VISION

To impart quality education through state-of-the-art technologies to achieve academic excellence for transforming students into innovators.

MISSION

To impart quality education through state-of-the-art technologies to achieve academic excellence for transforming students into innovators.

Creating a teaching-learning environment to produce industry ready and self-confident graduates. Motivate students to engage in creative projects throughout graduation.

To produce competitive graduates having creative skills and ethical values to succeed in their fields as well as the foundation for life-long learning.

PROGRAM EDUCATIONAL OBJECTIVES (PEO)

1. To provide students with strong basic and advanced programming concepts so that they can build solutions or systems for complex problems.
2. The program provides the fundamental and perspective to attain life-long learning in the thrust areas of Computer Programming.
3. To produce graduates who have ability to pursue research or have a successful career in academia or industries or as entrepreneurs.
4. The aim is inculcating technical knowledge of the programme and imbibe ethics with moral behaviour in the graduates.

PROGRAM SPECIFIC OBJECTIVES (PSO)

1. To acquire basic knowledge in hardware/software, algorithms, System Software, Computer graphics, Web design, Networking, and advanced computing for solving real-life and Research problems with the perspective of lifelong learning.
2. An ability to demonstrate Knowledge of data management systems like data acquisition and big data, Intelligent systems like AI, Data Science and Machine Learning, The techniques of data analytics like pattern recognition and knowledge discovery.
3. To develop skills which help to expand professional careers.

COURSE OUTCOMES (CO)

1. To identify the need for data science and solve basic problems using python basic concepts.
2. To learn the fundamentals of some of the widely used python packages and apply them into data analytics and visualization.
3. To prepare the data for design applications through various data pre-processing operations
4. To understand the concept of probabilistic and inferential concepts in data science.

~TABLE OF CONTENT~

Name:	Enrolment No:	Semester:

	"exam_data_list" with 5 list and each list stores all 4 key-value pairs for single student.				
03	Do the slicing of a given String to generate various substring by passing different index (like positive index, negative index, end index > string length, entire string), split this string into chunks of length 3 using list comprehension, split the string with specific character, iterate over the words of string. Apply trim, toupper, tolower, replace string and character, title, join and other operations on String. Perform add, union, intersection, difference, symmetric_difference, union, intersection_update, symmetric_difference_update, difference_update, discard, issubset, issuperset, isdisjoint, remove, pop and clear operations on Set. Import array module in python and perform all operations available in the module				
04	Download "heart_2020_cleaned.csv" dataset from "https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease "and perform all the descriptive statistics on above dataset using statistics module of python and scipy.stats package (Measures of central tendency, measure of dispersion/variation, measure of location, measure of shape and symmetry).				
05	Write following program on Pandas DataFrame: Create an array "rank" with 5 element (rank1, rank2,…,rank5). Create and display a DataFrame "exam" from a specified dictionary "exam_data_array" with "rank" as label. Also display a summary of basic information and its data. Perform following operations on DataFrame "exam": 1. Select the rows where the score is between 15 and 20 (inclusive). 2. Sort the data first by "score" in ascending order, then by "name" in descending order. 3. Replace the 'yes' and 'no' values from column "qualify" with True and False. 4. Display specified columns (columns: 2 and 4) and rows (row: 1,3 and 5). 5. Select the rows where number of attempts in the examination is less than 2 and score greater than 15. 6. Change the name 'James' to 'Suresh' in "name" column of the data frame. 7. Calculate the sum of the examination attempts by the students 8. Append one row. 9. Insert a new column "exam_name" and then Delete the "exam_name" column.				

	10. Convert a NumPy array, dictionary and first column of a DataFrame to a series Write following program on NumPy Array: 1. Create an array of all the even integers from 30 to 70. 2. Create an array of 10 zeros, other with 10 ones, and one more with10 fives. 3. Create a vector of length 10 with values evenly distributed between 5 and 50. 4. Create a 3x4 matrix filled with values from 10 to 21 and compute sum of all elements, sum of each column and sum of each row of a given array. 5. Create a 3x4 array and find the missing data in the array. 6. Calculate round, floor, ceiling, truncated and round (to the given number of decimals) of the input, elementwise of an array. 7. Find the maximum and minimum value, median, Weighted average, mean, standard deviation, variance, covariance matrix, of a given flattened array, minimum and maximum value along the second axis 8. Create a structured array from given student name, height, class and their data types. Now sort by class, then height if class are equal.				
06	Write following python programs on Beautiful Soup. Perform following operations on a HTML document. 1. Find the title tag from a given html document. 2. Count and retrieve all the paragraph tags and extract the text in the first paragraph tag. 3. Find the text of the first <a> tag and length of the text of the first <h2> tag. 4. Find the href of the first <a> tag. 5. Find the first tag with a given attribute value in an html document. Write a python program for MySQL Database connectivity (import sqlite3 module) Establish the connection with Education database named “Education” in SQLite, create a table named Student (with id_no, name, department, gander, total_marks) in Education database. Perform insert, update,select and delete operation on Student table.				
07	Write a python program to download appropriate dataset and explore random variable, Probability mass function, Probability density function, Cumulative distribution function, Discrete probability distribution and continuous probability distribution using scipy.stats , rv_discrete class and rv_continuous class .				

08	Write a python program to compute and explore normal distribution, central limit theorem, point estimate, interval estimation and hypothesis testing.				
09	1. Identify the column(s) of a given DataFrame which have at least one missing value, count the number of missing values in each column and drop the raws and columns with missing values. Check for the null values. Also remove the duplicate values from the DataFrame. Handle outliers in the Data Frame. 2. Access subset of data through indexing (Select data using labels (column headings)), Slicing(Extract range based subset, subset of rows, subset of columns, select a subset of rows and columns from our DataFrame using iloc method). 3. Perform other data processing on a given dataset.				
10	Write a python program to perform data visualization trough Matplotlib.				
11	Write python program for advanced data visualization through Seaborn.				
12	Exploring Google Data Analytics tool.				

MADHUBEN & BHANUBHAI PATEL INSTITUTE OF TECHNOLOGY
(A CONSTITUENT COLLEGE OF CVM UNIVERSITY)

DEPARTMENT OF COMPUTER ENGINEERING
A.Y. 2021-22, EVEN TERM
SUBJECT CODE: 102045603
SUBJECT NAME: PYTHON FOR DATA SCIENCE

~TABLE OF CONTENT~

Name:	Enrolment No:	Semester: