NOIDA INSTITUTE OF ENGG. & TECHNOLOGY, GREATER NOIDA, GAUTAM BUDDH NAGAR
(AN AUTONOMOUS INSTITUTE)

Affiliated to

DR. A.P.J. ABDUL KALAM TECHNICAL UNIVERSITY UTTAR PRADESH, LUCKNOW

Evaluation Scheme & Syllabus

For

Minor Degree / Specialization

in

Data Science

School of Computer Science in Emerging Technologies

(Effective from the Session: 2022-23)

NOIDA INSTITUTE OF ENGG. & TECHNOLOGY, GREATER NOIDA, GAUTAM BUDDH NAGAR
(AN AUTONOMOUS INSTITUTE)

Minor Degree / Specialization
Data Science
EVALUATION SCHEME

Sl. No.	Subject Codes	Subject Name	Periods			Evaluation Scheme				End Semester		Total	Credit	Sem
			L	T	P	AA	QZ	TOTAL	PS	TE	PE			
1	AMSDS0301	Introduction To Data Science	3	0	0	25	25	50		100		150	3	III
2	AMSDS0401	Artificial Intelligence & Machine Learning	3	0	0	25	25	50		100		150	3	IV
3	AMSDS0501	Analyzing, Visualizing, And Applying Data Science with Python	3	0	0	25	25	50		100		150	3	V
4	AMSDS0601	Web Data Mining	3	0	0	25	25	50		100		150	3	VI
5	AMSDS0701	Business Intelligence and Data Visualization	3	0	0	25	25	50		100		150	3	VII
6	AMSDS0351	Introduction To Data Science Lab	0	0	2				25		25	50	1	III
7	AMSDS0451	Artificial Intelligence & Machine Learning Lab	0	0	2				25		25	50	1	IV
8	AMSDS0551	Analyzing, Visualizing, And Applying Data Science with Python Lab	0	0	2				25		25	50	1	V
9	AMSDS0751	Capstone Project	0	0	2				50		50	100	2	VII
		GRAND TOTAL										1000	20	

Abbreviation Used:-

L: Lecture, T: Tutorial, P: Practical, AA: Assignment Assessment, QZ: Quiz, PS: Practical Sessional, TE: Theory End Semester Exam., PE: Practical End Semester Exam.

NOIDA INSTITUTE OF ENGG. & TECHNOLOGY, GREATER NOIDA, GAUTAM BUDDH NAGAR
(AN AUTONOMOUS INSTITUTE)

Branch wise Minor Degree / Specialization Details

S.no.	Name of Minor Degree/Specialization	Streams/Branches of B.Tech. Programs whose students are eligible to opt for the Minor Degree	Streams/Branches of B.Tech. Programs whose students are eligible to opt for the Specialization
1	Artificial Intelligence and Machine Learning	All Branches except CSE and EC related Branches	CSE and EC related Branches
2	Data Science	All Branches except CSE and EC related Branches	CSE and EC related Branches
3	E-mobility	All Branches except ME related Branches	Only ME Branch
4	VLSI Design	All Branches except EC related Branches	Only EC Branch

NOIDA INSTITUTE OF ENGG. & TECHNOLOGY, GREATER NOIDA, GAUTAM BUDDH NAGAR
(AN AUTONOMOUS INSTITUTE)

Guidelines for assessment of Minor Degree / Specialization Program

For Theory Paper

Internal (50)		External (100)
AA (25)	QZ (25)	
5 Assignments of 5 marks each	5 Quiz papers of 5 marks each	Theory Examination will be Conduct at the end of Semester

For Practical Paper

Internal (25)	External (25)
On the basis of continuous Assessment	Practical Examination will be Conduct at the end of Semester

Course code	AMSDS0301	L T P	Credits
Course title	INTRODUCTION TO DATA SCIENCE	3 0 0	3

Course objective: To Provide the knowledge and expertise to become a proficient data scientist; Demonstrate an understanding of statistics and machine learning concepts that are vital for data science Produce Python code to statistically analyse a dataset.

