AC_____
Item No:_____

UNIVERSITY OF MUMBAI

Syllabus for Approval

Sr. No.	Heading	Particulars
1.	Title of the Course	**B.Sc. (Data Science)**
2.	Eligibility for Admission	F.Y.B.Sc. Data Science / 3 years Diploma from MSBTE or equivalent
3.	Passing Marks	40%
4.	Ordinances / Regulations (if, any)	As applicable for all B.Sc. Courses
5.	Number of years / Semesters	Three years – Six Semesters
6.	Level	~~P.G.~~ / U.G. /~~Diploma / Certificate~~ (Strike out which is not applicable)
7.	Pattern	~~Yearly /~~ Semester, Choice Based (Strike out which is not applicable)
8.	Status	New / ~~Revised~~
9.	To be implemented from Academic year	From the Academic Year **2021 – 2022**

Date: June 10, 2021 **Signature: _____**

Name of the BoS Chairperson ~~/ Dean~~: **Dr. Jagdish Bakal**
(bakaljw@gmail.com)

Academic Council: _____

Item No: _____

UNIVERSITY OF MUMBAI

Syllabus for B.Sc. Data Science
Second Year
Semester III and IV
Programme: B.Sc.
Subject: Data Science
CHOICE BASED(REVISED)
with effect from the academic year 2021 – 2022

SEMESTER 3

Course Code	Course Type	Course Name	Credits	Marks
USDS301	**DSC**	Research Methods and Ethics	**2**	**100**
USDS3P1	**DSC**	Research Methods and Ethics Practical	**2**	**50**
USDS302	**DSC**	Data Structures and Algorithms using Python	**2**	**100**
USDS3P2	**DSC**	Data Structures and Algorithms using Python Practical	**2**	**50**
USDS303	**SEC**	Economics	**2**	**100**
USDS3P3	**SEC**	Economics Practical	**2**	**50**
USDS304	**DSC**	Data Warehousing and Mining	**2**	**100**
USDS3P4	**DSC**	Data Warehousing and Mining Practical	**2**	**50**
USDS305	**DSC**	Linear Algebra and Discrete Mathematics	**2**	**100**
USDS3P5	**DSC**	Linear Algebra and Discrete Mathematics Practical	**2**	**50**
		Total	**20**	**750**

SEMESTER 4

Course Code	Course Type	Course Name	Credits	Marks
USDS401	**DSC**	Testing of Hypothesis	**2**	**100**
USDS4P1	**DSC**	Testing of Hypothesis Practical	**2**	**50**
USDS402	**DSC**	Big Data	**2**	**100**
USDS4P2	**DSC**	Big Data Practical	**2**	**50**
USDS403	**SEC**	Fundamentals of Accounting	**2**	**100**
USDS4P3	**SEC**	Fundamentals of Accounting Practical	**2**	**50**
USDS404	**DSC**	Artificial Intelligence	**2**	**100**
USDS4P4	**DSC**	Artificial Intelligence Practical	**2**	**50**
USDS405	**DSC**	Numerical Methods	**2**	**100**
USDS4P5	**DSC**	Numerical Methods Practical	**2**	**50**
		Total	**20**	**750**

Table of Contents

SEMESTER 3

Research Methods and Ethics

B. Sc. (Data Science)		Semester – III	
Course Name: Research Methods and Ethics		Course Code: USDS301	
Periods per week (1 Period is 50 minutes)		5	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2½	75
	Internal	--	25

Course Objectives:

1. To impart analytical skill in solving complex problems.
2. To foster the ability to critically think in developing robust, extensible and highly maintainable solutions to simple and complex problems.
3. To explore the unknown and unlock new possibilities in different dimensions of the system.
4. To portray accurately the characteristics of a particular individual, situation or a group under study.

Unit	Details	Lectures
I	**Research Methodology-An Introduction:** Meaning of Research, Objectives of Research, Motivation in Research, Types of Research, Research Approaches, Significance of Research, Research Methods versus Methodology, Research and Scientific Method, Importance of Knowing How Research is Done, Research Process, Criteria of Good Research, Problems Encountered by Researchers in India **Defining the Research Problem:** What is a Research Problem?, Selecting the Problem, Necessity of Defining the Problem, Technique Involved in Defining a Problem, An Illustration **Research Design:** Meaning of Research Design, Need for Research Design, Features of a Good Design, Important Concepts Relating to Research Design, Different Research Designs, Basic Principles of Experimental Designs.	12
II	**Sampling Design:** Census and Sample Survey, Implications of a Sample Design, Steps in Sampling Design, Criteria of Selecting a Sampling Procedure, Characteristics of a Good Sample Design, Different Types of Sample Designs, How to Select a Random Sample?, Random Sample from an Infinite Universe, Complex Random Sampling Designs **Measurement and Scaling Techniques:** Measurement in Research, Measurement Scales, Sources of Error in Measurement, Tests of Sound Measurement, Technique of Developing Measurement Tools, Scaling,	12

	Meaning of Scaling, Scale Classification Bases, Important Scaling Techniques, Scale Construction Techniques **Methods of Data Collection:** Collection of Primary Data, Observation Method, Interview Method, Collection of Data through Questionnaires, Collection of Data through Schedules, Difference between Questionnaires and Schedules, Some Other Methods of Data Collection, Collection of Secondary Data, Selection of Appropriate Method for Data Collection, Case Study Method, (i) Guidelines for Constructing Questionnaire/Schedule (ii) Guidelines for Successful Interviewing (iii) Difference between Survey and Experiment	
III	**Processing and Analysis of Data:** Processing Operations, Some Problems in Processing, Elements/Types of Analysis, Statistics in Research, Measures of Central Tendency, Measures of Dispersion, Measures of Asymmetry (Skewness), Measures of Relationship, Simple Regression Analysis, Multiple Correlation and Regression, Partial Correlation, Association in Case of Attributes, Other Measures, Summary Chart Concerning Analysis of Data **Sampling Fundamentals:** Need for Sampling, Some Fundamental Definitions, Important Sampling Distributions, Central Limit Theorem, Sampling Theory, Sandler's *A*-test, Concept of Standard Error, Estimation, Estimating the Population Mean (m), Estimating Population Proportion, Sample Size and its Determination, Determination of Sample Size through the Approach Based on Precision Rate and Confidence Level, Determination of Sample Size through the Approach, Based on Bayesian Statistics **Testing of Hypotheses:** What is a Hypothesis?Basic Concepts Concerning Testing of Hypotheses, Procedure for Hypothesis Testing, Flow Diagram for Hypothesis Testing, Measuring the Power of a Hypothesis Test, Tests of Hypotheses, Important Parametric Tests, Hypothesis Testing of Means, Hypothesis Testing for Differences between Means, Limitations of the Tests of Hypotheses	**12**
IV	**Interpretation of Data and Paper Writing** – Layout of a Research Paper, Journals in Computer Science, Impact factor of Journals, When and where to publish ?, UGC-CARE, Web of Science, SCOPUS, Ethical issues related to publishing, Copyright, Data Privacy, Plagiarism and Self-Plagiarism, Software for detection of Plagiarism. Shodh Shudhhi (PDS), smallseotools.com Use of Encyclopedias, Research Guides, Handbook etc., Academic Databases for Computer Science and Information Technology Discipline. Google Scholar, shodhganga, IEEE Xplore, ResearchGate, IDELS, DASH **Use of tools / techniques for Research:** Chicago, Turabian, MLA and APA Style, Reference Management Software like EndNote, Zotero or Mendeley; Software for paper formatting like LaTeX/MS Office/ **Scrivener/ Open Office/Google Doc/ DropBox Paper**.	**12**

V	**Ethics in business research:** What Are Research Ethics? Ethical Treatment of Participants, Ethics and the Sponsor, Researchers and Team Members, Professional Standards, Resources for Ethical Awareness **Think like a Researcher:** The Language of Research, Concepts, Constructs, Definitions, Variables, Propositions and Hypotheses, Theory, Models, Research and the Scientific Method, Sound Reasoning for Useful Answers **E-Research**: Introduction, The Internet as object of analysis, Using websites to collect data from individuals. Virtual ethnography, Qualitative research using online focus groups, Qualitative research using online personal interviews, Online social surveys, Ethical considerations in e-research, The state of e-research	**12**

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1.	Research Methodology – Methods and techniques	C. R. Kothari	New Age International (P) Ltd., Publishers	---	---
2.	Business Research Methods	Donald R. Cooper Pamela Schindler	McGraw-Hill/Irwin	12th Ed	
3.	Business Research Methods	Allan Bryman Emma Bell	OXFORD University Press	---	---
4.	RESEARCH METHODOLOGY - a step by step guide for beginners	Ranjit Kumar	SAGE Publication Ltd	---	---
5.	Research Methods for Business Students	Mark Saunders Philip Lewis Adrian Thornhill	Pearson Education Limited	---	---

Course Outcome:

CO 1: Learner understands the reasons for doing research, the applications of research, characteristics and requirements of the research process, types of research and Research paradigms.

