Syllabus

Honours Programme in

Data Science and Analysis

(Offered by Department of Computer Engineering)

From

Academic Year 2024-25

(Revision-2)

Approved by FOET ------ and ------------

K J Somaiya College of Engineering, Mumbai-77

(A Constituent College of Somaiya Vidyavihar University)

Honours' Degree Programme in Data Science and Analysis

Offered by Department of Computer Engineering

Introduction:

In today's data-driven society, Data Science provides a foundation for problem solving that impacts many areas of the economy, including science, engineering, medicine, banking, finance, sports and the arts. Data science is an interdisciplinary field that focuses on analysing large amounts of data in various formats such as text, audio video etc., to extract information from inherent patterns, design and develop underlying models, and use data for real world applications for prediction, detection and decision.

Data processing and analytics converts raw data of different format into meaningful format which can be analysed and interpreted for a variety of purposes. Data Science focuses on data extraction and processing techniques and machine learning algorithms for inferring and representation of data for meaningful presentation for easy interpretation and decision making. Data Science focuses on innovative and intelligent ways of handling data in structured and unstructured format which may be available in large volume, and analyzing data for various purposes.

Data science and analytics programme is designed to prepare the students in in interdisciplinary domains for gaining hands-on experience on data analytics, machine learning and data visualization methods as it relates to multiple fields of interest. It is designed to empower the students to employ problem ideation, computational thinking and data science tools to solve practical business problems. The coursework consists of courses that cover the spectrum of Data Science to equip the students with knowledge of data analysis, Machine Learning and data visualization techniques for data-centric computation to address problems of real world.

Objectives:

- Applications of principles of Data Science for analysis of diverse domain problems.
- Use software tools and machine learning algorithms from the areas of statistics, mathematics, Computer Science and Artificial intelligence to model and analyze real-world data, communicate data insights, and effectively present results using data visualization techniques.
- Deployment of latest tools and technologies to analyse large amounts variety of Data.
- Understand ethical practices that are importantly and inevitably tied to data-driven prediction and decision-making.

Learning Outcomes of the Honours' Degree Programme:

At the successful completion of this programme an engineering graduates will be able to

- Understand foundational principles, key concepts, and ethical considerations in data science for recognizing its significance across various interdisciplinary domains.
- Perform data acquisition, cleaning, exploration, and statistical analysis for gaining proficiency in Data Science techniques.
- Apply Machine Learning algorithms to model, develop, and evaluate for solving real-world problems.
- Analyze data to create impactful visualizations for effective communication after assessing the data for recognising visual representation quality.

Assessment Methods: Evaluation is done by a variety of tools including Open book tests, MCQs (multiple choice questions), Study of research papers, Internal Assessment tools and End Semester Examinations etc. Mini-Projects are offered in courses also to encourage project based learning among students.

Acronyms used in syllabus document	
Acronym	**Definition**
CA	Continuous Assessment
ESE	End Semester Exam
IA	Internal Assessment
O	Oral
P	Practical
P&O	Practical and Oral
TH	Theory
TUT	Tutorial
TW	Term work

Acronyms used in syllabus document	
ISE	In-semester Examination
CO	Course Outcome

Acronyms used in Course code e.g. 116HxxC301

Position of Digit	Acronym	Definition
1	**2**	SUV 2023 Second Revision
2	**16**	KJSCE
3	**H**	**Honour Degree Program**
4	**03(xx)**	**Data Science and Analytics**
5	**C**	Core Course
	L	Laboratory Course
	T	Tutorial
	P	Project Based Course
6	**1/2/3/4**	Semester Number
7	**01/02/03--**	Course Number

