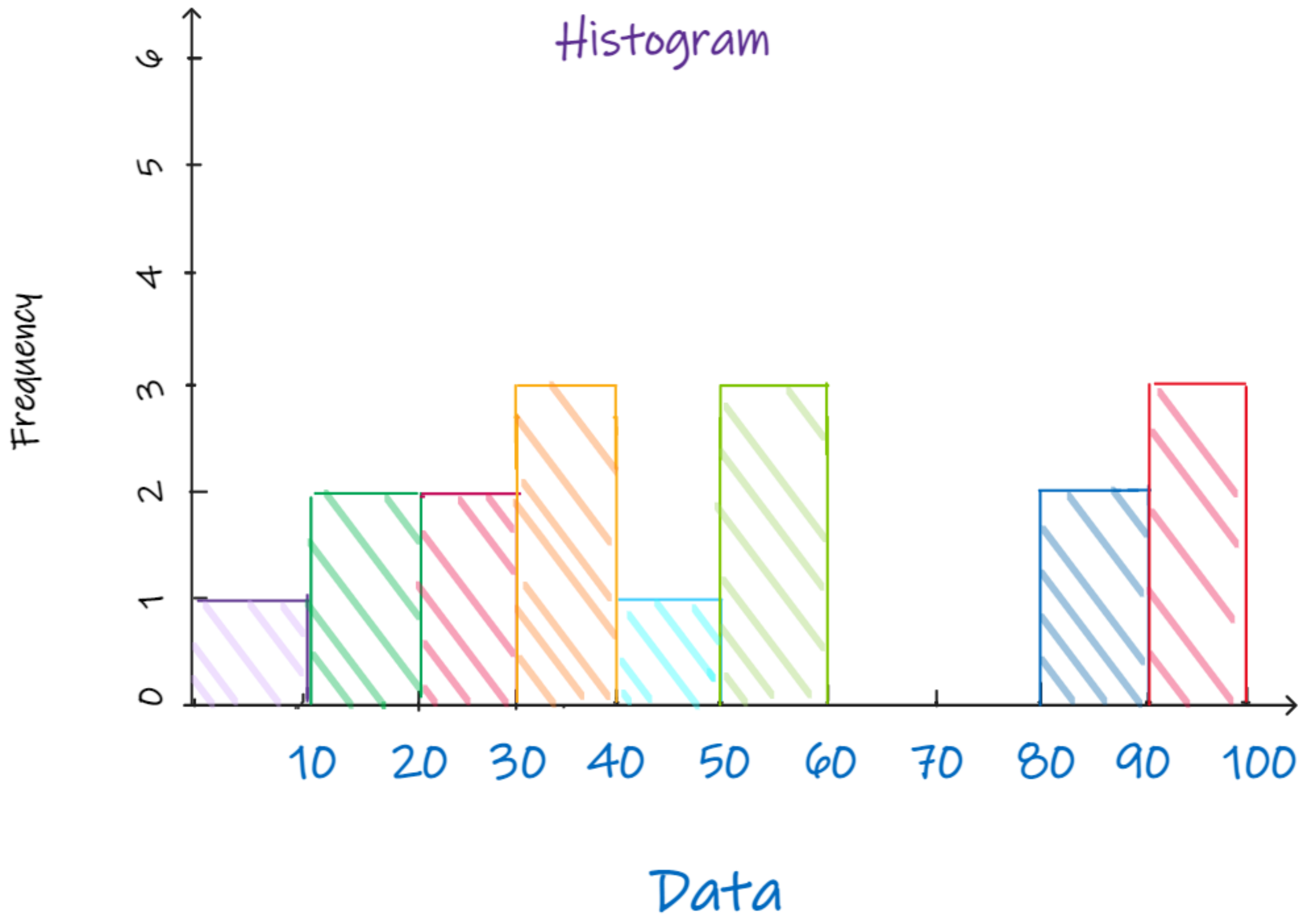
Answer 1 Plot a histogram:

10, 13, 18, 22, 27, 32, 38, 40, 45, 51, 56, 57, 88, 90, 92, 94, 99



Answer 2 In a quant test of the CAT Exam, the population standard deviation is known to be 100. A sample of 25 tests taken has a mean of 520. Construct an 80% CI about the mean.

