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)

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			
BDCP509	Cloud Computing Practical	2	BDM610	Marketing & Retail Analytics	2
			BDM611	Marketing & Retail Analytics Practical	2
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					
BDT514	Data Visualization using tableau	2	**SEC**	-	
FP			**OJT**		
BDPD515	Project Dissertation	2	BDPI514	Project Implementation	4
	Total	**22**			**22**

Dr. Yogeshwari Patil
Department Coordinator

Prof. Heena Chande
NEP Coordinator

Dr. Dilip Patil
Principal

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: VI**
Course Category/Vertical: **Major Mandatory**	
Name of the Dept: **Science and Technology**	
Course Title**: Exploratory Data Analysis**	
Course Code: BDE601	Course Level:5.5
Type: Theory	
Course Credit: 02	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

- To introduce the concepts and techniques of Exploratory Data Analysis (EDA) for understanding data characteristics, assumptions, and patterns using statistical summaries and graphical methods.
- To develop the ability to perform data preparation, univariate, bivariate, and multivariate analysis for discovering relationships, handling missing data, and identifying patterns in datasets.

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

CO1: apply EDA techniques to summarize, visualize, and interpret datasets using statistical measures and graphical representations.

CO2: analyze relationships among variables and perform data preparation and dimensionality reduction techniques to support data-driven decision making.

Description of the course:	This course introduces the fundamental concepts and techniques of Exploratory Data Analysis (EDA) for understanding and summarizing datasets. It covers data preparation, handling missing values, and identifying outliers, along with univariate, bivariate, and multivariate analysis using statistical measures and visualization techniques. The course also introduces correlation analysis and dimensionality reduction concepts such as PCA to discover patterns and relationships in data.

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

Unit No.	Content	Hours
I	**Introduction to Exploratory Data Analysis (EDA) :** Definition of EDA, difference between EDA with classical and **Bayesian Analysis**, comparison of EDA with Classical data summary measures, goals of EDA, Underlying assumptions in EDA, importance of EDA in data exploration techniques, introduction to different techniques to test the assumptions involved in EDA, role of graphics in data exploration, introduction to unidimensional, bidimensional and multidimensional graphical representation of data **Data Preparation:** Introduction to data exploration process for data preparation, data discovery, issues related with data access, characterization of data, consistency and pollution of data, duplicate or redundant variables, outliers and leverage data, noisy data, missing values, imputation of missing and empty places, with different techniques, missing pattern and its importance, handling non numerical data in missing places.	15
II	**Univariate Analysis :** Measures of Central Tendency (Mean, Median, Mode), Measures of Dispersion (Variance, SD, IQR),Shape of Data: Skewness (Pearson's coefficients) and Kurtosis, Distribution Analysis : Normal distribution, Binomial, and Poisson; Probability Density Functions (PDF) and Cumulative Distribution Functions (CDF). **Multivariate Analysis & Visualization :** Discovering relationships between variables and dimensionality reduction. **Bivariate Analysis: Continuous vs. Continuous:** Pearson and Spearman Correlation. Categorical vs. Categorical: Chi-square test and Contingency tables. Categorical vs. Continuous: Z-test, T-test, and ANOVA basics. Multivariate Exploration: Heatmaps for correlation matrices, Pair plots, and Parallel Coordinates. **Dimensionality Introduction:** The "Curse of Dimensionality," Intuition behind Principal Component Analysis (PCA) as an EDA tool.	15
	Total Hours	**30**

References:

1. Exploratory Data Analysis, Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley.
2. Python for Data Analysis, McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2nd ed.). Sebastopol, CA: O'Reilly Media.
3. Practical Statistics for Data Scientists, Bruce, P., Bruce, A., & Gedeck, P. (2020). Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python (2nd ed.). Sebastopol, CA: O'Reilly Media.
4. An Introduction to Statistical Learning, James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning with Applications in R (2nd ed.). New York, NY: Springer.

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: VI**
Course Category/Vertical: **Major Mandatory**	
Name of the Dept: **Science and Technology**	
Course Title**: Exploratory Data Analysis Practical**	
Course Code: BDEP602	Course Level:5.5
Type: Practical	
Course Credit: 02	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

- To introduce the concepts and techniques of Exploratory Data Analysis (EDA) for understanding data characteristics, assumptions, and patterns using statistical summaries and graphical methods.
- To develop the ability to perform data preparation, univariate, bivariate, and multivariate analysis for discovering relationships, handling missing data, and identifying patterns in datasets.

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

CO1: apply EDA techniques to summarize, visualize, and interpret datasets using statistical measures and graphical representations.

CO2: analyze relationships among variables and perform data preparation and dimensionality reduction techniques to support data-driven decision making.

Description of the course:	This course introduces the fundamental concepts and techniques of Exploratory Data Analysis (EDA) for understanding and summarizing datasets. It covers data preparation, handling missing values, and identifying outliers, along with univariate, bivariate, and multivariate analysis using statistical measures and visualization techniques. The course also introduces correlation analysis and dimensionality reduction concepts such as PCA to discover patterns and relationships in data.

Unit No.	Content (List of Practicals)	Hours
	1.Managing Data Frames with the dplyr package 2 Use dplyr Grammar for inbuilt data set car. 3 Use group by(), %>%,mutate(), rename(),arrange(), filter(), select() 4 Use the data set air quality from inbuilt data sets library. a) Use summary statistics and find the important key values from the output b) Use boxplot and find the interquartile range. Interpret the boxplot and inner and outerfencing of outliers c) Check the missing value in the data set and fine the suitable solution for the missing values. d) Using histogram, find the distribution of data and give proper comment over the dataset. 5 Use bar plot and identify the difference between bar plot and histograms. Conclude the appropriate use of bar plot and histogram. 6 Explore the two-dimensional data 7 Scatter plot between two variables 8 Five number summary in exploratory data analysis 9 Multiple histogram and multiple boxplots, Multiple scatter plots and colouring the graph 10 Lattice system in R environment and Graphical window in R and its uses	
	Total Hours	**60**

References:

1. Exploratory Data Analysis, Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley.
2. Python for Data Analysis, McKinney, W. (2018). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython (2nd ed.). Sebastopol, CA: O'Reilly Media.
3. Practical Statistics for Data Scientists, Bruce, P., Bruce, A., & Gedeck, P. (2020). Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python (2nd ed.). Sebastopol, CA: O'Reilly Media.

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: VI</td></tr>
<tr><td colspan="2">Course Category/Vertical: Major Mandatory</td></tr>
<tr><td colspan="2">Name of the Dept: Science and Technology</td></tr>
<tr><td colspan="2">Course Title: Deep learning</td></tr>
<tr><td colspan="2">Course Code: BDD603 Course Level:5.5</td></tr>
<tr><td colspan="2">Type: Theory</td></tr>
<tr><td colspan="2">Course Credit: 02</td></tr>
<tr><td colspan="2">Hours Allotted: 30 Hours</td></tr>
<tr><td colspan="2">Marks Allotted: 50 Marks</td></tr>
<tr><td colspan="2">Course Objectives:
To understand the mathematical foundations of Linear Algebra and their application in neural networks and deep learning models.
To explain the architecture, training process, and optimization techniques used in modern deep neural networks.</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of course learners will be able to
CO1: apply linear algebra concepts and implement neural network components such as activation functions, forward propagation, and loss functions.
CO2: analyze and explain deep learning models including feedforward networks, CNNs, and RNNs for solving real-world problems.</td></tr>
<tr><td>Description of the course:</td><td>This course introduces the mathematical foundations and core concepts of Deep Learning, including linear algebra, neural networks, and optimization techniques. It covers the structure and functioning of Artificial Neural Network, activation functions, and training methods such as backpropagation and gradient descent. The course also provides an overview of advanced architectures like Convolutional Neural Network and Recurrent Neural Network
used in modern AI applications.</td></tr>
</table>

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

Unit No.	Content	Hours
I	Applied Math and Machine Learning Basics: Linear Algebra: Scalars, Vectors, Matrices and Tensors, Multiplying Matrices and Vectors, Identity and Inverse Matrices, Linear Dependence and Span, norms, special matrices and vectors, Eigen decompositions. Biological neuron vs Artificial neuron, Structure of Neural Network (Input layer, Hidden layer, Output layer), Activation Functions: Sigmoid, ReLU, Tanh, Forward Propagation and Loss Function	15
II	Deep Neural Networks and Training, Feedforward Neural Network Backpropagation algorithm, Gradient Descent optimization Overfitting and underfitting, Regularization techniques (Dropout, L1, L2) Introduction to Convolutional Neural Network, Introduction to Recurrent Neural Network	15
	Total Hours	**30**

References:

Sr.No	Title	Author/s	Publisher	Edition	Year
1.	Deep Learning	*Ian Goodfellow, Yoshua Bengio, Aaron Courville.*		MIT Press	2016
2.	**F**undamentals of Deep Learning	O'Reilly Media			

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: VI**
Course Category/Vertical: **Major Mandatory**	
Name of the Dept: **Science and Technology**	
Course Title**: Deep Learning Practical**	
Course Code: BDDP604	Course Level:5.5
Type: Practical	
Course Credit: 02	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	
Course Objectives: • To understand the fundamental concepts of tensors, matrix operations, and neural network models used in Deep Learning. • To develop practical skills for implementing regression, classification, and image-based deep learning models using TensorFlow and Keras.	
Course Outcomes: After the completion of course learners will be able to **CO1 :** perform tensor operations, matrix computations, and implement basic machine learning models such as linear regression using TensorFlow. **CO2:** design and train deep neural networks for classification, prediction, and image segmentation tasks.	
Description of the course:	This course introduces practical implementation of **Deep Learning** concepts using **TensorFlow**.Students learn tensor operations, matrix computations, and model building for regression and classification tasks. The course also covers deep neural networks and applications such as image segmentation and prediction using modern deep learning techniques.

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

Sr. No.	Content (List of Practicals)	Hours
1	Create tensors with different shapes and data types.	
2	Perform basic operations like addition, subtraction, multiplication, and division on tensors.	
3	Reshape, slice, and index tensors to extract specific elements or sections.	
4	Performing matrix multiplication and finding eigenvectors and eigenvalues using TensorFlow	
5	Implement a simple linear regression model using TensorFlow's low-level API (or tf. keras).	
6	Train the model on a toy dataset (e.g., housing prices vs. square footage).	
7	Visualize the loss function and the learned linear relationship.	
8	Implementing deep neural network for performing binary classification task	
9	Using a deep feed-forward network with two hidden layers for performing multiclass classification and predicting the class.	
10	Write a program to implement deep learning Techniques for image segmentation.	
	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)

<table>
<tr><td colspan="2">Programme Name: B.Sc.(Data Science) Semester: VI</td></tr>
<tr><td colspan="2">Course Category/Vertical: Major Mandatory</td></tr>
<tr><td colspan="2">Name of the Dept: Science and Technology</td></tr>
<tr><td colspan="2">Course Title: Data Security & Compliance</td></tr>
<tr><td colspan="2">Course Code: BDK605 Course Level:5.5</td></tr>
<tr><td colspan="2">Type: Theory</td></tr>
<tr><td colspan="2">Course Credit: 02</td></tr>
<tr><td colspan="2">Hours Allotted: 30 Hours</td></tr>
<tr><td colspan="2">Marks Allotted: 50 Marks</td></tr>
<tr><td colspan="2">Course Objectives:
To introduce the fundamental principles of data security, cryptography, and privacy regulations relevant to data science applications.
To understand modern security practices and frameworks including security audits, zero trust architecture, and security challenges in AI/ML, IoT, cloud, and DevSecOps environments.</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of course learners will be able to
CO1: explain data security threats, cryptographic techniques, access control mechanisms, and privacy-preserving methods used in protecting data systems.
CO2: analyze and apply security strategies such as risk assessment, zero trust security, and secure practices in AI/ML, IoT, cloud computing, and DevSecOps environments.</td></tr>
<tr><td>Description of the course:</td><td>This course introduces the fundamental concepts of data security, cryptography, and privacy protection in data science environments. It covers security threats, encryption techniques, access control mechanisms, and global data privacy regulations. The course also explores modern security frameworks such as Zero Trust architecture, security auditing, and risk assessment. Additionally, it examines emerging security challenges in AI/ML systems, IoT devices, cloud platforms, and DevSecOps pipelines, preparing students to design and manage secure data-driven systems.</td></tr>
</table>

Unit No.	Content	Hours
I	**Foundations of Data Security** Introduction: Definition & Importance in Data Science, Cybersecurity vs. Data Privacy vs. Data Protection, Key Challenges & Threat Models Data Security Threats: Malware, Ransomware, Phishing, SQL Injection, Advanced Persistent Threats (APT), Social Engineering Attacks **Fundamentals of Cryptography:** Symmetric & Asymmetric Encryption, Encoding, Hashing, Digital Signatures, and Key Management, Applications in Secure Data Storage and Transfer **Access Control & Authentication:** Components of Access Control, ACL, Authentication VS Authorization, Multi-Factor Authentication (MFA), Biometric Security Mechanisms & Vulnerabilities, Types of Access Controls, Identity and Access Management (IAM) **Data Privacy** Data Privacy Regulations & Compliance: Overview of GDPR, CCPA, DPDP, HIPAA, PCI-DSS, ISO 27001, Privacy by Design & Default Principles, Data Governance & Accountability **Privacy-Preserving Data Science:** Differential Privacy & Homomorphic Encryption, Secure Multi-Party Computation (SMPC), Federated Learning Security & Data Leakage Prevention	15
II	**Security Audits & Risk Assessments:** Types of Security Audits, Security Compliance Metrics & Best Practices, Conducting Risk Assessments & Threat Modeling **Zero Trust Security Model:** Zero Trust Principles & Implementation, Micro-Segmentation & Continuous Authentication, Zero Trust in Cloud Environments **Security for AIML, IoT, Cloud, and DevSecOp** **AIML Security:** Threats to AI Models: Adversarial Attacks, Model Poisoning, Model Inversion, Bias, Fairness, & Explainability in AI Security, AI Model Governance & Secure AI Pipelines **IoT Security:** Threats to IoT Devices & Edge Computing, Frameworks related to IoT Security, AI-driven Threat Detection in IoT **Cloud Security:** Security Challenges in AWS, Azure and Google Cloud, Cloud-Native Security Tools (AWS Security Hub, Azure Sentinel) **DevSecOps:** Secure Software Development Life Cycle, Best Practices & CI/CD Pipeline Security	15
	Total Hours	**30**

References:

1. Jarmul, K. (2023). *Practical data privacy: Enhancing privacy and security in data.* O'Reilly Media.
2. Torra, V. (2022). *Guide to data privacy: Models, technologies, and solutions* Springer.
3. Solove, D. J. (2021). *Understanding privacy.* Harvard University Press.
4. Garfinkel, S. (2000). *Database nation: The death of privacy in the 21st century.* O'Reilly Media.

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)

Proramme Name: **B.Sc. (Data Science)**	**Semester : VI**
Course Category: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title: **Sports Analytics**	
Course Code: BDS606	Course Level:5.5
Type: Theory	
Course Credit: 02	
Hours Allotted: 30 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

- To introduce learners to cricket analytics using Python and the Cricpy package, enabling them to analyze individual and team performances across formats such as ODIs and T20s.
- To develop practical skills in data-driven sports performance evaluation, including batsman and bowler analysis, opponent-based comparisons, and granular player analytics using real match data.

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

CO1. perform statistical and performance analysis of cricketers (batsmen and bowlers) across ODI and T20 formats using Cricpy, including comparative and opponent-specific analysis.

CO2. conduct team-level and granular performance analytics, interpret cricket data insights, and generate meaningful visualizations and reports to support strategic decision-making in sports analytics.

Description of the course:	This course introduces learners to sports analytics with a focus on cricket using the Python package Cricpy. It covers performance analysis of batsmen and bowlers across ODI and T20 formats, including detailed evaluation of top players, opponent-based comparisons, and average run analysis against different teams. The course also explores team-level analytics and granular player performance insights. Through hands-on analysis and visualization, students will develop practical skills in applying data-driven techniques to evaluate cricket performance and support strategic decision-making in sports.

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

Unit No.	Content	Hours
I	Introduction, Cricket analytics with cricpy, Introducing cricpy:Importing Cricpy – Python, Getting help from cricpy python, Basic Analysis, A python package to analyze performances of cricketers, Cricpy takes a swing at the ODIs Analysis of Top 4 batsman	15
II	Cricpy takes guard for the Twenty20s, Analyzing batsmen and bowlers with cricpy template, Average runs against different opposing teams, Cricpy adds team analytics to its arsenal!!, Cricpy performs granular analysis of players	15
	Total Hours	30

References:

Sr.No.	Title	Author/s Publisher	Publisher	Edition	Year
1	Cricket analytics with cricketr and cricpy : Analytics harmony with R and Python	Tinniam V Ganesh	Paperback	4th	2023
2	Cricket 2.0: Inside the T20 Revolution	Tim Wigmore & Freddie Wilde	Treeshade	1st	2022
3	The Three Ws of West Indian Cricket: A Comparative Batting Analysis	Keith A. P. Sandiford & Arjun Tan	Paperback	1st	2002

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: VI**
Course Category: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title: **Sports Analytics Practical**	
Course Code: BDSP607	Course Level:5.5
Type: Practical	
Course Credit: 02	
Total Hours: 60 Hours	
Total Marks: 50 Marks	

Course Objectives:

- To develop analytical skills in sports data analysis by applying exploratory data analysis, statistical techniques, and visualization methods to evaluate player and team performance in cricket.
- To equip students with predictive modeling and business analytics skills for analyzing performance trends, injury impact, match outcomes, and sports revenue generation.

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

CO1: analyze and interpret cricket performance data by computing batting, bowling, team, and position-specific metrics, comparing player performances, and identifying trends affecting team success.

CO2: build and evaluate predictive models and generate analytical reports, providing data-driven insights and strategic recommendations for sports performance and revenue optimization.

Description of the Course:	This course provides a comprehensive introduction to sports analytics using cricket datasets. It covers exploratory data analysis (EDA) to examine key performance variables such as matches played, runs, wickets, and player statistics through statistical techniques and visualizations. Students will analyze batting and bowling performance metrics, compare player performances across seasons and teams, and evaluate position-specific contributions.

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

Sr No.	Content (List of Practicals)	Hours
1	**Exploratory Data Analysis** • Perform exploratory data analysis on a cricket dataset, analyzing variables such as number of matches, runs, not outs, wickets, etc. • Visualize the distribution of player performance metrics using histograms, box plots, or scatter plots.	
2	**Batting Performance Analysis** • Analyze batting performance in a Cricket dataset, calculating metrics such as batting average, strike rate, and runs scored by players.	
3	**Bowling Performance Analysis** • Analyze bowling performance in the Cricket dataset, calculating metrics such as bowling average, economy rate, and wickets taken by players.	
4	**Performance Comparison** • Compare the scoring averages of top-performing batsman in different seasons. Analyze the runs scoring (strike rate of Batting) of players from various teams in a specific league	
5	**Player Position Analysis** Calculate position-specific performance metrics and compare players within each position.	
6	**Injury Analysis** • Investigate the relationship between player injuries and their subsequentperformance using historical injury and performance data. Identify patterns and trends in the data to determine the impact of injuries on playerperformance and team success.	
7	**Team Analysis** • Analyze the impact of toss on a team's overall scoring and winning percentage. Study the relationship between batting averages of players and their team's win-loss record	
8	**Sports Revenue Analysis** • Analyze revenue generation in sports organizations by examining factors such as ticket sales, merchandise sales, and sponsorship deals. Identify key drivers of revenue and provide recommendations for maximizing financial performance	
9	**Predictive Modeling** • Build a regression model to predict the number of runs scored by players based ontheir historical performance data. Develop a classification model to predict the outcome of match based on team‟sstatistics.	

10	**Visualization and Reporting: (Mini-Project)** Prepare a comprehensive report summarizing the findings of the analysis and providing actionable insights for sports teams or organizations.	
	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: VI**
Course Category: **Major Elective**	
Name of the Dept: **Science & Technology**	
Course Title: **Healthcare Analytics**	
Course Code: BDH608	Course Level:5.5
Course Type: Theory	
Hours Allotted: 30 Hrs	
Course Credit: 02	Total Marks: 50

Course Objectives:

- To provide foundational knowledge of healthcare systems, electronic medical records, healthcare databases, and the role of computing and analytics in improving patient care and clinical decision-making.
- To develop understanding of machine learning techniques, predictive healthcare models, healthcare quality measurement systems, and emerging technologies in healthcare analytics.

Course Outcomes: After the completion of course learners will be able to
CO1: explain healthcare industry fundamentals, interpret electronic medical records and standardized coding systems and apply basic data handling techniques
CO2: apply machine learning algorithms to healthcare datasets for predictive modeling, evaluate healthcare quality measures, and analyze the impact of emerging technologies while considering ethical and practical limitations.

Description of the course:	This course introduces the fundamentals of Healthcare Analytics and its role in improving patient care and healthcare systems. It covers Electronic Medical Records (EMR), healthcare databases, and clinical coding standards such as International Classification of Diseases and Current Procedural Terminology. Students learn data preprocessing, visualization, and basic SQL for healthcare data management. The course also focuses on machine learning models for predictive analytics, including mortality and readmission prediction. Additionally, it discusses healthcare quality measurement, value-based programs, and emerging technologies like IoT and deep learning in healthcare.

Syllabus: NEP 2020 w.e.f 2025-26

Unit No.	Content	Hours
I	**Introduction to Healthcare Analytics** Definition and importance of healthcare analytics, Role of mathematics, statistics and computer science, Applications: patient care improvement, quality measurement, prediction, History and evolution of healthcare analytics, Healthcare delivery systems, Healthcare financing and policy, Basics of healthcare industry **Electronic Medical Records (EMR)** Structure of EMR (History, Physical Exam, Labs, Assessment & Plan), SOAP clinical notes, Standardized clinical code sets, International Classification of Diseases (ICD), Current Procedural Terminology (CPT), Logical Observation Identifiers Names and Codes (LOINC), Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT) **Computing and Data Foundations** Introduction to healthcare databases, Basics of SQL in healthcare data, Data preprocessing and visualization, Case example: Predicting mortality	15
II	**Machine Learning & Predictive Healthcare Models** Medical decision-making models, Bayes theorem and probabilistic reasoning, Machine learning pipeline, Logistic Regression, Random Forest, Neural Networks, Predictive analytics in healthcare, Readmission prediction models, Cardiovascular risk prediction (Framingham Risk Score), Cancer prediction basics **Measuring Healthcare Quality** Healthcare quality measures, Value-based programs, Performance measurement and benchmarking, Introduction to HEDIS **Emerging Technologies in Healthcare** Internet of Things (IoT) in healthcare, Deep learning applications, Ethical issues and limitations	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Healthcare Analytics Made Simple	Vikas Kumar	Packt Publishing	1st	2018
2.	Healthcare Analytics: Foundations and Frontiers	Ross M. Mullner Edward M. Rafalski	T&F / Routledge	1st	2020
3.	Hands-On Healthcare Data	Andrew Nguyen	Shroff / O'Reilly	1st	2022
4.	AI-First Healthcare: AI Applications	Kerrie L. Holley	Shroff / O'Reilly	1st	2021
5.	Healthcare Data Analytics	Chandan K. Reddy, Charu C. Aggarwal	Chapman and Hall/CRC	1st	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)

Programme Name: **B.Sc. (Data Science)**	**Semester: VI**
Course Category: **Major Elective**	
Name of the Dept: **Science & Technology**	
Course Title: **Healthcare Analytics Practical**	
Course Code: BDHP609	Course Level: 5.5
Course Type: Practical	
Hours Allotted: 60 Hrs	
Course Credit: 02	Total Marks: 50

Course Objectives:

- To develop practical skills in importing, cleaning, and preprocessing healthcare datasets using Python.
- To enable students to build and evaluate predictive models for healthcare data analysis.

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

CO1: import and preprocess healthcare data from CSV files and perform exploratory data analysis.

CO2: develop predictive healthcare models and evaluate their performance using appropriate evaluation metrics.

Description of the course:	This practical course introduces students to hands-on healthcare data analysis using Python. Students will learn to import healthcare datasets from CSV files, clean and preprocess real-world medical data, and perform exploratory data analysis. The course emphasizes building predictive models for healthcare applications such as disease prediction and risk analysis. Students will also evaluate model performance using appropriate metrics to support data-driven decision- making in healthcare systems.

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

Sr. No.	Content (List of Practicals)	Hours
1	**Healthcare Dataset Exploration** Load healthcare dataset, perform descriptive statistics, identify missing values, generate summary report.	
2	**Data Cleaning & Missing Value Treatment Handle missing** values, remove duplicates, correct data types	
3	**Data Encoding & Transformation** Convert categorical variables into numerical form (Label Encoding / One-Hot Encoding); normalization of numerical Data	
4	**Exploratory Data Analysis (EDA)** Perform correlation analysis; visualize age, disease, mortality trends using graphs	
5	**Feature Selection & Data Splitting** Select important features; split dataset into training and testing sets	
6	**Logistic Regression – Disease Prediction** Build Logistic Regression model using CSV dataset; evaluate accuracy & confusion matrix	
7	**Decision Tree – Mortality Prediction** Apply Decision Tree classifier; compare performance metrics	
8	**Random Forest – Readmission Prediction** Implement Random Forest model; calculate precision, recall, F1-score	
9	**Model Performance Evaluation** Evaluate models using confusion matrix, ROC curve, accuracy, precision, recall	
10	**Predictive Case Study (Healthcare Risk Analysis)** Perform Complete workflow: CSV import → preprocessing → model building → prediction → interpretation of results	
	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: VI**
Course Category: **Major Elective**	
Name of the Dept: **Science & Technology**	
Course Title: **Marketing and Retail Analytics**	
Course Code: BDM610	Course Level:5.5
Couse Type: Theory	
Hours Allotted: 30 Hours	
Course Credit: 02	Total Marks: 50

Course Objectives:

- To understand the fundamental concepts, components, and characteristics of time series data and forecasting techniques.
- To apply appropriate time series forecasting models for predicting future business trends and supporting data-driven decision making.

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

CO1: analyze time series data by identifying trend, seasonal, cyclical, and irregular components.

CO2: implement suitable forecasting methods such as Moving Average, Exponential Smoothing, and ARIMA to predict future values and interpret results effectively.

Description of the course:	This course introduces the fundamental concepts and techniques of Time Series Forecasting used for predicting future trends based on historical data. It covers the components of time series data such as trend, seasonality, cyclical and irregular variations, along with commonly used forecasting models including Moving Average, Exponential Smoothing, and ARIMA. The course emphasizes practical applications in sales forecasting, demand prediction, and business decision-making using statistical and analytical methods.

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

Unit No.	Content	Hours
I	**Introduction to Marketing Data** Types of Marketing Data (Structured & Unstructured), Sources of Marketing Data, Data Collection Methods, Role of Data in Marketing Decision Making **Introduction to Marketing Data and Forecasting** Pricing Strategies, Demand-Based Pricing, Cost-Based Pricing, Competitive Pricing, Price Optimization Techniques, Sales Forecasting Methods, Time Series Forecasting, Regression-Based Forecasting, Forecasting New Product Sales **Customer Analysis** Customer Behaviour Analysis, Customer Lifetime Value (CLV), Customer Profitability Analysis, Market Segmentation Techniques, Demographic, Geographic, Psychographic, Behavioural	15
II	**Retailing and Retail Analytics** Introduction to Retail Analytics, Retail Data Sources, Retail Performance Metrics, Importance of Geography and Demographics in Retail, Store Layout and Product Placement, In-Store Promotions, Store Operations Data, Inventory Analytics. **Advertising, Loyalty and Digital Marketing Analytics** Advertising Effectiveness Measurement, Marketing Research Tools, Internet and Social Media Marketing Analytics, Campaign Performance Metrics, Customer Loyalty Programs, Reward Systems, Retention Analytics, Churn Analysis.	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Marketing Analytics	Wayne L. Winston	John Wiley & Sons	-	-
2.	Retailing Analytics	Emmett Cox	John Wiley & Sons	-	-

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: VI**
Course Category/Vertical: **Major Elective**	
Name of the Dept: **Science and Technology**	
Course Title**: Marketing & Retail Analytics Practical**	
Course Code: BDM611	Course Level:5.5
Type: Practical	
Course Credit: 02	
Hours Allotted: 60 Hours	
Marks Allotted: 50 Marks	

Course Objectives:

- To learn working and analyzing with marketing data.
- To develop predictive marketing dashboard for organization.
- To Understand the concept of hypothesis testing and its role in assessing experiment

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

CO1. learn how to tabulate and summarize marketing data using R.
CO2. design and conduct experiments for marketing campaigns.
CO3. understand the concept of hypothesis testing and its role in assessing experiment outcomes.
CO4. calculate, predict, and apply Customer Lifetime Value.
CO5. analyze customer purchasing patterns and build a recommender system based on market basket analysis.

Description of the course:	This course focuses on practical implementation learn to organize and summarize marketing data using R programming. Develop skills to design and perform experiments for marketing campaigns. Understand hypothesis testing and evaluate experimental results. Calculate and apply Customer Lifetime Value (CLV) for business insights. Analyze customer purchasing behavior and build a recommender system using market basket analysis.

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

Sr. No.	Content (List of Practicals)	Hours
1	**Learn how to tabulate and summarize marketing data using R** -Clean and preprocess the marketing data. -Generate a simple histogram plot to visualize data distribution. -Use tabulation and summary functions to gain insights from the data. -Interpret the findings and discuss the implications for marketing analysis.	
2	**Gain proficiency in visualizing marketing data using R** - Understand the key elements of data visualization. - Create visualizations such as histograms, scatter plots, line plots, and bar charts using ggplot() in R. - Apply appropriate visualization techniques to communicate marketing insights.	
3	**Design and conduct experiments for marketing campaigns** -Learn about experimental design and its application in marketing. -Design experiments using marketing scenarios. -Implement randomization and sample splitting techniques. -Conduct experiments and collect data for analysis.	
4	**Hypothesis testing and experiment outcomes** -Explore the purpose of hypothesis testing. -Learn key terminologies related to hypothesis testing. -Learn hypothesis testing process and power calculation. -Conduct hypothesis testing using R.	
5	**Customer Lifetime Value (CLV)** -Calculate CLV using different frameworks. -Use linear regression and logistic regression for CLV prediction. -Assess accuracy and reliability of predictions.	
6	**CLV and cohort analysis** -Analyze CLV data and identify patterns. -Perform cohort analysis to segment customers. -Interpret results for marketing strategies.	
7	**Social media data extraction and analysis** -Use Python libraries like Beautiful Soup and Requests for scraping. -Clean and preprocess the data. -Perform sentiment analysis or engagement analysis. -Visualize results using charts.	
8	**Market basket analysis and recommender systems** -Identify frequent item sets using association rule mining. -Calculate support, confidence, and lift. -Build recommendation engine using collaborative filtering. -Evaluate recommender system performance.	

