Where's best?

A Geospatial Data Digestion Framework Prototyped in R with Shiny

Chao Shi

04/2017

Motivation

- Strengths of the R + Shiny ecosystem
- Personal appreciation of the power of maps
- Interest in decision making process
- Always try to solve a *type of* problem, record logical thinking, and develop reusable code

A few thoughts before we move on

- Too much data to handle — when facing fast growing ocean of data, user-friendly digestion solutions will be in high demand, from corporations to consumers.

- When a majority of service providers have superb reviews, consumers may turn to other more objective measurement to help making decisions

How do we make these decisions

- where to eat
- where to buy
- where to live

filtering system + score generator

Project Design, Scope and Deliverables

- Find a spatial grid complex enough in geometry, familiar enough for the targeted audience, while small enough in variable size for one laptop to handle

- Find 3 data sets (columns) to begin with

- Develop a data filtering and ranking engine in R

- Make proof-of-concept first version

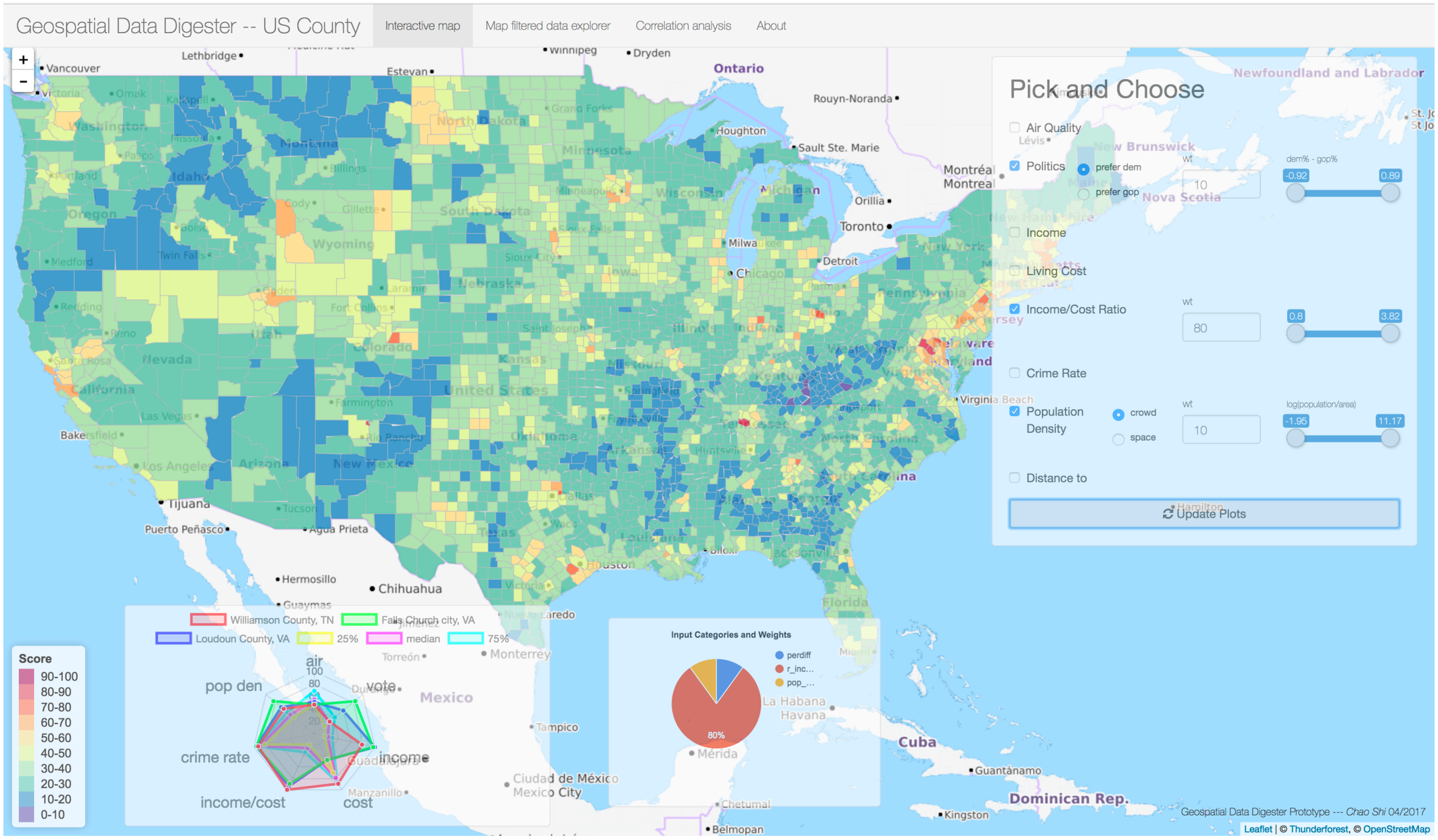
- Add data sets to enrich the complexity of the spatial decision making example

- Beautify UI with efficiency

- Document things along the way, make presentation and publish code

Live Demo

Key components



Near Future Steps

- A few features to develop before making a NYC where to buy app
    - categorical data
    - distance calculation -> travel time calculation

What have we built so far?

A flow chart like cartoon

The Future

- Imagine this as a service, how to make profit?
    - What kind of established companies might be interested?
    - This as a web start-up? What's the break even case?
- Scale up consideration
- Data with gaps — interpolation is rarely just a math problem
- Decision making based on data streams, real time digestion before they disappear

Acknowledgement

- NYC Data Science: Shu Yan, Zeyu Zhang
- Inspiration from the ‘SuperZip’ example by Joe Cheng
- Data cleaning example on datasciencenceriot.com
- Correlation Matrix app ‘shinyCorrplot’ by saurfang

Data Source

This is one small part of the documentation

- This talk — high level intro and discussion
- Blog post — expansion of the presentation
- Readme files on Github
- Comments in source code