VCAA "Dot Points"

Investigating data distributions, including:

• Pearson correlation coefficient, r, its calculation and interpretation

Pearson product-moment correlation coefficient (r)

Recall:

The **Pearson product-moment correlation coefficient** (r) is a measure of the linear correlation (dependence) between two variables X and Y.

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{(n - 1) \times s_x \times z_y}$$

Using the r value provides an accurate measure of both the strength and direction of a linear relationship between two numerical variables.

NB: The calculation of r is based on the following assumptions:

- Data s linearly distributed
- Data is numeric
- No outliers are present

Fortunately, the TI-Nspire CAS calculator can be used to calculate the Pearson product-moment correlation coefficient (r).

Recall:





Example 1

Use the TI-Nspire CAS calculator to calculate the Pearson product-moment correlation coefficient (r)for the ice cream sales v daily temperature data from Notes 3.1.12

ALC:	harmen	Burstein	6	15	12
=	oering.	0.99059	10		
1	14.2	215			
2	16.4	325			
3	11.9	185			
4	15.2	332			
8	18.5	406			

5	1: Actions	 Date 	Ψ.		912 (1) 🗵
1	2: Insert 3: Data	ales	100	. R.	
8	4: Statistics	 B1 	15RK C	alculation	
略:	5: Table	1 2	Distrit	utions	
72	16.4	3	Confid	lence Inter ests	vals 🕨
3	11.9	18	5		
4	15.2	. 33	2		
n.	18.5	40	6		
c.7	a score		-	10	4 3

Step.3 Menu/4/1/3	
1: One-Variable Statistics 2: Two-Variable Statistics	R+D 💮 🖬
2: Linear Regression (mod.). 4: Linear Regression (a+bx) 5: Median-Median Line 6: Quadratic Regression 7: Cubic Regression 8: Quartic Regression 9: Power Regression	ns e intervals



X List:	'sales	ĥ
V List	Temp	
Save RegEqn to:	13	1
Frequency List:	1) D	Ē
Category List:	- E	ľ
Include Categories;		

	1 12 1.	3 1 1	Dates	AND 📢
4	Esales	Č.	D	K.
=	-			=LinRegh
2	325		RegEan	m*x+b
3	185		m	0,030471
ŧ.	332		b	6.41281
5	406		1 ^d	0,916819
П.	522		i i	0.957507

r = 0.958 (3 decimal places)

: a strong, positive, linear association exists between the ice cream sales and the daily temperature.

Example 2

Use the TI-Nspire CAS calculator to calculate the Pearson product-moment correlation coefficient (r)for the driver accidents v driver age data from Notes 3.1.12

1.3	1.4 1.5	I ®ec∵p	PAG -	6
010	ige 8	accide 9	D	1
=				1
1	16	25		1
2	17	21		
3	18	19		
4	19	15		
芳.	20	15		

1:1	Actions		¶Dac 💬	A	AD 🕼
H 2:1	nsert. Data	1	cide 🤤	р.	
X 4: :	Statistics		1: Stat 1	Calculations	- 10
5:	able		2: Distri	butions	. *
2	17		3: Confi 4: Stat	idence Interv Tests	als •
2	18		19		
4	19		15		
5	20		15		

Linear Regression in	te+b)
X List:	bge 🚺
Y List:	accidents
Save RegEqn to:	14
Frequency Lists	1
Category List:	
Include Categories:	100

Step.5 Read answer

4	B accideC	P	E.
=			*LinRegh
2	21	RegEqn	m*x+b
3	19	m	-1.52121
4	15	b	46.6849
5	15	ra.	0.89841
H.	14	ir	-0.9478

Step.3 Menu/4/1/3

1: Actions 1: Stores	PM2 (11 🕅 🕅
1: One-Variable Statistics 2: Two-Variable Statistics	0
2: Linear Regression (meth) 4: Linear Regression (a+bx) 5: Median-Median Line 6: Quadratic Regression 7: Cubic Regression	ns e Intervals
8: Quartic Regression 9: Power Regression A:Exponential Regression	

r = -0.948 (3 decimal places)

: a strong, negative, linear association exists between the ice cream sales and the daily temperature.



Example 3

The following data was collected to investigate the association between the number of hours studied by students and the scores achieved upon a maths test.

Study	Score		RAD XII 🐼
(hours)	(Percentage)		LUND Y
0	50)
0	62	95	0
1	47		
1	54		0
2	67	vi 80 -	
2	72		
3	69		
3	78	°° 65 ● ●	
4	75		
4	86		
5	81	50 - •	
5	87		
6	88		7 0
6	94	0123450	/ 0
7	85	study	
7	98		

Task.1

8

Use your TI-Nspire CAS calculator to calculate the Pearson product-moment correlation coefficient (r) between the two variables.

4 E g	scores	ċ	D	E
=				=LinRegM
2	62		RegEqn	m*x+b
3	47		m	5,70605
4	54	<u>.</u>	b	54.2831
5	67		ra	0.841596
6	72		r	0.917385
EZ ="	Linear R	egrensi	on (mx+b)"	4)

95

Task.2

Use the r value to classify the association between the two variables.

r = 0.917 (3 decimal places)

 \therefore a strong, positive, linear association exists between the maths scores and the hours studied.



Example 4

The following data was collected to investigate the association between the number of siblings a student has and the scores achieved upon a maths test.

No of	Score		17	10	10	*0			R4	n 📶 👽
siblings	(Percentage)		1./	1.0	1.9	۰D	0C 🗢		150	ະ ນັ 🔽
0	98									
0	64			🔍						0
0	88		00	1						
1	72		90 -						0	
1	63	a	ر	ΙT					0	
1	59	- Lo	5 ·	1			0	•	ŏ	0
2	80	S	2 75 -						•	-
2	69	Ĩ	· .		C		8			
2	72						-			
3	64		60	🛡	Q	2		0		
3	82				C					
3	55			1				0		
4	90			00	<u> </u>	~	20	20	10	
4	78			0.0		0	2.0	3.0	4.0	5.0
4	83						sibli	ings		
5	79									
5	99									

Task.1

Use your TI-Nspire CAS calculator to calculate the Pearson product-moment correlation coefficient (r) between the two variables.

• S :	score	¢.	D	E.
=				=LinRegM
2	64		RegEqn	m*x+b
з	88		m	1.87308
4	72		b	71.7692
5	63		Ls.	0.058869
fő.	59		r	0.242629

Task.2

Use the r value to classify the association between the two variables.

r = 0.243 (3 decimal places)

 \therefore no association exists between the maths scores and the number of siblings

