Sheth T.J. Education Society's
Sheth N.K.T.T College of Commerce &
Sheth J.T.T College of Arts (Autonomous) Thane

T.Y.B.Sc (Data Science) 2026-27

Course Code	Semester V	Credits		Semester VI	Credits
Major (Mandatory)			**Major (Mandatory)**		
BDD501	Data Engineering	2	BDE601	Exploratory Data Analysis	2
BDDP502	Data Engineering Practical	2	BDEP602	Exploratory Data Analysis Practical	2
BDM503	Machine Learning	2	BDD603	Deep Learning	2
BDMP504	Machine Learning Practical	2	BDDP604	Deep Learning Practical	2
BDK505	IKS- A Data Driven Perspective	2	BDK605	Data Security and Compliance	2
Major (Elective)			**Major (Elective)**		
BDR506	Robotic Process Automation	2	BDS606	Sports Analytics	2
			BDSP607	Sports Analytics Practical	2
BDRP507	Robotic Process Automation Practical	2	BDH608	Healthcare Analytics	2
			BDHP609	Healthcare Analytics Practical	2
BDC508	Cloud Computing	2			2
BDCP509	Cloud Computing Practical	2	BDM610	Marketing & Retail Analytics	2
			BDM611	Marketing & Retail Analytics Practical	
BDA510	Applied Business Analytics	2			
BDAP511	Applied Business Analytics Practical	2			
Minor		2	**Minor**		
BDV512	Computer Vision	2	BDI612	Internet of Things	2
BDVP513	Computer Vision Practical		BDIP613	Internet of Things Practical	2
VSC	Data Visualization using		**SEC**		
BDT514	tableau	2		-	
FP			**OJT**		
BDPD515	Project	2	BDPI514	OJT	4
	Total	**22**			**22**

Dr. Yogeshwari Patil
Head of the Department

Prof. H.A. Chande
NEP Coordinator

Dr. Dilip Patil
Principal

Vision: Empowering learners through academic excellence, experiential learning and innovation

SHETH N.K.T.T COLLEGE OF COMM. & SHETH J.T.T COLLEGE OF ARTS AUTONOMOUS THANE

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

<table>
<tr><td colspan="2">Programme Name: B.Sc. (Data Science) Semester: V</td></tr>
<tr><td colspan="2">Course Category: Major Mandatory</td></tr>
<tr><td colspan="2">Name of the Dept: Science & Technology</td></tr>
<tr><td colspan="2">Course Title: Data Engineering</td></tr>
<tr><td colspan="2">Course Type: Theory</td></tr>
<tr><td colspan="2">Course Code: BDD501 Course Level: 5.5</td></tr>
<tr><td colspan="2">Course Credit: 02</td></tr>
<tr><td colspan="2">Hours Allotted: 30 Hours Total Marks: 50</td></tr>
<tr><td colspan="2">Course Objectives:
To understand the role of data engineering in building scalable data systems for analytics and machine learning.
To develop knowledge of data architecture, storage systems, ingestion methods, and data serving techniques.</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of course, learners will be able to
CO1: explain and design basic data engineering workflows including storage, ingestion, and transformation.
CO2: analyze data architecture choices and support analytics and machine learning use cases effectively.</td></tr>
<tr><td>Description of the course:</td><td>Data engineering refers to the building of systems to enable the collection and usage of data. This data is usually used to enable subsequent analysis and data science; which often involves machine learning. Making the data usable usually involves substantial compute and storage, as well as data processing and cleaning.</td></tr>
</table>

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Introduction to Data Engineering and Lifecycle** 1. Introduction to Data Engineering 2. Role of Data Engineers in an Organization 3. Data Engineering Skills and Activities 4. The Data Engineering Lifecycle 5. Key Undercurrents Across the Lifecycle **Data Architecture and Technology Selection** 1. What is Data Architecture? 2. Principles of Good Data Architecture 3. Types of Data Architecture 4. Choosing Technologies Across the Lifecycle	15
II	**Data Generation, Storage and Ingestion** 1. Data Generation in Source Systems 2. Types of Data Sources 3. Data Storage Systems 4. Storage Concepts 5. Data Ingestion **Data Modelling, Transformation and Serving** 1. Data Modelling Basics 2. Data Transformation Concepts 3. Serving Data for Analytics 4. Reverse ETL (Concept) 5. Security and Privacy in Data Engineering	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Fundamentals of Data Engineering	Joe Reis and Matt Housley	O'Reilly Media	First	2022
2.	Learning Spark: Lightning-Fast Data Analytics	Jules S. Damji, Brooke Wenig	O'Reilly Media	Second	2020
3.	Kafka: The Definitive Guide: Real-Time Data and Stream Processing at Scale	Neha Narkhede, Gwen Shapira & Todd Palino	O'Reilly Media	First	2018
4.	Data Engineering with Python	Paul Crickard	Packt Publishing	First	2020

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

<table>
<tr><td colspan="2">Programme Name: B.Sc. (Data Science) Semester: V</td></tr>
<tr><td colspan="2">Course Category: Major Mandatory</td></tr>
<tr><td colspan="2">Name of the Dept: Science & Technology</td></tr>
<tr><td colspan="2">Course Title: Data Engineering Practical</td></tr>
<tr><td colspan="2">Course Code: BDDP502 Course Level: 5.5</td></tr>
<tr><td colspan="2">Course Type: Practical</td></tr>
<tr><td colspan="2">Course Credit: 02</td></tr>
<tr><td colspan="2">Hours Allotted: 60 Hours Total Marks: 50</td></tr>
<tr><td colspan="2">Course Objectives:
• To understand the role of data engineering in building scalable data systems for analytics and machine learning.
• To develop knowledge of data architecture, storage systems, ingestion methods, and data serving techniques.</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of course, learners will be able to
CO1: explain and design basic data engineering workflows including storage, ingestion, and transformation.
CO2: analyze data architecture choices and support analytics and machine learning use cases effectively.</td></tr>
<tr><td>Description of the Course:</td><td>Data engineering refers to the building of systems to enable the collection and usage of data. This data is usually used to enable subsequent analysis and data science; which often involves machine learning. Making the data usable usually involves substantial compute and storage, as well as data processing and cleaning.</td></tr>
</table>

Syllabus: NEP 2020 w.e.f 2026-27

Sr No.	Content (List of Practicals)	Hours
1	**Practical on RDD (Resilient Distributed Dataset) with Scala Operations and transformations.**	
a)	**Use of next mentioned operations and functions:** Parallelize, Read text file, Read CSV, Create RDD, Actions, Pair Functions, Repartition and Coalesce, Shuffle Partitions, Broadcast Variables, Accumulator Variables and Convert RDD to DataFrame	
b)	**Use of next mentioned operations** : Read multiple text files into RDD, Read CSV file into RDD, Create an empty RDD, RDD Pair Functions and Generate DataFrame from RDD	
2	**Practical on the DataFrame operations**	
a)	**Demonstrate the use of next mentioned operations:** Create an empty DataFrame, Create an empty DataSet, use of Rename nested column, Adding or Updating a column on DataFrame, drop a column on DataFrame, Adding literal constant to DataFrame, changing column data type, Pivot and Unpivot a DataFrame, Create a DataFrame using StructType & StructField schema	
b)	**Use of next mentioned operations:** Selecting the first row of each group, Sort DataFrame, Union DataFrame, Drop Rows with null values from DataFrame, Split single to multiple columns, Concatenate multiple columns, Replace null values in DataFrame, Remove duplicate rows on DataFrame, Remove distinct on multiple selected columns, Spark UDF	
3	**Practical on the Spark Array and Map operations**	
a)	**Use of next mentioned operations:** Create an Array (ArrayType) column on DataFrame, Create a Map (MapType) column on DataFrame, Convert an Array to columns, Create an Array of struct column, Explode an Array and map columns, Explode an Array of structs, Explode an Array of map columns to rows	
b)	**Use of next mentioned operations:** Create a DataFrame with nested Array, Explode nested Arrays to rows, Flatten nested Array to single Array, Convert array of String to a String column	
4	**Spark Aggregate :** Group rows in DataFrame, Get Count distinct on DataFrame, Add row number to DataFrame, Select the first row of each group	
5	**Spark SQL Joins, Spark SQL Schema, StructType & SQL Functions**	
a)	**Use of next mentioned operations:** Use of Spark SQL Join, Join multiple DataFrames, Inner join two tables/DataFrame, Self join, Join tables on	

	multiple columns, Convert case class to a schema, Create array of struct column, Flatten nested column	
b)	**Use of next mentioned functions** : Date and Time Functions, String Functions, Array Functions, Map Functions, Aggregate Functions, Window Functions, Sort Functions, JSON Functions	
6	**Spark SQL :** Demonstrate the use of next functions : createDataFrame(), where() & filter(), withColumn(), withColumnRenamed(), drop(), distinct(), groupBy(), join(), map() vs mapPartitions(), foreach() vs foreachPartition(), pivot(), union(), collect(), cache() & persist(), udf()	
7	**Spark Data Source API**	
a)	**Use of next operations :**Process JSON from a Text file, Read & Write CSV file, Read and Write JSON file, Read & Write Parquet file, Read & Write XML file, Read & Write Avro files	
8	**Practical of Spark Streaming**	
a)	**Use of next operations :** OutputModes Append vs Complete vs Update, Read JSON Files From Directory with Scala Example, Read data From TCP Socket with Scala Example, Consuming & Producing Kafka messages in JSON format	
b)	**Use of next operations :**Consuming & Producing Kafka messages in Avro format, from_avro and to_avro functions, Avro data from Kafka topic using from_avro() and to_avro(), Batch Processing using Kafka Data Source	
9	**Spark MLlib : Demonstrate use of Estimator, Transformer, and Param**	
10	**Spark HDFS : Demonstrate the use of next operations :**Processing files from Hadoop HDFS (TEXT, CSV, Parquet, Avro, JSON), Processing TEXT files from Amazon S3 bucket, Processing JSON files from Amazon S3 bucket, Processing CSV files from Amazon S3 bucket, Processing Parquet files from Amazon S3 bucket, Processing Avro files from Amazon S3 bucket	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester V**
Course Category: **Major Mandatory**	
Name of the Dept: **Science & Technology**	
Course Title: **Machine Learning**	
Course Code: BDM503	Course Level:5.5
Type: **Theory**	
Course Credit: 02	
Hours Allotted: 30 Hours	
Total Marks: 50 Marks	

Course Objectives:

- To understand the fundamental concepts of Machine Learning
- To study and apply various Machine Learning algorithms
- To evaluate and improve Machine Learning models

Course Outcomes:

After the completion of course learners will be able to

CO1: understand fundamental Machine Learning concepts, implement various algorithms and clustering techniques for solving data-driven problems.

CO2: evaluate and improve Machine Learning models using performance metrics, dimensionality reduction techniques like PCA, and association rule mining methods.

Description of the course:	This course introduces the core concepts and techniques of Machine Learning, focusing on supervised and unsupervised learning methods. It covers key algorithms such as KNN, Decision Trees, Naïve Bayes, SVM, regression, and clustering techniques like K-means and hierarchical clustering. The course also includes model evaluation metrics, dimensionality reduction using PCA, and association rule mining for pattern discovery in data.

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Introduction to Machine Learning:** Machine Learning(ML), Need for Machine Learning, ML from Knowledge-driven to Data Driven, Applications of Machine Learning, Problems suitable for Machine Learning, Advantages, Disadvantages and Challenges of Machine Learning, Challenges of ML. General architecture of ML systems, **Underlying Concepts in Machine Learning**: Inductive Learning, Generalization, Bias and Variance, Overfitting and Underfitting, Parametric and Non Parametric algorithms . **Types of Machine Learning:** Supervised and Unsupervised Learning, Workflow, Semi supervised Learning, Reinforced Learning, **Introduction to Classification Algorithms:** Concept of Classification, Binary classification, Multi-Class Classification, Multi Label Classification **K-Nearest Neighbor Method:** need and Working of KNN, Computing Distance, Pros and Cons of KNN. **Decision Tree based Algorithm:** Terminologies and assumptions, Working of Decision Trees, ID3 Algorithm, Attribute selection Methods(Entropy, Gini Impurity, Information Gain)	15
II	**Performance Evaluation:** Classification Metrics-Accuracy, Sensitivity, Precision, F1 Score, ROC/AUC Curve, Cross Validation **Dimensionality Reduction:** Crieteria for Reduction, Feature Reduction and Selection, Principal Component Analysis(PCA) **Probabilistic Learning:** Introduction to Bayes Learning, Interpretation of Bayes Rule, Benefits and shortfalls of Bayesian Learning, Naïve Bayes Classifier, Characteristics of Naïve Bayes **Support Vector Machines:** Working of SVM, SVM Concepts :Support Vectors, Hard Margin, soft Margin, Kernels, Advantages and Disadvantages of SVM **Regression Methods:** Linear Regression Models, Logistic Regression **Unsupervised Learning:** Concept of unsupervised Learning, Importance and Challenges of unsupervised Learning, Clustering and its applications **Hierarchical Clustering:** Introduction, Types of Hierarchical Clustering, Issues with Hierarchical Clustering **Partition algorithm:** K-means Clustering, steps of K-means Clustering, Issues, Strength and Weakness of K-means clustering. **Association rule Mining:** Basic Concepts, Market Basket Analysis, and Apriori algorithm.	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Machine Learning: Concepts, Techniques and Applications	T V Geetha S, Sendhilkumar	CRC Press, Taylor and Francis	First	2023
2.	Machine Learning for Decision Sciences with Case Studies in Python	S. Sumathi, Suresh V. Rajappa	CRC Press, Taylor and Francis	First	2022
3.	Introduction to Machine Learning with Python	Andreas C. Müller, Sarah Guido	O'Reilly Media, Inc.	First	2016
4.	Machine Learning	S Sridhar	Oxford University Press	First	2021
5.	Machine Learning	Ruchi Doshi, Kamal Kant Hiran	BPB	First	2021

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category: **Major Mandatory**	
Name of the Dept: **Science & Technology**	
Course Title: **Machine Learning Practical**	
Course Code: BDMP504	Course Level:5.5
Type: Practical	
Course Credit:02	
Hours Allotted: 60 Hours	
Total Marks: 50 Marks	

Course Objectives:

- To implement and analyze Machine Learning algorithms and regression models for classification and prediction tasks.
- To apply clustering and dimensionality reduction techniques and PCA for data analysis and pattern discover and evaluate the performance of Machine Learning models using metrics.

Course Outcomes: After the completion of the course learners will be able to:
CO1: implement and apply Machine Learning algorithms and regression models to solve classification and prediction problems.
CO2: apply clustering and dimensionality reduction techniques and PCA for data analysis and evaluate Machine Learning models using performance metrics

Description of the course:	This course focuses on implementing Machine Learning algorithms such as KNN, Decision Trees, Naïve Bayes, SVM, regression, clustering methods, and evaluating models using performance metrics and PCA.

Sr. No.	Content (List of Practicals)	Hours
1.	Implementing a K-Nearest Neighbor (KNN) algorithm (e.g. to classify handwritten digits).	
2	Building a decision tree model using the ID3 algorithm (e.g. to predict whether a customer will churn or not).	
3.	Developing a Support Vector Machine (SVM) model (e.g. to classify email messages as spam or not spam).	
4.	Building a Naïve Bayes classifier (e.g. to classify movie reviews as positive or negative sentiment).	
5.	Implementing linear regression (e.g. to predict housing prices based on features such as size and location).	
6.	Using logistic regression (e.g. to predict whether a credit card transaction is fraudulent or not).	
7.	Evaluating a classification model using metrics such as accuracy, precision, recall, and F1 score.	
8.	Applying hierarchical clustering (e.g. to group customer segments based on their purchasing behavior).	
9.	Implementing the K-means clustering algorithm (e.g. to identify distinct clusters in a customer demographic dataset).	
10.	Utilizing Principal Component Analysis (PCA) for dimensionality reduction to improve the efficiency and interpretability of a model.	
	Total Hours	60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category: **Major Mandatory IKS**	
Name of the Dept: **Science & Technology**	
Course Title: **A Data-Driven Perspective**	
Course Code: BDK505	Course Level:5.5
Course Type: Theory	
Hours Allotted: 30 Hours	
Course Credit: 02	Total Marks: 50

Course Objectives:

- To analyze core methods of knowledge generation in the Indian Knowledge System and explore contributions from Nyāya, Vaiśeṣika, Vedic mathematics, and ancient computational traditions in modern data science.
- To integrate IKS-informed perspectives into data ethics and responsible AI while developing culturally rooted, multidisciplinary problem-solving skills for domain-specific innovation.

Course Outcomes: After the completion of the course learners will be able to:

CO1: apply classical Indian logical, mathematical, and algorithmic principles along with ethical values from Indian Knowledge Systems in data science practices.

CO2: design and develop culturally grounded, multidisciplinary data science solutions by integrating traditional Indian knowledge with modern approaches to address real-world challenges.

Description of the course:	Indian Knowledge Systems (IKS) explores India's interdisciplinary traditions integrating philosophy, mathematics, linguistics, medicine, and astronomy with data-centric thinking. It examines ancient concepts of observation, classification, algorithms, prediction, and ethics, linking classical texts and methods—such as zero, combinatorics, grammar, and astronomy—to modern data science, artificial intelligence, knowledge organization, and responsible, value-driven research.

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Foundations and Mathematical Traditions in Indian Knowledge Systems (IKS):** **Introduction to Indian Knowledge Systems:** Concept, scope, and characteristics of IKS, Interdisciplinary nature: integration of philosophy, mathematics, linguistics, astronomy, and medicine, **Knowledge, Data, and Observation in Indian Traditions:** Concept of knowledge (Pramāṇa) and its types ,Role of observation, inference, and empirical methods in knowledge creation **Sources of Structured Knowledge in Classical Texts:** Vedic and post-Vedic literature, Mathematical, astronomical, and linguistic treatises as organized knowledge systems. **Mathematical Foundations for Data Science:** Development of decimal place-value system, Concept of zero and its computational significance **Combinatorics and Early Computational Thinking:** Binary structures and enumeration in Pingala's Chandas Shastra , Recursive and combinatorial approaches. **Algorithmic Traditions in Indian Mathematics:** Computational methods of Aryabhata, Brahmagupta, and Bhaskara , Iterative and problem-solving techniques. **Relevance of IKS to Modern Data Science:** Connections between traditional methods and modern computational thinking	15
II	**Knowledge Organization, Analytics, and Applications of IKS in Data Science:** **Classification and Ontology in Indian Traditions:** Nyaya and Vaisheshika systems: categorization of reality, Ayurvedic classification: Doshas, diseases, and medicinal systems. **Pattern Recognition and Predictive Models:** Astronomical observations: cycles, periodicity, and pattern identification , Predictive techniques in works of Aryabhata and other scholars. **Analytical Reasoning and Inference:** Logical reasoning and inference in Indian philosophy (Nyaya) , Application of reasoning in knowledge validation, **Applications in Data Science and Digital Humanities:** Text analytics and computational study of Sanskrit literature, Digital tools for analysis of classical texts. **Knowledge Representation and Modern Technologies:** Ontologies, taxonomies, and knowledge graphs ,AI and machine learning applications in IKS. **Ethical Perspectives in Data Science:** Concepts of Dharma, balance, and responsibility , Ethical decision-making in data-driven technologies	15
	Total Hours	30

Reference Books

- Mahadevan, B., Bhat, Vinayak Rajat, and Nagendra Pavana R. Introduction to Indian Knowledge System: Concepts and Applications. PHI Learning, 2022.
- Kapil Kapoor and Avadesh Kumar Singh. Indian Knowledge Systems. D.K. Printworld.
- Kak, Subhash and Rao, T. R. N. Computing Science in Ancient India. Center for Advanced Computer Studies, University of Southwestern Louisiana, 1998.
- Bhanu Murthy, T. S. A Modern Introduction to Ancient Indian Mathematics. New Age International, 1992.
- Kim Plofker. Mathematics in India. Princeton University Press, 2009.
- C. K. Raju. Cultural Foundations of Mathematics. Pearson Education.
- Briggs, Rick. Knowledge Representation in Sanskrit and Artificial Intelligence. AI Magazine, 1985.

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

<table>
<tr><td colspan="2">Programme Name: B.Sc. (Data Science) Semester: V</td></tr>
<tr><td colspan="2">Course Category: Major Elective</td></tr>
<tr><td colspan="2">Name of the Dept: Science & Technology</td></tr>
<tr><td colspan="2">Course Title: Robotic Process Automation</td></tr>
<tr><td colspan="2">Course Code: BDR506 Course Level:5.5</td></tr>
<tr><td colspan="2">Type: Theory</td></tr>
<tr><td colspan="2">Course Credit: 2</td></tr>
<tr><td colspan="2">Hours Allotted: 30 Hours</td></tr>
<tr><td colspan="2">Total Marks: 50 Marks</td></tr>
<tr><td colspan="2">Course Objectives:
1. To learn the foundational knowledge of RPA, automation techniques, and the UiPath environment and build efficient automation solutions and UI automation techniques
2. implement exception handling, debugging, publishing, and orchestration to deploy and maintain reliable automation solutions</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of this course the learners will be able to
CO1: design and develop end-to-end automation workflows and OCR techniques.
CO2: deploy, and maintain event-driven RPA solutions, incorporating assistant bots, trigger-based automation, exception handling, debugging, logging, workflow optimization, and orchestration for effective bot management and monitoring.</td></tr>
<tr><td>Description of the course:</td><td>Robotic Process Automation (RPA) is a technology that uses software "bots" to automate repetitive, rule-based tasks by mimicking how humans interact with applications through their graphical user interfaces (GUIs). Unlike traditional automation, which relies on APIs and back-end programming, RPA observes and records user actions and then replicates them directly on the screen.
This makes it easier to automate
processes in systems that do not provide built-in automation capabilities.</td></tr>
</table>

Syllabus: NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Robotic Process Automation:** Scope and techniques of automation, About UiPath **Record and Play:** UiPath stack, Downloading and installing UiPath Studio, Learning UiPath Studio, Task recorder, Step-by-step examples using the recorder. **Sequence, Flowchart, and Control Flow:** Sequencing the workflow, Activities, Control flow, various types of loops, and decision making, Step by-step example using Sequence and Flowchart, Step-by-step example using Sequence and Control flow **Data Manipulation**: Variables and scope, Collections, Arguments – Purpose and use, Data table usage with examples File operation with step-by-step example, CSV/Excel to data table and vice versa (with a step-by-step example). **Taking Control of the Controls** : Finding and attaching windows, Finding the control, Techniques for waiting for a control, Act on controls – mouse and keyboard activities, Working with UiExplorer, Handling events, Revisit recorder, Screen Scraping, When to use OCR, Types of OCR available, How to use OCR	15
II	**Handling User Events and Assistant Bots**: What are assistant bots?, Monitoring system event triggers, Hotkey trigger, Mouse trigger, System trigger ,Monitoring image and element triggers, An example of monitoring email, Example of monitoring a copying event and blocking it, Launching an assistant bot on a keyboard event **Exception Handling, Debugging, and Logging**: Exception handling, Common exceptions and ways to handle them, Logging and taking screenshots, Debugging techniques, Collecting crash dumps, Error reporting **Managing and Maintaining the Code**: Project organization, Nesting workflows, Reusability of workflows, Commenting techniques, State Machine, When to use Flowcharts, State Machines, or Sequences, **Deploying and Maintaining the Bot**: Publishing using publish utility, Overview of Orchestration Server, Using Orchestration Server to control bots, Using Orchestration Server to deploy bots, License management, Publishing and managing updates	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Learning Robotic Process Automation .	Alok Mani Tripathi	Packt	First	2018
2.	Robotic Process Automation Tools, Process Automation and their benefits: Understanding RPA and Intelligent Automation	Srikanth Merianda	Createspace Independent Publishing	First	2018
3.	The Simple Implementation Guide to Robotic Process Automation (Rpa): How to Best Implement Rpa in an Organization	Kelly Wibbenmeyer	iUniverse	First	2018

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category: **Major Elective**	
Name of the Dept: **Science & Technology**	
Course Title: **Robotic Process Automation Practical**	
Course Code: BDRP507	Course Level:5.5
Type: Practical	
Course Credit: 02	
Hours Allotted: 60 Hours	
Total Marks: 50 Marks	
Course Objectives: **1.** To design and develop automation workflows, use the drag-and-drop interface, activities, variables, and conditions to create efficient and robust automation solutions. **2.** To develop advanced features and capabilities of UiPath and gain exposure to a wide range of UiPath functionalities to handle complex automation scenarios.	
Course Outcomes: After the completion of course the learners will be able to **CO1:** design and implement efficient automation workflows using UiPath Studio by applying activities, variables, conditions, and control flow to solve real-world business problems. **CO2:** develop advanced automation solutions using features and API integration to handle complex and dynamic automation scenarios.	
Description of the course:	The course will provides hands-on training in designing, developing, and deploying automation solutions using UiPath Studio. Students gain practical experience in workflow creation, control flow implementation, data manipulation, file and Excel automation, UI interaction, screen scraping, OCR, event handling, exception management, debugging, and bot deployment. The practical sessions focus on real-world automation scenarios to build industry-ready RPA development skills.

Syllabus : NEP 2020 w.e.f 2026-27

Sr. No.	Content (List of Practicals)	Hours
1.	**RPA Basics: Sequences and Flowcharts**	
a.	Create a simple sequence-based project.	
b.	Create a flowchart-based project.	
c.	Automate UiPath Number Calculation (Subtraction, Multiplication, Division of numbers).	
d.	Create an automation UiPath project using different types of variables (number, datetime, Boolean, generic, array, data table)	
2	**Decision making and looping**	
a.	Consider an array of names. We have to find out how many of them start with the letter "a". Create an automation where the number of names starting with "a" is. counted and the result is displayed.	
b.	Demonstrate switch statement with an example.	
c.	Create an automation to demonstrate use of decision statements (if)	
d.	Create an automation To Print numbers from 1 to 10 with break after the write line activity inside for each activity .	
3.	**Types of Recording**	
a.	Desktop Recording using Tool bar	
b.	Desktop Recording by creating a workflow	
c.	Web Recording e.g. Find the rating of the movie from imdb web site	
4.	**Excel Automation**	
a.	Automate the process to extract data from an excel file into a data table and vice versa	
b.	Create an automation To Write data to specific cell of an excel sheet.	
c.	Create an automation To Read data to specific cell of an excel sheet	
d.	Create an automation To append data to specific cell of an excel sheet.	
e.	Create an automation To sort a table of an excel sheet.	
5.	**Different controls in UiPath**	
a.	Implement the attach window activity.	
b.	Automate using Element Exists.	

c.	Automate using Find Children control.	
6.	**Keyboard and Mouse Events**	
a.	Demonstrate the following activities in UiPath: i. Mouse (click, double click and hover) ii. Type into iii. Type Secure text	
b.	Demonstrate the following events in UiPath: i. Element triggering event ii. Image triggering event iii. System Triggering Event	
c.	Automate the process of launching an assistant bot on a keyboard event.	
7.	**Screen Scraping and Web Scraping methods**	
a.	Automate the following screen scraping methods using UiPath: i) Full Text ii) Native iii) OCR	
b.	Demonstrate Data Scraping and display values in Message box.	
c.	Demonstrate Screen Scraping for a pdf, web page and image file.	
8.	**PDF Automation and Exception Handling**	
a.	Read PDF With OCR	
b.	Merge PDF's into one	
c.	Get PDF Total Page count Using Regex	
d.	Demonstrate Exception Handling using UiPath	
9.	**Email Automation**	
a.	Read Emails	
b.	Send Email with Attachment	
c.	Reply to Email	
10.	**Orchestrator management and mini project**	
a.	Deploy bots to Orchestrator	
b.	Run jobs from Orchestrator	
c.	Queue Introduction: i. Add items to Queue. ii. Get Queue item from Orchestrator	

d.	Build UiPath Chatbot using Google dialogflow	
	Total Hours	60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category/Vertical: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title**: Cloud Computing**	
Course Code: BDC508	Course Level:5.5
Type: **Theory**	
Course Credit: 2 credits	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

- To Analyze cloud computing architectures, vulnerabilities, and applications, and design workflows using the MapReduce programming model.
- To Apply virtualization concepts, cloud resource management, scheduling algorithms, cloud storage systems, and security measures to develop secure cloud applications while understanding their legal and societal impact.

Course Outcomes: After the completion of course, learners will be able to

CO1: understand and use cloud services, virtualization, task scheduling algorithms, and the MapReduce programming model for cloud computing applications.

CO2: develop and manage private cloud systems while understanding the legal and societal impact of cloud computing technologies.

Description of the course:	Cloud Computing syllabus introduces the basic concepts and technologies used to deliver computing services over the internet. It covers topics such as cloud architecture, service models (IaaS, PaaS, SaaS), deployment models (public, private, hybrid), virtualization, cloud storage, and security. The syllabus also explains benefits, challenges, and applications of cloud computing in modern IT systems.

Syllabus : NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Introduction to Cloud Computing –** Introduction, Historical developments, Building Cloud Computing Environments. **Principles of Parallel and Distributed Computing –** Eras of Computing, Parallel v/s distributed computing, Elements of Parallel Computing, Elements of distributed computing, Technologies for distributed computing. **Virtualization –** Introduction, Characteristics of virtualized environments, Taxonomy of virtualization techniques, Virtualization and cloud computing, Pros and cons of virtualization, Technology examples. Logical Network Perimeter, Virtual Server, Cloud Storage Device, Cloud usage monitor, Resource replication, Ready-made environment.	15
II	**Cloud Computing Architecture:** Introduction, Fundamental concepts and models, Roles and boundaries, Cloud Characteristics, Cloud Delivery models, Cloud Deployment models, Economics of the cloud, Open challenges. **Fundamental Cloud Security:** Basics, Threat agents, Cloud security threats, additional considerations. **Industrial Platforms and New Developments:** Amazon Web Services, Google App Engine, Microsoft Azure. Distributed software engineering, Distributed systems issues, Client-server computing, Architectural patterns for distributed systems, Software as a service	15
	Total Hours	**30**

References:

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1.	Mastering Cloud Computing Foundations and Applications Programming	Rajkumar Buyya, Christian Vecchiola, S. Thamarai Selvi	Elsevier	-	2013
2.	Cloud Computing Concepts, Technology & Architecture	Thomas Erl, Zaigham Mahmood , and Ricardo Puttini	Prentice Hall	-	2013
3.	Distributed and Cloud Computing, From Parallel	Kai Hwang, Jack Dongarra, Geoffrey Fox	MK Publishers	--	2012

	Processing to the Internet of Things				

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category/Vertical: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title**:** **Cloud Computing Practical**	
Course Code: BDCP509	Course Level:5.5
Type: Practical	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. To understand the basic concepts and services of cloud computing. 2. To gain hands-on experience in developing and deploying cloud-based applications and web services.	
Course Outcomes: After the completion of course, learners will be able to CO1: develop and deploy basic cloud-based applications and web services. CO 2: use cloud platforms and virtualization tools for storage and computing services.	
Description of the course:	Cloud Computing practical provides hands-on experience in creating and managing cloud environments. Students learn virtualization, cloud platforms, clustering, and deployment of applications on cloud services. These practical's help in understanding how cloud infrastructure, virtual machines, and distributed systems work in real-world applications.

Syllabus ; NEP 2020 w.e.f 2026-27

Sr No	Content (List of Practicals)	
1	Implement Windows / Linux Cluster.	
2	Developing application for windows Azure.	
3	Implementing private cloud with Xen Server.	
4	Implementing operating system on Xen Server.	
5	Implement Hadoop .	
6	Develop application using GAE.	
7	Implement VMWARE ESXI server.	
8	Native virtualization using Hyper V.	
9	Implementing operating system on Hyper V.	
10	Using OpenNebula to manage heterogeneous distributed data center infrastructure.	
	Total Hours	60

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

Programme Name: **B.Sc (Data Science)**	**Semester: V**
Course Category/Vertical: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title**: Applied Business Analytics**	
Course Code: BDA510	Course Level:5.5
Type: **Theory**	
Course Credit: 2 credits	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. To learn basics of statistical concepts and Business Intelligence Tools for Data Analysis.
2. To learn about the business analytics methods and Regression Analysis with Time Series Analysis and forecasting
3. To learn to understand various modelling techniques for Optimization and simulation.

Course Outcomes:

After the completion of course learners will be able to

CO1. understand basics of statistical concepts and experiment with Business Intelligence Tools for Data Analysis.

CO2. apply Regression Analysis with Time Series Analysis and forecasting and construct various modelling techniques for Optimization and simulation.

Description of the course:	The applied business analytics course is designed for students keen to advance their knowledge on data analytics. The dynamic curriculum in applied business analytics helps the participant to understand and apply analytics in the business context. Topics covered are fundamentals of probabilities, linear regression, and other business data analysis using Microsoft excel. Business cases from marketing analytics, HR analytics, financial analytics and operation analytics are covered as part of this course.

Syllabus : NEP 2020 w.e.f 2026-27

Unit No.	Content	Hours
I	**Introduction to business analytics -** Meaning and importance of Business Analytics, Types of Analytics: Descriptive, Predictive, Prescriptive, Role of data in business decision-making, Applications of analytics in business, describing the distribution of a variable, finding relationships among variables **Business Intelligence Tools** for Data Analysis, Probability and Probability Distributions, Decision making under uncertainty. **Regression Analysis -** Estimating Relationships, Regression Analysis: Statistical Inference, Time Series Analysis and forecasting	15
II	**Optimization modelling -** Introduction to Optimization Models, Types of Optimization Problems, Decision Variables, Objective Function, and Constraints, Linear Optimization Models, Integer Programming Concepts, Applications of Optimization in Business (resource allocation, cost minimization, profit maximization) **Simulation modelling -** Introduction to Simulation, Types of Simulation Models, Steps in Simulation Modelling, Monte Carlo Simulation, Random Number Generation, Applications of Simulation in Business Decision Making.	15
	Total Hours	**30**

References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1.	Business Analytics Data Analysis and decision making	S.Christian Albright and Wayne L. Winston	Cengage	7th	2020
2.	The Applied Business Analytics Casebook: Applications in Supply Chain Management, Operations Management, and Operations Research	Matthew Drake	O'Reilly	3rd	2013
3.	Applied Business Analytics: Integrating Business Process, Big Data, and Advanced Analytics	Nathaniel Lin	FT Press Analytics	1st	2015

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title**: Applied Business Analytics Practical**	
Course Code: BDAP511	Course Level:5.5
Type: Practical	
Course Credit: 02	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. To Acquire practical skills in data collection, data exploration, and descriptive statistics.
2. Apply statistical methods to analyze relationships between variables and make informed business decisions.
3. Utilize business intelligence tools and programming languages for data analysis and visualization.
4. Gain hands-on experience in applying various analytics techniques such as regression analysis, time series analysis, optimization modeling, and experimental design to solve business problems. .

Course Outcomes:

After the completion of course learners will be able to

CO1. analyze relationships among variables for informed decision-making and Utilize business intelligence tools for data analysis and visualization.

CO2. conduct statistical inference for making meaningful conclusions and apply advanced analytics techniques to solve complex business problems.

Description of the course:	This course " provides students with practical skills in using analytics techniques to solve real-world business problems. The course covers various topics, including data analysis, probability, hypothesis testing, regression analysis, time series analysis, optimization modelling, and experimental design. Students will learn how to collect, analyse, and interpret data to make informed business decisions. They will also gain proficiency in using business intelligence tools and programming languages such as Python for data analysis and visualization.

Sr. No.	Content (List of Practicals)	Hours
1	Introduction to Business Analytics: a. Collect data from a real-life business scenario and perform exploratory data analysis (EDA) to gain insights into the dataset. b. Analyze customer data to identify trends and patterns that can be used for business decision-making.	
2	Describing the Distribution of a Variable: a. Obtain a dataset and calculate descriptive statistics (mean, median, mode, variance, etc.) for a specific variable of interest. b. Create visualizations (histograms, box plots) to depict the distribution of a variable and analyze its characteristics	
3.	Finding Relationships Among Variables: a. Use a dataset with multiple variables and perform correlation analysis to determine the strength and direction of relationships between pairs of variables. b. Apply regression analysis to identify the relationship between an independent variable (e.g., advertising expenditure) and a dependent variable (e.g., sales revenue).	
4.	Business Intelligence Tools for Data Analysis: a. Utilize a business intelligence tool (e.g., Tableau, Power BI) to extract insights from a dataset and create interactive visualizations for effective data analysis.	
5.	Probability and Probability Distributions: a. Simulate a probability experiment (e.g., rolling dice) using programming and calculate the probabilities of different outcomes. b. Generate random numbers from various probability distributions (normal, uniform, exponential) and analyze their properties.	
6.	Decision Making under Uncertainty: a. Develop a decision tree model to make business decisions considering uncertainties and associated probabilities at each decision point. b. Apply the concept of expected value to evaluate different decision alternatives and select the optimal one.	
7.	Sampling and Sampling Distributions: a. Conduct a survey and collect data from a sample population, ensuring proper sampling techniques are employed. b. Use the Central Limit Theorem to analyze the sampling distribution of a sample mean and estimate population parameters.	
8.	Confidence Interval Estimation: a. Calculate confidence intervals for population means or proportions using sample data and interpret the results in a business context. b. Apply bootstrapping techniques to estimate confidence intervals for non-parametric statistics.	
9.	Hypothesis Testing: a. Formulate null and alternative hypotheses related to a business problem, conduct a hypothesis test using appropriate statistical tests, and interpret the results. b. Perform A/B testing on a website or marketing campaign to evaluate the effectiveness of different strategies and make data-driven decisions.	

10.	Regression Analysis and Time Series Analysis: a. Develop a regression model to predict future sales based on historical data, assess model performance, and interpret the significance of predictor variables. b. Apply time series analysis techniques (e.g., ARIMA, exponential smoothing) to forecast future demand for a product or service, and evaluate the accuracy of the forecasts.	
11.	Optimization Modeling and Simulation Modeling: a. Formulate an optimization model (e.g., linear programming, integer programming) to solve a real-world business problem and analyze the optimal solution. b. Use simulation modeling to evaluate different business scenarios, such as capacity planning, inventory management, or pricing strategies, and assess their impact on performance metrics.	
12.	Analysis of Variance and Experimental Design: a. Design and conduct an experiment to study the effects of different factors on a specific response variable, analyze the results using analysis of variance (ANOVA), and draw conclusions. b. Implement a factorial experiment and analyze the main effects and interaction effects of factors using statistical techniques.	
	Total Hours	**60**

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

Programme Name: **B.Sc. (Data Science)**	**Semester: V**
Course Category/Vertical: **Minor**	
Name of the Dept: **Science and Technology**	
Course Title**: Computer Vision**	
Course Code: BDV512	Course Level:5.5
Type: **Theory**	
Course Credit: 2 credits	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	
Course Objectives: 1. To introduce the fundamental concepts of image formation, image processing, and feature extraction used in computer vision systems. 2. To develop understanding of pattern recognition techniques, machine learning models, and practical computer vision applications such as object detection, medical imaging, and gesture recognition.	
Course Outcomes: After the completion of the course, the learners will be able to: **CO1:** apply image processing techniques to analyze digital images. **CO2:** implement computer vision models including pattern recognition methods, clustering, and neural networks for solving real-world vision problems.	
Description of the course:	This course introduces the fundamental concepts of computer vision and image processing. It covers topics such as image formation, image filtering, segmentation, feature extraction, and pattern recognition techniques. Students will learn how images are analyzed and interpreted using computational methods. The course also highlights machine learning approaches and real-world applications such as medical imaging, object tracking, and face recognition.

Unit No.	Content	Hours
I	**Image Formation and Image Processing :** Introduction to Computer Vision and Basic Concepts of Image Formation, Introduction and Goals of Computer Vision, Image Formation and Radiometry, Geometric Transformation, Geometric Camera Models, Image Reconstruction from a Series of Projections **Image Processing Concepts:** Fundamentals of Image Processing, Image Transforms, Image Filtering, Colour Image Processing, Mathematical Morphology, Image Segmentation **Image Features :** Image Descriptors and Features, Texture Descriptors, Colour Features, Edge Detection, Object Boundary and Shape Representations, Interest or Corner Point Detectors, Histogram of Oriented Gradients, Scale Invariant Feature Transform	15
II	**Recognition:** Fundamental Pattern Recognition Concepts, Introduction to Pattern Recognition, Linear Regression, Basic Concepts of Decision Functions, Elementary Statistical Decision Theory, Gaussian Classier, Parameter Estimation, Clustering for Knowledge Representation, Dimension Reduction, Template Matching, Artificial Neural Network for Pattern Classification, Convolutional Neural Networks **Applications of Computer Vision:** Applications in Medical, Image Segmentation, Motion, Estimation and Object Tracking, Face and Facial Expression, Recognition, Gesture Recognition, Image Fusion	15
	Total Hours	**30**

References:

1. Computer Vision and Image Processing, Bhuyan, M. K. (2020). Computer vision and image processing. CRC Press.
2. Computer Vision: A Modern Approach, Forsyth, D. A., & Ponce, J. (2012). Computer vision: A modern approach (2nd ed.). Pearson.
3. Machine Vision, Jain, R., Kasturi, R., & Schunk, B. G. (1995). Machine vision. McGraw-Hill.
4. Image Processing, Analysis, and Machine Vision, Sonka, M., Hlavac, V., & Boyle, R. (2007). Image processing, analysis, and machine vision (3rd ed.). Thomson Learning.
5. Computer and Robot Vision, Haralick, R. M., & Shapiro, L. G. (1993). Computer and robot vision. Addison-Wesley.

