PADP7120 DATA APPLICATIONS IN PUBLIC ADMINISTRATION
Thursdays @ 15:30-18:15 (FALL 2018)
Department of Public Administration & Policy
School of Public & International Affairs
The University of Georgia

Instructor: Tima T. Moldogaziev, Ph.D.	**Office Hours:** F 09:30-10:30; or
Office #: Baldwin Hall 280B	by appointment.
Course Materials: Posted on eLC	**Classroom:** Candler 214
E-mail: timatm@uga.edu or via eLC	**Course #:** 26207

Course Description:

PADP7120 DATA APPLICATIONS IN PUBLIC ADMINISTRATION takes a systematic approach to the exposition of the general linear model for continuous dependent variables. In addition to laying the theoretical foundations for linear econometric approaches, this course introduces students to the use of computerized statistical analysis using Excel and the software program *R*. Students are encouraged to think creatively about ways to use statistical methods in their own research.

This course provides students an opportunity to develop quantitative analysis skills that can be applied to public management and policy problems, program evaluations, and critical research questions. We will emphasize application of statistical techniques, interpretation of statistical results, the use of statistics in management decision-making, and implementation of statistical tools using computer software. This course emphasizes both statistical theory and software skills necessary to perform analysis.

To that end, during the semester, students meet once a week each week for a 2.5-hour session. Generally, one half of this session will be a lecture on statistical fundamentals, theory, applications, and related topics. The second half will be a lab exercise that focuses on computing methods and data analysis techniques necessary for completion of home assignments. ***(My preferred level of preparation for this course is at least one statistics/econometrics course at an undergraduate level or PADP7110 Research Methods in Public Administration.*** *Above that, there are no college-level mathematics prerequisites; beyond the typical high-school algebra and geometry/trigonometry courses, of course. Students are not expected to have a background in calculus, but facility with algebra and exposure to the rudiments of statistical distribution theory and hypothesis testing is expected. Please brush up your skills with the elements of linear and non-linear algebra, families of probability distributions, and the properties of normal and non-normal 'curves'. Students with prior computer programming exposure will have an advantage when working in R. Please use this open source statistics book written for HS students if you need a basic statistics textbook to polish up your econometrics fundamentals: https://www.openintro.org/stat/index.php?stat_book=aps.* ***)***

The course is organized into four sections. The first section of the course covers the fundamental statistical concepts that are the building blocks for regression analysis. The purpose of this section is both to refresh your memory and to provide a deeper, more formal presentation of familiar concepts. This section also introduces students to *R.* The second section focuses on the assumptions and mechanics of the classical linear regression model. At the end of the second section you will have a good mechanical knowledge of linear regression analysis. The third section includes a practical exposition of the general linear model as we begin to relax the assumptions of the classical linear regression model. At the end of the third section you will have a deeper theoretical and applied understanding of the flexibility and limitations of the general linear regression model. The final section presents an overview of topics in estimation for common problems in social science research, including an introduction to non-linear binary and/or count outcome models. The purpose of this brief section is to give you some exposure to more

complex models (beyond continuous dependent variables) rather than to ask you to develop sophistication with these techniques. Those students interested in taking a statistics course after this one should be able to enroll in categorical dependent variable, longitudinal/panel, time series, multilevel methods, and structural equation model courses at MPA or PhD levels.

Course Competency Objectives:

Upon successful completion of this course, students should be able to:

1. Identify and propose questions of analysis that are pertinent to contemporary public policy and the broader study of public administration.
2. Formulate a step-by-step approach for analyzing public management problems and policy questions, as well as identify and offer resolution to various issues often encountered during quantitative analysis.
3. Identify the most appropriate methodological tools for analysis of given research questions and available data.
4. Analyze and interpret data using each of the methodologies covered in the course.
5. Produce data analysis to effectively inform the public and other stakeholders.

Student progress on these learning objectives will be measured through homework assignments, computer applications, and a final research project.

Required Reading Materials/Application Tools:

Textbook: We will use an open source statistics book—you can download a free pdf copy here: *https://www.openintro.org/stat/index.php?stat_book=os*.

