3(a) Explain, what is meant by descriptive statistics and inferential statistics.

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. They form the basis of quantitative analysis of data.

Descriptive statistics describe what is or what the data shows. Inferential statistics help to reach conclusions that extend beyond the immediate data alone. For example, we use inferential statistics to try to infer from the sample data what the population might be. Or, we use inferential statistics to make judgments of the probable inflation. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in our data.

Descriptive Statistics are used to present quantitative descriptions in a manageable form. In a research study we may have lots of measures. Or we may measure a large number of people on any measure. Descriptive statistics help us to simplify large amounts of data in a sensible way. Each descriptive statistic reduces lots of data into a simpler summary. For instance, consider a simple number used to summarize how well a batsman is performing in cricket, the batting average. This single number is simply the score divided by the number of balls faced. The single number describes a large number of discrete events. If we consider the Grade obtained by a student in the Post Graduate Diploma examination, this single indicator describes the general performance of a student across wide range of course experiences such as Tutor Marked Assessment, oral presentation, Multiple Choice examination and the Final written examination.

Anyhow every time we try to describe a large set of observations with a single indicator we run the risk of distorting the original data or losing important detail. The batting average doesn't tell us whether the batsman hitting boundaries or singles. It doesn't tell whether he's been a steady bat. The Grade doesn't tell us whether the student was in difficult courses or easy ones, or whether they were courses in their major field or in other disciplines. Even given these limitations, descriptive statistics provide a powerful summary that may enable comparisons across people or other units.

The steps in descriptive statistics are to collect, classify, summarise and present data. To summarise the data we use graphs, tables and calculate values to represent the sample data. When we are summarizing a quantity like length or weight or age, it is common to use [arithmetic mean](http://en.wikipedia.org/wiki/Arithmetic_mean), the [median](http://en.wikipedia.org/wiki/Median), or the [mode](http://en.wikipedia.org/wiki/Mode_%28statistics%29).

Inferential statistics refer to the use of current information regarding a sample of subjects in order to make assumptions about the population at large and/or to make predictions about what might happen in the future. The goal of inferential statistics is to take what is known and make assumptions or inferences about what is not known.

For example, a Board of Examiners may want to compare the performance of 10,000 students that completed an examination. Of these, 5,500 students are girls and 4,500 students are boys. The 10,000 students represent our ‘population’. Whilst we are interested in the performance of all 10,000 students, girls and boys, it may be impractical to examine the marks of all of these students because of the time and cost required to collate all of their marks. Instead, we can choose to examine a ‘sample’ of these students and then use the results to make generalizations about the performance of all 10,000 students. For the purpose of our example, we may choose a sample size of 500 students. Since we are looking to compare boys and girls, we may select 275 girls and 225 boys in our sample (In reality, we should be more scientific and apply the instructions, in the statistical guide on [**sampling**](http://www.stats4students.com/Essentials/Sampling/Overview.php)to choose an appropriate sample size).

 3 (b) Differentiate a qualitative variable from a quantitative variable.

Variable is any quantity that varies. Any attribute, phenomenon or event that can have different values.

Variables that express a qualitative attribute are called qualitative variables.

Variables that are measured in terms of numbers are called quantitative variables.

Some examples of qualitative variables are colour of the skin, religion, favourite Sports and gender.

Some examples of quantitative variables are height, weight and shoe size.

3 (c) Differentiate between a discrete variable and a continuous variable

Variable with possible scores of discrete points on the scale such as number of students in a classroom are called [discrete variables](http://psych.rice.edu/online_stat/glossary/discrete_variables.html). A classroom could have thirty or thirty one students, but not 30.25 students.

Variable where the scale is continuous and not made up of discrete steps, such as "time taken to finish a 100 meter sprint" are [continuous variables](http://psych.rice.edu/online_stat/glossary/continuous_variables.html). The time taken by a certain athlete to finish a 100 meter dash could be 10.9 seconds, or it could be 10.897654321 seconds.

3 (d) Distinguish among nominal, ordinal, interval and ratio levels of measurement.

