

Economic Well-Being Through Data: All India's Debt & Investment Survey 2019

STATISTICAL ANALYSIS USING NSSO UNIT LEVEL DATA

FACULTY DEVELOPMENT PROGRAMME

19th January to 20th February 2026

7:30 pm – 9:00 pm (Monday to Friday)

Online (Webex)

Department of Humanities, Arts and Social Sciences,
National Institute of Technology – Calicut
Kozhikode (Kerala)

About the Workshop

The National Sample Survey Office (NSSO) of India's Ministry of Statistics and Programme Implementation conducts socioeconomic surveys across the country regularly and publishes results. The reports contain aggregate (macro) level information that is relevant to research and policy. The unit-level data from various rounds of the NSS surveys, on the other hand, is a rich set of information that is widely used in studies by researchers, corporate executives, and policymakers. NSSO conducts surveys on variables related to agriculture, employment, consumption, health, education, time-use, housing condition, disabilities, migration, unorganised manufacturing, debt and many other important segments.

The National Statistical Office (NSO) conducts the All India Debt & Investment (AIDIS) survey in the rural and urban areas of the country at periodic intervals. The current survey, latest in the AIDIS series, was conducted during the period January to December, 2019. Information in the survey was collected in two visits (Visit 1: January-August, 2019 and Visit 2: September - December, 2019) from the same set of sample households. The survey was spread over 5,940 villages covering 69,455 households in the rural sector and 3,995 blocks covering 47,006 households in the urban sector.

But due to the lack of a detailed manual on obtaining unit-level data from the NSSO and due to a lack of training in data extraction, merging and how to construct variables researchers often find it difficult to use the NSSO's unit-level data in their research. Because of that, only a small number of scholars are using these big national-level data sets in research. This is where this FDP will come in handy, as it will enable researchers with the applied skills to use NSSO's unit-level data in research. The objective of the present workshop is to provide a step-by-step process (methodology) of extracting, merging and analysing the unit-level data of NSSO surveys by using STATA software.

Objectives:

- To develop an understanding of national-level surveys on various types of economic and social variables of national importance.
- To aware participants on the utilization of the national level data for their respective field of research

- To increase the use of these unutilized data in research to find solutions for economic, social, geographical and environmental problems in future.
- To help researchers to apply statistical analysis tools on unit-level data.

Pedagogy:

A mix of Pedagogical tools will be used. Discussion and hands-on training with real data using statistical software. Participants will be guided with the relevant resources to facilitate the hands-on training*.

*STATA participants have to arrange for themselves.

Learning Outcomes:

- Participants will learn Basic STATA commands: gen, rename, label, egen, recode, tag, destring (with ignore and force), count, list, describe, substring, split, sort, browse.
- Participants will be introduced to official sources to download NSSO raw data and required files from ICSSR and MoSPI.
- Participants will understand the role of required documents and raw data files. Like, Read me file, Schedule, Instructions to field staff, Layout file and Reports.
- Learn how to extract the data of different levels.
- Participants will learn merging of files with different natures/combinations of files.
- Participants will understand the use of weight and cross-verification of results with the report.
- Participants will learn descriptive, correlation and inferential statistical methods using NSSO Unit Level Data.

Target Participants:

- Research Scholars, Faculty Members and Working Professionals in the area of economics, health, population studies, planning and related areas.

Number of Participants:

40 (Maximum)

Mode of Conduction:

Online (Webex)

Registration Fee:

Research Scholars -Rs. 3600

Faculty Member and Working Professionals- Rs. 4800

FDP Coordinator-

Dr. Mohamad Aslam

Assistant Professor

Department of Humanities, Arts and Social Sciences

NIT Calicut

Tentative Schedule-

	NSSO UNIT LEVEL DATA	STATISTICAL ANALYSIS
Week 1	<p>Extraction-</p> <p>How to apply the CONCATENATE command in the layout file.</p> <p>To make dct files in STATA using details from the layout file.</p> <p>To apply the command in Do-file Editor to extract level-wise data in STATA to get all the files in the dta format.</p> <p>About cross-verifying the data files with the details provided in the readme file, to test the accuracy of the extracted files.</p> <p>Append function for old rounds.</p>	<p>Introduction to the STATA software and essential STATA commands, such as gen, rename, label, egen, recode, label, tag, destring (with ignore and force), count, list, describe, substring, split, sort, browse.</p> <p>Learning about types of data (cross-sectional and time series), the Nature of variables (nominal/ordinal/interval/ratio scale), and tools for descriptive statistics and correlation.</p>
Week 2	<p>Merging and Matching</p> <ul style="list-style-type: none"> • To identify household and individual levels using a schedule, layout, questionnaire and instructions to field staff. • How to generate individual IDs and household IDs using common primary keys. • About merging one file with another for the following combinations 1:1, 1:m, m:1, and m:m. • Household with household, household with individual, individual with individual (with and without duplicates) • How to deal with duplicates. • How to generate and apply weights. • About cross-verifying the extracted data with the report, to validate the data. 	<p>Descriptive Statistic (qualitative): Learning about Tabulation and Graphical Methods for categorical data- Two-way graphs: Bar graph and Pie chart (one/two/three/four variable using by and over option), horizontal and vertical bar graph, segmented bar graph with stack format.</p> <p>Tabulation with one variable, two variables, three variables, and four variables. Option for groups, over options, by option, if option, controlling the categorical axis, controlling legends, changing the look, and graphing by groups.</p>
Week 3	<p>Construction of variables</p> <ul style="list-style-type: none"> • After extracting and merging the data files level-wise, one needs to calculate the variables as per the research objectives. • To illustrate these cases in detail, we teach how to calculate important indicators of NSSO round(AIDIS 2019). • Incidence of Indebtedness • Average Asset Holdings by Social Group 	<p>Numerical and Graphical Methods for Quantitative Data:</p> <p>Learning about Location and Dispersion Methods- Arithmetic, Geometric and Harmonic mean, Median, Mode, Range, Standard Deviation, Variance, Inter Quartile Range, Coefficient of Variation, Identify outliers, Skewness, and Kurtosis. Criteria for selecting mean or median as the centre and standard deviation or Inter Quartile Range as dispersion. Comparison across groups and categories. Learning about Graphical methods- Range plot,</p>