1. **Diez, David M., Christopher D. Barr, and Mine Çetinkaya-Ründel.** 2016. *OpenIntro Statistics*. 3rd Eds. www.openintro.org. (A hard copy can be purchased on Amazon for $25 or less.)
2. A more technical book, but relevant to public policy students—also free to UGA students via Springer e-library @ UGA library: **Monogan III, James E.** 2015. *Political Analysis Using R.* Springer. Available at: https://www.springer.com/us/book/9783319234458.

Data Applications: **Excel (or Numbers)**—it is expected that students already have access to this software. ***R or R-studio*** is available on most university work stations. It can also be downloaded for free (!!!) here: https://cran.r-project.org/. This is also a very good place for all the user written packages and FAQ regarding *R*.

> [Suggested *R* book—also free (!!!):
> **Weisberg, Sanford.** 2014. *Computing Primer for Applied Linear Regression Using R.* 4th Eds. http://z.umn.edu/alrprimer.]

Additional/occasional readings from other sources will be distributed in class or posted on eLC. Students are expected to have completed the readings before the class that the chapters are assigned for. Other necessary tools: Access to a computer station that has Microsoft Office 2007/2010/2013/10 package applications. **You will need a portable flash memory/thumb drive.** You will also need lots of paper and ink to print your assignments. (Please NEVER print your log/output files, but keep an electronic copy at all times--you will also upload it on eLC. MAC users: please convert the log files to pdf files prior to uploads. There are compatibility issues between Mac and Windows text files.)

Important disclaimer, one more time: The course assumes that students have the necessary skills in algebra + geometry and college level statistics/econometrics. Given the time frame and the goals of this course, it is neither possible nor efficient to spend time on math or statistics fundamentals. Recall that

PADP7110 *Research Methods in Public Administration* is a suggested prerequisite to this course. The course instructor may decide to substitute this prerequisite by another statistics/econometrics course/training that a student has completed elsewhere.

Assignments & Grading Scale:

Deliverables:
Students are required to complete 6 homework assignments during the course of the semester. Assignments 1 to 5 will closely follow the topics in the syllabus. Assignment 6 is a final GROUP project that should reflect what students have learnt in a collaborative setting. Each assignment is worth between 35-100 points. Each of the assignments includes data analysis exercises using *R* and the course data extracts provided. (BEWARE that there are minor differences in coding between Mac and Windows machines. No support will be provided for Linux machines in this course.) The earlier assignments will be structured, but towards the end of the semester you will be asked to choose or construct your own variables for analysis. You can use the course data extracts for these assignments, or you may discuss alternative data sources with me in advance. Note that you should design these analyses so that the results reveal interesting relationships.

There will also be 6 to 8 in-class data management/analysis lab assignments that will require the use of *R* in small groups. About 120-150 points will be reserved for these in-class lab assignments. Typically, the in-class assignments will be pre-tests for codes and concepts that will be useful for completing individual and/or group home assignments. You will do a great deal of your work for the course in computer labs. Laboratory sessions focus on computing methods and data analysis techniques. If you have questions about lectures, if you have concerns about what is required in order to answer a question on an assignment, if you are wondering how to interpret your results, see me. If you are having problems analyzing your data, be sure to bring a hardcopy listing of the command file and the output, along with an electronic version of the command file and the output file. It is impossible to diagnose error messages without these. If you send a question electronically, include the *R* log file.

Finally, there is a fully *R* based mid-term exam, which is designed to evaluate students' ability to manage data sets, complete descriptive and inferential analyses, and interpret statistical results. The midterm exam is 150 points. Overall, home assignments, in-class exercises, and the midterm add up to **750 points (=100%).** The following scale is used to determine the course grade:

A	93-100	B-	80-82	D+	67-69
A-	90-92	C+	77-79	D	63-66
B+	87-89	C	73-76	D-	60-62
B	83-86	C-	70-72	F	< 60

Deadlines and Late Penalties:
It is critical that you keep up with assignments. Assignments should be handed to me in class or in my office on the due date. Be sure to confirm with me if you need to make alternative arrangements or if you are handing in a late assignment. Late assignments will be penalized by 10 points if they are received within 24 hours of the time due, 20 points if they are received within 48 hours of the time due, 30 points if they are received within 72 hours, 40 points if they are received within 96 hours, and 50 points if received after 96 hours. In no case will assignments be accepted on or after the sixth calendar day after the due date.

