

II Year B.Tech. CSE (DS) – II Semester

L	T	P	C
2	0	0	2

(23A52402e) OPTIMIZATION TECHNIQUES

Course Objectives: The objectives of the course are to

- To provide the basic knowledge about Optimization, importance, application areas of in the industry, Linear Programming.
- To impart different optimization models under typical situations in the business organization like transportation, assignment.
- To understand the process of sequencing in a typical industry.
- To describe different game strategies under cut-throat competitive business environment
- To develop networks of activities of projects and to find out optimal modes of completing projects using network modelling evaluation techniques.

Course Outcomes:

COs	Statements	Blooms Level
CO1	Understanding Optimization and Formulation of Linear Programming Models	L1
CO2	Formulate and Solve Transportation & Assignment Models	L3
CO3	Sequencing of operations and optimizing	L2
CO4	Discuss the game theory and strategies	L2
CO5	Developing networks of activities and finding optimal mode of projects evaluation.	L3

UNIT - I

Introduction: Meaning, Nature, Scope & Significance of Optimization - Typical applications. The Linear Programming Problem – Introduction, Formulation of Linear Programming problem, Limitations of L.P.P, Graphical method, Simplex method: Maximization and Minimization model(exclude Duality problems), Big-M method and Two Phase method.

UNIT - II

Transportation Problem: Introduction, Transportation Model, Finding initial basic feasible solutions, Moving towards optimality, Unbalanced Transportation problems, Transportation problems with maximization, Degeneracy.

Assignment Problem – Introduction, Mathematical formulation of the problem, Solution of an Assignment problem, Hungarian Algorithm, Multiple Solution, Unbalanced Assignment problems, Maximization in Assignment Model.

UNIT - III

Sequencing – Job sequencing, Johnsons Algorithm for n Jobs and Two machines, n Jobs and Three Machines, n jobs through m machines, Two jobs and m Machines Problems.

UNIT - IV

Game Theory: Concepts, Definitions and Terminology, Two Person Zero Sum Games, Pure Strategy Games (with Saddle Point), Principal of Dominance, Mixed Strategy Games (Game without Saddle Point), Significance of Game Theory in Managerial Application.

UNIT - V

Project Management: Network Analysis – Definition –objectives -Rules for constructing network diagram- Determining Critical Path – Earliest & Latest Times – Floats - Application

of CPM and PERT techniques in Project Planning and Control – PERT Vs CPM. (exclude Project Crashing).

Textbooks:

1. Operations Research / R.Pannerselvam, PHI Publications.
2. Operations Research / S.D.Sharma-Kedarnath
3. Operations Research /A.M.Natarajan,P.Balasubramani,A. Tamilarasi/Pearson Education.
4. Engineering Optimization: Theory and practice / S.S.Rao, New Age International (P) Limited

Reference Books:

1. Quantitative Techniques in Management / ND Vohra, Tata McGraw Hill, 4th Edition, 2011.
2. ntroduction to O.R/Hiller &Libermann (TMH).
3. Operations Research: Methods & Problems / Maurice Saseini, ArhurYaspan& Lawrence Friedman. Pearson
4. Quantitative Analysis For Management/ Barry Render, Ralph M. Stair, Jr and Michael E. Hanna/
5. Operations Research / Wagner/ PHI Publications.

Online Learning Sources

https://onlinecourses.swayam2.ac.in/cec20_ma10/preview

https://onlinecourses.nptel.ac.in/noc20_ma23/preview

https://onlinecourses.nptel.ac.in/noc19_ma29/preview

II Year B.Tech. CSE (DS) – II Semester

L	T	P	C
3	0	0	3

(23A54402) STATISTICAL METHODS FOR DATA SCIENCE

(Common to CSE (Data Science), AI & DS, CSE (AIDS))

Course Outcomes: After successful completion of this course, the students should be able to:

COs	Statements	Blooms level
CO1	Understand the basic concepts of Statistics.	L2, L3
CO2	Analyze the data and draw conclusion about collection of data under study using Point estimation	L3, L5
CO3	Analyze data and draw conclusion about collection of data under study using Interval estimation.	L3
CO4	Analyze to test various hypotheses included in theory and types of errors for large samples.	L2, L3
CO5	Apply the different testing tools like t-test, F-test, chi-square test to analyze the relevant real life problems.	L3, L5

UNIT I : Basic Concepts

Random variables (discrete and continuous), probability density functions, properties, mathematical expectation. Probability distributions: Binomial, Poisson and Normal-their properties. Population, sample, parameter and statistic; characteristics of a good estimator; Consistency – Invariance property of Consistent estimator, Sufficient condition for consistency; Unbiasedness; Sufficiency.

UNIT II Point Estimation

Point Estimation- Estimator, Estimate, Methods of point estimation – Maximum likelihood method (the asymptotic properties of ML estimators are not included), Large sample properties of ML estimator (without proof)- applications, Method of moments, method of least squares, method of minimum chi-square and modified minimum chi-square-Asymptotic Maximum Likelihood Estimation and applications.

UNIT III Interval Estimation

Confidence limits and confidence coefficient; Duality between acceptance region of a test and a confidence interval; Construction of confidence intervals for population proportion (small and large samples) and between two population proportions(large samples); Confidence intervals for mean and variance of a normal population; Difference between the mean and ratio of two normal populations.

