Answer to the question no-1

Interval	1-2	2.3	3.4	4.5	5.6	Total
Frequency (Height)	0	0	3	14	23	
Mid value (x_i)	1.5	2.5	3.5	4.5	5.5	
$f_i x_i$	0	0	10.5	63	126.5	200
$\sum f_i (x_i - \bar{x})^2$	0	0	6.75	3.5	5.75	16

$$\bar{x} = \frac{\sum f_i x_i}{n} = \frac{200}{40} = 5$$

$$\sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{n} = \frac{16}{40} = 0.4$$

$$\sigma = 0.63$$

Now,

$$\bar{x} + 2\sigma = 5.00 + 2 \times (0.63) = 6.26$$

$$\bar{x} - 2\sigma = 5.00 - 2 \times (0.63) = 3.74$$

So, from my observation I told that 95% value one in between $(\bar{x} \pm 2\sigma)$ this range.
So, it will follow normal distribution.

Answer to the question no - 2

Gender	Positive	Negative	Suspected	Total
Male	7 / 7.5	15 / 11.88	3 / 3.75	25
Female	5 / 4.5	7 / 7.12	3 / 2.25	15
Total	12	19	6	40 (n)

H_0 : Gender and covid-19 information and are independend

H_1 : Gender and covid-19 information are dependent.

Now,

$$x^2 = \sum\sum \frac{O^2_{ij}}{E_{ij}} - n$$

$$= \frac{7^2}{7.5} + \frac{15^2}{11.88} + \frac{3^2}{3.75} + \frac{5^2}{4.5} + \frac{7^2}{7.12}$$

$$+ \frac{3^2}{2.25} - 40$$

$$= 4.31$$

At 5% Level of significance

$$X^2_{cal} < X^2_{tab}$$

$$df = (r-1)(c-1) \Rightarrow (2-1)(3-1) = 2$$

Here, $X^2_{tab} = 5.99$

So, Null hypothesis is accepted. The Gender and covid-19 information are not dependent.

Answer to the question no-3

Now, correlation coefficient

$$\Sigma x = 1144$$
$$\Sigma y = 1980$$
$$\Sigma x^2 = 42470$$
$$\Sigma y^2 = 95735$$
$$\Sigma xy = 62010$$

$$r = \frac{\Sigma xy \cdot \frac{\Sigma x \, \Sigma y}{n}}{\sqrt{\left(\Sigma x^2 - \frac{(\Sigma x^2)^2}{n}\right)\left(\Sigma y^2 - \frac{(\Sigma y)^2}{n}\right)}}$$

$$= \frac{62010 - \frac{1144 \times 1980}{40}}{\sqrt{\left(42970 - \frac{(1144)^2}{40}\right)\left(105735 - \frac{(1980)^2}{40}\right)}}$$

$$= \frac{5.382}{8679.35}$$

$$= 0.62$$

Test for the Significance.

$H_0 : \ell = 0$

$H_1 : \ell \neq 0$

$$t = \frac{r\sqrt{r-2}}{\sqrt{1-r^2}}$$

$$= \frac{0.62\sqrt{40-2}}{\sqrt{1-(0.62)^2}}$$

$$\therefore t = 4.87$$

$df = n - 2 = 38$

$df = N - 38 = 38 - 28 = 18$

at 5% level of significence

Acceplence region

$\Rightarrow -2.70 < t < 2.101$

Null hypothesis is rejected correlation is signicant.