Resolution No.: AC/II(24-25).3.RUS21

S. P. Mandali's

Ramnarain Ruia Autonomous College

(Affiliated to University of Mumbai)

RUIA COLLEGE
Explore • Experience • Excel

Syllabus for

Program Code For Semester 6:

Data Analysis Using Python(RUSACDA)

(Credit Based Semester and Grading System for academic year 2024-25)

GRADUATE ATTRIBUTES

GA	GA Description **A student completing Bachelor's/Master's Degree in science program will be able to:**
GA 1	Recall and explain acquired scientific knowledge in a comprehensive manner and apply the skills acquired in their chosen discipline. Interpret scientific ideas and relate its interconnectedness to various fields in science.
GA 2	Evaluate scientific ideas critically, analyse problems, explore options for practical demonstrations, illustrate work plans and execute them, organise data and draw inferences.
GA 3	Explore and evaluate digital information and use it for knowledge upgradation. Apply relevant information so gathered for analysis and communication using appropriate digital tools.
GA 4	Ask relevant questions, understand scientific relevance, hypothesize a scientific problem, construct and execute a project plan and analyse results.
GA 5	Take complex challenges, work responsibly and independently, as well as in cohesion with a team for completion of a task. Communicate effectively, convincingly and in an articulate manner.
GA 6	Apply scientific information with sensitivity to values of different cultural groups. Disseminate scientific knowledge effectively for upliftment of the society.
GA 7	Follow ethical practices at work place and be unbiased and critical in interpretation of scientific data. Understand the environmental issues and explore sustainable solutions for it.
GA 8	Keep abreast with current scientific developments in the specific discipline and adapt to technological advancements for better application of scientific knowledge as a lifelong learner.

PROGRAM OUTCOMES

PO	Description **A student completing Bachelor's Degree in science program in the subject of Elements of Operations Research (AC) (Semester 5) / Data Analysis using Python(AC)(Semester 6) will be able to:**
PO 1	Understand, condense, visualize, analyze and interpret various data types generated in various scenarios of scientific, industrial, or social problems.
PO 2	Apply Statistical tools for data analysis.
PO 3	Pursue their higher education programs leading to post-graduate and/or doctoral degrees in Statistics, Data Science, Business Analytics, Biostatistics, Econometrics, Management Studies.
PO 4	Compete globally to enter into promising careers.
PO 5	Make a pathway to a range of traditional avenues in Academia and Industry, Govt. Service, IAS, Indian Statistical/ Economic Services, Industries, Commerce, Investment Banking, Banks and Insurance Sectors, CSO and NSSO, Research Personnel/Investigator in Govt. organizations such as NCAER, IAMR, ICMR, Statistical and Economic Bureau & various PSUs., Market Research, Actuarial Sciences, Biostatistics, Demography etc.
PO 6	Seek employment or self-employment in different sectors like Stock trading, Pharmaceutical sector, Sports, Politics, Business, Financial services and Media Industry.

COURSE OUTLINE

YEAR	SEMESTER	COURSE CODE	COURSE TITLE	CREDITS
TYBSc	VI	RUSACDA601	DATA ANALYSIS USING PYTHON	2
TYBSc	VI	RUSACDAP601	PRACTICAL BASED ON RUSACDA601	2

Course Code: RUSACDA601

Course Title: DATA ANALYSIS USING PYTHON

Academic year 2024-25

COURSE OUTCOMES:

COURSE OUTCOME	DESCRIPTION A student completing this course will be able to:
CO 1	Understand the foundational concepts of Python programming, including syntax, data types, control structures, and functions.
CO 2	Develop the ability to write Python code to perform statistical analysis, including descriptive statistics, hypothesis testing.
CO 3	Utilize Python programming skills to address real-world statistical problems , also employ Python libraries and tools to manipulate data, conduct exploratory data analysis, and visualize data.