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous).Thane (W)

<table>
<tr><td colspan="2">Programme Name: B.Sc. (Data Science) Semester: V</td></tr>
<tr><td colspan="2">Course Category: Minor</td></tr>
<tr><td colspan="2">Name of the Dept: Science and Technology</td></tr>
<tr><td colspan="2">Course Title: Computer Vision Practical</td></tr>
<tr><td colspan="2">Course Code: BDVP513 Course Level:5.5</td></tr>
<tr><td colspan="2">Type: Practical</td></tr>
<tr><td colspan="2">Course Credit: 02</td></tr>
<tr><td colspan="2">Hours Allotted: 60 Hours</td></tr>
<tr><td colspan="2">Marks Allotted: 50 Marks</td></tr>
<tr><td colspan="2">Course Objectives:
To introduce the fundamental concepts of image formation, image processing, and feature extraction used in computer vision systems.
To develop understanding of pattern recognition techniques, machine learning models, and practical computer vision applications such as object detection, medical imaging, and gesture recognition.</td></tr>
<tr><td colspan="2">Course Outcomes:
After the completion of the course, the learners will be able to:
CO1: apply image processing techniques such as filtering, segmentation, feature extraction, and geometric transformations to analyze digital images.
CO2: analyze and implement computer vision models including pattern recognition methods, clustering, and neural networks for solving real-world vision problems.</td></tr>
<tr><td>Description of the course:</td><td>This course introduces the fundamental concepts of computer vision and image processing. It covers topics such as image formation, image filtering, segmentation, feature extraction, and pattern recognition techniques. Students will learn how images are analyzed and interpreted using computational methods. The course also highlights machine learning approaches and real-world
applications such as medical imaging, object tracking, and face recognition.</td></tr>
</table>

Syllabus : NEP 2020 w.e.f 2026-27

Unit No.	Content (List of Practicals)	Hours
	1 Basic operation on Image a) Program to change the Brightness of Image. b) To Flip the image around the vertical and horizontal line. c) Display the color components of the image. d) Display of gray scale images. e) To find the negative of an image. **2 Using histogram for image quality analysis** a) Calculate the Histogram of a given image. b) Histogram Equalization. **3 Program for Image Filtering** a) Low pass filter => 1)Average filter2)Weighted Average filter3)Median filter b) High pass filters using=>1) Sobel operator2) Laplacian operator c) Design non-linear filtering. **4 Edge detection with gradient and convolution of an Image** **5 Finding Threshold of Images** a) Program to find threshold of grayscale image. b) Program to find threshold of RGB image. **6 Program to estimate and subtract the background of an image.** **7 Program to convert color image to gray and hsv.** **8.** a) Determination of edge detection using operators. b) 2-D DFT and DCT. c) Filtering in Frequency domain. **9** a) Display of colour images. b) Conversion between colour spaces. **10** a) DWT of images b) Segmentation using watershed transform	
	Total Hours	**60**

References:

1. Computer Vision and Image Processing, Bhuyan, M. K. (2020). Computer vision and image processing. CRC Press.
2. Computer Vision: A Modern Approach, Forsyth, D. A., & Ponce, J. (2012). Computer vision: A modern approach (2nd ed.). Pearson.
3. Machine Vision, Jain, R., Kasturi, R., & Schunk, B. G. (1995). Machine vision. McGraw-Hill.
4. Image Processing, Analysis, and Machine Vision, Sonka, M., Hlavac, V., & Boyle, R. (2007). Image processing, analysis, and machine vision (3rd ed.). Thomson Learning.
5. Computer and Robot Vision, Haralick, R. M., & Shapiro, L. G. (1993). Computer and robot vision. Addison-Wesley.

Sheth T. J. Education Society's
Sheth N.K.T.T College of Commerce and
Sheth J.T.T College of Arts, (Autonomous). Thane (W)

Programme Name: **B.Sc (Data Science)**	**Semester: V**
Course Category: **Vocational Skill Course (VSC)**	
Name of the Dept: **Science and Technology**	
Course Title**: Data Visualization using Tableau Practical**	
Course Code: BDT514	Course Level:5.5
Type: Practical	
Course Credit: 2 credits	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

1. To familiarize students with the Tableau software and its key features.
2. To develop students' understanding of data preparation techniques and best practices in Tableau.
3. To enable students to create a variety of visualizations using different chart types and design principles in Tableau.
4. To introduce students to advanced analytics capabilities in Tableau for trend analysis, forecasting, and cluster analysis.
5. To equip students with the skills to share and present insights effectively using Tableau Online, Tableau Server, and Tableau Public.

Course Outcomes:

After the completion of course learners will be able to

CO1. install Tableau Desktop and navigate the Tableau workspace proficiently and demonstrate competence in data preparation,

CO2. design and create visually compelling and informative visualizations using various chart types and visualization techniques and utilize advanced analytics features

CO3. share their visualizations and insights using Tableau Online, Tableau Server, and Tableau Public, and demonstrate an understanding of best practices in data sharing and presentation.

Description of the course:	This course introduces students to Tableau, a powerful data visualization and analytics tool. From installing Tableau Desktop to creating interactive dashboards and sharing insights, students will learn data preparation, visualization techniques, advanced analytics, and how to leverage Tableau Online, Tableau Server, and Tableau Public. Additionally, they will gain proficiency in data cleaning, formatting, and combining using Tableau Prep. By the end of the course, students will have the skills to analyze data effectively and create impactful visualizations with Tableau.

Syllabus : NEP 2020 w.e.f 2026-27

Sr. No.	Content	Hours
1.	Introduction to Tableau - Install, prepare data, navigate workspace, create visualizations, save/share workbooks.	
2.	Adding Data Sources - Set up connectors, select tables, perform joins/unions, edit metadata, add hierarchies/calculated fields, optimize performance.	
3.	Creating Data Visualizations - Explore chart types, design bar/line/highlight/heatmap/bullet charts, understand visualization anatomy.	
4.	Aggregate Functions and Calculated Fields - Use aggregates/calculated fields, handle text/date fields, apply logical functions/parameters, search text fields.	
5.	Table Calculations and Level of Detail Calculations - Perform different calculations, apply quick/customized table calculations, implement level of detail expressions.	
6.	Maps in Tableau - Create symbol/filled/density maps, add layers/pie charts, use viz in tooltip, explore alternative map services, analyze spatial data.	
7.	Advanced Analytics in Tableau - Identify trends/forecasts/clusters, utilize analytics pane, incorporate lines/forecasts, perform cluster analysis.	
8.	Interactive Dashboards - Considerations for dashboard creation, create/place charts, add titles/navigation/buttons/actions, follow best practices.	
9.	Sharing Insights with Tableau - Utilize Tableau Online/Server, publish to Tableau Public, embed visualizations in websites.	
10.	Data Preparation with Tableau Prep - Connect to data, perform wildcard unions, inspect/clean/format data, remove unneeded fields, combine data with unions/joins, run/save flows.	
	Total Hours	60

Project Guidelines for TY B.Sc. Data Science, Sem V

2-Credit Project

1. Project Type

Students may undertake any one of the following:

1. Data Analysis Project
2. Data Visualization Project
3. Statistical Analysis Project
4. Machine Learning Project
5. Predictive Analytics Project
6. Business Intelligence / Dashboard Project
7. Social Data Science Project
8. Text/Data Mining Project
9. Time-Series Analysis Project
10. AI/Generative AI-based Data Science Project

The project should have a clear problem statement, dataset, methodology and measurable outcome.