Variables can be measured at different levels. In order to choose the proper statistics to examine data, we first have to figure out at what level each variable is measured.

A variable is measured by thinking of its categories. The categories are related to each other. Sometimes the categories are numbers, and sometimes they are words.  Sometimes the categories have an order, and sometimes they do not.

**The Nominal Level**

The least precise level of measurement is the nominal level.  Sex (male or female), ethnicity (Sinhalese, Tamils, Muslims, Burghers), Political Party Identification (SLFP, UNP, JVP, LSSP, Democrat, Republican, Independent, etc.) and Religion (Buddhist, Hindu, Catholic, Islam, Protestant, etc.) are categories that are referred by just names so that we can classify people by sex, ethnicity, or religion. The categories don't have any particular order from more to less or higher to lower.  That is, someone in the category Sinhalese does not have more or less "ethnicity" than a Muslim or a Tamil, just a different ethnicity.  A UNPer does not have more or less "party identification" than a SLFPer.  There is no order or ranking in the categories. It’s only a name. So we treat these variables as nominal-level

**The Ordinal Level**

The word "ordinal" means ‘in order’. If we had a variable whose categories have an order, we might have an ordinal-level variable. The variable "fear for height" with categories such as very afraid, somewhat afraid, and not afraid would be an example.  These categories have not only names but also have something more.  The categories have an order from more to less fear.  Another example is "social class," with categories such as lower class, working class, middle class, and upper class.

Almost any method of measuring attitudes results in ordinal-level variables, even if the variables include only two categories.  For example, we could categorize the variable "attitudes toward capital punishment" into those who favour and those who oppose; and those who favour capital punishment hold more favourable attitudes, while those who oppose hold less favourable attitudes.

The same variable could be treated as nominal as well as ordinal, depending on the situation. For example, if we are to categorize the reasons for calls made to ‘Sumithrayo’on a particular day, we might come up the following categories:  husband scolds the wife, husband hits the wife, getting suspended from work, chased out from home, fail in the examination, wife abandons husband and children and lives with another man, rape.  One way to think of these categories is as just names of problems, and the name of the variable could be "Type of Problem."  In this case, we would be conceiving of our variable as nominal-level.  But another way to think of the categories is to order them from least severe (husband scolds the wife) to most severe (rape).  The name of
the variable, as it is being conceived this time, would be "Severity of Problem," and it would be an ordinal-level variable.

**The Interval Level**

"Temperature in degrees," may be one of the best examples for an interval-level variable.  Temperature is measured in degrees, and the degrees are not words (cold, super-cold, warm, etc.), but numbers corresponding to levels of mercury in a thermometer.  The distance, or interval, between 1 degree and 2 degrees is exactly equal to the distance between 40 degrees and 41 degrees.  The intervals between any two adjacent categories are equal (exactly 1 degree).  It does not have a theoretical zero point.  Actually, a thermometer does have a zero, but the zero does not indicate a lack or absence of the variable, temperature.  Zero indicates "cold."  And one method of measuring temperature (e.g., Fahrenheit) has a different spot for zero than others (e.g., Celsius).  These are arbitrary zero points that are not intended to indicate a total lack of temperature.  It's really impossible to imagine a lack of temperature.  With no true zero point, temperature in degrees must be considered only an interval-level variable.

**The Ratio Level**

Variables measured at the ratio level have all the characteristics of nominal-, ordinal-, and interval-level measures (categories that have names, order, and equal intervals), and the categories include a true zero point.  Even though the sample we are examining may not include any cases in the category zero, zero is possible at least in theory.  An example is "income in dollars."  The categories have names (50,000 rupee, 75,000 rupee; 100,000 rupee; etc.); the categories follow an order from less income to more income; the intervals between the categories are equal (1 rupee;); and it is possible to have zero rupee.  Age is another example.  The categories have names (10 year old, 52 years old, etc.); the categories have an inherent order from youngest to oldest; the intervals between the categories are equal (1 year); and it is possible to be 0 years old.  Nnmber of children, number of pets, years of education, number of arrests, years on the job, number of marriages are a few examples.