Spatial Data Analysis with PySAL and GeoDaSpace

Sergio J. Rey and Luc Anselin

GeoDa Center for Geospatial Analysis and Computation
School of Geographical Sciences and Urban Planning
Arizona State University

Tutorial Proposal
NARSC 2014

Description

A unique feature of this tutorial is the use of Python based software tools for spatial data analysis. Python is an object oriented scripting language that is gaining rapid adoption in the computational sciences. To facilitate this adoption within the GIScience community, Rey and Anselin have collaborated on the creation of PySAL: Python Library for Spatial Analysis. Since its initial release in July 2010, PySAL has been downloaded over 50,000 times. This two-part tutorial will provide participants with an introduction to Python and related tools for spatial and regional analysis. In the second part of the tutorial participants will learn version 1.8 of PySAL as well as GeoDaSpace, a GUI application based on PySAL designed for spatial econometric analysis. The first component provides hands-on experience in the use of Python PySAL for exploratory spatial data analysis. In the second part of the tutorial the focus moves to practical spatial regression analysis using GeoDaSpace.

Instructors

Part 1: Introduction to Python and related tools - **Sergio J. Rey**
Part 2: PySAL and Spatial Regression/GeoDaSpace - **Sergio J. Rey and Luc Anselin**

Sergio J. Rey is Professor of Geographical Sciences and core faculty member of the GeoDa Center of Geospatial Analysis and Computation at Arizona State University. Prior to joining ASU he was Chair and Professor of Geography at San Diego State University where he was also the founding director of the Regional Analysis

Laboratory (REGAL) and co-director of the Center for Earth Systems Analysis Research (CESAR). He is a Fellow of the Spatial Econometrics Association.

Rey is project director for PySAL and the creator of the open source package STARS: Space-Time Analysis of Regional Systems. His research interests are in the area of geocomputation, exploratory space-time data analysis, spatial econometrics and urban/regional modeling with substantive applications to problems in regional economics, criminology, epidemiology, spatial demography and urban dynamics, among others. Rey has served as editor of the *International Regional Science Review* since 1999, editor of *Geographical Analysis* (July 2014) and as an editorial board member of *Computers, Environment and Urban systems, Geographical Analysis, Papers in Regional Science, Professional Geographer, Region et Devloppement, and Spatial Demography.*

Luc Anselin is Regents' Professor, Walter Isard Chair, and Director of the School of Geographical Sciences and Urban Planning at ASU where he also founded and directs the GeoDa Center for Geospatial Analysis and Computation. Prior to joining ASU, he held positions at The Ohio State University, the University of California, Santa Barbara, West Virginia University, the University of Texas, Dallas and the University of Illinois, Urbana-Champaign. Anselin is a member of the U.S. National Academy of Sciences and the American Academy of Arts and Sciences, a Fellow of the Regional Science Association International (RSAI) and of the Spatial Econometrics Association, and was awarded the Isard Prize and the Alonso Prize by RSAI.

Anselin is the creator of SpaceStat and GeoDA , co-founder of PySAL and directed the development of GeoDaSpace. His research interests are in the broad area of spatial data analysis, ranging from geovisualization to spatial and space-time econometrics. For over twenty years, he has taught workshops on spatial data analysis and spatial econometrics for a wide range of organizations, including CSISS, the World Bank, the Wharton School, as well as ten years at the University of Michigan's summer school of the Interuniversity Consortium for Social and Political Research (ICPSR). He has served as an editor of the *International Regional Science Review* since 1996 and is an editorial board member for several journals in regional science and quantitative geography.

Format

The tutorial is planned for a full day, broken into two half-day sessions. Each session is organized into two 80-minute components separated by a 20 minute break.

Objectives

This tutorial will offer participants the following:
- Introduction to Python for spatial data and regional analysis
- Introduction to PySAL for exploratory spatial data analysis
- Learning GeoDaSpace for spatial regression analysis

Outline

AM Session

1. Software and Tools Installation (80 min)
   a. Python Distributions
      i. Enthought Python
      ii. Anaconda Python
   b. Editors
   c. Terminals/Shells
   d. Cloud based environments
2. Python Primer (80 min)
   a. Data structures
   b. Control and Iteration
   c. Functions and Modules
   d. Files

PM Session

3. PySAL for ESDA (80 min)
   a. PySAL Overview and Setup (10 min)
   b. Spatial data processing with PySAL (40 min)
      i. Processing spatial data with PySAL
      ii. Spatial weights in PySAL
   c. ESDA with PySAL (30 min)
      i. Global spatial autocorrelation analysis
      ii. Local spatial autocorrelation analysis
4. GeoDaSpace for Spatial Regression (80 min)
   a. GeoDaSpace and PySAL Overview (10 min)
   b. Classic regression (40 min)
      i. Basic model setup
      ii. Estimation (OLS and 2SLS)
      iii. Diagnostics for spatial effects in OLS and 2SLS
   c. Basic spatial regression (30 min)
      i. Spatial error model
      ii. Spatial lag model

Audience

Regional scientists, GIScientists, researchers and students interested in learning Python and using PySAL for computational scripting in spatial analysis as well as GeoDaSpace for spatial econometrics.

Prerequisites

This tutorial is geared towards individuals who have a basic understanding of exploratory spatial data analysis and spatial regression.

Required Materials

- The tutorial will consist of computationally-based instruction. This will include basic scientific scripting (PySAL) as well as the use of a menu driven package (GeoDaSpace).
- Participants are encouraged to bring their own laptops. We will cover installation of all required software in the morning session.
- Software Requirements
  - Software will be made available on site.
- Sample data sets will be provided with the packages.

Expected Number of Participants

Based on previous experience, the tutorial should be capped at 25.