Proposed Credit Scheme

Course Code	Course Name	Teaching Scheme (Hrs.) TH – P – TUT	Total (Hrs.)	Credits Assigned TH – P – TUT	Total Credits	Suggested semester of Honours' degree
216H03C401	Introduction to Data Science	3 – 0 – 0	03	3 – 0 – 0	03	IV
216H03L401	Introduction to Data Science Laboratory	0 – 2 – 0	02	0 – 1 – 0	01	IV
216H03C501	Data Analysis	3 – 0 – 0	03	3 – 0 – 0	03	V
216H03L501	Data Analysis Laboratory	0 – 0 – 2	02	0 – 0 – 1	01	V
216H03C601	Machine Learning	3 – 0 – 0	03	3 – 0 – 0	03	VI
216H03L601	Machine Learning Laboratory	0 – 2 – 0	02	0 – 1 – 0	01	VI
216H03C701	Data Visualization	3 – 0 – 0	03	3 – 0 – 0	03	VII
216H03L701	Data Visualization Laboratory	0 – 2 – 0	02	0 – 1– 0	01	VII
216H03L702	Mini Project	0 – 2 – 0	02	0 – 0 – 2	02	VII
	Total	**12 – 8 – 2**	**22**	**12 – 4 – 1**	**18**	

<u>All courses are project/mini-project based</u>**

Proposed Examination Scheme

Course Code	Course Name	Examination Scheme				
		Marks				
		CA		ESE$	Lab/ Tut CA	Total
		ISE	IA			
216H03C401	Introduction to Data Science	30	20	50	--	100
216H03L401	Introduction to Data Science Laboratory	-	-	-	50	50
216H03C501	Data Analysis	30	20	50	--	100
216H03L501	Data Analysis Laboratory	-	-	-	50	50
216H03C601	Machine Learning	30	20	50	-	100
216H03L601	Machine Learning Laboratory	-	-	-	50	50
216H03C701	Data Visualization	30	20	50	-	100
216H03L701	Data Visualization Laboratory	-	-	-	50	50
216H03L702	Mini Project	-	-	-	50	50
	Total	**120**	**80**	**200**	**250**	**650**

Course Code	Name of the Course				
216H03C401	Introduction to Data Science				
Teaching Scheme (Hrs./Week)	TH	P	TUT	Total	
	03	-	-	03	
Credits Assigned	03	-	--	03	

Evaluation Scheme	Marks				
	LAB/TUT CA	CA (TH)		ESE	Total
		IA	ISE		
	-	20	30	50	100

Course prerequisites: Elementary mathematics, Basics of programming in Python.

Course Objectives:

1. Understand and Learn methodologies and significance of Data Science ethical considerations while working with various data types and the 5 V's of Data.

2 Comprehend mathematical foundation in probability, statistics, and basic linear algebra for exploratory data analysis.

3. Learn and implement data cleaning, transformation, and feature extraction and selection techniques to prepare knowledge base for data analysis.

4. Apply supervised and unsupervised machine learning techniques for real-world application for prediction, detection and decision.

5. Create effective and engaging data visualizations using visualization tools for presenting, inferring and for communication of insights.

Course Outcomes (CO):

CO 1	Understand the significance and Data Science processes in real world applications.
CO 2	Gain necessary mathematical in probability, statistics, and basic linear algebra.
CO 3	Learn data cleaning, transformation, and feature engineering techniques.
CO 4	Apply supervised and unsupervised machine learning techniques to solve real-world problem.
CO 5	Create compelling data visualizations for finding inferences and effective communication of application outcomes.

Module No.	Unit No.	Contents	No. of Hrs.	CO
1	**Introduction to Data Science**		08	CO1
	1.1	What is Data Science? Importance and applications of Data Science, Data Science workflow, The 5 V's of Data		
	1.2	Structured, semi-structured, and unstructured data. Challenges and considerations for handling different data types. Ethical Considerations in Data Science.		
2	**Mathematical Foundations for Data Science**		10	CO2
	2.1	Linear Algebra: Vectors and matrices. Matrix operations and transformations.		
	2.2	Probability Distributions: Gaussian distribution, Binomial distribution, Poisson distribution. Properties of Probability Distributions: Mean, variance, and standard deviation of a distribution. Moments and percentiles. Skewness and kurtosis.		
	2.3	Statistical analysis: Measures of central tendency (mean, median, mode). Measures of dispersion (range, variance, standard deviation). Quartiles, percentiles. Stem and leaf plots, Box plots Hypothesis testing		
3	**Exploratory Data Analysis and ETL**		09	CO3
	3.1	Data Cleaning and Transformation: Identifying Missing Data, Handling Missing Data, Handling Outliers and Data Types (categorical, numerical, ordinal), and methods for encoding categorical data, including one-hot encoding and label encoding. Importance of data type selection for modeling.		
	3.2	Introduction to Feature Engineering: Define feature engineering and its role in data pre-processing.		
	3.3	ETL Overview		
4	**Introduction to Machine Learning**		10	CO4
	4.1	Machine Learning Basics: What is machine learning? Types of machine learning: supervised, unsupervised, and reinforcement learning.		
	4.2	Supervised Learning: Linear Regression (simple linear regression, Least Squares Method, Model Evaluation Metrics – MSE & R-squared (R^2) coefficient of determination & Interpretation of R^2. Naïve Bayes Algorithm.		
	4.3	Unsupervised learning – k means clustering.		
		Self-learning: Logistic Regression, R libraries that implement the above algorithms		

