

Ispituje se težina suve materije paradajza uzgajanom na tlu sa različitom količinom đubriva:

Količina đubriva	12	13	15	17	18	19	21	22	23	25
Težina suve materije	2.9	3.1	4.7	4.5	4.7	5.3	6.1	6.5	6.2	6.8

- Nacrtati dijagram rasturanja.
- Odrediti jednačinu linearne regresije, odnosno parametre **a** i **b** iz jednačine (parametri regresije).
- Ucrtati liniju regresije na dijagramu na osnovu dobijene jednačine regresije.
- Odrediti standardnu grešku regresije i vrijednost koeficijenata korelacije i determinacije.
- Odrediti očekivanu težinu suve materije ako je količina đubriva 16?

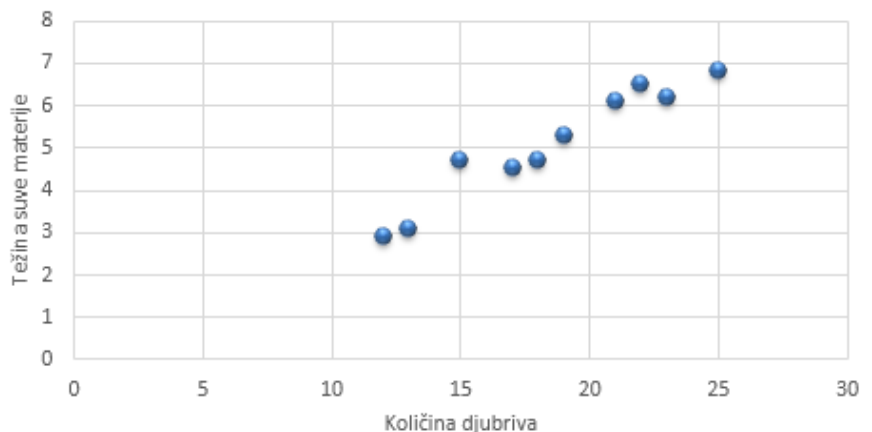
За решавање задатка потребно је обавити следеће:

Формирамо табелу у коју уносимо вриједности независне промјенљиве X (у овом случају је то количина ђубрива) и зависне промјенљиве Y (оvdје је то тежина суве материје).

i	X_i	Y_i
1	12	2,9
2	13	3,1
3	15	4,7
4	17	4,5
5	18	4,7
6	19	5,3
7	21	6,1
8	22	6,5
9	23	6,2
10	25	6,8

За дијаграм растурања у координатној равни уносимо тачке :
 $(x_1; y_1) = (12; 2,9)$ тако што на X -оси означимо 12, на Y -оси 2,9 и у пресјеку означимо тачку.
 $(x_2; y_2) = (13; 3,1)$ тако што на X -оси означимо 13, на Y -оси 3,1 и у пресјеку означимо тачку.

Dijagram rasturanja



За једначину линеарне регресије $y = ax + b$ допуњавамо табелу са колонама у којима рачунамо квадрате независне промјенљиве X , квадрате зависне промјенљиве Y и производе промјенљивих X и Y .

i	X_i	Y_i	X_i^2	Y_i^2	$X_i * Y_i$
1	12	2,9	144	8,41	34,8
2	13	3,1	169	9,61	40,3
3	15	4,7	225	22,09	70,5
4	17	4,5	289	20,25	76,5
5	18	4,7	324	22,09	84,6
6	19	5,3	361	28,09	100,7
7	21	6,1	441	37,21	128,1
8	22	6,5	484	42,25	143
9	23	6,2	529	38,44	142,6
10	25	6,8	625	46,24	170

У следећем кораку сумирамо вриједности по колонама

Добијамо

$$\sum_{i=1}^{10} X_i = 12 + \dots + 25 = 185$$

$$\sum_{i=1}^{10} Y_i = 2,9 + \dots + 6,8 = 50,8$$

$$\sum_{i=1}^{10} X_i^2 = 3591$$

$$\sum_{i=1}^{10} Y_i^2 = 274,68$$

$$\sum_{i=1}^{10} X_i * Y_i = 991,1$$

i	X_i	Y_i	X_i^2	Y_i^2	$X_i * Y_i$
1	12	2,9	144	8,41	34,8
2	13	3,1	169	9,61	40,3
3	15	4,7	225	22,09	70,5
4	17	4,5	289	20,25	76,5
5	18	4,7	324	22,09	84,6
6	19	5,3	361	28,09	100,7
7	21	6,1	441	37,21	128,1
8	22	6,5	484	42,25	143
9	23	6,2	529	38,44	142,6
10	25	6,8	625	46,24	170
	185	50,8	3591	274,68	991,1

У следећем кораку да би добили аритметичке средине колона добијене збирове дијелимо бројем тачака

Добијамо

$$\bar{X} = \frac{\sum_{i=1}^{10} X_i}{10} = 18,5$$

$$\bar{Y} = \frac{\sum_{i=1}^{10} Y_i}{10} = 5,08$$

$$\overline{X^2} = \frac{\sum_{i=1}^{10} X_i^2}{10} = 359,1$$

$$\overline{Y^2} = \frac{\sum_{i=1}^{10} Y_i^2}{10} = 27,468$$

$$\overline{X * Y} = \frac{\sum_{i=1}^{10} X_i * Y_i}{10} = 99,11$$

i	X_i	Y_i	X_i^2	Y_i^2	$X_i * Y_i$
1	12	2,9	144	8,41	34,8
2	13	3,1	169	9,61	40,3
3	15	4,7	225	22,09	70,5
4	17	4,5	289	20,25	76,5
5	18	4,7	324	22,09	84,6
6	19	5,3	361	28,09	100,7
7	21	6,1	441	37,21	128,1
8	22	6,5	484	42,25	143
9	23	6,2	529	38,44	142,6
10	25	6,8	625	46,24	170
	185	50,8	3591	274,68	991,1
	18,5	5,08	359,1	27,468	99,11