Pre-requisites: Statistics, Basics of Python.

Course Contents / Syllabus

UNIT-I	Introduction	7 Hours

Introduction to Data Science, Different Sectors using Data science, Purpose and Components of Python in Data Science.

UNIT-II	Data Analytics and Techniques	7 Hours

Data Analytics Process, Knowledge Check, Exploratory Data Analysis (EDA), EDA- Quantitative technique, EDA- Graphical Technique, Data Analytics Conclusion and Predictions.

UNIT-III	Data Extraction and Feature Generation	11 Hours

Feature Generation and Feature Selection (Extracting Meaning from Data)- Motivating application: user (customer) retention- Feature Generation (brainstorming, role of domain expertise, and place for imagination)- Feature Selection algorithms.

UNIT-IV	Data Visualization	10 Hours

Data Visualization- Basic principles, ideas and tools for data visualization, Examples of inspiring (industry) projects- Exercise: create your own visualization of a dataset.

UNIT-V	Applications & Ethics of Data Science	7 Hours

Applications of Data Science, Data Science and Ethical Issues- Discussions on privacy, security, ethics- A look back at Data Science- Next-generation data scientists.

Course outcome: After completion of this course students will be able to:

CO 1	Understand the purpose and components of Data Science.	K2
CO 2	Understand the techniques used in EDA	K2
CO 3	Apply various processes to extract features of data.	K3
CO4	Understand the key techniques and theory behind data visualization.	K2
CO 5	Understand key applications of data science that are commonly linked to ethical issues.	K2

Textbooks:

1. Business Analytics: The Science of Data-Driven Decision Making, U Dinesh Kumar, John Wiley & Sons.
2. Introducing Data Science: Big Data, Machine Learning, and More, Using PythonTools, Davy Cielen, John Wiley & Sons.
3. Joe l Grus, Data Science from Scratch, Shroff Publisher/ O'Reilly Publisher Media
4. Annalyn Ng, Kenneth Soo, Numsense! Data Science for the Layman, Shroff Publisher.

Reference Books:

1. Cathy O'Neil and Rachel Schutt Doing Data Science, Straight Talk from The Frontline. O'Reilly Publisher.
2. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.l, Cambridge University Press.
3. Jake Vander Plas, Python Data Science Handbook, Shroff Publisher/O'Reilly Publisher Media.
4. Philipp Janer t, Data Analysis wit h Open Source Tools, Shroff Publisher/ O'Reilly Publisher Media.

Links:

Unit 1	https://www.youtube.com/watch?v=X3paOmcrTjQ
Unit 2	https://www.youtube.com/watch?v=-o3AxdVcUtQ
Unit 3	https://www.youtube.com/watch?v=kEItYHtqQUg
Unit 4	https://www.youtube.com/watch?v=MiiANxRHSv4
Unit 5	https://www.youtube.com/watch?v=8Fz2nDfZinE

Course code	AMSDS0401	L T P	Credits
Course title	ARTIFICIAL INTELLIGENCE & MACHINE LEARNING	3 0 0	3

Course objective: To understand basics of machine learning in data science and to understand various basic machine learning algorithm that can be used with various type of data.

Pre-requisites: Basics of Machine learning

Course Contents / Syllabus

UNIT-I	Introduction to Machine Learning & Linear Regression	6 Hours

INTRODUCTION – Learning, Types of Learning, well-defined learning problems, Designing a Learning System, History of ML, Introduction of Machine Learning Approaches, Underfitting and Overfitting, Bias and Variance.

Linear Regression: Basic facts of linear regression, implementation of linear regression, and case studies of linear regression using the data set.

UNIT-II	Logistic Regression	8 Hours

Logistic Regression: Basic facts and implementation of logistic regression, solve a case study to predict output using an existing data set.

UNIT-III	Clustering	11 Hours

Clustering and Principal Component Analysis: Introduction, Types of clustering, Correlations and distances, K means and hierarchical clustering, and how to make market strategies using clustering.