UNIT IV Testing of hypotheses

Types of errors, power of a test, most powerful tests; Neyman-Pearson Fundamental Lemma and its applications; Notion of Uniformly most powerful tests; Likelihood Ratio tests: Description and property of LR tests - Application to standard distributions.

UNIT V Small sample tests

Student's t-test, test for a population mean, equality of two population means, paired t-test, F-test for equality of two population variances, Chi-square test for goodness of fit and test for independence of attributes, χ^2 test for testing variance of a normal distribution.

Textbooks:

1. Miller and Freunds, Probability and Statistics for Engineers,7/e, Pearson, 2008.
2. Manoj Kumar Srivastava and Namita Srivastava, Statistical Inference – Testing of Hypotheses, Prentice Hall of India, 2014

Reference Books:

1. S.C. Gupta and V.K. Kapoor, Fundamentals of Mathematical Statistics, 11/e, Sultan Chand & Sons Publications, 2012.
2. S. Ross, a First Course in Probability, Pearson Education India, 2002.
3. W. Feller, an Introduction to Probability Theory and its Applications, 1/e, Wiley, 1968.
4. Robert V Hogg, Elliot A Tannis and Dale L.Zimmerman, Probability and Statistical Inference, 9th edition, Pearson publishers,2013.

Online Learning Resources:

1. https://onlinecourses.nptel.ac.in/noc21_ma74/preview
2. https://onlinecourses.nptel.ac.in/noc22_mg31/preview

II Year B.Tech. CSE (DS) – II Semester

L	T	P	C
3	0	0	3

(23A32402T) DATA ENGINEERING**Course Objectives:**

- Explain basic concepts of Data Engineering
- Discuss bout Data Engineering Life Cycle
- How to design Good Data Architecture

Course Outcomes: By the end of the course students will be able to:

- Understand Data Engineering Life cycle
- Apply appropriate data modeling techniques for different types of data. (L3)
- Evaluate and select appropriate technologies and frameworks for specific data engineering tasks. (L5)
- Implement data quality checks and governance processes to ensure data reliability and compliance. (L5)

UNIT-I: Introduction to Data Engineering: Definition, Data Engineering Life Cycle, Evolution of Data Engineer, Data Engineering Versus Data Science, Data Engineering Skills and Activities,

Data Maturity, Data Maturity Model, Skills of a Data Engineer, Business Responsibilities, Technical Responsibilities, Data Engineers and Other Technical Roles.

UNIT-II: Data Engineering Life Cycle: Data Life Cycle Versus Data Engineering Life Cycle, Generation: Source System, Storage, Ingestion, Transformation, Serving Data.

Major undercurrents across the Data Engineering Life Cycle: Security, Data Management, Data Ops, Data Architecture, Orchestration, Software Engineering.

UNIT-III: Designing Good Data Architecture: Enterprise Architecture, Data Architecture, Principles of Good Data Architecture, Major Architecture Concepts.

Data Generation in Source Systems: Sources of Data, Files and Unstructured Data, APIs, Application Databases (OLTP), OLAP, Change Data Capture, Logs, Database Logs, CRUD, Source System Practical Details.

UNIT-IV: Storage: Raw Ingredients of Data Storage, Data Storage Systems, Data Engineering Storage Abstractions, Data warehouse, Data Lake, Data Lakehouse.

Ingestion: Data Ingestion, Key Engineering considerations for the Ingestion Phase, Batch Ingestion Considerations, Message and Stream Ingestion Considerations, Ways to Ingest Data

UNIT-V: Queries, Modeling and Transformation: Queries, Life of a Query, Query Optimizer, Queries on Streaming Data, Data Modelling, Modeling Streaming Data, Transformations, Streaming Transformations and Processing.

Serving Data for Analytics, Machine Learning and Reverse ETL: General Considerations for serving Data, Business Analytics, Operational Analytics, Embedded Analytics, Ways to serve data for analytics and ML, Reverse ETL.

Textbooks:

1. Joe Reis, Matt Housley, Fundamentals of Data Engineering, O'Reilly Media, Inc., June 2022, ISBN: 9781098108304

Reference Books:

1. Paul Crickard , Data Engineering with Python,Packt Publishing, October 2020.
2. Ralph Kimball, Margy Ross, The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Wiley, 3rd Edition, 2013
3. James Densmore, Data Pipelines Pocket Reference: Moving and Processing Data for Analytics, O'Reilly Media, 1st Edition, 2021

II Year B.Tech. CSE (DS)– II Semester

L	T	P	C
3	0	0	3

(23A05402T) DATABASE MANAGEMENT SYSTEMS

Course Objectives: The main objectives of the course is to

- Introduce database management systems and to give a good formal foundation on the relational model of data and usage of Relational Algebra
- Introduce the concepts of basic SQL as a universal Database language
- Demonstrate the principles behind systematic database design approaches by covering conceptual design, logical design through normalization
- Provide an overview of physical design of a database system, by discussing Database indexing techniques and storage techniques

Course Outcomes: After completion of the course, students will be able to

- Understand the basic concepts of database management systems (L2)
- Analyze a given database application scenario to use ER model for conceptual design of the database (L4)
- Utilize SQL proficiently to address diverse query challenges (L3).
- Employ normalization methods to enhance database structure (L3)
- Assess and implement transaction processing, concurrency control and database recovery protocols in databases. (L4)

UNIT I: Introduction: Database system, Characteristics (Database Vs File System), Database Users, Advantages of Database systems, Database applications. Brief introduction of different Data Models; Concepts of Schema, Instance and data independence; Three tier schema architecture for data independence; Database system structure, environment, Centralized and Client Server architecture for the database.

