GEO391 (Statistical Data Analysis for GIS)

Winter Quarter 2015 |Department of Geography |DePaul University

Time: Mon/Wed 4:20-5:50 pm **Location**: SAC 224 (GIS Lab), LPC

Instructor: Sungsoon (Julie) Hwang (shwang9@depaul.edu; (773)325-8668; 990 W Fullerton Room# 4513; office hours: Mon/Tue/Wed 2-3 pm or by appointment; home page: gis.depaul.edu/shwang)

Teaching Assistant: David Semitekol (email: DAVIDSEMITEKOL@GMAIL.COM; lab hours: TBA)

Course Description: This course covers statistical concepts and techniques that are applied to geographic problems. Topics include descriptive statistics and inferential statistics geared toward working with geographic data in GIS. Students will learn techniques for summarizing variables, testing difference of means, and relationships among variables through lectures and hands-on exercises using calculators, SPSS and ArcGIS. GEO 241 or PSC 201 is a prerequisite for this class.

Course Topics: through this course you will learn

- What statistics and graphs are used for summarizing variables (descriptive statistics)
- How probability can be calculated if data is normally distributed (probability)
- How sample is different from a population, and what sampling means to inference (sampling)
- How hypothesis testing is conducted (one-sample difference of means test)
- How to test difference of means among independent or related samples (independent sample t-test, matched-pair t-test, ANOVA)
- What spatial autocorrelation is, and how spatial autocorrelation is measured
- How to detect spatial hot spots and outliers statistically from areal (lattice) data
- How to determine how a variable is associated with another variable (correlation)

Learning Outcomes: After completing all requirements of GEO 391, you should be able to

- Calculate descriptive statistics using SPSS, and interpret them appropriately
- Summarize attributes using ArcGIS
- Calculate descriptive spatial statistics using ArcGIS, and interpret them appropriately
- Calculate probability based on normal distribution
- Generate spatial samples for survey using ArcGIS
- Test hypothesis (one-sample difference of means test)
- Conduct difference of means test (t-tests and ANOVA) appropriate for a given problem using SPSS, and interpret results appropriately
- Explore spatial autocorrelation of data using different techniques in ArcGIS
- Determine hot spots and spatial outliers from areal data in ArcGIS, and interpret results appropriately
- Conduct correlation analysis, and interpret results appropriately in SPSS and ArcGIS

Required Text: J. Chapman McGrew Jr., Arthur J. Lembo, Jr. Charles B Monroe (2014) *An Introduction to Statistical Problem Solving in Geography*, The Third Edition, Waveland Press (ISBN: 9781478611196)

Course Outlines and Tentative Schedules

Date	Topic	Reading	In-class group exercises
1.05	Course overview	Syllabus	
1.07	01. Descriptive statistics	Ch. 3	01. Descriptive statistics and graphs in SPSS
1.12			02. Descriptive statistics in ArcGIS
1.14	02. Descriptive spatial statistics	Ch. 4	03. Descriptive spatial statistics in ArcGIS
1.19	03. Probability	Ch. 6	04. Normal distribution - probability
1.21	04. Sampling	Ch. 7	05. Spatial sampling for error matrix in ArcGIS
1.26	**Test1**		
1.28	05. Hypothesis testing	Ch. 9	06. One-sample difference of means test
2.02	06. Two-sample difference tests	Ch. 10	
2.04	07. ANOVA	Ch. 11	07. Difference of means test (SPSS)
2.09	08. Spatial autocorrelation	Ch. 13	08. Explore spatial autocorrelation (ArcGIS)
2.11	09. Areal pattern analysis	Ch. 15	09. Spatial hot spot analysis for lattice data (ArcGIS)
2.16	10. Correlation	Ch. 16	10. Correlation analysis (SPSS)
2.18			
2.23	**Proposal** due		
2.25	**Test2**		
3.02	Project		
3.04	Project		
3.09	**Project presentation**		
3.11	**Project presentation**		
3.18	**Project report** due		

Grading Schemes

Components	Points (/100)	Description
Exercises	**30**	10 group exercises in class. Each exercise is worth 3 points
Tests	**32**	Two closed-book, non-cumulative tests. Each is worth 16 points.
Project	**28**	Proposal (4) + presentation (4) + report (20)
Participation	**10**	Attendance, class discussion, and participation to group exercises

Exercises: Three persons will be paired to work on exercises designed to reinforce concepts covered in readings and lectures. The format includes reflecting on concepts learned in the class, calculating statistics manually, interpreting analysis results in SPSS, and making maps that incorporate statistical concepts in ArcGIS. If groups can't complete exercises in the class, it is expected that groups complete exercises on their own. I suggest you work with a lab assistant during lab hours. Unless noted otherwise, exercises are due on D2L in a week from the date when exercises are handed out.

