NAACCR Webinar Series:
GIS Applications

May 6, 2010
9:00 am, Eastern Daylight Time
-- repeats at --
2:00 pm, Eastern Daylight Time

Questions

- Please use the Q&A panel to submit your questions
- Send questions to “All Panelists”

2

Fabulous Prizes

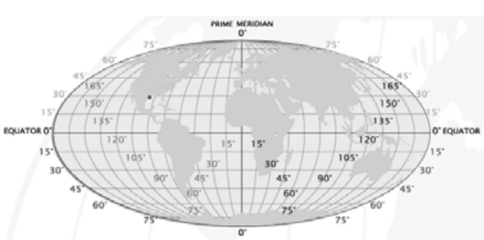
3

Topics

1. Geography basics
2. Geographic Information Systems (GIS) basics
3. Geocoding
   a. Individual records
   b. Batch geocoding
4. Keyhole Markup Language (KML)
5. Geographic accessibility
6. Geographic distances & time

1. Geography Basics

- Every point on the earth's surface can be described in terms of a *latitude* and *longitude*
  – Northern hemisphere = + latitude
  – Western hemisphere = - longitude



Galveston, TX

•29 degrees latitude

•-95 degrees longitude

Degrees

- Degrees/minutes/seconds
  - Example 30°30'0''
    - 60 seconds in a minute
    - 60 minutes in a degree
- Decimal degrees *(NAACCR standard)*
  - Example 30.500000°

How large is a degree?

- The metric system originally defined the *meter* as one-ten-millionth of the distance from the equator to the pole
  - Thus 90° of latitude=10 million meters, so 1° of latitude=111.11 km, about 69 miles
  - 1° of longitude is smaller than 1° of latitude (related by cosine function)

Latitude/longitude is customarily reported to six decimal places (NAACCR standard)

36.124579°, -75.034678°

69 miles
6.9 miles
0.69 miles
360 feet
36 feet
3.6 feet
4 inches

Important census subdivisions (US)

- 50 states + DC + territories
  - California 36 million, Wyoming 0.5 million, Palau 20,000
- 3,143 counties or county equivalents
  - Los Angeles 10 million, Loving County, TX 60
    - These change slightly over time
- 70 thousand census tracts
  - Average 4,000 people; most 2,500-8,000
    - These change every 10 years

Important census subdivisions (US)

- 200 thousand Block Groups
  - Subdivisions of census tracts, 1-9 per tract
  - Average population 1,500
- 8 million Blocks
  - Areas bounded by roads or water
  - Most have <50 people, many have 0

Other subdivisions of space

- ZIP codes
  - Census ZCTA (Zip Code Tabulation Area)
- Cities, towns, boroughs, villages
  - (Minor Civil Divisions)
- School districts
- Health districts
- Legislative districts
- Hospital service areas

Location

- Our data reflect the location of your residence at the time you get cancer
  - It says little about lifetime exposure and even recent exposure
  - "You are where you live"

2. GIS basics

*"Information systems used to manipulate, summarize, query, edit, visualize - generally, to work with information stored in computer databases." (*NCGIA)

"In the strictest sense, a GIS is a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information, i.e. data identified according to their locations. " (USGS)

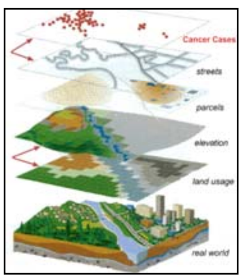
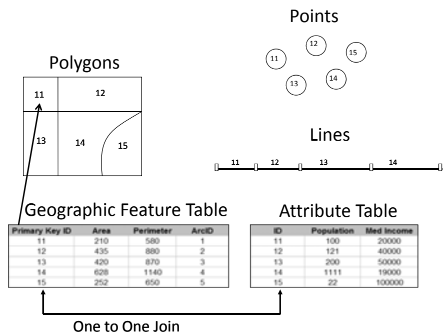


Image adapted from http://www.nysgis.state.ny.us/coordinationprogram/whatisgis.html

GIS software includes several components:

- Algorithms and methods for spatial operations (e.g. adjacency), querying, editing, and creating/capturing data, and visualization (e.g. overlay)
- Database management systems (e.g. Oracle, SQL, Access)
- Open software architecture allowing customization (e.g. Visual Basic, Java, Python)
- Supplemental tools for spatial analysis, spatial/geo statistics, network analysis, geocoding, geo-rectification

Representing Geographic Information



Primary Key ID	Area	Perimeter	ArcID
11	210	580	1
12	435	880	2
13	420	870	3
14	628	1140	4
15	252	650	5

ID	Population	Med Income
11	100	20000
12	121	40000
13	200	50000
14	1111	19000
15	22	100000

How is GIS useful for Epidemiology?

- Displays and stores data at various geographic scales
  - County, census tract, neighborhood
- Provides a systematic way to explore associations between health outcomes and various risk factors (environmental, social)
- Data from various sources and geographic scales can be combined to provide estimates of exposure or area-based characteristics

How is GIS Useful for Epidemiology?

- Analyzes spatial relationships (e.g. proximity of cases from treatment facility or screening facility)
- Provides tools for creating maps for communicating results

Why Map Cancer Data?

- Public and local cancer coalitions want information about local communities
- Public health officials (state, local federal) want to understand:
  - What areas are in need of intervention and/or prevention
  - Where to allocate financial resources
  - Track progress on cancer control and prevention activities

Why Map Cancer Data?

- Identify broad spatial patterns (descriptive)
- Hypothesis generation (cause of pattern)
- Define and compare patterns (health disparities)
- Changing patterns over time (cancer control)

Mapping software

- Online or stand-alone software products mostly for descriptive mapping
- Increasingly, online mapping products/services are adding additional methods analogous to GIS
  - Annotation
  - Data creation
  - Geocoding
  - Driving distances

Mapping software

- Freely available mapping tools
  - Google Earth
  - Microsoft: Virtual Earth; MapPoint
  - ESRI Arc Explorer

Google Earth

- Virtual globe, map and geographic information program
- Provides tools for viewing, editing, creating geographic layers, geocoding, and obtaining driving directions

Google Earth

- Internal coordinate system is geographic coordinates (latitude/longitude) on the World Geodetic System of 1984 datum
- Baseline resolutions for continental U.S. of 1 m
- JavaScript API
- Free

Google Earth

3a. Geocoding

- Geocoding (verb): act of transforming aspatial locationally descriptive text into a valid spatial representation using a predefined process.
  - Manual or automated process
- Geocoder (noun): a set of inter-related components in the form of operations, algorithms, and data sources that produce a spatial representation for locationally descriptive text
- A geocode (noun) is a spatial representation of a descriptive locational reference.

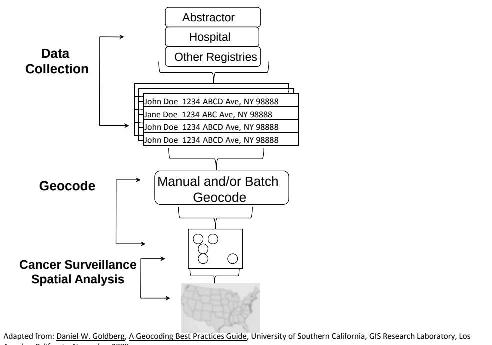
Source: Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008.