Entity Relationship Model: Introduction, Representation of entities, attributes, entity set, relationship, relationship set, constraints, sub classes, super class, inheritance, specialization, generalization using ER Diagrams.

Unit II: Relational Model: Introduction to relational model, concepts of domain, attribute, tuple, relation, importance of null values, constraints (Domain, Key constraints, integrity constraints) and their importance, Relational Algebra, Relational Calculus. BASIC SQL:Simple Database schema, data types, table definitions (create, alter), different DML operations (insert, delete, update).

UNIT III: SQL: Basic SQL querying (select and project) using where clause, arithmetic & logical operations, SQL functions(Date and Time, Numeric, String conversion).Creating tables with relationship, implementation of key and integrity constraints, nested queries, sub queries, grouping, aggregation, ordering, implementation of different types of joins, view(updatable and non-updatable), relational set operations.

UNIT IV: Schema Refinement (Normalization):Purpose of Normalization or schema refinement, concept of functional dependency, normal forms based on functional dependency Lossless join and dependency preserving decomposition, (1NF, 2NF and 3 NF), concept of surrogate key, Boyce-Codd normal form(BCNF), MVD, Fourth normal form(4NF), Fifth Normal Form (5NF).

UNIT V: Transaction Concept: Transaction State, ACID properties, Concurrent Executions, Serializability, Recoverability, Implementation of Isolation, Testing for

Serializability, lock based, time stamp based, optimistic, concurrency protocols, Deadlocks, Failure Classification, Storage, Recovery and Atomicity, Recovery algorithm.

Introduction to Indexing Techniques: B+ Trees, operations on B+Trees, Hash Based Indexing:

Textbooks:

1. Database Management Systems, 3rd edition, Raghurama Krishnan, Johannes Gehrke, TMH (For Chapters 2, 3, 4)
2. Database System Concepts, 5th edition, Silberschatz, Korth, Sudarsan, TMH (For Chapter 1 and Chapter 5)

Reference Books:

1. Introduction to Database Systems, 8th edition, C J Date, Pearson.
2. Database Management System, 6th edition, Ramez Elmasri, Shamkant B. Navathe, Pearson
3. Database Principles Fundamentals of Design Implementation and Management, Corlos Coronel, Steven Morris, Peter Robb, Cengage Learning.

Web-Resources:

1. <https://nptel.ac.in/courses/106/105/106105175/>
https://infyspringboard.onwingspan.com/web/en/app/toc/lex_auth_01275806667282022456_shared/overview

II Year B.Tech. CSE (DS)– II Semester

L	T	P	C
3	0	0	3

(23A30402) DIGITAL LOGIC & COMPUTER ORGANIZATION

Course Objectives: The main objectives of the course is to

- provide students with a comprehensive understanding of digital logic design principles and computer organization fundamentals
- Describe memory hierarchy concepts
- Explain input/output (I/O) systems and their interaction with the CPU, memory, and peripheral devices

Course Outcomes: After completion of the course, students will be able to

- Differentiate between combinational and sequential circuits based on their characteristics and functionalities. (L2)
- Demonstrate an understanding of computer functional units. (L2)
- Analyze the design and operation of processors, including instruction execution, pipelining, and control unit mechanisms, to comprehend their role in computer systems.(L3)
- Describe memory hierarchy concepts, including cache memory, virtual memory, and secondary storage, and evaluate their impact on system performance and scalability. (L3)
- Explain input/output (I/O) systems and their interaction with the CPU, memory, and peripheral devices, including interrupts, DMA, and I/O mapping techniques. (L3)
- Design Sequential and Combinational Circuits (L6)

UNIT – I:

Data Representation: Binary Numbers, Fixed Point Representation. Floating Point Representation. Number base conversions, Octal and Hexadecimal Numbers, components, Signed binary numbers, Binary codes

Digital Logic Circuits-I: Basic Logic Functions, Logic gates, universal logic gates, Minimization of Logic expressions. K-Map Simplification, Combinational Circuits, Decoders, Multiplexers

UNIT – II:

Digital Logic Circuits-II: Sequential Circuits, Flip-Flops, Binary counters, Registers, Shift Registers, Ripple counters

Basic Structure of Computers: Computer Types, Functional units, Basic operational concepts, Bus structures, Software, Performance, multiprocessors and multi computers, Computer Generations, Von- Neumann Architecture

UNIT – III:

Computer Arithmetic : Addition and Subtraction of Signed Numbers, Design of Fast Adders, Multiplication of Positive Numbers, Signed-operand Multiplication, Fast Multiplication, Integer Division, Floating-Point Numbers and Operations

Processor Organization: Fundamental Concepts, Execution of a Complete Instruction, Multiple-Bus Organization, Hardwired Control and Multi programmed Control

UNIT – IV:

The Memory Organization: Basic Concepts, Semiconductor RAM Memories, Read-Only Memories, Speed, Size and Cost, Cache Memories, Performance Considerations, Virtual Memories, Memory Management Requirements, Secondary Storage

UNIT – V:

Input /Output Organization: Accessing I/O Devices, Interrupts, Processor Examples, Direct Memory Access, Buses, Interface Circuits, Standard I/O Interfaces

Textbooks:

1. Computer Organization, Carl Hamacher, ZvonkoVranesic, SafwatZaky, 6th edition, McGraw Hill, 2023.
2. Digital Design, 6th Edition, M. Morris Mano, Pearson Education, 2018.
3. Computer Organization and Architecture, William Stallings, 11thEdition, Pearson, 2022.

