1. Given that,
   Daily average time spent by a user on a social media website = 50 minutes. "(μ), Hypothised or Known mean"
   Sample website users = 25 (n)
   Mean time spent by the sample users = 60 minutes "($\bar{X}$), sample mean of a random sample from a population"
   Sample Standard deviation = 30 minutes (S)

   Based on this information, the null and the alternative hypotheses will be:
   H_0 = the average time spent by the users is 50 minutes
   H_1 = the average time spent by the users is not 50 minutes

⇨ Here using t-test to test this hypothesis as comparison is in between sample mean (a random sample from a population) with the specific value (hypothesized or known mean of the population). Also population variance is unknown & sample standard deviation is given.

$t = (\bar{X} - \mu) / (S/\sqrt{n})$

Where, $\bar{X}$ is the sample mean, μ is the hypothesized population mean, S is the standard deviation of the sample and n is the number of observations in the sample.

From the above given data, t-value or t-test statistic & degree of freedom can be calculated as
$t = (60-50) / (30/\sqrt{25})$
$= (10) / (6)$
$= 1.67$

Degree of freedom = n-1
= 25-1
= 24

At a level of significance 0.05 & degree of freedom 24, the critical t-value comes out to be +/- 2.063899. Since the calculated t-value or t-test statistic of 1.67 is lesser than the critical t-value of 2.063899, the null hypothesis H_0 accepted.

Hence, the average time spent by the users is 50 minutes.

2.
First of all data is to be arranged in ascending or descending order for finding the median. Here, the given data set is arranged in ascending order.

160, 162, 162, 164, 168, 169, 170

The numbers of students (observations) are odd (Seven) in the given data set. Hence, the Median is equal to the observation present at $(n+1)/2^{th}$ position i.e., at $(7+1)/2=4^{th}$ position. The value present at 4^{th} position is 164.

Hence, the median of the height of seven students is 164 cm.

3.
Given observation of marks = {84, 85, 89, 92, 93, 89, 87, 89, 92}

Observations of Marks	Frequency
84	1
85	1
87	1
89	3
92	2
93	1

As we know that, mode is the value of observations which is occurred maximum no. of times in the given data set. From the above given data set, it may be seen that 89 is occurred 3 times whose frequency is highest.

Therefore, value of mode for the given data set is 89.

4.

Marks Xi	No. of students fi	fi*Xi
3	1	3
4	2	8
5	2	10
6	4	24
7	5	35
8	3	24
9	2	18
10	1	10
	∑fi=20	∑fi*Xi=132

Mean ($\bar{X}$ or E[X]) = (∑fi*Xi)/(∑fi)
= 132/20
= 6.6

i.e., Mean of marks obtained by 20 Students is 6.6.

5.
Let X be the random variable that represents the length of time. Here, to be find the probability that X is between 50 and 70 or $P(50<X<70)$.

Given that, random variable 'X' is normally distributed having mean (μ) 50 hours and standard deviation (σ) 15 hours.

Now converting the X to Z value to calculate the $P(50<X<70)$ using Z-table.

$Z= (X-\mu) / \sigma$

For X = 50, Z = (50 - 50) / 15 = 0
For X = 70, Z = (70 - 50) / 15 = 1.33 (rounded to 2 decimal places)

P(50< X<70) = P(0<Z<1.33) = [area to the left of Z = 1.33] - [area to the left of Z = 0]
= 0.9082 - 0.5 = 0.4082

i.e., the probability that John's computer has a length of time between 50 and 70 hours is equal to 0.4082.

6.
g = [10, 11, 12, 14, 15, 16, 17, 19, 21, 23]

Range= Largest value-Smallest Value
= 23-10
= 13

Hence, Range for the given data set is 13.

7.
Let A be the event that email detected as Spam & B be the event that email is Spam.

Given that, 50% of emails are Spam, i.e., P(B)=0.5,
Thus Probability that email is non-spam is P(B') = 1- P(B) = 1-0.5 = 0.5.

It is also given that a certain brand of software claims that it can detect 99% of spam emails,
i.e., P(A|B) = 0.99.
& the probability for a false positive (a non-spam email detected as spam) is 5%, i.e., P(A|B')=0.05.

Here, need to find the probability that the email is non-spam given that it is detected as spam, i.e., P(B'|A).

Using Bay's theorem required probability can be calculated as
P(B'|A) = P(A|B') P(B') / P(A)
= P(A|B') P(B') / (P(A|B)P(B) + P(A|B') P(B'))
= (0.05*0.5) / (0.99*0.5 + 0.05*0.5)
= 0.025 / (0.495 + 0.025)
= 0.025 / 0.52
= 0.0481

So, the required probability is 0.0481.

8.
For finding the quartile, first of all given data set is to be arranged in ascending or descending order. Here the given data is arranged in ascending order.

{10, 11, 12, 15, 16, 17, 19, 19, 21, 25} (n=10)

Lower quartile 'Q1' is at the position (n+1)/4 = (10+1)/4 = 2.75 of the ranked data. It means the value of lower quartile is equal to 'value at 2^{nd} position+0.75*(value at 3^{rd} Position-value at 2^{nd} position)'.

So, Lower quartile 'Q1' = 11+0.75*(12-11) = 11+0.75*1= 11.75.

9.

The measure of variability of the distribution is 'Variance'. Here the given distribution is binomial distribution, so the variance can be calculated as n*p*q. Where n is number of trials, p is probability of success & q is probability of failure.

