03. FINDING A RELATIONSHIP BETWEEN A DATASET

M5U3P3

WHAT IS A DATA SET?

DATA SET OR DATASET

- A data set (or dataset) is a collection of data
- The data set lists values for each of the variables, such as height and weight of an object, for each member of the data set
- Data sets can also consist of a collection of documents or files

DATASET RELATIONSHIP

- Analysis of the relationship between two datasets shows the **direction** of the relationship and the **degree** of the relationship
- Two data sets are know as pairwise

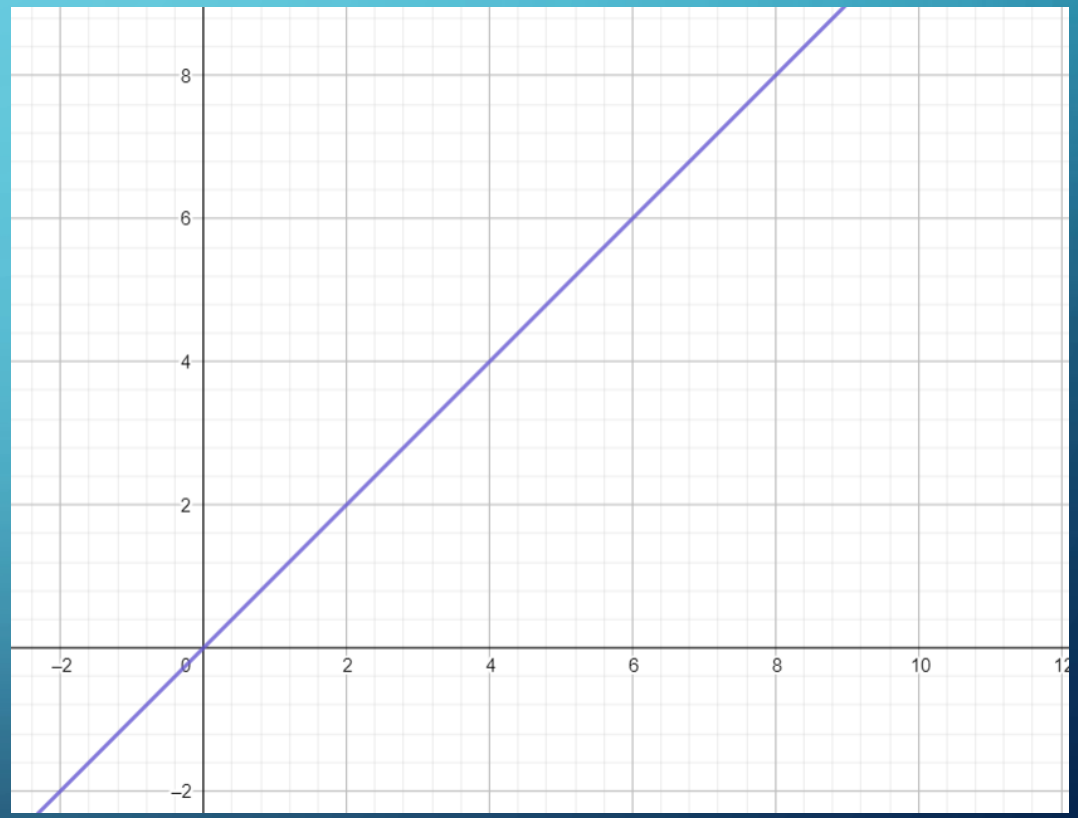
WHAT COULD BE THE DIRECTION OF A RELATIONSHIP?

DIRECTION OF A RELATIONSHIP

- It means in which direction is the relationship going ☺
- Direction can also be referred as correlation
- We have two types of directions
  - Positive direction
  - Negative direction

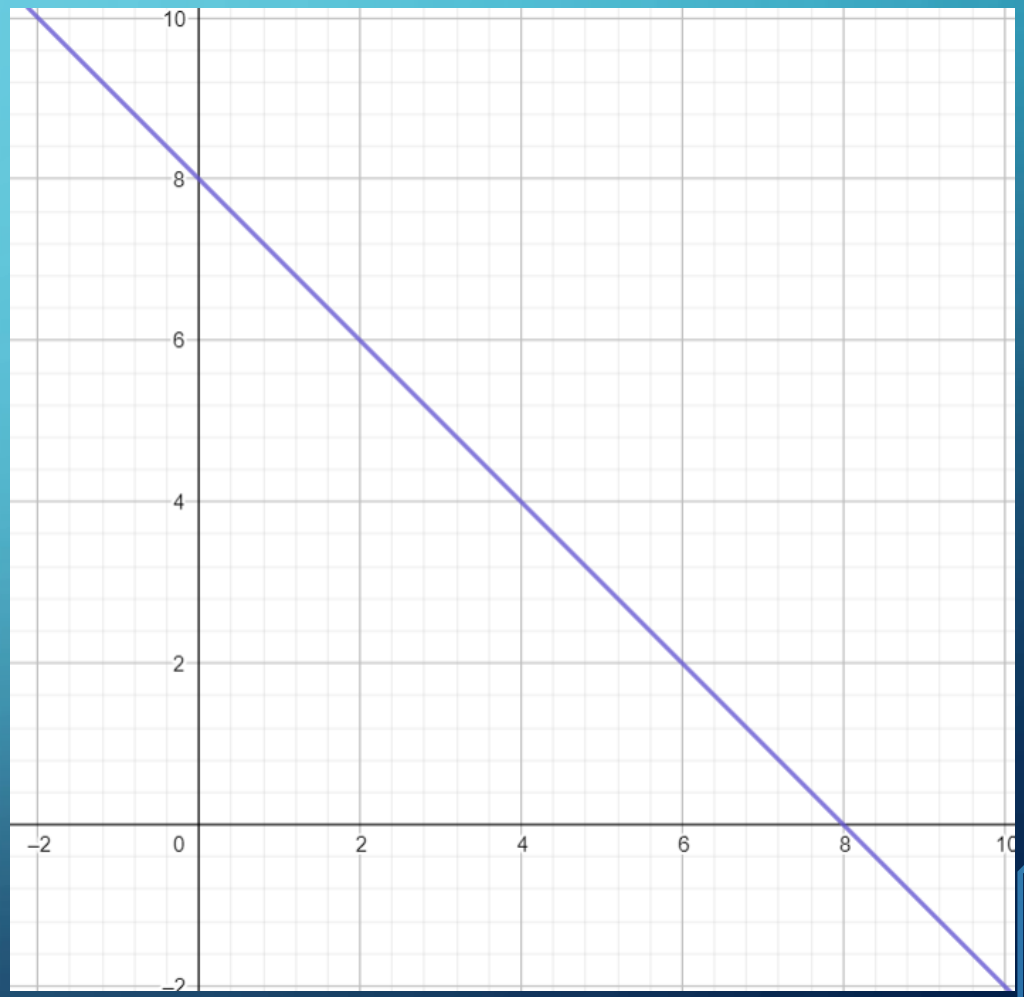
POSITIVE DIRECTION

- A positive direction means that both interests increase or decrease in the same direction
- For example the value of **x** is **increasing** ↑ and the value of **y** is also **increasing** ↑



NEGATIVE DIRECTION

- A negative direction means that both interests increase or decrease in opposite direction
- For example, if the value of **x increases** ↑ then the value of **y** will **decrease** ↓



WHAT IS THE DEGREE OF RELATIONSHIP?

DEGREE OF A RELATIONSHIP

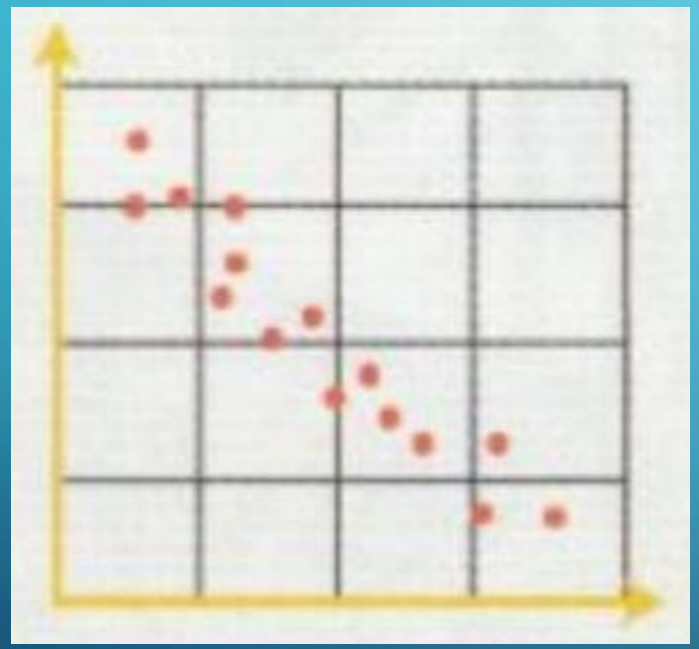
- In addition to the direction of relationship we also have to consider the degree of it
- We have 3 degrees or levels of a relationship:
  - Strong
  - Moderate
  - Weak

DEGREE OF A RELATIONSHIP

- Strong relationship – the data is aligned with each other
- Moderate relationship – the data isn't aligned perfectly
- Weak relationship – the data is not aligned at all, the data is placed everywhere

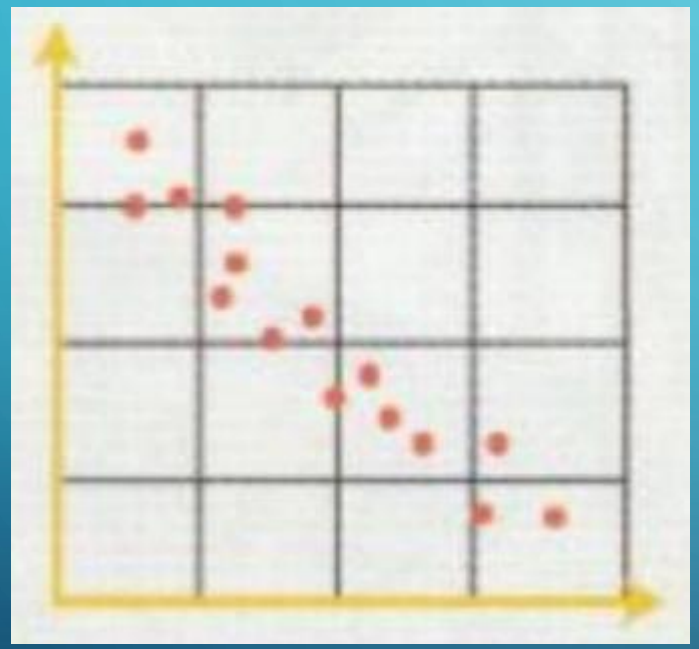
EXAMPLES

WHAT KIND OF RELATIONSHIP IS THIS ONE?



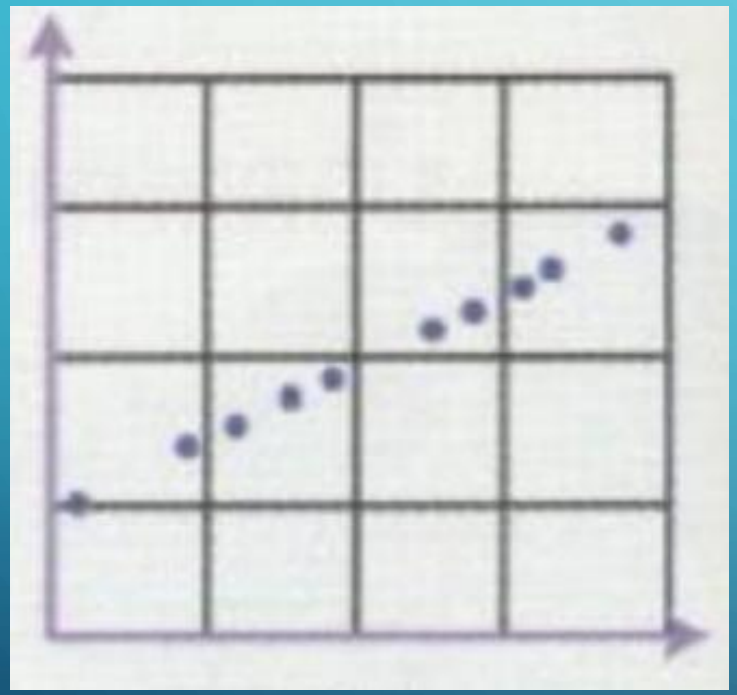
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



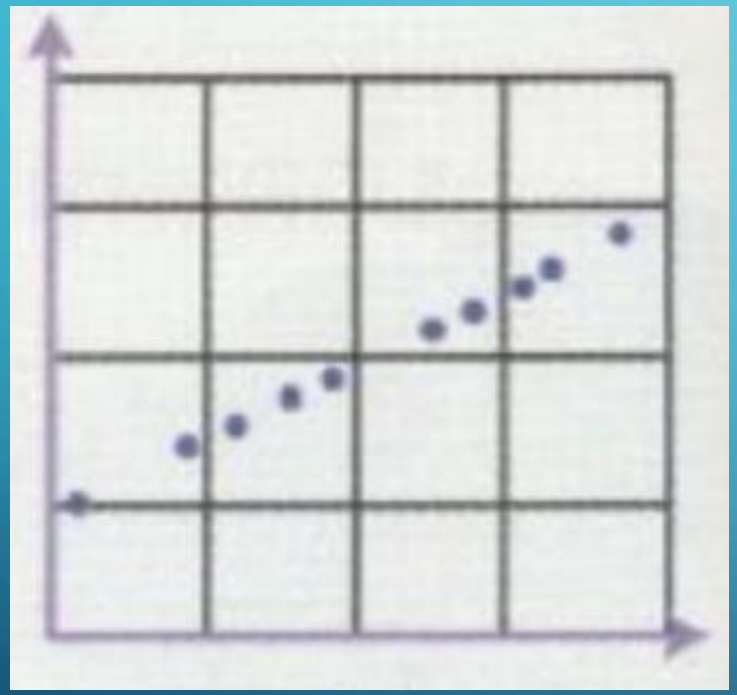
- Direction: Negative
- Degree: Moderate

WHAT KIND OF RELATIONSHIP IS THIS ONE?



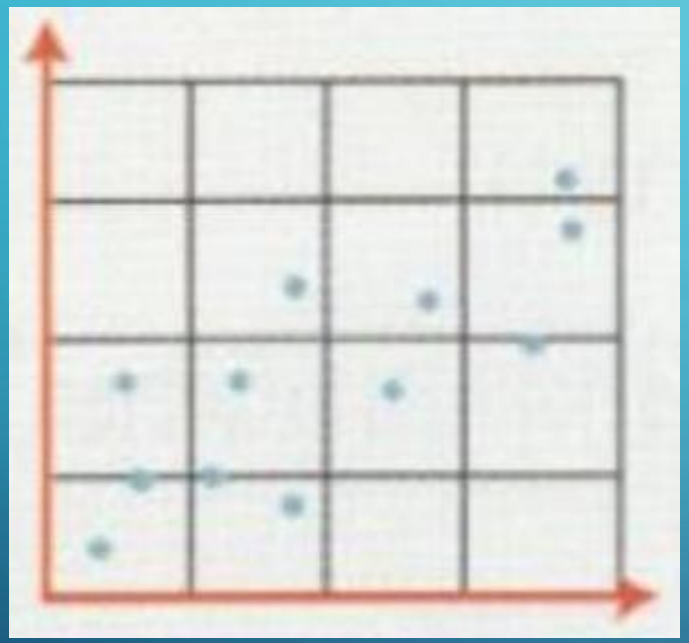
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



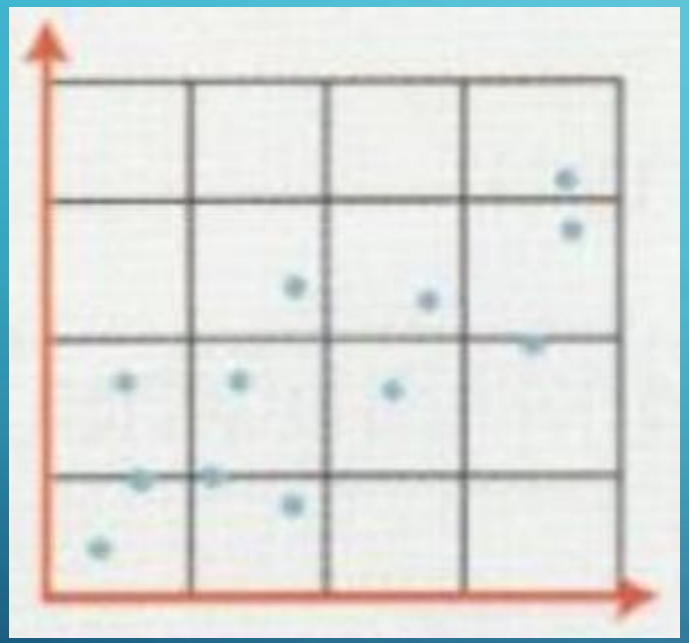
- Direction: Positive
- Degree: Strong

WHAT KIND OF RELATIONSHIP IS THIS ONE?



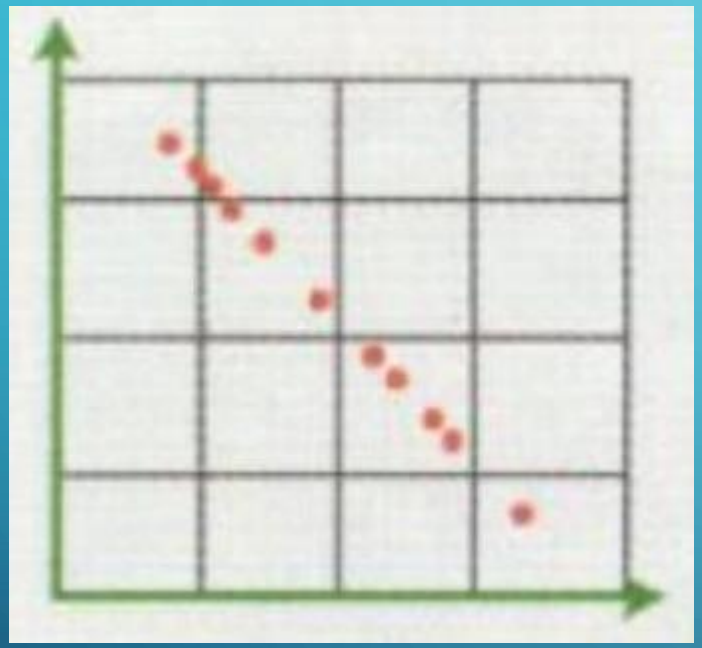
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



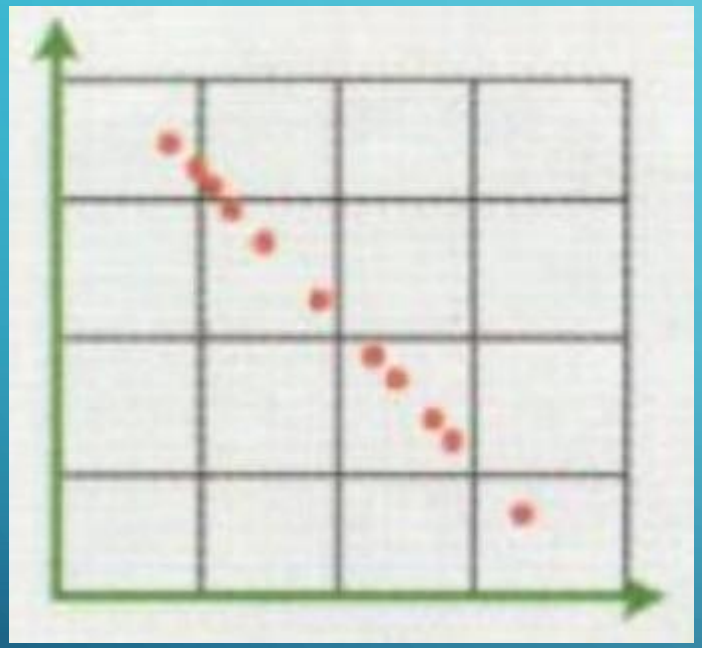
- Direction: Positive
- Degree: Weak

WHAT KIND OF RELATIONSHIP IS THIS ONE?



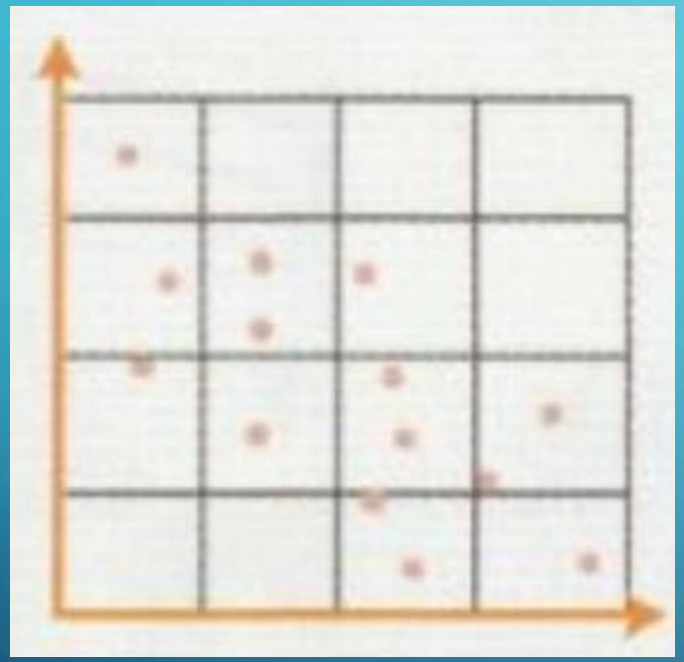
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



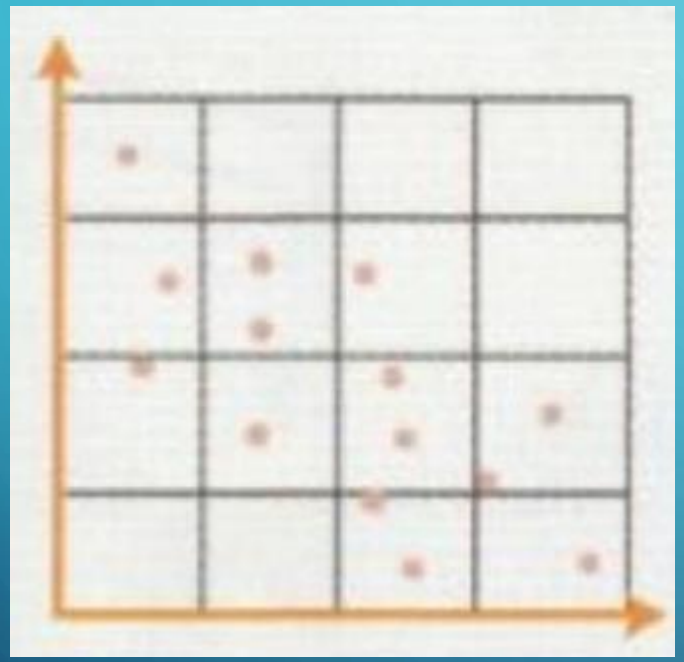
- Direction: Negative
- Degree: Strong

WHAT KIND OF RELATIONSHIP IS THIS ONE?



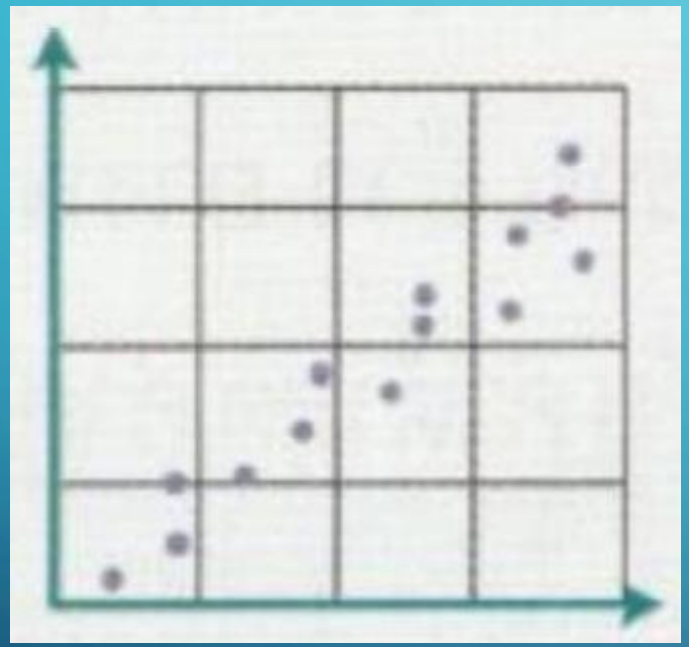
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



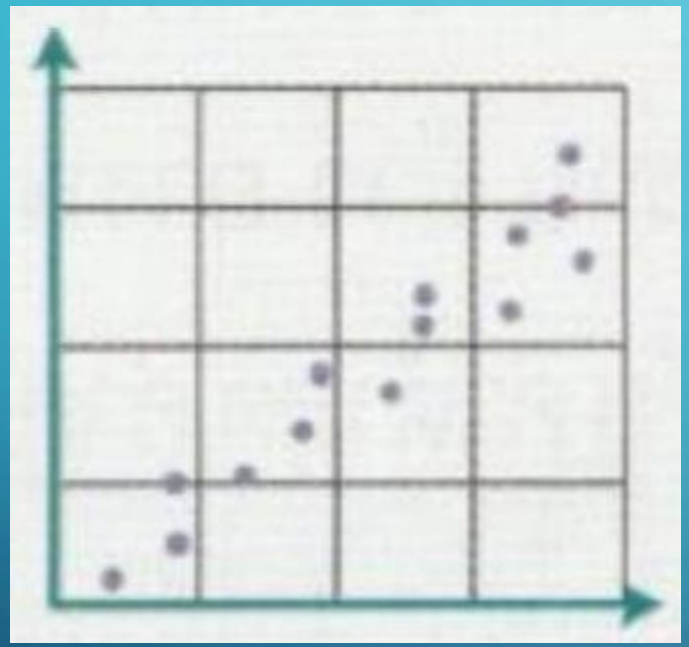
- Direction: Negative
- Degree: Weak

WHAT KIND OF RELATIONSHIP IS THIS ONE?



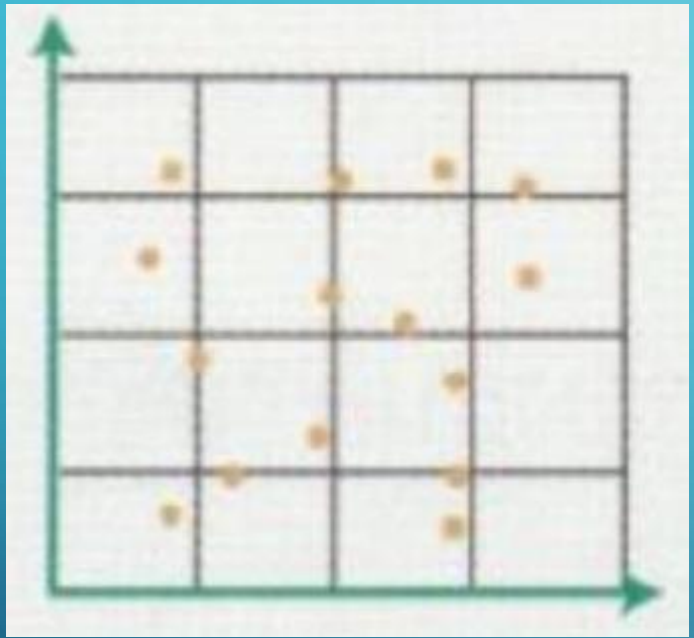
- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



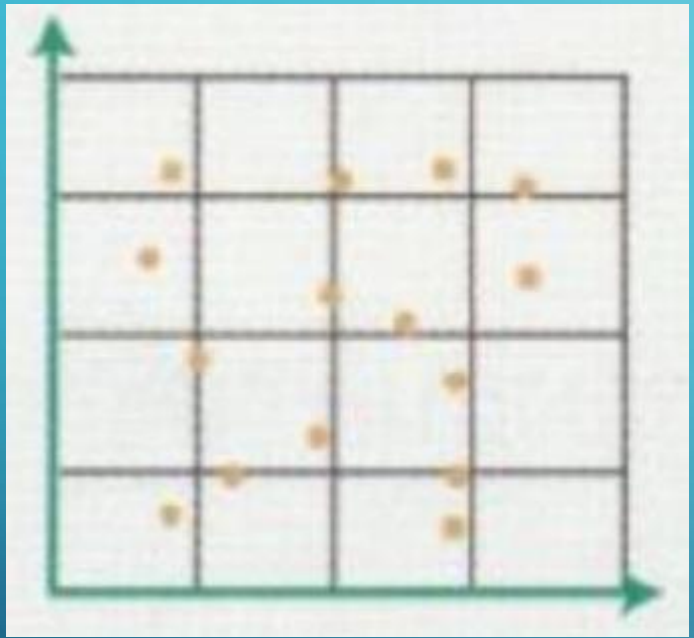
- Direction: Positive
- Degree: Moderate

WHAT KIND OF RELATIONSHIP IS THIS ONE?

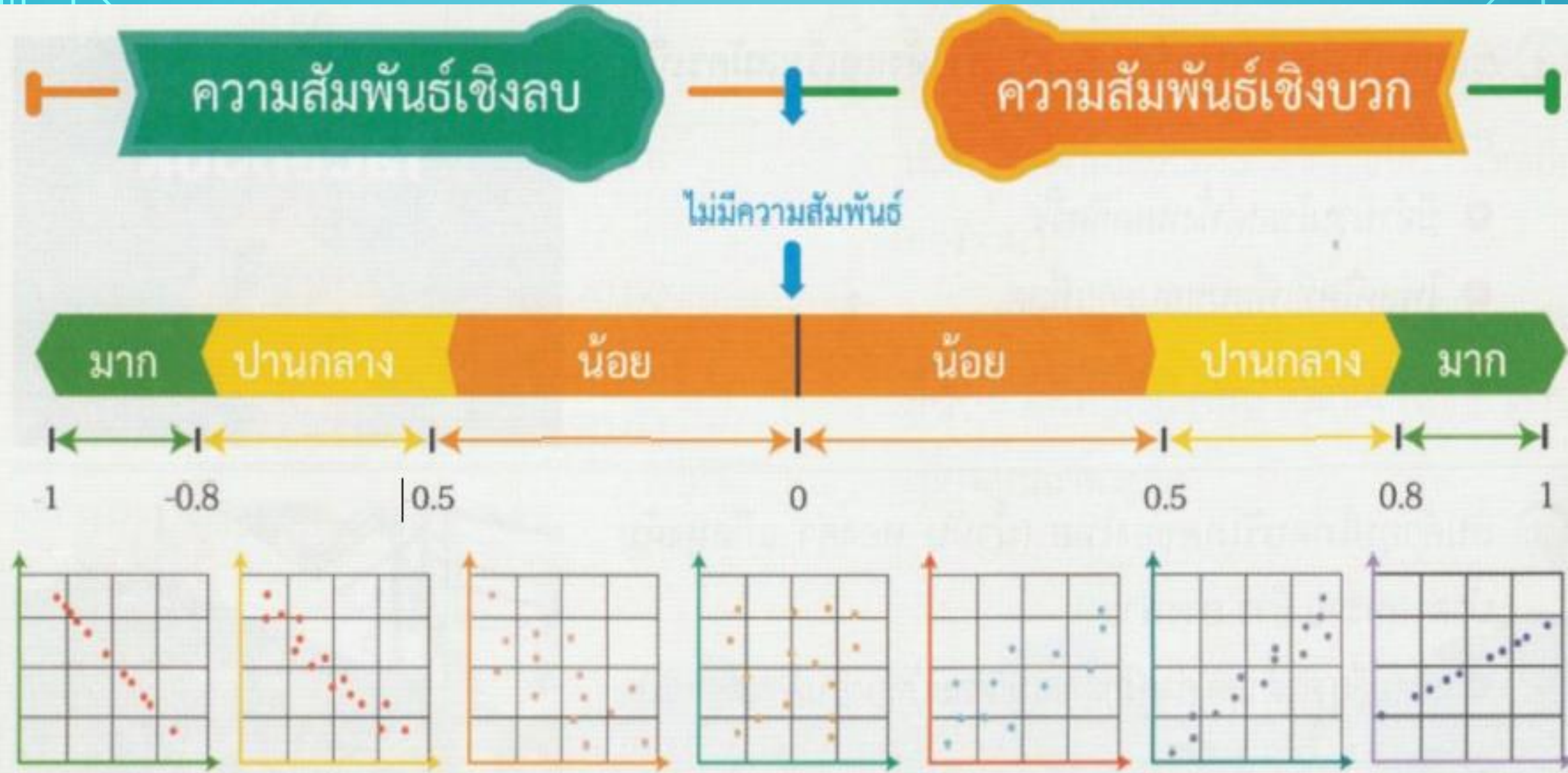


- Direction:
- Degree:

WHAT KIND OF RELATIONSHIP IS THIS ONE?



- Direction: Can't be defined
- Degree: Can't be defined
- Data is not related to each other


ความสัมพันธ์เชิงลบ
ความสัมพันธ์เชิงบวก
ไม่มีความสัมพันธ์
มาก
ปานกลาง
น้อย
น้อย
ปานกลาง
มาก
-1
-0.8
-0.5
0
0.5
0.8
1

ANY QUESTIONS?

CLASSWORK:

- Open this link:

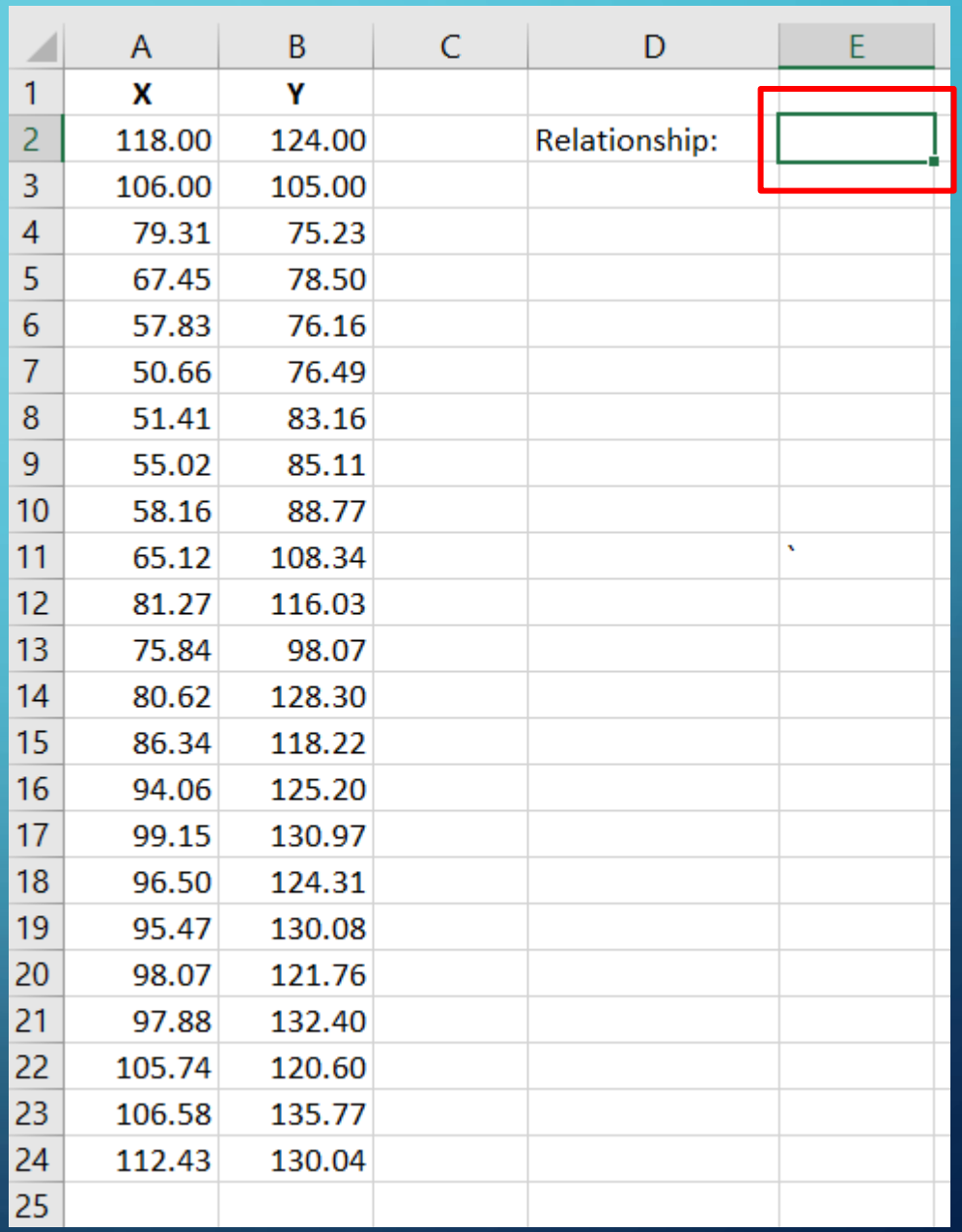
http://bit.ly/DatasetRel

TABLE:

	A	B
1	**X**	**Y**
2	118.00	124.00
3	106.00	105.00
4	79.31	75.23
5	67.45	78.50
6	57.83	76.16
7	50.66	76.49
8	51.41	83.16
9	55.02	85.11
10	58.16	88.77
11	65.12	108.34
12	81.27	116.03
13	75.84	98.07
14	80.62	128.30
15	86.34	118.22
16	94.06	125.20
17	99.15	130.97
18	96.50	124.31
19	95.47	130.08
20	98.07	121.76
21	97.88	132.40
22	105.74	120.60
23	106.58	135.77
24	112.43	130.04

STEP 1:

- Select an empty cell in the table



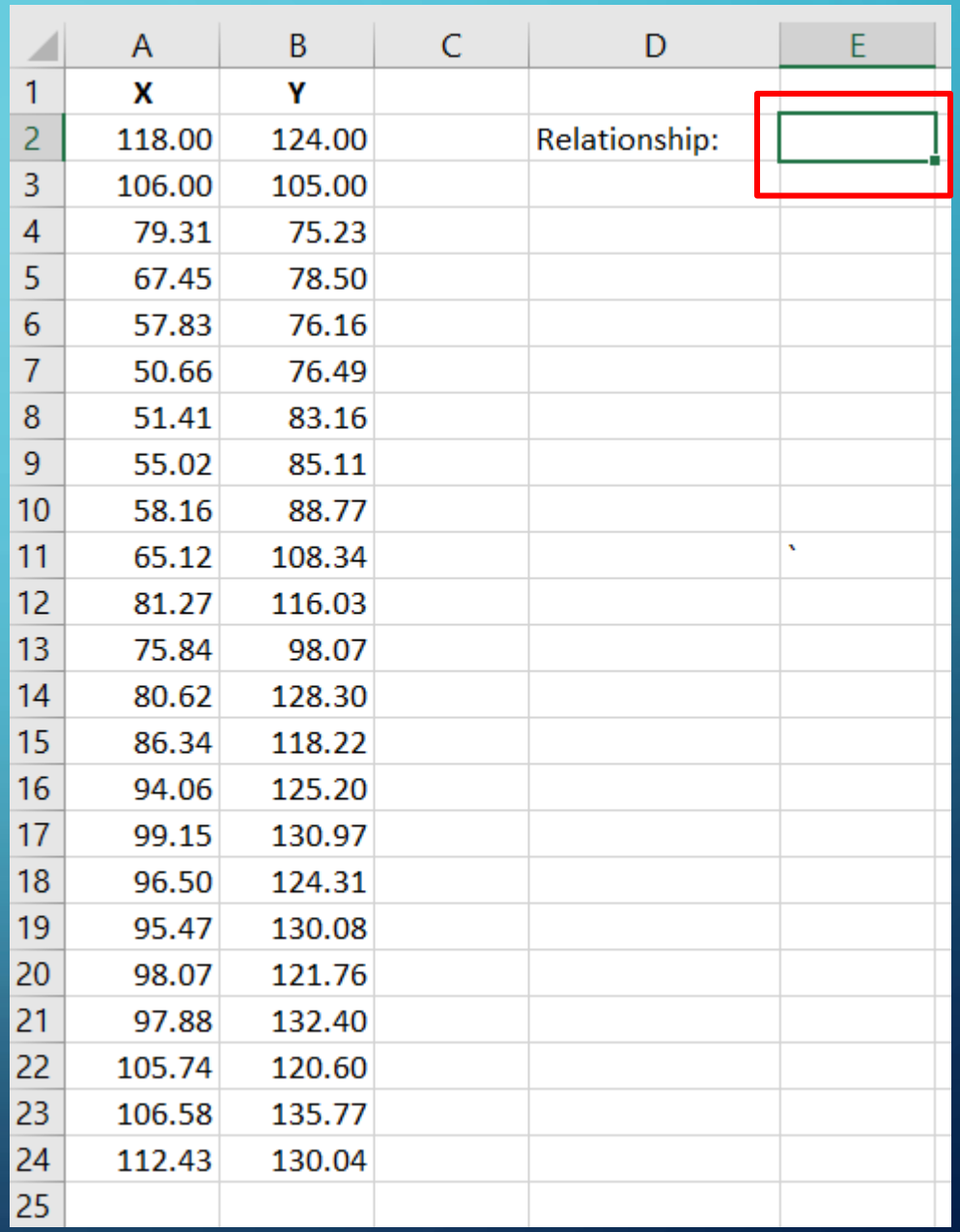
	A	B	C	D	E
1	**X**	**Y**			
2	118.00	124.00		Relationship:	
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

STEP 2:

- Write the equation:

  =CORREL(firstColumn,secondColumn)

- How can we select both columns?



	A	B	C	D	E
1	**X**	**Y**			
2	118.00	124.00		Relationship:	
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

STEP 2:

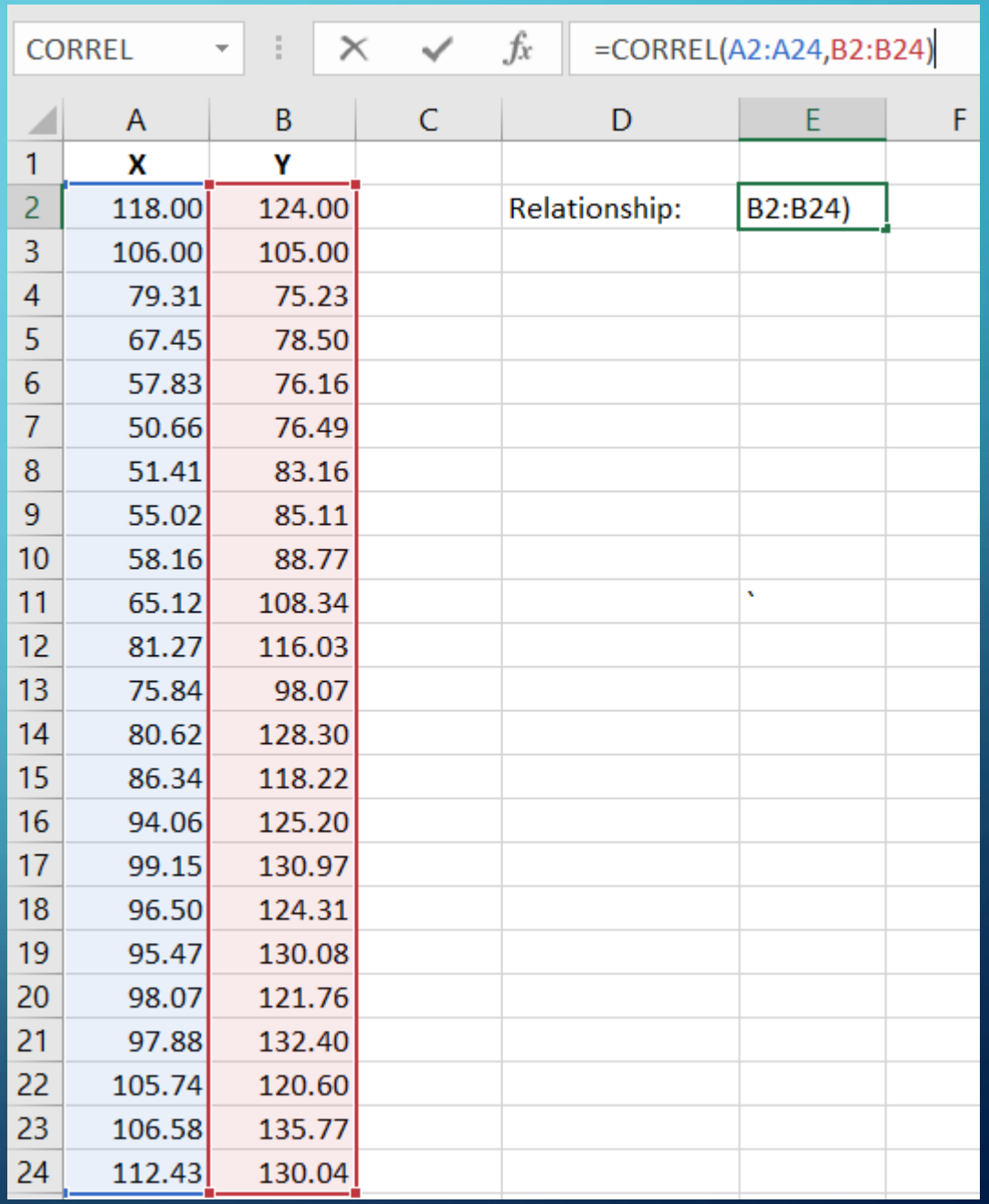
- Write the equation:

  =CORREL(A2:A24,B2:B24)

	A	B	C	D	E
1	**X**	**Y**			
2	118.00	124.00		Relationship:	
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

STEP 3:

- Press enter



	A	B	C	D	E	F
1	X	Y				
2	118.00	124.00		Relationship:	B2:B24)	
3	106.00	105.00				
4	79.31	75.23				
5	67.45	78.50				
6	57.83	76.16				
7	50.66	76.49				
8	51.41	83.16				
9	55.02	85.11				
10	58.16	88.77				
11	65.12	108.34				
12	81.27	116.03				
13	75.84	98.07				
14	80.62	128.30				
15	86.34	118.22				
16	94.06	125.20				
17	99.15	130.97				
18	96.50	124.31				
19	95.47	130.08				
20	98.07	121.76				
21	97.88	132.40				
22	105.74	120.60				
23	106.58	135.77				
24	112.43	130.04				

RESULT:

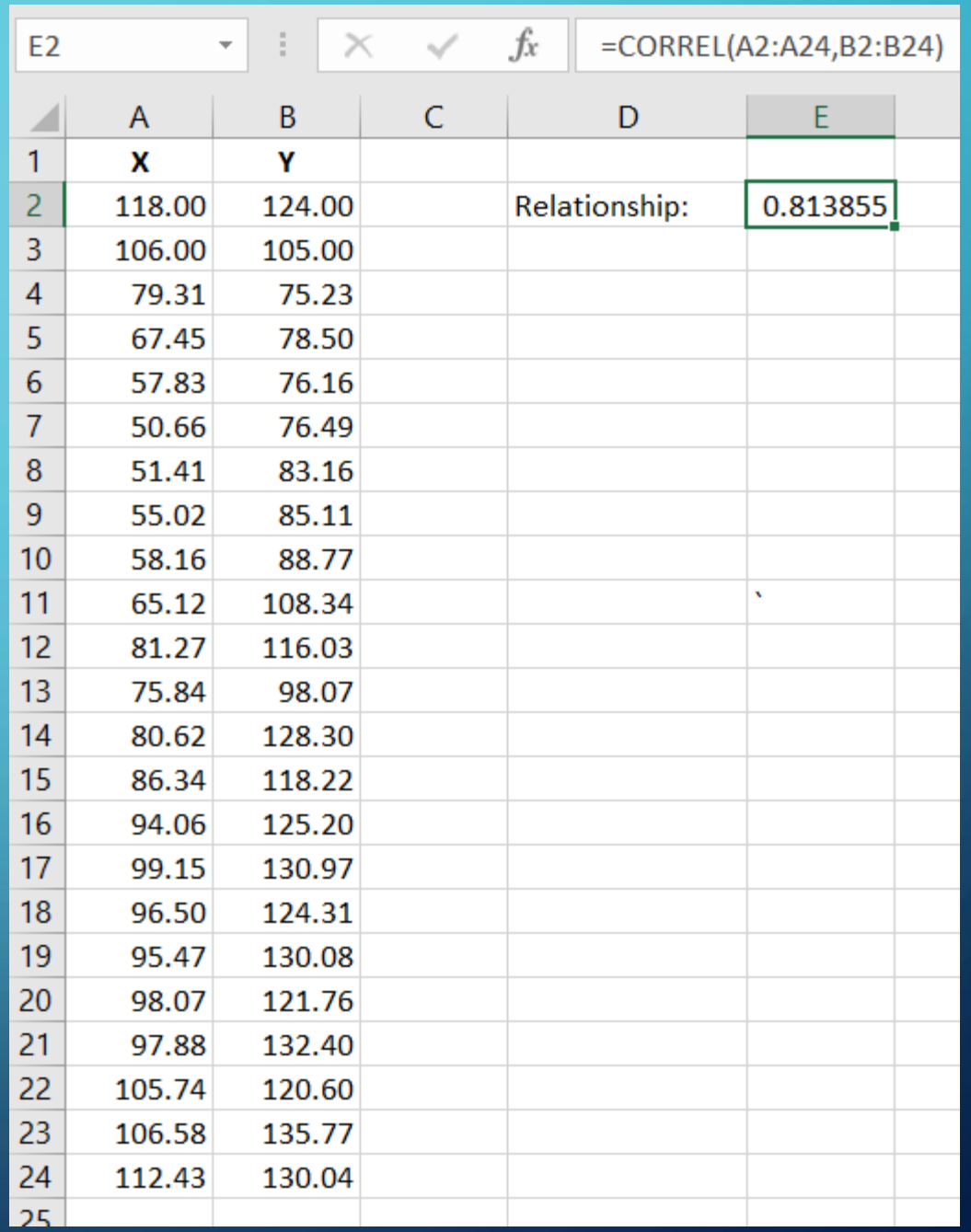
- As a result we got number 0.814
- What does that mean?

E2 =CORREL(A2:A24,B2:B24)

	A	B	C	D	E
1	**X**	**Y**			
2	118.00	124.00		Relationship:	0.813855
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

RESULT:

- As a result we got number 0.814
- What does that mean?
- What direction is our relationship?

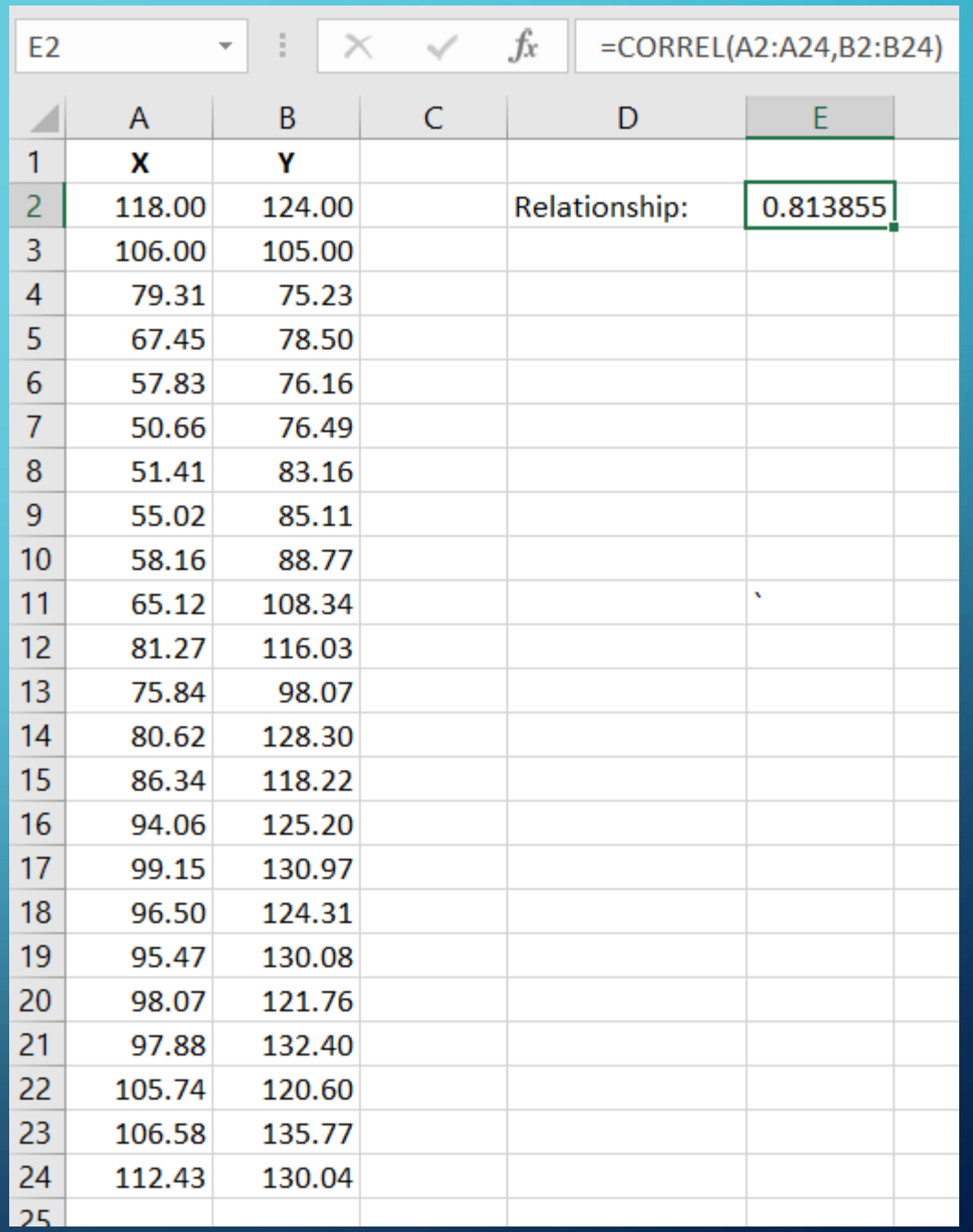


E2 =CORREL(A2:A24,B2:B24)

	A	B	C	D	E
1	X	Y			
2	118.00	124.00		Relationship:	0.813855
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

RESULT:

- As a result we got number 0.814
- What does that mean?
- What direction is our relationship?
  - Positive
- What degree is our relationship?



E2 =CORREL(A2:A24,B2:B24)

	A	B	C	D	E
1	X	Y			
2	118.00	124.00		Relationship:	0.813855
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

RESULT:

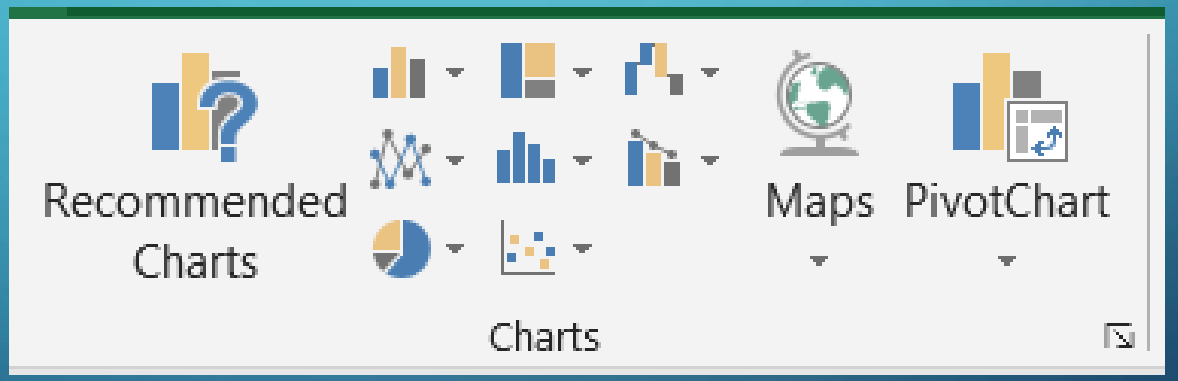
- As a result we got number 0.814
- What does that mean?
- What direction is our relationship?
  - Positive
- What degree is our relationship?
  - Strong

E2 =CORREL(A2:A24,B2:B24)

	A	B	C	D	E
1	X	Y			
2	118.00	124.00		Relationship:	0.813855
3	106.00	105.00			
4	79.31	75.23			
5	67.45	78.50			
6	57.83	76.16			
7	50.66	76.49			
8	51.41	83.16			
9	55.02	85.11			
10	58.16	88.77			
11	65.12	108.34			
12	81.27	116.03			
13	75.84	98.07			
14	80.62	128.30			
15	86.34	118.22			
16	94.06	125.20			
17	99.15	130.97			
18	96.50	124.31			
19	95.47	130.08			
20	98.07	121.76			
21	97.88	132.40			
22	105.74	120.60			
23	106.58	135.77			
24	112.43	130.04			
25					

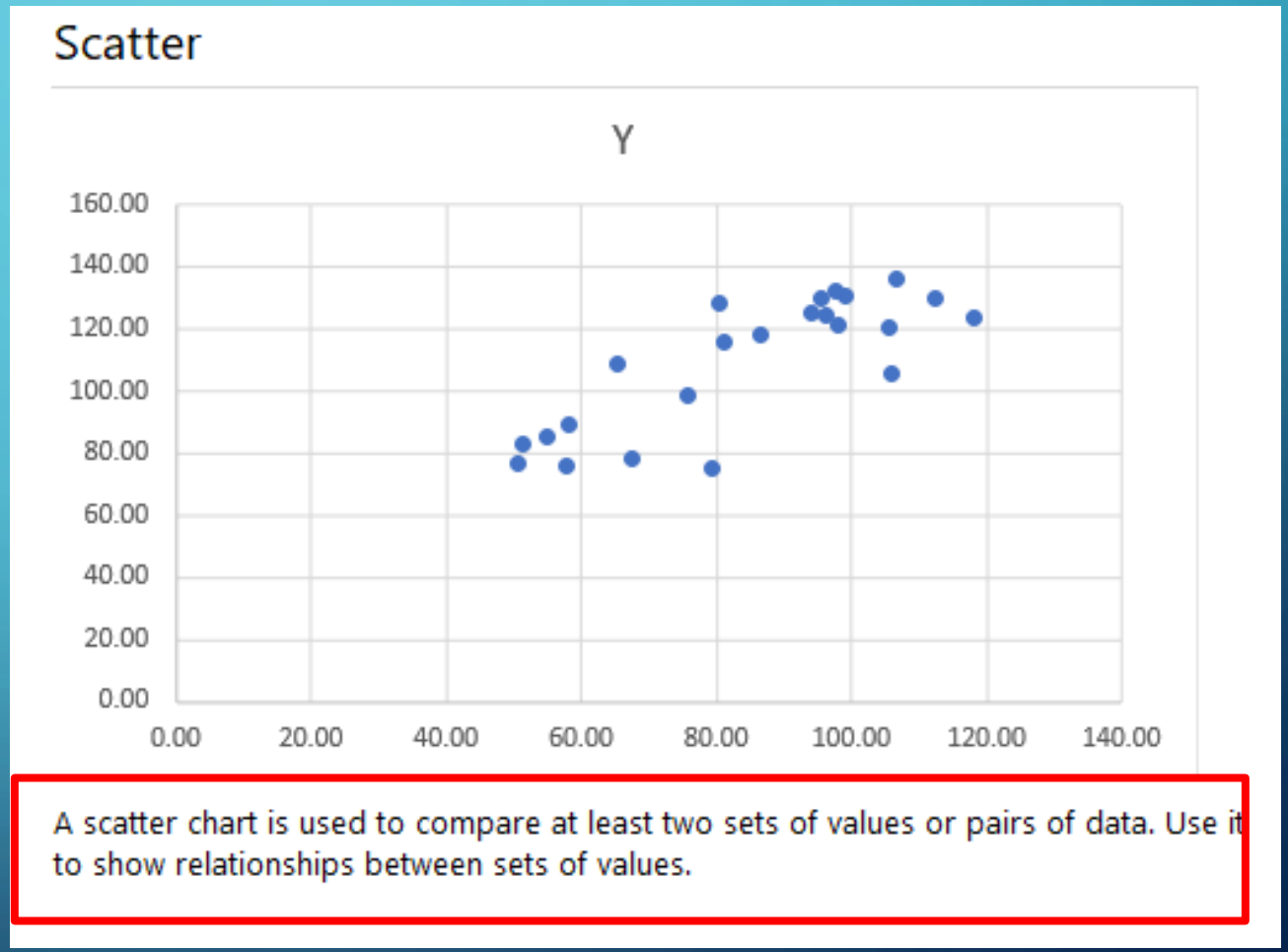
GRAPHS/CHARTS

- Which chart or graph will we use?