CO 2: Learner is applying major approaches to information gathering, the relationship between attitudinal and measurement scales Methods for exploring attitudes in research.

CO 3: Learner is able to analyze data in qualitative and quantitative studies, application of IT in data analysis.

CO 4: Learner is able to write a research report and use Information Technology in Research

CO 5: Learner is practicing ethical codes and practices of conduct research.

Research Methods and Ethics Practical

B. Sc. (Data Science)		Semester – III	
Course Name: Research Methods and Ethics Practical		**Course Code: USDS3P1**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1.	**Introduction to LaTex**
a.	**Report Writing:** report style having chapter, section and subsection, article style having section, subsection and subsubsection, Automatic generation of table of contents, toc file to store the information that goes into the table of contents, Automatic numbering of section numbers
b.	**Equations and Numbering Equations:** Creating an equation, writing multiple equations, Aligning multiple equations, creating matrices in Latex, label command, Cross referencing with ref command
c.	**Tables and Figures:** Tables and Figures Creating tables and figures in LaTeX
d.	**Bibliography:** Bibliography Creating Bibliography in LaTeX
2.	**Introduction to EndNote, Zotero or Mendeley**
a.	Integration with Word and adding citation and creating bibliographies
b.	Creating your own library
c.	Creating references from website
d.	Creating references manually
3.	Visit the college library or nearby research center or from internet collect 5 tittles of research papers/thesis and classify them according to types of research, Discuss how the problems are delineated, how they are relevant to scientific method etc.
4.	Identify 2 researchable problems relevant to your context and knowledge disciplines and justify the significance of their study.
5.	Preparation of a review article
6.	Identification of variables of a research study and their classification in terms of functions and level of measurement

7.	Preparation of a sampling design given the objectives and research questions/hypotheses of a research study
8.	Preparation of questionnaire for micro-level educational survey.
9.	Prepare 1 proposal on an identified research problem
10.	Checking and removing plagiarism using Plagiarism Detection Software

Data Structures and Algorithms Using Python

B. Sc. (Data Science)		Semester – III	
Course Name: Data Structures and Algorithms Using Python		**Course Code: USDS302**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

Course Objectives:

1. To learn the essential Python data structures.
2. To learn the significant Python implementation of popular data structures
3. To learn about various data structure algorithms and design paradigms
4. To acquire knowledge of how to create complex data structures.
5. To acquire basic understanding of complex data structures such as trees and graphs and their applications

Unit	Details	Lectures
I	**Python Objects & Object-Oriented Programming:** Goals, Principles, and Patterns, Overview of data types and objects, Classes and object programming, Class Definitions, Inheritance, Data encapsulation and properties, Namespaces and Object-Orientation, Shallow and Deep Copying **Python Data Types and Strutures:** Modules for data structures and algorithms- Collections, Deques, ChainMap objects Counter, Counter objects, Ordered dictionaries defaultdict, Learning about named tuples Arrays **Principles of Algorithm Design:** An introduction to algorithms, Algorithm design paradigms Recursion and backtracking, Backtracking, Divide and conquer - long multiplication The recursive approach Runtime analysis Asymptotic analysis Big O notation, Composing complexity classes Omega notation, Theta notation, Amortized analysis	**12**
II	**Lists and Pointer Structures**:Arrays-Pointer structures **Singly linked lists**-Singly linked list class, The append operation, A faster append operation, Getting the size of the list, Improving list traversal, Deleting nodes, List search, Clearing a list **Doubly linked lists**-A doubly linked list node Doubly linked list class Append operation The delete operation List search **Circular lists**-Appending elements, Deleting an element in a circular list, Iterating through a circular list **Stacks**: Stack implementation, Push operation, Pop operation, Peek	**12**

	operation, Bracket-matching application,	
III	**Queues**:List-based queues, Stack-based queues Node-based queues, Application of queues Media player queues **Trees**: Terminology, Tree nodes, **Tree traversal**,Depth-first traversal-In-order traversal and infix notation, Pre-order traversal and prefix notation, Post-order traversal and postfix notation, Breadth-first traversal, **Binary trees**-Binary search trees,Binary search tree implementation, Binary search tree operations, Finding the minimum and maximum nodes Inserting nodes Deleting nodes, Searching the tree, Benefits of a binary search tree, Balancing trees, Expression trees,Parsing a reverse Polish expression, **Heaps**, **Ternary search tree**	**12**
IV	**Hashing and Symbol Tables:** Hashing- Perfect hashing functions Hash tables-Storing elements in a hash table, Retrieving elements from the hash table, Testing the hash table, Using [] with the hash table, Non-string keys, Growing a hash table, Open addressing, Chaining, Symbol tables **Graphs and Other Algorithms:** Graphs-Directed and undirected graphs, Weighted graphs, Graph representations, Adjacency lists, Adjacency matrices, Graph traversals- Breadth-first traversal, Depth-first search, Other useful graph methods, Priority queues and heaps- Insert operation, Pop Operation, Selection Algorithm	**12**
V	**Sorting:** Sorting algorithms- Bubble sort algorithms, Insertion sort algorithms, Selection sort algorithms, Quick sort algorithms **Selection Algorithms:** Selection by sorting, Randomized selection-Quick Select, Deterministic selection-Pivot selection Median of medians Partitioning step **Pattern Matching Algorithms:** The brute-force algorithm, The Rabin-Karp algorithm	**12**

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1.	*Hands-On Data Structures and Algorithms with Python*	Basant Agarwal, Benjamin Baka	Packt Publishing	2nd	2018
2.	*Data Structure and algorithm Using Python*	Rance D. Necaise	Wiley India Edition		2016
3.	*Data Structure and Algorithm in Python*	Michael T. Goodrich, Robertom Tamassia, M. H. Goldwasser	Wiley India Edition		2016
4.	*Data Structure and Algorithmic Thinking with Python*	Narasimha Karumanchi	Careermonk Publications		2015
5.	Fundamentals of Python: Data Structures	Kenneth Lambert	Delmar Cengage Learning		2018

Course Outcomes:

CO 1: Learner is capable of choosing appropriate data structure in Python for specified problems and algorithms.
CO 2: Learner is able to implement Linked list and Stack data structure in various domains.
CO 3: Learner is able to implement Tree and Queue data structures and use their operation.
CO 4: Learner has ability to apply of Hashing techniques, Symbol Table and Graph Algorithms appropriately.
CO 5: Learner has skills to handle sorting, searching and pattern matching on various data structures.

