Jindal School of Government and Public Policy
India's First Public Policy School

O.P. JINDAL GLOBAL
INSTITUTION OF EMINENCE DEEMED TO BE
UNIVERSITY
A Private University Promoting Public Service

Advanced Topics in Statistics and Data Analysis with Stata
Spring 2026

Course Information

Course Duration: February through mid-May, 2026
Credit Hours: 4 credits
Meetings: TBD
Prerequisites: Statistics, basic econometrics

Instructor Information

Instructor: Pradeep Guin, PhD
Email: pguin@jgu.edu.in
Phone: +91-98181 78885
Office: FOB, 6th floor, 13 West
Office Hours: By appointment
Homepage: https://jgu.edu.in/jsgp/faculty/dr-pradeep-guin

Teaching Assistant Information

TA: TBD
Phone:
Email:

1. Course Description

Stata is a powerful statistical analysis software that allows us numerous ways to manipulate and present data. This course is designed to introduce this software, which can perform both basic as well as advanced statistical and econometric analyses. It is ***not*** intended to explain you the statistical and/or econometric methods. Students will be introduced to both basic and advanced features of Stata (e.g., learning how to use a *do-file*) leading to efficient data management and analysis skills. We will include several topics on statistical analyses, performing regression analyses, as well as delving into the territory of graphics.

2. Course Intended Learning Objectives (Aim)

Course Intended Learning Outcomes	Teaching and Learning Activities	Assessment/Activities
Basic understanding of Stata as a data management tool	Class lectures and group discussion	• In-class lecture • One research paper • One class presentation based on the research paper • One poster based on the research paper • Class participation
Use Stata to perform statistical and regression analyses and create graphs		

3. Scheme of Evaluation and Grading[1]

Evaluation is comprised of internal (70%) and external (30%) assessments.

Internal evaluation is comprised of:

- **Attendance (5%)**
- **Classroom participation (5%), and**

Group-based assignment:

- **End-term research paper (45% for UG and 40% of PG):** A maximum of 12/15-page (double space, 12 font Times New Roman, default margin on each side) research paper based on testing a research hypothesis. The paper should include the following components:
  1. Introduction/background: Present your research topic, question, and hypothesis.
  2. Literature review: Include a discussion from the existing literature including various aspects of your topic, including gaps in literature and your research topic's contribution.
  3. Data & method: Talk about the dataset(s), how you have merged (if you have merged), what method(s) have you used to analyze the data, etc.
  4. Results: You are required to use your Stata skills to do the analysis and present the results. Include some of the most important tables/graphs in your paper; you may include the remaining ones as part of the Annexure section (this is outside of the prescribed page limit).

[1] Evaluation scheme is indicative and may change.

5. Discussion: Present a discussion of your analysis. Include a sub-section on limitations of your research work.
6. Conclusion

Please note that the filename of your paper should follow ***groupXX_respaper.docx*** pattern (XX represents your group number, i.e., 01, 02, and so on).

To be able to work on the paper, each group would do the following:

1. Identify a research question that you would want to work on. There is no limitation to this, you can pick any topic.
2. Based on the topic, develop the hypothesis that you would want to test. You will apply statistical and econometric techniques (that you have already learned as part of your courses) to test the hypothesis.
3. You will identify dataset(s) that has information regarding variables that you would use to test the hypothesis. It may be the case that information is available across several datasets, in which case you will merge those datasets to create a final dataset that you would eventually use to conduct the analysis. You may seek help from the course instructor to identify dataset(s).
4. You will analyze the data using techniques that you have learned as part of this course, as well as any other techniques that you may have learned elsewhere.
5. You will make a final presentation and submit a poster (only for graduate-level students) based on this research work.
6. You are required to submit your STATA *do* and/or *smcl* file(s) that you have used as part of your analysis plan.

- **Presentation (15% for UG and 10% of PG):** Groups will make a power-point presentation in the class based on their research work. This is a timed activity. Each group will have a total of 15 minutes, which included 10 minutes of presentation time and 5 minutes of Q&A.

