METHODS IN EPIDEMIOLOGICAL DATA ANALYSIS

1. Credit Distribution, Eligibility and Pre-Requisites of the Course

2. Course title and Code	Credits	Credit distribution of the course			Eligibility Criteria	Prerequisite of the Course
		Lecture	Tutorial	Practical		
Methods in Epidemiological Data Analysis	2	0	0	2	Class XII	Nil

2. Learning Objectives

The Learning Objectives of this course are as follows:

- To gain practical experience through hands-on training with 'R', a free software environment for statistical computing and graphics with particular reference to epidemiological data.
- To acquire skills in presenting epidemiological data in different formats through tables and graphs.
- To learn about different epidemiological and health parameters related to communicable and non-communicable diseases.
- To understand relationships between different epidemiological variables using correlation and regression analyses.
- To learn how to design hypotheses and analyze epidemiological data to draw statistically significant conclusions.

3. Learning Outcomes

Upon successful completion of the course, students will be able to:

- Install R and execute various commands and functions in R to analyse epidemiological data.
- Generate tables and graphs to organize, stratify and present epidemiological data using R.
- Calculate various epidemiological parameters like prevalence, incidence, and vital statistics, and measures such as morbidity, mortality, DALYs, and fertility rates.
- Perform correlation and regression analyses
- Conduct various tests of significance and make statistical inferences.

4. Main Course Structure

Unit I: Introduction to 'R' statistical package (5 weeks/ 20 Hours)

Laboratory 1: Installation of R and RStudio, performing essential operations & commands, and

exploring basic data types and their functions in R.

Laboratory 2: Importing and exporting different types of data (CSV, Excel) and manipulating key data structures in R (vectors, data frames, lists).

Laboratory 3: Performing data cleaning tasks including handling missing data, removing duplicates and transforming variables.

Laboratory 4: Performing basic data manipulations such as subsetting, merging, filtering, sorting, etc.

Unit II: Descriptive Epidemiology and Visualization using R (5 weeks/ 20 Hours)

Laboratory 5: Summarizing epidemiological data with tables and stratify data based on key variables, such as age, gender, or risk factors to account for confounding variables.

Laboratory 6: Calculation of different statistical parameters including measures of central tendency (mean, mode, median and partition values) and measures of dispersion (range, standard deviation, variance, coefficient of variance and covariance).

Laboratory 7: Representation of epidemiological data as graphs (scatter plots, bar plots, histograms, pie charts, box plots, epidemic curve, etc.).

Laboratory 8: Calculation of epidemiological parameters and health indicators like prevalence, incidence, vital statistics (rates, ratios, and proportions), morbidity, mortality, life expectancy, hospitalization rate, DALYs (Disability-Adjusted Life Years), fertility rates, attack rate, case fatality rate, herd immunity threshold, etc. for communicable and non-communicable diseases.

Unit III: Inferential Statistical Analysis of Epidemiological Data Using R (5 weeks/ 20 Hours)

Laboratory 9: Correlation and linear regression analysis on epidemiological data.

Laboratory 10: Tests of significance for single mean and difference of means for large (z-test) and small samples (Student's t-test for independent and dependent samples).

Laboratory 11: Chi-square tests for independence, homogeneity, and goodness-of-fit to analyze categorical epidemiological data.

Laboratory 12: F-test, one-way and two-way ANOVA on epidemiological data.

5. Teaching Methodology/Activities in the classroom

Hands-on activities using real or simulated datasets, Online Databases and Tools, Videos, Project based learning, Workshops, etc.

6. Assessment Pattern for each Unit/practical.

1. Maintenance of practical records (10 Marks)
2. Viva Voce (10 marks)

Unit I: Introduction to 'R' statistical package

1. Installation of R and RStudio. Execution of basic operations and importing/ exporting of data. (10 Marks)
2. Problem solving activity based on use of various commands and functions in R for data cleaning and manipulation on provided datasets. (10 Marks)

Unit II: Descriptive Epidemiology and Visualization using R

1. Presentation of descriptive analysis of epidemiological data with appropriate graphs and tables using R. (10 Marks)
2. Class activity focused on interpreting epidemiological parameters related to any disease outbreak. (10 Marks)

Unit III: Inferential Statistical Analysis of Epidemiological Data Using R

1. Practical test for execution of various inferential tests such as correlation, regression, T-tests, chi-square tests, and ANOVA in R. (8 Marks)
2. Group report/ presentation to critically analyze/ review the statistical analysis performed in a published epidemiological study. (12 marks)

7. Mapping with the next suggestive course

DSE :Research Methodology (Sem VI/ VII)

8. Prospective Job Roles after a particular course

Epidemiological Data Analyst, Biostatistical Data Analyst, Survey Coordinator, Public Health/Disease Surveillance Assistant, Community Health/Social/NGO Worker, Project Assistant in academic and research labs.

