**Mandatory Requirements**

**Prerequisites**

1. Courses to Learn

         Coursera - The Data Scientist’s Toolbox

         Coursera - R Programming

         Coursera - Programming for Everybody (Getting Started with Python)

         Coursera - Statistical Inference

         Coursera - Practical Machine Learning

1. Report of outcome near about 50 pages
2. Presentation on outcome and learning near about 15 slides
3. Student coordinator from every group who will communicate with concerned mentor
4. Faculty coordinator: Prof. Hradesh Kumar

**Company expectation, please read**

As output, the least that is expected to be produced for a certificate to be issued is –

         a Word document detailing “what, why & how?” of the use case, elaborating on the understanding and how the challenge is being addressed, a project report if you will; am expecting this to be no more than 50 pages, and must list, besides usual things, URLs / books / research papers etc. that have been referenced during the execution;

o   Request you to get a template created basis similar experience share it with the teams so that it is consistent across teams;

o   This document is expected to be built progressively over the 6-week period;

         a summary in the form of a presentation with no more than 15 slides which would mainly be based on the Word document mentioned above; this would be shown during the final review;

         Running code, with appropriate visualizations as that will demonstrate the solution developed for the use case;

o   Would suggest using Python or R or Java as these are probably the most popular ones for data sciences and would make the development easier; but feel free to take a call in consultation with the mentors as required;

Pl. use GitHub (<https://github.com/>) or something similar as the repository for storing all the documents & code and ensure that the mentors are given full access to these repositories.