5	**Data Visualization**		08	CO5
	5.1	Principles of Effective Data Visualization, Case studies of data visualization		
	5.2	Tools for Creating Visualizations: Matplotlib, Seaborn		
		Self-Learning : ggplot2 (R), Plotly & Tableau.		
		Total	**45**	--

Reference Books

Sr. No	Name/s of Author/s	Title of Book	Publisher	Edition/ Year
1	Jiawei Han, Micheline Kamber, Jian Pei	*Data Mining: Concepts and Techniques*	Morgan Kaufmann	3rd edition, 2012
2	Sheldon M. Ross	*Introductory Statistics*	Academic Press	4th edition, 2017
3	Avrim Blum, John Hopcroft, and Ravindran Kannan	Foundations of Data Science	ONLINE	2014
4	Jake VanderPlas	*Python Data Science Handbook*	O'Reilly	2016
5	Cathy O'Neil and Rachel Schutt	*Doing Data Science, Straight Talk From The Frontline*	O'Reilly	2014, Edition 1

Course Code	Name of the Course			
216H03L401	Introduction to Data Science Laboratory			

Teaching Scheme (Hrs./Week)	TH	P	TUT	Total
	-	02	-	02
Credits Assigned	-	01	--	01

Evaluation Scheme	Marks				
	LAB/TUT CA	CA (TH)		ESE	Total
		IA	ISE		
	50	-	-	-	50

Course prerequisites: Elementary mathematics, Basics of programming in Python.

Course Objectives:

1. Understand the significance, workflow, and ethical considerations of Data Science while recognizing various data types and the 5 V's of Data.
2. Develop a solid foundation in probability, statistics, and basic linear algebra for data analysis.
3. Learn data cleaning, transformation, and feature engineering techniques to prepare data for analysis.
4. Apply supervised and unsupervised machine learning techniques to real-world data for prediction and pattern discovery.
5. Create effective and engaging data visualizations using principles and tools for clear communication of insights.

Course Outcomes (CO):

CO 1	Understand the significance and Data Science processes in real world applications.
CO 2	Gain necessary mathematical in probability, statistics, and basic linear algebra.
CO 3	Learn data cleaning, transformation, and feature engineering techniques.
CO 4	Apply supervised and unsupervised machine learning techniques to solve real-world problem.
CO 5	Create compelling data visualizations for finding inferences and effective communication of application outcomes.

Laboratory experiments will be based on the entire syllabus of the course 216H03C401, 'Introduction to Data Science'. Students will be graded based on continuous assessment during laboratory.

Course Code	Name of the Course				
216H03C501	Data Analysis				
Teaching Scheme (Hrs./Week)	TH	P	TUT		Total
	03	--	--		03
Credits Assigned	03	--	--		03
Evaluation Scheme	Marks				
	LAB/TUT CA	CA (TH)		ESE	Total
		IA	ISE		
	--	20	30	50	100

Course pre-requisites:

- Concepts of DBMS, Probability and statistics, knowledge of programming language (C/C++/Java/ Python).

