Section A

1. i. a. Inductive C. Specific to General
    b. Knowledge A. Declarative Knowledge
    c. Bloom B. Intellectual Skill
    d. Who what when & how D. Descriptive Research

ii. Survey research are four types Questionnaire, Case Study, Interview, Opinion Poll

iii. Generalization is associated with Scientific Research

iv. Variance: The Variance is defined as:
The average of the squared differences from the Mean.

Example:
You and your friends have just measured the heights of your dogs (in millimeters):
The heights (at the shoulders) are: 600mm, 470mm, 170mm, 430mm and 300mm.
Find out the Mean, the Variance, and the Standard Deviation.
Your first step is to find the Mean:

Answer:
Mean = \( \frac{600 + 470 + 170 + 430 + 300}{5} = \frac{1970}{5} = 394 \) so the mean (average) height is 394 mm.

To calculate the Variance, take each difference, square it, and then average the result:

\[
\text{Variance: } \sigma^2 = \frac{206^2 + 76^2 + (-224)^2 + 36^2 + (-94)^2}{5}
\]

\[
= \frac{42,436 + 5,776 + 50,176 + 1,296 + 8,836}{5}
\]

\[
= \frac{108,520}{5} = 21,704
\]

So, the Variance is 21,704.

v. The full form of ANOVA is Analysis of Variance.

vi. Case study starts with sample population

vii. Sample mean is mean calculated from the sample. The mean (or average) is the most popular and well known measure of central tendency. It can be used with both discrete and continuous data, although its use is most often with continuous. The mean is equal to the sum of all the values in the data set divided by the number of values in the data set. So, if we have \( n \) values in a data set and they have values \( x_1, x_2, ..., x_n \), the sample mean, usually denoted by \( \bar{x} \) (pronounced x bar), is:

\[
\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}
\]

viii. The null hypothesis (H0), stated as the null, is a statement about a population parameter, such as the population mean, that is assumed to be true. The null hypothesis is a starting point. We will test whether the value stated in the null hypothesis is likely to be true.

ix. Webometric analysis

x. Probability and non-probability sampling
Section – B

2. The various mode of thinking are Authority centered, Speculative, tenacious, positivistic. Students have to explain each mode of thinking with suitable example. Here one thinking process is give as example.

**Positivistic Thinking**
Thinking as based on scientific methodology and scientific views, such as physicalism, realism, reductionism, empiricism, positivism, etc, and specifically on *scientism*, which is defined as:
(a) The sciences are more important than the arts for an understanding of the world in which we live, or, even, all we need to understand it.
(b) Only a scientific methodology is intellectually acceptable.
(c) Philosophical problems are scientific problems and should only be dealt with as such.

Students have to explain steps in each type of thinking, its characteristics with examples.

3. Scientific method: The **scientific method** is a way to ask and answer scientific questions by making observations and doing experiments. The scientific method is an ongoing process, which usually begins with observations about the natural world. Human beings are naturally inquisitive, so they often come up with questions about things they see or hear and often develop ideas (hypotheses) about why things are the way they are and finally trying to prove.

![Scientific Method Diagram](http://nopr.niscair.res.in/bitstream/123456789/28515/1/ALIS%204%281%29%2019-32.pdf)

The scientific method is well explained in [http://nopr.niscair.res.in/bitstream/123456789/28515/1/ALIS%204%281%29%2019-32.pdf](http://nopr.niscair.res.in/bitstream/123456789/28515/1/ALIS%204%281%29%2019-32.pdf).

Students may know the detail of this method from the article of Dr. S.R. Ranganathan.

4. Here the relation of knowledge and Five laws of library science need to be explained.

**Speculative Thinking**
Thinking based on the philosophical notion of *speculation*, which is defined as:
Speculation or speculative thinking designates a knowledge or cognition that transcends experience and is directed at the spiritual, super-natural and divine, fundamental to experience. Kant (*Logic, Introd. IV*): "Cognition of the general *in abstracto* is speculative cognition; cognition of the general *in concreto* is common cognition. Philosophy is speculative cognition and it therefore begins where the common use of reason sets out to make attempts at cognition of the general *in abstracto."

Students have to explain the characteristics, steps and example of Speculative mode of thinking in detail.

5. **Sampling Errors**
**Sampling errors** occur because only a sample of the population is investigated. The particular sample used in this survey is one of a large number of possible samples of the same size that could have been selected using the same sampling method. It is clear that estimates based on different samples will
differ from one another, and almost all of them will differ from the value obtained had a complete census been taken – “the census value”.

The estimate $X'$ is the value, estimated from the specific sample of this survey, of the corresponding value $X$ that would have been obtained had a complete census been conducted.

The sampling error $\sigma(X')$ is a measure of the variability between the different values of the estimate that would have been obtained from all possible samples of the same size and derived by the same sampling method and the census value obtained under the same conditions of data collection.

In view of the variability of sample estimates, it may be preferable to consider a range of values where the census value is likely to be found, with a given probability (or level of confidence) rather than to rely on the specific value of the estimate obtained from the sample. This can be done using the sampling error.

