APSTA-GE 2015 Applied Spatial Statistics (2 pts.)
Marc Scott

[TOPIC ORDER SUBJECT TO CHANGE]

Meeting time: Fri 1:00-4:30 Lecture/Lab
Important: no lecture week 6 (see below)
Location: 7 E. 12th St. LL 33

Office: 207W Kimball Hall
Hours: Fridays 9:30-10:30am (tent.)
Phone: 212-992-9407 (email advised)
Email: tm11@nyu.edu

Main text: Chun & Griffith, Spatial Statistics & Geostatistics (Sage 2013)
Software: R version 3.5 (freely distributed). This course will use NYU Classes.

COURSE OVERVIEW: Spatial data arise when information is collected on units that reside in different locations. Common examples include geology, criminology and epidemiology, where the goal may be to identify patterned or clusters (‘hot spots’) in the outcomes across the terrain being examined. In the social sciences, a similar set of questions and techniques are required, for example in studies of homelessness, poverty, environmental justice, and education. However, spatial data present a novel set of exploratory and modeling challenges, given the unique way in which outcomes are related (correlated) with each other through proximity. This course is an overview of the methods needed to analyze data for which it is suspected that the spatial component plays an important role.

PREREQUISITES: an advanced course in statistics in which hierarchical or heteroscedastic linear models has been presented. At least one of courses that follow should be sufficient: APSTA-GE 2004, PSYCH-GA 2248, POL-GA 2251 (Quant III), SOC-GA 2312, or APSTA-GE 2110 AND PADM-GP 2902. Consult with the instructor before registering if another course is to be used as a prerequisite.

R language: it is assumed that you are comfortable using the R programming language/environment. If you need a crash course, contact me immediately to gain access to certain R modules in “datacamp.”

COURSE REQUIREMENTS:
Participation: 20% You are expected to attend class and participate in class discussions; there will be opportunities to ask questions in an online forum, and these will count toward the participation grade.

Data Analysis Project: 50% There will be a multi-component data analysis problem that will require computing, analysis, and interpretation. Completed in teams/groups.

Proposal: 30% There will be a 3-5 page writeup that includes the following components:
• Description of a spatial dataset to which one potentially could gain access (the feasibility matters if one is continuing on to the practicum in the second half of the term)
• A set of research question(s) one could address using these data
• An outline – research plan – of the methods they would use to do this, based on the material covered in class
• A self-critique of the choices (limitations, alternatives, etc.)

Students with an existing spatial data set that they wish to analyze are encouraged to register additionally for the 1-credit MLM Practicum (APSTA-GE 2041).

COURSE TOPICS AND READINGS FOR EACH CLASS:

Notes: Each of the first four classes will consist of two lectures, spaced apart; this “front loads” the essential material for doing projects.
Every class involves discussion of the readings (including lecture notes) & looking at code associated with the reading; we quickly dedicate more and more time to lab/group work on the Stop&Frisk dataset, which is the foundation of the project.
Text: the course text provides highly scaffolded examples and covers some topics more or less than we do.
The lecture notes form the core of the course, with the text functioning more as workbook and complementary viewpoint.
Week 1 (9/7):
Reading: Chapters 1-3
Lectures: Introduction to spatial data and spatial correlation; spatial visualization.
Code: Chapters 2-3
Lab: Intro to Stop&Frisk – assignment into groups for project

Week 2 (9/14):
Reading: Chapter 4
Lectures: Spatial Heterogeneity (Do My Data Cluster?)
Code: Chapter 4
Lab: Explore Stop&Frisk data in groups with the goal of forming a research proposal

Week 3 (9/21):
Reading: Chapter 6
Lectures: Region-based analyses; Variogram models
Code: Chapter 6
Assignment due: research proposal for Stop&Frisk data
Lab: Explore Discuss/tweak research proposals as a class

Week 4 (9/28):
Reading: Chapter 7
Lectures: Kriging; Spatial Autoregression Models
Code: Chapter 7 (selections)
Assignment due: revised research proposal for Stop&Frisk data
Lab: Work on Stop&Frisk Data in groups – troubleshoot problems that arise

Week 5 (10/5):
Reading: Chapter 5
Lecture: NONE
Code: Chapter 5 (selections)
Assignment due: Independent research proposal (30% of grade)
Lab: Work on Stop&Frisk Data in groups – troubleshoot problems that arise

Week 6 (10/12):
Reading: NONE
Lecture: NONE
Code: NONE
Lab: Work on Stop&Frisk Data in groups – INDEPENDENTLY

Week 7 (10/19):
Reading: Chapter 8
Lecture: NONE
Code: Chapter 8 (selections)
Assignment due: draft research report for Stop&Frisk data

Assignment due October 26: FINAL research report for Stop&Frisk data (50% of grade; completed in groups)