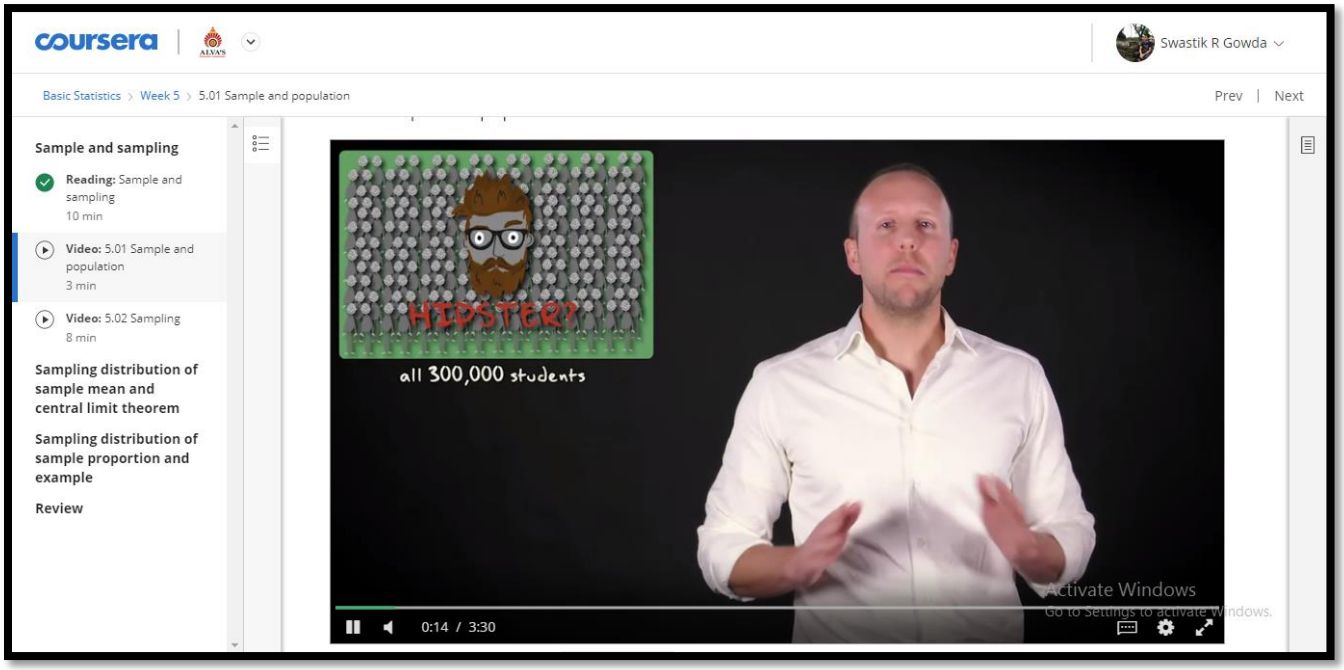
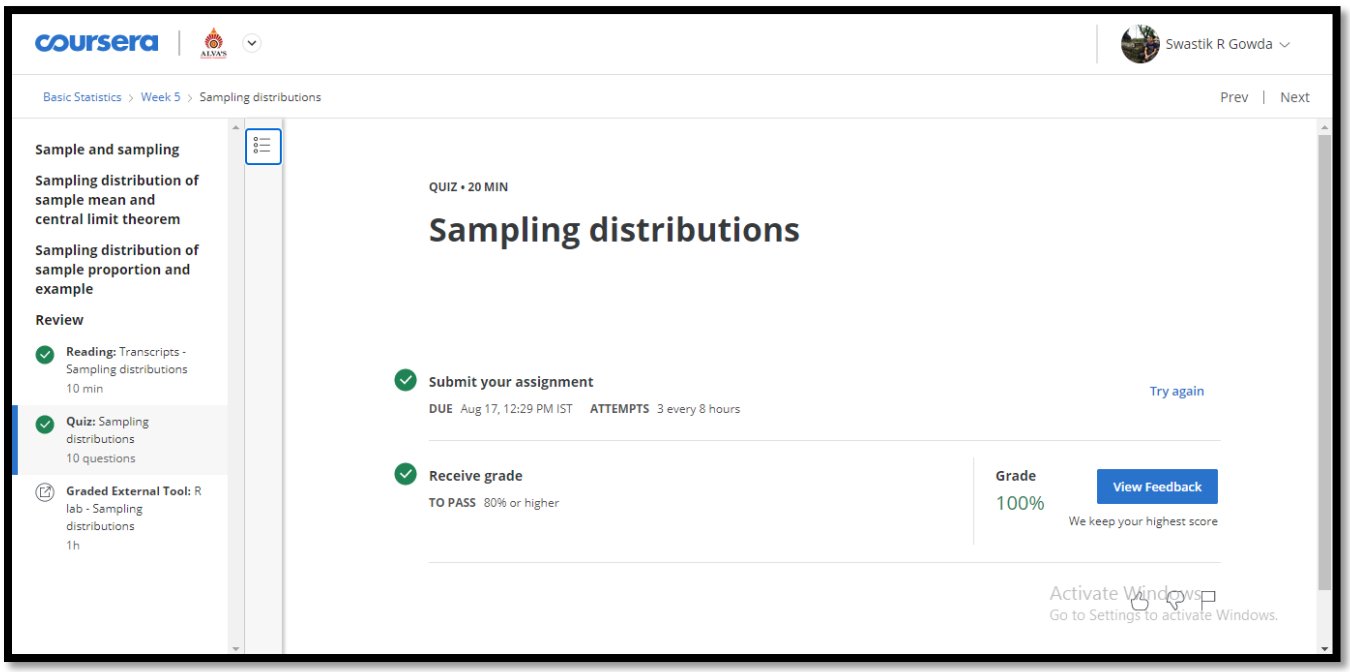
DAILY ASSESSMENT REPORT