UNIT-IV	Support Vector Machine	9 Hours

Support Vector Machine: basics of SVM and its application to detect spam emails and recognize alphabets, SVM for classification and regression problems.

UNIT-V	Advance regression	8 Hours

Model Selection and advanced regression: use of Lasso and Ridge

Course outcome: After completion of this course students will be able to:

CO 1	Understand various types of machine learning approaches.	K2
CO 2	Demonstrate logistic regression to predict class of a dataset	K3
CO 3	Understand the role of clustering in real life dataset.	K2
CO4	Classify dataset using Support Vector Machine	K2
CO 5	Understand the model selection and advance regression process	K2

Text books:

1. Machine Learning using Python , U Dinesh Kumar and Manaranjan Pradhan, Jo hn Wiley &Sons.
2. Advance d Data Analytics Using Python: With Machine Learn ing , Deep Learning by By Sayan Mukhopadhyay
3. Practical Data Mining" by Monte F. Hancock , Auerbach Publication.

Reference Books:

1. "Machine Learning for Absolute Beginners: A Plain English Introduction (Second Edition)" by Oliver Theoba ld.
2. Practical Data Science with R, Nina Zume l, John Wiley &Sons.
3. Python for Data Science for Dummies , John Paul Mue ller, Luca Massaro n, Jo h n Wiley &Sons.
4. Big Data and Analytics, Seema Acharya and Subhas hini CheLlappan, Wiley Publication.

Links:

Unit 1	https://www.youtube.com/watch?v=lzGKRSvs5HM
Unit 2	https://www.youtube.com/watch?v=yIYKR4sgzI8
Unit 3	https://www.youtube.com/watch?v=4cxVDUybHrI
Unit 4	https://www.youtube.com/watch?v=H9yACitf-KM
Unit 5	https://www.youtube.com/watch?v=cJpWQkoe4BA

Course code	AMSDS0501	L T P	Credits
Course title	ANALYZING, VISUALIZING, AND APPLYING DATA SCIENCE WITH PYTHON	3 0 0	3

Course objective: To learn how to use tools and libraries of python for data science, and way to import, clean and prepare data for analysis. To familiarize with Pandas DataFrames, and SciKit libraries to work with various datasets and Load, manipulate, analyze, and visualize datasets with pandas.

Pre-requisites: Basics of Python.

Course Contents / Syllabus

UNIT-I	INTRODUCTION TO LIBRARIES IN PYTHON	6 Hours

Data Analysis libraries: will learn to use Pandas Data Frames, Numpy multi-dimensional arrays, and SciPy libraries to work with a various dataset.

UNIT-II	PANDAS	8 Hours

Pandas, an open-source library: load, manipulate, analyze, and visualize various datasets. Series and Data Frames, Grouping, aggregation, Merge Data Frames, Generate summary tables, Group data into logical pieces, Manipulation of data

UNIT-III	SCIKIT	10 Hours

Scikit- learn : build smart models and make predictions, various parameters that can be used to compare various parameters, Data Representation, Estimator API, Conventions, Linear Modelling, extended Linear Modeling. Anomaly Detection, KNN Learning.

UNIT-IV	DATA ANALYSIS & PREDICTION	10 Hours

Descriptive Statistics, Basic of Grouping, ANOVA, Correlation, Polynomial Regression and Pipelines, R-squared and MSE for In-Sample Evaluation, Prediction and Decision Making.

UNIT-V	MODEL EVALUATION	10 Hours

Grid Search, Model Refinement, Binning, Indicator variables, Model Evaluation, Over-fitting, Under-fitting and Model Selection, Ridge Regression.

Course outcome: After completion of this course students will be able to:

CO 1	Understand basic data analysis python libraries.	K2
CO 2	Apply the various techniques used in pandas' library.	K3
CO 3	Apply machine learning models using scikit-learn	K3
CO4	Understand the role of ANOVA in prediction and analysis of data.	K2

CO 5	Identify the importance of model evaluation and data model refinement techniques.	K2

Textbooks:

1. Data Visualization with Python and JavaScript, Kyran Dale, Shro ff Publisher/ O'ReillyPublisher Publication.
2. Data Science Using Python and R by Chanta l D. Larose and Daniel T. Larose , Wiley Publication.

