Hindi Vidya Prachar Samiti's

Ramniranjan Jhunjhunwala College

of Arts, Science & Commerce

(Empowered Autonomous College)

Affiliated to

UNIVERSITY OF MUMBAI

Syllabus of Statistics VSC for the T. Y. B.Sc. Semester V (under NEP)

Program: B.Sc. STATISTICS

Program Code: RJSUSTA

(REVISED in 2025-2026 in alignment with the NEP2020 facilitating the inter-and multidisciplinary learning and multiple entry and exit of the students)

(CBCS 2025-2026)

Preamble

The National Education Policy 2020 aims at imparting skill-based learning and caters to the multiple entry and exit facility for the students thus empowering them to acquire knowledge at their pace. In the three-year UG program, the student has two exit options. Students also have the option of choosing the Honors program of four years study in a given discipline and later converting it to a five-year integrated PG degree program. As an undergraduate student, he/she learns the core subject (Major), subject complementing the core subject (Minor), a course from another discipline (OEC or GEC), Vocational and Skill Enhancement course from the Major (VSEC). The remaining verticals under NEP 2020 are IKS (Indian Knowledge System), AEC (Ability Enhancement Course), VEC (Value Enhancement Course) and with progressive three years of UG, student also completes at different levels OJT (On Job Training), FP (Field Projects), CEP (Community Engagement Program), RP (Research Project) which helps him/her in understanding their roots, application of the knowledge for the benefit of self and the society. Vertical CC (Co-curricular activities and activities related to yoga and human well-being) helps in preparing youth with good character and interpersonal relationships.

T.Y.B.Sc. Semester V Statistics VSC Syllabus

Credit Structure for B Sc Semester V as per NEP 2020 Implemented

from the academic year 2025-2026
Course Code: RJVSCSTAP

Semester V
Level 5.5 Major Statistics

Courses	Credits	Total Credits	Course Code
Discipline Specific Core -I Probability and Distribution Theory	4	12	RJDSCSTA351
Discipline Specific Core -II Theory of Estimation	4		RJDSCSTA352
Discipline Specific Core -III Practical on DSC-I and DSC-II	4		Practicals related to Theory I and II RJDSCSTAP351
Discipline Specific Elective-I Biostatistics	4	4	RJDSESTA351
Discipline Specific Elective-II Mathematical Economics			RJDSESTA352
Vocational Skill Course (VSC) Regression Analysis Using R software	4	4	RJVSCSTA351
Vocational Skill Course (VSC) Practical on VSC			Practicals related to VSC RJVSCSTAP351
Field Project (FP)	2	2	RJFPSTA351
Total Credits	22	22	

*As per University Grid

Credit Structure for VSC Statistics Semester V as per NEP 2020 Implemented from the academic year 2025-2026.

Vertical	Course Code	Nomenclature	Topics	Credits
VSC (4 Credits)	Paper Code : RJVSCSTA351	Regression Analysis Using R Software	Unit I: Fundamentals of R and Simple Linear Regression Model Unit II: Multiple Linear Regression Model and Validity of Assumptions	02
	Paper Code : RJVSCSTAP351		Practical	02

SEMESTER	:	V (Vocational Skill Course)
TITLE OF THE SUBJECT/COURSE	:	**Regression Analysis Using R Software**
COURSE CODE	:	**RJVSCSTA351**
CREDITS	:	02 (Theory 02)
DURATION	:	30 Hours

LEARNING OBJECTIVES	
1.	To develop the technique of model building using software.
2.	To understand how predictions can be made using regression models.

COURSE OUTCOME NUMBER	On completing the course, the student will be able to:	PSO Addressed	BLOOMS LEVEL
CO1	understand fundamental concepts and skills necessary to efficiently manage and analyze data, create charts and graphs.	3,4,5	BT level I and II remember and understand
CO2	develop the student's ability to deal with numerical and quantitative variables.	3,4	BT level I, II and III remember ,understand and apply
CO3	have a proper understanding of statistical investigations on data analytics.	3,4,5	BT level I, II and III remember ,understand and apply
CO4	enable the use of software packages on data manipulation and its interpretation.	3,4,5	BT level II, III and IV understand,apply and Analyze

	SEMESTER V		**L**	**Cr**
REGRESSION ANALYSIS USING R SOFTWARE		**Paper Code: RJVSCSTA351**	30	2
	UNIT I - FUNDAMENTALS OF R AND SIMPLE LINEAR REGRESSION MODEL		15	
1	Introduction to R, features of R, installation of R, Starting and ending R session, getting help in R , Value assigning to variables. Basic Operations : +, -, *, ÷, ^, sqrt . Numerical functions : log 10, log , sort, max, unique, range, length, var, prod, sum, summary, dim, sort, etc . Data Types : Vector, list, matrices, and data frame. Variable Type : logical, numeric, integer, complex, character and factor. Data Processing : Data import and export, setting working directory, checking structure of data :Str(), Class(), Changing type of variable (for eg as.factor, as.numeric).			
2	Data Manipulation : Selecting random N rows, removing duplicate row(s), dropping a variable(s) , Renaming variable(s), subsetting data, creating a new variable(s), selecting of random fraction of row(s), appending of row(s) and column(s), simulation variable. Data Visualization using ggplot: Simple bar diagram, subdivided bar diagram, multiple bar diagram pie diagram, Box plot for one and more variables, histogram, frequency polygon, scatter plot eg plot().			
3	Implementing t-Tests in R : One-sample, two-sample, and paired t-tests, Interpretation of R output. Implementing Chi-Square Tests in R : chi-square goodness-of-fit and independence tests, Interpretation of R output. Implementing ANOVA in R : One way and two way classification, Interpretation of R output.			
4	Assumptions of the model, Derivation of ordinary least square (OLS) estimators of regression coefficients for simple, Properties of least square estimators (without proof), Coefficient of determination R2 and adjusted R2. Procedure of testing a) Overall significance of the models b) Significance of individual coefficients c) Confidence intervals for the regression coefficients Data.			

	Pre-processing: Detection and treatment of missing value(s) and outliers, Variable selection and Model building, Interpretation of output produced by lm command in R. Polynomial Regression Models.		
***UNIT II : MULTIPLE LINEAR REGRESSION MODEL AND* VALIDITY OF ASSUMPTIONS**		15	
1	Derivation of ordinary least square (OLS) estimators of regression coefficients for multiple regression models, Coefficient of determination R2 and adjusted R2. Procedure of testing a) Overall significance of the models b) Significance of individual coefficients c) Confidence intervals for the regression coefficients. Derivation of ordinary least square(OLS) estimators of regression coefficients for a simple linear regression model with 2 explanatory variables. Autocorrelation: Concept and detection using Durbin Watson Test, Interpretation of output produced by DW-test function in R, Heteroscedasticity: Concept and detection using Breusch –Pagan-Godfrey Test, Interpretation of output produced by bptest function in R. Multicollinearity: Concept and detection using i) R2 and t-ratios ii) pairwise correlation between repressors iii) Variance Inflation Factor (VIF), Interpretation of output produced by mctest function in R, Consequences of using OLS estimators in presence of Autocorrelation, Heteroscedasticity and Multicollinearity, Remedial measures.		

References:

1. Purohit, S.G.; Gore, S.D. and Deshmukh, S.R. (2015). Statistics using R, second edition. Narosa Publishing House, New Delhi.
2. Verzani, J. (2005). Using R for Introductory Statistics, Chapman and Hall /CRC Press, New York.
3. Asha Jindal (Ed.)(2018), Analysing and Visualising Data with R software- A Practical Manual, Shailja Prakashan, K.C. College.
4. Nina Zumel and John Mount (2014): Practical Data Science with R- Manning Publications.
5. Damodar, Guajarati and Sangeetha (2007) : Basic Econometrics, Tata McGrew Hill Education Private Limited

SEMESTER	:	V (Vocational Skill Course)
TITLE OF THE SUBJECT/COURSE	:	**VSC Practical**
COURSE CODE	:	**RJVSCSTAP351**
CREDITS	:	02
DURATION	:	60 Hours

LEARNING OBJECTIVES	
1.	To develop the technique of model building using software.
2.	To understand how predictions can be made using regression models.