Data Structures and Algorithms Using Python Practical

B. Sc. (Data Science)		Semester – III	
Course Name: Data Structure and Algorithm Using Python Practical		**Course Code: USDS3P2**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

List of Practical:	
1	**General Python Programs**
A	Write Python Program to demonstrate the use of various Python Data Types and Structures
B	Write Python Program to demonstrate OOP Concepts including Class, Objects, Inheritance and encapsulation.
C	Write Python Program to implement array and operations of arrays.
2	**List and Pointer Structure**
A	Write Python Program to create singly linked list and various operations on it..
B	Write Python Program to create doubly linked list and various operations on it.
c	Write Python Program to create circular linked list and various operations on it.
3	**Stacks and Queues**
a.	Write Python Program to implement stack and demonstrate push, pop and peek operations.
b.	Write Python Program to implement stack for Bracket-matching application
c.	Write Python Program to implement list based queues and demonstrate various operations on it.
d.	Write Python Program to implement stack based queues and demonstrate various operations on it.
e.	Write Python Program to implement Node based queues and demonstrate various operations on it.
f.	Write Python Program to implement queue data structure for simulating media player playlist queues.
4	**Trees**
a.	Write Python Program to implement tree data structure and demonstrate depth-first traversal
b.	Write Python Program to implement tree data structure and demonstrate breadth-first traversal
c.	Write Python Program to implement binary search tree to find the minimum node.

d.	Write Python Program to implement binary search tree to find the minimum node.
e.	Write a Python implementation to demonstrate the insert and delete method to add/delete the nodes in the BST.
f.	Python implementation to search the node in the BST
g.	Write a python program build up a tree for an expression written in postfix notation and evaluate it.
5	**Hashing and Symbol Tables**
a.	Write a Python Program to demonstration of computing Hash for given strings.
b.	Write a Python program to implement hash table for storing and searching values from it.
c.	Write aa Python Program to create Symbol Table
6	**Graphs**
a.	Write a Python program to store and display Graph data structure using adjacency matrix.
b.	Write a Python Program to implement Graph traversal (BFS/DFS) based on above practical.
c.	Write a Python program to implement priority queue and heap operations
7	**Searching**
a.	Write a Python Program for implementation in Python for the linear search on an unordered list of items
b.	Write a Python Program for implementation in Python for the linear search on an ordered list of items
c.	Write a Python Program for implementation of the binary search algorithm on an ordered list of items
d.	Write a Python Program for implementation of implementation of the interpolation search algorithm
8	**Sorting**
a.	Write a Python Program for implementing Insertion Sort.
b.	Write a Python Program for implementing Bubble Sort.
c.	Write a Python Program for implementing Quick Sort.
d.	Write a Python Program for implementing Selection Sort.
9	**Selection Algorithms**
a.	Write a Python Program to implement Randomized Selection
b.	Write a Python Program to implement Deterministic Selection
10	**Application**
a.	Write a Python Program to create an application for storing Polynomial
b.	Write a Python Program to create an application for adding two Polynomials

Economics

B. Sc. (Data Science)		Semester – III	
Course Name: Economics		Course Code: USDS303	
Periods per week (1 Period is 50 minutes	Lectures	5	
	Credits	2	
		Hours	Marks
Evaluation System	Theory Examination	2½	75
	Theory Internal	--	25

Course Objectives:

1. To understand Fundamental economic ideas and the operation of the economy on a national scale.
2. Basic Understanding of production, distribution and consumption of goods and services, the exchange process, the role of government, the national income and its distribution, GDP, consumption function, savings function, investment spending and the multiplier principle
3. Acquire basic knowledge of the influence of government spending on income and output.
4. Develop ability to analyze monetary policy, including the banking system and the Federal Reserve System.

Unit	Details	Lectures
I	**First Principles:** What is the difference between macro and micro economics? The central choices of economic decision making: what, how and for whom to produce? The participants in the market economy Key concepts used in economic analysis: Scarcity, choice, opportunity cost Marginal analysis and choice Ceteris Paribus or 'everything else held constant.' Positive and normative economics and using theories and models to measure economic events Criteria for evaluation of economic policy and policy proposals Economic systems – the market economy, mixed economies & command economies Review of expressing relationships between economic variables using graphs **Economic Models:** Trade-offs and Trade: Defining the resources used in the production of goods and services The production possibility frontier applied to the concept of opportunity cost/tradeoffs and to marginal costs and benefits; increasing marginal opportunity costs. Productive efficiency; inefficient choices and unattainable choices, Using the frontier to illustrate economic growth, attainment of new resources, technological change, and more efficient production. Comparative advantage and the gains from trade The circular flow of income, product and services in the economy	12

II	**Supply and Demand: Product and Resource Markets** – Role of households (consumers) and firms What is a market? Consumer demand and the "Law of Demand" Law of Demand: the inverse relationship between price and quantity demanded Change in quantity demanded vs. shift in demand: the concept of "ceteris paribus" Causes of a shift in demand: changes in income, expectations, number of consumers, tastes and preferences; Normal and inferior goods Law of Supply: The positive relationship between price and quantity supplied. Change in quantity supplied vs. a shift in supply Causes of a shift in Supply: changes in cost of resources, prices of related goods, technology, expectations of producers, number of producers Applications (examples) of Demand and Supply graphs; Market demand, market supply and market equilibrium Government price controls: price ceilings, price floors **Macroeconomics:** Theory and Policy: The Business Cycle in Market Economies; short-term vs. long-term growth trend Expansion, peak, decline, trough Emergence of modern-day macroeconomic policy to moderate effects of recessions: Keynesian policy/government spending and taxation to stimulate aggregate demand Components of aggregate demand and aggregate supply Shifts in the AD and AS curves: What do they show? The roots of macroeconomics: John Maynard Keynes and the Great Depression Classical vs. Keynesian economics; the short-run vs. long run model of macroeconomic equilibrium The Keynesian short-run model and the classical economists' long-run model Keynes' challenge to Say's Law: the Demand Driven Economy Wage and Price inflexibility; The role of Government Concerns of Inflation (boom times) and deflation (severe economic downturns) The impact of recession on trade imbalances	**12**
III	**Unemployment and Inflation:** How is the labor force defined? Who is in the labor force? Measuring employment and unemployment. Who is not counted in the Government's official count of the unemployed? Types of unemployment; cyclical unemployment and the business cycle. The difference between the 'household survey' (the civilian labor force) and the 'establishment survey' (number of payroll jobs added by employers). The labor force participation rate Unemployment and the changes in the global economy **Gross Domestic Product:** Measuring the economy's output of goods and services; Government Sector: federal state and local government in the economy The financial sector; the international sector The three markets: goods and service, labor market, money market Nominal and real GDP; The difference between GNP and GDP Expenditure Measure of GDP: consumption by households, businesses, government and the rest of the world (Net exports) Income Measure of GDP: Income from labor, rent, interest, proprietors' income, profit Value added approach vs. measure of final goods and services produced What GDP Does Not Include; alternative measures of GDP	**12**

IV	**Measuring inflation the consumer price index:** What does it say about the state of the economy? Real vs. nominal income and earnings Real and Nominal rates of interest Costs and causes of inflation **Fiscal policy:** Defining fiscal policy: taxation and spending to achieve macroeconomic goals The role of government in the U.S. economy Fiscal policy and the Recession of 2007 – 2009 The Employment Act of 1946 A history of U.S. fiscal policy since the early 20th century The multiplier effect Government spending and taxation Automatic stabilizers: the income tax, unemployment insurance Discretionary tax and spending policy Progressive, proportional and regressive taxes and their impacts Fiscal Policy Lags The circular flow diagram with government spending and taxation Budget deficits and surpluses; Government debt and deficits: Are they the same thing?	**12**
V	**Money, Banking and the Federal Reserve System:** What is money? Commodity and fiat monies; the barter system Money as a medium of exchange; Money supply defined: M1 and M2 Gold and the money supply: Why was the gold standard adopted (1873) and why was it later eliminated (1971)? Monetary role of banks; Establishment of bank reserves; The T-account (assets and liabilities) Bank regulation: the FDIC deposit insurance; capital requirements; the discount window at the Fed. **Monetary Policy:** The structure of the Federal Reserve System How the Fed regulates the money supply: reserve requirements, the discount rate, open market operations; the goals of monetary policy The federal funds rate; fed funds market Banking legislation and deregulation since the 1980's Growth of the "Shadow Banking System" and the financial crisis of 2007-2009 The role of credit, debit cards and electronic money in the money supply Role of financial intermediaries – modern depository institutions Savings and Loan crisis of the late 1980's The financial crisis of 2008 and the Federal Reserve's policy response How the banking system creates and expands money in circulation The difference between treasury bonds and bonds issued by the Fed Fed Policies during the 2007 – 2009 Recession	**12**

,

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
01	Macroeconomics	Krugman and Wells, Eds.,	Worth Publishers	3 rd	2012
02	Macroeconomics	Leeds, Michael A., von Allmen, Peter and Schiming, Richard C	Pearson Education	1st	2006

03	Lectures in Quantitative Economics with Python	Thomas J. Sargent and John Stachurski	----	----	2019

Course Outcome

CO1: Learner understands the basic economic decisions that underline the economic process: What and how to produce goods and services and how they are distributed.