Reference Books:

1. Python for Data Science and Visualization -Beginners to Pro, Udemy.

Links:

Unit 1	https://www.youtube.com/watch?v=0IbkMZHOsC0
Unit 2	https://www.youtube.com/watch?v=UB3DE5Bgfx4
Unit 3	https://www.youtube.com/watch?v=0Lt9w-BxKFQ
Unit 4	https://www.youtube.com/watch?v=TTCshtsdRuU
Unit 5	https://www.youtube.com/watch?v=08-ml-TGLLY

Course code	AMSDS0601	L T P	Credits
Course title	WEB DATA MINING	3 0 0	3

Course objective: This course covers concepts and methods used to search the web and other sources of unstructured text from a human-centred standpoint. These include document indexing, crawling, HITS algorithm; distance metrics; analysing streaming data, such as social media; link analysis; and system evaluation. To learn how to extract data from the Web and to understand how to analyse collected data to derive the most information.

Pre-requisites: Concepts of Data Warehousing and Data Mining Concepts

Course Contents / Syllabus

UNIT-I	INTRODUCTION TO DATA MINING	6 HOURS

Introduction to internet and WWW, Data Mining Foundations, Data Reduction:-Data Cube Aggregation, Dimensionality reduction, Data Compression, Numerosity Reduction, Discretization and Concept hierarchy generation, Decision Tree.

UNIT-II	ASSOCIATION RULES	8 HOURS

Mining Class Association Rules, Basic Concepts of Sequential Patterns, Mining Sequential Patterns on GSP, Mining Sequential Patterns on Prefix Span, Generating Rules from Sequential Patterns', Apriori Algorithm.

UNIT-III	WEB SPAMMING	10 HOURS

Concepts of Information Retrieval, IR Methods, Boolean Model, Vector Space Model and Statistical Language Model, Relevance Feedback, Evaluation Measures, Text and Web Page Pre-processing, Stop word Removal , Stemming, Duplicate Detection, Inverted Index and Its Compression, Index Compression, Latent Semantic Indexing , Singular Value Decomposition , Query and Retrieval, Web Search, Meta Search, Web Spamming.

UNIT-IV	CRAWLERS	10 HOURS

Link Analysis, Social Network Analysis, Co-Citation and Bibliographic Coupling, Page Rank Algorithm, HITS Algorithm, CommModuley Discovery, Problem Definition, Bipartite Core CommModuleies, Web Craw ling, *A* Basic Crawler Algorithm - Breadth First Crawlers , Preferential Crawlers , Implementation Issues - Fetching, Parsing, Link Extraction, Spider Traps, Page Repository, Universal Crawlers , Focused Crawlers, Topical Crawlers , Crawler Ethics and Conflicts.

UNIT-V	Classification	8 Hours

Opinion Mining, Sentiment Classification, Classification based on Sentiment Phrases, Classification Using Text Classification Methods, Feature based Opinion Mining and Summarization, Problem Definition, Object feature extraction, Comparative Sentence and Relation Mining, Opinion Search and Opinion Spam.

Course outcome: After completion of this course students will be able to:

CO 1	Explain data reduction and data compression of Web Text data.	K2
CO 2	Extract and analyze data and information from the webpages.	K4
CO 3	Understand the concepts of web spamming.	K2
CO4	Understand a crawler application to collect and index documents from the web.	K2
CO 5	Understand the classification of web text data using various techniques.	K2

Text books:

1. Mining the Web: Discovering Knowledge from Hypertext Data, Soumen Chakrabarti, Morgan Kaufmann Publishers.
2. Bing Liu, Web Data Mining: Exploring Hyperlinks, Contents , and Usage Data, Springer Publications, 2011.

