3. Assume A = X, B = Y

X	50	60	50	60	80	50	80	40	70
Y	30	60	40	50	60	30	70	50	60
$X - \bar{X}$	-10	0	-10	0	20	-10	20	-20	10
$Y - \bar{Y}$	-20	10	-10	0	10	-20	20	0	10
$(X-\bar{X})^2$	100	0	100	0	400	100	400	400	100
$(Y-\bar{Y})^2$	400	100	100	0	100	400	400	0	100
$(X-\bar{X}).(Y-\bar{Y})$	200	0	100	0	200	200	400	0	100

$\bar{X} = 60$, $\bar{Y} = 50$, $\sum(X-\bar{X})^2 = 1600$, $\sum(Y-\bar{Y})^2 = 1600$, $\sum(X-\bar{X}).(Y-\bar{Y}) = 1200$, $n=9$

$r = \sqrt{bxy \,.\, byx} = \sqrt{0.75 \times 0.75} = \pm 0.75$

$bxy = byx = 1200/1600 = 0.75$

X on Y -> $x - 60 = 0.75\,(y - 50)$

$x = 0.75y + 22.5$

Y on X -> $y - 50 = 0.75\,(x - 60)$

$y = 0.75x + 5$

Your Answers

1.

X	57	42	40	33	42	45	42	44	40	56	44	43
Y	10	60	30	41	29	27	27	19	18	19	31	29
X- X̄	13	-2	-4	-11	-2	1	-2	0	-4	12	0	-1
Y- Ȳ	-18.3	31.7	1.7	12.7	0.7	-1.3	-1.3	-9.3	-10.3	-9.3	2.7	0.7
(X- X̄)^2	169	4	16	121	4	1	4	0	16	144	0	1
(Y- Ȳ)^2	334.89	1004.89	2.89	161.29	0.49	1.69	1.69	86.49	106.09	86.49	7.29	0.49
(X- X̄). (Y- Ȳ)	-273.9	-63.4	-6.8	-139.7	-1.4	-1.3	2.6	0	41.2	-111.6	0	-0.7

X̄ = 44, Ȳ = 28.3, ∑(X- X̄)^2 = 480, ∑(Y- Ȳ)^2 = 1749.68, ∑(X- X̄). (Y- Ȳ) = -555, n=12

$$r = \frac{\sum (X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum (X-\bar{X})^2 \sqrt{(Y-\bar{Y})^2}}}$$

r = -555 / √ [480 x √1749.68] = -0.598

X & Y are negatively correlated, as X increases Y decreases & vice versa with ~60%

2.

X	97.8	99.2	98.8	98.3	98.4	96.7	97.1	80
Y	73.2	85.5	78.9	75.8	77.2	87.2	83.8	85
Xi	5	1	2	4	3	7	6	8
Yi	8	2	5	7	6	1	4	3
di = Xi - Yi	-3	-1	-3	-3	-3	6	2	5
di ^2	9	1	9	9	9	36	4	25

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

∑d^2 = 102, n = 8

r = 1 - (6 x 102) / (8 x 63) = -0.214

X & Y are negatively correlated, as X increases Y decreases & vice versa with ~21%

4.

x	10	12	15	23	20
y	14	17	23	25	21
x ^2	100	144	225	529	400
y ^2	196	289	529	625	441
xy	140	204	345	575	420

∑x = 80, ∑y = 100, ∑x ^2 = 1398, ∑y ^2 = 2080, ∑xy = 1684

SSxx = ∑x^2 - (∑x)^2 / n = 1398 - (6400/5) = 118

SSxy = ∑xy - ∑x . ∑y / n = 1684 - 8000 / 5 = 84

SSyy = ∑y^2 - (∑y)^2 / n = 2080 - (10000 / 5) = 80

r = SSxy / √SSxx . SSyy = 84 / √118 x 80 = 0.865

5. The Chi-square test analyzes categorical data. It means that the data has been counted and divided into categories. It will not work with parametric or continuous data. It tests how well the observed distribution of data fits with the distribution that is expected if the variables are independent.

6. $X^2 = \sum (O - E)^2 / E$,

where O is Observed value, E is Expected value.

7. $X^2 = \sum (O - E)^2 / E$

Day	O	E	O-E	(O-E) ^2	X^2
Sunday	6	8	-2	4	0.5
Monday	4	8	-4	16	2
Tuesday	9	8	1	1	0.125
Wednesday	7	8	-1	1	0.125
Thursday	8	8	0	0	0
Friday	10	8	2	4	0.5
Saturday	12	8	4	16	2

- H0 : Days & no. of accidents are independent.
- H1 : Days & no. of accidents are not independent.
- Calculated value (CV): $\sum X^2 = 5.25$
- Table value (TV): DoF = 7-1 = 6 is 12.592
- Inference: CV < TV, H0 acceptable i.e., Days & no. of accidents are independent.