CO2: Learner is able to apply of the concepts of scarcity, choice and opportunity cost to analyze the workings of a market economy.

CO3: Learner is able to demonstrate a firm knowledge of the interrelationships among consumers, government, business and the rest of the world in the U.S. macroeconomy.

CO4: Learner is able to identify the process of how the nation's output of goods and services is measured through the national income and product accounts; clearly comprehend the income and expenditure approaches to measuring national output and national income.

CO5: Learner is capable to clearly illustrate the specific roles and functions of monetary and fiscal policy in the economy and explain how these are applied to the process of shaping economic policy and stabilizing the economy, specifically regarding controlling inflation, promoting full employment and facilitating economic growth.

Economics Practical

B. Sc. (Data Science)		**Semester – III**	
Course Name: Economics Practical		**Course Code: USDS3P3**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1	Application to Asset Pricing using Geometric series for elementary economics in R/python/scilab/matlab**.**
2	Cass-Koopmans Optimal Growth Model using R/python/scilab/matlab.
3	a. Job Search and Separation using R/python/scilab/matlab**.**
	b. Modeling Career Choice using R/python/scilab/matlab.
4	Consumption and Tax Smoothing with Complete and Incomplete Markets using R/python/scilab/matlab.
5	A Lake Model of Employment and Unemployment using R/python/scilab/matlab.
6	Cattle Cycles model using R/python/scilab/matlab.
7	Von Neumann Growth Model using R/python/scilab/matlab.
8	The Lucas Asset Pricing Model using R/python/scilab/matlab.
9	Implement the optimal government plan model using R/python/scilab/matlab**.**
10	Credible Government Policies in Chang Model R/python/scilab/matlab.

Data Warehousing and Mining

B. Sc. (Data Science)		Semester – III	
Course Name: Data Warehousing and Mining		**Course Code: USDS304**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

Course Objectives:

1. To understand business intelligence for an enterprise and review data warehouse with architectural types and architectural building blocks
2. To discuss and understand changing dimensions and learn about aggregate tables and determine their usage.
3. To learn basics of data mining, understand the need and the process of data mining in contrast with machine learning.
4. To study the use of classification and clustering techniques for Data Mining.
5. To appreciate the use of various data mining algorithms and learn about their specific applications.

Unit	Details	Lectures
I	THE COMPELLING NEED FOR DATA WAREHOUSING: Escalating Need for Strategic Information, Failures of Past Decision-Support Systems, Operational Versus Decision-Support Systems, Data Warehousing—The Only Viable Solution, Data Warehouse Defined, The Data Warehousing Movement, Evolution of Business Intelligence DATA WAREHOUSE: The Building Blocks: Defining Features, Data Warehouses and Data Marts, Architectural Types, Overview of The Components, Metadata in The Data Warehouse TRENDS IN DATA WAREHOUSING: Continued Growth in Data Warehousing, Significant Trends, Emergence of Standards, Web-Enabled Data Warehouse ARCHITECTURAL COMPONENTS: Understanding Data Warehouse Architecture, Distinguishing Characteristics, Architectural Framework, Technical Architecture, Architectural Types THE SIGNIFICANT ROLE OF METADATA: Why Metadata Is Important, Metadata Types By Functional Areas, Business Metadata, Technical Metadata, How To Provide Metadata	**12**
II	PRINCIPLES OF DIMENSIONAL MODELING: From Requirements to Data Design, The Star Schema, Star Schema Keys, Advantages of The Star Schema, Star Schema: Examples DIMENSIONAL MODELING: ADVANCED TOPICS: Updates to The Dimension Tables, Miscellaneous Dimensions, The Snowflake Schema, Aggregate Fact Tables, Families of Stars	**12**

	DATA EXTRACTION, TRANSFORMATION, AND LOADING: ETL Overview, ETL Requirements and Steps, Data Extraction, Data Transformation, Data Loading, ETL Summary, Other Integration Approaches	
III	INTRODUCTION TO DATA MINING: Introduction to Data Mining, Need of Data Mining, What Can Data Mining Do and Not Do? Data Mining Applications, Data Mining Process, Data Mining Techniques, Difference between Data Mining and Machine Learning BEGINNING WITH WEKA AND IRIS DATASET IN R: About Weka, Installing Weka, Understanding Fisher's Iris Flower Dataset, Preparing the Dataset, Understanding A RFF, Working with a Dataset in Weka, Working with the Iris dataset in R Data Preprocessing: Need for Data Preprocessing, Data Preprocessing Methods CLASSIFICATION: Introduction to Classification, Types of Classification, Input and Output Attributes, Guidelines for Size and Quality of the Training Dataset, Introduction to the Decision Tree Classifier, Naive Bayes Method, Understanding Metrics to Assess the Quality of Classifiers	**12**
IV	CLUSTER ANALYSIS: Introduction to Cluster Analysis, Applications of Cluster Analysis, Desired Features of Clustering, Distance Metrics, Major Clustering Methods/Algorithms, Partitioning Clustering, HIERARCHICAL CLUSTERING ALGORITHMS: Web Mining and Search Engines: Introduction, Web Content Mining, Web Usage Mining, Web Structure Mining, Hyperlink Induced Topic Search algorithm, Introduction to Modern Search Engines, Working of a Search Engine, PageRank Algorithm, Precision and Recall	**12**
V	INTRODUCTION TO ASSOCIATION RULE MINING: Defining Association Rule Mining, Representations of Items for Association Mining, The Metrics to Evaluate the Strength of Association Rules, The Naive Algorithm for Finding Association Rules, Approaches for Transaction Database Storage THE APRIORI ALGORITHM, Closed and Maximal Itemsets, The Apriori–TID Algorithm for Generating Association Mining Rules, Direct Hashing and Pruning (DHP), Dynamic Itemset Counting (DIC), Mining Frequent Patterns without Candidate Generation (FP Growth)	**12**

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1.	DATA WAREHOUSING FUNDAMENTALS FOR IT PROFESSIONALS	PAULRAJ PONNIAH	Wiley	Second	2010
2.	Data Mining and Data Warehousing : Principles and Practical Techniques	Parteek Bhatia	Cambridge University Press	First	2019

3.	The Data Warehouse Toolkit	Ralph Kimball Margy Ross	Wiley	Third	2013
4.	Encyclopedia of Data Warehousing and Mining	John Wang	Information Science Reference	Second	2008
5.	Data Mining and Data Warehousing	S.K. Mourya Shalu Gupta	Alpha Science International Ltd	First	2013

Course Outcomes:

CO1: Learner is able to demonstrate knowledge of business intelligence, data warehouse with clear understanding of architectural types and will be able to establish the relationship between architectural building blocks.
CO2: Learner is able to elaborate changing dimensions with respect to current trends & using aggregate tables.
CO3: Learner is able to handle the processes of data preprocessing, data transformation and data reduction.
CO4: Learner has knowledge of using various Data Mining techniques for classification and clustering.
CO5: Learner is able to align the Data Mining techniques for analyzing the datasets using tools like Weka, R or Python

Data Warehousing and Mining Practical

B. Sc. (Data Science)		Semester – III	
Course Name: Data Warehousing and Mining Practical		**Course Code: USDS3P4**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1.	**Data warehouse design**
a.	Design dimension tables.
b.	Design fact tables.
c.	Create an indexed view and rebuild columnstore indexes.
2.	**Data Warehouse with Azure**
a.	Create an Azure SQL Data Warehouse Project.
b.	Develop tables in Azure SQL Data Warehouse.
c.	Migrate Data Warehouse to Azure.
d.	Pause and remove Azure data warehouse.
3.	**Data Warehouse implementation and use**
a.	Cleanse data with SQL Server Data Quality Services.
b.	Create custom knowledge base.
c.	Install Master Data Services and IIS.
d.	Configure MDS and deploy sample MDS model.
e.	Install MDS excel add-in and Update master data in excel.
f.	Consume the data from the warehouse.
4.	**Working with Data and Data Preprocessing**
a.	Demonstrate the use of ARFF files taking input and display the output of the files.
b.	Create your own excel file. Convert the excel file to .csv format and prepare it as ARFF files.
c.	Preprocess and classify Customer dataset. http://archive.ics.uci.edu/ml/
d.	Perform Preprocessing, Classification techniques on Agriculture dataset. (http://archive.ics.uci.edu/ml/)
e.	Preprocess and classify Weather dataset. http://archive.ics.uci.edu/ml/
f.	Perform data Cleansing of customer dataset. http://archive.ics.uci.edu/ml/ www.kdnuggets.com/datasets/
5.	**Performing classification on data sets**
a.	Building a Decision Tree Classifier in Weka
b.	Applying Naïve Bayes on Dataset for classification

c.	Creating the Testing Dataset
d.	Decision Tree Operation with R
e.	Naïve Bayes Operation using R
f.	Classify the dataset using decision tree. www.kdnuggets.com/datasets/
6.	**Simple Clustering**
a.	Perform Clustering technique on Customer dataset. http://archive.ics.uci.edu/ml/
b.	Perform Clustering technique on Agriculture dataset. http://archive.ics.uci.edu/ml/
c.	Perform Clustering technique on Weather dataset. http://archive.ics.uci.edu/ml/
7.	**Implementing Clustering with Weka and R**
h.	Clustering Fisher's Iris Dataset with the Simple k-Means Algorithm
i.	Handling Missing Values
j.	Results Analysis after Applying Clustering
k.	Classification of Unlabeled Data
l.	Clustering in R using Simple k-Means
8.	**Implementing Apriori Algorithm with Weka and R**
a.	Applying Predictive Apriori in Weka
b.	Applying the Apriori Algorithm in Weka on a Real World Dataset
c.	Applying the Apriori Algorithm in Weka on a Real World Larger Dataset
d.	Applying the Apriori Algorithm on a Numeric Dataset
9.	**Implementing Association Mining with R**
a.	Applying Association Mining in R
b.	Application of Association Mining on Numeric Data in R
c.	Perform Association technique on Agriculture dataset. http://archive.ics.uci.edu/ml/,www.kdnuggets.com/datasets/
d.	Perform Association technique on Agriculture dataset. http://archive.ics.uci.edu/ml/ www.kdnuggets.com/datasets/
e.	Perform Association technique on Weather dataset.
10.	**Web Mining**
a.	Implement Hyperlink Induced Topic Search (HITS) algorithm
b.	Implement PageRank Algorithm

Linear Algebra and Discrete Mathematics

B. Sc. (Data Science)		**Semester – III**	
Course Name: Linear Algebra and Discrete Mathematics		**Course Code: USDS305**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

Course Objectives

1. To analyze the solution set of a system of linear equations.
2. To interpret existence and uniqueness of solutions geometrically.
3. To formulate, solve, apply, and interpret properties of linear systems.

Unit	Details	Lectures
I	**Matrices and Gaussian Elimination:** Introduction, The Geometry of Linear Equations, An Example of Gaussian Elimination, Matrix Notation and Matrix Multiplication, Triangular Factors and Row Exchanges, Inverses and Transposes, Special Matrices and Applications **Vector Spaces:** Vector Spaces and Subspaces, Solving $Ax = 0$ and $Ax = b$, Linear Independence, Basis, and Dimension, The Four Fundamental Subspaces, Graphs and Networks, Linear Transformations	**12**
II	**Orthogonality:** Orthogonal Vectors and Subspaces, Cosines and Projections onto Lines, Projections and Least Squares, Orthogonal Bases and Gram-Schmidt, The Fast Fourier Transform **Determinants:** Introduction, Properties of the Determinant, Formulas for the Determinant, Applications of Determinants	**12**
III	**Eigenvalues and Eigenvectors:** Introduction, Diagonalization of a Matrix, Difference Equations and Powers A^k, Differential Equations and e^{At}, Complex Matrices, Similarity Transformations	**12**
IV	**Positive Definite Matrices:** Minima, Maxima, and Saddle Points, Tests for Positive Definiteness, Singular Value Decomposition, Minimum Principles, The Finite Element Method **Computations with Matrices:** Introduction, Matrix Norm and Condition Number, Computation of Eigenvalues, Iterative Methods for $Ax = b$	**12**
V	**Linear Programming and Game Theory:** Linear Inequalities, The Simplex Method, The Dual Problem, Network Models, Game Theory	**12**

Books and References:					
Sr. No.	**Title**	**Author/s**	**Publisher**	**Edition**	**Year**
1	Linear Algebra and Its Applications	Gilbert Strang	**Cengage Publication**	Fourth Edition	**----**
2.	Advanced Linear Algebra	David Surowski			
3.	Linear Algebra, Theory and Applications	Kenneth Kuttlet			

Course Outcome:

CO 1: Learner is able to perform common matrix operations such as addition, scalar multiplication, multiplication, and transposition.
CO 2: Learner is able to describe how the determinant of a product of matrices relates to the determinant of the individual matrices.
CO 3: Learner expresses clear understanding of the concept of a 'solution to a *game'* and also the limitations on the applicability of the theory.

Linear Algebra and Discrete Mathematics Practical

B. Sc. (Data Science)		Semester – III	
Course Name: Linear Algebra and Discrete Mathematics Practical		**Course Code: USDS3P5**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1	**Matrices and Gaussian Elimination**
a	Multiplication and transpose of matrix using R/python/scilab/matlab.
b	Inverses of matrix in R/python/scilab/matlab without using any inbuilt package.
c	Inverses of matrix in R/python/scilab/matlab using any inbuilt package like numpy.
d	Linear equation with n unknowns using Gauss Elimination Method using R/python/scilab/matlab.
2	**Vector**
a	Addition, subtraction, multiplication and division of vector using R/python/scilab/matlab.
b	dot product & cross product of vector using R/python/scilab/matlab
c	Visualising vector Linear Transformations using R/python/scilab/matlab
3	
a	Computes the orthonormal vectors using the GS algorithm using R/python/scilab/matlab.
b	Projections and Least Squares using R/python/scilab/matlab.
c	Fast Fourier Transform using R/python/scilab/matlab.
4	
a	Finding determinant of matrix in R/python/scilab/matlab without using any inbuilt package.
b	Finding determinant of matrix in R/python/scilab/matlab using any inbuilt package.
5	
a	Compute the eigenvalues and right eigenvectors of a given square array using R/python/scilab/matlab.

b	Program to test diagonalizable matrix using R/python/scilab/matlab.
6	
a	Tests for Positive Definiteness using R/python/scilab/matlab.
b	Singular Value Decomposition using R/python/scilab/matlab.
c	The Finite Element Method using R/python/scilab/matlab. **(Only Demonstration)**
7	Simplex Method using R/python/scilab/matlab. **(Only Demonstration)**
8	The Dual Problem using R/python/scilab/matlab. **(Only Demonstration)**
9	Implementing Network Models using R/python/scilab/matlab. **(Only Demonstration)**
10	Implementing Game Theory using R/python/scilab/matlab. **(Only Demonstration)**

SEMESTER-4

Testing of Hypothesis

B. Sc. (Data Science)		Semester – IV	
Course Name: Testing of Hypothesis		Course Code: USDS401	
Periods per week (1 Period is 50 minutes)		5	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2½	75
	Internal	--	25

Course Objectives:

1. To impart statistical significance in solving complex problems.
2. To critically test in developing robust, extensible and highly maintainable solutions to simple and complex problems.
3. To implement various statistical functions using suitable programming languages and packages.
4. To scientifically test the unknown and unlock possibilities in different dimensions of the system.
5. To write the reports of analytical results generated by the system.

Unit	Details	Lectures
I	**Introduction to Hypothesis Testing:** Hypothesis Tests, Stating a Hypothesis, Types of Errors and Level of Significance, Statistical Tests and P-Values, Making a Decision and Interpreting the Decision, Strategies for Hypothesis Testing, Characteristics of a good hypothesis, Steps for hypothesis testing **Hypothesis Testing for the Mean (σ Known):** Using P-Values to Make Decisions, Using P-Values for a z-Test, Rejection Regions and Critical Values, Using Rejection Regions for a z-Test, Critical Values in a t-Distribution, The t-Test for a Mean μ, Using P-Values with t-Tests, Sums and case studies **Packages used for Hypothesis testing:** Introduction to statistical functions in R / Python / Excel, Packages used for finding P-value to make decision and hypothesis testing.	12
II	**Goodness of fit tests:** Anderson-Darling, Chi-square test, Kolmogorov-Smirnov, Ryan-Joiner,Shapiro-Wilk ,Jarque-Bera, Lilliefors **Variance tests:** Chi-square test of a single variance, F-tests of two variances, Tests of homogeneity **Wilcoxon rank-sum/Mann-Whitney U test, Sign test** **Contingency tables:** Chi-square contingency table test, G contingency table test, Fisher's exact test, Measures of association, McNemar's test **Packages used for Hypothesis testing:** Packages used for finding goodness of fit test, variance test, Wilcoxon rank-sum / Mann-Whitney U test and sign test, Using Contingency table in R / Python / Excel.	12

III	**Analysis of variance and covariance:**ANOVA**,** Single factor or one-way ANOVA, Two factor or two-way and higher-way ANOVA , MANOVA, ANCOVA **Non-Parametric ANOVA:** Kruskal-Wallis ANOVA, Friedman ANOVA test, Mood's Median **Packages used for Hypothesis testing:** Packages used for finding Anova, Manova, Ancova and Non-Parametric Anova in R / Python / Excel.	**12**
IV	**Regression and smoothing:** Least squares, Ridge regression, Simple and multiple linear regression, Polynomial regression, Generalized Linear Models (GLIM), Logistic regression for proportion data, Poisson regression for count data, Non-linear regression, Smoothing and Generalized Additive Models (GAM), Geographically weighted regression (GWR), Spatial series and spatial autoregression- SAR models, CAR models, Spatial filtering models **Time series analysis and temporal autoregression:** Moving averages, Trend Analysis, ARMA and ARIMA (Box-Jenkins) models, Spectral analysis	**12**
V	**Communicating and Documenting the Results of Analyses:** Introduction, The Difficulty of Good Communication, Communication Hurdles: Graphical Distortions, Communication Hurdles: Biased Samples & Sample Size, Preparing Data for Statistical Analysis, Guidelines for a Statistical Analysis and Report, Documentation and Storage of Results ,Supplementary Exercise **Data Storytelling:** What is a Data Story?, The Art and Science of Storytelling, Planning the Data Story, Elements of the Data Story, Parts of the Data Story, Framing and Formatting of the Data Story, False Narratives and Data Storytelling **Infographics:** What is an Infographic?, Why are Infographics Useful? Types of Infographics, Infographic Design Elements, Steps in Designing an Infographic, Best Practices in Designing an Infographic	**12**

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	Hypothesis Testing	---	Pearson Higher Education	---	---
2	Statistical Analysis Handbook	Dr. Michael J de Smith	The Winchelsea Press, Drumlin Security Ltd, Edinburgh	2018 Ed	2018

3	An Introduction to Statistical Methods and Data Analysis	R. Lyman Ott & Michael Longnecker	Thomson Learning	---	---

Course Outcome:

CO 1: Learner is developing null and alternative hypotheses to test for a given situation.
CO 2: Learner is able to differentiate one- and two-tailed hypothesis tests.
CO 3: Learner is able to do sampling a normal distribution and random sampling.
CO 4: Learner is using statistical models and their associations in performing hypothesis testing.
CO 5: Lerner is writing the reports and interpreting the data using the various programming languages and packages.

Testing of Hypothesis Practical

B. Sc. (Data Science)		Semester – IV	
Course Name: Testing of Hypothesis Practical		**Course Code: USDS4P1**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
Practical can be performed using R / Python / scilab / matlab / SPSS / MS Excel	
1	**Hypothesis Testing for the Mean**
a	Perform testing of hypothesis using one sample t-test.
b	Perform testing of hypothesis using two sample t-test.
c	Perform testing of hypothesis using paired t-test.
d	Perform testing of hypothesis using Z-test.
2	**Goodness-of-fit test**
a	Perform goodness-of-fit test using chi-squared test.
b	Perform goodness-of-fit test using KS-test.
c	Perform testing of hypothesis using chi-squared Test of Independence
3	Variance Testing
a.	Using Chi-square test of a single variance
b.	Using F-tests of two variances
c.	Testing of homogeneity
4	**Analysis of variance and covariance**
a.	Perform testing of hypothesis using one-way ANOVA.
b.	Perform testing of hypothesis using two-way ANOVA.
c.	Perform testing of hypothesis using Multivariate ANOVA (MANOVA)
d.	Perform testing of hypothesis using one-way ANOVA.
5	**Regression**
a.	Perform simple linear regression
b.	Perform multiple linear regression
c.	Perform polynomial regression
6	Perform spatial series and spatial auto-regression
7	Perform time series analysis using Moving averages
8	Perform time series analysis using Trend Analysis
9	Perform Spectral analysis
10	Creating “Infographics” using secondary data available on internet. (Use Canva / Adobe Spark / Prezi / Vennage

Big Data

B. Sc. (Data Science)		Semester – IV	
Course Name: Big Data		**Course Code: USDS402**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

Course Objectives

1. To develop core abilities to make data-driven decisions through big data.
2. To provide an overview of an exciting growing field of big data analytics.
3. To introduce the tools required to manage and analyze big data like Hadoop, NoSql Map Reduce.
4. To teach the fundamental techniques and principles in achieving big data analytics with scalability and streaming capability.

Unit	Details	Lectures
I	Introduction to Big Data Analytics: Defining Big Data analytics: Discovering value from large data sets, Exploiting data to optimize decision-making Planning your analytics life cycle project: Outlining steps in the life cycle, Contrasting traditional analytics with Big Data analytics Representing Big Data with R and Rattle: Preparing the data: Loading data for knowledge discovery, Spotting outliers in the data, Transforming and summarizing data Visualizing data characteristics: Revealing changes over time, Displaying proportions within your data, Leveraging charts to display relationships, Displaying relationships across categories	**12**
II	*Modeling and Predictive Data Analysis:* Categorizing analytic approaches: *Predictive vs. descriptive analytics, Supervised vs. unsupervised learning* Applying appropriate mining techniques: Discovering unknown groups through clustering, Detecting relationships with association rules, Uncovering decision tree classifications, Identifying patterns with time series analysis	**12**
III	Leveraging Analytics with RHadoop Expanding the analytic capabilities of your organization • Exploring the MapReduce and Hadoop architecture, Creating and executing HadoopMapReducejobs, Integrating R and Hadoop with RHadoop, Examining the components of RHadoop, Creating modules for RHadoopjobs, ExecutingRHadoopjobs, Monitoring job execution flow	**12**

	Building a Recommendation Framework: Streamlining business decisions • Considering motivations for a recommender engine, Leveraging recommendations based on collaborative filtering, Exploring the architecture of the recommendation framework, Building programming components, Executing the recommendation model, Performing tradeoff analysis	
IV	Mining Unstructured Data Investigating business value within unstructured data • Making a business case for unstructured data mining, Extending business intelligence with mining tools Implementing text mining and social network analysis • Analyzing the structure of text mining, Evaluating mining approaches, Building a text mining framework, Inspecting social network interactions	**12**
V	*Planning and Implementing a Complete Data Analytics Solution* Transforming business objectives to analytic projects • Arguing your business case for analytics, Mapping analytics models to business objectives, Identifying performance metrics targets Implementing the analytics life cycle • Finding core data sets, Preparing the data for analysis, Modeling the data, Executing the model, Communicating results *Ensuring a Successful Data Analytics Solution* • Identifying barriers to Big Data analytics, Managing and mitigating risks, Employing an implementation checklist	**12**

Reference Books:

1. **VigneshPrajapati, "Big Data Analytics with R and Hadoop", Packt Publishing House.**
2. **Bart Baesens, Analytics in a Big Data World: The Essential Guide to Data Science and its Applications (Wiley and SAS Business Series), Wiley**
3. **ArvindSathi, Big Data Analytics Disruptive Technologies For Changing The Game, MC Press LLC**
4. **Viktor Mayer Schonberger, Kenneth Cukier, Big Data,**
5. **Adam Jorgensen, James Rowland-Jones, John Welch and Dan Clark, "Microsoft Big Data Solutions", Wiley**
6. **http://www.bigdatauniversity.com**

Course Outcome:

CO 1: Learner understands the key issues in big data management and its associated applications in intelligent business and scientific computing.

CO 2: Lerner is acquiring fundamental techniques and algorithms like Hadoop, Map Reduce and NO SQL in big data analytics.

CO 3: Learner is able to interpret business models and scientific computing paradigms, and apply software tools for big data analytics.

CO 4: Learner understands adequate perspectives of big data analytics in various applications like recommender systems, social media applications etc.

Big Data Practical

B. Sc. (Data Science)		Semester – IV	
Course Name: Big Data Practical		**Course Code: USDS4P2**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1 a	Install, configure and run Hadoop and HDFS
b	Implement word count/ frequency program using MapReduce
2	Implement an Mapreduce program that process a weather dataset
3	Exploring Hadoop Distributed File System (HDFS)
4	Implement an application that store big data in Hbase/ Mongodb/ Pig using Hadoop/R
5	Implement a program in Pig
6	Configure the Hive and implement the application in Hive
7	Illustrate the working of Jaql
8 a	Implement Decision tree classification technique
b	Implement SVM Classification technique
9 a	**Regression Model:** Import a data from web storage. Name the dataset and do Logistic Regression to find out relation between variables that are affecting the admission of a student in an institute based on his or her GRE score, GPA obtained and rank of the student. Also check the model is fit or not require (foreign), require (Mass)
b	MULTIPLE REGRESSION MODEL: Apply multiple regressions, if data have a continuous independent variable. Apply on above dataset.
10 a	CLASSIFICATION MODEL: a. Install relevant package for classification. b. Choose classifier for classification problem. c. Evaluate the performance of classifier.
b	CLUSTERING MODEL a. Clustering algorithms for unsupervised classification. b. Plot the cluster data using R visualizations.

Fundamentals of Accounting

B. Sc. (Data Science)		Semester – IV	
Course Name: Fundamentals of Accounting		Course Code: USDS403	
Periods per week (1 Period is 50 minutes)		5	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2½	75
	Internal	--	25

Course Objectives

1: To be able to track a company's finances in their numerous forms, from credits, debits, and profitability to payroll and tax filing.

2: To analyzing the organization's financial health and apply data science principles/practices on that information to plot current and future strategies for growth.

Unit	Details	Lectures
I	**Accounting in Action:** Identify the activities and users associated with accounting, Explain the building blocks of accounting: ethics, principles, and assumptions, State the accounting equation, and define its components, Analyze the effects of business transactions on the accounting equation, Describe the four financial statements and how they are prepared **The Recording Process:** Describe how accounts, debits, and credits are used to record business transactions, Indicate how a journal is used in the recording process, Explain how a ledger and posting help in the recording process, Prepare a trial balance	12
II	**Adjusting the Accounts:** Explain the accrual basis of accounting and the reasons for adjusting entries, Prepare adjusting entries for deferrals, Prepare adjusting entries for accruals, Describe the nature and purpose of an adjusted trial balance **Completing the Accounting Cycle:** Prepare a worksheet, Prepare closing entries and a post-closing trial balance, Explain the steps in the accounting cycle and how to prepare correcting entries, Identify the sections of a classified balance sheet	12
III	**Accounting for Merchandising Operations:** Describe merchandising operations and inventory systems, Record purchases under a perpetual inventory system, Record sales under a perpetual inventory system, Apply the steps in the accounting cycle to a merchandising company, Compare a multiple-step with a single-step income statement **Inventories:** Discuss how to classify and determine inventory, Apply inventory cost flow methods and discuss their financial effects, Indicate	12

	the effects of inventory errors on the financial statements, Explain the statement presentation and analysis of inventory **Accounting Information Systems:** Explain the basic concepts of an accounting information system, Describe the nature and purpose of a subsidiary ledger, Record transactions in special journals	
IV	**Fraud, Internal Control, and Cash:** Discuss fraud and the principles of internal control, Apply internal control principles to cash, Identify the control features of a bank account, Explain the reporting of cash. **Accounting for Receivables:** Explain how companies recognize accounts receivable, Describe how companies value accounts receivable and record their disposition, Explain how companies recognize notes receivable, Describe how companies value notes receivable, record their disposition, and present and analyze receivables **Plant Assets, Natural Resources, and Intangible Assets:** Explain the accounting for plant asset expenditures, Apply depreciation methods to plant assets, Explain how to account for the disposal of plant assets, Describe how to account for natural resources and intangible assets, Discuss how plant assets, natural resources, and intangible assets are reported and analyzed	**12**
V	**Current Liabilities and Payroll Accounting:** Explain how to account for current liabilities, Discuss how current liabilities are reported and analyzed, Explain how to account for payroll **Accounting for Partnerships:** Discuss and account for the formation of a partnership, Explain how to account for net income or net loss of a partnership, Explain how to account for the liquidation of a partnership	**12**

Books and References:

Sr. No.	Title	Author/s	Publisher	Edition	Year
1	ACCOUNTING PRINCIPLES	Jerry J. Weygandt Paul D. Kimmel Donald E. Kieso	WILEY	12th	
2	MS-EXCEL FOR CHARTERED ACCOUNTANTS	CA SUNIL B GABHAWALLA	WIRC - ICAI	----	----

Course Outcome:

CO 1: Lerner understands the laws governing the business, typical business administration schemes, and the ethics of accountancy, statistics, and accounting theory.

CO 2: Learner understands the record keeping of financial transactions and further implementations in relevant area.

Fundamentals of Accounting Practical

B. Sc. (Data Science)		**Semester – IV**	
Course Name: Fundamentals of Accounting Practical		**Course Code: USDS4P3**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
Practical can be performed using MS Excel / Google Sheet / WPS Office	
1	**Introduction to Spreadsheet**
a	Data entry using spreadsheet : Text, Number, Formula, Function, Auto fill, Auto Correct and Data Validation
b	Using Total and Subtotal: +, sum() , Quick sum, subtotal(), sumif(), conditional sums, sorting of data
2	**Advance functions**
a	Lookup(), HLOOKUP(), VLOOKUP(), date functions, numeric functions, string functions, Index(), Match()
b	Financial Functions
3	**Using paste special**
a	To demonstrate different types of paste options available in paste special.
4	**Analysing Data**
a	Data tables
b	Scenarios
c	Goal Seek
5	**Pivot Table**
a	Creating a Pivot Table
b	Layout of the PivotTable
6	**Auditing Tools**
a	Auditing Toolbar
b	Documenting a Sheet
c	Migrating Data from Other Software
d	Common Audit Techniques
7	**Application of spreadsheet**
a	Creating Balance Sheet and Balance Sheet Summary

b	Creating Income Statement and Income Statement Summary
c	Creating Cash Flow Statement and Cash Flow Statement Summary
8	Creating a General Ledger
9	Create a Loan Amortization table using the PMT function
10	Automate common tasks using Macros
a	Using Global Macros (download and use Currency conversion, Number to Word conversion etc)
b	Creating own Macros (Income tax & GST Macros)

Artificial Intelligence

B. Sc. (Data Science)		Semester – IV	
Course Name: Artificial Intelligence		**Course Code: USDS404**	
Periods per week (1 Period is 50 minutes)		**5**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Theory Examination**	**2½**	**75**
	Internal	**--**	**25**

Course Objectives:

1. To introduce and appreciate use of AI and the theory underlying for solving problems.
2. To Learn about Rational Intelligent Agent and Agent types to solve problems
3. To learn about representing difficult real life problems as state space representation and solving them using AI techniques.
4. To understand the basic issues of knowledge representation and develop skills for reasoning and handling uncertainty
5. To introduce advanced topics of AI for solving complex problems.

