UNIVERSITY SCHOOL
OF
INFORMATION AND COMMUNICATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

PROGRAMME STRUCTURE

B.TECH. COMPUTER SCIENCE AND ENGINEERING
SPECIALIZATION IN DATA SCIENCE
2022-2026

30.1.5
08.09.23

GAUTAM BUDDHA UNIVERSITY
GAUTAM BUDH NAGAR, GREATER NOIDA, UP, INDIA

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

SEMESTER I

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CS101	Fundamentals of Computer Programming	3	1	0	4	CC1 / FC
2	CD101	Fundamentals of Data Science and MS Excel	2	0	0	2	CC2 / FC
3	MA101	Engineering Mathematics-I	3	1	0	4	GE1
4	PH102	Engineering Physics	3	1	0	4	GE2
5	EC101	Basic Electronics Engineering	3	1	0	4	GE3 /FC
6	EN101	English Proficiency	2	0	0	2	OE1 / AECC
7	CE103	Engineering Graphics Lab	1	0	2	2	GE-L1
8	PH104	Engineering Physics Lab	0	0	2	1	GE-L2
9	CS181	Computer Programming Lab	0	0	2	1	CC-L1 / SEC
10	EC181	Basic Electronics Engineering Lab	0	0	2	1	GE-L3
11	GP	General Proficiency	Non Credit				
		Total Hours and Credits	17	4	8	25	

SEMESTER II

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD102	Introduction to Python	2	0	0	2	CC3 / FC
2	CD104	Computer Organistaion and Architecure	3	0	0	3	CC4 / SEC
3	MA102	Engineering Mathematics-II	3	1	0	4	GE4
4	EE102	Basic Electrical Engineering	3	1	0	4	GE5
5	ME101	Engineering Mechanics	3	1	0	4	GE6
6	ES101	Environmental Studies	3	1	0	4	OE2 / AECC
7	CD182	Python Programming Lab	0	0	2	1	CC-L2 / SEC
8	EE104	Basic Electrical Engineering Lab	0	0	2	1	GE-L4
9	ME102	Workshop Practice	1	0	2	2	GE-L5
10	GP	General Proficiency	Non Credit				
		Total Hours and Credits	16	4	6	25	

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

Annexure 30.1.5

SEMESTER III

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD201	Internet Technology	3	0	0	3	CC5 / SEC
2	CD203	Operating Systems	3	0	0	3	CC6
3	CD205	Data Structure & Algorithms	3	0	0	3	CC7 / SEC
4	CD207	Optimization Problem for Data Science	3	0	0	3	CC8
5	CD209	Introduction to R Programming	3	0	0	3	CC9
6	MA201	Engineering Mathematics-III	3	1	0	4	GE7
7	CD281	R Programming Lab	0	0	3	2	CC-L3 / SEC
8	CD283	Data Structure & Algorithms Lab	0	0	3	2	CC-L4 / SEC
9	CD285	Internet Technology Lab	0	0	3	2	CC-L5 / SEC
10	GP	General Proficiency	Non Credit				
		Total Hours and Credits	18	1	9	25	

SEMESTER IV

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD202	Software Engineering	3	0	0	3	CC10
2	CD204	Database Management System	3	0	0	3	CC11 / SEC
3	CD206	Java Programming	3	0	0	3	CC12
4	CD208	Artificial Intelligence	3	0	0	3	CC13
5	CD210	Theory of Automata	3	0	0	3	CC14
6	CD212	Data Handling and Visualization	3	1	0	4	CC15/SEC
7	CD282	Database Management System Lab	0	0	3	2	CC-L6 / SEC
8	CD284	Java Programming Lab	0	0	3	2	CC-L7 / SEC
9	CD286	Data Handling and Visualization Lab	0	0	3	2	CC-L8 / SEC
10	GP	General Proficiency	Non Credit				
		Total Hours and Credits	18	1	9	25	