$\boldsymbol{\sigma} = 100$

$n \ = \ 25$

$\bar{x} \ = 520$

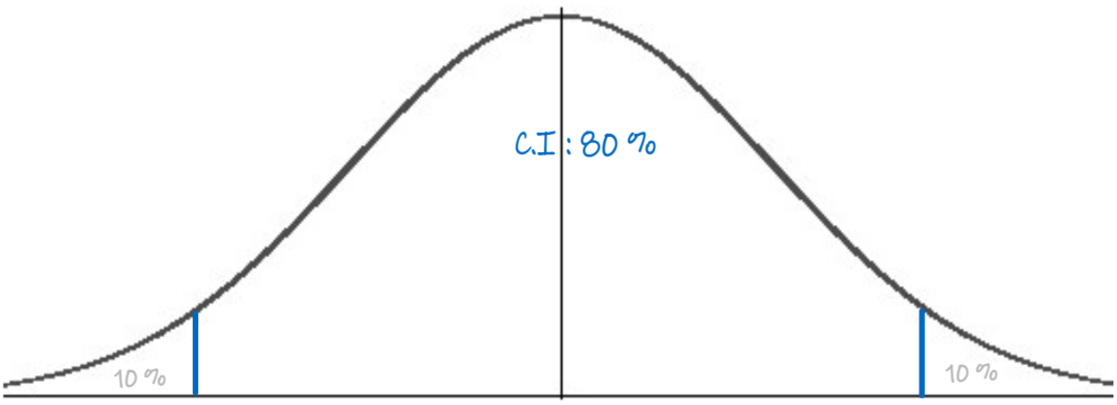
$CI \ \ = \ 80\,\%$

With the use of the sample mean (520), which is the point estimate, we attempt to estimate the population mean.

It would be a two-tailed test because it concern whether there is a difference between the two means

Even though our sample size n is less than 30, I'll still apply the z-test since the population standard deviation, a parameter from the population data, is provided.

If I consider that distribution of the sample mean is approximately normally distributed as shown below :



And by using an 80 % confidence interval, I can calculate the range of the population mean μ, as well as the lower and upper limits within which the mean may fall.

Confidence Interval = [lower limit, upper limit]

Confidence Interval = *Point estimate* $\pm$ *Margin of error*.

Point estimate is $\bar{x} = 520$

C.I = $\bar{x} \pm\ z_{\alpha/2} * \sigma/\sqrt{n}$

α is the significance value = 1- C.I = 1-0.8

α = 0.2

C.I = $520 \pm\ Z_{0.2/2} * 100/\sqrt{25}$

$Z_{0.2/2} = Z_{0.1}$

If $\alpha/2 = 0.1$, then the area under the curve of the Gaussian distribution for our expected range will be 0.9 ($Z_{0.1}$ = 1-0.1 = 0.9), So I find the z-score of about 1.28 in the Z-table (Even if I chose 1.29, the result will be nearly identical to 1.28, with only minor differences in decimals.)

Z	0.08	0.09
1.2	0.8997	0.9015

Lower limit = $520 - \ 1.28 * 100/\sqrt{25} = 494.4$

Upper limit = $520 + \ 1.28 * 100/\sqrt{25} = 545.6$

The sample mean is 520 and the confidence interval : [494.4 to 545.6], so the conclusion is :

I estimate that the population means of the quant test will fall between 494.4 and 494.4 with a 80 % confidence interval.

Answer 3 A car believes that the percentage of citizens in city ABC that owns a vehicle is 60% or less. A sales manager disagrees with this. He conducted a hypothesis testing surveying 250 residents & found that 170 residents responded yes to owning a vehicle.

a) *State the null & alternative hypothesis*.

The claim is 60% or less which is the null hypothesis, but the sales manager disagrees with this which means that the alternative hypothesis is greater than 60%, So:

- ➔ Null hypothesis (H_0): 60% of citizens own a vehicle or less.
- ➔ Alternative hypothesis (H_1): greater than 60% of citizens own a vehicle.

H_0: p_0 <= 0.6 and H_1: p_0 > 0.6

b) *At a 10% significance level, is there enough evidence to support the idea that vehicle ownership in ABC city is 60% or less*.

$n = 250$ residents

x value = 170 said yes (170 residents out of 250 own a car)

$\hat{p} = x/n = 170/250$

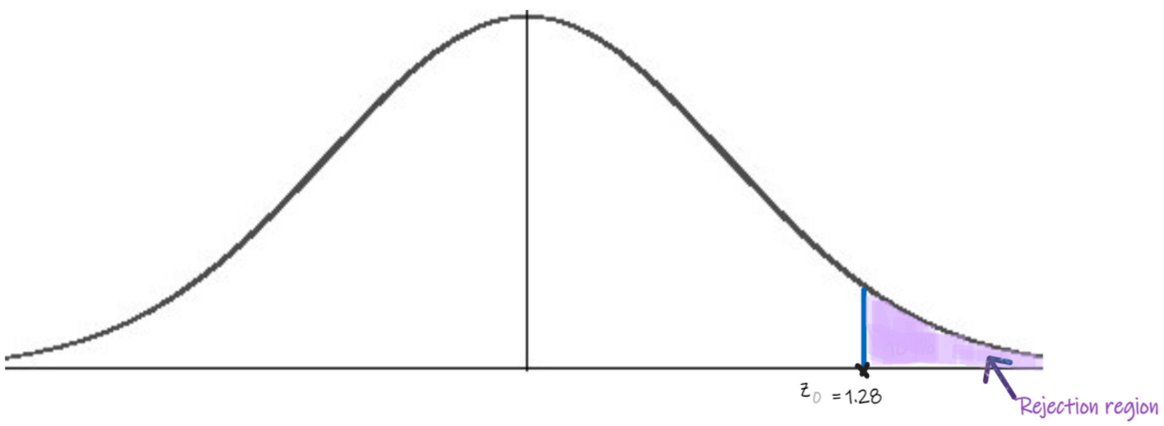
$\hat{p} = 0.68$

$\alpha = 0.1$

Since n is greater than 30, the z-test for proportions is suitable in this case:

This is a one tailed test: H_1 is greater than 0.6 so the rejection region is on the right side and to find out the right-tailed value:

Since $\alpha = 0.1$, the area under the curve is $1 - \alpha = 1 - 0.1 = 0.9$. So our critical z-value is 1.28. As shown by the curve below:



Calculate test statistic:

$$z = \hat{p} - p0 \ / \sqrt{p0(1 - p0)/n}$$

$$z = 0.68 - 0.6 / \sqrt{0.6(1 - 0.6)/250}$$

$z \approx$ 2.58

<u>Conclusion from rejection region</u> :

The found z statistic value 2.58 is above the critical value 1.28, so we reject the null hypothesis (H_0 *<= 60%*). No there is not enough evidence to support the idea that vehicle ownership in ABC city is 60% or less

Calculate *P-value* :

The value 2.58 in the normal distribution's z-table is about 0.9951.

0.9951 is the area in the left side, Since the entire area is equal to 1 then the area in the left tail is *1-0.9951 ≈ 0.0049*.

So P-value ≈ 0.0049.

<u>Conclusion from P-value</u>:

P-value is less than α (*0.0049 < 0.1*), So we reject the null hypothesis and accept the salesman hypothesis, which states that more than 60% of residents in the city ABC own a vehicle.

Answer 4 The value of the 99 percentile

2,2,3,4,5,5,5,6,7,8,8,8,8,8,9,9,10,11,11,12

Value = 99/100*(n+1) = 0.99*21 = 20.79

As we have precisely 20 values, the 99th percentile equals **12**.

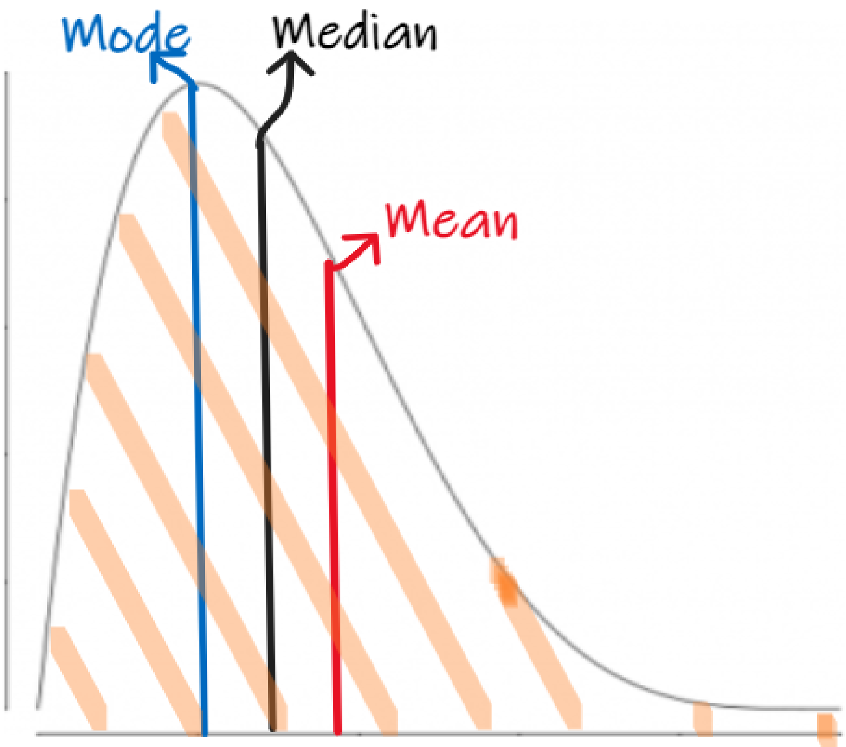
Answer 5 In left & right-skewed data, what is the relationship between mean, median & mode?

In a right-skewed distribution:

The mean is to the right of the median, indicating that it is greater than the median.

The mode appears at the highest frequency on the peak of the curve, somewhere in the left side before the median, meaning the mode is less than the median.

Right skewed graph: (Mode **<** Median **<** Mean), as shown below:



In a Left-skewed distribution:

The mean is to the left of the median, indicating that it is lower than the median.

The mode occurs on the top of the curve, the highest frequency in the left skewed is somewhere in the right side of the graph before the median, which means that the mode is greater than the median.

Left skewed graph: (Mode **>** Median **>** Mean), as shown below: