



Kot Bhalwal, Jammu



Model Institute of Engineering
& Technology (Autonomous)
Dr. Arun K. Gupta Teaching-Learning Centre

Department of CSE

Details of Lesson Plan

S.No.	Particulars	Details
1.	Course Name	Introduction to Data Sciences
2.	Course Code	COM - 501
3.	Academic Year	2024- 2025
4.	Semester	5 th
5.	Number of Lesson plans	48
6.	Faculty Assigned	Dr. Palvi Sharma

Faculty Signature

Lesson Plan No. 1	Course Name: Introduction to Data Science Topic: Data Structures (Structured, Semi-structured, and Unstructured Data)	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the different types of data structures. Recognize the characteristics and use cases of structured, semi-structured, and unstructured data. Identify examples of each type of data structure in real-world applications. Describe the importance of handling different data structures in big data analytics.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask questions: <ul style="list-style-type: none"> What types of data do you think exist in the digital world? Can you provide examples of data you encounter daily? Discussion Points <ul style="list-style-type: none"> Brief overview of data structures in the context of big data. Importance of understanding different data structures for effective data analysis. Course Outcome <ul style="list-style-type: none"> Introduction to the lesson objectives and expected outcomes. Development (30 minutes) <ol style="list-style-type: none"> Structured Data: <ul style="list-style-type: none"> -Definition and characteristics: Highly organized, easily searchable. -Examples: SQL databases, spreadsheets. -Use cases: Financial systems, customer relationship management (CRM) systems. -Show examples of SQL queries and how structured data is stored and retrieved. Semi-structured Data: <ul style="list-style-type: none"> -Definition and characteristics: Partially organized, lacks a fixed schema. -Examples: XML files, JSON documents. -Use cases: Web data, data interchange between systems. -Demonstrate parsing JSON and XML data. Unstructured Data: <ul style="list-style-type: none"> -Definition and characteristics: Lacks any predefined structure, not easily searchable. -Examples: Text files, images, videos. -Use cases: Social media content, multimedia archives. -Show how tools like Hadoop handle unstructured data.

	<p>d. Importance of Handling Different Data Structures:</p> <ul style="list-style-type: none"> -Discuss the challenges and benefits of analyzing various data types. -Explain how big data tools and technologies (e.g., Hadoop, Spark) manage different data structures. -Highlight the significance of data integration and processing in deriving meaningful insights. <p>e. Future Trends in Data Science</p> <ul style="list-style-type: none"> -Discuss emerging trends such as real-time analytics, integration of AI in data processing. -Examples of how these trends impact the handling and analysis of different data structures.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson correlating with learning outcomes. 2. Encourage students to explore further reading and online resources on data structures and their applications in big data.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> 1. What are the main differences between structured, semi-structured, and unstructured data? 2. Why is it important to manage and analyze different data structures in big data analytics? 3. Can you think of real-world examples where each type of data structure is used? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>



Lesson Plan No. 2	Course Name: Introduction to Data Science Topic: Analyst Perspective on Data Repositories	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: a. Understand the role and importance of data repositories in data science. b. Identify different types of data repositories and their characteristics. c. Explain the key considerations for choosing and managing data repositories. d. Recognize the applications of data repositories in various industries.
Teaching Aids (if any)	a. Presentation slides. b. Projector c. You Tube video
Teaching Development	<ol style="list-style-type: none">Introduction (5 minutes)<ul style="list-style-type: none">- Discuss what students understand by the term "data repository."- Ask how they think data repositories are used in everyday applications.- Explain the relevance of data repositories in storing and managing vast amounts of data.Development (30 minutes)<ol style="list-style-type: none">What is a Data Repository?<ul style="list-style-type: none">- Definition: Explain a data repository as a centralized place where data is stored and managed.- Types of Data Repositories: Discuss different types, such as databases, data warehouses, data lakes, and data marts.Importance of Data Repositories:<ul style="list-style-type: none">- Highlight the role of data repositories in ensuring data availability, consistency, and security.- Explain how they support data analysis and decision-making processes.Types of Data Repositories:<ul style="list-style-type: none">● Databases: Used for storing structured data, typically managed by a Database Management System (DBMS).● Data Warehouses: Designed for analytical processing, integrating data from multiple sources.● Data Lakes: Store large volumes of raw data in its native format, often used for big data analytics.● Data Marts: Subsets of data warehouses tailored for specific business functions or departments.Choosing a Data Repository:<ul style="list-style-type: none">- Key considerations: Discuss factors like data volume, data type, performance requirements, and budget.



	<ul style="list-style-type: none">- Examples: Provide examples of when to use a database vs. a data lake vs. a data warehouse.e. Managing Data Repositories:<ul style="list-style-type: none">- Best practices: Discuss best practices for data governance, data quality, and data security.- Tools and technologies: Introduce tools like SQL, Hadoop, and Apache Spark for managing data repositories.f. Applications of Data Repositories:<ul style="list-style-type: none">- Healthcare: Managing patient records, supporting clinical research, and improving patient outcomes.- Retail: Tracking sales, managing inventory, and understanding customer behaviour.- Finance: Storing transaction data, supporting fraud detection, and risk management.- Social media: Analysing user behaviour, supporting targeted advertising, and sentiment analysis. <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none">- Ask students to summarize the examples of Big Data analysis discussed and their significance.- Discuss the answers with the class using Nearpod to collect responses.
Closure	<ol style="list-style-type: none">1. Summarize the lesson learning outcomes and affirm understanding with students.2. Suggested Reading<ul style="list-style-type: none">- "Data Science and Big Data Analytics" by EMC Education Services (Editor). <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none">1. How would you define a data repository in your own words?2. What are the different types of data repositories and their main characteristics?3. Provide examples of how data repositories are used in different industries. <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>



Lesson Plan No. 3	Course Name: Introduction to Data Science Topic: State of the Practice in Analytics: Business Intelligence (BI) vs. Data Science	Course No.: COM-501
-------------------	---	---------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the differences between Business Intelligence (BI) and Data Science. Identify the key components and tools associated with BI and Data Science. Explain the roles and responsibilities of BI analysts and data scientists. Recognize the applications and implications of BI and Data Science in various industries.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Discuss what students understand by the terms "Business Intelligence" and "Data Science." Ask how they think these fields are applied in everyday business operations. Explain the relevance of BI and Data Science in making data-driven decisions. Development (30 minutes) <ol style="list-style-type: none"> What is Business Intelligence (BI)? <ul style="list-style-type: none"> -Definition: Explain BI as a set of strategies and tools for analysing business information to support decision-making. -Components: Discuss components such as reporting, online analytical processing (OLAP), dashboards, and data mining. What is Data Science? <ul style="list-style-type: none"> - Definition: Explain Data Science as a field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from data. - Components: Discuss components such as data cleaning, data analysis, machine learning, and predictive modeling. Key Differences Between BI and Data Science: <ul style="list-style-type: none"> ● Purpose: <ul style="list-style-type: none"> - BI: Focuses on descriptive and diagnostic analysis to understand past and current performance. - Data Science: Focuses on predictive and prescriptive analysis to forecast future trends and make recommendations. ● Tools:



	<ul style="list-style-type: none">- BI: Common tools include Power BI, Tableau, QlikView, and traditional SQL databases.- Data Science: Common tools include Python, R, Apache Spark, and Jupyter notebooks.● Data Handling:<ul style="list-style-type: none">- BI: Typically deals with structured data.- Data Science: Handles both structured and unstructured data.● Output:<ul style="list-style-type: none">- BI: Generates reports, dashboards, and visualizations.- Data Science: Produces predictive models, algorithms, and actionable insights. <p>d. Roles and Responsibilities:</p> <ul style="list-style-type: none">● BI Analysts:<ul style="list-style-type: none">- Role: Focus on creating reports and dashboards, performing data analysis, and supporting business decision-making.- Skills: Proficiency in BI tools, SQL, data visualization, and understanding of business processes.● Data Scientists:<ul style="list-style-type: none">- Role: Develop predictive models, perform complex data analysis, and design algorithms to solve business problems.- Skills: Proficiency in programming languages (Python, R), machine learning, statistical analysis, and data engineering. <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none">- Ask students to summarize the key differences between BI and Data Science.- Discuss the answers with the class using Nearpod to collect responses.
Closure	<ol style="list-style-type: none">1. Summarize the lesson learning outcomes and affirm understanding with students.2. Suggested Reading<ul style="list-style-type: none">- "Data Science and Big Data Analytics" by EMC Education Services (Editor).- "Data Science for Business" by Foster Provost and Tom Fawcett.
Evaluation	<ol style="list-style-type: none">1. How would you differentiate between Business Intelligence and Data Science?2. What are the main tools used in BI and Data Science?3. Provide examples of how BI and Data Science are applied in different industries. Provide examples of how data repositories are used in different industries.



Spend 5 minutes to evaluate student assimilation of the lesson contents



Lesson Plan No. 4	Course Name: Introduction to Data Science Topic: Current Analytics Architecture and Drivers of Big Data	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the current architecture used in analytics. Identify the key components of analytics architecture. Explain the primary drivers of Big Data. Recognize the implications of Big Data drivers on analytics architecture.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Discuss what students understand by the term "analytics architecture." Ask how they think Big Data impacts analytics architecture. Explain the relevance of having a robust analytics architecture in handling Big Data. Development (30 minutes) <ol style="list-style-type: none"> What is Analytics Architecture? <ul style="list-style-type: none"> Definition: Explain analytics architecture as the framework that defines how data is collected, stored, processed, and analysed. Components: Discuss key components such as data sources, data storage, data processing, data analysis, and data visualization. Key Components of Analytics Architecture: <ul style="list-style-type: none"> Data Sources: Include databases, data lakes, streaming data, and external data sources. Data Storage: Discuss storage solutions such as data warehouses, data lakes, and cloud storage. Data Processing: Explain processing frameworks like ETL (Extract, Transform, Load), real-time processing, and batch processing. Data Analysis: Introduce tools and methods for data analysis including machine learning, statistical analysis, and predictive analytics. Data Visualization: Discuss visualization tools and techniques to present data insights, such as dashboards and reports. Drivers of Big Data: <ul style="list-style-type: none"> Volume: The sheer amount of data generated from various sources. Velocity: The speed at which data is generated and processed.



	<ul style="list-style-type: none"> - Variety: The different types of data (structured, unstructured, semi-structured). - Veracity: The uncertainty and reliability of data. - Value: The potential insights and benefits derived from analysing Big Data. <p>d. Implications of Big Data Drivers on Analytics Architecture:</p> <ul style="list-style-type: none"> - Scalability: Need for scalable storage and processing solutions to handle large volumes of data. - Real-Time Processing: Importance of real-time data processing to manage high-velocity data. - Data Integration: Challenges of integrating diverse data types and sources. - Data Quality: Ensuring data accuracy and reliability despite the high veracity. - Cost Efficiency: Balancing the cost of storing and processing Big Data with the value derived. <p>e. Examples of Current Analytics Architecture:</p> <ul style="list-style-type: none"> - Healthcare: Utilizing real-time patient monitoring and predictive analytics for better patient outcomes. - Retail: Implementing customer data platforms to personalize marketing and optimize inventory. - Finance: Leveraging fraud detection systems that process and analyse transaction data in real-time. - Transportation: Using data from IoT devices to optimize routes and manage vehicle maintenance. <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none"> - Ask students to summarize the key components of analytics architecture and the drivers of Big Data. - Discuss the answers with the class using Nearpod to collect responses.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson learning outcomes and affirm understanding with students. 2. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none"> 1. What are the key components of analytics architecture? 2. What are the primary drivers of Big Data? 3. Provide examples of how Big Data drivers impact analytics architecture in different industries <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>



Kot, Bhalwal, Jammu

Model Institute of Engineering & Technology (Autonomous) Lesson Plan



Dr. Arun K. Gupta Teaching-Learning Centre

Version 1.1



Please Do Not Print Unless Necessary



Lesson Plan No. 5	Course Name: Introduction to Data Science Topic: Emerging Big Data Ecosystem	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: a. Understand the concept of the big data ecosystem. b. Identify key components and technologies within the big data ecosystem. c. Explain new approaches to analytics within the context of big data. d. Recognize the applications and implications of big data analytics in various industries.
Teaching Aids (if any)	a. Presentation slides. b. Projector c. You Tube video
Teaching Development	<ol style="list-style-type: none">Introduction (5 minutes)<ul style="list-style-type: none">- Discuss what students understand by the term "Big Data."- Ask how often they encounter Big Data applications in their daily lives.- Explain the relevance of Big Data in today's data-driven world.Development (30 minutes)<ol style="list-style-type: none">What is Big Data?<ul style="list-style-type: none">- Definition: Explain Big Data as large, complex datasets that traditional data-processing software cannot manage effectively.- Characteristics: Discuss the 4 V's of Big Data: Volume, Velocity, Variety, and Veracity.Importance of Big Data Analysis:<ul style="list-style-type: none">- Highlight how Big Data analysis helps in uncovering hidden patterns, correlations, and insights.- Explain its role in improving decision-making processes across various industries.Examples of Big Data Analysis: - Amazon AWS<ul style="list-style-type: none">● Healthcare:<ul style="list-style-type: none">- Discuss how Big Data is used in predictive analytics to improve patient outcomes, manage healthcare costs, and personalized medicine.- Case Study: Use of Big Data in predicting disease outbreaks and patient readmission rates.● Retail:<ul style="list-style-type: none">- Explain how retailers use Big Data to understand customer behaviour, optimize inventory, and improve the shopping experience.- Case Study: How Big Data helped a major retailer boost sales through personalized marketing.● Finance:



	<ul style="list-style-type: none">- Discuss the use of Big Data in fraud detection, risk management, and algorithmic trading.- Case Study: Big Data in detecting credit card fraud and market trends.● Social Media:<ul style="list-style-type: none">- Explain the role of Big Data in sentiment analysis, targeted advertising, and user behaviour analysis.- Case Study: Analysing social media trends to forecast public opinion and marketing success.● Transportation:<ul style="list-style-type: none">- Discuss the application of Big Data in optimizing routes, predicting maintenance, and improving safety.- Case Study: How Big Data is used by ride-sharing companies to optimize routes and reduce wait times. <p>d. Big Data Tools and Techniques:</p> <ul style="list-style-type: none">- Introduce common Big Data tools like Hadoop, Spark, and NoSQL databases.- Briefly explain the role of machine learning in Big Data analysis. <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none">- Ask students to summarize the examples of Big Data analysis discussed and their significance.- Discuss the answers with the class. <p>Use Nearpod to collect responses and discuss the answers.</p>
Closure	<ol style="list-style-type: none">1. Summarize the lesson learning outcomes and affirm understanding with students.2. Suggested Reading<ul style="list-style-type: none">- "Data Science and Big Data Analytics" by EMC Education Services (Editor). <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none">1. How would you define Big Data in your own words?2. What are the main characteristics of Big Data?3. Provide examples of Big Data analysis in different industries. <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>



Lesson Plan No. 6	Course Name: Introduction to Data Science Topic: New Approach to Analytics	Course No.: COM-501
-------------------	---	---------------------

Objectives	At the end of the lesson the student shall be able to: <ul style="list-style-type: none"> a. Understand the new approaches and methodologies in analytics. b. Identify the key differences between traditional analytics and new analytics approaches. c. Explain the benefits and challenges of adopting new analytics approaches. d. Recognize the applications of new analytics approaches in various industries.
Teaching Aids (if any)	<ul style="list-style-type: none"> a. Presentation slides. b. Projector c. You Tube video
Teaching Development	<ol style="list-style-type: none"> 1. Introduction (5 minutes) <ul style="list-style-type: none"> - Discuss what students understand by "new approach to analytics." - Ask how they think analytics has evolved over time. - Explain the relevance of adopting new approaches in analytics to stay competitive and derive more value from data. 2. Development (30 minutes) <ol style="list-style-type: none"> a. Traditional Analytics vs. New Analytics Approaches: <ul style="list-style-type: none"> ● Traditional Analytics: <ul style="list-style-type: none"> - Definition: Explain traditional analytics as methods focused on descriptive and diagnostic analysis using structured data and conventional tools. - Tools: Common tools include SQL databases, Excel, and basic reporting tools. - Limitations: Discuss limitations such as handling only structured data, lack of scalability, and slower processing times. ● New Analytics Approaches: <ul style="list-style-type: none"> - Definition: Explain new analytics approaches as methods that leverage advanced technologies and methodologies to perform predictive, prescriptive, and real-time analysis on diverse data types. - Components: Discuss key components such as big data technologies, machine learning, artificial intelligence, and real-time processing. b. Key Differences: <ul style="list-style-type: none"> ● Data handling: <ul style="list-style-type: none"> - Traditional: Structured data from relational databases. - New: Structured, semi-structured, and unstructured data from various sources.



	<ul style="list-style-type: none"> ● Processing Speed: <ul style="list-style-type: none"> - Traditional: Batch processing with longer times. - New: Real-time and near-real-time processing. c. Benefits of New Analytics Approaches: <ul style="list-style-type: none"> ● Scalability: Ability to handle large volumes of data from multiple sources. ● Real-Time Insights: Gaining immediate insights for timely decision-making. ● Enhanced Accuracy: Improved accuracy through machine learning and AI. ● Flexibility: Adapting to various data types and changing business needs. d. Challenges of New Analytics Approaches: <ul style="list-style-type: none"> ● Complexity: Higher complexity in implementation and management. ● Cost: Increased costs associated with advanced technologies and skilled personnel. ● Data Privacy: Ensuring data privacy and security with larger and diverse datasets. ● Skill Gap: Need for specialized skills in data science, machine learning, and big data technologies. <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none"> - Ask students to summarize the key differences between traditional and new analytics approaches. - Discuss the answers with the class. <p>Use Nearpod to collect responses and discuss the answers.</p>
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson's learning outcomes and affirm understanding with students. 2. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none"> 1. How would you differentiate between traditional analytics and new analytics approaches? 2. What are the main benefits and challenges of new analytics approaches? 3. Provide examples of how new analytical approaches are applied in different industries. <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>



Kot, Bhalwal, Jammu

Model Institute of Engineering & Technology (Autonomous) Lesson Plan



Dr. Arun K. Gupta Teaching-Learning Centre

Version 1.1



Please Do Not Print Unless Necessary



Lesson Plan No. 7	Course Name: Introduction to Data Science Topic: Key Roles for the New Big Data Ecosystem	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the key roles in the new Big Data ecosystem. Identify the responsibilities and skills required for each role. Explain the importance of each role in managing and leveraging Big Data. Recognize the interdependencies between different roles in the Big Data ecosystem.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Discuss what students understand by the term "Big Data ecosystem." Ask what roles they think are critical in managing Big Data. Explain the relevance of having specialized roles to effectively manage and utilize Big Data. Development (30 minutes) <ol style="list-style-type: none"> Key Roles in the Big Data Ecosystem: <ul style="list-style-type: none"> Data Engineer: <ul style="list-style-type: none"> Design, construct, install, and maintain large-scale processing systems. Ensure systems are secure, reliable, and scalable. Develop, construct, test, and maintain architectures such as databases and large-scale processing systems. Proficiency in programming languages like Python, Java, and Scala. Experience with big data tools like Hadoop, Spark, and Kafka. Strong understanding of database systems and data warehousing solutions. Data Scientist: <ul style="list-style-type: none"> Analyse and interpret complex data to provide actionable insights. Develop algorithms and models to predict future trends. Work with large datasets to solve specific business problems. Proficiency in statistical analysis and machine learning. Experience with programming languages like Python and R. Strong analytical and problem-solving skills. Data Analyst:



	<ul style="list-style-type: none">- Analyse data and generate reports to support business decision-making.- Use data visualization tools to present data insights.- Perform data cleansing and transformation tasks.- Proficiency in SQL and data visualization tools like Tableau and Power BI.- Strong analytical skills and attention to detail.- Ability to communicate insights effectively to non-technical stakeholders. <p>d. Data Architect:</p> <ul style="list-style-type: none">- Design and maintain the overall data architecture for an organization.- Ensure data integrity, security, and availability.- Work with other IT and data professionals to implement data solutions.- Deep understanding of data modeling and database design.- Experience with data integration and data governance.- Strong knowledge of cloud platforms and big data technologies. <p>e. Machine Learning Engineer:</p> <ul style="list-style-type: none">- Design and develop machine learning models and algorithms.- Implement and maintain machine learning pipelines.- Collaborate with data scientists to deploy models into production.- Proficiency in machine learning frameworks like TensorFlow and PyTorch.- Experience with software engineering and DevOps practices.- Strong programming skills in Python or similar languages. <p>f. Chief Data Officer (CDO):</p> <ul style="list-style-type: none">- Develop and implement data strategy and governance policies.- Oversee data management and ensure compliance with regulations.- Lead data initiatives to drive business growth and innovation.- Strong leadership and strategic planning skills.- Deep understanding of data management and data governance.- Ability to communicate and align data strategy with business goals. <p>g. Interdependencies Between Roles:</p> <ul style="list-style-type: none">- Explain how these roles collaborate to manage the end-to-end data lifecycle.
--	--



	<p>- Highlight the importance of communication and collaboration between data engineers, data scientists, and analysts to ensure data quality and usability.</p> <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none">- Ask students to summarize the responsibilities and skills required for each key role.- Discuss the answers with the class. <p>Use Nearpod to collect responses and discuss the answers.</p>
Closure	<ol style="list-style-type: none">1. Summarize the lesson's learning outcomes and affirm understanding with students.2. Suggested Reading<ul style="list-style-type: none">- "Data Science and Big Data Analytics" by EMC Education Services (Editor). <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none">1. What are the key roles in the Big Data ecosystem?2. What are the main responsibilities and skills required for each role?3. Provide examples of how these roles interact and collaborate in an organization. <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>



Lesson Plan No. 8	Course Name: Introduction to Data Science Topic: Example of Big Data Analytics	Course No.: COM-501
-------------------	---	---------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand what Big Data analytics is and its significance. Identify the components and process of Big Data analytics. Explain a real-world example of Big Data analytics. Recognize the impact of Big Data analytics in various industries.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> - Discuss what students understand by "Big Data analytics." - Ask how they think Big Data analytics is applied in real-world scenarios. - Explain the relevance of Big Data analytics in uncovering insights and driving decision-making. Development (30 minutes) <ol style="list-style-type: none"> What is Big Data Analytics? <ul style="list-style-type: none"> ● Case Study: Healthcare Analytics <ul style="list-style-type: none"> - Scenario: Using Big Data analytics to predict patient outcomes and improve healthcare services. - Data Collection: Explain how patient data is collected from electronic health records (EHRs), wearable devices, and medical imaging. - Data Storage: Discuss the use of cloud storage and data lakes to store vast amounts of healthcare data. - Data Processing: Explain the processing of structured and unstructured data using Hadoop and Spark. - Data Analysis: Describe the use of machine learning algorithms to predict disease outbreaks, patient readmission rates, and personalize treatment plans. - Outcome: Highlight the benefits, such as improved patient care, reduced healthcare costs, and enhanced predictive capabilities. Impact of Big Data Analytics in Various Industries: <ul style="list-style-type: none"> - Retail: Personalizing customer experiences, optimizing supply chains, and improving inventory management. - Finance: Detecting fraud, managing risks, and performing algorithmic trading. - Manufacturing: Predicting equipment failures, optimizing production processes, and managing supply chains.



	<p>- Transportation: Optimizing routes, predicting maintenance needs, and improving safety.</p> <p>3. Exercise (5 minutes) –</p> <ul style="list-style-type: none">- Ask students to summarize the process and components of Big Data analytics.- Discuss the answers with the class. <p>Use Nearpod to collect responses and discuss the answers.</p>
Closure	<ol style="list-style-type: none">1. Summarize the lesson's learning outcomes and affirm understanding with students.2. Suggested Reading<ul style="list-style-type: none">- "Data Science and Big Data Analytics" by EMC Education Services (Editor).- "Data Science for Business" by Foster Provost and Tom Fawcett. <p>Spend 5 minutes to wrap up and consolidate the learnings</p>
Evaluation	<ol style="list-style-type: none">1. What are the main components of Big Data analytics?2. Describe the process of Big Data analytics.3. Provide a real-world example of Big Data analytics and its impact. <p>Spend 5 minutes to evaluate student assimilation of the lesson contents</p>

Lesson Plan No. 1	Course Name: Introduction to Data Science Topic: Lifecycle Overview	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	<p>At the end of the lesson the student shall be able to:</p> <ol style="list-style-type: none"> Understand the fundamental stages of the data science lifecycle. Identify the key tasks involved in each stage of the data science process. Explain the importance of each stage in deriving insights from data. Recognize the practical applications of the data science lifecycle in real-world scenarios.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "data science lifecycle"? Can you give examples of how data is used in decision-making in your daily life? Explain the relevance of understanding the lifecycle of data science in modern business and technology. Development (30 minutes) <ol style="list-style-type: none"> What is the Data Science Lifecycle? <ul style="list-style-type: none"> Definition: Explain the data science lifecycle as a series of iterative steps involved in data analysis, from data collection to actionable insights. Importance: Discuss the importance of a structured approach to data analysis to ensure accuracy, reproducibility, and actionable results. Stages of the Data Science Lifecycle: <ul style="list-style-type: none"> Data Collection: <ul style="list-style-type: none"> Explain the process of gathering raw data from various sources. Discuss different types of data sources such as databases, web scraping, and APIs. Data Preparation: <ul style="list-style-type: none"> Define data cleaning and preprocessing, including handling missing values, outliers, and data normalization. Emphasize the importance of data quality for accurate analysis. Data Exploration: <ul style="list-style-type: none"> Describe exploratory data analysis (EDA) techniques like visualization and summary statistics.