9 10	**RFM analysis** -Segment customers using clustering algorithms like K-means or hierarchical clustering. -Perform descriptive analysis of segments. -Develop targeted marketing strategies. **A/B testing for marketing campaigns** -Collect relevant data and perform statistical analysis. -Calculate metrics like conversion rate, click-through rate, and revenue. -Interpret results and provide optimization recommendations.	
	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)

<table>
<tr><td colspan="2">Programme Name: B.Sc. (Data Science) Semester: VI</td></tr>
<tr><td colspan="2">Course Category: Minor</td></tr>
<tr><td colspan="2">Name of the Dept: Science & Technology</td></tr>
<tr><td colspan="2">Course Title: Internet of Things</td></tr>
<tr><td colspan="2">Course Code: BDI612 Course Level:5.5</td></tr>
<tr><td colspan="2">Course Type: Theory</td></tr>
<tr><td colspan="2">Hours Allotted: 30 Hours</td></tr>
<tr><td colspan="2">Course Credit: 02 Total Marks: 50</td></tr>
<tr><td colspan="2">Course Objectives:
To understand the fundamental architecture, communication protocols, and design principles of IoT systems.
To develop knowledge of IoT hardware platforms, communication mechanisms, and deployment strategies.</td></tr>
<tr><td colspan="2">Course Outcomes: After the completion of course learners will be able to
CO1: explain IoT architecture, protocols, and embedded platforms used in real-world applications.
CO2: analyze IoT implementation challenges including security, deployment, and ethical considerations.</td></tr>
<tr><td>Description of the course:</td><td>This course introduces the fundamental concepts and architecture of the Internet of Things (IoT). It covers internet communication principles, embedded systems, and prototyping of connected devices. Students will learn about IoT hardware platforms like Arduino and Raspberry Pi along with cloud and API integration basics. The course also discusses IoT security, business models, and ethical considerations in connected systems.</td></tr>
</table>

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

Unit No.	Content	Hours
I	**Introduction and Internet Principles** 1. Overview of Internet of Things 2. Design Principles for Connected Devices 3. Internet Communication Basics **Embedded Systems and Prototyping Basics** 1. Basics of Embedded Systems 2. IoT Hardware Platforms 3. Introduction to Prototyping	15
II	**IoT Communication and Cloud Integration** 1. APIs and Web Services in IoT 2. Real-Time IoT Communication 3. Basics of Cloud Integration in IoT **Deployment, Security and Ethics in IoT** 1. Moving to Manufacturing 2. Security in IoT 3. IoT Business Models (Overview) 4. Ethics in IoT 5. Responsible Innovation	15
	Total Hours	30

References:

Sr. No	Title	Author	Publisher	Edition	Year
1.	Designing the Internet of Things	Adrian McEwen, Hakim Cassimally	Wiley	First	2014
2.	Internet of Things – Architecture and Design	Design Raj Kamal	McGraw Hill	First	2017
3.	Getting Started with the Internet of Things	Cuno Pfister	O'Reilly	Sixth	2018
4.	Getting Started with Raspberry Pi	Matt Richardson and Shawn Wallace	SPD	Third	2016

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: VI**
Course Category: **Minor**	
Name of the Dept: **Science & Technology**	
Course Title: **Internet of Things Practical**	
Course Code: BDIP613	Course Level:5.5
Course Type: Practical	
Hours Allotted: 60 Hours	
Course Credit: 02	Total Marks: 50

Course Objectives:

- To develop practical skills in configuring Raspberry Pi and interfacing sensors, actuators, and external devices for IoT applications.
- To enable students to implement cloud-based monitoring and web-controlled IoT systems using ESP8266 and Raspberry Pi.

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

CO1: design and implement IoT applications using Raspberry Pi, ESP8266, sensors, RFID, and camera modules**.**

CO2: develop real-time IoT systems with cloud integration, remote access, and basic network configuration.

Description of the course:	This practical course provides hands-on experience in developing IoT applications using Raspberry Pi and ESP8266. Students learn device configuration, GPIO interfacing, sensor integration, and cloud connectivity. The course covers real-time monitoring, web-based control, RFID, camera interfacing, and IoT communication using platforms like ThingSpeak and Telegram. It emphasizes practical implementation of embedded systems and cloud-enabled IoT solutions.

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

Sr No.	Content (List of Practicals)	Hours
1	Starting Raspbian OS, Familiarizing with Raspberry Pi Components and interface, Connecting to ethernet, Monitor, USB.	
2	Displaying One LED blinking with Raspberry Pi.	
3	Displaying different LED patterns with Raspberry Pi.	
4	Displaying Time over 4-Digit 7-Segment Display using Raspberry Pi.	
5	Raspberry Pi Based Oscilloscope.	
6	Controlling Raspberry Pi with Telegram.	
7	Interfacing Raspberry Pi with RFID.	
8	Controlling LED using Web Browser (ESP8266)	
9	IoT-Based Temperature Monitoring using ThingSpeak (Cloud Integration)	
10	Interfacing Pi Camera with Raspberry Pi	
11	Installing Windows 10 IoT Core on Raspberry Pi (Demo Practical)	
	Total Hours	**60**

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