Course Objectives:

Introduction to the fundamental concepts of Data Analytics , analyse real world case studies by applying mining algorithms and visualization for decision-making in Geospatial, social media ,healthcare and text mining business applications

Course Outcomes (CO):

At the end of successful completion of the course the student will be able to

CO1	Understand basic concepts of data analytics to solve real world case studies
CO2	Apply the data analysis on geospatial system
CO3	Perform the graph data analysis
CO4	Perform Time series Analysis
CO5	Apply the text data analytics in the field of Health care

Module No.	Unit No.	Details	Hrs.	CO
1	**Introduction to Data Analysis**		**05**	**CO1**
	1.1	Introduction to Data Analytics, Different types of data analytics: Descriptive analytics, Diagnostics Analytics, Predictive Analysis, Prescriptive Analysis		
		# Self-Learning: LinkedIn Analysis, Netflix Analysis, Cricket and FIFA Analysis.		
2	**Spatial Data Analysis**		**10**	**CO2 CO1**
	2.1	Contents and characteristics of Spatial data, Spatial data formats, spatial data bases		
	2.2	Introduction, Definition of GIS, Evolution of GIS , components of GIS Spatial Association rule mining, spatial hierarchal clustering, Set based decision and knowledge recovery		
	2.3	Case study: GIS application for Spatial data mining using open source software tools		
		# Self-learning: QGIS, Hadoop, GeoSpark R PostgreSQL, PostGIS, Python		
3	**Graph Analysis**		**10**	**CO3**
	3.1	Introduction to the Social Network, Clustering of Social-Network Graphs, Direct Discovery of Communities		
	3.2	Partitioning of Graphs, Finding Overlapping Communities, Simrank, Counting Triangles, Neighborhood Properties of Graphs		
	3.3	# Self-learning: GraphX tools of Apache.		
4	**Time series Analysis for prediction and forecasting**		10	**CO4**
	4.1	Introduction, Finding and Wrangling Time Series Data, Exploratory Data Analysis for Time Series, Simulating Time Series Data, Storing Temporal Data,		
	4.2	Statistical Models for Time Series, State Space Models for Time Series, forecasting methods, Testing for randomness, Regression based trend model :AR,MA,ARIMA, random walk model, moving average forecast, exponential smoothing forecast, seasonal models,		
5	**Data Analysis in Health Care Application**		**10**	**CO5**
	5.1	Introduction, Components of HER, Benefits of EHR-		
	5.2	Natural Language Processing and Analysis Of Clinical Text : Introduction , report analyser, text analyser, Core NLP Components Morphological Analysis , Lexical Analysis , Syntactic Analysis ,Semantic Analysis , Data Encoding .		
	5.3	Mining Information from Clinical Text: Rule-Based Approaches. Pattern-Based Algorithms ,Machine Learning Algorithms		

Module No.	Unit No.	Details	Hrs.	CO
		Self-learning: Introduction to Social media Analysis for healthcare,		
		Total	**45**	

 # Students should prepare all Self Learning topics on their own. Self-learning topics will enable students to gain extended knowledge of the topic. Assessment of these topics may be included in IA and Laboratory Experiments.

Reference Books:

Sr. No.	Name/s of Author/s	Title of Book	Name of Publisher with country	Edition and Year of Publication
1.	Michael J. de Smith, Michael F. Goodchild and Paul A. Longley	Geospatial Analysis: A Comprehensive Guide to Principles, Techniques, and Software Tools,	Wiley, Second Edition	2019
2.	Anil Maheshwari	*Data Analysis*	Mc Graw Hill	2017
3.	James, G., Witten, D., Hastie, T., Tibshirani, R.	*An introductionto statistical learning with applications in R*	Springer	2013
4.	Chandan K. Reddy and Charu C Aggarwal	*Healthcare data Analysis*	Taylor & Francis	2015
5	U. Dinesh Kumar	*Business Analysis*	Wiley	2017
6	Li, Deren., Wang, Shuliang., Li,Deyi	Spatial Data Mining: Theory and Application.	Spatial Data Mining: Theory and Application.	2016
7.	Albright and Winston	Bussiness Analysis	Cengage Publication	5th edition, 2015
8.	Aileen Nielsen	Practical Time Series Analysis	O'Reilly Media, Inc.	1st edition October 2019

*In addition to printed books, faculty can suggest (authentic) urls or e-books, e-contents etc.