Tests: There will be two closed-book tests in the classroom. You will be allowed to bring one sheet of paper in which you write down formulae. You will be provided with information necessary to solve problems, including normal table, student distribution, and F table, etc. Test1 covers the first half to be held at week 4, and Test2 covers the second half to be held in week 8 unless noted otherwise. The format of tests will vary from manual calculation, short essay, to conducting analysis in SPSS and ArcGIS.

Project: You are to propose and conduct a mini geographic research amenable to quantitative analysis during the last two weeks of the quarter. You should submit one-page research proposal, and present your work before classmates. By week 11, submit a three-page report (excluding figures) that addresses the purpose of a study, research hypotheses, data collection, data analysis, and findings. Detailed guidelines will be provided later.

I recommend that you (a) choose a familiar topic so that you don't have to spend bulk of time in reviewing literature and familiarizing yourself with data; (b) focus on research questions that can be tested. For example, it will be easier to examine change in temperature than in religious views because temperature is more robust to measurement than religious views and temperature data is easy to collect (i.e., available from NOAA website). Project can be done either individually or in group.

In the previous quarter, students have written about whether diversity has increased or decreased by newly drawn wards boundary in Chicago, how abortion rates are associated with political affiliation in US States, whether Wicker Park has gentrified, and so on. Note that all these papers test one or more research hypotheses based on empirical data by employing an appropriate statistical technique.

Participation: assigned according to the criteria below

- A (9 - 10) = Student is present in all or nearly class meetings, and prepared, at all times, to respond to questions. Student is an active participant in small group activities, in and out of class, and in class-time activities stays on task.
- B (8 - 9) = Student participates as above, 75% of the time.
- C (6.5 - 8) = Student does not volunteer comments; responses demonstrate vague familiarity with course readings. Student is a passive member of small group activities and/or does not stay on task during class-time activities.
- D (5 – 6.5) = Student never volunteers, cannot respond to direct questions, keeps silent during class discussions and is unable to summarize readings if asked.
- F (0-5) = Student misses many class sessions and/or sits silently in classes when present, or is disruptive and non-participatory in the classroom.

Grading scale: A = 93-100%; A- = 90-92.99%; B+ = 87-89.99%; B = 83-86.99%; B- = 80-82.99%; C+ = 77-79.99%; C = 73-76.99%; C- = 70-72.99%; D+ = 60-69.99%; D = 50-59.99%; F = 0-49.99%

Late Work Policy: Late work will be accepted with 20% of the total grade deducted for each day being late. Extensions can be requested if needed, but will be only granted if excused. This policy applies to exercises only. Although I will accommodate to ongoing progress in the class, if group exercises are not completed in class, it is expected that your group will complete exercises in time through collaboration outside of the class.

Makeup Exam/Incomplete Grade Policy: A makeup exam or an incomplete grade can be arranged or granted only when credible dire and documented medical or family situations arise and these circumstances are communicated in a timely fashion.

Attendance/Absentee Policy: Consistent with university's policy, all students are expected to attend class meetings.Unless absence is explained on medical or compassionate grounds (documentation is required), absence from any classes is grounds for a grade adjustment.

Academic Honesty and Plagiarism: Academic honesty and integrity are expected at all times. Academic dishonesty, such as cheating or copying during exams, will be punished severely. Plagiarism – using someone else's work without acknowledgment and, therefore, presenting their ideas or quotations as your own work – is strictly forbidden. DePaul University officials will be informed of any instance of academic dishonesty and notification will be placed in your file. Please read the DePaul Academic Integrity Resources page (http://academicintegrity.depaul.edu/Resources/index.html) for definitions and explanations of plagiarism and the University's Academic Integrity expectations for students. Cutting and pasting text taken directly from a web-site without appropriate referencing and quotation marks is plagiarism and is forbidden. Submitting work that has any part cut and pasted directly from the internet is grounds for an automatic grade of zero.

Accommodations: Any student who requires assistance is asked to contact the University's Center for Students with Disabilities (CSD) (Phone 773/325-1677, TTY 773/325-7296, Fax 773/325-

7396,http://studentaffairs.depaul.edu/studentswithdisabilities).They will be able to assist both student and faculty. If you have a condition that requires accommodation from the Productive Learning Strategies program (PLuS Program) please contact them at the Student Center room 370 (Phone 773/3251677 or online: http://studentaffairs.depaul.edu/plus/

University Center for Writing-Based Learning: Collaborates with writers from all disciplines, backgrounds, levels of expertise, and roles within the University community. Their goal is to help develop better writers along with better writing and reflection through continual revision. If you need assistance with writing assignments, they can be contacted at: 773.325.4272 (LPC) or wcenter@depaul.edu

Department of Geography Learning Goals—GEO 391 addresses 1), 5), 6), and 7).