	<ul style="list-style-type: none"> - Highlight tools such as histograms, scatter plots, and correlation matrices. ● Data Modeling: <ul style="list-style-type: none"> - Explain the process of selecting and applying statistical or machine learning models to the prepared data. - Discuss different types of models such as regression, classification, clustering, etc. ● Model Evaluation: <ul style="list-style-type: none"> - Define metrics for evaluating model performance like accuracy, precision, recall, and F1 score. - Discuss cross-validation and the importance of preventing overfitting. ● Model Deployment: <ul style="list-style-type: none"> - Explain how models are integrated into production environments to provide ongoing predictions or insights. - Discuss challenges and best practices for deployment. ● Model Monitoring and Maintenance: <ul style="list-style-type: none"> - Describe the need for continuous monitoring of model performance. - Discuss updating models with new data to maintain accuracy over time. <p>c. Applications of the Data Science Lifecycle:</p> <ul style="list-style-type: none"> ● Real-World Applications: <ul style="list-style-type: none"> - Discuss examples like predictive maintenance in manufacturing, recommendation systems in e-commerce, and fraud detection in finance. - Highlight the role of data science in enhancing business strategies and operations. <p>d. Case Study:</p> <ul style="list-style-type: none"> - Walk through a real-world example illustrating the complete data science lifecycle, from data collection to model deployment and monitoring. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key stages of the data science lifecycle and their functions. - Discuss the answers with the class.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson learning outcomes and get affirmation from students on these. 2. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). - "Data Science for Business" by Foster Provost and Tom Fawcett, Chapter 2, pp. 25-32.
Evaluation	<ol style="list-style-type: none"> 1. How would you define the data science lifecycle in your own words? 2. What are the key stages of the data science lifecycle?

	<p>3. Why is data preparation essential in the data science process?</p> <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>
--	---

Lesson Plan No. 2	Course Name: Introduction to Data Science Topic: Discovery Phase	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the key components of the discovery phase in data science. Identify the importance of problem definition and data collection. Explain the significance of data exploration and understanding. Recognize the initial steps to prepare data for analysis.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "discovery phase" in data science? How do you approach solving a new problem or project? Explain the relevance of the discovery phase as the foundation for any data science project, setting the stage for successful data analysis and insights. Development (30 minutes) <ol style="list-style-type: none"> What is the Discovery Phase? <ul style="list-style-type: none"> -Definition: Introduce the discovery phase as the initial step in a data science project where the problem is defined, and relevant data is collected and explored. -Importance: Discuss how the discovery phase sets the Direction for the entire project, emphasizing its role in ensuring that the problem is well-understood and that data is appropriate and sufficient for the analysis. Key Components of the Discovery Phase: <ul style="list-style-type: none"> ● Problem Definition: <ul style="list-style-type: none"> -Explain the importance of clearly defining the problem or question to be addressed. Highlight techniques like stakeholder interviews and requirement gathering. ● Data Collection: <ul style="list-style-type: none"> -Discuss various sources of data, including internal databases, external datasets, APIs, and web scraping. Emphasize the importance of understanding data sources and ensuring data quality. ● Data Exploration: <ul style="list-style-type: none"> -Introduce techniques for initial data exploration, including summary statistics, data visualization, and identifying data types and distributions. ● Understanding Data:



	<p>- Explain how understanding the context and content of the data helps in making informed decisions about data cleaning and preparation.</p> <p>c. Tools and Techniques for Data Exploration:</p> <ul style="list-style-type: none"> ● Summary Statistics: <ul style="list-style-type: none"> - Describe how summary statistics (mean, median, mode, standard deviation) provide a quick overview of the data. ● Data Visualization: <ul style="list-style-type: none"> - Highlight the use of visual tools like histograms, scatter plots, and box plots to understand data distributions and relationships. ● Data Cleaning: <ul style="list-style-type: none"> - Briefly introduce data cleaning techniques such as handling missing values, removing duplicates, and correcting data types. <p>d. Applications of the Discovery Phase:</p> <ul style="list-style-type: none"> ● Real-World Examples: <ul style="list-style-type: none"> - Discuss examples of data science projects, such as customer segmentation, predictive maintenance, and sentiment analysis, where the discovery phase played a critical role. ● Case Study: <ul style="list-style-type: none"> - Walk through a simple case study illustrating the steps of the discovery phase in a real-world scenario. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key steps in the discovery phase and their importance. - Discuss the answers with the class.
<p>Closure</p>	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). - "Data Science for Business" by Foster Provost and Tom Fawcett, Chapter 2.
<p>Evaluation</p>	<ol style="list-style-type: none"> 1. How would you define the discovery phase in your own words? 2. What are the basic components of the discovery phase? 3. Why is the discovery phase essential in data science projects? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 3	Course Name: Introduction to Data Science Topic: Data Preparation Phase	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the significance of the data preparation phase in data science. Identify the key steps involved in data preparation. Apply various techniques for data cleaning, transformation, and integration. Recognize the challenges and best practices in preparing data for analysis.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "data preparation" in the context of data science? Why is data preparation crucial before conducting any data analysis? Explain the relevance of well-prepared data in ensuring accurate and reliable data science outcomes. Development (30 minutes) <ol style="list-style-type: none"> What is Data Preparation? <ul style="list-style-type: none"> - Definition: Explain data preparation as the process of collecting, cleaning, transforming, and organizing data for analysis. Emphasize its role in improving data quality and ensuring meaningful analysis. Importance of Data Preparation: <ul style="list-style-type: none"> - Discuss how proper data preparation can prevent errors, save time, and enhance the quality of insights derived from data. Highlight the importance of having clean, well-structured data for effective machine learning models. Key Steps in Data Preparation: <ul style="list-style-type: none"> ● Data Collection: <ul style="list-style-type: none"> - Gathering data from various sources (databases, web scraping, surveys). - Discuss challenges like data availability, accessibility, and volume. ● Data Cleaning: <ul style="list-style-type: none"> - Removing or correcting errors, duplicates, and inconsistencies. - Handling missing values by imputation or deletion.

	<ul style="list-style-type: none"> - Example: Use a dataset with missing values and demonstrate different imputation techniques. ● Data Transformation: <ul style="list-style-type: none"> - Normalizing or scaling numerical features. - Encoding categorical variables (one-hot encoding, label encoding). - Example: Demonstrate normalization and encoding on a sample dataset using a Jupyter Notebook. ● Data Integration: <ul style="list-style-type: none"> - Combining data from multiple sources to create a cohesive dataset. - Discuss issues like schema matching and data redundancy. - Example: Merge two datasets and resolve conflicts. ● Data Reduction: <ul style="list-style-type: none"> - Reducing the volume of data while maintaining its integrity (feature selection, dimensionality reduction). - Example: Apply PCA (Principal Component Analysis) to reduce the number of features. d. Real-World Applications: <ul style="list-style-type: none"> - Discuss examples like preparing data for predictive modeling in healthcare, financial analysis, and customer segmentation. - Highlight the use of data preparation in modern technologies such as AI and machine learning. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key steps in data preparation. - Discuss their summaries and clarify any misunderstandings.
Closure	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Reinforce the importance of each step in the data preparation phase. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you define data preparation in your own words? 2. What are the basic steps involved in data preparation? 3. Why is data cleaning essential in data science? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 4	Course Name: Introduction to Data Science Topic: Model Planning Phase	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	<p>At the end of the lesson the student shall be able to:</p> <ol style="list-style-type: none"> Understand the importance and objectives of the model planning phase in data science. Identify various techniques and tools used during the model planning phase. Explain the process of selecting and preparing data for modeling. Recognize the significance of model validation and evaluation.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Begin with a question: "What do you think are the steps involved in developing a data science model?" Discuss the role of the model planning phase in the overall data science lifecycle. Emphasize the relevance of careful planning to ensure successful model deployment. Development (30 minutes) <ol style="list-style-type: none"> What is the Model Planning Phase? <ul style="list-style-type: none"> Definition: Explain the model planning phase as the stage where the approach for building the model is decided. This includes selecting techniques, algorithms, and tools. Importance: Discuss how this phase lays the foundation for model development and its success. Highlight the need for planning to avoid common pitfalls and to ensure the model meets business objectives. Key Components of the Model Planning Phase: <ul style="list-style-type: none"> Selecting Techniques and Algorithms: <ul style="list-style-type: none"> Overview of different modeling techniques (e.g., regression, classification, clustering). Criteria for selecting appropriate algorithms based on the problem at hand. Data Preparation: <ul style="list-style-type: none"> Importance of data cleaning, transformation, and feature selection. Tools and techniques used for data preparation (e.g., Python libraries like Pandas, Scikit-learn). Model Validation and Evaluation: <ul style="list-style-type: none"> Discuss various methods for validating models (e.g., cross-validation, train-test split).

	<ul style="list-style-type: none"> - Importance of evaluating model performance using metrics like accuracy, precision, recall, F1 score. <p>c. Case Study:</p> <ul style="list-style-type: none"> - Walk through a simple case study where students can see the application of model planning in a real-world scenario. - Include steps from data selection to choosing the appropriate algorithm and evaluating the model. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to outline the steps they would take in the model planning phase for a given data science problem. - Discuss the answers with the class, highlighting different approaches and best practices.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson by reviewing the key points: <ul style="list-style-type: none"> - Importance of model planning - Techniques and tools used - Steps in data preparation and model validation 2. Confirm understanding by asking students to reflect on what they learned. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you describe the model planning phase in your own words? 2. What are the essential steps in preparing data for modeling? 3. Why is it important to validate and evaluate a model? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 5	Course Name: Introduction to Data Science Topic: Model Building Phase	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the key stages in the model-building phase of data science. Identify different types of models used in data science. Explain the importance of model evaluation and selection. Apply basic techniques for model training and validation.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "model building" in data science? How do you think models impact decision-making in real-world scenarios? Explain the relevance of model building in extracting insights from data. Development (30 minutes) <ol style="list-style-type: none"> What is Model Building in Data Science? <ul style="list-style-type: none"> Definition: Explain model building as the process of creating algorithms that learn patterns from data to make predictions or decisions. Importance: Discuss how model building is crucial for applications like predicting customer behavior, diagnosing diseases, and personalizing recommendations. Stages of Model Building: <ul style="list-style-type: none"> Data Preparation: <ul style="list-style-type: none"> Data Cleaning: Removing noise and handling missing values. Data Transformation: Scaling, normalizing, and encoding data. Feature Selection: Identifying the most relevant variables for the model. Model Selection: <ul style="list-style-type: none"> Types of Models: Supervised (e.g., regression, classification) and Unsupervised (e.g., clustering). Criteria for Selection: Performance, interpretability, and computational efficiency. Model Training: <ul style="list-style-type: none"> Training Algorithms: How models learn from data.