The confidence interval for an estimate is the range containing the census value $X$ with a given level of confidence. Using both the estimate $X'$ and its estimated sampling error $\sigma'(X')$, it is possible to construct a confidence interval that contains the census value $X$ with the given level of confidence.

Confidence intervals are usually given at a 95% level. These may be obtained as the range between $X' - 2\sigma'(X')$ and $X' + 2\sigma'(X')$. It may be asserted with 95% confidence that the census value lies in this range.

Diagram is the way to present the data in pictorial form. Sometimes data can not well explain by table or text. Diagram can however depict the things in well manner. Students have to explain the use of diagram in LIS research.

6. In DDC main division are arranged in inverted Baconian form. Therefore, Philology, Poesy and History are the three main divisions of DDC. Further the whole universe is divided into 10 divisions and under each division ideas are arranged in sub-divided form.

Here students have to explain the mapping of main class, sub-class, use of add device, tables etc with suitable example. How classes are constructed in DDC in schedule to other schedule, schedule to same schedule and schedule to table are to be explained.

7. Data:

**Different kinds of data**

The type of data you have determines what sort of test you apply and so you need to know what kind yours is. It is by far the best policy to decide what sort of data you will be collecting before you actually go out and collect it. (See "Advice to distressed non-statisticians").

There are two types of data: **qualitative** and **quantitative**. The way we typically define them, we call data 'quantitative' if it is in numerical form and 'qualitative' if it is not. Notice that qualitative data could be much more than just words or text. Photographs, videos, sound recordings and so on, can be considered qualitative data.

Further data may be:

<table>
<thead>
<tr>
<th>Qualitative data</th>
<th>is data that is not given numerically;</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. favourite colour, place of birth, favourite food, type of car.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quantitative data</th>
<th>is numerical. There are two types of quantitative data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discrete data</td>
<td>can only take specific numeric values;</td>
</tr>
<tr>
<td>e.g. shoe size, number of brothers, number of cars in a car park.</td>
<td></td>
</tr>
<tr>
<td>Continuous data</td>
<td>can take any numerical value;</td>
</tr>
<tr>
<td>e.g. height, mass, length.</td>
<td></td>
</tr>
</tbody>
</table>
**Interval Data**
These are measurements of things like weight and length. A length of 10m it is 2m smaller than a length of 12m. If you can say how far apart one measurement is from another, then you have interval level data.

**Ordinal Data**
You might have measurements where you do not know how far apart one measurement is from another. For instance let’s say you are concerned with facial topiary in humans. You don’t want to upset your subjects by measuring directly so you invent a scale of beardiness thus:

1 = naked chin ...2 = stubbly...3 = toothbrush...4 = medium hair...5 = very hairy...6 = badger concealingly hairy

You cannot say that level 4 is twice as bearded as level 2 but you can put your measurements in order. They are ordinal level measurements.

**Categorical Data**
Quite often in biology you might be concerned with putting things into categories. For example it might be that you are interested in testicle possession in humans. You can put each human into one of two categories: Either with testicles or without testicles. (There is the possibility of a third category of uni-testicular individuals but let us ignore them). You could then use this data to prove that the mean number of testicles per human was one.

Or maybe you are looking at shore crabs and their parasitic barnacle *Saculina*. You would have two categories: With the parasite and without the parasite.

**Matched and Unmatched Data**
Your data is matched if a piece of data from one set goes with only one piece of data from the other set. For example you might be measuring temperature of the sea with depth. A specific temperature recording would only be associated with one specific depth.

Your data is unmatched if there is no reason to associate a piece of data from one set with any particular piece of data from the other set. For example you might be measuring the heights of vegetation on trampled and untrampled parts of a path. There is no connection between any of the measurements from the trampled part and the untrampled part.

**Questionnaires**
Where an enumerator poses questions directly, questionnaires refer to forms filled in by respondents alone. Questionnaires can be handed out or sent by mail and later collected or returned by stamped addressed envelope. This method can be adopted for the entire population or sampled sectors.

Questionnaires may be used to collect regular or infrequent routine data, and data for specialised studies. While the information in this section applies to questionnaires for all these uses, examples will concern only routine data, whether regular or infrequent. Some of the data often obtained through questionnaires include demographic characteristics, fishing practices, opinions of stakeholders on fisheries issues or management, general information on fishers and household food budgets.

A questionnaire requires respondents to fill out the form themselves, and so requires a high level of literacy. Where multiple languages are common, questionnaires should be prepared using the major languages of the target group. Special care needs to be taken in these cases to ensure accurate translations.

In order to maximise return rates, questionnaires should be designed to be as simple and clear as possible, with targeted sections and questions. Most importantly, questionnaires should also be as short as possible. If the questionnaire is being given to a sample population, then it may be preferable to prepare several smaller, more targeted questionnaires, each provided to a sub-sample. If the questionnaire is used for a complete enumeration, then special care needs to be taken to avoid overburdening the respondent. If, for instance, several agencies require the same data, attempts should be made to co-ordinate its collection to avoid duplication.