- **Poster (10% for PG):** Post-graduate level students will create a poster based on their research work. This poster should follow the quality standards of a professional conference.

External assessment (30%) will include an in-class exam. The date of this exam is TBD.

Grade definition

The following grading scheme will be applicable.

Letter Grade	Percentage of Marks	Grade Points	Interpretation
O	80 and above	8	Outstanding
A+	75 – 79	7.5	Excellent
A	70 – 74	7	Very good

Letter Grade	Percentage of Marks	Grade Points	Interpretation
A–	65 – 69	6	Good
B+	60 – 64	5	Fair
B	55 – 59	4	Acceptable
B–	50 – 54	3	Marginal
P1	45 – 49	2	Pass 1
P2	40 – 44	1	Pass 2
F	Below 40	0	Fail
P*	Pass		Opting for pass/fail grading system over CGPA
I	Incomplete		Incomplete

* This may (or may not) be applicable.

4. Academic Integrity

Academic Honesty, Cheating, and Plagiarism: As per University policy.

Participation/Attendance Policy: As per University policy. You are required to have a minimum of 75% attendance to be able to not obtain a "no-bar" status for this course. You will receive an attendance ***only*** if you attend the entire class. Please refrain from coming up with any excuse for a late show-up in the class. Documented medical conditions (self and family) and VC/Registrar/Dean/Vice-Dean's written excuse will ***only*** be considered as valid reason(s) for your absence. Please provide such documents within one week from your missed class.

Use of phone/texting/laptop: As a courtesy to your instructor and fellow colleagues you are expected **not** to use mobile phone devices for any purposes, unless there is an emergency call/text that you need to attend. You may use a laptop judiciously, solely for the purposes of in-class requirements.

5. Keyword Syllabus

Stata, data management, data analyses, graphics, statistical and econometric techniques.

6. Course Material

A. Books: We will use the following books. You are required to ***skim through*** the readings.

I. Bittmann, F. (2019). *Stata: A really short introduction*. Berlin/Boston: Walter de Gruyter GmbH. **[FB]**

II. Acock, A. C. (2018). *A gentle introduction to Stata* (6th ed.). College Station, TX: Stata Press. **[ACC]**
III. Mitchell, M. N. (2010). *Data management using Stata: A practical handbook.* College Station, TX: Stata Press. **[MNM]**
IV. Baum, C. F. (2016). *An introduction to Stata programming* (2nd ed.). College Station, TX: Stata Press. **[CFB]**
V. Cox, N. J. (2014). *Speaking Stata graphics*. College Station, TX: Stata Press. **[NJC]**

B. Web links:

https://stats.idre.ucla.edu/stata/

http://data.princeton.edu/stata/

7. Session plan[2]

[2] Session plans are suggestive, which may be modified if necessary.

Week	General Topic	Readings	Approach/Pedagogy

1	• Introduction: course goal and objectives
	• The graphical user interface (GUI)
	• The basic steps
	  o Opening Stata files
	  o Do-files
	  o Manually entering data
	  o Using Stata in-built data
	  o Saving a file
		FB, ACC & MNM	Lectures, discussions, class presentations

2 – 3	• ***Each group to submit their research topic***
	• Reading and preparing datasets
	  o Reading various types of data files
	  o Importing various types of data files
	  o Preparing a dataset for analysis
		FB, ACC & MNM	

4 – 5	• ***Discuss your dataset(s) and variables to be used*** • Describing data o Basic statistical analysis	https://www.stata.com/features/basic-statistics/	
6 – 7	• ***Discuss your data analysis method(s)*** • Graphics	NJC	
8 – 11	• Analyses using various data types and topics	CFB & ACC	
12 – 13	• Reporting results: tables and graphs	FB	
14 – 15	• Presentations and course overview • Submission of poster • Submission of end-term research paper		