Reference Books:

1. Computer Systems Architecture, M.Moris Mano, 3rdEdition, Pearson, 2017.
2. Computer Organization and Design, David A. Paterson, John L. Hennessy, Elsevier, 2004.
3. Fundamentals of Logic Design, Roth, 5thEdition, Thomson, 2003.

Online Learning Resources:

<https://nptel.ac.in/courses/106/103/106103068/>

II Year B.Tech. CSE (DS)– II Semester

L	T	P	C
0	0	3	1.5

(23A32402P) DATA ENGINEERING LAB**Course Objective:**

- The main objective of this course is to teach how build data engineering infrastructure and data pipelines.

Course Outcomes: After completion of the course, students will be able to

- Analyze and troubleshoot data engineering problems using systematic approaches. (L5)
- Work effectively in teams to solve data engineering challenges and deliver projects on time. (L6)

Experiments:

- Installing and configuring Apache NiFi, Apache Airflow
- Installing and configuring Elasticsearch, Kibana, PostgreSQL, pgAdmin 4
- Reading and Writing files
 - Reading and writing files in Python
 - Processing files in Airflow
 - NiFi processors for handling files
 - Reading and writing data to databases in Python
 - Databases in Airflow
 - Database processors in NiFi
- Working with Databases
 - Inserting and extracting relational data in Python
 - Inserting and extracting NoSQL database data in Python
 - Building database pipelines in Airflow
 - Building database pipelines in NiFi
- Cleaning, Transforming and Enriching Data
 - Performing exploratory data analysis in Python
 - Handling common data issues using pandas
 - Cleaning data using Airflow
- Building the Data Pipeline
- Building a Kibana Dash Board
- Perform the following operations
 - Staging and validating data
 - Building idempotent data pipelines
 - Building atomic data pipelines
- Version Control with the NiFi Registry
 - Installing and configuring the NiFi Registry
 - Using the Registry in NiFi
 - Versioning your data pipelines
 - Using git-persistence with the NiFi Registry
- Monitoring Data Pipelines
 - Monitoring NiFi in the GUI
 - Monitoring NiFi using processors
 - Monitoring NiFi with Python and the REST API
- Deploying Data Pipelines
 - Finalizing your data pipelines for production
 - Using the NiFi variable registry

- c. Deploying your data pipelines
- 12. Building a Production Data Pipeline
 - a. Creating a test and production environment
 - b. Building a production data pipeline
 - c. Deploying a data pipeline in production

Reference Books:

- 1. Paul Crickard , Data Engineering with Python,Packt Publishing, October 2020.

II Year B.Tech. CSE (DS) – II Semester

L	T	P	C
0	0	3	1.5

(23A05402P) DATABASE MANAGEMENT SYSTEMS LAB

Course Objectives: This Course will enable students to

- Populate and query a database using SQL DDL/DML Commands
- Declare and enforce integrity constraints on a database
- Writing Queries using advanced concepts of SQL
- Programming PL/SQL including procedures, functions, cursors and triggers.

Course Outcomes: After completion of the course, students will be able to

- Utilizing Data Definition Language (DDL), Data Manipulation Language (DML), and Data Control Language (DCL) commands effectively within a database environment (L3)
- Constructing and execute queries to manipulate and retrieve data from databases. (L3)
- Develop application programs using PL/SQL. (L3)
- Analyze requirements and design custom Procedures, Functions, Cursors, and Triggers, leveraging their capabilities to automate tasks and optimize database functionality (L4)
- Establish database connectivity through JDBC (Java Database Connectivity) (L3)

Experiments covering the topics:

- DDL, DML, DCL commands
- Queries, nested queries, built-in functions,
- PL/SQL programming- control structures
- Procedures, Functions, Cursors, Triggers,
- Database connectivity- ODBC/JDBC

Sample Experiments:

1. Creation, altering and dropping of tables and inserting rows into a table (use constraints while creating tables) examples using SELECT command.
2. Queries (along with sub Queries) using ANY, ALL, IN, EXISTS, NOTEXISTS, UNION, INTERSET, Constraints. Example:- Select the roll number and name of the student who secured fourth rank in the class.
3. Queries using Aggregate functions (COUNT, SUM, AVG, MAX and MIN), GROUP BY, HAVING and Creation and dropping of Views.
4. Queries using Conversion functions (to_char, to_number and to_date), string functions (Concatenation, lpad, rpad, ltrim, rtrim, lower, upper, initcap, length, substr and instr), date functions (Sysdate, next_day, add_months, last_day, months_between, least, greatest, trunc, round, to_char, to_date)
5.
 - i. Create a simple PL/SQL program which includes declaration section, executable section and exception –Handling section (Ex. Student marks can be selected from the table and printed for those who secured first class and an exception can be raised if no records were found)
 - ii. Insert data into student table and use COMMIT, ROLLBACK and SAVEPOINT in PL/SQL block.
6. Develop a program that includes the features NESTED IF, CASE and CASE expression. The program can be extended using the NULLIF and COALESCE functions.