Date:	**28-July-2020**	**Name:**	**Swastik R Gowda**
Course:	**Coursera – Basic Statistics**	**USN:**	**4AL17EC091**
Topic:	**Week - 5**	**Semester & Section:**	**6th Sem 'B' Sec**
Github Repository:	**swastik-gowda**		

FORENOON SESSION DETAILS

Image of session





Report – Report can be typed or hand written for up to two pages.

Sample and sampling:

- By now we know that we can do all kinds of unit variety analyses (e.g., compute modes, means, and standard deviations) and bivariate analyses (e.g., compute Pearson's r correlation coefficients or do regression analyses).
- Usually, all these analyses are fully based on your sample. In general, the methods for analyzing sample data are called methods of descriptive statistics.
- Yet in real life we're often not so much interested in samples but in populations. We therefore often use data obtained from a sample to draw conclusions about an entire underlying population.
- If we employ sample data to draw inferences about a population we are using methods of inferential statistics.
- We use the computed sample statistics to draw inferences about the corresponding population parameters.
- Almost all statistical studies are based on samples. Imagine you want to know to what extent students in London identify themselves as hipsters. It's almost impossible to ask all students.
- So you decide to draw a sample of say, 200 respondents, and to assess to what extent they see themselves as hipsters.
- The great thing about statistics is that it can help you to draw conclusions about all students in London, which is the population, based on the analysis of only these 200 respondents, which is the sample.
- The central question now is what the mean hipsterness score in the wider population is?
- You know the relevant statistic in your example, x-bar equals 3.12, but what you actually want to know is what the mean hipsterness score in the wider population is.
- You want to know, in other words, the value of population parameter mu. Methods of inferential statistics can help us to answer such questions. So if you want to know more about that or about hipsters, watch the videos in this module very carefully.

Sampling:

- Inferential statistics refers to methods used to draw conclusions about a population, based on data coming from a sample.
- You can imagine that in order to understand methods of inferential statistics, it is of essential importance that you know how you should draw samples.
- A sample is nothing more than a subset of a population, yet for methods of inferential statistics, not every sample is appropriate. What you want is a representative samples.
- What you want, in other words, is that your sample is a micro version of the entire population.
- A good way to achieve that is to draw a simple random sample. That means that you make sure that each subject in your population has the same chance of being selected.
- Imagine you want to know to what extent students in London identify themselves as hipsters. You decide to draw a sample of 200 respondents.
- The mean hipsterness score, which could range from 0 to 10, is 3.12. The population here consists of all students in London.
- The parameter you're interested in is the population mean mu. The sample consists of the 200 selected students.

Sampling distribution of sample mean and central limit theorem:

- Researchers often use a sample to draw inferences about a population that sample is from. To do that, they make use of a probability distribution that is very important in the world of statistics.
- Pay close attention because the sampling distribution is the link that helps researchers to draw conclusions about a population on the basis of only one sample.
- We've seen that researchers often use a sample to draw conclusions about the population their sample is from.
- To do so, they make use of a probability distribution that is very important in the world of statistics: the sampling distribution.
- The sampling distribution of the sample mean is the distribution that you get if you draw an infinite number of samples from your population and compute the mean of all the collected sample means.
- The central limit theorem says that, provided that the sample size is sufficiently large, the sampling distribution of the sample mean has an approximately normal distribution.
- The mean of the sampling distribution equals the population mean, and the standard deviation of the sampling distribution equals the standard deviation in the population divided by the square root of the sample size.
- As a researcher you have to carefully distinguish between the sampling distribution, the data distribution and the population distribution.
- Many social, political, and religious groups have their own sacred text. Hipsters, also have an almost holy canonical text. It's a book called On the Road. We are interested in the question, how much time hipsters in New York have spent reading On the Road.
- Assume we know that in the entire population, the mean time hipsters have spent reading the book is 943 minutes. You also know, that the population standard deviation equals 212 minutes.
- The standard deviation in the population, symbolized by sigma, measures the variability of all the individual reading times in the population, around the mean reading time. It equals 212 minutes. The second distribution is the data or sample distribution. It is the distribution of the sample data.

Sampling distribution of sample proportion and example:

- What the sampling distribution of a population proportion looks like. The sampling distribution of a population proportion is bell-shaped if you have at least 15 positive cases and 15 negative cases, and the mean of the distribution is equal to the population proportion. With a rather simple formula we can also obtain the standard deviation of the sampling distribution of the sample proportion.
- You know that 10% of the students in Paris identifies as a hipster, this means that the population proportion, which is symbolized by pi, equals 0.10.
- Now imagine that we draw a sample of n equals 200 from this population. The sample proportion, which is symbolized by p, will be a number close to 0.10. It could be for instance 0.09 or 0.12.