The information that can be obtained through questionnaires consists of almost any data variable. For example, catch or landing information can be collected through questionnaire from fishers, market middle-persons, market sellers and buyers, processors etc. Likewise, socio-economic data can also be obtained through questionnaires from a variety of sources. However, in all cases variables
obtained are an opinion and not a direct measurement, and so may be subject to serious errors. Using
direct observations or reporting systems for these sorts of data is more reliable.
Questionnaires, like interviews, can contain either structured questions with blanks to be filled in,
multiple choice questions, or they can contain open-ended questions where the respondent is
encouraged to reply at length and choose their own focus to some extent.
To facilitate filling out forms and data entry in a structured format, the form should ideally be
machine-readable, or at least laid out with data fields clearly identifiable and responses pre-coded.
In general, writing should be reduced to a minimum (e.g. tick boxes, multiple choices), preferably
being limited to numerals. In an open-ended format, keywords and other structuring procedures
should be imposed later to facilitate database entry and analysis, if necessary.

8. Central Tendency - Mean
The mean (or average) of a set of data values is the sum of all of the data values divided by the
number of data values. That is:
\[
\text{Mean} = \frac{\text{Sum of all data values}}{\text{Number of data values}}
\]
Symbolically,
\[
\bar{x} = \frac{\sum x}{n}
\]
where \(\bar{x}\) (read as 'x bar') is the mean of the set of \(x\) values,
\(\sum x\) is the sum of all the \(x\) values, and
\(n\) is the number of \(x\) values.

Example 1
The marks of seven students in a mathematics test with a maximum possible mark of 20 are given
below:
15  13  18  16  14  17  12
Find the mean of this set of data values.
Solution:
\[
\text{Mean} = \frac{\text{Sum of all data values}}{\text{Number of data values}} = \frac{15+13+18+16+14+17+12}{7} = \frac{105}{7} = 15
\]
So, the mean mark is 15.
Symbolically, we can set out the solution as follows:
\[
\bar{x} = \frac{\sum x}{n} = \frac{15+13+18+16+14+17+12}{7} = \frac{105}{7} = 15
\]
So, the mean mark is 15.

**Median**
The **median** of a set of data values is the middle value of the data set when it has been arranged in ascending order. That is, from the smallest value to the highest value.

**Example 2**
The marks of nine students in a geography test that had a maximum possible mark of 50 are given below:

47 35 37 32 38 39 36 34 35

Find the median of this set of data values.

**Solution:**
Arrange the data values in order from the lowest value to the highest value:

32 34 35 36 37 38 39 47

The fifth data value, 36, is the middle value in this arrangement.

∴ Median = 36

**Note:**

The number of values, \( n \), in the data set = 9

\[
\text{Median} = \frac{1}{2}(n+1) \text{ th value}
\]

= 5th value

= 36

**In general:**

\[
\text{Median} = \frac{1}{2}(n+1) \text{ th value}, \text{ where } n \text{ is the number of data values in the sample}
\]

If the number of values in the data set is even, then the **median** is the average of the two middle values.

**Example 3**
Find the median of the following data set:

12 18 16 21 10 13 17 19

**Solution:**
Arrange the data values in order from the lowest value to the highest value:

10 12 13 16 17 18 19 21

The number of values in the data set is 8, which is even. So, the median is the average of the two middle values.

\[
\therefore \text{Median} = \frac{4\text{th data value} + 5\text{th data value}}{2}
\]

\[
= \frac{16+17}{2}
\]

\[
= \frac{33}{2}
\]

\[
= 16.5
\]
Alternative way:
There are 8 values in the data set.
\[ \therefore \ n = 8 \]
Now, median = \[\left(\frac{n+1}{2}\right)\text{th value}\]
= \[\left(\frac{8+1}{2}\right)\]
= \[\frac{9}{2}\]
= 4.5\text{th value}

The fourth and fifth scores, 16 and 17, are in the middle. That is, there is no one middle value.
\[ \therefore \text{Median} = \frac{16+17}{2} \]
= \[\frac{33}{2}\]
= 16.5

Note:
- Half of the values in the data set lie below the median and half lie above the median.
- The median is the most commonly quoted figure used to measure property prices. The use of the median avoids the problem of the mean property price which is affected by a few expensive properties that are not representative of the general property market.

Mode

The mode of a set of data values is the value(s) that occurs most often.
The mode has applications in printing. For example, it is important to print more of the most popular books; because printing different books in equal numbers would cause a shortage of some books and an oversupply of others.
Likewise, the mode has applications in manufacturing. For example, it is important to manufacture more of the most popular shoes; because manufacturing different shoes in equal numbers would cause a shortage of some shoes and an oversupply of others.

Example 4

Find the mode of the following data set:
48 44 48 45 42 49 48

Solution:
The mode is 48 since it occurs most often.

Note:
- It is possible for a set of data values to have more than one mode.
- If there are two data values that occur most frequently, we say that the set of data values is bimodal.
- If there is no data value or data values that occur most frequently, we say that the set of data values has no mode.