	<p>Cross-Validation: Technique to ensure the model generalizes well to unseen data.</p> <ul style="list-style-type: none"> ● Model Evaluation: <ul style="list-style-type: none"> - Metrics: Accuracy, precision, recall, F1-score, RMSE, etc. - Validation Techniques: Holdout method, k-fold cross-validation. ● Model Tuning: <ul style="list-style-type: none"> - Hyperparameter Optimization: Methods like Grid Search and Random Search to find the best model parameters. <p>c. Applications of Model Building:</p> <ul style="list-style-type: none"> ● Real-World Examples: <ul style="list-style-type: none"> - Predicting stock prices, fraud detection, and recommendation systems. - Case Study: Walk through the creation of a simple classification model for predicting customer churn. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key stages in the model-building phase. - Discuss the answers with the class.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson learning outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you define model building in your own words? 2. What are the key stages of model building in data science? 3. Why is cross-validation important in the model-building process? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 6	Course Name: Introduction to Data Science Topic: Communicated Result Phase	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	<p>At the end of the lesson the student shall be able to:</p> <ol style="list-style-type: none"> Understand the concept of the communicated result phase in the data science lifecycle. Identify the key elements and techniques used in communicating results. Explain the significance of effectively communicating data science results to stakeholders. Recognize various tools and best practices for presenting data findings.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Question to Class: What do you think is the importance of communication in data science? Discuss the role of the communicated result phase in the overall data science lifecycle. Explain the relevance of clear and effective communication in ensuring data-driven decisions. Development (30 minutes) <ol style="list-style-type: none"> Definition and Importance: <ul style="list-style-type: none"> Definition: Introduce the communicated result phase as the stage where data findings are summarized, interpreted, and presented to stakeholders. Importance: Emphasize how this phase bridges the gap between data analysis and decision-making, making complex data insights accessible and actionable. Key Elements of Communicated Results: <ul style="list-style-type: none"> Data Visualization: Discuss the use of charts, graphs, and other visual tools to make data understandable. Storytelling with Data: Explain how to craft a narrative around data to make it compelling and relevant to the audience. Report Writing: Highlight the structure and components of an effective data science report. Techniques and Best Practices: <ul style="list-style-type: none"> Choosing the Right Visuals: Guide students on selecting appropriate visual representations based on the data type and audience. Simplifying Complex Information: Techniques for breaking down complex data into digestible insights.

	<ul style="list-style-type: none"> - Interactive Dashboards: Introduction to tools like Tableau or Power BI for creating interactive data presentations. <p>d. Case Study and Example:</p> <ul style="list-style-type: none"> - Real-World Example: Walk through a case study where data communication played a critical role in business decisions. - YouTube Video: Show a video on data storytelling to illustrate how effective communication can transform data insights. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Activity: Ask students to create a simple data visualization using sample data provided. - Discussion: Have students present their visualizations and discuss the choices they made.
Closure	<ol style="list-style-type: none"> 1. Summarize Lesson: Recap the key points about the importance of the communicated result phase, key elements, techniques, and best practices. 2. Affirmation: Ensure students understand the lesson outcomes and are able to apply them. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you define the communicated result phase in your own words? 2. What are the key elements of effective data communication? 3. Why is storytelling important in presenting data findings? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 7	Course Name: Introduction to Data Science Topic: Operationalize	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the key concepts of the operationalize phase in data science. Identify different steps involved in operationalizing a data science project. Explain the importance of deploying, monitoring, and maintaining data science models.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "operationalize" in the context of data science? Have you encountered any deployed data science models in real-world applications? Explain the relevance of the operationalize phase in ensuring data science projects deliver continuous value. Development (30 minutes) <ol style="list-style-type: none"> What is the Operationalize Phase? <ul style="list-style-type: none"> Definition: Explain the operationalize phase as the process of deploying data science models into production environments where they can be used by end-users or systems. Emphasize its role in making models actionable and valuable. Steps in the Operationalize Phase: <ul style="list-style-type: none"> ● Model Deployment: <ul style="list-style-type: none"> Describe different deployment options (e.g., batch deployment, real-time deployment, etc.). Discuss the tools and platforms commonly used for deployment (e.g., Docker, Kubernetes, cloud platforms). ● Monitoring and Maintenance: <ul style="list-style-type: none"> Explain the importance of monitoring model performance and accuracy over time. Discuss common metrics for monitoring models (e.g., accuracy, precision, recall). Highlight the need for regular maintenance and updates to models. ● Scaling and Optimization: <ul style="list-style-type: none"> Talk about the need to scale models to handle increased loads. Discuss techniques for optimizing model performance.

	<p>c. Importance of the Operationalize Phase:</p> <ul style="list-style-type: none"> ● Ensuring Continuous Value: <ul style="list-style-type: none"> - Discuss how deployed models provide ongoing value through automation and improved decision-making. ● Feedback Loop: <ul style="list-style-type: none"> - Highlight the importance of feedback loops in improving models over time based on real-world data. ● Real-World Impact: <ul style="list-style-type: none"> - Provide examples of operationalized models in various industries (e.g., healthcare, finance, retail). <p>d. Challenges and Best Practices:</p> <ul style="list-style-type: none"> ● Common Challenges: <ul style="list-style-type: none"> - Discuss potential issues such as model drift, data privacy concerns, and integration complexities. ● Best Practices: <ul style="list-style-type: none"> - Provide guidelines on best practices, including thorough testing before deployment, continuous monitoring, and maintaining clear documentation. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to outline the steps involved in operationalizing a data science project. - Discuss the answers with the class, highlighting key points.
Closure	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you define the operationalize phase in data science in your own words? 2. What are the basic steps involved in operationalizing a data science project? 3. Why is monitoring essential after deploying a model? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 8	Course Name: Introduction to Data Science Topic: Case Study - Global Innovation Network and Analysis	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept of a global innovation network. Analyze the key components and functions of a global innovation network. Identify and discuss real-world examples of global innovation networks. Evaluate the impact of global innovation networks on business strategy and operations.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> What do you understand by the term "global innovation network"? Can you think of any companies that operate on a global scale to innovate? Explain the importance of innovation in today's globalized business environment. Development (30 minutes) <ol style="list-style-type: none"> What is a Global Innovation Network? <ul style="list-style-type: none"> Definition: A global innovation network is a system of interconnected entities (companies, research institutions, governments) across different countries working collaboratively to drive innovation. Importance: Discuss how global innovation networks facilitate the flow of knowledge, ideas, and technologies across borders, enhancing innovation capabilities. Key Components of a Global Innovation Network: <ul style="list-style-type: none"> Nodes: The participants in the network (e.g., firms, universities, research centers). Links: The connections between nodes through collaborations, partnerships, and joint ventures. Flows: The exchange of information, resources, and technologies within the network. Governance: The policies and rules governing the network's operations. Real-World Examples: <ul style="list-style-type: none"> Example 1: IBM's Global Research Network, Overview of IBM's research centers across the globe.

	<ul style="list-style-type: none"> - Discuss the collaborative projects and innovations stemming from this network. ● Example 2: The Pharmaceutical Industry <ul style="list-style-type: none"> - How pharmaceutical companies collaborate globally for R&D. - Impact of these networks on drug discovery and development. d. Impact on Business Strategy and Operations: <ul style="list-style-type: none"> - Competitive Advantage:How being part of a global innovation network can provide a competitive edge. - Efficiency and Speed:Enhancing operational efficiency and accelerating time-to-market. -Market Expansion:Facilitating entry into new markets and access to new customer bases. e. Case Study Analysis (15 minutes) <ul style="list-style-type: none"> - Present a case study of a successful global innovation network. - Walk through the components and functions as outlined above. - Discuss the outcomes and benefits observed from the case study. <p>3. Exercise (10 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key components and functions of a global innovation network. - Discuss their summaries and clarify any doubts.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson's learning outcomes and confirm understanding with students. 2. Highlight the importance of global innovation networks in modern business strategy. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor).
Evaluation	<ol style="list-style-type: none"> 1. How would you define a global innovation network in your own words? 2. What are the key components of a global innovation network? 3. Why are global innovation networks essential in today's business environment? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No.1	Course Name: Introduction to Data Science Topic: Introduction to R	Course No.: COM-501
-------------------------	---	----------------------------

Objectives	<p>At the end of the lesson the student shall be able to:</p> <ol style="list-style-type: none"> Understand the basic concepts and features of the R programming language. Install and set up R and RStudio. Write and execute basic R scripts. Understand and use basic data structures in R (vectors, data frames, lists, and matrices). Perform simple data manipulation and visualization using R.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What programming languages have you heard of or used before? How do you think programming can help in data analysis? Introduce R as a powerful tool for statistical computing and graphics. Briefly discuss the importance of R in data science and its growing popularity. Development (30 minutes) <ol style="list-style-type: none"> What is R? <ul style="list-style-type: none"> -Definition:R is a programming language and software environment used for statistical computing, data analysis, and graphical representation. -Features:Open-source,Extensive library of packages,Strong graphical capabilities Setting Up R and RStudio: <ul style="list-style-type: none"> -Steps:Guide students through the installation process of R and RStudio.Demonstrate the Rstudio interface, including the console, script editor, environment, and plots pane. Basic R Syntax and Operations: <ul style="list-style-type: none"> ● Writing and Executing Scripts: <ul style="list-style-type: none"> -Show how to create a new script and write simple commands. -Execute the script and view the results in the console. Data Structures in R: <ul style="list-style-type: none"> -Vectors:Explain and demonstrate the creation and manipulation of vectors. -Data Frames:Introduce data frames as tables for storing data and perform basic operations. -Lists:Explain lists and show how they can store different types of data.

	<p>-Matrices: Demonstrate creating and manipulating matrices.</p> <p>e. Basic Data Manipulation and Visualization:</p> <p>-Data Manipulation: Introduce functions like subset(), merge(), and basic data transformation techniques.</p> <p>-Visualization: Demonstrate simple plots using the plot() function.</p> <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to write and execute a simple R script that creates a vector and plots it. - Discuss the output and answer any questions.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson learning outcomes and get affirmation from students on these. 2. Encourage students to explore R further using online resources and tutorials. 3. Suggested Reading: <ul style="list-style-type: none"> - “Introduction to Data Science: Data Analysis and Prediction Algorithms with R” by Rafael A. Irizarry. 4. Homework: <ul style="list-style-type: none"> - Write a short R script that creates a data frame, adds a new column, and generates a plot of one of the columns. Upload the script on the CAMU..
Evaluation	<ol style="list-style-type: none"> 1. How would you describe the R programming language in your own words? 2. What are the basic data structures in R and their uses? 3. Why is R a valuable tool for data analysis? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 2	Course Name: Introduction to Data Science Topic: R Graphical User Interface	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept of graphical user interfaces (GUIs) in R. Identify different types of GUIs available for R. Explain the functionality and features of popular R GUIs. Utilize R GUIs to perform basic data analysis and visualization tasks.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "graphical user interface"? How often do you use GUIs for data analysis in your daily work? Explain the importance of GUIs in simplifying data analysis and visualization tasks in R. Development (30 minutes) <ol style="list-style-type: none"> What are R Graphical User Interfaces? <ul style="list-style-type: none"> -Definition: Explain R GUIs as software applications that provide a visual interface for R, allowing users to perform data analysis tasks without extensive coding. -Importance: Discuss how R GUIs make data analysis more accessible to beginners and enhance productivity for experienced users. Types of R Graphical User Interfaces: <ul style="list-style-type: none"> ● R Commander: <ul style="list-style-type: none"> -Overview: A basic-statistics GUI for R. -Features: Data import/export, statistical analysis, and graph creation. -Demonstration: Walk through a simple data analysis task using R Commander. ● Rattle: <ul style="list-style-type: none"> -Overview: A GUI for data mining using R. -Features: Data manipulation, statistical modeling, and evaluation. -Demonstration: Show a basic data mining workflow using Rattle. ● RStudio: <ul style="list-style-type: none"> -Overview: An integrated development environment (IDE) for R. -Features: Script editor, data viewer, package management, and plotting capabilities.