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

SEMESTER V

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD301	Compiler Design	3	0	0	3	CC16 / AECC
2	CD303	Soft Computing Techniques	3	0	0	3	CC17
3	CD305	Analysis and Design of Algorithms	3	0	0	3	CC18
4	CD307	Big Data Analytics	3	0	0	3	CC19
5	CD309	Machine Learning	3	1	0	4	CC20 / SEC
6		Elective 1	3	0	0	3	E1 / DSE
7	CD381	Analysis and Design of Algorithms Lab	0	0	3	2	CC-L9 / SEC
8	CD383	Big Data Analytics Lab	0	0	3	2	CC-L10 / SEC
9	CD385	Machine Learning Lab using Python	0	0	3	2	CC-L11 / SEC
10	GP	General Proficiency	Non Credit				
		Total Hours and Credits	18	1	9	25	

SEMESTER VI

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD302	Web Development using PHP	3	0	0	3	CC21
2	CD304	Introduction to Statistical Learning	3	0	0	3	CC22
3	CD306	Operation Research in Data Science	3	1	0	4	CC23
4	CD308	Cloud Computing	3	0	0	3	CC24
5	CD310	Data Privacy and Database Security	3	0	0	3	CC25 / SEC
6		Elective 2	3	0	0	3	E2 / DSE
7	CD382	Web Development using PHP Lab	0	0	3	2	CC-L12
8	CD384	Statistical Learning Lab	0	0	3	2	CC-L13
9	CD386	Data Privacy and Database Security Lab	0	0	3	2	CC-L14
10	GP	General Proficiency	Non Credit				
		Total Hours and Credits	18	1	9	25	

Industrial Training will be done by candidate individually after third year during the summer break and it will be of minimum 4 weeks. It will be evaluated as per University Examination in VII semester.

4/6

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

Annexure 30.1.5

SEMESTER VII

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD401	Parallel processing and CUDA programming	3	1	0	4	CC26
2	CD403	Cryptography and Network Security	3	0	0	3	CC27
3	CD405	Data Analytics using R	2	0	0	2	CC28 / SEC
4		Elective 3	3	0	0	3	E3 / DSE
5		Elective 4	3	0	0	3	E4 / DSE
6	CD481	Data Analytics using R Lab	0	0	3	2	CC-L15
7	CD491	Minor Project	0	0	10	5	MP1 / E
8	CD493	Industrial Traning	0	0	6	3	IT1 / E
9	GP	General Proficiency	Non Credit				
		Total Hours and Credits	14	1	19	25	

SEMESTER VIII

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD490	Seminar	0	0	3	2	S / E
2	CD492	Major Project	0	0	16	8	MP2 / E
3	CD494	Intenship	0	0	30	15	I / E
4	GP	General Proficiency	Non Credit				
		Total Hours and Credits	0	0	49	25	

GRAND TOTAL OF CREDITS = 200

In the **Seminar**, student need to study and present individually, on latest research paper of their specialized area and It will be evaluated as per University Examination Rules.

The **Internship** in Industry will be done by candidate individually during the 8th semester and it will be for a minimum of 4 (-6) months. It will be evaluated as per University Examination Rules.

Minor and Major Project will be in a group and It will be evaluated as per University Examination Rules.

USICT will provide a mentor/supervisor for industrial training, seminar, internship, minor and major projects.