Reference Books:

1. Anthony Scime, Web Mining: Applications and Techniques, 2005.
2. Kowalski , Gerald , Mark T Maybury: Information Retrieval Systems: Theory and Implementation.
3. Mathew Russell , Mining the Social Web 2nd Edition , Shroff Publisher/O'ReillyPublisher Publication.
4. Data Mining and Data Warehousing Principles and Practical Techniques, Parteek Bhatia, Cambridge University Press.

Links:

Unit 1	#1 Introduction To Data Mining, Types Of Data |DM| - YouTube
Unit 2	Apriori Algorithm Explained | Association Rule Mining | Finding Frequent Itemset | Edureka - YouTube
Unit 3	Search Engine Working | How Search Engines Work: Crawling, Indexing, and Ranking - YouTube
Unit 4	PageRank Algorithm - Crawling The Web With BFS - YouTube Search Engine Working | How Search Engines Work: Crawling, Indexing, and Ranking - YouTube

Course code	AMSDS0701	L T P	Credits
Course title	BUSINESS INTELLIGENCE AND DATA VISUALIZATION	3 0 0	3

Course objective: This course covers fundamental concepts of Business Intelligence tools, techniques, components and its future. As well as a bit more formal understanding of data visualization concepts and techniques. The underlying theme in the course is feature of Tableau, its capabilities.

Pre-requisites: Basic Knowledge of Business intelligence.

Course Contents / Syllabus

UNIT-I INTRODUCTION TO BUSINESS INTELLIGENCE — 8 HOURS

Business Intelligence (BI), Scope of BI solutions and their fitting into existing infrastructure, BI Components and architecture, BI Components, Future of Business Intelligence, Functional areas of BI tools, End user assumptions, setting up data for BI, Data warehouse, OLAP and advanced analytics, Supporting the requirements of senior executives including performance management, Glossary of terms and their definitions specific to the field of BI and BI systems.

UNIT-II ELEMENTS OF BUSINESS INTELLIGENCE SOLUTIONS — 8 HOURS

Business Query and Reporting, Reporting, Dashboards and Scorecards Development, Development, Scorecards, Metadata models, Automated Tasks and Events, Mobile Business Intelligence, Software development kit (SDK). Stages of Business Intelligence Projects, Project Tasks, Risk Management and Mitigation, Cost justifying BI solutions and measuring success, BI Design and Development, Building Reports, Building a Report, Drill-up, Drill-down Capabilities.

UNIT-III TABLEAU — 8 HOURS

Introductions and overview: What Tableau can and cannot do well, Debug and troubleshoot installation and configuration of the software.

Creating Your First visualization: Getting started with Tableau Software, Using Data file formats, connecting your Data to Tableau, creating basic charts (line, bar charts, Tree maps), Using the Show me panel

Tableau Calculations: Overview of SUM, AVR, and Aggregate features Creating custom calculations and fields, Applying new data calculations to your visualization.

Formatting Visualizations: Formatting Tools and Menus, formatting specific parts of the view, Editing and Formatting Axes.

UNIT-IV DATA VISUALIZATION — 8 HOURS

Manipulating Data in Tableau: Cleaning-up the data with the Data Interpreter, structuring your data, Sorting, and filtering Tableau data, Pivoting Tableau data.

Advanced Visualization Tools: Using Filters, Using the Detail panel Using the Size panels, customizing filters, Using and Customizing tooltips, Formatting your data with colours.

Creating Dashboards & Stories: Using Storytelling, creating your first dashboard and Story, Design for different displays, Adding interactivity to your Dashboard

Distributing & Publishing Your Visualization: Tableau file types, Publishing to Tableau Online, sharing your visualization, Printing, and exporting.