	<p>-Demonstration: Highlight key features of RStudio and how it can be used for data analysis.</p> <p>c. Applications of R GUIs:</p> <ul style="list-style-type: none"> ● Real-World Applications: <ul style="list-style-type: none"> -Discuss examples like data cleaning, statistical analysis, and creating visualizations. -Highlight the use of R GUIs in industries such as healthcare, finance, and academia. ● Case Study: <ul style="list-style-type: none"> -Walk through a real-world example using R GUIs to analyze a dataset and generate insights. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the key features of R Commander, Rattle, and RStudio. - Discuss the answers with the class.
Closure	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 3. Homework: <ul style="list-style-type: none"> - Write a short essay on how R GUIs can enhance productivity in data analysis. - <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>
Evaluation	<ol style="list-style-type: none"> 1. How would you define R graphical user interfaces in your own words? 2. What are the main features of R Commander, Rattle, and RStudio? 3. Why are GUIs essential for performing data analysis in R? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No. 3	Course Name: Introduction to Data Science Topic: Data Import and Export	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	<p>At the end of the lesson the student shall be able to:</p> <ol style="list-style-type: none"> Understand the basic concepts of data import and export in R. Identify different file formats and their functions in data handling. Demonstrate the ability to import data from various file types into R. Demonstrate the ability to export data from R into various file types. Recognize the applications of data import and export in data analysis and research.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "data import and export" in programming? How often do you work with different data file types in your projects? Explain the relevance of data import and export in data analysis. Development (30 minutes) <ol style="list-style-type: none"> What is Data Import and Export? <ul style="list-style-type: none"> Definition: <ul style="list-style-type: none"> -Explain data import as the process of bringing external data into R for analysis and manipulation. -Explain data export as the process of saving data from R to external files for sharing or further use. -Emphasize their importance in data analysis, allowing for the integration and dissemination of data. Common Data Formats: <ul style="list-style-type: none"> CSV Files: <ul style="list-style-type: none"> -Discuss how to import and export CSV files using read.csv() and write.csv(). -Example: data <- read.csv("data.csv"), write.csv(data, "output.csv") Excel Files: <ul style="list-style-type: none"> -Discuss how to import and export Excel files using readxl and writexl packages. -Example: data <- read_excel("data.xlsx"), write_xlsx(data, "output.xlsx") Text Files:

	<p>-Discuss how to import and export text files using read.table() and write.table().</p> <p>-Example: data <- read.table("data.txt"), write.table(data, "output.txt")</p> <ul style="list-style-type: none"> ● R Data Files: <p>-Discuss how to save and load R data files using save() and load().</p> <p>-Example: save(data, file = "data.RData"), load("data.RData")</p> ● Other Formats: <p>-Briefly mention JSON, XML, and SQL databases and their handling in R using respective packages (jsonlite, XML, DBI).</p> <p>c. Practical Demonstration:</p> <p>-Walk through importing a CSV file into R and performing basic data manipulation.</p> <p>-Walk through exporting the manipulated data into a new CSV file.</p> <p>-Showcase the use of RStudio for data import/export operations.</p> <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to import a sample CSV file into R and summarize the data. - Ask students to export a modified version of the data into a new CSV file.
Closure	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). - "Hands-On Programming with R" by Garrett Golemund (Chapter 7: Data Import and Export)
Evaluation	<ol style="list-style-type: none"> 1. How would you define data import and export in your own words? 2. What are the common file formats used in data import and export in R? 3. Demonstrate the process of importing a CSV file and exporting a modified version of it in R. <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No.4	Course Name: Introduction to Data Science Topic: Attributes and Data Types	Course No.: COM-501
-------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the basic attributes in R. Identify different data types in R. Explain the significance of data types in data analysis. Recognize the applications of various data types in R programming.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<p>1. Introduction (5 minutes)</p> <ul style="list-style-type: none"> - Ask the class: "What do you understand by the term 'data types'?" - Discuss the importance of data types in programming and data analysis. - Introduce the topic by explaining how R handles different types of data and their attributes. <p>2. Development (30 minutes)</p> <p>a. Attributes in R:</p> <ul style="list-style-type: none"> -Definition: Attributes are metadata, or data about data, which describe the properties of objects in R.Example: names, dim, class, length, etc. -Importance: Explain how attributes provide additional information about R objects that can be used in data manipulation and analysis. <p>b. Data Types in R:</p> <ul style="list-style-type: none"> ● Vectors: <ul style="list-style-type: none"> -Definition: The most basic data type in R, representing a sequence of elements of the same type. -Types of vectors: <ul style="list-style-type: none"> Numeric: Contains numbers, e.g., c(1, 2, 3) Character: Contains text, e.g., c("a", "b", "c") Logical: Contains TRUE or FALSE values, e.g., c(TRUE, FALSE) Integer: Contains integer values, e.g., c(1L, 2L, 3L) Complex: Contains complex numbers, e.g., c(1+2i, 3+4i) ● Matrices: <ul style="list-style-type: none"> -Definition: Two-dimensional arrays where elements are arranged in rows and columns.

	<p>-Usage: Used for mathematical computations and data manipulation.</p> <ul style="list-style-type: none"> ● Arrays: <ul style="list-style-type: none"> -Definition: Multi-dimensional generalization of matrices. -Usage: Suitable for data with more than two dimensions. ● Data Frames: <ul style="list-style-type: none"> -Definition: Table-like structures where columns can contain different types of data. -Usage: Commonly used for storing datasets. ● Lists: <ul style="list-style-type: none"> -Definition: Collection of elements that can be of different types. -Usage: Flexible structure that can hold various data types including other lists. ● Factors: <ul style="list-style-type: none"> -Definition: Used for categorical data. -Usage: Useful in statistical modeling and data analysis. <p>c. Applications of Data Types:</p> <ul style="list-style-type: none"> -Discuss how different data types are used in real-world data analysis. -Example: Numeric data types for statistical calculations, character data types for text analysis, data frames for storing and manipulating datasets. <p>3. Exercise (10 minutes)</p> <ul style="list-style-type: none"> - Hands-on Activity: Provide a simple dataset and ask students to identify and modify the data types and attributes. - Discussion: Go over the students' findings and discuss any challenges they faced.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson's learning outcomes and confirm understanding with the students. 2. Suggested Reading <ul style="list-style-type: none"> - “Introduction to Data Science: Data Analysis and Prediction Algorithms with R” by Rafael A. Irizarry.
Evaluation	<ol style="list-style-type: none"> 1. How would you define attributes in R? 2. What are the basic data types in R? 3. Why are data types important in data analysis? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>



Lesson Plan No. 5	Course Name: Introduction to Data Science Topic: Descriptive Statistics	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the basic concepts of descriptive statistics. Calculate measures of central tendency (mean, median, mode). Compute measures of variability (range, variance, standard deviation). Interpret and present statistical data using tables, charts, and graphs.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "descriptive statistics"? How can descriptive statistics be used in everyday life? Explain the importance of descriptive statistics in summarizing and understanding data. Development (30 minutes) <ol style="list-style-type: none"> What is Descriptive Statistics? <ul style="list-style-type: none"> -Definition: Descriptive statistics involves methods for summarizing and organizing data to make it easily understandable. -Importance: These statistics provide simple summaries about the sample and the measures. They form the basis for virtually every quantitative analysis of data. Measures of Central Tendency: <ul style="list-style-type: none"> -Mean: Explain how to calculate the mean (average) of a dataset. -Median: Define and demonstrate how to find the median (the middle value) in a dataset. -Mode: Describe and identify the mode (the most frequently occurring value) in a dataset. Measures of Variability: <ul style="list-style-type: none"> -Range: Calculate the range (difference between the highest and lowest values) of a dataset. -Variance and Standard Deviation: Explain and compute variance and standard deviation to understand data spread and variability. Data Presentation: <ul style="list-style-type: none"> • Tables and Charts: <ul style="list-style-type: none"> -Show how to organize data in tables and represent it using bar charts, histograms, and pie charts. • Graphical Interpretation:

	<p>-Emphasize the importance of visual data representation for better interpretation.</p> <p>e. Practical Application:</p> <p>-Case Study:Analyze a dataset using descriptive statistics measures.</p> <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to calculate the mean, median, mode, range, variance, and standard deviation for a given dataset. - Discuss the answers with the class.
Closure	<ol style="list-style-type: none"> 1. Summarize the key learning outcomes: understanding measures of central tendency and variability, and the ability to present data graphically. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 3. Homework: <ul style="list-style-type: none"> - Complete a worksheet involving the calculation and interpretation of descriptive statistics for a given dataset.
Evaluation	<ol style="list-style-type: none"> 1. How would you define descriptive statistics in your own words? 2. What are the key measures of central tendency and variability? 3. Why is it important to present data graphically? 4. <p>Spend 5 minutes reviewing and consolidating the lesson's learning outcomes.</p>

Lesson Plan No. 6	Course Name: Introduction to Data Science Topic: Exploratory Data Analysis (EDA)	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept and importance of Exploratory Data Analysis. Identify and use various EDA techniques to summarize data. Utilize graphical and non-graphical methods for data exploration. Apply EDA techniques to uncover patterns, spot anomalies, and test hypotheses.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "Exploratory Data Analysis"? How can EDA help in understanding the structure and patterns of a dataset? Explain the relevance of EDA in the data analysis process and its role in data-driven decision-making. Development (30 minutes) <ol style="list-style-type: none"> What is Exploratory Data Analysis? <ul style="list-style-type: none"> Definition: <ul style="list-style-type: none"> -Explain EDA as a statistical approach used to analyze datasets to summarize their main characteristics, often using visual methods. -Highlight the importance of EDA in the initial stages of data analysis to make sense of data before applying more complex techniques. Importance of Exploratory Data Analysis: <ul style="list-style-type: none"> Discuss how EDA helps in: <ul style="list-style-type: none"> -Identifying data patterns and relationships -Detecting anomalies and outliers -Formulating hypotheses -Guiding further data processing and analysis steps Basic Techniques of EDA: <ul style="list-style-type: none"> Summary Statistics: <ul style="list-style-type: none"> -Mean, Median, Mode: Central tendency measures to understand the dataset's distribution. -Standard Deviation and Variance: Measures of data dispersion.

	<p>-Percentiles and Quartiles: To understand the distribution spread.</p> <ul style="list-style-type: none"> ● Graphical Methods: <ul style="list-style-type: none"> -Histograms: For visualizing the distribution of a single variable. -Box Plots: To identify outliers and understand data spread. -Scatter Plots: For identifying relationships between two variables. -Pair Plots: To visualize relationships across multiple variables. -Heatmaps: To show the correlation between variables. <p>d. Real-World Applications of EDA:</p> <ul style="list-style-type: none"> -Discuss examples such as market analysis, healthcare data assessment, and social media trend analysis. -Highlight case studies where EDA provided insights leading to successful outcomes. <p>e. Case Study:</p> <ul style="list-style-type: none"> ● Walk through an EDA example using a sample dataset. Demonstrate using R or Python to: <ul style="list-style-type: none"> -Load the dataset -Perform summary statistics -Create various plots (histograms, box plots, scatter plots) -Interpret the results <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to perform a simple EDA on a given dataset and summarize their findings. - Discuss the answers and different insights obtained by students.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson learning outcomes and affirm the students' understanding of EDA concepts. 2. Encourage students to practice EDA on different datasets to gain deeper insights and improve their analytical skills. 3. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 4. Homework: <ul style="list-style-type: none"> - Write a short report on performing EDA on a chosen dataset and upload it to the learning management system.
Evaluation	<ol style="list-style-type: none"> 1. How would you define Exploratory Data Analysis in your own words? 2. What are the basic techniques used in EDA? 3. Why is EDA crucial in the data analysis process? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>



Kot Bhalwal, Jammu

Model Institute of Engineering & Technology (Autonomous) Lesson Plan



Dr. Arun K. Gupta Teaching-Learning Centre

Version 1.1

श्रेष्ठ श्रम नवीनता

Please Do Not Print Unless Necessary

Lesson Plan No. 7	Course Name: Introduction to Data Science Topic: Statistical Methods for Evaluation	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the basic concepts of statistical evaluation methods. Identify different statistical methods used for data evaluation. Apply statistical methods to real-world datasets. Interpret results obtained from statistical analyses.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> What do you understand by the term "statistical evaluation"? Can you recall a situation where statistical evaluation was used to make a decision? Explain the importance of statistical methods in research and data analysis. Development (30 minutes) <ol style="list-style-type: none"> What are Statistical Methods for Evaluation? <ul style="list-style-type: none"> -Definition: Explain statistical methods as techniques used to collect, analyze, interpret, and present data. Emphasize their role in making informed decisions. -Importance: Discuss how statistical methods help in understanding trends, testing hypotheses, and making predictions. Highlight their significance in various fields like healthcare, business, and social sciences. Basic Statistical Methods: <ul style="list-style-type: none"> -Descriptive Statistics: Methods for summarizing and organizing data (e.g., mean, median, mode, standard deviation). -Inferential Statistics: Techniques for making generalizations from a sample to a population (e.g., hypothesis testing, confidence intervals). -Regression Analysis: Understanding relationships between variables and making predictions. -ANOVA (Analysis of Variance): Comparing means across multiple groups to see if there are significant differences. Chi-Square Test: Testing relationships between categorical variables. Applications of Statistical Methods: <ul style="list-style-type: none"> -Real-World Applications: Discuss examples like evaluating the effectiveness of a new drug, analyzing market trends, and assessing student performance. -Case Study: Walk through a real-world dataset and apply various statistical methods to analyze and interpret the data. Exercise (5 minutes)