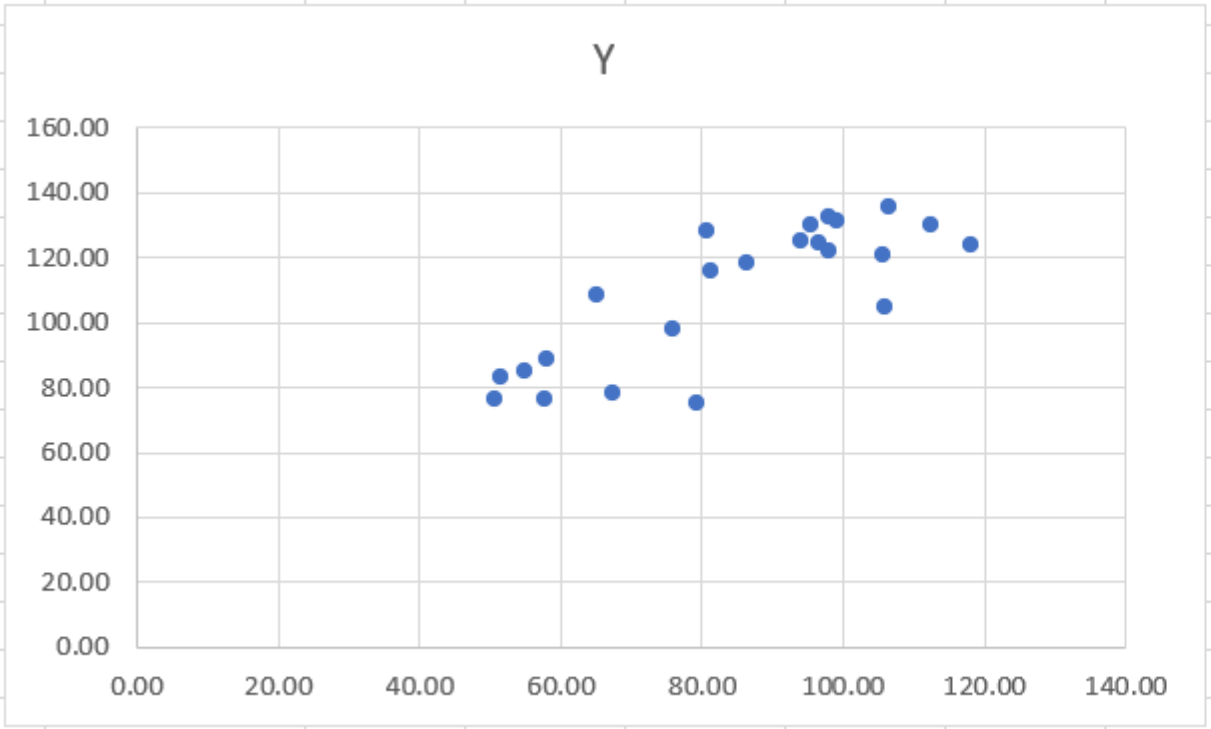
GRAPHS/CHARTS

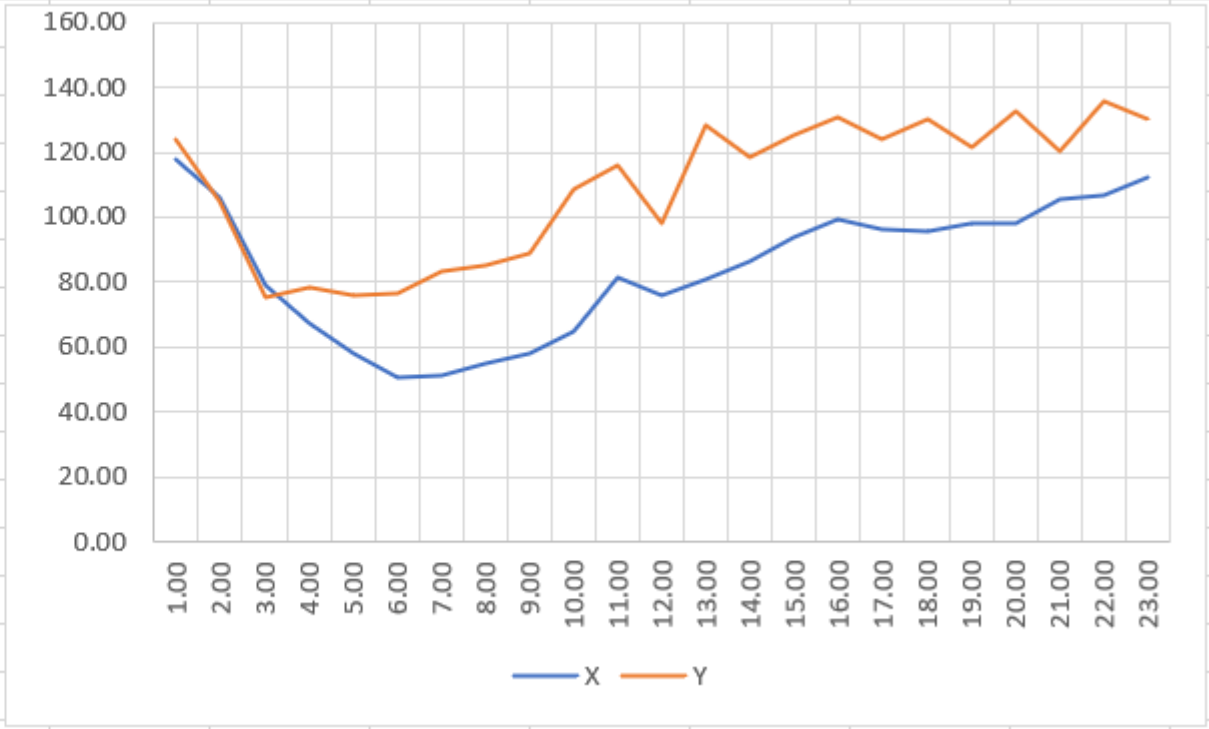
- We will use the scatter chart



GRAPHS/CHARTS

- For a better visualization we can also use a line chart





TASK2: DO IT ON YOUR OWN

- What is the relationship?

  What is the direction and degree?

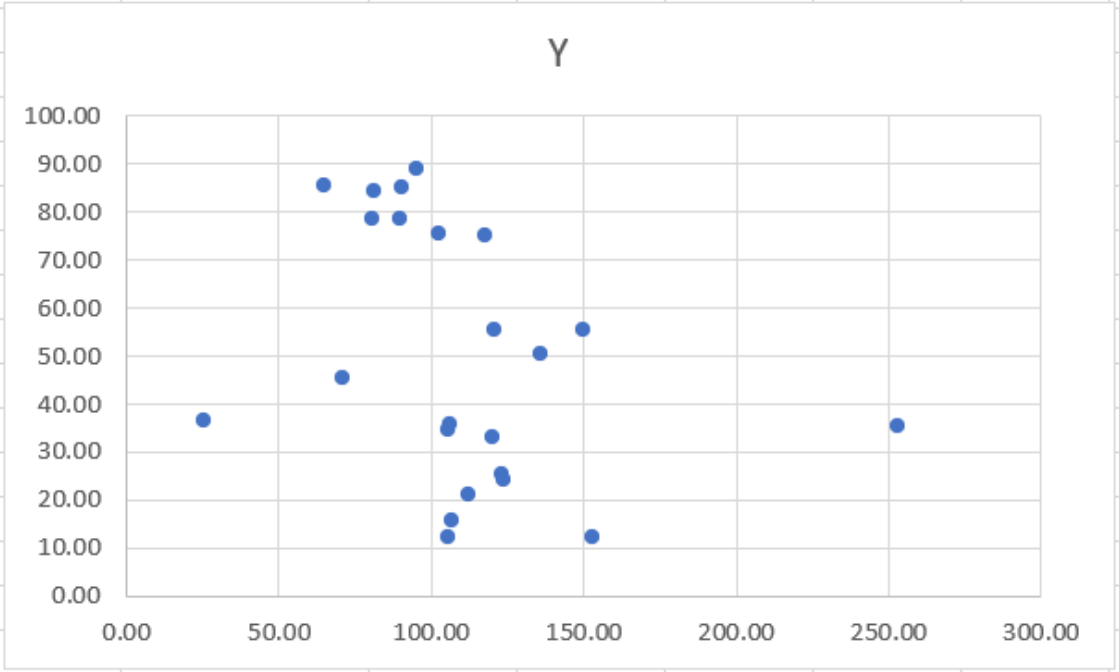
  Make both scatter and line chart.

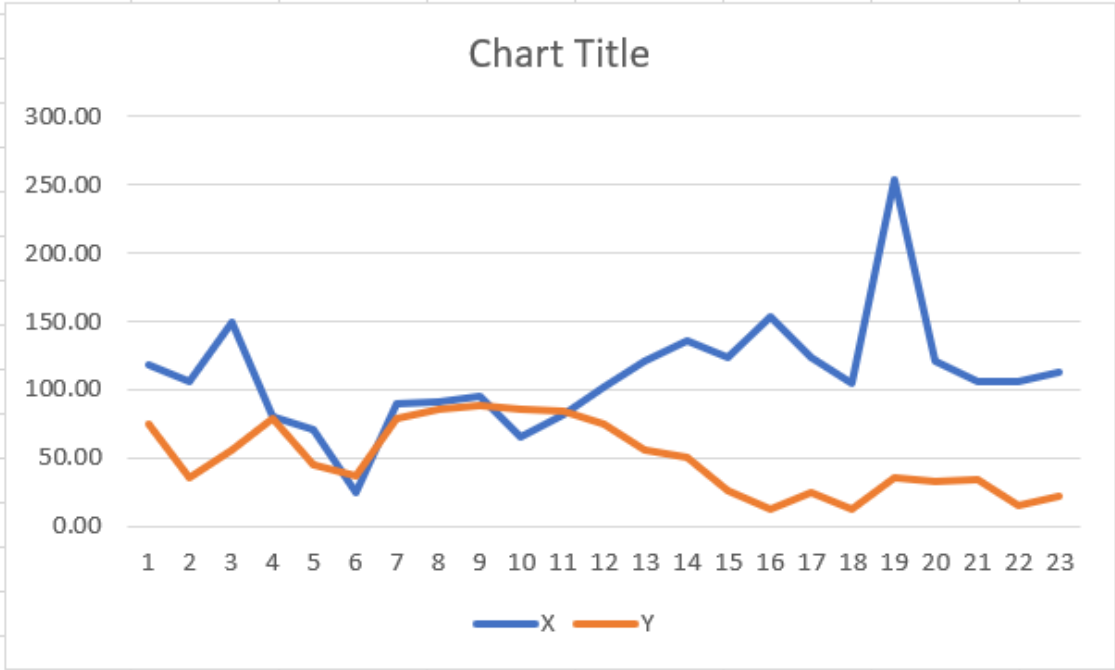
A	B	C	D
X	**Y**		
118.00	75.20		Relationship:
106.00	35.60		
150.00	55.36		
80.45	78.50		
70.83	45.53		
25.30	36.60		
90.00	78.63		
90.50	85.11		
95.12	88.77		
65.12	85.63		
81.27	84.20		
102.36	75.32		
120.50	55.60		
136.00	50.30		
123.25	25.50		
153.00	12.30		
123.60	24.30		
105.30	12.30		
253.30	35.30		
120.30	33.00		
105.74	34.50		
106.58	15.63		
112.43	21.30		

RESULT

X	Y
118.00	75.20
106.00	35.60
150.00	55.36
80.45	78.50
70.83	45.53
25.30	36.60
90.00	78.63
90.50	85.11
95.12	88.77
65.12	85.63
81.27	84.20
102.36	75.32
120.50	55.60
136.00	50.30
123.25	25.50
153.00	12.30
123.60	24.30
105.30	12.30
253.30	35.30
120.30	33.00
105.74	34.50
106.58	15.63
112.43	21.30

Relationship: -0.32235





HOMEWORK

- Worksheet on e-learning platform

THE END