08/09/23

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

Annexure 30.1.5

ELECTIVES FROM DCSE

S.No.	Course Code	Course Name	L	T	P	Credits	Types
1	CD311	Pattern Recognition	3	0	0	3	E1
2	CD313	Deep Learning	3	0	0	3	E1
3	CD315	Data Science Life Cycle	3	0	0	3	E1
4	CD317	Data Storage Technologies and Networking	3	0	0	3	E1
5	CD319	Internet of Things	3	0	0	3	E1
6	CD312	Big Data Platforms	3	0	0	3	E2
7	CD314	Research Techniques for Data Science	3	0	0	3	E2
8	CD316	High Performance Computing	3	0	0	3	E2
9	CD318	Data Mining	3	0	0	3	E2
10	CD320	Information Retrieval Systems	3	0	0	3	E2
11	CD407	Business Intelligence	3	0	0	3	E3
12	CD409	Computer Vision with Machine Learning	3	0	0	3	E3
13	CD411	Digital Image Processing	3	0	0	3	E3
14	CD413	Mobile and Wireless Network Security	3	0	0	3	E3
15	CD415	Quantam Computing	3	0	0	3	E3
16	CD417	SAS Programming	3	0	0	3	E4
17	CD419	Biomedical Image and signal processing	3	0	0	3	E4
18	CD421	AI Enabled Data Science	3	0	0	3	E4
19	CD423	Web Analytics	3	0	0	3	E4
20	CD425	Social Media Analytics and Techniques	3	0	0	3	E4

CD	Computer Science & Engineering / Data Science for Course Code		**SEC**	Skill Enhancement Course
	Core Course from USICT for course type		**E**	Elective from USICT
GE	General Elective from related discipline of other Deptt./School		**CC-L**	Core Course Lab from USICT
GE L	General Elective Lab from related discipline of other Deptt./School		**IT1**	Industrial Training
OE	Open Elective from other discipline of other Deptt./School		**MP**	Minor / Major Project
AECC	Ability Enhancement Compulsary Course		**S**	Seminar
DSE	Discipline Specific Course		**I**	Internship

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

Annexure 30.1.5

B.Tech (CSE) Specialization: Data Science

DATA HANDLING AND VISUALIZATION			
Course Code:	CD212	Course Credits:	4
Course Category:	CC	Course (U / P)	U
Course Year (U / P):	2U	Course Semester (U / P):	4U
No. of Lectures + Tutorials (Hrs/Week):	03 + 01	Mid Sem. Exam Hours:	1
Total No. of Lectures (L + T):	45 + 00	End Sem. Exam Hours:	3

COURSE OBJECTIVES
1 Understand basic of data handling
2.Understand the various visualization technologies
3. Understand and verify the underlying assumptions of a particular analysis
4. Understanding & Visualizing Bar, grouped Plots & stacked plots
5.Understand histograms, distribution analysis, statistics analysis.

COURSE OUTCOMES
At the end of the course the students should be able to:
1. Understand basics of Data Visualization.
2. Implement visualization of distributions
3. Write programs on visualization of time series, proportions & associations
4. Apply visualization on Trends and uncertainty
5. Explain principles of proportions

UNIT I INTRODUCTION TO VISUALIZATION
Visualizing Data-Mapping Data onto Aesthetics, Aesthetics and Types of Data, Scales Map Data Values onto Aesthetics, Coordinate Systems and Axes- Cartesian Coordinates, Nonlinear Axes, Coordinate Systems with Curved Axes, Color Scales-Color as a Tool to Distinguish, Color to Represent Data Values ,Color as a Tool to Highlight, Directory of Visualizations- Amounts, Distributions, Proportions, x−y relationships, Geospatial Data

UNIT II VISUALIZATION TECHNIQUES AND ASSOCIATIONS
Visualizing Amounts-Bar Plots, Grouped and Stacked Bars, Dot Plots and Heatmaps, Visualizing Distributions: Histograms and Density Plots- Visualizing a Single Distribution, Visualizing Multiple Distributions at the SameTime, Visualizing Distributions: Empirical Cumulative Distribution Functions and Q-Q Plots-Empirical Cumulative Distribution Functions, Plots, Visualizing Many Distributions at Once- Visualizing Distributions Along the Vertical Axis, Visualizing Proportions-A Case for Pie Charts, A Case for Side-by-Side Bars, A Case for Stacked Bars and Stacked Densities, Visualizing Proportions Separately as Parts of the Total ,Visualizing Nested Proportions- Nested Proportions Gone Wrong, Mosaic Plots and Treemaps, Nested Pies ,Parallel Sets. Visualizing Time Series and Other Functions of an Independent Variable-Individual Time Series , Multiple Time Series and Dose–Response Curves, Time Series of Two or More Response Variables