Given that, number of trials (n) = 25
Probability of success (p) = 0.3
So, Probability of failure (q) = 1-p
= 1-0.3
= 0.7

Hence, as per question, Variance will be 25*0.3*0.7 = 5.25.

i.e., Variability of the given binomial distribution is 5.25.

10.

Let A to be the event of drawing a red ball and X & Y be the events that the ball is from the bag-I and bag-II, respectively.

By the question, probability of choosing a bag for drawing a ball is 1/2, i.e., P(X) = P(Y) = 1/2

Since there are 7 red balls out of a total of 9 balls in the bag-I, therefore, P(drawing a red ball from the bag-I) = P(A|X) = 7/9

Similarly, P(drawing a red ball from bag-II) = P(A|Y) = 5/14

As per question, need to determine the probability that the ball drawn is from the bag-I given that it is a red ball, i.e., P(X|A). Using Bayes theorem, P(X|A) can be calculated as

P(X|A) = P(A|X)*P(X) / P(A)
= P(A|X)*P(X) / [P(A|X)P(X) + P(A|Y)P(Y)]
= (7/9)*(1/2) / [(7/9)(1/2) + (5/14)(1/2)]
= 0.39 / [0.39 + 0.17]
= 0.39 / 0.56
= 0.69

Hence, Probability that the ball drawn is from the bag-I given that it is a red ball, i.e., P(X|A) = 0.69.

11. Given data set is g = [10, 23, 12, 21, 14, 17, 16, 11, 15, 19, 12]

⇨ Mean ($\bar{X}$) = Sum of observations / Number of observations
= (10+23+12+21+14+17+16+11+15+19+12) / 11
= 170 / 11
= 15.45

So, mean of the given data set is 15.45.

⇨ For finding of median, first of all the given data set is to be arranged in ascending or descending order. Here, arranged in ascending order.

g = [10, 11, 12, 12, 14, 15, 16, 17, 19, 21, 23], n=11

The numbers of observations are odd (Eleven) in the given data set. Hence, the Median is equal to the observation present at $(n+1)/2^{th}$ position i.e., at $(11+1)/2=6^{th}$ position. The value present at 6^{th} position is 15.

Hence, the median of the given data set is 15.

⇨ The mode is the value that appears most frequently in the given data set. Here, 12 occur most frequently (two times which is most of the times in the given data set).

So, mode of the given data set is 12.

12.
Given that, Sample Size (n) = 100
Mean height of the sample ($\bar{X}$) = 160
Standard deviation of Sample (S) = 10
Mean height of Population (μ) = 165 (to be checked the reasonability) [hypothesized population mean]

Let's assume, 5% significance level, i.e., $\alpha = 0.05$

H_0 = the mean height of population is 165 (I.e., Null Hypothesis H_0: $\mu = 165$)
H_1 = the mean height of population is not 165 (I.e., Alternative Hypothesis H_1: $\mu \neq 165$)

From the above given data, t-value or t-test statistic & degree of freedom can be calculated as

$t = (\bar{X} - \mu) / (S/\sqrt{n}) = (160-165) / (10/\sqrt{100}) = -5$

Degree of freedom = n-1 = 100-1 = 99

Where, $\bar{X}$ is the sample mean, μ is the hypothesized population mean, S is the standard deviation of the sample and n is the number of observations in the sample.

For level of significance $\alpha = 0.05$ & degree of freedom 99, the critical t-value comes out to be +/-1.984217. Since the calculated t-value or t-test statistic of $|-5|=5$ is greater than the critical t-value of 1.984217, the null hypothesis H_0 rejected.

Hence, the mean height of population may not be 165.

13.
Let's assume, A be the event that mammogram result is positive, B be the event that tumor is benign, C be the event that having risk of cancer.

Note that B' = C. Given that, P(C) = 0.01, so P(B) = 1 − P(C) = 0.99.

By the question,

Conditional probabilities P(A|C) =0.80 and P(A'|B) = 0.90, where the event A' is the complement of A, thus P(A|B) =0 .10
By using the Bayes' formula,

$$P(C|A) = P(A|C)P(C) / (P(A|C)P(C) + P(A|B)P(B))$$
$$= 0.80 \times 0.01 / (0.80 \times 0.01 + 0.10 \times 0.99)$$
$$= 0.075$$

So the chance would be 7.5%.

Hence, it is not agree that probability of cancer to be about 75%.

14.
Let E1, E2 & E3 be the event of choosing a card 'Red on both sides', 'Black on both sides' and 'Red on one side & Black on other side'.

Let X be the event that drawn card shows red on upper side. Then,

P(X|E1) = 2/2 = 1
P(X|E2) = 0/2 = 0
P(X|E3) = 1/2

Probability of choosing a card is P(E1)=P(E2)=P(E3)=1/3

By the question, need to calculate the probability that other side is colored black, if the upper side of chosen card is red means P(E3|X) is to be calculated.

By using the Baye's Formula,

$$P(E3|X) = P(X|E3)P(E3) / [P(X|E1)P(E1) + P(X|E2)P(E2) + P(X|E3)P(E3)]$$
$$= (1/2)*(1/3) / [1*(1/3) + 0*(1/3) + (1/2)*(1/3)]$$
$$= (1/6) / [(1/3) + (1/6)]$$
$$= (1/6) / [1/2]$$
$$= 1/3$$

Hence, the required probability is 1/3=0.33.