		Distribution plot, Box plot, Dot plot, and Histogram.
Week 4	<ul style="list-style-type: none"> • IOI by type of Credit agency • Institutional credit by Nature of Interest • Cash debt by Purpose • Other required functions, like reshape and data filters, to handle the data. 	Correlation Analysis: For Nominal, Ordinal and Ratio scale variables- Direction of association, and Degree of association. Nominal scale with nominal scale, Likert scale with nominal scale, ordinal scale with ordinal scale, ratio scale with ratio scale. Graphical analysis with scatter plot. Numeric Parametric and Non-Parametric methods (Karl Pearson, Spearman's rank correlation, Kendall's rank correlation, Pearson's chi-squared with Cramer's V). Discussing common mistakes in correlation analysis and the relevance of correlation analysis for regression analysis.
Week 5		Inferential Statistics: Learning about distributions, properties, and probabilities for usage in hypothesis testing. Process of hypothesis testing with confidence interval/ p -value/test of significance methods. Formulation of the null and alternative hypotheses, Selection of the Level of significance (with Type I and Type II errors), Selection of the test statistic, Critical Value, Decision Rule, one-tailed/two-tailed, and role of the Central Limit Theorem. Applications with test of hypothesis (t -test, z - test, proportion test, and variance comparison test), One sample, two sample using group, two sample using variables and paired, One way ANOVA, Two-way ANOVA and Normality test (Skewness and Kurtosis test and Shapiro Wilk normality test). Non-parametric methods and Test of Hypothesis-Wilcoxon Matched Pairs signed-rank test, Wilcoxon rank-sum test, Kruskal Wallis rank test.

Resource Persons-

Dr. Subhajit Dutta, Associate Professor, Indian Statistical Institute (ISI), Kolkata

Dr Subhajit Dutta is currently working as an Associate Professor at the Indian Statistical Institute, Kolkata. Prior to that, he worked in the Department of Mathematics and Statistics at IIT Kanpur. He received his Ph.D. in Statistics from the Indian Statistical Institute, Kolkata, in 2013 under the supervision of Prof. Probal Chaudhuri, and later worked as a Postdoctoral Fellow at KAUST with Prof. Marc G. Genton. His research interests include discriminant analysis, data depth-based inference, characterization of multivariate distributions, and classification of sequence data.

Dr. Thomas Felix, Assistant Professor, Institute for Social and Economic Change (ISEC), Bengaluru

Dr. Thomas Felix is currently working as an Assistant Professor in ADRTC, Institute for Social and Economic Change (ISEC), Bengaluru. Prior to that, he worked as an Assistant Professor of Economics at Tamil Nadu National Law University (TNNLU), Tiruchirappalli. He completed his B.Sc. (Agri) from Thanthai Roever Institute of Agriculture and Rural Development, Preamble, M.Sc. (Agri) with the specialisation of Agricultural Economics from Agricultural College and Research Institute, Madurai and Ph.D. (Economics) from Tamil Nadu Agricultural University, Coimbatore. He got the UGC-MANF fellowship award for his doctoral programme, and he also got the ICSSR Post-Doctoral Fellowship. After his doctoral programme, he worked as a Research Associate at the Madras Institute of Development Studies (ICSSR Institute) in Chennai. Before joining TNNLU, Dr. Thomas Felix was engaged as a Post-Doctoral Fellow at Tamil Nadu Agricultural University, Coimbatore. He has cleared ASRB-NET, UGC-NET and TN-SLET. He has published 20 research articles in reputed journals, presented more than 25 papers in national and international conferences and participated in 10 national workshops. His areas of interest include Agricultural Economics, Law and Economics, Natural Resource Economics, Environmental Economics and International Trade and Marketing.

Dr. Mohamad Aslam, Assistant Professor, Department of Humanities, Arts and Social Sciences, National Institute of Technology, Calicut.

He is currently working as an Assistant Professor in the Department of Humanities, Arts and Social Sciences at NITC. Prior to that, he worked as a Postdoctoral Fellow at IISER, Mohali. His Ph.D is from Panjab University. He has completed his M.Phil. from Jawaharlal Nehru University. His research interests lie primarily in agrarian studies, labour market, economic transformation among tribals, education, and inequality. Currently, he is working on economic changes in the Himalayan States using NSSO unit-level data. And on the challenges of tribals in occupational transformation in the North Western Himalayas. He has teaching experience of more than half a decade in theoretical and applied statistics and econometrics. He has already conducted several such workshops and FDPs with Institutes of National Importance, including national seminars and many other courses to train scholars and working professionals in applied skills.

Ms. Gaibul Preet, Co-Founder and Program Coordinator at Stats2econo Applied Research Pvt. Ltd., Palampur, Himachal Pradesh.

An Economist, Educator, and Assessment Developer, with a unique blend of subject expertise, academic training, and instructional design experience. Holds a Master's in Economics and a degree in Education, and has qualified for key national-level teaching exams including UGC-NET (Economics), CET, and PTET. Her journey spans roles in academia, content development,

and EdTech, all driven by a passion for making complex ideas accessible and engaging. As Co-Founder and Program Coordinator at Stats2Econo, she contributed to course design and now oversees the smooth execution of live, interactive programs in Econometrics, Statistics, and Data Analysis.