Courses in the Department of Geography teach students:

1) Understand spatial patterns and processes of modification of the Earth's physical and cultural landscapes
    a. As social constructions.
    b. As systems that link the Earth with human society in interdependent, dialectical relationships, and
    c. Through mapping and visualization.
2) Understand the concept of scale as a spatial phenomenon that ties the local, the regional, the national, the transnational, and the global in a system of interaction.
3) Understand the phenomenology of the discipline of Geography—most importantly, "space", "place", "landscape," "region," and "location".
4) Distinguish that spaces, places, and so on, may have both objective and subjective/symbolic dimensions.
5) Develop research and writing competences that would allow you to:
    a. Formulate a cogent research question about the spatial character of a physical, socio-cultural, or environment-societal phenomenon,
    b. Write about it in ways that reflect analytical and critical thinking, and
    c. Ethical concern over social and environmental justice, consistent with the University's social mission.
6) Engage competently in qualitative and quantitative spatial analysis, and with exercises that are concerned with explaining spatial regularities (for example, the spatial calculus behind the location of retail commerce in Chicago, or transnational flows of capital).
7) Learn the basic utility and use competently one or more of the information technologies that are now redefining the logistical limits of spatial analysis: geographic information systems (GIS) and remote sensing.
8) Achieve greater general knowledge of the world, its regions, its physical systems, its cultures, and political-territorial divisions.

Learning Objectives and Outcomes by Modules

Module 1: Descriptive Statistics (Read chapter 3)

- Describe measures of central tendency, including *mode*, *median*, and *mean*
- Describe measures of dispersion, including *interquartile range*, *variance*, and *standard deviation*

➢ Ex 01: Calculate descriptive statistics using SPSS, and interpret them appropriately
➢ Ex 02: Summarize attributes in a table using ArcGIS

Module 2: Descriptive Spatial Statistics (Read chapter 4)

- Describe spatial measures of central tendency, that is *mean center*
- Describe spatial measures of dispersion, that is *standard distance*

➢ Ex03: Calculate descriptive spatial statistics using ArcGIS, and interpret them appropriately

Module 3: Probability (Read chapter 6—focus on normal distribution)

- Describe characteristics and appropriate use of the *normal distribution*
- Describe *z score*
- ➢ Ex04: Calculate *probability* of a given condition based on normal distribution

Module 4: Sampling (Read chapter 7)

- Distinguish between *population* and *sample*
- Describe when to use different types of *sampling strategy* with spatial data in mind
- ➢ Ex05: Generate spatial samples for survey (error matrix) using ArcGIS

Module 5: Hypothesis Testing (Read chapter 8, 9)

- Understand what *standard error* of estimates is
- Describe basic concepts in hypothesis testing—*null hypothesis*, *type I error*, *significance level*, *test statistics*, and *p-value*
- ➢ Ex06: Test hypothesis (one-sample difference of means test) using *p-value approach*

Module 6: Difference of Means Tests (Read chapter 10)

- Describe objective and test statistic of *two-sample difference of means test* (aka. Independent samples t-test)
- Describe objective and test statistic of *matched-pairs difference test* (aka. Paired samples t-test)

Module 7: ANOVA (Read chapter 11)

- Describe objective and test statistics of *ANOVA*
- ➢ Ex07: Conduct difference of means test (t-tests and ANOVA) appropriate for a given problem using SPSS, and interpret results appropriately

Module 8: Spatial Autocorrelation (Read chapter 13, 15 on Moran's I)

- Understand concepts of *spatial autocorrelation*
- Describe methods for exploring spatial autocorrelation, including *variogram* and *Moran's I*
- ➢ Ex08: Explore spatial autocorrelation of data using different techniques in ArcGIS

Module 9: Areal Pattern Analysis (Read chapter 15)

- Describe *Local Moran's I*
- Describe *Getis-Ord Gi**
- Ex 09: Determine hot spots and spatial outliers from area data in ArcGIS, and interpret results appropriately

Module 10: Correlation (Read chapter 16)

- Describe different ways to conduct correlation analysis—graphing (*scatter plot*) and *correlation coefficient*
- Describe an objective and inferential test statistic of *Pearson correlation coefficient*
- ➢ Ex10: Conduct correlation analysis, and interpret results appropriately in SPSS and ArcGIS