

just few/brave ones: introduce yourself, and:

- - 1) what are you researching/analyzing?
 - 2) what data are you using?
 - 3) what do you expect from this workshop?

what is there?

- GIS: Geographic Information Systems
 - Geographic: Cities, Roads, Rivers, Countries, etc
 - Information Systems: data, software, programming,
- GIS=CS(graphics, database/sys adm, coding)+geography
- geographic=geospatial=spatial

past and future

- much of the gis has been (still is) done with ArcGIS/ArcMap
 - this is more of a dinosaur, however
- the future is open source software like qgis
- and internet companies like Google

rules

- i'll go slowly as computer skills likely vary a lot!
- do interrupt and ask questions
- help your neighbor! [will see how it goes, may add Straso]
- many slides have refs as urls; at the end i list more refs
- communicate outside of the classroom:
 - bother me, email everyone (emails in mass email)

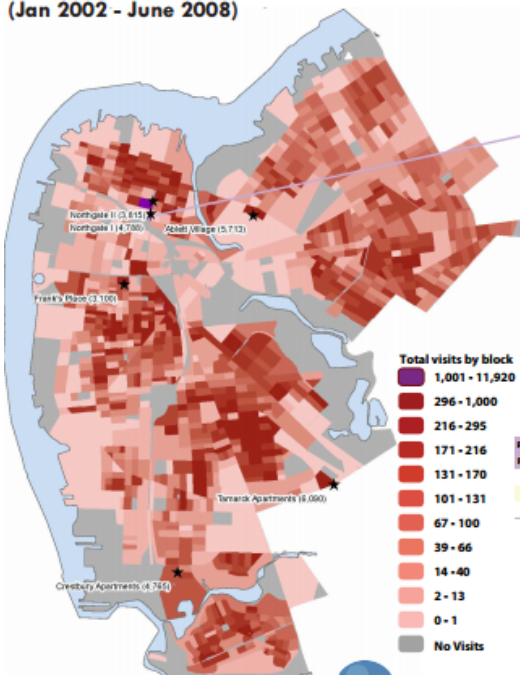
setup

- don't have much time: skip the intros
- we'll be mostly displaying data on maps using colors (thematic/choropleth)

why? discovery! just put it on a map

- Dick De Veaux: blackboard: US map with loc of faulty devices
- and Cooper's dr Brenner on next slide

Inpatient and Emergency Room Visits in Camden, NJ (Jan 2002 - June 2008)



Northgate I Public Housing

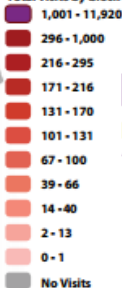


	Visits	Patients	Charges	Receipts	Collected
Cooper	3,172	749	\$42,144,897	\$4,994,658	12%
Loardies	811	337	\$7,848,889	\$1,028,611	13%
Virtua	805	331	\$1,742,467	\$345,092	20%
2005	838	370	\$10,834,420	\$1,269,373	12%
2006	738	355	\$6,967,995	\$683,549	13%
2007	790	369	\$7,979,262	\$901,181	11%
ED	3882	978	\$6,150,592	\$664,019	14%
Inpatient	906	408	\$45,584,781	\$5,504,342	12%
Total	4,788	1,070	\$51,735,374	\$6,368,361	12%

Primary Diagnosis

Rank	ED	Inpatient
1	abdominal pain (789.0)	live birth (V3X.0)
2	acute URI NOS (465.9)	chest pain (786.5)
3	chest pain (786.5)	congestive heart failure NOS (428.0)

Total visits by block



say you have housing prices

- the “traditional” (non-gis) data in excel from
<http://www.zillow.com/research/data/>
- reposted:
https://sites.google.com/site/adamokuliczkozaryn/gis_int/NJ-counties-Zillow-Home-Value-Index-TimeSeries.xls
- note: we have geography! county! this is our key to map!

geographic (map) data to match our spreadsheet

- now need to find map (geographic) data to match our spreadsheet
- let's search for what we need: NJ counties!
- just goog 'your geography' + 'shapefile'
= 'nj counties shapefile'
- reposted: <https://docs.google.com/uc?id=1xJDhcRCkgv7k4tNCa720og5bohV6dTB2&export=download>
- download it, unzip it, and put into project folder and keep there
- there are couple files, keep them intact
- don't rename, don't change location within project folder

ADJUST HERE TO UNZIPPING LOL

load data into qgis!

- first icon on the left “line with nodes”: nj_counties.shp
 - or just drag it over and drop
- can zoom in and out
 - either click the map with “+” “-” tool
 - or draw a rectangle to achieve appropriate zoom
- grab map and move around with “hand” pointer
- layer listed the left: right click-
 - Open Attribute Table [can select/highlight in table or map]
 - all right! we have county variable!
 - note how it looks like!
 - either upper or proper case (w 'county' string)

your spreadsheet and geo data must have same ID

- “Camden county” \neq “Camden”
- “Camden” \neq “CAMDEN”
- “08012” \neq “8012”

adjusting and cleaning up spreadsheets

- adjust ID: make counties uppercase
 - (or could drop 'County' from COUNTY LABEL variable)
- always clean up the spreadsheet:
 - one row header (I dropped first row)
 - make col (variable) names brief: say <5 alphanumeric chars
 - drop excessive columns you wont need, keep it clean
 - important! leave only plain numbers!
 - drop all special chars from vals: “#” “\$” “,” etc
- save as csv (just one sheet); reposted:
`https://sites.google.com/site/adamokuliczkozaryn/gis_int/all_homes.csv`
- note missing value! and save in project folder

references

- https://geodacenter.github.io/workbook/1_datascience/lab1.html
- just search for 'merge'
- merging in geoda <https://www.youtube.com/watch?v=6ihK4xVTl00>

joining (merging)

- start geoda by searching for it at the bottom-left
- input file: just navigate to nj counties
- Table-Merge: csv: all_homes.csv
- current table key: COUNTY
- import table key: UPPER
- hit ' ' to mv everything to 'Include'
- and hit Merge
- (accept proposed changes for var names)
- then hit table icon to have a look at the table and compare with input csv

now can map

- Map-Quantile Map-5: 'Dec2012'
- change color for 'undefined': right-click: Color for category...and pick say white
- can pick a basemap, say Carto Light
- right-click-Save image as: map1.png
- keep it open, can have many windows at the same time

and let's map POPDEN2010

- Map: Quantile Map: 5
- Map: Percentile Map
- Map: Equal Intervals: 5
- what differences do you see?
- lets discuss :)

explore more

- Explore-Scatter Plot
 - X: POPDEN2010
 - Y: Dec2012
- and click right most and top most points

- meh guess later! another example: columbus; see my old class on geoda!
- advertising: my class, contact info: send students, take my class, hire our students, hire me