Learning Outcomes:

Students should be able to

- Handle data files in Python
- Describe Numpy, Pandas, Strings, List, Tuples and Dictionaries in Python
- Express different decision making statements and Functions
- Draw various types of graphs and diagrams using python
- Apply python to small sample test and large sample test

DETAILED SYLLABUS

Unit		Title: – Data Analysis using Python	No. of lectures
Unit 1		**Introduction to PYTHON Software**	**15**
	1.1	Python Setup Python Arithmetic: Basic operators	
	1.2	Basic Data Types, Variables, Lists, Tuples and Strings, Dictionaries and sets Derive new variable/function Summary statistics	
Unit 2		**Numpy, Pandas and Data Exploration**	**15**
	2.1	**numpy arrays:** Creating arrays crating n-dimensional arrays using np.array and array operations(indexing and slicing, transpose, mathematical operations) **pandas data frames:** Creating series and data frames and	

		Operations on series and data frames **Reading and writing data:** From and to Excel and CSV files. Merging, sorting, sub-setting of Data files	
	2.2	**Control statements:** if, if-else, if-elif, while loop, for loop **Defining functions:** def statement **Text data operations:** len, upper, lower, slice, replace, contains, Frequency Tables	
Unit 3		**Descriptive statistics and Statistical Methods**	**15**
	3.1	**Plotting:** using "matplotlib"(Histograms, Box plots, Scatter plot, Barplot, Line plot) **Descriptive Statistics:** mean, median, mode, min, max, quantile, standard deviation , variance, skew, kurtosis, correlation **Probability distributions:** (using scipy.stats) computations of probabilities, Cumulative probabilities, quantiles and drawing random sample using functions for following distributions:	
	3.2	Simulation from distributions, Binomial, Poisson, Hyper geometric, Normal, Exponential, Uniform, Graphs of pmf/pdf by varying parameters for above distributions	
Unit 4		**Inferential Statistics**	**15**
	4.1	**Hypothesis testing and T-Tests:** (using scipy.stats, math) Large sample test , ttest_1sample, ttest_ind(2 sample test), ttest_rel(paired), Type I and Type II error	
	4.2	**Chi-square tests:** (using scipy.stats) chisquare, chi2 **ANOVA:** (using scipy.stats) f_oneway	
	4.3	**Linear regression:** from sklearn import linear model and use linear model. Linear regression function.	

DISTRIBUTION OF TOPICS FOR PRACTICALS

Course Code RUSACDAP601	
Sr. No	**Practical based on course**
1	Python basics on data types
2	Descriptive Statistics
3	Probability Distributions: Discrete

4	Probability Distributions: Continuous
5	Data visualization
6	Testing of Hypothesis
7	ANOVA
8	Regression analysis

REFERENCES

- Python for Data Analysis by O'Reilly Media (Second Edition) (2017)
- How to think like a computer scientist learning with Python by Allen Downey. (2002)
- Python for Data Analysis by Armando Fernandgo. (2017)

Modality of Assessment

Theory Examination Pattern:

A) Internal Assessment- 40%- 40 Marks

Sr No	Evaluation type	Marks
1	Class Test/ Project / Assignment / Presentation	20
2	Class Test/ Project / Assignment / Presentation	20
	TOTAL	**40**

B) External Examination- 60% - 60 Marks

Semester End Theory Examination:

1. Duration - These examinations shall be of **two hours** duration.
2. Theory question paper pattern:

Paper Pattern:

Question	Options	Marks	Questions Based on
1	Any TWO subparts out of Three subparts	16	Unit I
2	Any TWO subparts out of Three subparts	14	Unit II
3	Any TWO subparts out of Three subparts	16	Unit III

4	Any TWO subparts out of Three subparts	14	Unit IV
	TOTAL	**60**	

Semester End Practical Examination Pattern:

Particulars	Marks
Journal	20
Practical	**80**
Total	**100**

External Examination: 80 Marks

There will be Two question with 4 parts each. Each part will be based on one unit for 20 marks. Student will attempt ANY ONE question.

Duration - These examinations shall be of **THREE HOURS** duration.

Overall Examination & Marks Distribution Pattern

Semester VI

Course	RUSACDA601		
	Internal	**External**	**Total**
Theory	**40**	**60**	**100**
Practical	**--**	**100**	**100**
