



Internship on Data Science for Humanities and Social Sciences: Tools, Techniques, and Applications

(A Multi-Disciplinary Internship Program [Teaching and Project] under NEP 2020 for Undergraduate Students with No Prior Knowledge of Mathematics and Statistics; Students will be exposed to hands on training with Excel, STATA and Basic techniques / tools of Artificial Intelligence)

Course Duration & Mode

- **Duration:** 22nd November, 2026 to 31st December, 2026
- **Total Duration: 60 Hours (2 Credits)**
 - **Teaching Component:** 30 Hours
 - **Project Component:** 30 Hours
- **Mode:** Online classes through **Google Meet**, from **6:00 PM onwards**.
- **Requirement:** Participants must have access to a **laptop or desktop computer** for attending classes and completing project work.

Teaching Structure:

- **Number of Lectures:** 20 (Online Mode)
- **Duration per Lecture:** 1 hour 30 minutes

Project Component:

- **Duration:** **30 Hours**
- **Mode:** **Individual Projects**

Students will work individually on real-world datasets relevant to social science research

Who can apply:

- UG students in Languages, Literature, and Social Sciences, with no prior background in Mathematics or Statistics.

Course Fee for UG Students

- **Regular Fee:** ₹2,500/-
- **Fee Waiver:**
 - **₹500 waiver** for students recommended by the Principal/Head of the Institution.
 - **Additional ₹500 waiver**, subject to approval, for economically weaker students with financial constraints certified by the Principal/Head of the Institution.



Internship on Data Science for Humanities and Social Sciences: Tools, Techniques, and Applications

How to Apply

Apply online through the official application page:

www.starebs.com → [Data Science for Social Sciences](#)

Alternatively, applicants may apply directly through the Google Application Form:

<https://forms.gle/nnrS7Fq9dThMPtEY8>

- Read the instructions and open the application form.
- Fill in the required details.
- Submit the form online.

For Fee Waiver: UG students seeking a fee waiver must upload a recommendation letter from the Principal/Head of the Institution. Students seeking an additional waiver under the Economically Weaker category must submit a letter certifying their financial constraint.

Course Fee Structure

- **UG Students (Regular): ₹2,500/-**
- **UG Students (with Institutional Recommendation): ₹2,000/-**
Upload a recommendation letter from the Principal/Head of the Institution on official letterhead before payment.
- **UG Students (with Financial Constraints): ₹1,500/-**
Upload a recommendation letter from the Principal/Head of the Institution confirming the student's financial constraint.

Commencement of the Course: 22nd November, 2026

Last Date of Application: 18th November, 2026

For Queries:

 starebsconnect@gmail.com |  +91-9002016612



Internship on Data Science for Humanities and Social Sciences: Tools, Techniques, and Applications

COURSE OUTLINE

Module 1: Understanding Data in Humanities and Social Sciences

Data Science, Statistics and Data Analytics; role of data in Humanities and Social Sciences; quantitative, qualitative and mixed-method research; numerical, categorical, textual, image, audio, archival and historical data; cross-sectional, time-series and panel data; primary and secondary data; structured and unstructured data; government, Census, institutional, archival and digital data sources. Practical: Identifying research questions and appropriate datasets.

Module 2: Data Cleaning, Management and Visualization Using Excel

Research data organisation and coding; data cleaning, missing and duplicate observations, validation and transformation; Excel formulas and functions, frequency tables, pivot tables and summary statistics; data visualization and appropriate chart selection; AI-assisted data cleaning, analysis and visualization. Practical: Cleaning, analysing and visualizing a real dataset.

Module 3: Statistics Without Fear

Population, sample, variables and observations; descriptive and inferential statistics; frequency distributions, mean, median, mode, percentages, quartiles and range; variance and standard deviation; correlation and causation; hypothesis testing, p-values and statistical significance. Practical: Interpreting statistical results without mathematical derivations.

Module 4: Introduction to STATA for Humanities and Social Science Research

STATA interface and dataset structure; importing and managing data; creating, recoding and labelling variables; missing-value management; descriptive statistics, cross-tabulations and graphs; correlation and introductory regression; interpretation of coefficients, significance and R^2 ; AI-assisted STATA coding, debugging and verification. Practical: Basic empirical analysis using a Social Science dataset.

Module 5: Text as Data — Introduction to Digital Humanities

Digital Humanities and computational approaches to Humanities research; textual data, digitisation, OCR and corpus construction; word frequency, keywords, word clouds, concordance, n-grams and co-occurrence; comparative and computational literary analysis; historical newspaper and document analysis; AI-assisted text analysis. Practical: Analysing a literary or historical text corpus.

Module 6: Artificial Intelligence for Humanities and Social Science Research

Artificial Intelligence, Machine Learning, Generative AI and Large Language Models; AI-assisted literature review, research questions, conceptual frameworks, questionnaires and data analysis; classification, prediction and clustering; statistical and textual analysis using AI; hallucinations, bias, privacy, copyright, academic integrity, human verification and AI disclosure. Practical: Using AI as a research assistant.



Internship on Data Science for Humanities and Social Sciences: Tools, Techniques, and Applications

Module 7: Surveys, Historical Data and Social Research

Census and sample surveys; population, sampling frame and sampling methods; questionnaire design, open and closed questions and Likert scales; avoiding bias; AI-assisted questionnaire design and response coding; historical, archival and Census data; construction of historical datasets; social and development indicators. Practical: Designing a questionnaire or constructing a social/historical dataset.

Module 8: From Research Question to Data-Driven Research Project

Research problem, research gap, research questions and objectives; research design, data requirements, sources, variables and methodology; data analysis, visualization and interpretation; research report writing, presentation and viva preparation.

30-Hour Project Component

Students will undertake research project under one of the following tracks:

English / Literature: Literary text analysis, author comparison, gender representation, sentiment analysis, vocabulary change and character analysis.

Digital Humanities: Newspaper text mining, historical documents, literary corpus analysis, digital archives, text visualization and mapping.

History: Census and population analysis, migration, historical newspapers, education, and historical social and economic change.

Sociology: Youth attitudes, education and inequality, gender and employment, migration, social media and survey-based social attitudes.

Economics: Inflation and price trends, unemployment and labour markets, poverty and inequality, consumption, education and earnings, financial inclusion, digital payments, informal enterprises and economic development.

Geography: Population distribution, migration, urbanisation, land-use change, climate and environmental trends, regional disparities, spatial analysis, GIS and geospatial data.

Commerce: Consumer behaviour, digital payments and e-commerce, financial literacy, investment behaviour, MSMEs, business performance, corporate performance and market trends.

Multidisciplinary: Education and inequality, migration and employment, digitalisation and social change, gender and economic participation, urbanisation and livelihoods, climate and development, health and socioeconomic outcomes, and technology and financial inclusion.