

[theaok.github.io/shortGIS](https://theaok.github.io/shortGIS)  
very short GIS course: 1.5hrs in geoda

adam okulicz-kozaryn  
[adam.okulicz.kozaryn@gmail.com](mailto:adam.okulicz.kozaryn@gmail.com)

this version: Tuesday 8<sup>th</sup> October, 2019 16:19

## 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

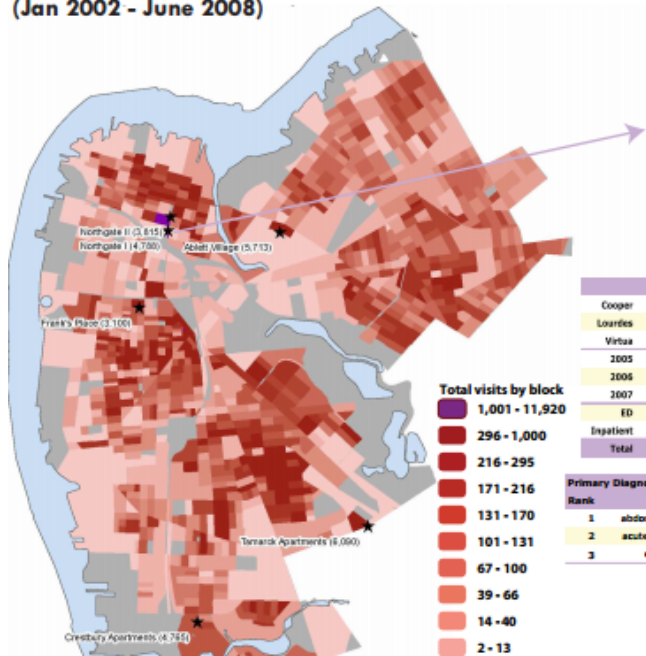
- much of the gis has been (still is) done with ArcGIS/ArcMap
  - this is more of a dinosaur, however
- the future is open source software like geoda or qgis
- and internet companies like Google

## rules

- i'll go slowly as computer skills likely vary a lot!
- do interrupt and ask questions: no student left behind
- help your neighbor!
- slides have refs as urls, see them if time/at home
- do email me after today 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%
Loardies	811	337	\$7,848,809	\$1,028,611	13%
Virbaa	805	331	\$1,742,467	\$345,092	20%
2005	838	370	\$10,834,420	\$1,268,373	12%
2006	738	355	\$6,867,995	\$881,549	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	1,070	\$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)



## 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](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 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, just navigate to nj counties (.shp)
- open attribute table

## 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](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](https://geodacenter.github.io/workbook/1_datascience/lab1.html)
- just search for 'merge'
- merging in geoda <https://www.youtube.com/watch?v=6ihK4xVT100>

## 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