7. Program development using WHILE LOOPS, numeric FOR LOOPS, nested loops using ERROR Handling, BUILT –IN Exceptions, USE defined Exceptions, RAISE-APPLICATION ERROR.
8. Programs development using creation of procedures, passing parameters IN and OUT of PROCEDURES.
9. Program development using creation of stored functions, invoke functions in SQL Statements and write complex functions.
10. Develop programs using features parameters in a CURSOR, FOR UPDATE CURSOR, WHERE CURRENT of clause and CURSOR variables.
11. Develop Programs using BEFORE and AFTER Triggers, Row and Statement Triggers and INSTEAD OF Triggers
12. Create a table and perform the search operation on table using indexing and non-indexing techniques.
13. Write a Java program that connects to a database using JDBC
14. Write a Java program to connect to a database using JDBC and insert values into it
15. Write a Java program to connect to a database using JDBC and delete values from it

Text Books/Suggested Reading:

1. Oracle: The Complete Reference by Oracle Press
2. Nilesh Shah, "Database Systems Using Oracle", PHI, 2007
3. Rick F Vander Lans, "Introduction to SQL", Fourth Edition, Pearson Education, 2007

II Year B.Tech. CSE (DS)– II Semester

L	T	P	C
0	1	2	2

(23A52403) EXPLORATORY DATA ANALYSIS WITH PYTHON**Course Objectives:**

- This course introduces the fundamentals of Exploratory Data Analysis
- It covers essential exploratory techniques for understanding multivariate data by summarizing it through statistical methods and graphical methods.

Course Outcomes: At the end of this course, the students will be able to:

CO1: Understand the fundamentals of exploratory data analysis. (L2)

CO2: Implement the data visualization using Matplotlib. (L5)

CO3: Perform univariate data exploration and analysis.

CO4: Apply bivariate data exploration and analysis. (L3)

CO5: Use Data exploration and visualization techniques for multivariate and time series data (L3)

UNIT-I

Exploratory Data Analysis Fundamentals: Understanding data science, The significance of EDA, Steps in EDA, Making sense of data, Numerical data, Categorical data, Measurement scales, Comparing EDA with classical and Bayesian analysis, Software tools available for EDA, Getting started with EDA.

Sample Experiments:

1. a) Download Dataset from Kaggle using the following link :
<https://www.kaggle.com/datasets/sukhmanibedi/cars4u>
 b) Install python libraries required for Exploratory Data Analysis (numpy, pandas, matplotlib, seaborn)
2. Perform Numpy Array basic operations and Explore Numpy Built-in functions.
3. Loading Dataset into pandas dataframe
4. Selecting rows and columns in the dataframe

UNIT-II

Visual Aids for EDA: Technical requirements, Line chart, Bar charts, Scatter plot using seaborn, Polar chart, Histogram, Choosing the best chart

Case Study: EDA with Personal Email, Technical requirements, Loading the dataset, Data transformation, Data cleansing, Applying descriptive statistics, Data refactoring, Data analysis.

Sample Experiments:

5. Apply different visualization techniques using sample dataset
 a) Line Chart b) Bar Chart c) Scatter Plots d) Bubble Plot
6. Generate Scatter Plot using seaborn library for iris dataset
7. Apply following visualization Techniques for a sample dataset
 a) Area Plot b) Stacked Plot c) Pie chart d) Table Chart
8. Generate the following charts for a dataset.
 a) Polar Chart b) Histogram c) Lollipop chart
9. Case Study: Perform Exploratory Data Analysis with Personal Email Data

UNIT-III

Data Transformation: Merging database-style dataframes, Concatenating along with an axis, Merging on index, Reshaping and pivoting, Transformation techniques, Handling missing data, Mathematical operations with NaN, Filling missing values, Discretization and binning, Outlier detection and filtering, Permutation and random sampling, Benefits of data transformation, Challenges.

Sample Experiments:

10. Perform the following operations

- a) Merging Dataframes
- b) Reshaping with Hierarchical Indexing
- c) Data Deduplication
- d) Replacing Values

11. Apply different Missing Data handling techniques

a)NaN values in mathematical Operations

- b) Filling in missing data
- c) Forward and Backward filling of missing values
- d) Filling with index values
- e) Interpolation of missing values

12. Apply different data transformation techniques

a) Renaming axis indexes
 b) Discretization and Binning
 c) Permutation and Random Sampling
 d) Dummy variables

UNIT-IV

Descriptive Statistics: Distribution function, Measures of central tendency, Measures of dispersion, Types of kurtosis, Calculating percentiles, Quartiles, Grouping Datasets, Correlation, Understanding univariate, bivariate, multivariate analysis, Time Series Analysis

Sample Experiments:

13. Study the following Distribution Techniques on a sample data

- a) Uniform Distribution
- b) Normal Distribution
- c) Gamma Distribution
- d) Exponential Distribution
 - e) Poisson Distribution
 - f) Binomial Distribution

14. Perform Data Cleaning on a sample dataset.

15. Compute measure of Central Tendency on a sample dataset

- a) Mean
- b) Median
- c) Mode

16. Explore Measures of Dispersion on a sample dataset

- a) Variance
- b) Standard Deviation
- c) Skewness
- d) Kurtosis

17. a) Calculating percentiles on sample dataset

- b) Calculate Inter Quartile Range(IQR) and Visualize using Box Plots

18. Perform the following analysis on automobile dataset.

- a) Bivariate analysis
- b) Multivariate analysis

19. Perform Time Series Analysis on Open Power systems dataset

UNIT-V

Model Development and Evaluation: Unified machine learning workflow, Data preprocessing, Data preparation, Training sets and corpus creation, Model creation and training, Model evaluation, Best model selection and evaluation, Model deployment

Case Study: EDA on Wine Quality Data Analysis

Sample Experiments:

20. Perform hypothesis testing using statsmodels library
 - a) Z-Test
 - b)T-Test
21. Develop model and Perform Model Evaluation using different metrics such as prediction score, R2 Score, MAE Score, MSE Score.
22. Case Study: Perform Exploratory Data Analysis with Wine Quality Dataset

Textbook:

1. Suresh Kumar Mukhiya, Usman Ahmed, Hands-On Exploratory Data Analysis with Python, Packt Publishing, 2020.

References:

1. Ronald K. Pearson, Exploratory Data Analysis Using R, CRC Press, 2020
2. RadhikaDatar,HarishGarg, Hands-On Exploratory Data Analysis with R: Become an expert in exploratory data analysis using R packages, Ist Edition, Packt Publishing, 2019

Web References:

1. <https://github.com/PacktPublishing/Hands-on-Exploratory-Data-Analysis-with-Python>
2. <https://www.analyticsvidhya.com/blog/2022/07/step-by-step-exploratory-data-analysis-eda-using-python/#h-conclusion>
3. <https://github.com/PacktPublishing/Exploratory-Data-Analysis-with-Python-Cookbook>

II Year B.Tech. CSE (DS) – II Semester

L	T	P	C
1	0	2	2

(23A99401) DESIGN THINKING & INNOVATION**Course Objectives:**

The objective of this course is to familiarize students with design thinking process as a tool for breakthrough innovation. It aims to equip students with design thinking skills and ignite the minds to create innovative ideas, develop solutions for real-time problems.

Course Outcomes:

- Define the concepts related to design thinking. (L1, L2)
- Explain the fundamentals of Design Thinking and innovation (L1, L2)
- Apply the design thinking techniques for solving problems in various sectors. (L3)
- Analyse to work in a multidisciplinary environment (L4)
- Evaluate the value of creativity (L5)
- Formulate specific problem statements of real time issues (L3, L6)

UNIT I Introduction to Design Thinking

Introduction to elements and principles of Design, basics of design-dot, line, shape, form as fundamental design components. Principles of design. Introduction to design thinking, history of Design Thinking, New materials in Industry.

UNIT II Design Thinking Process

Design thinking process (empathize, analyze, idea & prototype), implementing the process in driving inventions, design thinking in social innovations. Tools of design thinking - person, costumer, journey map, brainstorming, product development

Activity: Every student presents their idea in three minutes, Every student can present design process in the form of flow diagram or flow chart etc. Every student should explain about product development.

UNIT III Innovation

Art of innovation, Difference between innovation and creativity, role of creativity and innovation in organizations- Creativity to Innovation- Teams for innovation- Measuring the impact and value of creativity.

Activity: Debate on innovation and creativity, Flow and planning from idea to innovation, Debate on value-based innovation.

UNIT IV Product Design

Problem formation, introduction to product design, Product strategies, Product value, Product planning, product specifications- Innovation towards product design- Case studies

Activity: Importance of modelling, how to set specifications, Explaining their own product design.

UNIT V Design Thinking in Business Processes

Design Thinking applied in Business & Strategic Innovation, Design Thinking principles that redefine business – Business challenges: Growth, Predictability, Change, Maintaining Relevance, Extreme competition, Standardization. Design thinking to meet corporate needs- Design thinking for Startups- Defining and testing Business Models and Business Cases- Developing & testing prototypes.

Activity: How to market our own product, About maintenance, Reliability and plan for startup.

Textbooks:

1. Tim Brown, Change by design, Harper Bollins (2009)
2. Idris Mootee, Design Thinking for Strategic Innovation, 2013, John Wiley & Sons.

Reference Books:

1. David Lee, Design Thinking in the Classroom, Ulysses press
2. Shruti N Shetty, Design the Future, Norton Press
3. William Lidwell, Universal Principles of Design- Kritinaholden, Jill Butter.
4. Chesbrough.H, The Era of Open Innovation – 2013

Online Learning Resources:

<https://nptel.ac.in/courses/110/106/110106124/>
<https://nptel.ac.in/courses/109/104/109104109/>
https://swayam.gov.in/ndl_noc19_mg60/preview

COMMUNITY SERVICE PROJECT

.....Experiential learning through community engagement

Introduction

- Community Service Project is an experiential learning strategy that integrates meaningful community service with instruction, participation, learning and community development.
- Community Service Project involves students in community development and service activities and applies the experience to personal and academic development.
- Community Service Project is meant to link the community with the college for mutual benefit. The community will benefit with the focused contribution of the college students for the village/ local development. The college finds an opportunity to develop social sensibility and responsibility among students and emerge as a socially responsible institution.