UNIT III PRINCIPLE OF PROPORTIONALINK
The Principle of Proportional Ink-Visualizations Along Linear Axes, Visualizations Along Logarithmic Axes, Direct Area Visualizations, Handling Overlapping Points-Partial Transparency and Jittering, 2D Histograms, Contour Lines, Common Pitfalls of Color Use-Encoding Too Much or Irrelevant Information ,Using Nonmonotonic Color Scales to Encode Data Values, Not Designing for VisionDeficiency Color

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

UNIT IV VISUALIZING UNCERTAINTY

Visualizing Trends-Smoothing, Showing Trends with a Defined Functional Form, Detrending and Time- Series Decomposition, Visualizing Geospatial Data-Projections, Layers, Choropleth Mapping, Cartograms, Visualizing Uncertainty-Framing Probabilities as Frequencies, Visualizing the Uncertainty of Point Estimates, Visualizing the Uncertainty of Curve Fits, Hypothetical Outcome Plots

UNIT V : DATA HANDLING AND VISUALIZATION USING TABLEAU

Introduction to tableau , Tableau products suite , file type , Connection to data source , Creating basic charts andgraphs , handling filter data , sorting grouping data in tableau , working with dates , waterfall chart and bump chart in tableau , heat and tree map in tableau

Text Books

1. Claus Wilke, "Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures", 1st edition, O'Reilly Media Inc, 2019.
2. Ryan Sleeper "Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master ", O'Reilly Media

Reference Books

1. Tony Fischetti, Brett Lantz, R: Data Analysis and Visualization,O'Reilly ,2016
2. Ossama Embarak, Data Analysis and Visualization Using Python: Analyze Data to Create Visualizations for BI Systems,Apress, 2018
3. Joshua N. Milligan : Learning Tableau

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

B.Tech (CSE) Specialization: Data Science

DATA HANDLING AND VISUALIZATION LAB			2
Course Code:	CD286	Course Credits:	U
Course Category: CC	CC	Course(U/P)	
Course Year(U/P):U	2U	Course Semester(U/P):	4U
No. of Labs (Hrs/Week):	02(3 hrs)		
Total No. of Lectures(L+T):30	10	End Sem. Exam Hours:	3

COURSE OBJECTIVES

1. Introduce the main concepts of visual analytics

2. leading self-service data visualization tool.

3. aims to learn about how to create effective charts and interactive dashboards will provide the student with a very useful skill applicable in many business scenarios

4. Create ad-hoc reports, data visualizations, and dashboards using Tableau Desktop

5. Publish the created visualizations to Tableau Server and/orTableau Public

COURSE OUTCOMES

At the end of the course the students should be able to:

1. Creating several different charts using Tableau

2. Preparing data for visualization

3. Identifying stories and insights in data

4. How to recognize good (and bad) data visualizations

5. How to interpret a data visualization

Program List:

1. Introduction to Tableau
- Dataviz best practices
- Getting started with Tableau Desktop
- Connecting to the tutorial dataset

2. Creating the first charts
- Filtering and sorting data

3. Common charts
- Creating common visualizations (bar charts, line charts etc.)
- Assembling a dashboard layout
- Using dashboard filters

4. Transform the data
- Dataviz best practices
- Creating simple calculations in Tableau
- Using table calculations

5. Interactions
- Interactivity with text and visual tooltips

6. Interactivity with actions (filter, highlight, URL)
- Drilldown between dashboards

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)

7. Advanced visualizations
• Dataviz best practices

8. Creating more advanced chart types
• Using multiple source tables

9Data Storytelling

• Intro to data storytelling
10.Creating a data story in Tableau

11. Overview of the Tableau ecosystem
• Further learning opportunities

School of ICT
Gautam Buddha University
Greater Noida, (U.P.)