Course Code	Name of the Course			
216H03L501	Data Analysis Laboratory			
Teaching Scheme (Hrs./Week)	TH	P	TUT	Total
	-	02	-	02
Credits Assigned	-	01	--	01

Evaluation Scheme	Marks				
	LAB/TUT CA	CA (TH)		ESE	Total
		IA	ISE		
	50	–	–	–	50

Course Objectives

- Comprehend methodologies of data analysis

Course Outcomes

At the end of the course students will be able to

CO 1	Apply and Implement spatial data analysis
CO 2	Apply and Implement graph data analysis
CO 3	Apply and Implement time seriesl data analysis
CO 4	Apply and Implement data analysis for inferring the results

Term-Work:

Term work will consist of minimum 8 experiments/ tutorials covering entire syllabus of the course ‘Data Analysis Laboratory. Students will be graded based on continuous assessment of their term work.

Course Code	Name of the Course			
216H03C601	Machine Learning			
Teaching Scheme (Hrs./Week)	TH	P	TUT	Total
	03	-	-	03
Credits Assigned	03	-	--	03

Evaluation Scheme	Marks				
	LAB/TUT CA	CA (TH)		ESE	Total
		IA	ISE		
	-	20	30	50	100

Course prerequisites: Basic data science concepts, Basics of programming in Python.

Course Objectives:

1. To establish a strong foundation in machine learning fundamentals.
2. To enable students to apply supervised learning techniques to solve real-world problems.
3. To introduce students to unsupervised learning methods for data exploration and pattern discovery.
4. To provide a foundational understanding of deep learning and artificial neural networks.
5. To introduce students to the concepts and principles of reinforcement learning.

Course Outcomes (CO):

After completing the course, student will be able to

CO 1	Recall and understand fundamental machine learning concepts,
CO 2	Apply supervised learning techniques to analyze and evaluate real-world applications.
CO 3	Employ unsupervised learning methods to explore, categorize data and differentiate patterns within datasets.
CO 4	Understand & describe mathematical foundation behind deep learning architectures.
CO 5	Gain a basic understanding of Reinforcement Learning (RL) principles and its potential applications.

Module No.	Unit No.	Contents	No. of Hrs.	CO
1	**Machine learning foundation:**		06	CO1
		What is Machine Learning? Types of learning, applications, Bias, variance, overfitting, under-fitting, cross validation and feature engineering.		
2	**Supervised Learning**		10	CO2
		Non Linear Regression, Multivariable Linear Regression, gradient descent learning algorithm and its variations & Ridge Regression and Lasso Regression (Regularization). Logistic regression for binary classification, Maximum Likelihood Estimation (MLE), L1 and L2 regularization for logistic regression & hyper-parameter tuning		
		Decision Trees: Structure of decision trees, Tree building algorithms (e.g., ID3, CART, C4.5), Ensemble learning, bagging and Random Forests.		
		Support Vector Machines: SVM for classification and regression, Margin and decision boundary, Soft margin vs. hard margin SVM, Kernel functions (e.g., linear, polynomial, radial basis function) & SVM for non-linear data separation		
3	**Unsupervised Learning**		10	CO3
	3.1	Dimension reduction using Principle Component Analysis		
	3.2	Hierarchical clustering, Agglomerative vs. divisive clustering & Density-based clustering (DBSCAN)		
	3.3	Introduction to Gaussian Mixture Models, Expectation-Maximization (EM) algorithm for clustering.		
	3.4	Performance Metrics (Accuracy, Precision, Recall, F1-score)		
4	**Introduction to Deep learning**		10	CO4
	4.1	Introduction to Artificial Neural Networks (ANN), Feed-forward and Back-propagation.		
	4.2	Activation Functions (ReLU, Sigmoid, Tanh & Softmax)		
	4.3	Introduction to Deep Learning Architectures: CNN		
5	**Reinforcement learning and applications**		09	CO5
	5.1	What is Reinforcement Learning? Key components: Agent, Environment, Reward, Policy, Value Function		
	5.2	Q-Learning		
		Total	45	--

