YES Spatial Analysts Team: Learning outcomes discussion

Eric Darsow, Computer department faculty member, Community College of Allegheny County

Youth Enrichment Services Pb (lead) Exposure Project | Summer 2017

The following table describes the core learning areas explored during the Lincoln-Lemington Pb (lead) exposure project coordinated by Youth Enrichment Services (YES) based in East Liberty, Pittsburgh, PA. The far right column labeled “Relative Priority” is a subjective assessment of the degree of energy invested and growth achieved in several learning areas over the course of the project.

Note that this table’s data is drawn from my experience working with only the four mapping students for 5-8 hours per week during July and August, 2017. Their work with me was part of a larger research project involving close to 30 students also divided into groups, all expertly coordinated by YES staff. Hence, the relative priority “scores” should not be viewed as an assessment of the progress of the larger YES research expedition but rather a way to sense where energies and resources were allocated within the mapping subgroup. The column “Future Growth Areas” describes ideas for a learning trajectory in each area were the project to be continued.

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Academic Research Process			**########**
	Internalization of core research question	*Core question: How does understanding the built and social environments of Lincoln-Lemington increase our understanding of Pb exposure in young children?* Students could express this question in their own words and provide examples of features of the built and social environments. With guidance, students could revisit this question during various work phases and discuss how the chunk of data they were working with might lend insight into the question.	The research question was determined by the Pb Exposure Project Leaders prior to the beginning of my work with the mappers. Students would probably have internalized a research question more holistically if they construct it themselves with skilled guidance. (Note: Given the constraints of this project, a pre-constructed research question made eminent sense.)	######

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Research Method design	Students brainstormed specific practices to use during their field research endeavors, such as the way they were planning to convert a one-hour count of folks going into the corner store into an estimated total daily use measurement.	Due to time pressures, students were not actively considering alternative research methods and choosing the most appropriate for the question at hand. With more time, this important part of the research process could receive more resources.	##
	Literature review	The students and I “looked over” a few academic studies relating to neighborhood health measures and Pb exposure published in journals. We discussed the author’s use of side-by-side maps to illustrate relative intensities of various measures in the same geographic extent.	A more thorough literature review involving Pb exposure research would support a more informed decision making process regarding map creation and symbolization. Students would benefit from careful guidance in the areas of source assessment and academic search best practices to facilitate a more thorough and relevant literature review.	###

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Analyzing field data and drawing conclusions/ making recommendations	Core analysis tools used by students: * **Graduated symbol size and graduated symbol coloring** followed by visual pattern analysis. Using this method, the students concluded, for example, that the age of housing stock in Lincoln-Lemington is substantially aging, with most units built prior to 1978—the year Pb was banned from non-commercial paint. * **Kernel density smoothing ("heat mapping")** with experimentally optimized kernel weighting radius values. Students visualized adult and child use patterns in parks and two retail establishments in Lincoln-Lemington and concluded that adult frequent use is clustered around the stores and child use is clustered around playgrounds. * **Choropleth symbolization** of US Population census data (Census File 1). Students produced a map of Census blocks in Lincoln-Lemington and used the block boundaries to create a choropleth map of of child population by census block. When point-locations of the parks in the neighborhood were overlayed on the Census data, students drew general conclusions about park proximity to high density block of children.	Depth of analysis can be gained by increasing the number of observations made during the public facility use enumeration step (counting people going into stores, parks, etc.). Since the students were using primarily data they gathered themselves, record volume was low, and the geospatial analysis didn't add particularly useful insights to the patterns that existed (more adults go to stores and more kids go to parks). Since meaningful analysis only occurs following data collection and compilation, time devoted to this process felt squeezed by firm project deadlines, such as the YES Symposium. With increased observation volume and a few more weeks of work time, students can iterate on analytical ideas by applying several analysis tools to the maps and carefully reviewing the results.	#####

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Documenting and presenting findings to an audience	Each of the four students on the team created and presented an individual academic project poster at the YES Community Symposium on 17 August, 2017. The team collectively presented a slide deck to an audience of a few hundred people—with great reception.	With an expanded project time line, students can practice their final presentations with their instructors a few more times to grow more comfortable explaining the technical content fluently.	#####
	Field Research			**####**
	Data schema creation *(data schema = a plan for collecting and storing data, often schemas are represented as a data table)*	Students created and adjusted several types of data schemas (tables) during the project: group member information schema, community gathering places schema, human use frequency schema, research zone schema. Students were able to revise their initial method of collecting and organizing property condition survey data after initial attempts to use the tool with other project members. Their experience using a tool in "the field" followed by an adjustment before another return to the field allowed students to sense the value of iterative design based on field testing.	As the volume and specificity of data increases, the demands of the data schemas used increases. Students can benefit from working with a larger data set which requires more systematization to clean and process. For example, the students designed a google spreadsheet for recording lot-level condition data. Training of other students doing field assessments of properties was limited to the most basic of instructions. Students would benefit from serving as the data gathering coordinators who review field gathered data and provide feedback to the data gathering team.	##########
	Collecting data in field	The group collectively conducted a public site use frequency study by counting and visually classifying folks who visited four high-use areas in Lincoln-Lemington. One group member also conducted a playground equipment condition census.	Due to time and environmental conditions, field data gathering time was limited. It would be very useful to build in time for the mapping facilitator to dedicate time to join students in the field to provide real-time guidance as the systems are internalized.	#######