	<ul style="list-style-type: none"> - Ask students to summarize the key statistical methods discussed and their applications. - Discuss the answers with the class to reinforce understanding.
Closure	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading: <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). - "Introduction to Data Science: Data Analysis and Prediction Algorithms with R" by Rafael A. Irizarry. 3. Homework <ul style="list-style-type: none"> - Analyze a given dataset using descriptive and inferential statistical methods. Write a report on your findings and submit it on the CAMU.
Evaluation	<ol style="list-style-type: none"> 1. How would you define statistical evaluation in your own words? 2. What are the basic statistical methods used for data evaluation? 3. Why is regression analysis important in understanding relationships between variables? <p>Spend 5 minutes to wrap up and consolidate the learnings.</p>

Lesson Plan No.8	Course Name: Introduction to Data Science Topic: Hypothesis Testing	Course No.: COM-501
-------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept of hypothesis testing. Identify the null and alternative hypotheses. Explain the process of conducting a hypothesis test. Recognize different types of errors in hypothesis testing. Apply hypothesis testing to real-world scenarios.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Begin with a discussion: What do you understand by the term "hypothesis"? Ask: How would you test a claim or theory in real life? Introduce the relevance of hypothesis testing in making data-driven decisions. Development (30 minutes) <ol style="list-style-type: none"> What is Hypothesis Testing? <ul style="list-style-type: none"> -Definition: Explain hypothesis testing as a statistical method that uses sample data to evaluate a hypothesis about a population parameter. -Importance: Discuss the significance of hypothesis testing in research and decision-making across various fields such as medicine, economics, and social sciences. Basic Components of Hypothesis Testing: <ul style="list-style-type: none"> ● Null Hypothesis (H_0): <ul style="list-style-type: none"> -Define the null hypothesis as a statement of no effect or no difference. -Example: "There is no difference in average test scores between two groups." ● Alternative Hypothesis (H_1): <ul style="list-style-type: none"> -Define the alternative hypothesis as a statement indicating the presence of an effect or difference. -Example: "There is a difference in average test scores between two groups." ● Significance Level (α): <ul style="list-style-type: none"> -Explain the significance level as the probability of rejecting the null hypothesis when it is actually true. -Commonly used values: 0.05, 0.01. ● Test Statistic: <ul style="list-style-type: none"> -Describe the test statistic as a standardized value used to determine the probability of observing the sample data if the null hypothesis is true.



	<ul style="list-style-type: none"> ● P-value: <ul style="list-style-type: none"> -Explain the p-value as the probability of obtaining a test statistic at least as extreme as the one observed, assuming the null hypothesis is true. ● Decision Rule: <ul style="list-style-type: none"> -Discuss how to use the p-value to make a decision: If p-value $\leq \alpha$, reject H₀; otherwise, fail to reject H₀. <p>c. Types of Errors:</p> <ul style="list-style-type: none"> ● Type I Error (α): <ul style="list-style-type: none"> -Define Type I error as the incorrect rejection of a true null hypothesis. ● Type II Error (β): <ul style="list-style-type: none"> -Define Type II error as the failure to reject a false null hypothesis. <p>d. Steps in Hypothesis Testing:</p> <ul style="list-style-type: none"> -Formulate the null and alternative hypotheses. -Choose the significance level. -Collect and summarize the data. -Calculate the test statistic and p-value. -Make a decision and interpret the results. <p>e. Applications of Hypothesis Testing:</p> <ul style="list-style-type: none"> ● Real-World Applications: <ul style="list-style-type: none"> -Discuss examples like testing the effectiveness of a new drug, comparing average incomes between two regions, or assessing the impact of an educational intervention. ● Case Study: <ul style="list-style-type: none"> -Walk through a simple example using sample data to test a hypothesis. <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to formulate null and alternative hypotheses for a given scenario. - Discuss the answers with the class.
<p>Closure</p>	<ol style="list-style-type: none"> 1. Summarize the Lesson Learning Outcomes and get affirmation from students on these. 2. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 3. Homework <ul style="list-style-type: none"> - Analyze a provided dataset and conduct a hypothesis test. Write a short report on your findings and upload it on the learning platform.
<p>Evaluation</p>	<ol style="list-style-type: none"> 1. How would you define hypothesis testing in your own words? 2. What are the basic components of hypothesis testing? 3. Why are significance levels important in hypothesis tests?



Kot Bhalwal, Jammu

Model Institute of Engineering
& Technology (Autonomous)
Lesson Plan

	Spend 5 minutes to wrap up and consolidate the learnings.
--	---

Lesson Plan No. 9	Course Name: Introduction to Data Science Topic: Type 1 and Type 2 Errors	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concepts of Type 1 and Type 2 errors. Differentiate between Type 1 and Type 2 errors. Explain the implications of these errors in hypothesis testing. Identify examples of Type 1 and Type 2 errors in real-world scenarios.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Ask students: "What do you know about hypothesis testing?" Introduce the concepts of Type 1 and Type 2 errors. Explain the importance of understanding these errors in the context of statistical analysis and decision-making. Development (30 minutes) <ol style="list-style-type: none"> Definition: <ul style="list-style-type: none"> Type 1 Error: Occurs when the null hypothesis is true, but we reject it. Also known as a false positive. Type 2 Error: Occurs when the null hypothesis is false, but we fail to reject it. Also known as a false negative. Significance: <ul style="list-style-type: none"> Discuss the implications of making Type 1 and Type 2 errors in research and data analysis. Emphasize the balance between these errors and the role of significance level (alpha) and power of the test (beta). Examples and Real-World Applications: <ul style="list-style-type: none"> Provide examples of Type 1 and Type 2 errors in different fields such as medicine, quality control, and criminal justice. <ul style="list-style-type: none"> Medicine: Type 1 error could mean incorrectly diagnosing a patient as having a disease when they do not. Type 2 error could mean failing to diagnose a patient who actually has the disease. Quality Control: Type 1 error could mean rejecting a good product. Type 2 error could mean accepting a defective product. Criminal Justice: Type 1 error could mean convicting an innocent person. Type 2 error could mean acquitting a guilty person. Exercise (5 minutes) <ul style="list-style-type: none"> Ask students to provide their own examples of Type 1 and Type 2 errors in pairs or small groups.

	<ul style="list-style-type: none"> - Discuss these examples with the class and clarify any misunderstandings.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson and reinforce the key points: definitions, significance, and examples of Type 1 and Type 2 errors. 2. Confirm that students have achieved the lesson objectives through a quick recap and question-answer session. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 4. Homework <ul style="list-style-type: none"> - Write a short essay on the potential consequences of Type 1 and Type 2 errors in a chosen field (e.g., healthcare, manufacturing, finance) and upload it to the learning platform.
Evaluation	<ol style="list-style-type: none"> 1. How would you define Type 1 and Type 2 errors in your own words? 2. Provide examples of both types of errors. 3. Explain the significance of understanding these errors in hypothesis testing. <p>Spend the last 5 minutes consolidating the learning outcomes and addressing any remaining questions.</p>

Lesson Plan No.10	Course Name: Introduction to Data Science Topic: Analysis of Variance (ANOVA)	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept and purpose of ANOVA. Identify different types of ANOVA (one-way, two-way, etc.). Perform ANOVA calculations manually and using software. Interpret the results of ANOVA and understand their significance in research.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector You Tube video
Teaching Development	<ol style="list-style-type: none"> Introduction (5 minutes) <ul style="list-style-type: none"> Ask: "What do you understand by the term 'Analysis of Variance'?" Discuss the relevance of ANOVA in comparing multiple group means. Briefly introduce the history and application of ANOVA in various fields. Development (30 minutes) <ol style="list-style-type: none"> What is ANOVA? <ul style="list-style-type: none"> -Definition: Explain ANOVA as a statistical method used to compare means of three or more groups to see if at least one group mean is significantly different from the others. -Importance: Discuss the necessity of ANOVA in research when dealing with multiple groups, as it helps in understanding the overall variance within the data. Types of ANOVA: <ul style="list-style-type: none"> -One-Way ANOVA: Used when comparing the means of three or more independent groups. -Two-Way ANOVA: Used when examining the influence of two different categorical independent variables on one continuous dependent variable. -Repeated Measures ANOVA: Used when the same subjects are used for each treatment (e.g., in a longitudinal study). Performing ANOVA: <ul style="list-style-type: none"> -Manual Calculation: Walkthrough of the ANOVA table: Sources of variance (Between Groups, Within Groups), Sum of Squares, Degrees of Freedom, Mean Squares, F-Statistic. -Using Software: Demonstrate how to perform ANOVA using statistical software like R or SPSS. Interpreting Results: <ul style="list-style-type: none"> -F-Statistic and P-Value: Explain the F-Statistic and its role in determining the ratio of variances.

	<p>-Discuss the significance of the P-Value in hypothesis testing. -Post-Hoc Tests: Introduce post-hoc tests (e.g., Tukey's HSD) to determine which specific groups differ.</p> <p>e. Applications of ANOVA: -Provide real-world examples of how ANOVA is used in fields like psychology, medicine, agriculture, and business.</p> <p>3. Exercise (5 minutes)</p> <ul style="list-style-type: none"> - Ask students to summarize the steps of performing ANOVA. - Discuss the answers with the class to reinforce understanding.
Closure	<ol style="list-style-type: none"> 1. Summarize the lesson and ensure students understand the learning outcomes. 2. Affirm students' understanding by asking them to recall key points. 3. Suggested Reading <ul style="list-style-type: none"> - "Data Science and Big Data Analytics" by EMC Education Services (Editor). 4. Homework <ul style="list-style-type: none"> - Analyze a given dataset using ANOVA and write a report on the findings, including interpretation of results and conclusions.
Evaluation	<ol style="list-style-type: none"> 1. How would you define ANOVA in your own words? 2. What are the basic steps involved in performing ANOVA? 3. Why is ANOVA essential in comparing multiple group means?

Lesson Plan No. 1	Course Name: Introduction to Data Science Topic: Overview of Clustering	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> At the end of the lesson, the student shall be able to: Understand what clustering is and its importance in data science. Differentiate between supervised and unsupervised learning. Recognize real-world scenarios where clustering is applied. Describe the basic concepts behind clustering algorithms.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Projector. Examples from research papers or industry use cases.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Discuss general data analysis goals. Introduce clustering as a method for grouping data. Development (30 minutes) <ul style="list-style-type: none"> Overview of clustering methods: hierarchical, partitioning, density-based, and model-based. Advantages and challenges of clustering techniques. Examples of clustering applications in marketing, biology, and image segmentation. Discussion Points: <ul style="list-style-type: none"> Overview of clustering in the context of unsupervised learning. Importance of clustering in discovering hidden patterns. Course Outcome: <ul style="list-style-type: none"> Understanding the basics of clustering and its applications.
Closure	<ol style="list-style-type: none"> Recap types of clustering and application examples.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> What is clustering, and why is it important? Can you identify a real-world application of clustering? What are some challenges faced in clustering? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 2	Course Name: Introduction to Data Science Topic: K-Means: Use Cases	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Identify and understand common use cases for K-means clustering. Recognize the relevance of K-means in real-world applications across various fields, such as customer segmentation, image compression, and anomaly detection. Apply basic K-means clustering to synthetic data to explore its practical implications.
Teaching Aids (if any)	<ol style="list-style-type: none"> Case studies and examples Whiteboard and markers Synthetic dataset for demonstration Presentation slides Laptops or computers with data analysis software (e.g., Python, Jupyter Notebook, or R)
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Define K-means: Start by defining K-means clustering as a method that partitions a dataset into K clusters, where each data point belongs to the cluster with the nearest mean. Explain that K-means is an unsupervised learning algorithm commonly used for clustering. Business Contexts: Highlight its popularity in business for applications such as customer segmentation and market analysis, emphasizing its efficiency and scalability with large datasets. Discuss how K-means helps organizations make data-driven decisions. Development (30 minutes) <ol style="list-style-type: none"> Use Case 1: Customer Segmentation <ul style="list-style-type: none"> Explain customer segmentation in marketing, where K-means is used to group customers with similar purchasing behaviors. Examples: Retailers may use it to identify high-value customers or tailor marketing strategies for different customer groups. Activity: Walk through a simple customer segmentation example with income and spending data. Use Case 2: Image Compression <ul style="list-style-type: none"> Describe how K-means can reduce the number of colors in an image by clustering pixels with similar colors. Examples: Applications in digital image storage and processing to reduce memory usage. Demo: Show a visual example of an image compressed with K-means to illustrate how colors are grouped and reduced. Use Case 3: Anomaly Detection <ul style="list-style-type: none"> Explain how K-means can identify anomalies by spotting data points that do not fit into any cluster. Examples: Fraud detection in banking, network security monitoring for identifying unusual activities.