Reference Books

Sr. No	Name/s of Author/s	Title of Book	Publisher	Edition/ Year
1	Tom M.Mitchell	Machine Learning	McGraw Hill	2017
2	M. Gopal	Applied Machine Learning	McGraw Hill	2018
3	Ian Goodfellow, Yoshua Bengio, Aaron Courville	Deep Learning	An MIT Press book	2016
4	Aurélien Géron	Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition	O'Reilly	2019
5	Sebastian Raschka Vahid Mirjalili	Python Machine Learning, Second Edition: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow	Packt	2017

<table>
<tr><th>Course Code</th><th colspan="6">Name of the Course</th></tr>
<tr><td>216H03L601</td><td colspan="6">Machine Learning Laboratory</td></tr>
<tr><td rowspan="2">Teaching Scheme (Hrs./Week)</td><td>TH</td><td colspan="2">P</td><td colspan="2">TUT</td><td>Total</td></tr>
<tr><td>-</td><td colspan="2">02</td><td colspan="2">-</td><td>02</td></tr>
<tr><td>Credits Assigned</td><td>-</td><td colspan="2">01</td><td colspan="2">--</td><td>01</td></tr>
<tr><td rowspan="4">Evaluation Scheme</td><td colspan="6">Marks</td></tr>
<tr><td rowspan="2">LAB/TUT CA</td><td colspan="3">CA (TH)</td><td rowspan="2">ESE</td><td rowspan="2">Total</td></tr>
<tr><td>IA</td><td colspan="2">ISE</td></tr>
<tr><td>50</td><td>-</td><td colspan="2">-</td><td>-</td><td>50</td></tr>
</table>

Laboratory experiments will be based on the entire syllabus of the course 216H03C601, ‘Machine Learning’. Students will be graded based on continuous assessment during laboratory.

Course Code	Name of the Course				
216H03C701	**Data Visualization**				
Teaching Scheme (Hrs./Week)	**TH**	**P**		**TUT**	**Total**
	03	-		-	03
Credits Assigned	03	-		--	03
Evaluation Scheme	**Marks**				
	LAB/TUT CA	**CA (TH)**		**ESE**	**Total**
		IA	**ISE**		
	-	20	30	50	100

***Onscreen examination**

Course prerequisites (if any):

Basics of statistics, database and data analysis

Course Objectives

- Employ best practices in data visualization to develop charts, maps, tables, and other visual representations of data
- Use visualization tools to conduct data analysis, especially exploration of an unfamiliar dataset.
- Create compelling, interactive dashboards to combine several visualizations into a cohesive and functional whole.
- Use data visualizations, dashboards and Stories to support relevant communication for diverse audiences.

Course Outcomes

At the end of successful completion of the course the student will be able to

CO1	Learn how to locate and download datasets, extract insights from that data and present their findings in a variety of different formats
CO2	Detect and understand the stories within datasets and its applications.
CO3	Apply data visualization best practices
CO4	Design static charts, interactive Dashboards and data stories

Module No.	Unit No.	Details	Hrs.	CO
1		**Introduction Data Visualization**	**06**	**CO1 CO2**
	1.1	Introduction to data visualization and its need. Data analysis lifecycle. A Visual Revolution, Various types of visualization with its best practices.From Visualization to Visual Data Storytelling: An Evolution, From Visual to Story: Bridging the Gap		
	1.2	Data Fundamentals, Collecting data, Preparing Data		
	1.3	Gestalt Design principles		
	1.4	Data cleaning using Excel		
2		**Visualization Foundations**	**09**	**CO 3**
	2.1	Induction to Data foundation		
	2.2	Addressing the design of visualizations Human Perception and Information Processing		
	2.3	Geospatial Displays : Connecting to Geographic Data Assigning Geographic Roles Creating Geographic Hierarchies Proportional Symbol Maps, Choropleth Map		
3		**Visualization Techniques**	**10**	**CO3**
	3.1	Visualization Techniques for Spatial Data, implicit or explicit spatial or spatiotemporal attribute, mapping of the data attributes to graphical attributes		
	3.2	Visualization Techniques for Time-Oriented Data, Handling the temporal dimension. time and time-oriented data. TimeBench, visual analytics of time-oriented data.		
	3.3	Visualization Techniques for Multivariate Data		
	3.4	Visualization Techniques for Trees, Graphs, and Networks		
	3.5	Text and Document Visualization		
4		**Exploratory Visualization**	**10**	**CO3**
	4.1	Data Joins Best Practices o Sorting, Top N, bottom N ,Filtering , Maps		
	4.2	Visual Analytics		
	4.3	Optimal visualization types , Binning values, Calculated fields , Table calculations , Level of Detail calculations		
5		**Storytelling and Dashboards**	**10**	**CO4**
	5.1	Storytelling Multivariate displays The Science of Storytelling The Power of Stories Context in action Exploratory versus Explanatory Analysis Structuring Stories Audience Analysis for Storytelling Steps to Visual Data Storytelling The Important Role of Feedback		
	5.2	Create dashboard Working with dashboard Publishing through dashboard		