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Data entry and cleaning	Field data was transferred from paper notes and google spreadsheet cells. Simple formulas were constructed to multiply hourly traffic counts into daily values. The spreadsheet was exported as a comma separated value (CSV) file and loaded into the attribute tables in QGIS for visualization. Very little cleaning was required because the volume of data was so low.	A clear and valuable next step would be to use a tabular join to connect flat file data with shapes on the map—instead of entering field data into the attribute table directly. Higher volumes of data would require a more robust transfer method and would therefore be great experience for students who wish to continue engaging in field research.	####
	Technical tools			**########**
	Human and physical mapping fundamentals	Students internalized the concept of features on a map existing as points, lines, or polygons. They worked fluently with data layers.	Since we were working on both research methods and technical skills, some core mapping ideas remained to be taught: spatial references, raster vs. vector data, tabular joins, scaling principles, geo processing algorithm basics, and file format knowledge.	#####
	Navigating geographic information system software	The students became fluent at a decent range of basic digital mapping procedures in the QGIS package, including: panning, zooming, and setting spatial bookmarks, adding layers from data files, exporting and importing data as ESRI Shapefles, adjusting properties and symbology on each layer, editing layers in tabular form.	The QGIS package includes literally hundreds of individual functions and tools and many remain to be internalized. The tools the students did learn were selected based on the project needs at that stage in the investigation. Careful planning can be done to layer on more tool competencies as analytical needs increase.	#########

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Digitizing features of a paper map and attaching attribute fields to features	Students used a paper map of Lincoln's four research zones as a guide for digitizing the zone boundaries in QGIS. This required developing comfort with the half dozen or so digitizing tools (add feature, add/subtract vertex, move vertex, delete feature) and proficiency using the pan and zoom tools while in feature editing mode. Students revised their initial digitizing after a first round of field research when they realized that their initial zone boundaries were not precise enough to feed into a select by location tool for grabbing all the features in a particular zone. The team willingly revised their initial digitization to provide more accurate zone boundaries.	Digitizing has both artistic and scientific skills needs and students were well on their way to developing a "sense" about how to transfer lines on paper into digital shapes. They worked mostly with polygons and points whose location only needed to be approximate. Future growth involves digitizing features requiring a greater degree of accuracy or specific spatial relationships to other features.	##########
	Creating map layouts and choosing appropriate symbolization	Students used the print layout feature on QGIS to create printable maps with scale bars, compass roses, legends, and carefully positioned map frames. Students learned how to select layout elements from a list and manipulate the properties of those elements using the complicated properties dialog box for each element. By the conclusion of the program, students were assessing the features that create easy-to-read maps and improving their layouts to conform to those ideals.	Print layouts can become more complicated by including several map frames on the same layout.	##############

	Knowledge/Skill area	Description of student outcomes	Future growth areas	Relative Priority *less* *more*
	Auxiliary tool use	Auxiliary tools include: spreadsheets, internet data repositories, file explorers, and on-line cloud drives.	Developing competencies in spreadsheet use is the highest priority in this learning axis: creating exportable data schemas in the spreadsheet that easily are interpreted by the mapping software, generating formulas for repetitive calculations, editing data tables by adding rows and columns, and many more.	########
	Troubleshooting and debugging	Anybody working with complicated software packages like QGIS encounter bugs and issues that are not easy to be solved by "just clicking around." I helped students generate search phrases that would yield useful results via search engine requests when they got "stuck." Most of them could follow the steps listed in a tutorial that I identified as pertaining to their particular problem, but I did a lot of the driving of the troubleshooting process given our time and learning volume constraints.	Finding out how to get useful help on a technical issue is a high-level thinking activity because it requires a foundation of technical knowledge of a tool to determine if a possible solution is relevant to one's current issue. Students would be well served by developing this competency more explicitly: give students time to read tutorials and StackOverflow posts, try them out on the system, and evaluate the usefulness of the attempt at solving the core problem.	##