	<ul style="list-style-type: none"> - Activity: Present a synthetic dataset with normal and anomalous data points and demonstrate how K-means clusters them. <p>c) Class Activity (5 min)</p> <ul style="list-style-type: none"> - Ask students to think of and discuss additional potential use cases in small groups, encouraging them to consider fields like healthcare, e-commerce, and social media.
Closure	<ol style="list-style-type: none"> 1. Summarize the discussion by reiterating the flexibility and widespread applicability of K-means in various domains, especially in scenarios that require grouping similar data points or detecting patterns. 2. Mention K-means limitations (e.g., sensitivity to the number of clusters, centroid initialization), which will be explored in future lessons.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> 1. Ask students to propose specific examples where K-means could be applied, such as in e-commerce or social media analytics. 2. Encourage them to explain why K-means would be a useful approach in those contexts. <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 3	Course Name: Introduction to Data Science Topic: K-Means: Overview of the Method	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the K-means clustering algorithm, including its basic steps and mathematical foundation. Describe how K-means partitions data into clusters and why it minimizes the distance between points within clusters. Perform a basic K-means clustering operation and interpret the results. Recognize potential challenges, such as choosing the initial centroids and defining the optimal number of clusters.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides with visual aids of clustering steps Whiteboard and markers A simple, visualizable dataset (e.g., two-dimensional data points) Access to software (e.g., Python with Jupyter Notebook or RStudio) for hands-on demonstration
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Begin with the Concept of Clustering: Ask students to recall previous discussions on clustering applications. Briefly recap how clustering organizes data into groups based on similarity. Introduce K-means Clustering: Define K-means as an algorithm that seeks to divide data into K distinct, non-overlapping clusters, with each data point belonging to the cluster with the nearest mean, called the centroid. Why K-means?: Explain that K-means is a simple yet powerful algorithm widely used for pattern recognition, data analysis, and machine learning due to its computational efficiency. Development (30 minutes) <ol style="list-style-type: none"> Step 1: Choosing K <ul style="list-style-type: none"> Explain that K represents the number of clusters to form and that choosing K can significantly impact results. Mention that selecting K will be explored in more detail in a future lesson. Step 2: Initializing Centroids <ul style="list-style-type: none"> Describe how K-means randomly selects initial centroids from the dataset. Discuss the importance of centroid initialization and the potential for varied results with different starting points. Step 3: Assigning Data Points to Nearest Centroid <ul style="list-style-type: none"> Explain how each data point is assigned to the nearest centroid, forming preliminary clusters. Use a visual example to demonstrate the initial cluster formation Step 4: Updating Centroids <ul style="list-style-type: none"> Explain how each data point is assigned to the nearest centroid, forming preliminary clusters. Use a visual example to demonstrate the initial cluster formation Step 5: Repeating Steps 3 and 4 Until Convergence

	<ul style="list-style-type: none"> - Explain that the algorithm repeats the assignment and updating steps until the centroids stabilize (i.e., there is minimal change between iterations). - Note that convergence indicates the clusters are now final. <p>c) Class Activity (5 min)</p> <ul style="list-style-type: none"> - Briefly discuss the concept of distance minimization (e.g., Euclidean distance) and how K-means aims to minimize the sum of squared distances within clusters.
Closure	<ol style="list-style-type: none"> 1. Recap the five main steps of the K-means algorithm, reinforcing the iterative process and convergence. 2. Emphasize the balance between simplicity and effectiveness that makes K-means a popular choice for clustering tasks.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> 1. Ask students to explain the purpose of each step in the K-means process. 2. Prompt them to consider how initial centroid selection might affect clustering results. 3. Challenge students to brainstorm ways to determine the best K value. 4. Assign students a small dataset to perform K-means clustering independently, encouraging them to experiment with different K values and initializations. <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 4	Course Name: Introduction to Data Science Topic: Diagnostics in Clustering	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the concept of clustering diagnostics. Evaluate the quality of clusters using metrics like Dunn Index and Davies-Bouldin Index. Compare the performance of clustering algorithms using diagnostic tools.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python libraries for clustering evaluation (e.g., scikit-learn). Example datasets for hands-on practice.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How can we measure the quality of clustering? Why is it important to evaluate clusters after performing clustering? Introduce K-means Clustering: <ul style="list-style-type: none"> Introduction to clustering diagnostics. Importance of evaluating the strength and validity of clusters. Why K-means? <ul style="list-style-type: none"> Explain that K-means is a simple yet powerful algorithm widely used for pattern recognition, data analysis, and machine learning due to its computational efficiency. Development (30 minutes) <ol style="list-style-type: none"> K- Metrics <ul style="list-style-type: none"> Dunn Index: Ratio of the smallest inter-cluster distance to the largest intra-cluster distance. Davies-Bouldin Index: Measures similarity between clusters. Lower values indicate better clustering. Practical Demonstration <ul style="list-style-type: none"> Python implementation of clustering diagnostics using real-world data. Comparison of Algorithm <ul style="list-style-type: none"> Evaluate K-means vs. DBSCAN on a dataset. Discuss the strengths and weaknesses of each. Step 5: Repeating Steps 3 and 4 Until Convergence <ul style="list-style-type: none"> Explain that the algorithm repeats the assignment and updating steps until the centroids stabilize (i.e., there is minimal change between iterations). Note that convergence indicates the clusters are now final. Class Activity (5 min) <ul style="list-style-type: none"> Briefly discuss the concept of distance minimization (e.g., Euclidean distance) and how K-means aims to minimize the sum of squared distances within clusters.
Closure	<ol style="list-style-type: none"> Recap the importance of clustering diagnostics. Encourage students to explore other evaluation metrics.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss.

- | | |
|--|--|
| | <ol style="list-style-type: none">1. What are the Dunn Index and Davies-Bouldin Index?2. How do clustering diagnostics help in evaluating algorithms?3. Can you think of any other ways to test the quality of clusters? <p>Spend 5 minutes to wrap up and consolidate the learnings</p> |
|--|--|

Lesson Plan No. 5	Course Name: Introduction to Data Science Topic: Apriori Algorithm	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the working of the Apriori algorithm in generating association rules. Apply the concept of frequent itemsets to derive rules. Identify scenarios where the Apriori algorithm is useful.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python implementation of the Apriori algorithm. Example datasets for hands-on practice.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How do you think businesses discover relationships between products? Can you identify patterns between commonly purchased products? Discussion Points <ul style="list-style-type: none"> Overview of the Apriori algorithm and its role in association rule mining. Development (30 minutes) <ol style="list-style-type: none"> Algorithm Steps <ul style="list-style-type: none"> Generating frequent itemsets using support threshold. Deriving association rules from frequent itemsets. Evaluating rules based on confidence and lift. Hands on example <ul style="list-style-type: none"> Use a small dataset (e.g., a grocery store's transactions). Demonstrate finding frequent itemsets and generating rules using Python. Applications <ul style="list-style-type: none"> Discuss real-world use cases (e.g., market basket analysis, customer segmentation).
Closure	<ol style="list-style-type: none"> Recap the steps and key metrics of the Apriori algorithm. Encourage students to try implementing the algorithm on larger datasets.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> What is the purpose of the Apriori algorithm? How are frequent itemsets and association rules derived? Can you think of a business problem where the Apriori algorithm could be applied? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>



Lesson Plan No. 6	Course Name: Introduction to Data Science Topic: Apriori Algorithm	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the working of the Apriori algorithm in generating association rules. Apply the concept of frequent itemsets to derive rules. Identify scenarios where the Apriori algorithm is useful.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python implementation of the Apriori algorithm. Example datasets for hands-on practice.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How do you think businesses discover relationships between products? Can you identify patterns between commonly purchased products? Discussion Points <ul style="list-style-type: none"> Overview of the Apriori algorithm and its role in association rule mining. Development (30 minutes) <ol style="list-style-type: none"> Algorithm Steps <ul style="list-style-type: none"> Generating frequent itemsets using support threshold. Deriving association rules from frequent itemsets. Evaluating rules based on confidence and lift. Hands on example <ul style="list-style-type: none"> Use a small dataset (e.g., a grocery store's transactions). Demonstrate finding frequent itemsets and generating rules using Python. Applications <ul style="list-style-type: none"> Discuss real-world use cases (e.g., market basket analysis, customer segmentation).
Closure	<ol style="list-style-type: none"> Recap the steps and key metrics of the Apriori algorithm. Encourage students to try implementing the algorithm on larger datasets.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> What is the purpose of the Apriori algorithm? How are frequent itemsets and association rules derived? Can you think of a business problem where the Apriori algorithm could be applied? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 7	Course Name: Introduction to Data Science Topic: Evaluation of Candidate Rules	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the metrics for evaluating candidate rules (support, confidence, lift). Analyze the strength and validity of rules using these metrics. Apply evaluation techniques to select the best rules for decision-making.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Examples of rule evaluation metrics. Python libraries for rule evaluation.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> What makes one rule better than another? Why is it important to evaluate association rules? Discussion Points <ul style="list-style-type: none"> Importance of metrics in identifying strong and useful rules. Development (30 minutes) <ol style="list-style-type: none"> Key Metrics <ul style="list-style-type: none"> Support: Minimum occurrence threshold. Confidence: Measures reliability of the rule. Lift: Identifies non-random relationships. Hands on example <ul style="list-style-type: none"> Evaluate rules generated by the Apriori algorithm. Use Python libraries like mlxtend to calculate metrics. Real Life Examples <ul style="list-style-type: none"> Show rules with varying levels of support, confidence, and lift, and discuss their relevance.
Closure	<ol style="list-style-type: none"> Summarize the metrics and their role in rule evaluation. Highlight the importance of selecting actionable rules for decision-making.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> What are the key metrics used to evaluate candidate rules? Why is lift an important metric in association rule mining? Can you describe a scenario where a rule might have high confidence but low lift? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 8	Course Name: Introduction to Data Science Topic: Applications of Association Rules	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Recognize various applications of association rules in business and research. Understand how association rules are used in retail, healthcare, and e-commerce. Analyze examples of successful implementations of association rules.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Case studies from industries. Group activity: Brainstorming applications in real-world scenarios.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> Can you think of situations where identifying associations might be helpful? How do you think Amazon or Netflix uses association rules? Discussion Points <ul style="list-style-type: none"> Overview of the versatility of association rules. Development (30 minutes) <ol style="list-style-type: none"> Retail and E-commerce <ul style="list-style-type: none"> Example: Market basket analysis for cross-selling products. Discuss how rules improve recommendation systems. Healthcare <ul style="list-style-type: none"> Identifying relationships between symptoms and diseases. Finance <ul style="list-style-type: none"> Fraud detection based on transaction patterns. Class Activity <ul style="list-style-type: none"> Divide students into groups to brainstorm other use cases
Closure	<ol style="list-style-type: none"> Summarize key applications across different domains. Highlight the impact of association rules on decision-making.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> How are association rules used in retail and healthcare? Can you identify an area where association rules might not work well? Why? Which domain, in your opinion, benefits most from association rule mining? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 9	Course Name: Introduction to Data Science Topic: Examples of Association Rules	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Analyze real-world examples of association rules. Interpret and validate the derived association rules. Relate examples to practical decision-making scenarios.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Example datasets. Python or R implementations for generating rules.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> Have you observed any patterns in the way items are grouped in stores or websites? What factors might influence these patterns? Discussion Points <ul style="list-style-type: none"> Introduction to real-world examples of association rules. Development (30 minutes) <ol style="list-style-type: none"> Retail <ul style="list-style-type: none"> Rule: “If a customer buys bread, they are likely to buy butter.” Dataset: Analyze a supermarket dataset to demonstrate this relationship. E-Commerce <ul style="list-style-type: none"> Rule: “If a user watches a sci-fi movie, they are likely to watch action movies.” Example: Netflix or Amazon recommendation systems. Healthcare <ul style="list-style-type: none"> Rule: “If a patient has symptom A and symptom B, they are likely to have condition C.” Dataset: Discuss medical datasets for patterns in diagnosis. Hands on demonstration <ul style="list-style-type: none"> Use Python to generate association rules on a small dataset. Discuss support, confidence, and lift for the rules derived. Highlight how these rules influence marketing strategies, recommendations, and diagnostics.
Closure	<ol style="list-style-type: none"> Recap the examples and their interpretations. Discuss how understanding association rules enhances data-driven decision-making.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> Can you describe an association rule and its business impact? What metrics are used to validate the strength of an example rule? How might the rules vary based on different datasets or industries? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 10	Course Name: Introduction to Data Science Topic: Validation and Testing of Association Rules	Course No.: COM-501
---------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the importance of validating association rules. Learn techniques for testing the reliability of rules. Apply validation techniques to ensure robust decision-making.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Validation techniques and sample datasets. Statistical tools like cross-validation and holdout methods.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> Why do you think it's important to test association rules? What could go wrong if rules are applied without validation? Discussion Points <ul style="list-style-type: none"> Overview of the need for validation and testing in association rule mining. Development (30 minutes) <ol style="list-style-type: none"> Validation Techniques <ul style="list-style-type: none"> Cross-validation: Splitting the dataset into training and testing sets. Statistical significance tests for rules. Test for overfitting <ul style="list-style-type: none"> Ensure the rules generalize well across datasets. Hands on example <ul style="list-style-type: none"> Split a dataset, derive rules from the training set, and test them on the holdout set. Discuss the results and how validation improves reliability. Industry use cases <ul style="list-style-type: none"> Discuss examples of where rule validation prevents costly decisions.
Closure	<ol style="list-style-type: none"> Summarize the importance of validation and testing for association rules. Encourage students to explore different datasets and validation techniques.
Evaluation	<p>Reflective Questions (What Why Who?). Allow students to answer and discuss.</p> <ol style="list-style-type: none"> Why is it important to validate association rules? What techniques can be used to test the reliability of rules? Can you identify an example where rule testing might prevent errors in decision-making? <p>Spend 5 minutes to wrap up and consolidate the learnings</p>