Objective

Community Service Project should be an integral part of the curriculum, as an alternative to the 2 months of Summer Internships / Apprenticeships / On the Job Training, whenever there is an exigency when students cannot pursue their summer internships. The specific objectives are;

- To sensitize the students to the living conditions of the people who are around them,
- To help students to realize the stark realities of society.
- To bring about an attitudinal change in the students and help them to develop societal consciousness, sensibility, responsibility and accountability
- To make students aware of their inner strength and help them to find new /out of box solutions to social problems.
- To make students socially responsible citizens who are sensitive to the needs of the disadvantaged sections.
- To help students to initiate developmental activities in the community in coordination with public and government authorities.
- To develop a holistic life perspective among the students by making them study culture, traditions, habits, lifestyles, resource utilization, wastages and its management, social problems, public administration system and the roles and responsibilities of different persons across different social systems.

Implementation of Community Service Project

- Every student should put in 6 weeks for the Community Service Project during the summer vacation.
- Each class/section should be assigned with a mentor.
- Specific Departments could concentrate on their major areas of concern. For example, Dept. of Computer Science can take up activities related to Computer Literacy to different sections of people like - youth, women, housewives, etc
- A logbook must be maintained by each of the students, where the activities undertaken/involved to be recorded.
- The logbook has to be countersigned by the concerned mentor/faculty in charge.
- An evaluation to be done based on the active participation of the student and grade could be awarded by the mentor/faculty member.
- The final evaluation to be reflected in the grade memo of the student.

- The Community Service Project should be different from the regular programs of NSS/NCC/Green Corps/Red Ribbon Club, etc.
- Minor project reports should be submitted by each student. An internal Viva shall also be conducted by a committee constituted by the principal of the college.
- Award of marks shall be made as per the guidelines of Internship/apprentice/ on the job training.

Procedure

- A group of students or even a single student could be assigned for a particular habitation or village or municipal ward, as far as possible, in the near vicinity of their place of stay, to enable them to commute from their residence and return back by evening or so.
- The Community Service Project is a twofold one –
 - First, the student/s could conduct a survey of the habitation, if necessary, in terms of their own domain or subject area. Or it can even be a general survey, incorporating all the different areas. A common survey format could be designed. This should not be viewed as a duplication of work by the Village or Ward volunteers, rather, it could be another primary source of data.
 - Secondly, the student/s could take up a social activity, concerning their domain or subject area. The different areas, could be like –
 - Agriculture
 - Health
 - Marketing and Cooperation
 - Animal Husbandry
 - Horticulture
 - Fisheries
 - Sericulture
 - Revenue and Survey
 - Natural Disaster Management
 - Irrigation
 - Law & Order
 - Excise and Prohibition
 - Mines and Geology
 - Energy
 - Internet
 - Free Electricity
 - Drinking Water

EXPECTED OUTCOMES

BENEFITS OF COMMUNITY SERVICE PROJECT TO STUDENTS

Learning Outcomes

- Positive impact on students' academic learning
- Improves students' ability to apply what they have learned in "the real world"
- Positive impact on academic outcomes such as demonstrated complexity of understanding, problem analysis, problem-solving, critical thinking, and cognitive development.
- Improved ability to understand complexity and ambiguity

Personal Outcomes

- Greater sense of personal efficacy, personal identity, spiritual growth, and moral development
- Greater interpersonal development, particularly the ability to work well with others, and build leadership and communication skills.

Social Outcomes

- Reduced stereotypes and greater inter-cultural understanding
- Improved social responsibility and citizenship skills
- Greater involvement in community service after graduation

Career Development

- Connections with professionals and community members for learning and career opportunities
- Greater academic learning, leadership skills, and personal efficacy can lead to greater opportunity.

Relationship with the Institution

- Stronger relationships with faculty
- Greater satisfaction with college
- Improved graduation rates

BENEFITS OF COMMUNITY SERVICE PROJECT TO FACULTY MEMBERS

- Satisfaction with the quality of student learning
- New avenues for research and publication via new relationships between faculty and community
- Providing networking opportunities with engaged faculty in other disciplines or institutions
- A stronger commitment to one's research.

BENEFITS OF COMMUNITY SERVICE PROJECT TO COLLEGES AND UNIVERSITIES

- Improved institutional commitment.
- Improved student retention
- Enhanced community relations

BENEFITS OF COMMUNITY SERVICE PROJECT TO COMMUNITY

- Satisfaction with student participation
- Valuable human resources needed to achieve community goals.
- New energy, enthusiasm and perspectives applied to community work.
- Enhanced community-university relations.

SUGGESTIVE LIST OF PROGRAMMES UNDER COMMUNITY SERVICE PROJECT

The following the recommended list of projects for Engineering students. The lists are not exhaustive and open for additions, deletions, and modifications. Colleges are expected to focus on specific local issues for this kind of project. The students are expected to carry out these projects with involvement, commitment, responsibility, and accountability. The mentors

of a group of students should take the responsibility of motivating, facilitating, and guiding the students. They have to interact with local leadership and people and appraise the objectives and benefits of this kind of project. The project reports shall be placed in the college website for reference. Systematic, Factual, methodical and honest reporting should be ensured.