Given a case study: Perform Interactive Data Visualization with Tableau

UNIT-V INTRODUCTION TO POWER BI — 8 HOURS

Describe the Power BI ecosystem, Define Power BI and its relationship with Excel, Discuss the Power BI suite of products, Describe how the Power BI products integrate, Explain the typical analytics process flow, Differentiate between the various data sources, Connect Power BI to a data source, Clean and transform data to ensure data quality, Load the data to the Power BI Data Model, Describe the Power BI ecosystem, Define Power BI and its relationship with Excel, Discuss the Power BI suite of products, Describe how the Power BI products integrate, Explain the typical analytics process flow.

Course outcome: After completion of this course students will be able to

CO 1	Apply quantitative modelling and data analysis techniques to the solution of real-world business problems	K3
CO 2	Understand the importance of data visualization and the design and use of many visual components	K2
CO 3	Understand as products integrate defining various analytical process flow.	K2
CO 4	Learn the basics of troubleshooting and creating charts using various formatting tools.	K6
CO 5	Learn basics of structuring data and creating dashboard stories adding interactivity dashboard stories.	K6

Textbooks:

1. Efraim Turban, Ramesh Sharda, Dursun Delen, "Decision Support and Business Intelligence Systems", 9th Edition, Pearson 2013.
2. Learning Tableau 10 - Second Edition: Business Intelligence and data visualization that brings your business into focus" by Joshua N. Milligan
3. Tableau Your Data! - "Daniel G. Murray and the Inter Works BI Team"-Wiley

Reference Books:

1. Larissa T. Moss, S. Atre, "Business Intelligence Roadmap: The Complete Project Lifecycle of Decision Making", Addison Wesley, 2003.
2. Carlo Vercellis, "Business Intelligence: Data Mining and Optimization for Decision Making", Wiley Publications, 2009.
3. David Loshin Morgan, Kaufman, "Business Intelligence: The Savvy Manager"s Guide", Second Edition, 2012.

NPTEL/ Youtube/ Faculty Video Link:

Unit 1	Introduction to Business Intelligence - YouTube
Unit 2	Business Intelligence Tutorial - YouTube
Unit 3	What Is Power BI? | Introduction To Microsoft Power BI | Power BI Training | Edureka - YouTube
Unit 4	https://www.tableau.com/academic/students
Unit 5	Top 10 Data Visualization Tools in 2020 | Best Tools for Data Visualization | Edureka - YouTube Learn Data Visualization Using Tableau | Tableau Tutorial | Tableau | Edureka Live - YouTube RNN W2L09 : Sentiment Classification - YouTube Understanding Cluster Analysis for Customer Segmentation and Targeting - YouTube

Course code	AMSDS0351	L T P	Credit
Course title	INTRODUCTION TO DATA SCIENCE LAB	0 0 2	1

The suggested list of Experiments

Sr. No.	Name of Experiment	CO
1.	Python Environment installation/setup and Essentials.	CO1
2.	Implement Basic statistics functions (mean, mode, average, etc.) using numpy library	CO1
3.	Print multiplication table of a given number.	CO1
4.	Python Program to Find the Sum of Natural Numbers	CO1
5.	Python Program to Convert Celsius To Fahrenheit	CO1
6.	Given a list, iterate it, and display numbers divisible by five, and if you find a number greater than 150, stop the loop iteration. list1 = [12, 15, 32, 42, 55, 75, 122, 132, 150, 180, 200]	CO1
7.	Given a list, iterate it, and display numbers divisible by five, and if you find a number greater than 150, stop the loop iteration.	CO1
8.	Write a Pandas program to split the following dataframe into groups based on all columns and calculate Groupby value counts on the dataframe. Test Data: Id type book 1 10 Math 2 15 English 1 11 Physics 1 20 Math 2 21 English 1 12 Physics 2 14 English	CO2

9.	Write a Pandas program to partition each of the passengers into four categories based on their age Note: Age categories (0, 10), (10, 30), (30, 60), (60, 80)	CO2
10.	Write a Python program to plot two or more lines on same plot with suitable legends of each line.	CO1
11.	Write a Python program to plot two or more lines with legends, different widths and colours.	CO1
12.	Write a NumPy program to create a 3x3 matrix with values ranging from 2 to 10.	CO2
13.	Write a NumPy program to get help on the add function	CO2
14.	Write a Python program to create a 2-D array with ones on the diagonal and zeros elsewhere. Now convert the NumPy array to a SciPy sparse matrix in CSR format.	CO2