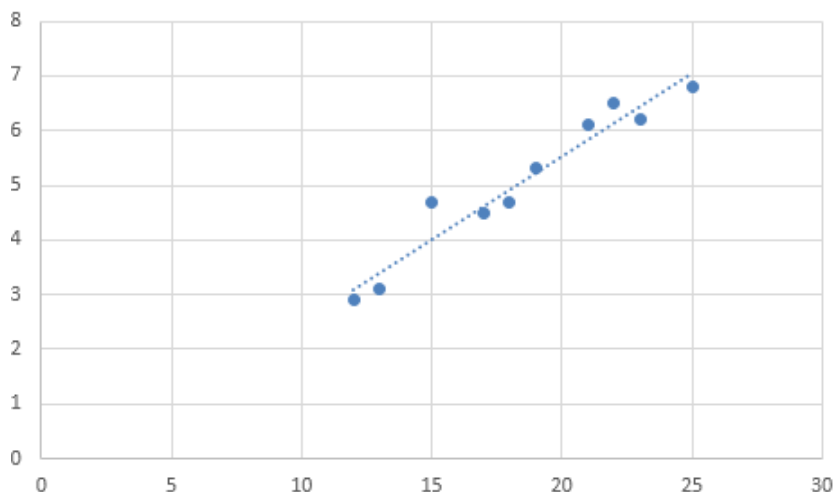
Коефицијенте a и b из једначине $y = ax + b$ одређујемо на следећи начин :

$$a = \frac{\overline{X * Y} - \overline{X} * \overline{Y}}{\overline{X^2} - \overline{X}^2} = \frac{99,1 - 18,5 * 5,08}{359,1 - 18,5^2} = 0,304$$

$$b = \overline{Y} - a * \overline{X} = 5,08 - 0,304 * 18,5 = -0,544$$

Једначина линеарне регресије је $y = ax + b = 0,304x - 0,544$

За цртање криве линеарне регресије можемо изабрати двије вриједности независне промјенљиве X и за њих израчунати вриједности зависне промјенљиве Y
 $y(10) = 0,304 * 10 - 0,544 = 2,496$
 $y(20) = 0,304 * 20 - 0,544 = 5,536$
 На графику треба означити те двије тачке и кроз њих повући праву линију.



Стандардну грешку регресије одређујемо по формули $S_e = \sqrt{\frac{\sum_{i=1}^n Y_i^2 - b \sum_{i=1}^n Y_i - a \sum_{i=1}^n X_i Y_i}{n - 2}}$.

$$\text{У овом задатку је то } S_e = \sqrt{\frac{274,68 + 0,544 * 50,8 - 0,304 * 991,1}{8}} = 0,3572$$

Коефицијент корелације одређујемо по формули $r = \frac{\overline{X * Y} - \overline{X} * \overline{Y}}{\sqrt{\overline{X^2} - \overline{X}^2} \sqrt{\overline{Y^2} - \overline{Y}^2}}$.

$$\text{У овом задатку је то } r = \frac{\overline{X * Y} - \overline{X} * \overline{Y}}{\sqrt{\overline{X^2} - \overline{X}^2} \sqrt{\overline{Y^2} - \overline{Y}^2}} = \frac{99,1 - 18,5 * 5,08}{\sqrt{359,1 - 18,5^2} \sqrt{27,468 - 5,08^2}} = 0,969$$

Како је вриједност коефицијента корелације веома близу вриједности 1 то закључујемо да постоји јака зависност тежине суве материје од количине ђубрива.

Коефицијент детерминације је $r^2 = 0,969^2 = 0,939$ или $r^2 = 93,9\%$

Очекивану тежину суве материје ако је количина ђубрива 16 добијамо помоћу једначине линеарне регресије $y = ax + b = 0,304x - 0,544$ тако што за вриједност независне промјенљиве узимамо $x = 16$ и добијамо $y(16) = 0,304 * 16 - 0,544 = 4,32$

На исти начин се могу урадити задаци :

1. Ispituje se kako nivo fizičkih aktivnosti utiče na efikasnost u poslu.

Fizičke aktivnosti (h)	1	2	2	3	3	2	4	4
Efikasnos (h)	6	5,5	5	5	4,5	6	4	3,5

- Nacrtati dijagram rasturanja.
- Odrediti jednačinu linearne regresije, odnosno parametre **a** i **b** iz jednačine (parametri regresije).
- Ucrtati liniju regresije na dijagramu na osnovu dobijene jednačine regresije.
- Odrediti standardnu grešku regresije i vrijednost koeficijenata korelacije i determinacije.
- Odrediti očekivanu efikasnost na poslu ako je fizička aktivnost radnika 5h?

1. Prikupljeni su podaci o prosječnom dohodku po članu domaćinstva i potrošnji benzina za neka slučajno izabrana domaćinstva.:

Dohodak po članu domaćinstva (00 €)	2	3	3	4	5	5	6	6	8	10
Potrošnja benzina	4	3	5	6	7	8	9	12	10	14

- Nacrtati dijagram rasturanja.
- Odrediti jednačinu linearne regresije, odnosno parametre **a** i **b** iz jednačine (parametri regresije).
- Ucrtati liniju regresije na dijagramu na osnovu dobijene jednačine regresije.
- Odrediti standardnu grešku regresije i vrijednost koeficijenata korelacije i determinacije.
- Odrediti očekivanu potrošnju benzina ako je dohodak domaćinstva 200 €.