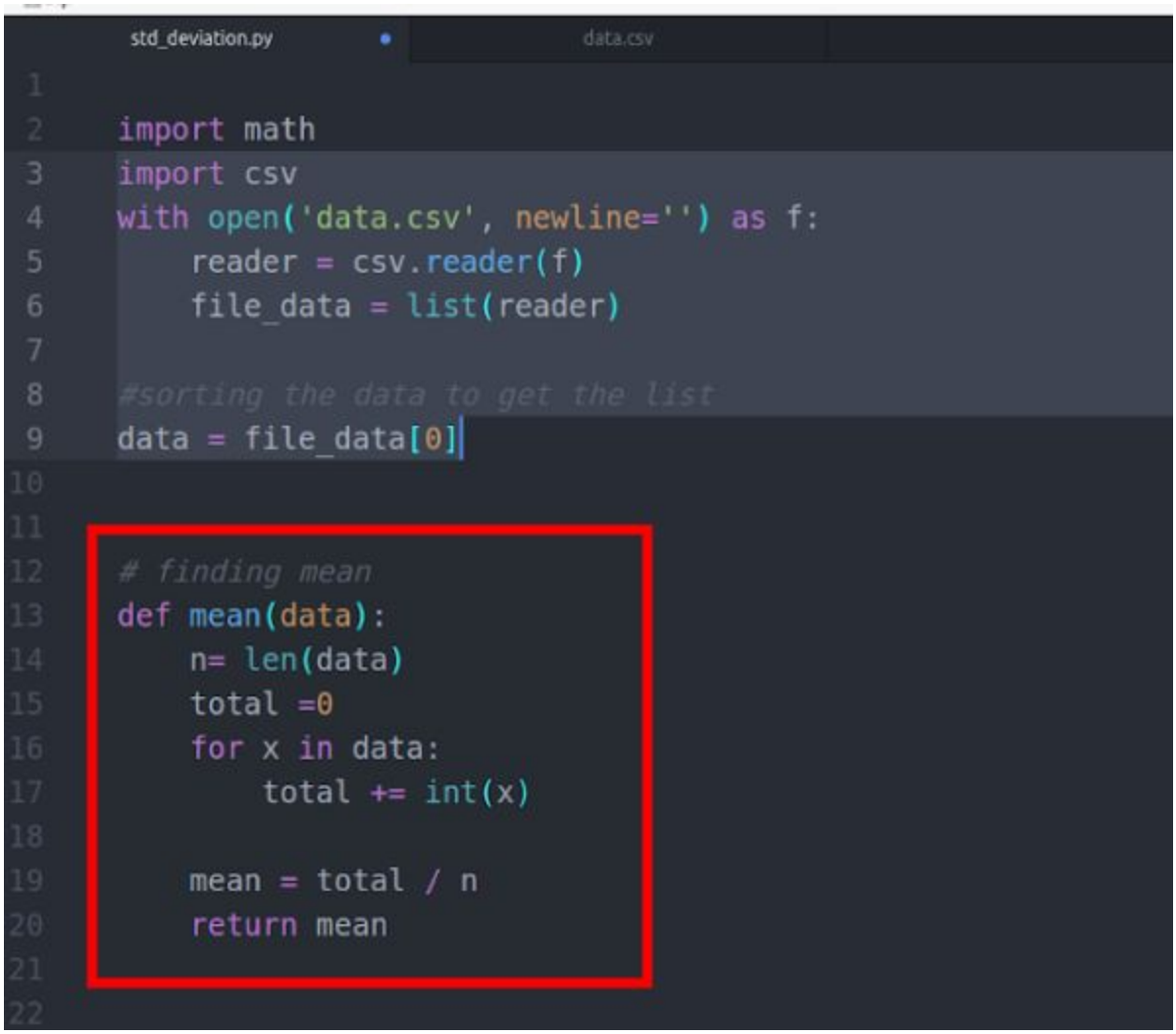
F	G
sum/n-1	(sum/n-1)sqrt
292.177777	=SQRT(F2)
292.1777778	
292.1777778	
292.1777778	
292.1777778	
292.1777778	
292.1777778	
292.1777778	
292.1777778	
292.1777778	

F	G
sum/n-1	(sum/n-1)sqrt
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853
292.1777778	17.09320853

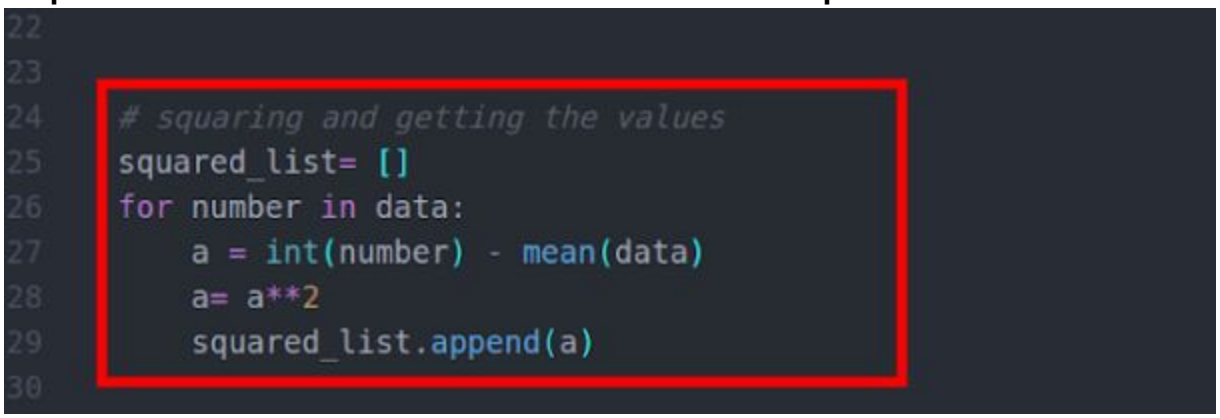
Then we wrote code to find standard deviation.
Imported the file and read it.



Found the mean



Step 3. Subtracted the mean from all the values and squared it.



Step 4. Got the sum of the squared list

```
30
31    #getting sum
32    sum =0
33    for i in squared_list:
34        sum =sum + i
35
```

Step5. Divided the sum by the number of values.

```
    #getting sum
    sum =0
    for i in squared_list:
        sum =sum + i

    #dividing the sum by the total values
    result = sum/ (len(data)-1)
```

Step6. Found the square root of the results.

```
    #dividing the sum by the total values
    result = sum/ (len(data)-1)

    # getting the deviation by taking square root of the result
    std_deviation = math.sqrt(result)
    print(std_deviation)
```

We ran and tested the code.

```
[redacted]$ python3 std_deviation.py
17.093208527885505
[redacted]$
```

What's next?

In the next class, we will learn to find correlation..