Module No.	Unit No.	Details	Hrs.	CO
	5.3	Design principles for creating effective dashboards Implementing interactivity and storytelling elements Presenting insights effectively using Tableau and Power BI features		
		Total	**45**	

Recommended Books:

Sr. No.	Name/s of Author/s	Title of Book	Name of Publisher with country	Edition and Year of Publication
1.	Lindy Ryan	*Visual Data Storytelling with Tableau*	Pearson Education	First edition, 2018
0.	Cole Nussbaumer Knaflic	Storytelling with Data	Wiley	First edition, 2015
0.	Alberto Ferrari and Marco Russo	Introducing Microsoft Power BI	Microsoft Press	2016
0.	Nisal Mihiranga	Power Bi Data Modelling	BPB	First Edition 2022
0.	Brett Powell	Mastering Microsoft Power BI	Packt	First Edition , 2018

<table>
<tr><th>Course Code</th><th colspan="5">Name of the Course</th></tr>
<tr><td>216H03L701</td><td colspan="5">Data Visualization Laboratory</td></tr>
<tr><td rowspan="2">Teaching Scheme (Hrs./Week)</td><td>TH</td><td colspan="2">P</td><td>TUT</td><td>Total</td></tr>
<tr><td>-</td><td colspan="2">02</td><td>-</td><td>02</td></tr>
<tr><td>Credits Assigned</td><td>-</td><td colspan="2">01</td><td>--</td><td>01</td></tr>
<tr><td rowspan="4">Evaluation Scheme</td><td colspan="5">Marks</td></tr>
<tr><td rowspan="2">LAB/TUT CA</td><td colspan="2">CA (TH)</td><td rowspan="2">ESE</td><td rowspan="2">Total</td></tr>
<tr><td>IA</td><td>ISE</td></tr>
<tr><td>50</td><td>-</td><td>-</td><td>-</td><td>50</td></tr>
</table>

Laboratory experiments will be based on the entire syllabus of the course 216H03C701, ‘Data Visualization’. Students will be graded based on continuous assessment during laboratory.

Course Code	Name of the Course			
216H03L702	**Mini Project**			
Teaching Scheme (Hrs./Week)	**TH**	**P**	**TUT**	**Total**
	-	02	-	02
Credits Assigned	-	01	--	01

Evaluation Scheme	**Marks**				
	LAB/TUT CA	**CA (TH)**		**ESE**	**Total**
		IA	**ISE**		
	50	-	-	-	50

Course Objectives

- Skill development for design and development of data science and data analytics applications.

Course Outcomes

At the end of the course students will be able to

CO 1	Apply Data gathering techniques
CO 2	Apply and implement Data preprocessing techniques
CO 3	Apply and implement Data analysis techniques to get data insight
CO 4	Apply Machine Learning techniques
CO 5	Apply data analysis vVisualization techniques on results

Term-Work:

Term work will consist of a Mini Project with a topic of mini project selection will be based on Data Science applications covering the entire syllabus of all the courses ‘Data Science and Analytics’ Honours Degree. Students will be graded based on presentations and reports of ISE and ESE.