Lesson Plan No. 1	Course Name: Introduction to Data Science Topic: Linear Regression: Concepts and Use Cases	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the basic concept and assumptions of linear regression. Apply linear regression to predict numerical outcomes. Explore real-world use cases, such as sales prediction and trend analysis.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python or Excel for demonstration. Sample dataset (e.g., housing prices).
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> What is the relationship between advertising expenditure and sales? Can you think of situations where numerical predictions are needed? Discussion Points <ul style="list-style-type: none"> Define linear regression and its importance in predicting continuous variables. Development (30 minutes) <ol style="list-style-type: none"> Concepts of Linear Regression <ul style="list-style-type: none"> Line of best fit, assumptions, and formula. Key terms: coefficients, residuals, R^2 score. Hands on Example <ul style="list-style-type: none"> Use a dataset to predict housing prices based on features like square footage and location. Use Cases <ul style="list-style-type: none"> Sales forecasting, financial analysis, and environmental predictions.
Closure	<ol style="list-style-type: none"> Recap key concepts and use cases. Highlight how linear regression fits into advanced analytics..
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> What is the main goal of linear regression? How can linear regression be applied in retail and finance? Why is it important to check for assumptions in linear regression? Spend 5 minutes to wrap up and consolidate the learnings

Lesson Plan No. 2	Course Name: Introduction to Data Science Topic: Logistic Regression: Concepts and Use Cases	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the working and assumptions of logistic regression. Apply logistic regression for classification problems. Explore use cases such as customer churn prediction and medical diagnosis.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python demonstration using sklearn. Dataset for classification (e.g., heart disease dataset).
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How do we predict whether a customer will purchase a product? What might be a simple way to classify outcomes into "yes" or "no"? Discussion Points <ul style="list-style-type: none"> Overview of logistic regression for binary classification. Development (30 minutes) <ol style="list-style-type: none"> Concepts of Logistic Regression <ul style="list-style-type: none"> Logistic function, odds ratio, and decision boundary. Key metrics: accuracy, precision, recall, AUC-ROC. Hands on Example <ul style="list-style-type: none"> Predict whether a patient has a disease based on symptoms. Use Cases <ul style="list-style-type: none"> Fraud detection, customer segmentation, credit scoring.
Closure	<ol style="list-style-type: none"> Recap of logistic regression and its practical applications. Link logistic regression with other classification algorithms.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> What is the key difference between linear and logistic regression? Name one practical application of logistic regression. Why is logistic regression suitable for binary outcomes? Spend 5 minutes to wrap up and consolidate the learnings



Kot Bhalwal, Jammu



Lesson Plan No. 3	Course Name: Introduction to Data Science Topic: Decision Trees: Concepts and Use Cases	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: a. Understand the structure and working of decision trees. b. Apply decision trees to solve classification problems. c. Explore practical applications in marketing and healthcare.
Teaching Aids (if any)	a. Visual diagrams of decision trees. b. Python demonstration (sklearn). c. Sample dataset (e.g., student performance).
Teaching Development	<ol style="list-style-type: none">Introduction (10 minutes)<ul style="list-style-type: none">Ask Questions:<ul style="list-style-type: none">- How would you classify students into high or low performance based on attendance?- Can you think of scenarios where splitting decisions helps?Discussion Points<ul style="list-style-type: none">- Overview of decision trees for classification and regression.Development (30 minutes)<ol style="list-style-type: none">Concepts of Decision Tree<ul style="list-style-type: none">- Root, branches, leaves, entropy, and Gini index.Hands on Example<ul style="list-style-type: none">- Predict whether a customer will buy based on income and age.Use Cases<ul style="list-style-type: none">- Credit risk assessment, patient diagnosis, and targeted marketing.
Closure	<ol style="list-style-type: none">Recap of decision tree concepts and use cases.Discuss advantages and limitations of decision trees.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none">What are the key components of a decision tree?Name one example where decision trees can be used in business.How does a decision tree decide on a split? Spend 5 minutes to wrap up and consolidate the learnings

Lesson Plan No. 4	Course Name: Introduction to Data Science Topic: Bayes' Theorem and Naïve Bayes Classifier	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand Bayes' Theorem and its relevance to classification. Explore the working of the Naïve Bayes algorithm. Apply Naïve Bayes for text classification tasks.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Example calculations of Bayes' Theorem. Python demonstration (email spam detection).
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How do email services detect spam? Can probability be used to classify outcomes? Discussion Points <ul style="list-style-type: none"> Introduction to Bayes' Theorem and Naïve Bayes. Development (30 minutes) <ol style="list-style-type: none"> Bayes' Theorem <ul style="list-style-type: none"> Formula and its components. Application in simple probability problems. Naïve Bayes Classifier <ul style="list-style-type: none"> Assumptions of independence. Steps in applying Naïve Bayes. Use Cases <ul style="list-style-type: none"> Classify emails as spam or not using a text dataset.
Closure	<ol style="list-style-type: none"> Recap of Bayes' Theorem and Naïve Bayes. Link Naïve Bayes with real-world text analytics tasks.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> What is the key assumption of Naïve Bayes? Name one application of Naïve Bayes in text classification. How does Bayes' Theorem relate to classification tasks? Spend 5 minutes to wrap up and consolidate the learnings

Lesson Plan No. 5	Course Name: Introduction to Data Science Topic: Smoothing Techniques in Naïve Bayes	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand the need for smoothing in Naïve Bayes classification. Learn about Laplace Smoothing and its application. Apply smoothing techniques to handle unseen data in classification problems.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python demonstration
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> What happens when a category in a dataset has no occurrences during training? Why might this lead to classification errors? Discussion Points <ul style="list-style-type: none"> Importance of smoothing in handling zero-frequency problems. Development (30 minutes) <ol style="list-style-type: none"> Concept of Smoothing <ul style="list-style-type: none"> Define Laplace Smoothing and explain its formula. Illustrate with a small dataset example. Hands on Example <ul style="list-style-type: none"> Use a text classification dataset with rare categories. Demonstrate the difference with and without smoothing. Real World Application <ul style="list-style-type: none"> Spam detection, language modeling in NLP.
Closure	<ol style="list-style-type: none"> Recap the importance of smoothing in Naïve Bayes. Encourage students to explore other smoothing techniques.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> Why is smoothing important in Naïve Bayes? What does Laplace Smoothing address? Can you think of other scenarios where smoothing might be useful? Spend 5 minutes to wrap up and consolidate the learnings

Lesson Plan No. 6	Course Name: Introduction to Data Science Topic: Diagnostics of Naïve Bayes Classifier	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand evaluation metrics for Naïve Bayes classifiers. Perform diagnostics to evaluate model performance. Interpret classification metrics such as accuracy, precision, recall, and F1-score.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python demonstration with sklearn. Confusion matrix visualization tool.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> How do we know if a classifier performs well? What metrics might indicate that the model is making correct predictions? Discussion Points <ul style="list-style-type: none"> Overview of evaluation metrics for classification models. Development (30 minutes) <ol style="list-style-type: none"> Evaluation Metrics <ul style="list-style-type: none"> Accuracy: When to use and its limitations. Precision, Recall, and F1-Score: Define and explain scenarios for use. Confusion Matrix Analysis <ul style="list-style-type: none"> Demonstrate confusion matrix components (TP, FP, TN, FN). Hands on Example <ul style="list-style-type: none"> Evaluate a Naïve Bayes model on a text classification dataset. Generate metrics using sklearn's classification report.
Closure	<ol style="list-style-type: none"> Recap the metrics and their significance in diagnostics. Link evaluation with model improvement strategies.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> What does a confusion matrix tell us about model performance? Why might recall be more important than precision in some cases? How can we use diagnostics to improve a Naïve Bayes classifier? Spend 5 minutes to wrap up and consolidate the learnings

Lesson Plan No. 7	Course Name: Introduction to Data Science Topic: Diagnostics of Decision Trees	Course No.: COM-501
--------------------------	---	----------------------------

Objectives	At the end of the lesson the student shall be able to: <ol style="list-style-type: none"> Understand evaluation methods for decision trees. Interpret overfitting, underfitting, and tree pruning techniques. Evaluate decision trees using accuracy and cross-validation.
Teaching Aids (if any)	<ol style="list-style-type: none"> Presentation slides. Python demonstration using sklearn's DecisionTreeClassifier.
Teaching Development	<ol style="list-style-type: none"> Introduction (10 minutes) <ul style="list-style-type: none"> Ask Questions: <ul style="list-style-type: none"> What happens when a decision tree is too complex? How can we improve a decision tree's performance? Discussion Points <ul style="list-style-type: none"> Importance of diagnostics in decision tree models. Development (30 minutes) <ol style="list-style-type: none"> Overfitting and Under fitting <ul style="list-style-type: none"> Explain concepts with visual examples of decision boundaries. Pruning Techniques <ul style="list-style-type: none"> Introduce pre-pruning and post-pruning methods. Hands on Example <ul style="list-style-type: none"> Train a decision tree and evaluate using cross-validation. Demonstrate pruning to improve performance.
Closure	<ol style="list-style-type: none"> Recap the importance of diagnostics in decision trees. Discuss practical use cases where diagnostics prevent overfitting.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none"> Why is cross-validation important for decision tree evaluation? What is the purpose of tree pruning? How can overfitting be identified in a decision tree? Spend 5 minutes to wrap up and consolidate the learnings



Kot Bhalwal, Jammu



Lesson Plan No. 8	Course Name: Introduction to Data Science Topic: Applications of Linear and Logistic Regression	Course No.: COM-501
--------------------------	--	----------------------------

Objectives	At the end of the lesson the student shall be able to: a. Identify real-world problems solved by linear and logistic regression. b. Analyze datasets to apply appropriate regression techniques. c. Compare the use of linear and logistic regression in different scenarios.
Teaching Aids (if any)	a. Presentation slides. b. Python demonstration with datasets from healthcare and business.
Teaching Development	<ol style="list-style-type: none">Introduction (10 minutes)<ul style="list-style-type: none">Ask Questions:<ul style="list-style-type: none">- Can regression help predict customer behavior?- What types of problems suit linear or logistic regression?Discussion Points<ul style="list-style-type: none">- Overview of linear and logistic regression applications.Development (30 minutes)<ol style="list-style-type: none">Linear Regression Use Cases<ul style="list-style-type: none">- Predict housing prices, sales trends, and stock market values.Logistic Regression Use Cases<ul style="list-style-type: none">- Predict disease outcomes, customer churn, and fraud detection.Hands on Example<ul style="list-style-type: none">- Apply both regression methods to a multi-feature dataset.
Closure	<ol style="list-style-type: none">Recap the key differences and applications of the two methods.Encourage students to explore industry datasets for practice.
Evaluation	Reflective Questions (What Why Who?). Allow students to answer and discuss. <ol style="list-style-type: none">What kind of problems require logistic regression instead of linear regression?Name one example each for linear and logistic regression applications.How can understanding these techniques enhance business decision-making? Spend 5 minutes to wrap up and consolidate the learnings