

**Massachusetts Institute of Technology
Department of Urban Studies and Planning**

11.188: Urban Planning and Social Science Laboratory

11.205: Introduction to Spatial Analysis

In-Lab TEST - April 6, 2020

Test Instructions

- This test starts at the 2:35 PM. You 24 hours to finish, even though it is the same test for which we had planned 2.5 hours in lab. Any Stellar submissions tagged after 2:30 pm on April 7 will be marked as late. Please do not stay up all night working on the test. The 24 hours is needed to accommodate your classmates spread across the globe in many time zones. Given all that is happening (and P/NE grading this semester), do not lose sleep because of this test.
- This is an open-book, open-note test. However, do not contact any "live" non-staff person via electronic or other means while you take this test. We will be in our zoom classroom during classtime. If you have questions, we can address them privately in a zoom breakout room. Otherwise, you can post them on piazza or email the two of us.
- Please create your answers in electronic format - just as you have for lab assignments - and save your work on your personal computer (or on a USB drive, one-drive, etc.). You can edit your answers directly into your copy of this test if you have an HTML editor. For your convenience we have also provided MS-Word and PDF versions of the test in case you find these formats easier to use.
- Even where we don't require an explanation, you may want to say a few words about what you did so that **you can get partial credit** if some of your numerical answers and names are incorrect but your method has merit.
- Turn in your test (both the text and the PDF exports of your maps) by uploading your answer documents to Stellar (just as you have done for homework). If you wish, you may open the answer sheet in MS-Word and include your PDF images within your answer sheet as a single document.
- *Remember to turn in your test before logging out and check to be sure that Stellar has uploaded it successfully. We strongly recommend that you retain the file containing your answers until we return the graded exams to you.*
- Make sure to include your name and Athena user ID near the beginning of each document that you turn in.
- *Finally, don't spend all your time on one or two questions. Start by looking over all the questions to get a general sense of what, and how much, is being asked, and the grade points associated with each question--then start work on your answers. Move on to the next question if you've spent more than 10 minutes on any short answer question; mapping questions will likely take longer.*

Good luck!

Datasets for the Test

For this test, we will use US county-level data from the November 2016, US presidential election (between Donald Trump and Hillary Clinton).. The relevant data for the test is bundled together in a

folder called **test20data** in the AFS class locker. The folder is also compressed into a 'zip' file called, **test20data.zip**, in our '11.188 2020 DATA' Dropbox folder. Copy **test20data.zip** to a writable space on your local drive (e.g., C:\TEMP) and extract all its files into a local folder: C:\temp\test20data.

Look into your local copy of **test20data** for an ArcMap document called **11.188_test20_start.mxd** that has already included some of the shapefiles and tables that you will need. This ArcMap Document utilizes one new Metropolitan Statistical Area shapefile from the US Census plus several shapefiles that were obtained from the MIT Library Geodata Repository and then projected to a North_America_Albers_Equal_Area_Conic projection (NAD 1983) projection (in meters). (The standard EPSG code for this coordinate system is 102008 so I have included it in the shapefile names, but you will not need to use this fact.) For each of these shapefile, only the contiguous 48 United States were saved (the "lower 48").

Filename	Description
us_a1city_2000_lower48_epsg102008.shp	Point feature shapefile showing the location of major US cities (within the lower 48 states) having 2000 US Census population of at least 50,000.
us_f7states_2006_lower48_epsg102008.shp	Polygon feature shapefile showing state boundaries for the lower 48 states.
us_e25msa_2010_lower48_epsg102008.shp	Polygon feature shapefile from US Census Bureau of the US 2010 Census boundaries of metropolitan statistical areas (MSAs). The original shapefile has been limited to the lower 48 states and projected to the EPSG=102008 coordinate system.
us_f7counties_1996_lower48_epsg102008.shp	Polygon feature shapefile from US Census Bureau of the 1996 County boundaries within the lower 48 states together with the state, county name, Federal Information Processing Standard code (FIPS), and estimated 1990 and 1996 population.
2016_US_County_Level_Presidential_Results.csv	(Optional) A comma-separate-value (CSV) text file containing the 2016 US Presidential Election results by County showing the votes for Donald Trump and Hillary Clinton, the total votes, the percentage of votes for these two candidates, the difference in votes, and the percentage point difference in their votes. The same data are duplicated within the personal geodatabase (below) in the election16_county table so you do not need to use this CSV file.
election16.mdb	An MS-Access database (that is accessible from ArcMap as a personal geodatabase) containing two key tables: (1) election16_county with county-level presidential election results and a few 5-year 2015 ACS census variables, and (2) election16_county_data_dictionary explaining the meaning of each field in election16_county . The rest of the tables (GDB...) are extra geodatabase tables that ArcMap utilizes.

Within the ArcMap document, **11.188_test20_start.mxd**, the several state, county, MSA, and city shapefiles have already been added. The counties are shown thematically using the default (Jenks Natural Breaks) classification broken into 10 categories based on the 1996 population of each county. The county boundaries are omitted in order to avoid map clutter and make the map more readable. Data are shown for the 3111 counties in the lower 48 states.

In addition to the shapefiles, spreadsheet, CSV file,, and ArcMap document, the **test20data** folder

also contains an MS-Access database, `election16.mdb`, with the `ELECTION16_COUNTY` table of county-level presidential election results along with selected census data extracted from the 5-year 2015 ACS census. For your information, the MS-Access database also contains a data-dictionary, `election16_county_meta`, explaining the columns in the `election16_county` table. Even if you do not have MS-Access installed on your local machine, you may connect to these tables from ArcMap.

The 2016 election dataset was obtained from https://github.com/tonmcg/County_Level_Election_Results_12-16. Portions of the CSV file with election data were then imported into MS-Access and joined with ACS census data to produce `ELECTION16_COUNTY`. The 5-year 2015 ACS census data included with the election data come from the S1701 table on Poverty Status and the S2301 table on Employment and have been downloaded from the ACS website: <https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk>. In addition, we have added population counts for years 2000 and 2010 to the MSA shapefile, `us_e25msa_2010_lower48_epsg102008.shp`. These population data were obtained from an Excel spreadsheet downloaded from: <https://www.census.gov/population/www/cen2010/cph-t/cph-t-7.html>. [In order to join the population data for Los Angeles with the most similar shapefile polygon, we had to change the CBSA_code for Los Angeles-Long_Beach-Santa_Ana from 31100 to 31080 (Los_Angeles-Long_Beach-Anaheim). This editing issue is not relevant for the test.]

Part I: Short Queries (32 Points)

Question I-1 (28 points total, 4 points each part) - *Queries of election16_county*

- Which county had the largest number of votes for Donald Trump? `fips` = _____, County name = _____.
- How many votes did Trump get in that county: _____? How many votes did Clinton get in that county: _____?
- How many counties are there in Michigan? _____
- How many of the Michigan counties favored Clinton over Trump? _____
- Among all Michigan counties, which county had the largest vote difference (positive or negative) between Trump and Clinton? What was the vote difference in this county? `diff` = _____, `fips` = _____.
- What is the smallest total number of votes received by Hillary Clinton in any one of the 48 US states (plus the District of Columbia)? _____ State = _____?
- How many votes did Trump receive in that state? _____ State = _____?

Question I-2 (4 points total, 2 points each) - *Queries of us_alcity_2000_lower48_epsg102008.shp*

- How many cities in the shapefile have a reported population (POP) greater than 250000: _____?
 - What is the 8th largest city: Name = _____, POP = _____?
-

Part II: Mapping Concepts (6 Points)

Question II-1 (6 points total, 3 points each)

When you open the map document, the county map is displayed using a `North_America_Albers_Equal_Area_Conic` projection (NAD 1983) rather than in geographic latitude-longitude coordinates. As we know, local planning agencies typically use projected coordinate systems. For

example, the MassGIS mapfiles that we have used are generally saved in the Massachusetts (mainland) State Plane coordinate system.

Part II-1a (3 points): Suppose we were to view the same US County map in latitude-longitude coordinates. Would the size of Washington state (in the northwest corner) appear to be larger, smaller, or the same relative to the size of Florida? Explain Briefly. [Note: Feel free to change the coordinate system to see the results. However, there is no need to change the coordinate system and submit a map. You need only provide a brief answer and explanation.]

Part II-1b (3 points): Now suppose we were to view the same US County map in the Massachusetts (mainland) State Plane projection. Would the map be rotated clockwise or counter-clockwise relative to the original orientation when you opened the ArcMap document? Explain Briefly. [Notes: there is no need to change the coordinate system and submit a map. You need only provide a brief answer and explanation.]

Part III: Election Map (30 Points)

Let's develop a choropleth map showing the percentage of votes for Donald Trump in each county. Add the `election16_county` table to ArcMap, join it to the shapefile of 'lower-48' counties and design your map. Note that `election16_county` lists the county FIPS code as text (field=`fips_txt`) as well as an integer (field=`combined_fips`). We added the `fips_txt` field for your convenience. It is not in the original CSV file. Be careful to select the appropriate data type when you join the table to your shapefile.

There are 3139 records in `election16_county` table and 3111 records in `us_f7counties_1996_lower48_epsg102008.shp`. Note, that `election16_county` data is missing from a few counties (such as Dade County, FL, `fips12025`) and includes counties in states like Alaska and Hawaii that are not part of the 'lower 48'). So every row in `election16_county` will *not* match a row in `us_f7counties_1996_lower48_epsg102008.shp` and vice versa. After you join `election16_county` to the shapefile, you should see 3105 non null values for the election results among the lower-48 counties.

In addition, please notice that `election16_county` has many columns, most of which are not needed for the test. We have made these available just in case you want to explore this data beyond the test. This may be interesting for some of you, depending on your research interests.

Question III-1 (14 points total)

MAP #1: Prepare and submit an ArcMap layout of the counties within the lower 48 states and shade the counties based on the percentage of votes that were for Donald Trump. Be sure to:

- Include the outline of the States, and the usual scale bar, source, legend, north arrow, and an appropriate title,
- Indicate your choice of classification method,
- Overlay a proportional point map of major cities (in the `us_alcity_2000_lower48_epsg102008` shapefile) with the size of the city 'point' proportional to the population of the city. Only include cities that have POP >- 250000,
- Adjust color ramps, choice of background and foreground colors and outline, etc. to improve the readability of the map

NOTE: This question is worth 15 points and provides an opportunity to demonstrate the cartographic skill that you have developed.

Question III-2 (6 points total)

Part III-2a (3 points): Briefly discuss your choice of classification scheme and number of categories for your election results map.

Part III-2b (3 points): Briefly discuss any spatial pattern that you observe in your map regarding Trump's percentage of votes among the US counties.

Question III-3 (10 points total)

Part III-3a (2 points): What is average (mean) value of the percentage of votes for Trump (`per_gop`) among all those 'lower 48' counties for which you have the voting results joined to your map: _____?

Part III-3b (4 points): Select those counties that contain the cities that have $POP \geq 250000$. How many counties are selected: _____? What is the average (mean) value of the percentage of votes for Trump (`per_gop`) among these counties with the big cities: _____?

Part III-3c (4 points): Using the county-level ACS census data joined to your county shapefile, determine the average (mean) percentage of persons below the poverty level (`pct_below_pov`) for those counties containing the larger cities _____ and for those counties outside the larger cities _____.

Part IV: Proximity to Big Cities and Metro Areas (32 Points)

Next, let's examine the presidential voting pattern for counties in and around the larger cities.

Question IV-1 (6 points total)

Create a buffer of 50 km radius around the cities with $POP \geq 250000$. Select all the counties that intersect your buffer. In particular, select the counties that have their centroid within your buffered large cities.

Part IV-1a (4 points): How many counties fall within your buffered large cities: _____? If, instead, your overlap criteria were having any or all of the county area within your buffer, how many counties would satisfy that criterion: _____?

Part IV-1b (2 points): Among those counties with centroids within the large city buffer, what is the total number of votes for each candidate (i.e., the sum of `votes_gop`, and the sum of `votes_dem`):
Trump votes: _____, and Clinton votes: _____?

Question IV-2 (14 points total)

Part IV-2a (4 points): Based on the 2010 MSA population (`pop2010`) field in the `us_e25msa_2010_lower48_epsg102008` shapefile, select those Metropolitan Statistical Areas that had a population of more than one million persons. How many MSAs met this criterion: _____? Which MSA in this million-plus group had the smallest population: `fips_txt = _____`? and

Name = _____ ?

Part IV-2b (2 points): Next, select those 'lower-48' counties that *have their centroid* within those MSAs that have a population of one million or more. How many counties meet this criterion:
_____ ?

Part IV-2c (2 points): Among these counties that *have their centroid* within the larger MSAs, what is the average (mean) value of the percentage of votes for Trump (per_gop): _____? and for Clinton (per_dem): _____?

Part IV-2d (2 points): Among those counties that *have their centroid* within the larger MSAs, what is the total number of votes for each candidate (i.e., the sum of votes_gop, and the sum of votes_dem):
Trump votes: _____, and Clinton votes: _____?

Part IV-2e (4 points): Upon examining the last two questions, we notice that the average vote percentage (in the counties within the larger MSAs) is higher for Trump, but Part IV-2d shows that Clinton earned more votes in total among those counties within the larger MSAs. Explain briefly what is going on that allows this to occur:

Question IV-3 (12 points)

MAP #2: Prepare and submit a second ArcMap layout of the counties within the lower 48 states and once again shade the counties based on the percentage of votes that were for Donald Trump. This time, be sure to:

- Include the outline of the States, and the usual scale bar, source, legend, north arrow, and an appropriate title,
- Indicate your choice of classification method,
- Overlay your 50 km buffer of the larger cities with enough transparency to allow the choropleth map to be visible underneath.
- Include only those counties that you associated with the larger (million-plus population) MSAs and shade them using a hash pattern so the choropleth map is visible underneath.
- Adjust color ramps, choice of background and foreground colors and outline, etc. to improve the readability of the map
- Instead of showing all 48 states, zoom in to an area in the mid-west in order to see the details more clearly in that region. For this purpose, consider the mid-west to (at least) include all of Illinois, Ohio, Kentucky, and West Virginia.

That's all for the test, but feel free to keep the election data and explore some of the election patterns further when you have time. We do not have time on the test to examine further the relationship between voting, demographics, and health exchange participation and subsidies.

Please note:

- Before creating the PDFs, be sure to include your name and Athena ID somewhere on the text file and the map layouts
- You should test your PDF files with Adobe Acrobat Reader before submitting them to Stellar in order to be sure they are readable.

- You should keep a copy of your text output and PDFs file in your Athena locker until your test is graded
 - You should confirm with an instructor that your test files were received.
-

Last modified 31 March 20020. (jf)

Back to the [11.188 Home Page](#).

Back to the [CRON Home Page](#).
