**Summary Notes**

* Techniques for minimizing cognitive burden
  + Sparklines are tiny charts that show data trends.
  + A small multiple (also trellis chart, lattice chart, grid chart, or panel chart) is a series of similar graphs or charts using the same scale and axes, allowing them to be easily compared to show data trends over time.
* Elements of data science
  + Statistics
  + Programming
  + Communication and Visualization
  + Substantive Knowledge
* Regarding the distinction between space vs. place, per Chris Prener refer to pages 508-509 of the Logan reading.
* Space (and spatial thinking) is the process/idea of locating something on the earth’s surface, and understanding that something’s location relative to other things.
* Place is the social meaning tide to particular phenomena that are spatially rooted.
* Example:
  + We can find the latitude and longitude of a pebble in a parking lot (i.e. define it spatially) but that does not given inherent social meaning.
  + Whereas we can talk about places - like SLU’s campus or Morrissey Hall - as having both a spatial definition and a particular meaning.

**Additional Questions of Interest**

1. None

**Summary Notes**

* As of today, we have access to Morrissey Hall 24 hours a day, 7 days a week.
  + GIS Mapping Lab access code is 1553.
  + If police ask just tell them you’re on the approved list for after-hours access; it they insist that you vacate, leave the premises and inform Chris Prener.
* Push to the cloud; pull from the cloud.
* More commits generally provides a better chain of evidence.
* Fewer commits results in a de facto black box.
* Rule of thumb is to over document and include narrative to explain program code.
* Use of file types
  + .Rmd files are used for performing actions.
  + .md files are what Chris Prener and Brandon Syracuse review.
  + .html files are used to preview the output.
* For next week
  + We’ll discuss types of data.
  + Per Chris Prener, don’t worry about Part 2 (reproducible examples) of Lab-01.

**Summary Notes**

* GitHub Desktop changes identified by color:
  + Green indicates new file
  + Red indicates deleted file
  + Yellow indicates modified file
* String data is more useful for mapping (i.e., ArcGIS); not useful for statistical analysis.
* Logical data
  + 0 indicates False
  + 1 indicates True
* ArcGIS uses the term “qualitative data” for categorical data.
* ArcGIS uses the term “quantitative data” for ordinal data.
* Raster data are actually images.
* Vector data is the building block of mapping (i.e., used the most in mapping applications).
* Spatial autocorrelation is when some aspect of the data is correlated to itself.
  + Relates to Tobler’s First Law of Geography.
* We obtain boundary-related data primarily from the U.S. Census (i.e., Tiger Database).
* We obtain topographical data primarily from the U.S. Geological Survey.
* When mapping quantitative data, normalize it by expressing it variable of interest as a percent or ratio of some uniform variable.
* ArcGIS requires short variable names (i.e., less than 8 characters).

**Summary Notes**

* You have to use data from an R package when creating a reprex.
* At the local government level, the most accurate data tends to come from agencies that collect money.
* Centroid is the geographic middle.
* Latitude and longitude are defined in decimal degrees.
* Mapping presents problems when moving between projection/coordinate systems.
* Use ArcGIS to move shapefiles rather than the computer system file manager to avoid missing components.
* Shapefiles generally are open standard.
* Geodatabases (.gdb) are generally closed standard.
* Handrails are elements of a map that help map readers orient themselves.
* [www.colorhexa.com](http://www.colorhexa.com) provides color hex values.
* R draws maps top to bottom as listed in the geom function.
* ArcMap dataframe is a group of layers that make up a single map.
* Use ArcCatalog to manage data and copy files.
* Use ArcMap to create refined maps.

Final Project Workgroup18-08

* Met via Google Hangouts on Sunday, February 18, 2018.
* Decided:
  + Mike Markee will take responsibility for team member 1 work package.
  + Malcolm Townes will take responsibility for team member 2 work package.
  + Eleanor Bergquist will take responsibility for team member 3 work package.

**Summary Notes**

* Graduated symbols are tools of last resort when creating maps.
* Color palettes are also called color ramps.
* Ground level is used to orient the map reader.
  + Use lower contrast/neutral colors.
* Figure level contains the important features of the map to be studied.
  + Use higher contrast/bright colors.
* Achieve visual contrast through changes in color and pattern.
* Use high value hues for ground layer.
* Use low value hues for figure layer.
* Warm colors rise to into the foreground.
* Cool colors recede into the background.
* Every RGB has corresponding hex value.
* DO NOT use rainbow or ggplot2 default color ramps.
* ColorBrewer and viridis have good color ramps for people with color impairment.

**Summary Notes**

* Per Chris Prener, annotated bibliography is due after spring break.
* GISc Public Policy discussion
  + Idea for using GIScc to encourage and facilitate public participation
    - How might we:  
      Use GISc to enable crowdsourcing of potential construction and economic development projects by have citizens identify new business establishments and services needed in their communities and neighborhoods.
* Most cartographers use Adobe Illustrator for fine grain control map development (e.g., labels, etc.)
* For poster maps, use PowerPoint or Keynote to make the legend.
* Make posters using PowerPoint or Keynote.
* 1 pt approx. 1/72 inch
* It’s better to make scale indicators in ArcGIS.
* It’s hard for readers to distinguish more than 5 classes on a map.
* The ArcGIS paradigm is for printed maps.
* For the final project, make the reference maps in ArcGIS.
* Map layouts in ArcGIS
  + Double click dataframe to get to the coordinate system.
  + Use the Layout View to make the map layout after you’re satisfied with the data.
  + Remove neat lines from the dataframes.

**Summary Notes**

* No class on Monday of Spring Break (i.e., 3/12/2018).
  + No lecture prep
  + No problem set
  + No policy discussion
* We DO have class the Monday of Easter Break (i.e., 4/2/2017)
* Character data (string data)
  + Missing data is NOT the same as empty data.
  + Recode empty data as missing data (NA).
* miss\_case\_summary identifies the observations with the most missing data.
  + Different path for arriving at much the same place as miss\_var\_summary.
* There are three (3) different types of Esri geodatabases
  + Personal 🡪 DO NOT USE
  + File 🡪 Use for class
  + Enterprise 🡪 DO NOT USE
* To convert a shapefile to a geodatabase
  + Export > To Geodatabase (single)…
  + Output feature class is simply the name of the feature class
  + This is an irrevocable process so its best practice to keep a copy of the raw shapefile
* Use geodatabases to perform analysis, not shapefiles
  + Geodatabases are more efficient, robust
* R functions
  + str\_detect() is case sensitive.
  + str\_replace() replaces only the first instance of a string in each observation.
  + str\_replace\_all() replaces all instances of a string in each observation.

**Summary Notes**

* NO CLASS – spring break
* Viewed lecture videos posted on YouTube.

**Summary Notes**

* There is less chance of making unintended changes to the data when you manipulate character data.
* If variables in a join are different types, it’s recommend that you change the numeric variable to a character variable.
* When joining data tables in R, it doesn’t matter if the data is sorted.
* Eliminate unnecessary variables in the data before joining to avoid data bloat.