GIS for OHDSI

- Forming workgroup to propel development
- Geographic variables such as environmental exposures and distance to care can be important health predictors. We aim to add this functionality to the OHDSI set of tools
- Current project name is *Gaia* (greek personification of Earth)
- Written in R (open to alternatives)

Where does it fit?

The project looks to contribute to OHDSI analysis in two ways:

1. Data exploration through visualization
   “Are there hotspots for diagnosis x?”

2. Construction of geospatial data from existing CDM locations
   “For each location, how many grocery stores are within 10 minutes travel time?”

Work so far...

1 Data Preparation

Manipulate CDM
-add lat/lon columns

Develop R package to:

- Clean / Standardize location data
- Geocode:
  - Address to coordinates
  - Coordinates to GEOID

2 Data Extraction

Extract and aggregate data from the CDM into region tables

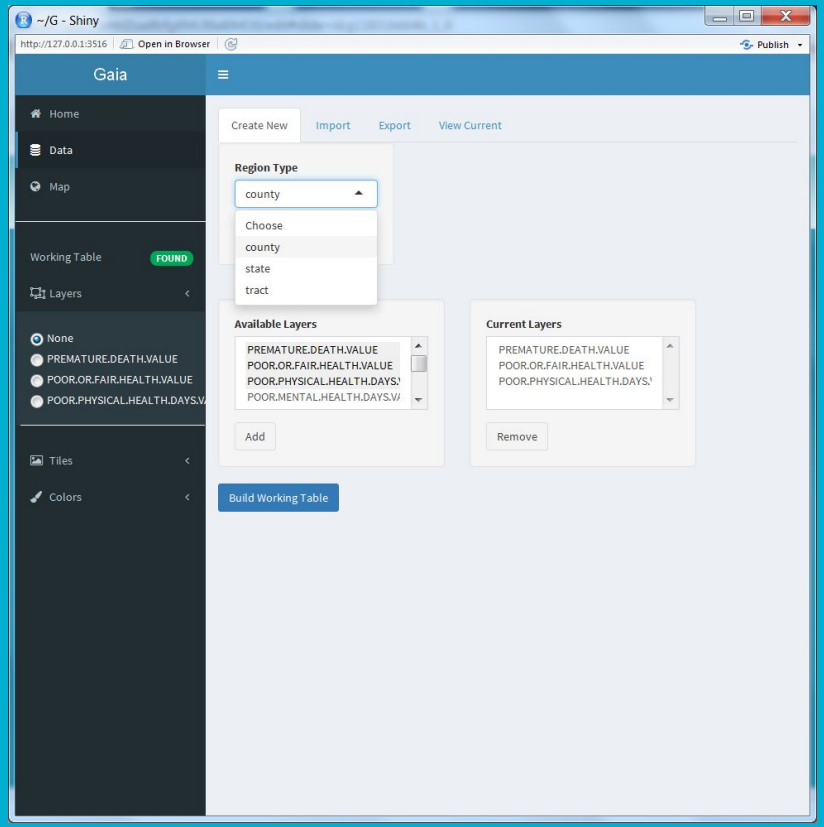
- R package
- Access CDM directly
- Cohort as input
- e.g. Density of diabetes diagnosis for each county

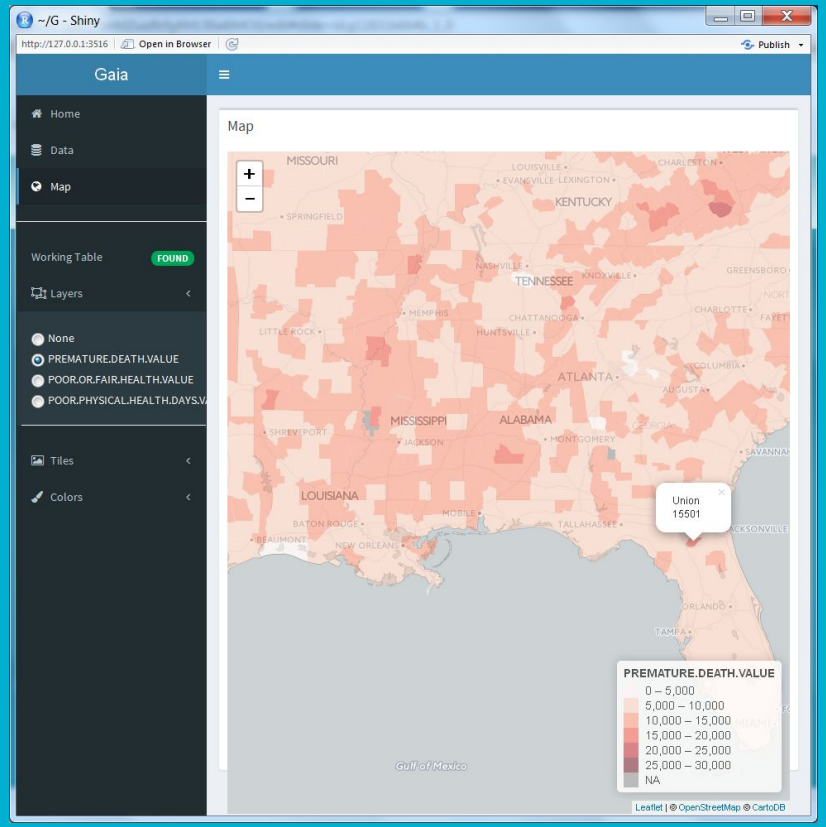
3 Visualization

Mapping tool prototype

- R - Shiny/ Leaflet
- Build dynamically from aggregate region tables

Unknown whether this is needed or existing mapping software (e.g. Qgis, Geoserver) is sufficient


~/G - Shiny
http://127.0.0.1:3516
Open in Browser
Publish
Gaia
Home
Data
Map
Working Table
FOUND
Layers
None
PREMATURE.DEATH.VALUE
POOR.OR.FAIR.HEALTH.VALUE
POOR.PHYSICAL.HEALTH.DAYS.V
Tiles
Colors
Create New
Import
Export
View Current
Region Type
county
Choose
county
state
tract
Available Layers
PREMATURE.DEATH.VALUE
POOR.OR.FAIR.HEALTH.VALUE
POOR.PHYSICAL.HEALTH.DAYS.
POOR.MENTAL.HEALTH.DAYS.V
Add
Current Layers
PREMATURE.DEATH.VALUE
POOR.OR.FAIR.HEALTH.VALUE
POOR.PHYSICAL.HEALTH.DAYS.
Remove
Build Working Table


~/G - Shiny
http://127.0.0.1:3516
Open in Browser
Publish
Gaia
Home
Data
Map
Working Table
FOUND
Layers
None
PREMATURE.DEATH.VALUE
POOR.OR.FAIR.HEALTH.VALUE
POOR.PHYSICAL.HEALTH.DAYS.V
Tiles
Colors
Map
MISSOURI
KENTUCKY
TENNESSEE
ATLANTA
MISSISSIPPI
ALABAMA
LOUISIANA
Union
15501
PREMATURE.DEATH.VALUE
0 – 5,000
5,000 – 10,000
10,000 – 15,000
15,000 – 20,000
20,000 – 25,000
25,000 – 30,000
NA
Gulf of Mexico
Leaflet | © OpenStreetMap © CartoDB

Problems to be tackled

1. **Where and how does the data fit together?**
   Polygon data, extracted/aggregate data, geospatial variables, geographic identifiers, etc

2. **Geocoding**
   How do we want to implement standardization?
   Which geocoding services do we want to support? Confidentiality concerns?

3. **Data Extraction**
   How can we leverage other portions of OHDSI?
   How should the user specify the data to be extracted and method of aggregation?

4. **Mapping / Visualization Software**
   Should we continue to develop our own? (is the customization worth it?)
   Does it make more sense to utilize existing software? If so, what is the best fit?

Interested?

Contact Information

Email :

robert.miller@maine.edu

OHDSI Forum User:

rtmill

See you at the symposium!