Lab Course Outcome: After completion of this course students will be able to:

CO1	Implement basic statistics functions in python and a variety of plots using matplotlib.	K3
CO2	Apply the fundamentals of the Pandas and Scipy library in Python	K3

Course code	AMSDS0451	L T P	Credit
Course title	ARTIFICIAL INTELLIGENCE & MACHINE LEARNING LAB	0 0 2	1

Suggested list of Experiments

Sr. No.	Name of Experiment	CO
1.	Use python to predict employee attrition in a firm and help them plan their manpower. (take data set from kaggle).	CO1
2.	Create customer clusters using different market strategies on a data set.	CO2
3.	Make a movie recommendation system.	CO2
4.	Develop a prediction mechanism to predict which employee can go on leave in a company in near future.	CO1
5.	Recognizing alphabets using SVM.	CO2
6.	Write a program to perform various types of regression (Linear & Logistic).	CO2
7.	Write a program to implement k-Nearest Neighbour algorithm to classify the iris dataset. Print both correct and wrong predictions. Python ML library classes can be used for this problem.	C02

Lab Course Outcome: After completion of this course students will be able to:

CO1	Apply linear and logistic regression models	K3
CO2	Apply Machine Learning algorithms to solve real world problems.	K3

Course code	AMSDS0551	L T P	Credit
Course title	ANALYSING, VISUALIZING AND APPLYING DATA SCIENCE WITH PYTHON LAB	0 0 2	1

Suggested list of Experiments

Sr. No.	Name of Experiment	CO
2.	Apply basic statistics function of python on New York City- 311 Complaints and Housing datasets.	CO1
3.	Visualize Iris dataset using matplotlib library.(bar, histogram, pie chart, boxplot, etc.)	CO2
4.	Write a program to predict the class of a flower based on various features of iris dataset.	CO2
5.	Write a Python program to add, subtract, multiple and divide two Pandas Series.	CO1
6.	Write a Pandas program to split the following dataframe into groups based on all columns and calculate Groupby value counts on the dataframe. Test Data: Id type book 0 1 10 Math 1 2 15 English 2 1 11 Physics 3 1 20 Math 4 2 21 English 5 1 12 Physics 6 2 14 English	CO1
7.	Write a Pandas program to partition each of the passengers into four categories based on their age Note: Age categories (0, 10), (10, 30), (30, 60), (60, 80)	CO1
8.	/Write a Pandas program to create a) Date time object for Jan 15 2012. b) Specific date and time of 9:20 pm. c) Local date and time. d) A date without time. e) Current date. f) Time from a date time. g) Current local time.	CO1
9.	Write a Pandas program to create a date from a given year, month, day and another date from a given string formats.	CO1
10.	Write a Pandas program to print the day after and before a specified date. Also print the days between two given dates.	CO1
11.	Write a Pandas program to create a time series using three months frequency.	CO1
12.	Write a Pandas program to create a sequence of durations increasing by an hour.	CO1
13.	Write a Pandas program to check if a day is a business day (weekday) or not.	CO1
14.	Write a Pandas program to create a Pivot table with multiple indexes from a given excel sheet	CO1

15.	Write a Pandas program to create a Pivot table and find the total sale amount region wise, manager wise	CO1
16.	Write a Pandas program to create a Pivot table and count the manager wise sale and mean value of sale amount.	CO2
17.	Write a Pandas program to create a Pivot table and find the maximum and minimum sale value of the items	CO2
18.	Write a Python program using Scikit-learn to print the keys, number of rows-columns, feature names and the description of the Iris data.	CO2

Lab Course Outcome: After completion of this course students will be able to:		
CO1	Understand the basic libraries in python and its implementation.	K2
CO2	Apply predictive analytics on dataset and make predictions.	K3