Geocoding

- The NAACCR GIS Committee edited a Geocoding Best Practices Guide in 2008

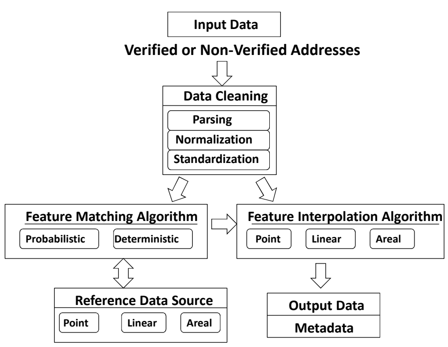
Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008. http://www.naaccr.org/filesystem/pdf/Geocoding_Best_Practices.pdf

Possible geocoding workflow at central registry



Adapted from: Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008.

Geocoding process details



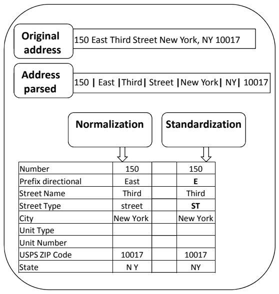
Adapted from: Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008.

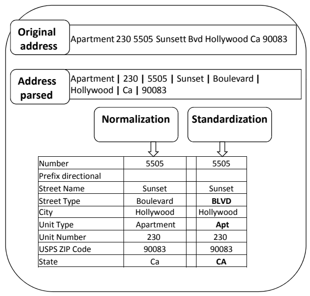
Geocoding: Input data types

Name	Type	Usage	Best/Worst output Resolution
The University of Southern California	Named place	County Counts	Parcel-level/ Non-matchable
The University of Southern California GIS Research Lab	Named Place	Screening, Disparities	Sub Parcel-level/ Non-matchable
The northeast corner of Vermont Avenue and 36th Place	Relative intersection	Screening, Disparities	Intersection-level/ Non-matchable
3620 South Vermont Ave, Los Angeles, CA 90089	Street Address	Screening, Disparities	Building-level/ Street-level
USPS ZIP Code 90089-0255	USPS ZIP Code	County Counts	Building-level/ USPS ZIP Code-level
34.022351, -118.291147	Geographic coordinates	Screening, Disparities	Sub Parcel-level/ Non-matchable

Source: Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008.

Data cleaning





Reference data source

- Reference dataset: geographic database containing features and address data that the geocoder uses to generate geographic outputs

Source	Description	Coverage
U.S. Census Bureau's TIGER/Line files (United States Census Bureau 2008c)	Street centerlines	US
NAVTEQ Streets,(NAVTEQ 2010)	Street centerlines	Worldwide
Tele Atlas Dynamap, MultiNet (2010)	Street centerlines	Worldwide
Los Angeles (LA) County Assessor Parcel Data (Los Angeles County Assessor 2008)	Vector polygon file	Los Angeles
US ZIP Code	Point	US
Canadian postal code	Point	Canada

Adapted from: Daniel W. Goldberg, A Geocoding Best Practices Guide, University of Southern California, GIS Research Laboratory, Los Angeles, California, November 2008.

Geocode Metadata

- Metadata: Descriptions associated with data that provide insight into its attributes.
  - Geocoder software/vendor
  - Dataset name (e.g TIGER/Line files)
  - Dataset year, version
  - Feature-matching strategy (e.g. deterministic or probabilistic)

Geocode Metadata

- Metadata: Descriptions associated with data that provide insight into its attributes(cont).
  - Dataset type (points, lines, polygons)
  - Feature interpolator strategy (e.g Address-range interpolation; lot interpolation)
  - Attribute matching weights
  - Date geocoded
  - Coordinate quality (NAACCR Standard)

NAACCR V11 Data Standards and Data Dictionary

GIS Coordinate Quality Codes	
00	Coordinates derived from local government-maintained address points, which are based on property parcel locations, not interpolation over a street segment's address range
01	Coordinates assigned by Global Positioning System (GPS)
02	Coordinates are exact match of house number and street, and based on property parcel location
03	Coordinates are exact match of house number and street, interpolated over the matching street segment's address range
04	Coordinates are street intersections
05	Coordinates are at midpoint of street segments (missing or invalid building number)
06	Coordinates are address ZIP code+4 centroid
07	Coordinates are address ZIP code+2 centroid

NAACCR V11 Data Standards and Data Dictionary

GIS Coordinate Quality Codes	
08	Coordinates were obtained manually by looking up a location on a paper or electronic map
09	Coordinates are address 5-digit ZIP code centroid
10	Coordinates are point ZIP code of Post Office Box or Rural Route
11	Coordinates are centroids of address city (where address ZIP code is unknown or invalid, and there are multiple ZIP codes for the city)
12	Coordinates are centroid of county
98	Latitude and longitude are assigned, but coordinate quality is unknown
99	Latitude and longitude are not assigned, but geocoding was attempted; unable to assign coordinates based on available information
Blank	GIS Coordinate Quality not coded

Exercise 1:
Using Google Earth to Geocode Individual Cases

Geocoding exercises with individual cases – Summary

- Accurate and precise
- Time-intensive
- Not practical for a large number of cases
- What's the best way to geocode thousands of cases all at once?

3b. Batch geocoding

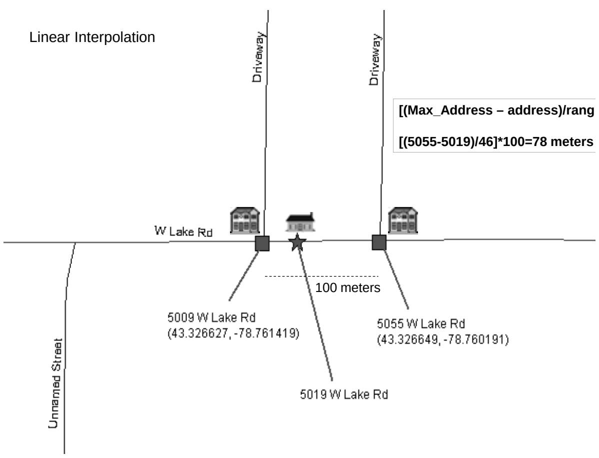
- A geocoding process that operates in an automated fashion and processes more than a single record.
- Traditionally, this has been done by interpolating along a street segment.

Geocode 5019 W Lake Rd, Wilson NY 14172

ID	637913638
FromLeft	5009
ToLeft	5055
FromRight	5008
ToRight	5054
Prefix	W
Name	LAKE
Type	RD
Suffix	
ZIPLeft	14172
ZIPRight	14172

StateLeft	36
StateRight	36
CountyLeft	063
CountyRight	063
TractLeft	0243.02
TractRight	0243.02
BlockLeft	1000
BlockRight	1002
FromLat	43.326627
ToLat	43.326649
FromLon	-78.761419
ToLon	-78.760191

Linear Interpolation



Probabilistic versus deterministic approaches to match to a street segment

Deterministic matching

- The deterministic approach requires an *exact match* on all fields being compared
- Different passes can be made that compare different subsets of fields

Geocode 28 LAURE PL, Spring Valley, NY 10977

- *Pass 1*: Match on number, name, type, and ZIP code
  - Result: no matches
- *Pass 2*: Match on number, name, and ZIP code only
  - Result: no matches

Goal: Geocode 28 LAURE PL, Spring Valley, NY 10977

- *Pass 3*: Match on number, name soundex, type, and ZIP code
  - Result: one match – 28 LAURA PL

  (Soundex is a function that considers most of the consonants and ignores the vowels in a string of text. The soundex of Laure is L600).

  http://en.wikipedia.org/wiki/Soundex

Geocode 28 LAURE ST, Spring Valley, NY 10977

- *Pass 1*: Match on number, name, type, and ZIP code
  – Result: no matches
- *Pass 2*: Match on number, name, and ZIP code only
  – Result: no matches
- *Pass 3*: Match on number, name sndx, type, and ZIP code
  – Result: no matches
- *Pass 4*: Match on number, name sndx, and ZIP code
  – Result: 2 matches
    - 28 LAURA PL
    - 28 LAURA LN

