3600 Morrissey Hall Access Code: 1553

Course philosophy

* The result of using the scientific method is NOT absolute truth.
* More interested in your position and opinion rather than just the correct answer.
* Quality of work is more important that arbitrary due dates.
* Course synthesizes statistics and demography.

Trends

* A lot of traditional social science assumes that space doesn’t matter.
* Spatial demography is becoming computational.

Spatial demography

* Creating maps is a prerequisite for developing the spatial model.

Statistics concepts

* Goal is to create dependent variables that are interval or ratio whenever possible.
  + Allows use of the most sophisticated spatial statistics (i.e., high statistical power).
* Distributions
  + Methods are based on normal distribution
  + Real world data is rarely normally distributed
  + Income distribution is an example of positively skewed data (right skewness)
  + High school education is an example of negatively skewed data (left skewness)
* Ideal situation
  + Normal distribution
  + Small standard deviation
  + 68-95-99 rule
* Errors
  + Type I error is error of commission
  + Type II error is error of omission
  + Type II error is preferred to Type I error
  + Type I errors lead to retractions of journal articles
  + U.S. legal system
    - H0: person is NOT guilty (not the same as innocent)
      * Evidence is insufficient to reject the null hypothesis
    - HA: person is guilty
* Generally, cannot achieve a BLUE model.

Demography concepts

* The scientific study of human populations
* Elements of demography
  + Mathematical knowledge of populations
  + General movement of populations
  + The physical, civil, intellectual, and moral state of populations

The Importance of Space

* Y indicates dependent variable
* X indicates independent variable
* The goal is integrating space as part of the model
* More interdisciplinary research
* Many disciplines are afraid of integration of social sciences

Software and Infrastructure

* ArcMap will be going away; will be replaced by ArcPro and ArcGIS online
* QGIS is open source

Spatial Perspectives

* Functional distance
  + Time and effort required to move from point A to point B
  + Often must create additional data
* Spatial position
  + Comparative advantage and disadvantage of one space to another space
  + e.g., location of new MLS stadium
    - St. Charles residents will be the primary users
    - St. Louis City residents will absorb the costs
* Spatial order
  + Spatial hierarchy within a region
  + Relevant to class and stratification

Spatial Data

* Geodatabases are easier to use than shapefiles
* Quality of shapefile is important
* Can produce errors in spatial analysis
* Dr. Sandoval recommends obtaining shapefiles from U.S. Census

Data Sources

* U.S. Census ([www.census.gov](http://www.census.gov))
* Social Explorer ([www.socialexplorer.com](http://www.socialexplorer.com))

Homework

* Choose an MSA region other than St. Louis MSA or choose variables for St. Louis MSA that are different from the lab.
* Dependent variable should be interval/ratio or dichotomous.

Lab

* Download data from SocialExplorer.com
  + Treat the category Hispanic or Latino as a race
  + Download for files STATA
    - Tab delimited data
    - STATA .dct file (i.e., dictionary for how to read text files)
    - STATA .do file (i.e., program)
    - Data dictionary .txt file
  + Download for RStudio
    - Comma separated data .csv file
    - Data dictionary .txt file
* Creating ratio variables using STATA
  + Double click .do file to open it
  + \* used to add comments
  + Add folder file path in front of .txt filename in code
  + Add code before infile line code
  + Add line code capture log close
  + Add line code set more off
  + Add line code clear
  + Add line code log using analysis.log, replace after infile line code
  + Copy code for calculating ratio variables from lab .do file on Blackboard
  + Paste code for calculating ratio variable into .do file from data download
  + Change folder file path in
    - line code for exporting to Excel file
    - line code for creating new STATA database
  + Run program
    - Control-A to highlight the program
    - Click Execute(do) menu button
  + Creates new files in lab01/data folder
    - part01.xls
    - stl\_part01.dta
* Download base shapefiles
  + Download shapefiles from U.S. Census using TIGER/line shapefiles database
    - 2019 ACS 5-year estimate
  + Export ZIP files into lab01/data subfolders named for each shapefile
    - Illinois Census tract 17
    - Missouri Census tract 29
    - Entire USA
* Create new shapefile in ArcGIS
  + Open new map in ArcMap
  + Add each shapefile as a layer
    - Select layers for Illinois and Missouri to activate them
    - Deactivate layer for Entire USA
  + Create new file geodatabase in lab01 folder
    - Right Click on folder 🡪 New 🡪 File Geodatabase
    - Rename geodatabase (e.g., lab01)
  + Merge Illinois and Missouri shapefiles
    - Geoprocessing 🡪 Merge
    - Drag and drop Illinois and Missouri layers
    - Set output file path to newly created file geodatabase
    - Enter name for new shapefile (e.g., states01)
      * Use feature class file type
  + Select Entire USA shapefile layer to activate it
    - Deactivate other layers
  + Export St. Louis MSA shapefile
    - Select tract
    - Right click on layer 🡪 Data 🡪 Export Data
    - Export to file geodatabase
    - Name shapefile (e.g., stl\_msa)
      * Use file and personal geodatabase feature class file type
  + Clip census track shapefile to St. Louis MSA shapefile
    - Select stl\_msa shapefile layer
    - Select states01 shapefile layer with census tracts
    - Geoprocessing 🡪 Clip
    - Input feature is states01
    - Clip feature is stl\_msa
    - Output to file geodatabase
    - Name shapefile (e.g., stl\_msa\_ct\_00)
      * Use feature class file type
      * GEOID is unique to each case
  + Merge ratio variables data created using STATA into shapefile for St. Louis MSA
    - Select St. Louis MSA census tracts shapefile layer to activate it
      * Deactivate all other layers
    - Add ratio variables data as a new layer
      * Click on Add Data button
      * Select part01.xls file
      * Select Sheet1$
    - Right click on St. Louis MSA census tracts shapefile layer
      * Join and Relates 🡪 Join
      * Choose GEOID as the basis for the shapefile layer
      * Choose Sheet1$ as the table
      * Use FIPS as the basis for the table
  + Convert to permanent shapefile
    - Right click on St. Louis MSA census tracts shapefile layer
    - Data 🡪 Export Data
    - Output to file geodatabase
    - Name shapefile (st\_msa\_ct\_01)