Working Together:
Students are encouraged to discuss homework assignments and data preparation with each other. In particular, when cleaning data and constructing new variables for the early assignments it is a good idea to compare your data with one or two other students before beginning your write-up. Students are also

encouraged to share their *R* programs or 'command files'. The final product (with the exception of group assignments), however, must reflect your own work. On individual computer assignments that require that you choose variables for analysis, everyone is expected to use different variables. If you are aware that someone else is using the same variables that you are using, one or both of you need to change variables.

Attendance & Participation:
It is expected that students attend all classes. We will do many applied problems in labs every week, individually and in groups. If you miss a class, you may find it very difficult to complete the home assignments on your own.

The surest way to learn is to participate. The best way to participate is to join class discussions and ask questions. When students are in the classroom it is expected that they participate; it's an integral part of your 'job description'.

Note 1: I do understand that we all have the days when we are late. Should you be late, don't be upset, you are very welcome to join the class. However, **chronic lateness** will be considered as negative participation and will be graded correspondingly (after a short while, it becomes obvious who is chronically late).
Note 2: Leaving class early (without a prior notice; i.e. before class) is not tolerated. If you leave in the middle of the class without any substantive justification, your action will be considered as negative participation.
Note 3: Please participate in class discussions by using the widely expected and accepted norms of civility. Please adhere to the norms of university student conduct. If you are not sure what these are, please study the link of the Office of Student Conduct: https://conduct.uga.edu/content_page/welcome-to-student-conduct-content-page.

Students that accumulate three instances of negative participation will see a 200-point deduction from their total grade.

Academic Honesty:

The University of Georgia requires all members of the University community to be responsible for knowing and understanding the policy on academic honesty. In addition, every student must agree to abide by the University of Georgia's academic honesty policy and procedures when applying for admission to the University of Georgia.

The University of Georgia defines academic honesty as "performing all academic work without plagiarism, cheating, lying, tampering, stealing, giving or receiving unauthorized assistance from any other person, or using any source of information that is not common knowledge without properly acknowledging the source." Academic dishonesty is defined as "performing, attempting to perform, or assisting any other person in performing any academic work that does not meet this standard of academic honesty."

According to the policy's prohibited conduct, "No student shall perform, attempt to perform, or assist another in performing any act of dishonesty on academic work to be submitted for academic credit or advancement. A student does not have to intend to violate the honesty policy to be found in violation. For example, plagiarism, intended or unintended, is a violation of this policy." The policy also states that, "Any behavior that constitutes academic dishonesty is prohibited."

ANY INSTANCE OF ACADEMIC DISHONESTY WILL RESULT IN A GRADE OF F FOR THIS COURSE. In addition, the instructor reserves the right to pursue further academic disciplinary action. It is your responsibility to adhere to the University of Georgia's policies concerning academic honesty. See

the Office of the Vice President for Instruction for policies regarding academic honesty: https://ovpi.uga.edu/academic-honesty/academic-honesty-policy.

Students with Disabilities:

Students who have a disability that requires accommodations should contact the Disability Resource Center to discuss their needs and obtain appropriate paperwork. I cannot make special accommodations for students with disabilities unless students have completed the appropriate paperwork to register with the Disability Resource Center. For further details, please see: https://drc.uga.edu/.

eLC:

This syllabus, necessary reading materials, and homework materials will be posted on the course on-line pages--eLC. More on this will be discussed in the classroom throughout the semester.

Other…

It is the student's responsibility to keep all copies of graded/returned assignments for this course. This will protect all the parties involved should any misunderstandings arise. Electronic devices must be switched off during the class time. No texting will be tolerated. Should the student need to keep such a device switched on for any important reason, the course instructor should be consulted before the class starts. The course instructor reserves the right to define what an "important reason" constitutes. Finally, **laptops and computers can be used only for the purposes of completing lab assignments.** Uses for other purposes may be allowed IF AND ONLY IF one has a required need to use a computer or an electronic device due to a particular medical disability.

