theaok.github.io/shortGIS
very short GIS course: 1.5hrs in geoda

adam okulicz-kozaryn

adam.okulicz.kozaryn@gmail.com

this version: Tuesday 8th October, 2019 16:34

outline

just few/brave ones: introduce yourself:

- 1) what are you researching/analyzing?
- 2) what data are you using?
- 3) any GIS experience, any software?
- 4) 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 geoda or qgis
- and internet companies like Google

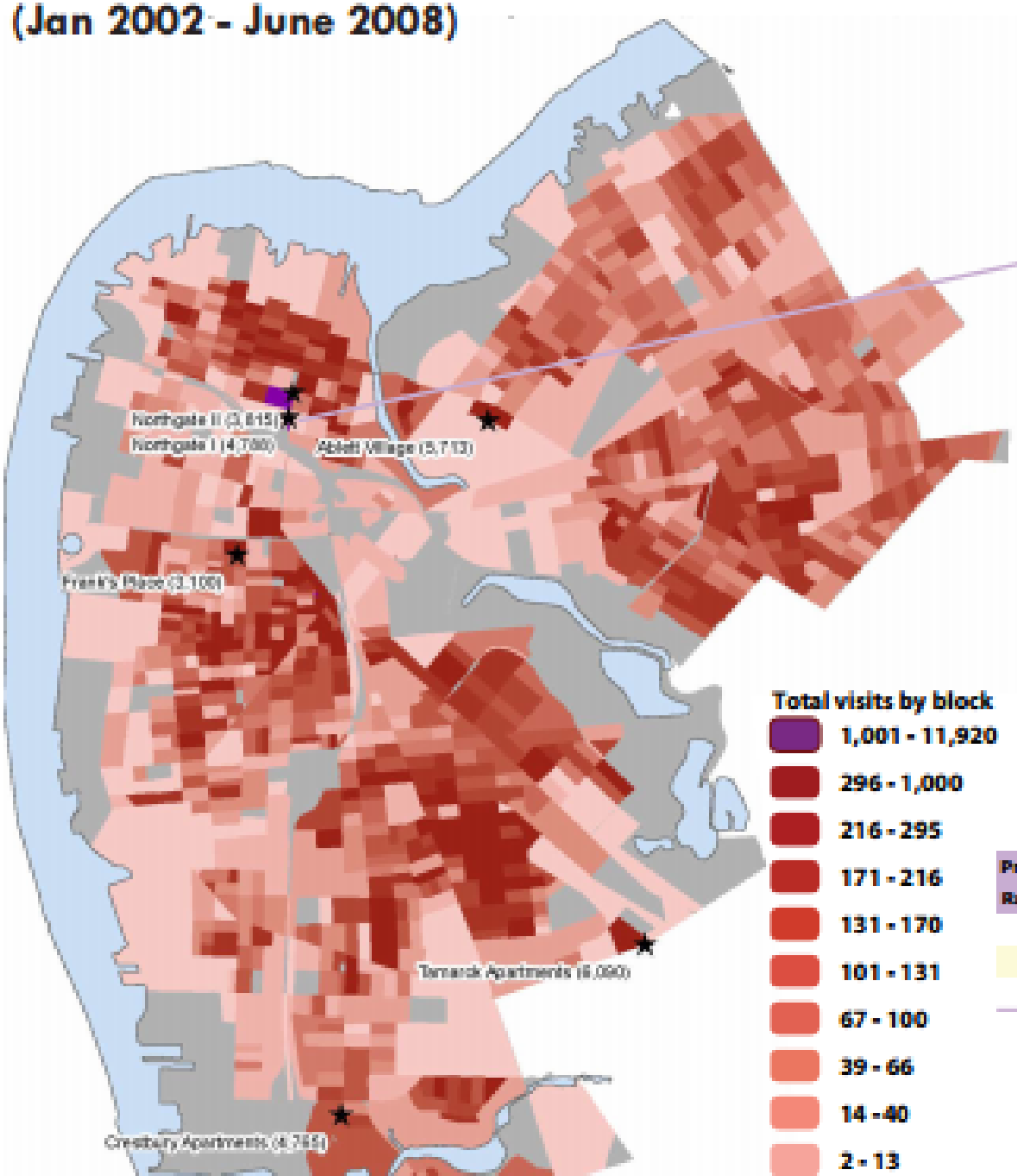
rules

- i'll go slowly as computer skills likely vary a lot!
- do interrupt and ask questions: i'll walk around and help
- help your neighbor!
- do email me after class with questions/comments, etc
- don't have much time: just get straight to it!
- display your own spreadsheet data on maps using colors (thematic/choropleth)
- why? discovery! just put it on a map

why? discovery! just put it on a map

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%
Lourdes	811	337	$7,848,839	$1,028,611	13%
Virtua	805	333	$1,742,467	$345,092	20%
2005	838	370	$10,834,420	$1,269,373	12%
2006	738	355	$8,867,995	$883,549	13%
2007	790	369	$7,979,262	$901,181	11%
ED	3882	978	$6,150,592	$864,019	14%
Inpatient	906	408	$45,584,781	$5,504,342	12%
Total	4,788	1,079	$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)

let's do it! say you have housing prices

- the "traditional" (non-gis) data in spreadsheet format
  `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=1xJDhcRCkgv7k4tNCa72Oog5bohV6dTB2&export=download`
- download it and unzip it
  - there are couple files, keep them intact and in one place
  - don't rename, don't change location within project folder

load shapefile into geoda

- start geoda by searching for it at the bottom-left
  - (no need to update to latest version)
- input file: ESRI Shapefile, navigate to nj counties (.shp)
- open attribute table

your spreadsheet and geo data must have same ID

- have csv and shapefile opened side by side
- “Camden county” $\neq$ “Camden”
- “Camden” $\neq$ “CAMDEN”
- “08012” $\neq$ “8012”
- so need to adjust ID: make counties uppercase
  - (or could drop ’County’ from COUNTY LABEL variable)

cleaning up spreadsheets

- almost always need to clean up the spreadsheet
  - one row header (I dropped first row)
  - make col (variable) names brief: say $<$10 alphanumeric chars

    drop excessive columns you wont need, keep it clean
  - important! leave only plain numbers!
  - drop all special chars from vals: "#" "$" "," etc
- download as csv (just one sheet); reposted:

  `https://sites.google.com/site/adamokuliczkozaryn/gis_int/all_homes.csv`
  - note missing value!

references

- https://geodacenter.github.io/workbook/1_datascience/lab1.html
  - just search for 'merge'
- merging in geoda https://www.youtube.com/watch?v=6ihK4xVTl00
  - and much more!
  - open source software like geoda has excellent online documentation

joining (merging)

- 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
  - important to always check your join

now can map

- Map-Quantile Map-5: 'Dec2012'
- change color for 'undefined': right-click: Color for category...and pick say white
- right-click and 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

the end

- advertising:)
- keep in touch, keep me posted
  - `https://theaok.github.io`
- see my full fledged class
  - `https://theaok.github.io/gis`
  - take my class, send students, hire our students
- don't waste money on ArcGIS/MAP
  - go with opensource: geoda, qgis, python