Worksheet 4 Solution

Fred Azizi

2023-10-03

1. The following data has mean income and housing for 10 cities in Florida. Values are in dollars ($) and rounded to the nearest thousand.

City	Income ($\boldsymbol{x}$)	Housing (y)
A	26	109
B	29	97
C	25	115
D	28	99
E	38	122
F	32	145
G	25	100
H	22	76
I	29	113
J	42	144

   a. Calculate the correlation coefficient between x and y. What can you conclude about the relationship between the 2 variables?

   b. Calculate the least square line.

   c. Calculate the coefficient of variation.

Answer:

a.

x_i	y_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
26	109	-3.6	12.96	-3	9	10.8
29	97	-0.6	0.36	-15	225	9
25	115	-4.6	21.16	3	9	-13.8
28	99	-1.6	2.56	-13	169	20.8
38	122	8.4	70.56	10	100	84
32	145	2.4	5.76	33	1089	79.2
25	100	-4.6	21.16	-12	144	55.2
22	76	-7.6	57.76	-36	1296	273.6
29	113	-0.6	0.36	1	1	-0.6
42	144	12.4	153.76	32	1024	396.8
$\bar{x} = 29.6$	$\bar{y} = 112$		346.4		4066	915

$$SD(x) = \sqrt{\frac{346.4}{9}} = 6.2039$$
$$SD(y) = \sqrt{\frac{4066}{9}} = 21.255$$

Correlation: $r = \frac{s_{xy}}{s_x s_y} = \frac{\frac{1}{9}\times 915}{6.2039\times 21.255} \approx 0.77$. Correlation is greater than zero and close to 1. Hence, as x increases, y increases as well.

b.

$$b_1 = \frac{s_{xy}}{s_x^2} = \frac{\frac{1}{9}\times 915}{6.2039^2} \approx 2.641491$$

$$b_0 = \bar{y} - b_1\bar{x} = 112 - 2.641491 \times 29.6 \approx 33.81$$

Least square line can be written as $\hat{y} = 33.81 + 2.64x$.

c.

coefficient of variation:

$$CV(x) = \frac{S_x}{\bar{x}} \times 100 = \frac{6.2039}{29.6} \times 100 = 20.96\%$$
$$CV(y) = \frac{S_y}{\bar{y}} \times 100 = \frac{21.255}{112} \times 100 = 18.98\%$$

2. A sample of 30 observations has a standard deviation of 4. Find the sum of squared deviations from the sample mean.

   Answer:

   We note that $n = 30$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$
$$4 = \sqrt{\frac{\sum (x_i - \bar{x})^2}{30-1}}$$
$$16 = \frac{\sum (x_i - \bar{x})^2}{29}$$
$$\rightarrow \sum (x_i - \bar{x})^2 = 16 \times 29 = 464$$

3. Following observations are given for two variables.

y	**x**
5	2
8	12
18	3
20	6
22	11
30	19
10	18
7	9

a. Compute and interpret P_{86}.

Answer:

ordered y: 5 7 8 10 18 20 22 30

$L_{86} = (8+1)\frac{86}{100} = 7.74$. Therefore, $P_{86} = 22 + (30 - 22) \times 0.74 = 27.92$

ordered x: 2 3 6 9 11 12 18 19

$L_{86} = (8+1)\frac{86}{100} = 7.74$. Therefore, $P_{86} = 18 + (19 - 18) \times 0.74 = 18.74$

b. Compute and interpret the correlation coefficient.

Answer:

Using a calculator, the correlation coefficient is approximately 0.345. This indicates a positive and moderately weak relationship between x and y.