Unit	Details	Lectures
I	**Intelligent Systems and Intelligent Agents:** Introduction to AI, AI Problems and AI techniques, Solving problems by searching, Problem Formulation. State Space Representation Structure of Intelligent agents, Types of Agents, Agent Environments PEAS representation for an Agent.	**12**
II	**Searching Techniques:** Uninformed Search: DFS, BFS, Uniform cost search, Depth Limited Search, Iterative Deepening. Informed Search: Heuristic functions, Hill Climbing, Simulated Annealing, Best First Search, A* Constraint Satisfaction Programming: Crypto Arithmetic, Map Coloring, N-Queens. Adversarial Search: Game Playing, Min-Max Search, Alpha Beta Pruning	**12**
III	**Knowledge and Reasoning:** Knowledge Based Agent, Overview of Propositional Logic, First Order Predicate Logic, Inference in First Order Predicate Logic: Forward and Backward Chaining, Resolution.	**12**
IV	**Uncertainity and Reasoning:** Uncertainly, Representing Knowledge in an Uncertain Domain, Bayesian Network, Conditional Probability, Joint Probability, Bayes' theorem, Belief Networks, Simple Inference in BeliefNetworks. Sequential decsion problems.	**12**
V	**Machine Learning:** Forms of Learning, Supervised Learning, Learning Decision Trees, Evaluating and Choosing the Best Hypothesis, Theory of Learning, Regression and Classification with Linear Models, Artificial Neural Networks, Nonparametric Models, Support Vector Machines, Ensemble Learning, Statistical Learning, Introduction to deep learning concepts	**12**

Books and References:					
Sr. No.	**Title**	**Author/s**	**Publisher**	**Edition**	**Year**
1	Artificial Intelligence: A Modern Approach	Stuart J. Russell and Peter Norvig	Pearson	Fourth Edition	2020
2	Artificial Intelligence: Foundations of Computational Agents	David L Poole, Alan K. Mackworth	Cambridge University Press	Second Edition	2017
3	Artificial Intelligence	Kevin Knight and Elaine Rich	McGraw Hill	3rd Edition	2017
4	The Elements of Statistical Learning	Trevor Hastie, Robert Tibshirani and Jerome Friedman	Springer		2013
5	A First Course in Artificial Intelligence	Deepak Khemani	TMH	1st Edition	2017

Course Outcome:

CO 1: Leaner understands building blocks of AI.

CO 2: Learner is analyzing problem and solving it by implementing suitable techniques.

CO 3: Learner is applying logic based techniques to solve examples**.**

CO 4: Learner is able to implement Bayesian approaches.

CO 5: Learner is using machine learning concepts for solving problems

Artificial Intelligence Practical

B. Sc. (Data Science)		**Semester – IV**	
Course Name: Artificial Intelligence Practical		**Course Code: USDS4P4**	
Periods per week (1 Period is 50 minutes)		**3**	
Credits		**2**	
		Hours	**Marks**
Evaluation System	**Practical Examination**	**2½**	**50**
	Internal	**--**	

List of Practical:	
1.	**Generate the state-space possiblities for the folling problems**
a.	Water Jug Problem
b.	Number Puzzel
2.	**Write the program to compute the following Uninformed Search Search Algorithms for suitable problem**
a.	Depth First Search
b.	Breadth First Search
3.	**Write the program to compute the following Informed Search Search Algorithms for suitable problem**
a.	Hill Climbing
b.	Simulated Annealing
c.	A* Algorithm
4.	**Write the program to compute the following Algorithms for suitable problem**
a.	Simulate solution for 4-Queen / N-Queen problem
b.	Constraint satisfaction problem: Map Coloring
5.	**Write the program to compute the following Search Algorithms for suitable problem**
a.	simulation of tic – tac – toe game using Min-Max Search
b.	Alpha Beta Pruning
c.	Water Jug Problem
6.	**Write the program to compute the following Algorithms for suitable problem**
a.	Missionaries and Cannibals
b.	Simple Inferencing
7.	**Write the program to implement decsion tree for suitable problem.**

a.	Two Class decsion
b.	Multi class decsion
8.	**Write the program to compute the following Algorithms for suitable problem**
a.	Linear Regression
b.	Classification Problems
9.	**Write the program to solve the following problems using suitable AI method.**
a.	Traveling salesman problem
b.	Number Puzzel Problem
10.	**Write the program to demonstrate the following**
a.	Simple neural network
b.	Support Vector Machine

Numerical Methods

B. Sc. (Data Science)		Semester – IV	
Course Name: Numerical Methods		Course Code: USDS405	
Periods per week (1 Period is 50 minutes)		5	
Credits		2	
		Hours	Marks
Evaluation System	Theory Examination	2½	75
	Internal	--	25

Course Objectives

1. To be able to precisely solve problems using mathematical modeling.
2. To be able to find solution for a solvable to unsolvable problems.
3. To find an answer or solution close to answer, without even knowing what the answer is.

Unit	Details	Lectures
I	**Background and Introduction:** Differential Equations, Matrix Analysis, Matrix Eigenvalue Problem, Errors and Approximations, Iterative Methods **Numerical Solution of Equations of a Single Variable:** Numerical Solution of Equations, Bisection Method, Regula Falsi Method, Fixed-Point Method, Newton's Method, Secant Method, Equations with Several Roots	12
II	**Numerical Solution of Systems of Equations:** Linear Systems of Equations, Numerical Solution of Linear Systems, Gauss Elimination Method, LU Factorization Methods, Iterative Solution of Linear Systems, Ill-Conditioning and Error Analysis, Systems of Nonlinear Equations	12
III	**Curve Fitting and Interpolation:** Least-Squares Regression, Linear Regression, Linearization of Nonlinear Data, Polynomial Regression, Polynomial Interpolation, Spline Interpolation, Fourier Approximation and Interpolation **Numerical Differentiation and Integration:** Numerical Differentiation, Finite-Difference Formulas for Numerical Differentiation, Numerical Integration: Newton–Cotes Formulas, Numerical Integration of Analytical Functions: Romberg Integration, Gaussian Quadrature, Improper Integrals	12
IV	**Numerical Solution of Initial-Value Problems:** Introduction, One-Step Methods, Euler's Method, Runge–Kutta Methods, Multistep Methods, Systems of Ordinary Differential Equations, Stability, Stiff Differential Equations	12

V	**Numerical Solution of Boundary-Value Problems:** Second-Order BVP, Boundary Conditions, Higher-Order BVP, Shooting Method, Finite-Difference Method **Matrix Eigenvalue Problem:** Matrix Eigenvalue Problem, Power Method: Estimation of the Dominant Eigenvalue, Inverse Power Method: Estimation of the Smallest Eigenvalue, Shifted Inverse Power Method: Estimation of the Eigenvalue Nearest a Specified Value, Shifted Power Method, Deflation Methods, Householder Tridiagonalization and QR Factorization Methods, A Note on the Terminating Condition Used in HouseholderQR, Transformation to Hessenberg Form (Nonsymmetric Matrices)	**12**

Books and References:					
Sr. No.	**Title**	**Author/s**	**Publisher**	**Edition**	**Year**
1. 1	**Numerical Methods for Engineers and Scientists Using MATLAB**	**Ramin S. Esfandiari**	CRC Press	**Second Edition**	2017

Course Outcome:

CO 1: Learner implementing Numerical Methods to solve the problems.

CO 2: Learner is computing the numerical results using raw data.

Numerical Methods Practical

B. Sc. (Data Science)		Semester – IV	
Course Name: Numerical Methods Practical		Course Code: USDS4P5	
Periods per week (1 Period is 50 minutes)		3	
Credits		2	
		Hours	Marks
Evaluation System	Practical Examination	2½	50
	Internal	--	

List of Practical:	
1	**Write a program using R/Python/Scilab/Matlab**
a	Bisection Method
b	Regula Falsi Method
c	Newton's Method
2	**Write a program using R/Python/Scilab/Matlab**
a	Gauss Elimination Method
b	LU Factorization Methods
3	**Write a program using R/Python/Scilab/Matlab**
a	Numerical Differentiation
b	Newton–Cotes Formulas
4	**Write a program using R/Python/Scilab/Matlab**
a	Romberg Integration
b	Gaussian Quadrature
5	**Write a program using R/Python/Scilab/Matlab**
a	Euler's Method
b	Runge–Kutta Methods
6	**Write a program using R/Python/Scilab/Matlab**
a	Second-Order BVP
b	Higher-Order BVP
7	**Write a program using R/Python/Scilab/Matlab for Finite difference method**
8	**Write a program using R/Python/Scilab/Matlab for Estimation of the Dominant Eigenvalue**
9	**Write a program using R/Python/Scilab/Matlab for Estimation of the Smallest Eigenvalue**
10	**Write a program using R/Python/Scilab/Matlab for Estimation of the Eigenvalue Nearest a Specified Value**