For Engineering Students

1. Water facilities and drinking water availability
2. Health and hygiene
3. Stress levels and coping mechanisms
4. Health intervention programmes
5. Horticulture
6. Herbal plants
7. Botanical survey
8. Zoological survey
9. Marine products
10. Aqua culture
11. Inland fisheries
12. Animals and species
13. Nutrition
14. Traditional health care methods
15. Food habits
16. Air pollution
17. Water pollution
18. Plantation
19. Soil protection
20. Renewable energy
21. Plant diseases
22. Yoga awareness and practice
23. Health care awareness programmes and their impact
24. Use of chemicals on fruits and vegetables
25. Organic farming
26. Crop rotation
27. Floury culture
28. Access to safe drinking water
29. Geographical survey
30. Geological survey
31. Sericulture
32. Study of species
33. Food adulteration
34. Incidence of Diabetes and other chronic diseases
35. Human genetics
36. Blood groups and blood levels
37. Internet Usage in Villages
38. Android Phone usage by different people
39. Utilisation of free electricity to farmers and related issues
40. Gender ration in schooling lvel- observation.

Complimenting the community service project the students may be involved to take up some awareness campaigns on social issues/special groups. The suggested list of programs

Programs for School Children

1. Reading Skill Program (Reading Competition)
2. Preparation of Study Materials for the next class.
3. Personality / Leadership Development
4. Career Guidance for X class students
5. Screening Documentary and other educational films
6. Awareness Program on Good Touch and Bad Touch (Sexual abuse)
7. Awareness Program on Socially relevant themes.

Programs for Women Empowerment

1. Government Guidelines and Policy Guidelines
2. Women's Rights
3. Domestic Violence
4. Prevention and Control of Cancer
5. Promotion of Social Entrepreneurship

General Camps

1. General Medical camps
2. Eye Camps
3. Dental Camps
4. Importance of protected drinking water
5. ODF awareness camp
6. Swatch Bharath
7. AIDS awareness camp
8. Anti Plastic Awareness
9. Programs on Environment
10. Health and Hygiene
11. Hand wash programmes
12. Commemoration and Celebration of important days

Programs for Youth Empowerment

1. Leadership
2. Anti-alcoholism and Drug addiction
3. Anti-tobacco
4. Awareness on Competitive Examinations
5. Personality Development

Common Programs

1. Awareness on RTI
2. Health intervention programmes
3. Yoga
4. Tree plantation
5. Programs in consonance with the Govt. Departments like –
 - i. Agriculture
 - ii. Health
 - iii. Marketing and Cooperation
 - iv. Animal Husbandry
 - v. Horticulture
 - vi. Fisheries
 - vii. Sericulture
 - viii. Revenue and Survey
 - ix. Natural Disaster Management
 - x. Irrigation

- xi. Law & Order
- xii. Excise and Prohibition
- xiii. Mines and Geology
- xiv. Energy

Role of Students:

- Students may not have the expertise to conduct all the programmes on their own. The students then can play a facilitator role.
- For conducting special camps like Health related, they will be coordinating with the Governmental agencies.
- As and when required the College faculty themselves act as Resource Persons.
- Students can work in close association with Non-Governmental Organizations like Lions Club, Rotary Club, etc or with any NGO actively working in that habitation.
- And also, with the Governmental Departments. If the program is rolled out, the District Administration could be roped in for the successful deployment of the program.
- An in-house training and induction program could be arranged for the faculty and participating students, to expose them to the methodology of Service Learning.

Timeline for the Community Service Project Activity

Duration: 8 weeks

1. Preliminary Survey (One Week)

- A preliminary survey including the socio-economic conditions of the allotted habitation to be conducted.
- A survey form based on the type of habitation to be prepared before visiting the habitation with the help of social sciences faculty. (However, a template could be designed for different habitations, rural/urban).
- The Governmental agencies, like revenue administration, corporation and municipal authorities and village secretariats could be aligned for the survey.

2. Community Awareness Campaigns (One Week)

- Based on the survey and the specific requirements of the habitation, different awareness campaigns and programmes to be conducted, spread over two weeks of time. The list of activities suggested could be taken into consideration.

3. Community Immersion Programme (Three Weeks)

Along with the Community Awareness Programmes, the student batch can also work with any one of the below-listed governmental agencies and work in tandem with them. This community involvement programme will involve the students in exposing themselves to experiential learning about the community and its dynamics. Programs could be in consonance with the Govt. Departments.

4. Community Exit Report (One Week)

- During the last week of the Community Service Project, a detailed report of the outcome of the 8 weeks' works to be drafted and a copy shall be submitted to the local administration. This report will be a basis for the next batch of students visiting that habitation. The same report submitted to the teacher-mentor will be evaluated by the mentor and suitable marks are awarded for onward submission to the University.

Throughout the Community Service Project, a daily logbook need to be maintained by the students batch, which should be countersigned by the governmental agency representative and the teacher-mentor, who is required to periodically visit the students and guide them.