When there are no matches, or 2 or more equal matches, the results should be manually reviewed (the correct answer is 28 Laurel St)

Probabilistic matching

- Each field being compared has a weight.
- If the fields match, a score equal to the weight is assigned. If they do not match, the score is zero.
- The highest match above a certain score is the one taken.
- Can be done in a single pass, or multiple passes

Geocode 28 LAURE PL, Spring Valley, NY 10977

- Match on number, name, type, and ZIP code
  - Suppose the calculated weights are: number 6.9, name 7.5, type 1.5, ZIP code 6.2

Address	Score
28 LAURA PL, 10977	20.1 (highest)
128 LAURA PL, 10977	13.2
28 LAURA PL, 11213	13.9
43 RAMSEY PL, 12208	1.5

- Usually, weights are specific to each address
  - Yznaga has a higher weight than Maple because it is less common

- However, some geocoding software gives the same weight to every field

Probabilistic matching results in more matches…both correct and incorrect

Reasons For Non Matches

- Abbreviated & Misspelled addresses

Street Name Submitted to Registry	Correct Street Name
Bevrbrk	Beaverbrook
32 A Hadon Rd	32 Haddon, Apartment A
122 Stkhom Road	122 Stockholm Road
400 Prin-Lawre Road	400 Princeton Lawrenceville Road
23 A Green Street	23 Green Street, Apt A
Unit B, 1200 Comm Ave.	1200 Commonwealth Ave

- Apartment numbers or street addresses included in the building field (unit designations belong in a separate field).
- The address is missing one or more components (building number is missing, street type or direction is missing or incorrect).

Reasons For Non Matches

- Post office box is given rather than a street address.
- Rural route address is given.
- A building or facility name is given instead of an address (e.g., just the nursing home or apartment tower name is given).
- Missing, incorrect, retired or new Zip Codes (US Postal Bulletin reports changes)

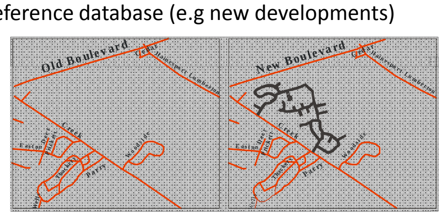
Old	332205	07801	NJ	Dover	Morris	Main Office	Post Office	7-1-95	Establish a new ZIP Code for a delivery area. Use Mine Hill NJ 07803 as last line of address for a portion of the deliveries previously in ZIP Code 07801.
New	332205	07803	NJ	Dover	Morris	Mine Hill	Place Name	7-1-95	

USPS Postal Bulletin, February, 1995

Postal Bulletin: http://www.usps.com/cpim/ftp/bulletin/archive.htm

Reasons for Non Matches

- Street name changes and missing streets from reference database (e.g new developments)



2002 2004

Geocoding software/services

Name	Free Limit/Cost	API	Reference Data	Type	Security
TeleAtlas EZ-Locate	100 free, 25,000-$410 - 100,000-$1,535	yes	TeleAtlas	Online	encrypted
ThinkGeo		yes	Tiger 2009	Online or PC-Based	encrypted
Spatial Insights	Setup fee: $2,500.00, monthly- $200.00-15,000	no		Online	encrypted
iBegin	100 free, $.01/geocode after that; min charge of $250	yes		Online	encrypted
TerraPages		yes	G-NAF	Online	
ArcGIS Desktop Geocoding	cost of ArcGIS license $1500	no	Tiger or GDT	PC-Based	local

Geocoding software/services

Name	Free Limit/Cost	API	Reference Data	Type	Security
IMapTools	$3000; $300 updates	no	Tiger	PC-Based	local
Geolytics	1 million/$1195 - unlimited $5975	yes		Online or PC-Based	local/encrypted
USC Geocoder	free -2500 batches; $.002 per geocode	yes	Tiger	Online	encrypted
Maps Online			Tiger 2000	SAS geocoder	local
Geocoder.us	50,000/day	yes	TIGER		unsecure
Yahoo	5,000/day	yes	Navteq		unsecure
Google	15,000/day	yes	TeleAtlas	GDT	Unsecure
Geobase		no	Tiger or GDT	PC-Based	local

Adapted from: https://webgis.usc.edu/Services/Geocode/About/GeocoderList.aspx

Exercise 2:
Batch Geocoding

Batch Geocoding

- Summary
  - Knowledge of record linkage theory is helpful
  - Before selecting a vendor evaluate geocode match accuracy and review available metadata
  - Some software and vendor solutions are expensive
  - Some cases will still require individual clerical review

4. Keyhole Markup Language (KML)

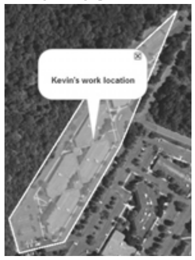
- An XML-based open-standard markup language for displaying geographic data and annotation on Internet-based Earth browsers such as Google Earth.
- KML is an open standard officially named the OpenGIS® KML Encoding Standard . It is maintained by the Open Geospatial Consortium, Inc.
- KML specifies a set of features (placemarks, images, polygons, 3D models, textual descriptions) in an Earth browser

KML example (placemark)

```
<?xml version="1.0" encoding="UTF-8"?>
<kml  xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
    <name>Simple placemark KML example</name>
    <description>
      Attached to the ground. Intelligently places itself
at the height of the underlying terrain.
    </description>
    <Point>
      <coordinates>-74.688705,
40.2552900</coordinates>
    </Point>
  </Placemark>
</kml>
```

KML example (polygon)

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
 <Document>
  <name> KmlFile</name>
  <Style id="transGreenPoly">
   <LineStyle>
    <width> 2.5 </width>
   </LineStyle>
   <PolyStyle>
    <color>7d00ff00</color>
   </PolyStyle>
  </Style>
  <Placemark>
   <name> Kevin's work location </name>
   <visibility> 1 </visibility>
   <styleUrl> #transGreenPoly </styleUrl>
   <Polygon>
    <tessellate> 1 </tessellate>
    <altitudeMode> relativeToGround </altitudeMode>
    <outerBoundaryIs>
     <LinearRing>
      <coordinates>
      -74.687419534125822,40.25720434943663,0
      -74.68915852154247,40.25597308390642,0
      -74.69063268048805,40.25416475220034,0
      -74.69044354892967,40.25366463956237,0
      -74.68957402002546,40.25382824263983,0
      -74.68850305962293,40.25523256980062,0
      -74.68741953412582,40.25720434943663,0   </coordinates>
     </LinearRing>
    </outerBoundaryIs>
   </Polygon>
  </Placemark>
 </Document>
</kml>
```

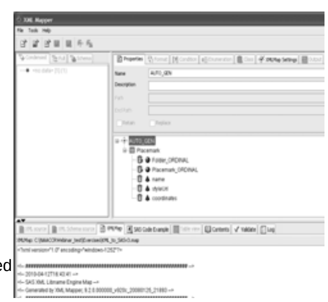


Exporting geographic data from SAS to KML

```
data _null_;
    file  'C:\NAACCRWebinar_test\studentdata\RIHOSPITALS.kml';
    set Rihospsitals end=last;
        if _n_ = 1 then do;
            put '<?xml version="1.0" encoding="UTF-8"?>';
            put '<kml xmlns="http://www.opengis.net/kml/2.2" xmlns:gx="http://www.google.com/kml/ext/2.2"
            xmlns:kml="http://www.opengis.net/kml/2.2" xmlns:atom="http://www.w3.org/2005/Atom">';
            put '<Document>';
        end;
 points = '<coordinates>'||trim(left(LONG2))||','||trim(left(LAT2))||',0 </coordinates>';
            put '<Placemark>';
            put '<Style>';
            put '<IconStyle>';
            put '<scale>1.2</scale>';
            put '<Icon>';
            put '<href>http://maps.google.com/mapfiles/kml/pal3/icon46.png</href>';
            put '</Icon>';
            put '<hotSpot x="32" y="1" xunits="pixels" yunits="pixels"/>';
            put '</IconStyle>';
            put '</Style>';
            put '<gx:balloonVisibility>1</gx:balloonVisibility>';
            put '    <Point>';
            put points;
            put ' </Point>';
            put ' </Placemark>';
        if last then do;
            put '</Document>';
            put '</kml>';
        end;
 run;quit;
```

Importing KML to SAS

- Drag-and drop SAS interface that generates syntax describing the tables and variables contained within an XML (KML) file
- Syntax is contained in SXLE file
- SXLE file needs to be called from SAS



SAS code calling SXLEMAP called KML_TO_SAS.map