Class schedule: August 16th through December 6th, 2018 (class schedule is subject to adjustments; any changes will be announced in advance and/or posted on the eLC.)

Weeks (Dates)	THEMES/READINGS/EXTRA MATERIALS Finish ALL readings PRIOR to class.	ASSIGNMENTS OUT/DUE
Week 1 (Aug 16)	**Introduction & Data Concepts** Readings: Chapter 1	
Week 2 (Aug 23)	**Sources & Nature of Data** Readings: Textbook, Chapter 1	**HW #1 assigned**
Week 3 (Aug 30)	**Working in *R*** Readings: Monogan, Chapters 1 & 2 Weisberg, Chapter 0 Lab *R* code will be provided.	The entire class session will be lab-based.
Week 4 (Sep 6)	**Refresher: Probability & Distributions** Readings: Textbook, Chapter 2 & 3 Monogan, Chapter 3	****HW #1 DUE**** **HW #2 assigned**
Week 5 (Sep 13)	**Refresher: Univariate Inference & Hypothesis Testing** Readings: Textbook, Chapter 4, 5 & 6 Monogan, Chapter 4	
Week 6 (Sep 20)	**Linear Regression & Binary Hypothesis Testing** Readings: Textbook, Chapter 7 Monogan, Chapter 5 Weisberg, Chapters 1 & 2	

Week 7 (Sep 27)	**Multiple Regression** Readings: Textbook, Chapter 8 Weisberg, Chapter 3	****HW #2 DUE**** **HW #3 assigned**
Week 8 (Oct 4)	**OUT-OF-CLASS ASSIGNMENT** **[**Examples using Regression Analysis: **Reading and making sense of regression reports. AFTER READING THESE: Select group project members (3-members per group).]** **FINAL GROUP ASSIGNMENT** Brainstorm about a final project similar to readings; write a (1-page) research proposal. **Important NOTE:** GROUP HW #6 is a substitute for your Final Examination. Please treat it as a final research project. **Start work NOW.**	**GROUP HW #6—FINAL project assigned** HW #6 has a presentation component on November 29.
Week 9 (Oct 11)	**Model Functional Forms; Log-Lin, Lin-Log, and Log-Log Models; Indicator Variables** Readings: Textbook, Chapter 8 Monogan, Chapter 6 Weisberg, Chapters 4 & 8	****Final Group Project Proposal DUE**** Bring to class for a discussion. Cost = 20 points.
Week 10 (Oct 18)	**Multicollinearity; Heteroskedasticity** **Model Selection Criteria and Tests** Readings: Textbook, Chapter 8 Monogan, Chapter 6 Weisberg, Chapter 9	
Week 11 (Oct 25)	**Time-Series & Autocorrelation** Readings: Textbook, Chapter 8 Monogan, Chapter 9	****HW #3 DUE**** **HW #4 assigned**
Week 12 (Nov 1)	**Working in *R*** Completing a multiple regression exercise in *R*	The entire class session will be lab-based.
Week 13 (Nov 8)	**Take-Home MIDTERM EXAM (Cost = 150): Lab-based analysis**	Open for completion between 03:30—11:30PM.
Week 14 (Nov 15)	**Non-Linear Outcomes:** **probit & logit; count models** Readings: Textbook, Chapter 8 Monogan, Chapter 7 Weisberg, Chapters 11 & 12	****HW #4 DUE**** **HW #5 assigned**
Week 15 (Nov 22)	**No Class: Happy Thanksgiving Day!**	
Week 16 (Nov 29)	**RESEARCH/REPORT PRESENTATIONS**	****HW 6 Presentations****
Week 17 (Dec 6) THURSDAY	**EXAM WEEK** **Submit HW #6 via eLC on Dec 6, 2018 by 6PM** [incorporate comments from the presentation].	****HW #5 DUE**** ****GROUP HW #6—FINAL project DUE****
Grades Available:	The course grades will be submitted to the Registrar's Office after the finals week.	**You are all done here. Good luck elsewhere!**