**Course:** BTech **Semester:** 5

**Prerequisite:** Data analytics and Data analysis, Data visualization techniques and Statistical measures, Basics of Programing Languages, Understanding of Python.

**Course Objective:** Data Analytics helps small and large organizations maximize the value of their data, unearth insights, build plans and respond in real-time to customer demand.

## Teaching and Examination Scheme

| Teaching Scheme | | | | | Examination Scheme | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| Lecture Hrs/Week | Tutorial Hrs/Week | Lab Hrs/Week | Hrs/Week | Credit | Internal Marks | | | External Marks | | |
| | | | | | T | CE | P | T | P | |
| 3 | 0 | 0 | - | 3 | 20 | 20 | - | 60 | - | 100 |

**SEE** - Semester End Examination, **T** - Theory, **P** - Practical

## Course Content

**W** - Weightage (%) , **T** - Teaching hours

| Sr. | Topics | W | T |
|---|---|---|---|
| 1 | **Introduction to Data Analytics:** Introduction, Data and its importance, Data analytics and its types, Why data analytics is important, Data analysis Vs Data analytics, Classification of data analytics, Elements of Data analytics, Data analyst Vs. Data scientist | 25 | 9 |
| 2 | **Introduction to Python Fundamentals and Statistics:** Introduction, Importance of Python, Levels of Data measurement, Central tendency and Dispersion, Distribution of Sample Means, Population and Variance, Confidence interval estimation | 15 | 8 |
| 3 | **Probability and Types of Testing:**Probability and Probability distribution, Sampling and Sampling distribution, Hypothesis testing, Anova test, Chi-square test | 20 | 9 |
| 4 | **Regression, Classification and Clustering:** Linear and Logistic regression, Clustering: K-Means clustering and Hierarchical clustering, Classification: Decision tree, Confusion matrix | 25 | 10 |
| 5 | **Data Visualization Using PowerBI:** Introduction to visualization and analytic tool: Power BI, Getting Data from different sources, data transformations, introduction to data modeling, types of data visualizations in PowerBI, Publishing and sharing reports, Use cases of Dashboard and Analytical Reports Creation. | 15 | 9 |

## Reference Books


| | |
|---|---|
| 1. | **Data Analytics using Python By Bharati Motwani, Wiley Publications. (TextBook)** |
| 2. | **Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.(TextBook) Wiley Publications** |
| 3. | **Statistics 101: From Data Analysis and Predictive Modeling to Measuring Distribution and Determining Probability, Your Essential Guide to Statistics By David Borman, Adams Media** |
| 4. | **Machine Learning, A Probabilistic Approach. By Kevin P. Murphy** |

**Course Outcome**

| **After Learning the Course the students shall be able to:** |
| --- |
| 1. Explain basics of data analytics lifecycle and visualization.<br>2. Compare different analytics techniques and visualization using Python.<br>3. Apply various testing methods and techniques using probability<br>4. Apply different regression, classification, clustering techniques.<br>5. Create an interactive data visualization using PowerBI. |