Name of the Programme: MSc Integrated
Course Code: IMC- 503
Title of the Course: Data Science Toolkit
Number of Credits: 4(2L-0T-2P)
Effective from AY: 2022-23

Prerequisites for the course:	Knowledge of data science and data analytics	
Objective:	The aim of this course is to provide an introduction to the main tools and ideas in the data scientist's toolbox.	
Content Theory::	**Excel for Data Visualization**: Predefined, custom number and conditional data format for cells; macros; sorting and filtering data; plotting charts and graphs; working across sheets in excel file; creating interactive dashboards, Pivot table, lookup functions **Numeric and Statistical Computing:** Programming and functions; strings, lists, arrays, matrices and data frames; R packages; working with data (e.g. csv, excel, xml, json); plot graphs and charts; R statistical functions and models **Markdown**: Document structure; basic text formatting; paragraphs; headings; lists; links and images; code blocks; escape characters; HTML elements; converting markdown to html web pages **Source Version Control:** Version Control; introduction to SVN and Git; Git repositories; Git cloning, forks and branches; Git stash; Git pull requests; resolving Git merge conflicts; maintaining your Git pages	6 hours 6 hours 6 hours 6 hours
Content Practical:	**Suggested Lab Assignments** (1) Sample Assignments using Excel (a) Using a provided sample dataset excel file (containing office supplies dats, or food sales), format the columns for different currency (currency unit and thousands' delimiter) based on geo location mentioned (2) Sample Assignments using R (3) Sample Assignments using Markdown (4) Sample Assignments using Git	4 * 12 = 48 hours
Pedagogy:	Lectures/tutorials/practical assignments/self-study	
References/Readings	1. Alexander, Kusleika, Walkenbach, "Excel Bible", Wiley 2. Wickham, Grolemund, "R for Data Science", O'Reilly 3. Matt Cone, "The Markdown Guide" 4. Chacon, Straub, "Pro Git", Apress	
Course Outcomes	1. Create a Github repository 2. Explain essential study design concepts 3. Set up R, R-Studio, Github and other useful tools 4. Understand the data, problems, and tools that data analysts work with.	