

Regressione

Soluzioni

1. (a) Il diagramma di dispersione è riportato in Figura 1.

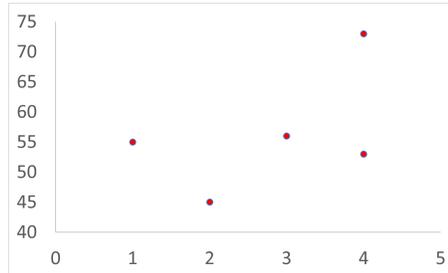


Figura 1: Diagramma di dispersione.

- (b) Per individuare la retta di regressione si svolgono i calcoli sotto riportati dopo aver individuato $\bar{x} = 2,8$ e $\bar{y} = 56,4$.

X	Y	$x_i - \bar{x}$	$y_i - \bar{y}$	Prodotto	$(x_i - \bar{x})^2$
1	55	-1,8	-1,4	2,52	3,24
4	53	1,2	-3,4	-4,08	1,44
3	56	0,2	-0,4	-0,08	0,04
2	45	-0,8	-11,4	9,12	0,64
4	73	1,2	16,6	19,92	1,44
Tot				27,4	6,8

Dunque

$$\hat{\beta}_1 = \frac{27,4}{6,8} = 4,03$$

e

$$\hat{\beta}_0 = 56,4 - 4,03 \cdot 2,8 = 45,12$$

L'equazione della retta di regressione è dunque

$$\hat{Y} = 45,12 + 4,03X$$

La retta di regressione è riportata in Figura 2.

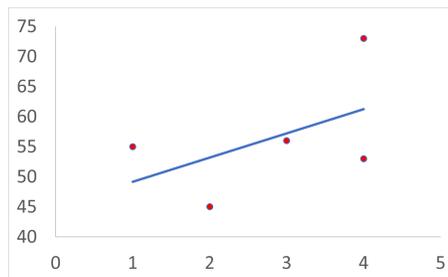


Figura 2: Diagramma di dispersione e retta di regressione.

- (c) Poiché $\hat{\beta}_1 > 0$, tra X e Y esiste concordanza (ovvero una relazione diretta). All'aumentare del numero di consorzi, ci si aspetta un aumento della percentuale di terreni coltivati con metodi organici.

(d) Per calcolare l'indice R^2 bisogna disporre di SQE (la somma dei quadrati dei residui) e di SQT (la somma dei quadrati degli scarti dalla media della variabile Y).

X	Y	$y_i - \bar{y}$	\hat{y}_i	e_i	e_i^2	$(y - \bar{y})^2$
1	55	-1,4	49,147	5,853	34,257	1,96
4	53	-3,4	61,235	-8,235	67,820	11,56
3	56	-0,4	57,206	-1,206	1,454	0,16
2	45	-11,4	53,176	-8,176	66,855	129,96
4	73	16,6	61,235	11,765	138,408	275,56
Tot					308,794	419,200

Dunque

$$R_{XY}^2 = 1 - \frac{SQE}{SQT} = 1 - \frac{308,794}{419,200} = 0,263$$

La bontà di adattamento è bassa, il 26% delle variabilità della percentuale di terreni coltivati con metodi organici è dovuta alla variabilità del numero di consorzi finalizzati alla diffusione delle colture basate su metodi organici.

2. (a) Il diagramma di dispersione è riportato in Figura 3.

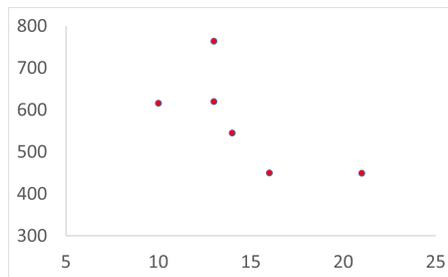


Figura 3: Diagramma di dispersione.

(b) Per individuare la retta di regressione si svolgono i calcoli sotto riportati dopo aver individuato $\bar{x} = 14,5$ e $\bar{y} = 574$.

X	Y	$x_i - \bar{x}$	$y_i - \bar{y}$	Prodotto	$(x_i - \bar{x})^2$
10	616	-4,5	42	-189	20,25
13	620	-1,5	46	-69	2,25
14	545	-0,5	-29	14,5	0,25
16	450	1,5	-124	-186	2,25
13	764	-1,5	190	-285	2,25
21	449	6,5	-125	-812,5	42,25
Tot				-1527	69,5

Dunque

$$\hat{\beta}_1 = \frac{-1527}{69,5} = -21,97$$

e

$$\hat{\beta}_0 = 574 - (-21,97) \cdot 14,5 = 892,58$$

L'equazione della retta di regressione è dunque

$$\hat{Y} = 892,58 - 21,97X$$

La retta di regressione è riportata in Figura 4.

(c) Poiché $\hat{\beta}_1 < 0$, tra X e Y esiste discordanza (ovvero una relazione inversa). All'aumentare del numero di articoli scandalistici, ci si aspetta una riduzione del numero di copie vendute.

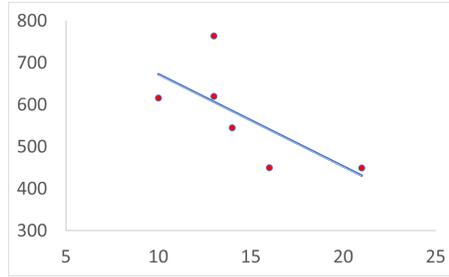


Figura 4: Diagramma di dispersione e retta di regressione.

(d) Con un numero di articoli scandalistici pari a 11, le vendite attese sono

$$\hat{Y}(11) = 892,58 - 21,97 \cdot 11 = 650,91$$