```
filename  RIHOSPIT 'C:\NAACCRWebinar_test\studentdata\RIHOSPITALS_EDITS.kml';
filename  SXLEMAP 'C:\NAACCRWebinar_test\exercises\KML_TO_SAS.map';
libname   RIHOSPIT xml xmlmap=SXLEMAP access=READONLY;
```

Exercise 3:
Creating and Editing Geographic data in Google Earth

5. Geographic Accessibility

- Access: Ability to obtain appropriate health services when required
- Components of Access (Ricketts & Savitz, 1994) :
  - Availability
  - Accommodation
  - Affordability
  - Acceptability
  - Accessibility

Accessibility

- Relative ease by which health services can be reached, taking account of resources (e.g. transportation, physicians) and travel time, distance and cost.
- Types of Accessibility:
  - Revealed (or Realized)
    - Actual use of a service
  - Potential
    - Probable utilization of a service

Accessibility examples in cancer research

- Revealed:
  - For cancer patients increased travel distance/time has been shown to be associated with:
    - A greater risk of presenting with advanced stage cancer
    - Decreased utilization of breast-conserving therapy
    - Lower enrollment in clinical trials
- Potential:
  - Hayanda et al. (2009) found that each percentage point increase in the African American population within a county was associated with a decrease in the number of colorectal surgeons and gastroenterologists.

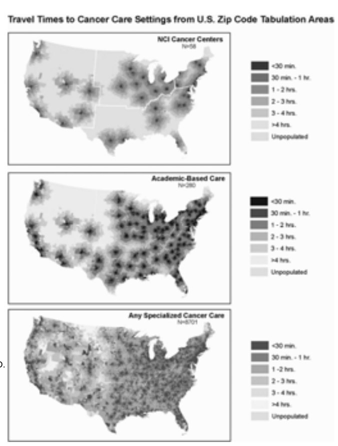
Racial clustering and access to colorectal surgeons, gastroenterologists, and radiation oncologists by African Americans and Asian Americans in the United States: a county-level data analysis.
Hayanga AJ, Waljee AK, Kaiser HE, Chang DC, Morris AM,Arch Surg. 2009 Jun;144(6):532-5.

Accessibility examples in cancer research

- Potential:
  - Onega et al (2008) found:
    - For the total continental U.S. population, the median travel time to the nearest NCI Cancer Center is 78 minutes.
    - Native Americans and nonurban residents have relatively longer travel times to NCI Cancer Center, Academic care and specialized cancer care.
    - Travel times of 1 hour were estimated for 45.2% of the population to the nearest NCI Cancer Center, 69.4% to the nearest academic-based care, and 91.8% to any specialized cancer care.

Geographic access to cancer care in the U.S. Onega T, Duell EJ, Shi X, Wang D, Demidenko E, Goodman D. Cancer. 2008 Feb 15;112(4):909-18.

- Onega et al (2008)



Geographic access to cancer care in the U.S.
Onega T, Duell EJ, Shi X, Wang D, Demidenko E, Goodman D.
Cancer. 2008 Feb 15;112(4):909-18.
c

Commonly used measures of geographic accessibility

1. Provider-to-population ratios:
   - Physicians per # of population
   - Does not include distance or time
   - Does not account for patient border crossing (e.g. county)

Commonly used measures of geographic accessibility

2. Travel distance or time to the closest service
3. Number of services within *n* miles or minutes
4. Average travel distances or times to all or *n* closest services
5. Floating catchment or Gravity-based methods
   - Account for supply & demand

6. Distance Measures

- Great Circle Distance (GCD)
  - *"Shortest distance between any two points on the surface of a sphere measured along a path on the surface of the sphere "*
  - More accurate than straight-line (linear) distance
  - Prior to SAS 9.2 Haversine formula used to calculate GCD
  - 9.2 uses Vincenty's formula (more precise than Haversine formula)

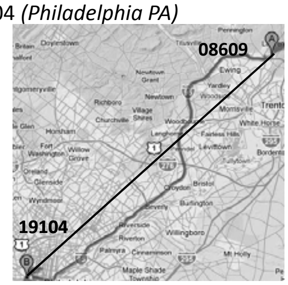
Distance Measures

- Driving Distance and Travel Time
  - Employ geometric network files with speed limits, road direction, and turn restrictions
  - Search algorithm necessary to find shortest path on a road network while minimizing the travel cost (e.g. Dijkstra algorithm)

Distances from two locations

- 08609 (*Hamilton, NJ) to* 19104 *(Philadelphia PA)*

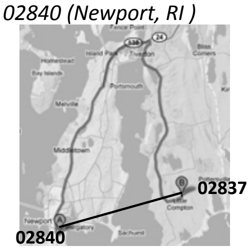
•HAVERSINE	30.08 MILES
•GEODIST	29.97 MILES
•DRIVING DISTANCE	46 MILES
•DRIVING TIME	49 MINUTES



Distances from two locations

- 02837 *(*Little Compton, RI) *to 02840 (Newport, RI)*

•HAVERSINE	7.61 MILES
•GEODIST	7.58 MILES
•DRIVING DISTANCE	25 MILES
•DRIVING TIME	42 MINUTES



Resources to calculate distances

- NAACCR Great Circle Distance Calculator
  - A SAS program that calculates the GCD distance between the locations of cases at the time of diagnosis and the locations of treatment facilities.
  - Designed to be used with the NAACCR v10 or v11 record layout file
  - Can use either source (unconsolidated) or consolidated case records as input.

Resources to calculate distances

- NAACCR/Komen road network travel distance/time application
  - Calculates shortest driving distance and travel time from origin to destination for batches of records for the entire US and Canada
  - Source data: NAVTEQ street shapefiles
  - Web service allows users to upload encrypted data containing origin and destination latitude and longitude coordinates
  - Outputs shortest travel distance and travel time

Resources to calculate distances

- Google Maps
  - Obtain shortest driving distance and travel time from origin to destination.
  - Possible to obtain distance/travel time based on public transportation and walking routes.
  - Programs can be written to send and receive batches of records for the entire US and Canada
    - SAS or Python
  - *Data not encrypted—best for public use data
    - Distance from census block group to closest hospital

Exercise 4:
Calculating Distances Between Locations

Example:
Census Tract Centroids To Hospitals

NAACCR Central Registry Webinar:
GIS Applications

•Wrap-up
•Questions??

Next Month…

- Collecting Cancer Data: Esophagus and Stomach
  - June 3,2010

…July Webinar

- Using CINA Data in Cancer Surveillance Activities
  - July 8, 2010

81

2010-2011 Webinar Series

- Available for purchase
- https://www.regonline.com/naaccr_webinar_series_2010_2011