Correlation

Ds 1a

$$r = \frac{n\Sigma xy - \Sigma x\,\Sigma y}{\sqrt{[n\times\Sigma x^2 - (\Sigma x)^2]\,[n\times\Sigma y^2 - (\Sigma y)^2}}$$

Table

X	Y	XY	X^2	Y^2
7	9	63	49	81
8	20	160	64	400
6	8	48	36	64
8	22	176	64	484
14	22	308	196	484
12	26	312	144	676
10	27	270	100	729
10	27	270	100	729
$\Sigma x = 75$	$\Sigma y = 161$	$\Sigma xy = 1607$	$\Sigma x^2 = 753$	$\Sigma y^2 = 3647$

$$r = \frac{8\times1607 - 75\times161}{\sqrt{(8\times753 - 75^2)(8\times3647 - 161^2)}}$$

$$= \frac{781}{\sqrt{1298745}}$$

$$= 0.6853$$

Ds 1b

The correlation coefficient r is positive. This implies a positive correlation exists between number of arguments and sensation seeking

The correlation coefficient is close to 1 implying the correlation is strong. Therefore, the correlation coefficient suggests that there is a positive linear relationship between Number of arguments and sensation seeking.

Ds 1c

$$r^2 = 0.6853^2$$
$$= 0.4697$$

$$0.4697 \times 100\% = 46.97\%$$

46.97% of variability in sensation seeking can be explained by number of arguments.

Ds 2 : Significance of a Correlation

Ds 2 a.

1. Hypotheses

$H_0 : \rho = 0$

$H_a : \rho \neq 0$

$\alpha = 0.05$

2. Test-statistic

$$t = \frac{0.6853 \times \sqrt{8-2}}{\sqrt{1-0.6853^2}}$$

$$= 2.305$$

3. P-value

$df = 8 - 2 = 6$

$$P\text{-value} = P(|t| > 2.305)$$

$$= 0.0607$$

4. Conclusion.

Since the p-value (0.0607) is greater than $\alpha = 0.05$, we fail to reject the null hypothesis. It is concluded that there is not enough evidence to claim that the population correlation ρ is different than 0.

Qs 2b

A type 2 error. Since we failed to reject H_0, a type 2 error is a possibility

Qs3 : Regression

Regression equation : $\hat{y} = bx + a$

Where : $$b = \frac{n \times \Sigma xy - \Sigma x \times \Sigma y}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$a = \frac{\Sigma y - b\Sigma x}{n}$$

Thus,

$$b = \frac{8 \times 1607 - 75 \times 161}{8 \times 753 - 75^2}$$

$$= \frac{781}{399} \approx 1.957$$

$$a = \frac{161 - 1.957 \times 75}{8}$$

$$= 1.774$$

Regression equation : $\hat{y} = 1.957x + 1.774$

Ds 3b

$X = 7$

$\hat{y} = 1.957 X 7 + 1.774$

$= 15.5$

Ds 3c

The above prediction is valid because the prediction value lies in the given data

Ds 3d

$X = 15$

$\hat{y} = 15 X 1.957 + 1.774$

$= 31.1$

Ds 3e

This is not a valid prediction because the prediction value lies outside the given data.