9. Essential readings:

- Park, K. (2021), 26th Edition. Park's Textbook of Preventive and Social Medicine. Banarsidas Bhanot Publisher, ISBN: 9789382219163.
- A. Stewart (2022), 5th Edition. Basic Statistics and Epidemiology: A Practical Guide. ISBN: 9781003148111.
- Daniel, W.W. and Cross, C.L. (2019), 11th Edition. Biostatistics: A foundation for analysis in the health sciences. New York, USA: John Wiley & Sons. ISBN: 9781119588825.
- Website for 'R': www.r-project.org
- Aho, K. A. (2023). Foundational and Applied Statistics for Biologists Using R. United States: CRC Press. ISBN: 9781032477411.
- Carstensen, B. (2021), 1st Edition Epidemiology with R. United Kingdom: Oxford University Press. ISBN: 9780198841326.

10. Suggestive readings:

- Christiansen-Lindquist, L., Christiansen-Lindquist, L., Wall, K. M., Wall, K. M. (2024). Fundamentals of Epidemiology. United States: Springer Publishing Company. ISBN: 978-0826166937.
- Rahman, A., Abdulla, F., Hossain, M. M. (2024). Scientific Data Analysis with R: Biostatistical Applications. United States: CRC Press. ISBN: 9781040146972.
- Webb, P., Bain, C., Page, A. (2024). Essential Epidemiology: An Introduction for Students and Health Professionals (5th ed.). Cambridge: Cambridge University Press. ISBN: 9781009415361.

- Welham, S. J., Mead, A., Clark, S. J., Gezan, S. A. (2024). Statistical Methods in Biology: Design and Analysis of Experiments and Regression. United States: CRC Press LLC. ISBN: 9780826166944.
- Quinn, G. P., Keough, M. J. (2023). Experimental Design and Data Analysis for Biologists. United Kingdom: Cambridge University Press. ISBN: 9781107036710.
- Triola, M.M., Triola M.F., Roy J. (2019). 2nd Edition. Biostatistics for Biological and Health Sciences. Harlow, UK: Pearson Education Ltd. ISBN: 9789353436537.
- A. Aschengrau and G. R. Seage, (2018), 4th Edition. Essentials Of Epidemiology In Public Health Ann Aschengrau and George R. Seage. ISBN:97812841283s52
- Hui, E.J.M. (2018). 1st Edition. Learn R for Applied Statistics With Data Visualizations, Regressions, and Statistics Hui. Springer New York, ISBN: 9781484242018.
- Pagano, M. and Gauvrean, K. (2018). 2nd Edition. Principles of Biostatistics. California, USA: Duxbury Press. ISBN-13: 9781138593145.
- Bertram K.C. Chan (2016), 1st Edition. Biostatistics for Epidemiology and Public Health Using R. ISBN: 9780826110268
- Norman, G.R. and Streiner, D.L. (2014). 4th Edition. Biostatistics: The bare essentials, New York, USA: McGraw-Hill Medical. ISBN: 978-1607951780.
- Zar, J.H. (2014). 5th Edition. Biostatistical analysis. USA: Pearson. ISBN-13: 9789332536678.
- Katz, D.L., Elmore, J.G., Wild, D. Lucan, S.C. (2013). 4th Edition. Jekel's epidemiology, biostatistics, preventive medicine and public health. Philadelphia, USA: Elsevier Saunders. ISBN: 978-1455706587.
- Glantz, S. (2012). 7th Edition. Primer of biostatistics. New York, USA: McGraw-Hill Medical. ISBN-13: 9780071781503.
- Bonita, R., Beaglehole, R. and Kjellstrèom, T. (2006). 2nd Edition. Basic epidemiology. Geneva, Switzerland: World Health Organization. ISBN-13: 978-9241547079.