Standard Deviation
- measurement of variation.

Class Exercise 12

Student's Name: ___Andrew Graff___

Measures of variation

Q1. Consider the following three data sets A, B and C.

mean data		
A = {9,10,11,7,13}	50/5 (a) 10	
B = {10,10,10,10,10}	50/5 (b) 10	
C = {1,1,10,19,19}	50/5 (c) 10	

Sum
↓

to calculate $\bar{x} = \dfrac{\sum x}{n}$ ← how many
(mean)

a) Calculate the mean of each data set. → ⟨a= 10 b = 10 c = 10⟩
(averages)

b) Calculate the standard deviation of each data set.
⟨a = 2 b = 0 c = 8·04⟩

$$S = \sqrt{\frac{\sum (x-\bar{x})^2}{2}}$$

c) Which set has the largest standard deviation? → ⟨Set C (8·04)⟩

d) Is it possible to answer question c) without calculations of the standard deviation? **No.**

(x) 9-10
(B)

Set A	(x-x̄)	(x-x̄)²
9	−1	1
10	0	0
11	1	1
7	−3	9
13	3	9
50		20

example
9-10 (mean)

$\sum (x-\bar{x})^2 = 20$

$= \sqrt{\dfrac{20}{5}}$

$= 1.414$

$= \sqrt{4}$

⟨S = 2⟩

Set B	(x-x̄)	(x-x̄)²
10	0	0
10	0	0
10	0	6
10	0	6
10	0	0

⟨S = 0⟩

0.

Set C	(x-x̄)	(x-x̄)²
1	−9	81
1	−9	81
10	0	0
19	9	81
19	9	81
50		324

$\sum (x-\bar{x})^2 = 324$

$= \sqrt{\dfrac{324}{5}}$

$= \sqrt{64.8}$

⟨S = 8·04⟩

Q2. The frequency table of the monthly salaries of 20 people is shown below.

salary(in $)	frequency	Sal x freq	$(x-\bar{x})$	$(x-\bar{x})^2$
3500	5	17500	13545	183467025
4000	8	32000	28045	786522025
4200	5	21,000	17045	290532025
4300	2	8600	4645	21576025
	20	79100		

Fold

79100 /20 = 3955 x

a) Calculate the mean of the salaries of the 20 people.

b) Calculate the standard deviation of the salaries of the 20 people.

$$\sqrt{\dfrac{1,282,097,100}{4}}$$

$$= 320524275$$

$$= \sqrt{320524275}$$

(b) $S = 17,903.1917$

Q3. The following table shows the grouped data, in classes, for the heights of 50 people.

height (in cm) - classes	frequency	Midpoint	Midpoint × frequency
120 - 130	2	125	250
130 - 140	5	135	675
140 - 150	25	145	3625
150 - 160	10	155	1550
160 - 170	8	165	1320
Total	55		7420

a) Calculate the mean of the height of 50 people.
b) Calculate the standard deviation.

Ttl she see

Mean 7420/~~55~~ 50

mean = $\underline{134.90}$ ~

(a) $\underline{148.40 \sim 148 \, cm}$

Midpoint × frequency	$(x - \bar{x})$	$(x - \bar{x})^2$
Height		
250	~~104.7~~	13248.01
675	~~540.7~~	291708.01
3625	~~3490.7~~	12180798.01
1550	~~1415.70~~	2602508.01
1320	~~1185.70~~	1404462.01

total

		$x - \bar{x}$	$(x - \bar{x})^2$
125 - 148.40		-23	529
135 - 148		-13	169
148 - 148		-3	9
155 - 148		7	49
165 - 148		16	256
			1012

OR

$(x - \bar{x})$ height $(x - \bar{x})\sqrt{}$

250 - (148.40) | 101.60 | 10322.56 |
                 | 526.60 | 277307.56 |
                 | 3476.60 | 12086744.56 |
                 | 1401.60 | 1964482.56 |
                 | 1171.60 | 1372648.56 |
                 | 15711503.7 | total

$S = \sqrt{\dfrac{\cancel{1.77215} \; 1012}{\cancel{5} \; 5}}$

$= S\sqrt{202.40}$

(b) $S = 14.22 \, cm$

$\sqrt{\dfrac{15711503.7}{5}}$

$\sqrt{3142300.74}$

$= 1772.65$

Q4. Nine friends each guessed the number of marbles in a jar.

When the answer was revealed they found they had guessed well (and one was the winner!)

Here is how close they each got:

-9, -7, -4, -1, 0, 2, 7, 9, 12

(A negative number shows an underestimate, a positive number shows an overestimate.)

What was the standard deviation of their errors?

$(x-\bar{x})$	$(x-\bar{x})^2$
-9	-10	100
-7	-8	64
-4	-5	25
-1	-2	4
0	-1	1
2	1	1
7	6	36
9	8	64
12	11	121

Total 9

416

Mean $\frac{9}{9}$

$= 1$

$\sum (x-\bar{x})^2 = 416$

$S = \sqrt{\frac{\sum (x-\bar{x})^2}{x}}$

$S = \sqrt{\frac{416}{9}}$

$S = 6.7986$

Q.5 What is the variance of the first 10 numbers of the Fibonacci sequence {0, 1, 1, 2, 3, 5, 8, 13, 21, 34} ?

sample variance
$$S^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$$

$\downarrow$

size of sample

$0, 1, 1, 2, 3, 5, 8, 13, 21, 34$

-8.8

Data	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
0	-8.8	77.44
1	-7.8	60.84
1	-7.8	60.84
2	-6.8	46.24
3	-5.8	33.64
5	-3.8	14.44
8	-0.8	0.64
13	4.2	17.64
21	12.2	148.84
34	25.2	635.04
88		1095.60

$$\bar{x} = \frac{\sum x_i}{n} = \frac{88}{10} = 8.8$$

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

$$S^2 = \frac{1095.60}{10}$$

$$\boxed{= 109.56}$$

Variance.

high variance
~ higher spread
of data.

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} \quad S^2 = \sqrt{S^2}$$
$$S = \sqrt{2 \cdot 5}$$

Q6. Daniel did a survey of the <u>number</u> of <u>pets</u> owned by his classmates, with the following results:

$$x \quad \overset{-8}{x-\bar{x}} \quad (x-\bar{x})^2$$

Standard Deviation — measure how spread out numbers are

symbol σ

$\rightarrow$ is the $\sqrt{}$ of the <u>variance</u>

Number of pets	Frequency			
0	4	4	~4	16
1	12	12	4	16
2	8	16	8	64
3	2	6	~2	4
4	1	4	~45	16
5	2	10	2	4
6	1	6	~2	4

data points → 30   58   124

$$\sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$
$$\Sigma'(x-x)^2$$

Population standard deviation

sample of population

variance = average of squared differences from mean

$\downarrow$

what mean

$\downarrow$

subtract mean and square result.

— average of square differences

S Determine the standard deviation.

$Mean = \dfrac{58}{7}$

$= 8.28 \sim 8$

$S = \sqrt{\dfrac{124}{7}}$

$S = 4.208 \sim 4$ pets.

$\sigma^2 = $ variance   n = data points. $x_i = $ ith data point and m is the mean.

variance sample $\rightarrow S^2 = \dfrac{1}{8}\Sigma(x-\bar{x})^2)$

$= \dfrac{124}{7-1}$

$= \dfrac{124}{6}$

$S^2 = 20.66$

$S = \sqrt{S^2}$

$s = 4.54$