Question 1. There is an assumption that there is no significant difference between boys and girls with respect to intelligence. Tests are conducted on two groups and the following are the observations.

	Mean	Standard Deviation	Size
Girls	89	4	50
Boys	82	9	120

Validate the claim with 5% LoS (Level of Significance).

Answer:- To validate the claim that there is no significant difference between boys and girls with respect to intelligence, we perform a hypothesis test for difference in means using the Z-Test (because the sample sizes are larger: n1=50, n2=120).

Null Hypothesis (H_0): $\mu_1 = \mu_2$
(It means there is no significant difference in intelligence between boys and girls)

Alternative Hypothesis (H_1): $\mu_1 \neq \mu_2$
(It means there is significant difference in intelligence between boys and girls)

So this is a two tailed test at 5% LoS(Level of Significance $\alpha = 0.05$).

The Formula for comparing two sample:

Z = (x_bar1−x_bar2) / sqrt((sigma1^2/n1)+(sigma2^2/n2))

Z = (89-82) / sqrt((4^2/50)+(9^2/120))

Z = 7 / sqrt((16/50)+(81/120))

Z = 7 / sqrt(0.32+0.675)

Z = 7 / sqrt(0.995) = 7 / 0.997

Z = 7.02

In a two-tailed test, $\alpha = 0.05$ is split between both tails:

- Each tail has 0.025(i.e., 0.05/2)

Z-value that leaves 2.5% in the upper tail, which means:

- $p(Z>z) = 0.025 \Rightarrow p(Z<z) = 0.975$

Calculated Z-Value = 7.02
Critical Z-Value = NORM.S.INV(0.975) = 1.96

Calculated Z-Value > Critical Z-Value
Here 7.02 > 0.05, We reject the null hypothesis.

Final Conclusion: There is a significant difference between the intelligence of boys and girls at 5% Level of Significance. Hence, the assumption that there's no significant difference is not valid.

Question 2. Analyze the below data and tell whether you can conclude that smoking causes cancer or not?

Category	Diagnosed as Cancer	Without Cancer	Total
Smokers	220	230	550
Non-Smokers	350	640	990
Total	680	910	1590

Answer:- To determine whether smoking is associated with cancer, we can use the Chi-Square Test of independence. This helps us evaluate whether there's a significant association between two categorical variables: **smoking status and cancer diagnosis.**

Null Hypothesis(H_0):- Smoking and cancer are independent(i.e., smoking does not cause cancer).

Alternative Hypothesis(H_1):- Smoking and cancer are not independent(i.e., smoking may be associated with cancer).

<u>Calculated Expected Values</u>:-

Formula for expected values:

- **E = (Row Total) * (Column Total) / Grand Total**

Expected Values Table:

- **E(smokers,cancer) = 550*680/1590 = 235.22**
- **E(smokers, non-cancer) = 550*910 / 1590 = 314.78**
- **E(non-smokers,cancer) = 990*680 / 1590 = 423.39**
- **E(non-smokers,non-cancer) = 990*910 / 1590 = 566.60**

<u>Now Compute Chi-Square Test(X^2)</u>:

X^2 = Σ (O−E)^2 / E

X^2 = ((220-235.22)^2 / 235.22) + ((230-314.78)^2 / 314.78) + ((350-423.39)^2 / 423.39) + ((640-566.60)^2 / 566.60)

X^2 = ((-15.22)^2 / 235.22) + ((-84.78)^2 / 314.78) + ((-73.39)^2 / 423.39) + ((73.4)^2 / 566.60)

X^2 = (231.65/235.22) + (7187.65/314.78) + (5386.09/423.39) + (5387.56/566.60)
X^2 = 0.9848 + 22.8339 + 12.7213 + 9.5086

Calculated X^2 = 46.0486

Degree of Freedom:
df = (rows - 1) * (columns - 1) = (2-1) * (2-1) = 1
α = 0.05, df =1

Critical X^2 =CHISQ.INV.RT(0.05,1)
Critical X^2 = 3.84

Calculated X^2 > Critical X^2
46.0486 > 3.84
So, we reject the null hypothesis.

Final Conclusion:- We can conclude that smoking is associated with a higher risk of cancer.