COURSE OUTCOME NUMBER	On completing the course, the student will be able to:	PSO Addressed	BLOOMS LEVEL
CO1	efficiently manage and analyze data and create clear charts and graphs.	3,4,5	BT level I and II remember and understand
CO2	enhance students' ability to work with numerical and quantitative variables.	3,4	BT level I, II and III remember ,understand and apply
CO3	fit different models like simple linear regression, Multiple linear regression, Anova etc using linear models.	3,4,5	BT level I, II and III remember ,understand and apply

Semester V			
Statistics Practical	**Paper Code: RJVSCSTAP351**	**L 60 hr**	**Cr 2**
Practicals based on VSC			
1. Fundamentals of R			
2. Graphs and Diagram			
3. t- Test			
4. Chi- Square test			
5. ANOVA test			
6. Simple Linear Regression using R			
7. Multiple Linear Regression using R			
8. Weighted Least Square using R			

Scheme of Examinations

1. Evaluation will be for 100 marks based on Theory and Practical assessment.
2. Internal (Theory) examination of 20 Marks based on MCQ/ True or false/ Short answers .
3. External (Theory) examination (Semester End Examination) of 30 marks based on all units.
4. External (Practical) examination (Semester End Examination) of 50 marks.
5. Minimum marks for passing the examination is 40 %.
6. A candidate will be allowed to appear for the practical examinations if he/she submits a certified Journal of T.Y.B.Sc. Statistics or a certificate from the Head of the Department / Institute to the effect that the candidate has completed the practical course of T.Y.B.Sc. Statistics as per the minimum requirements.
7. In case of loss of journal, a candidate must produce a certificate from the Head of the Department /Institute that the practicals for the academic year were completed by the student. However, such a candidate will be allowed to appear for the practical examination, but the marks allotted for the journal (if any) will not be granted.
8. HOD's decision, in consultation with the Principal, shall remain final and abiding to all.

Evaluation and Assessment
(Based on the centralised guidelines given by EC under NEP 2020)

Internal examination : 20 Marks

Internal Assessment	Max Marks	Duration	Evaluation Particulars
1	20	30 Minutes	MCQ / True or False / Short Answers/Projects / Assignments / Seminar

External (Theory) examination : 30 Marks

Duration: 1 hour.

Question no.	Max Marks	Question based on
1	10 (with internal option)	Based on all units
2	10 (with internal option)	
3	10 (with internal option)	
Total	30	

Key to set effective Question paper:

Question	Knowledge	Understanding	Application and Analysis	Total marks- Per unit
1	03	05	02	10
2	03	05	02	10
3	03	04	03	10

-TOTAL- **Per objective**	9	14	7	**30**
% WEIGHTAGE	30%	46%	24%	**100%**

Practical examination : **50 marks**
Duration : 1 hours 30 mins

Question no.	**Max Marks**
1	40 (with internal option)
Journal	05
Viva	05
Total	50

Mapping of the course to Local/ Regional/National/International relevance

Class	Course Name	Course Code	Local relevance	Regional relevance	National relevance	International relevance
T Y B Sc Statistics VSC	Regression Analysis using R Software	RJVSCSTA351	Enable to analyze local data trends.	Enable to analyze regional economic trends using Statistics.	Enhance productivity for data management, analysis, and reporting.	Gain analytical skills for better presentation and decisions.
T Y B Sc Statistics VSC Practical	Practical	RJVSCSTAP351	Enables data-driven decisions in local industries using R for regression analysis.	Supports evidence-based decision-making across key sectors in the region through regression analysis using R.	Enhance practical skills in data management, analysis, and reporting to improve productivity.	Develop hands-on analytical skills to improve data presentation and decision-making.

Mapping of the course to Employability/ Entrepreneurship/Skill development

Class	Course Name	Course Code	Topic focussing on Employability/ Entrepreneurship /skill development	Employability/Entrepreneurship/Skill development	Specific activity
T Y B Sc Statistics VSC	Regression Analysis using R Software	RJVSCSTA351	All topics	Data Analysis and Interpretation.	Conducting statistical analysis using R software

Integration of Cross cutting Issues

Class	Course Name	Course Code	Cross Cutting Issues
T Y B Sc Statistics VSC	Regression Analysis using R Software	RJVSCSTA351	Professional Ethics, Social accountability UNSDG 4, 8, 9 NEP2020 Interdisciplinary