2. Suggested Project Areas

Students can select projects from areas such as:

- Education and student performance
- Healthcare and public health
- Banking and finance
- Marketing and customer analytics
- E-commerce
- Social media analytics
- Crime and safety analytics
- Sports analytics
- Environmental data analysis
- Traffic and transportation
- Agriculture
- Tourism
- Human resource analytics
- Retail analytics
- Cybersecurity data analytics
- Financial literacy
- Consumer behaviour
- Employment and placement analytics

3. Recommended Technology Stack

Students may use one or more of the following:

Programming

- Python
- R

Python Libraries

- Pandas
- NumPy
- Matplotlib
- Seaborn
- Scikit-learn
- Plotly

Database

- MySQL / SQLite

Visualization

- Power BI
- Tableau
- Excel

Other Tools

- Jupyter Notebook / Google Colab
- GitHub (recommended)

Students should mention the software, libraries and tools used in the project report.

4. Project Development Stages

Stage 1 – Problem Identification

Students should clearly identify:

- Project title
- Background of the problem
- Problem statement
- Need/significance of the study
- Objectives
- Scope of the project

Stage 2 – Literature Review

Students should refer to relevant:

- Research papers
- Articles
- Reports
- Government/open datasets
- Existing analytical studies

A minimum of 5 relevant references is recommended.

Stage 3 – Data Collection

Students should identify the source of data.

Possible sources:

- Kaggle
- Government open-data portals
- Public datasets
- Institutional data, wherever permitted
- Primary survey using Google Forms
- APIs

Students must clearly mention the source and nature of the dataset.

Stage 4 – Data Preprocessing

Students should demonstrate appropriate preprocessing, such as:

- Handling missing values
- Removing duplicates
- Handling inconsistent data
- Data type conversion
- Outlier identification
- Encoding categorical variables
- Feature selection/engineering, wherever required

Stage 5 – Exploratory Data Analysis

Students should perform appropriate EDA using:

- Descriptive statistics
- Frequency distributions
- Correlation analysis
- Group-wise analysis
- Trend analysis

Stage 6 – Data Visualization

Students should present meaningful visualizations such as:

- Bar charts
- Line charts
- Pie charts, where appropriate
- Histograms
- Scatter plots
- Box plots
- Heatmaps
- Interactive dashboards

Students should interpret the visualizations instead of merely presenting charts.

Stage 7 – Data Science Technique

Depending on the project, students may apply:

Statistical techniques

- Mean, median, standard deviation
- Correlation
- Hypothesis testing

Machine Learning

- Linear regression
- Logistic regression
- Decision tree
- Random forest
- K-Means clustering
- Classification techniques

Students are not required to use complex machine-learning algorithms unless they are relevant to the problem.

Stage 8 – Results and Interpretation

Students should explain:

- Major findings
- Patterns identified
- Model performance, if applicable
- Important insights
- Practical implications

Stage 9 – Conclusion

The conclusion should summarize:

- What was achieved
- Major findings
- Limitations
- Future scope

5. Suggested Project Report Format

Chapter 1 – Introduction

- Background
- Problem statement
- Objectives
- Scope
- Significance

Chapter 2 – Literature Review

- Existing studies
- Research gap/problem identification

Chapter 3 – Research Methodology

- Research design
- Data source
- Dataset description
- Tools and technologies
- Methodology/workflow

Chapter 4 – Data Preparation and Analysis

- Data cleaning
- Preprocessing
- EDA
- Statistical analysis

Chapter 5 – Data Visualization / Model Development

- Visualizations/dashboard
- Machine-learning model, if applicable
- Model evaluation

Chapter 6 – Results and Findings

- Key findings

- Interpretation
- Discussion

Chapter 7 – Conclusion and Future Scope

- Conclusion
- Limitations
- Future improvements

References

Annexure

- Questionnaire, if applicable
- Sample dataset
- Important code snippets
- Additional charts/screenshots

Additional Formatting Instructions

1. The project report should be typed on A4-size pages.
2. All body text should use Times New Roman, 12 pt.
3. The text should be justified with 1.5 line spacing.
4. Each chapter should begin on a new page.
5. Chapter titles should be centre aligned, bold and 16 pt.
6. Main section headings should be 14 pt, bold and left aligned.
7. Sub-headings should be 12 pt, bold and left aligned.
8. Tables and figures should be numbered sequentially and referred to in the text.
9. All tables and figures should have appropriate captions.
10. Page numbers should be placed at the bottom centre of each page.
11. The report should have consistent formatting throughout.
12. Students should submit the final report in PDF format, along with the source code and dataset.

6. Individual or Group Project

For a 2-credit project, a group of 2–4 students may be permitted.
However, each student must have a clearly defined role and demonstrate individual contribution.
Individual projects may also be permitted depending on departmental policy.

Scheme of Examination

Course with Credit	External Examination	Internal Examination	Total
Credit 4	**60 marks**	**40 marks**	**100 marks**
Credit 2	**30 marks**	**20 marks**	**50 marks**

Internal Examination Structure(Theory)

Internal examination	40 marks	20 marks
Project Presentation/Case Study /Quiz/Group Discussion	10 marks	5 marks
Assignment /Active class Participation/Attendance	10 marks	5 marks
Class test	20 marks	10 marks
Total	40 marks	20 marks

Structure for Class Test(For 10 Marks)

For 10 Marks	
Q1. Multiple Choice Questions	**5 Marks**
Q2. Answer in one or two sentences	**5 Marks**

Structure for Class Test(For 20 Marks)

For 20 Marks	
Q1. Multiple Choice Questions	**10 Marks**
Q2. Answer in one or two sentences	**10 Marks**

External Examination (For 60 Marks)

Q. No.	External	Marks: 60
Q .1	Answer the following questions (Any 3) A B C D E F	15 Marks
Q. 2	Answer the following questions (Any 3) A B C D E F	15 Marks
Q. 3	Answer the following questions (Any 3) A B C D E F	15 Marks
Q. 4	Answer the following questions (Any 3) A B C D E **F**	15 Marks

External Examination (For 30 Marks)

Q. No.	External	Marks: 30
Q .1	Answer the following questions (Any 3) A B C D E F	15 Marks
Q. 2	Answer the following questions (Any 3) A B C D E F	15 Marks

Practical External Exam: 50 marks
A Certified copy journal is essential to appear for the practical examination.

1	Practical Question 1	20
2	Practical Question 1	20
3	Journal	5
4	Viva Voce	5

OR

1	Practical Question 1	40
2	Journal	5
3	Viva Voce	5