CDSA - 101 Foundations of Data Science and Data Analysis Tools

Total Credit- 06 **Max. Marks- 100**

Block 1-

Unit 1- Introduction to Data Science- I

[What is Data Science? Why Data Science? Data Science Components, Data Manipulation and Visualization, etc.]

Unit 2- Introduction to Data Science- II

[Tools for Data Science, Challenges of Data Science Technology, Types of Data Science Jobs, Qualities of a Data Scientist, etc.]

Unit 3- Introduction to Data Science- III

[What is Data? structured and unstructured Data, Data Science & Python, what is Data Analytics? Types of Data Analytics, Role of Data Analytics, Data analytics methods and techniques, what is Big Data? Applications of Big Data.]

Unit 4- Basics of R programming- I

[Features of R, why use R? Applications of R Programming, Alternatives to R programming, Downloading and Installing R, Run R Programming on Your Computer, R's Help System, Understanding Errors.]

Block 2-

Unit 5- Basics of R programming- II

[The Terminal, Working with Vectors, Sub-setting Vectors - the magic "[]", Other Useful Functions, R Syntax, Loops in R]

Unit 6- Basics of Python- I

Why should you learn to write programs? Computer hardware architecture, understanding programming, Conversing with Python, Terminology: Interpreter and compiler, what is a program? writing a program, The building blocks of programs]

Unit 7- Basics of Python- II

Why you should learn Python? How to download and install Python, basic syntax, etc.

Unit 8- Brief Overview of data analysis tools using Python

[Jupyter Notebook, pandas and Bokeh]

Block 3-

Unit 9- Data Science Applications

Real World Applications of Data Science, Some suggestive case studies.

Unit 10 Tools for Data Science

Brief Introduction to data science tools

Unit 11- Mathematical Computations with MS Excel

[Overview of Microsoft Excel, Formatting and Data Analysis, Formulas, Introductory Statistical Functions, Functions for Personal Finance, Preparing to Print]

Unit 12- Formulas, Functions, Logical and Lookup Functions in MS Excel

[More on Formulas and Functions, Logical and Lookup Functions, Conditional Formatting, etc.]

Block 4-

Unit 13- Presenting Data with Charts in MS Excel

[Choosing a Chart Type, Formatting Charts, Using Charts with MS Excel]

Unit 14- Data preparation in MS Excel

[Importing data, data validation, Filtering, Sorting, etc.]

Unit 15- Advance Excel- A Tool for data analysis

Data analysis using MS excel, i. e. Rank and Percentile, Correlation, etc.

Unit 16- Introduction to Database concepts

[Basics of Database, Usefulness in digital marketing strategy, Database models in brief, Sample database designing and rules, basic understanding of ER diagram, Introduction to SQL, frequently asked SQL queries]

Block 5-

Unit 17- Introduction to Statistical Package for the Social Sciences (SPSS)

[Overview, entering data in SPSS, cleaning and checking SPSS database, recording data: collapsing SPSS data, constructing scales and checking their reliability]

Unit 18- Univariate analysis in SPSS

[Descriptive statistics and Graph]

Unit 19- Examining relationships among variables in SPSS

[Cross tabulations, measures of associations and correlation, Chi square test of independence]

Unit 20- Examine differences between groups in SPSS

[T- test, ANOVA: One way]