Q-1. What **is** central limit theorem **and** why **is** it important?

Ans :-

    The CLT gives us a certain distribution over our estimations.
    We can utilize this to pose an inquiry about the probability
    of an estimate that we make
    The Central Limit Theorem **is** important **for** statistics because it
allows us
    to safely assume that the sampling distribution of the mean will
be normal **in** most cases.
    This means that we can take advantage of statistical techniques
that assume a normal distribution,
    as we will see **in** the next section.

Q-2. What **is** sampling? How many sampling methods do you know?

Ans :-

    Sampling means selecting the group that you will actually collect
data from **in** your research.
    Two types of sampling method.
    Probability sampling involves random selection, allowing you to
make strong statistical inferences
    about the whole group.
    Non-probability sampling involves non-random selection based on
convenience **or** other criteria,
    allowing you to easily collect data.

Q-3. What **is** the difference between type1 **and** typeII error?

Ans :-
    Type – 1 **and** type – 2 errors are inversely related, i.e., **if** one
increases, the other decreases.
    Type -1 error **is** also known as false-positive, whereas type – 2
error **is** known as false-negative.
    By decreasing **or** increasing the significance level, we can reduce
the type – 1 **or** type – 2 errors.

Q-4. What do you understand by the term Normal distribution?

Ans :-

    Normal distribution, also known as the Gaussian distribution,
    is a probability distribution that **is** symmetric about the mean,
    showing that data near the mean are more frequent **in** occurrence
    than data far from the mean.

Q-5. What **is** correlation **and** covariance **in** statistics?

Ans :- Covariance **is** a statistical term that refers to a systematic relationship
between two random variables **in** which a change **in** the other reflects a
change **in** one variable.
correlation **is** a measure that determines the degree to which two **or** more
random variables move **in** sequence

Q-6. Differentiate between univariate ,Biavariate,**and** multivariate analysis.

Ans :-

1. Univariate data —
This type of data consists of only one variable.

2. Bivariate data —
This type of data involves two different variables.

3. Multivariate data —
When the data involves three **or** more variables

Q-7. What do you understand by sensitivity **and** how would you calculate it?

Ans :-
It **is** commonly used by investors who takes into consideration
the conditions that affect their potential investment to test,
predict **and** evaluate result.

Calculation of sensitivity.
For the calculation of Sensitivity Analysis,
go to the Data tab **in** excel **and** then select
What **if** analysis option

8. What **is** hypothesis testing? What **is** H0 **and** H1? What **is** H0 **and** H1 **for** two-tail test?

Ans :-
In statistics, a method **for** testing how accurately a mathematical model based
on one set of data predicts the nature of other data sets generated by the same process.

$H0$ **is** usually referred to as null hypothesis **and** $H1$ **is** an alternative hypothesis.

Q-9. What **is** quantitative data **and** qualitative data?

Ans :-
   Quantitative data are measures of values **or** counts **and** are
expressed as numbers.
   Quantitative data are data about numeric variables (e.g. how many;
how much; **or** how often).
   Qualitative data are measures of types **and** may be represented by a
name, symbol, **or** a number code.

Q-10. How to calculate range **and** interquartile range?

Ans :- To calculate the range, you need to find the largest observed
value of a variable (the maximum)
   and subtract the smallest observed value (the minimum).
   The range only takes into account these two values **and** ignore the
data points
   between the two extremities of the distribution.

Q-11. What do you understand by bell curve distribution ?

Ans :-

   Bell curve refers to the bell shape that **is** created when a line **is**
plotted
   using the data points **for** an item that meets the criteria of
normal distribution.

12. Mention one method to find outliers.

Ans :-
   Sorting method **is** a method to find outliers.
   You can sort quantitative variables from low to high **and** scan **for**
extremely
   low **or** extremely high values. Flag any extreme values that you
find.

Q-13. What **is** p-value **in** hypothesis testing?

A p-value **is** a statistical measurement used to validate a hypothesis
against observed data.

Q-14. What **is** the Binomial Probability Formula?

Q-15. Explain ANOVA **and** it's applications.

Ans :-

   ANOVA **is** a statistical formula used to compare variances across
the means (**or** average) of different groups.

A range of scenarios use it to determine **if** there **is** any difference between the means of different groups.