Spatial Analysis using QGIS and Other Applications

Felipe Gonzalez

University of California, Berkeley

Overview

1. Introduction to QGIS
   i. Loading files
   ii. Understanding files
   iii. Creating shapefiles
   iv. Personalizing maps
2. Applications using QGIS
   i. Creating maps
   ii. Digitalization of maps
   iii. Raster data to polygons
   iv. Shapefiles from Shapefiles
3. Other applications
   i. ~~Buffers~~
   ii. Network Analysis
   iii. Dijkstra's algorithm

Overview

1. Introduction to QGIS
2. Applications using QGIS
3. Other applications

Introduction to QGIS

- Download QGIS from `www.qgis.com`
- QGIS is a free software with a friendly interface that helps you perform spatial analysis
- Interface:
  1. Layers
  2. Map window
  3. Horizontal menu (shortcuts)
  4. Vertical menu (shortcuts)
  5. Menu bar

Loading Files

- QGIS reads shapefiles (`.shp`): Polygons, Lines, Points
- QGIS can also read rasters (e.g. `.tiff`)
- You can find shapefiles and rasters for free in many places in the internet. One example is `www.diva-gis.org`
- To open a new file go to `Layer/Add Vector Layer` to add a new shapefile, or `Layer/Add Raster Layer` to add a raster
- The layer you added should now appear in the Layers window.

Understanding Files

- You can add as many layers you want
- To see the data behind each layer `Open Attribute Table`
  - Each row is either a polygon, a point, or a line
  - Columns are attributes of rows

Creating Shapefiles

- You can also create your own shapefiles:
  - Polygons
  - Lines
  - Points
- Suggestion: keep your newly created shapefiles in a different folder to avoid confusion with existing shapefiles.

Personalizing Maps

- Double click on the layer's name to modify map's appearance
- You can do many many things
  - Change colors
  - Add names
  - Change styles
  - Give color to polygons using attributes
  - etc...

Notes

- Your QGIS window is now a “project”, a combination of shapefiles and your added preferences
- You should save this project and it will stored in your chosen folder as `NAME.qgs`

Overview

1. Introduction to QGIS
2. Applications using QGIS
3. Other applications

Applications using QGIS

1. Creating maps
2. Digitalization of maps
3. Raster data to polygons
4. Shapefiles from Shapefiles

Creating Maps

- We have a survey of Mexican immigrants in the United States and we want to know the spatial distribution of these individuals
  - We want want to make sure there are observations in the Midwest
- First, we need to assign each observation in the survey a geographic location. This can be done using the command `geocode` in Stata
  - You don't need to look for many locations in google maps, this is what the command `geocode` does automatically for you

Do-file to geo-reference locations

```
1  use aux/cities.dta, clear
2
3  * Creates variable to use in geocode
4  g dir = city + "," + " " + state + "," + " " + "United ...
       States"
5
6  * Geocodes cities to latitude and longitude
7  geocode3, address(dir)
8  keep jobplace g_lat g_lon
9
10 * Save
11 outsheet using aux/Cities_Geo.csv, comma replace
```

Creating Maps

- Second, we need a shapefile of the U.S. We can download one from DIVA (country, states, counties)
- Always check that you have the right shapefile by looking at the date it was created and checking basic statistics (e.g. number of states)

Figure: Geographic distribution of observations in dataset

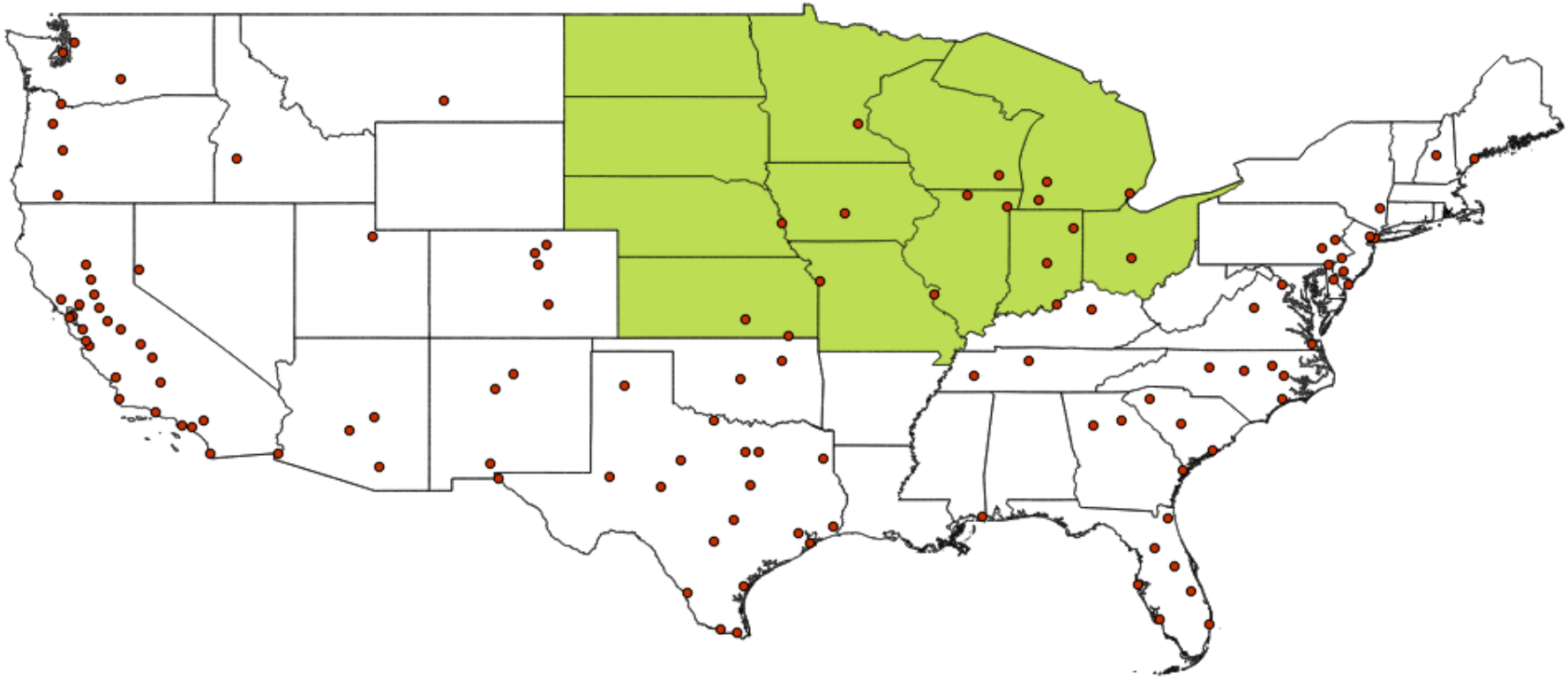
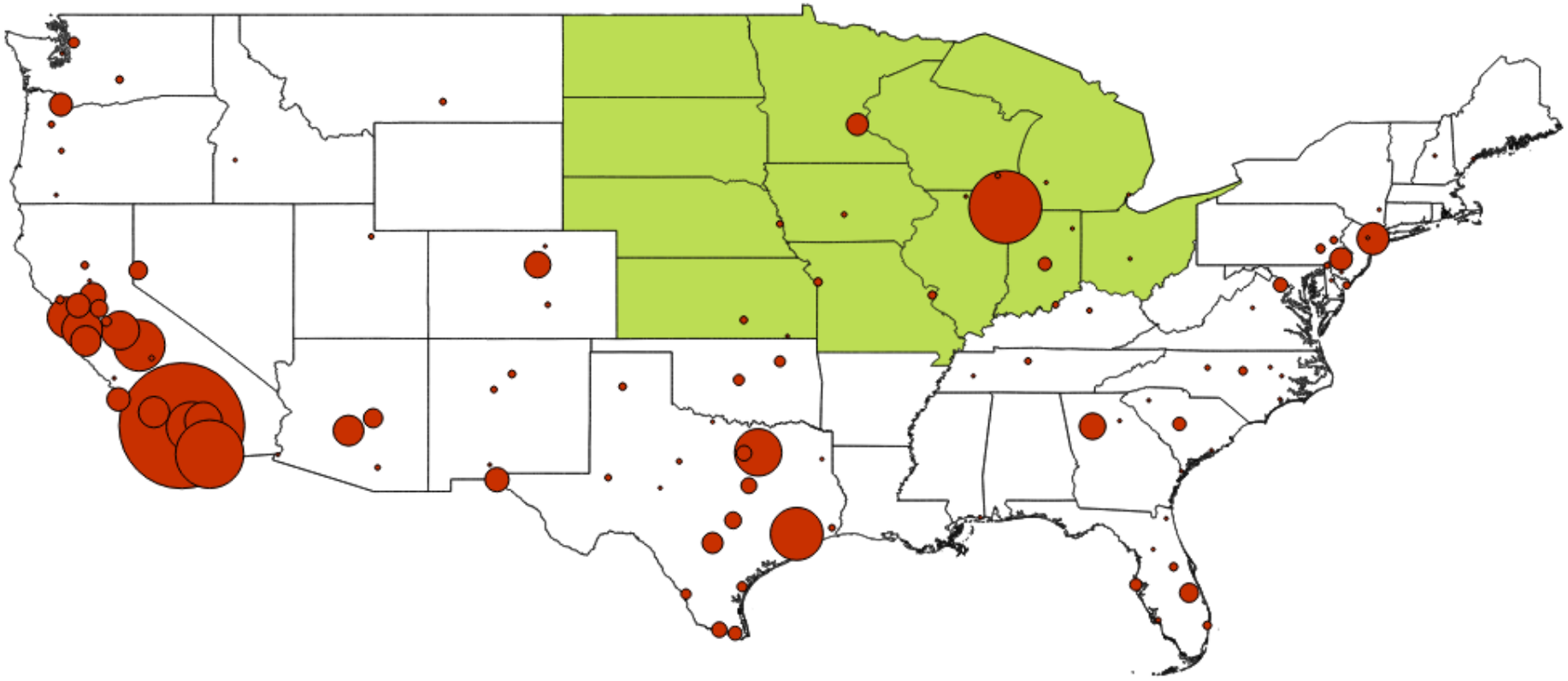


Figure: Weighted by population

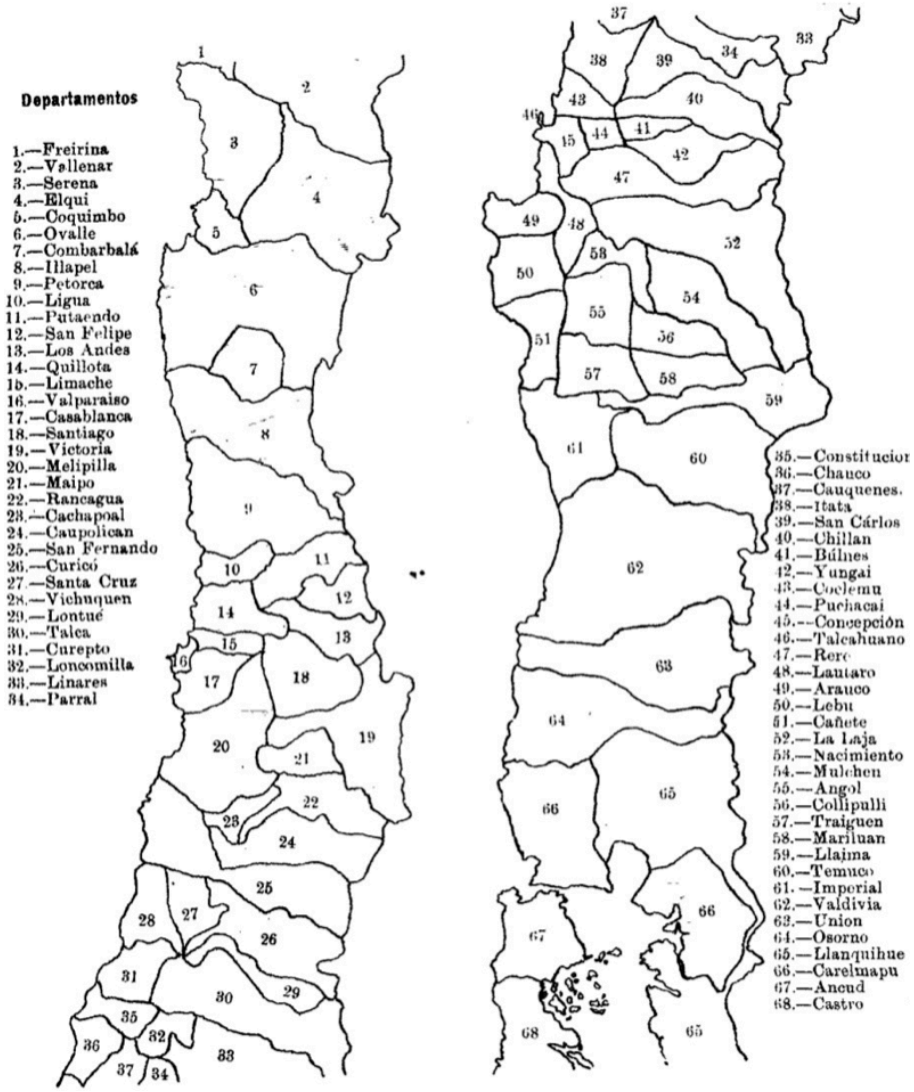


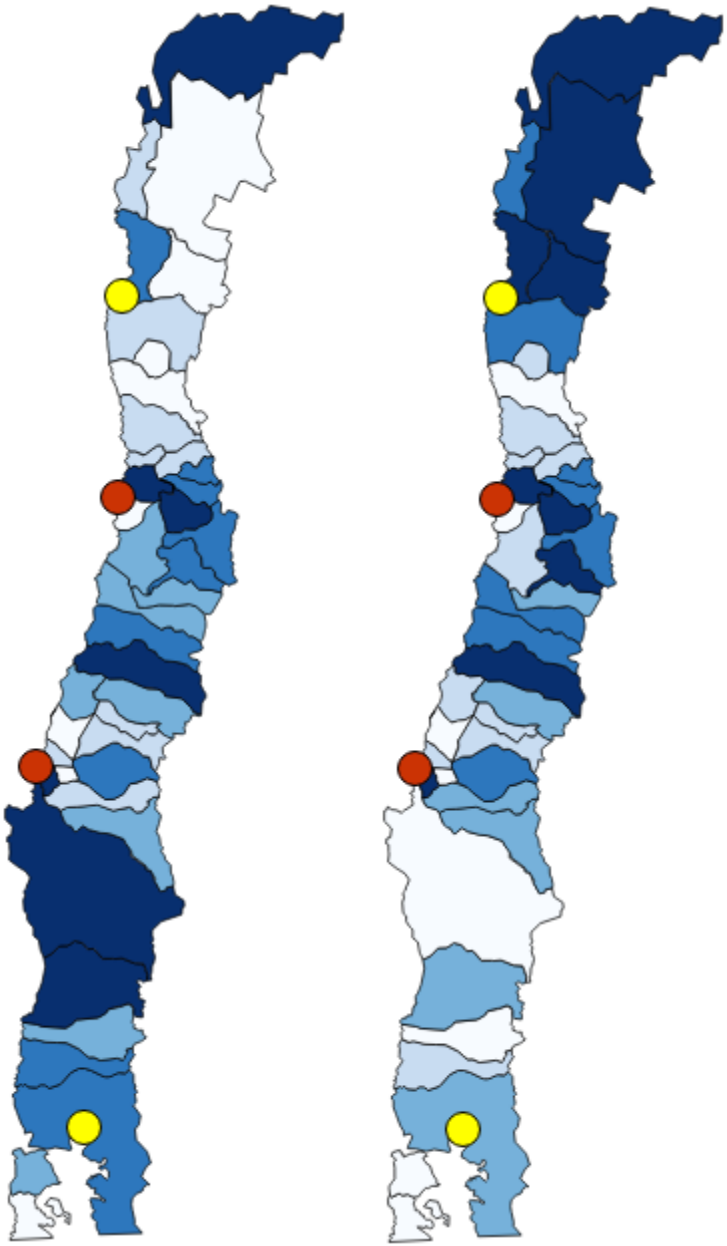
Digitalization of Maps

Map from Chile in the 1907 Census showing Provinces.

Our objective is to digitize this map to show geographic immigration patters.

We will use `Georeferencer`





The map on the left shows the distribution of European immigrants, while the one on the right shows the distribution of other immigrants.

Ports of entry are marked by circles ● and ●

Raster Data to Polygons

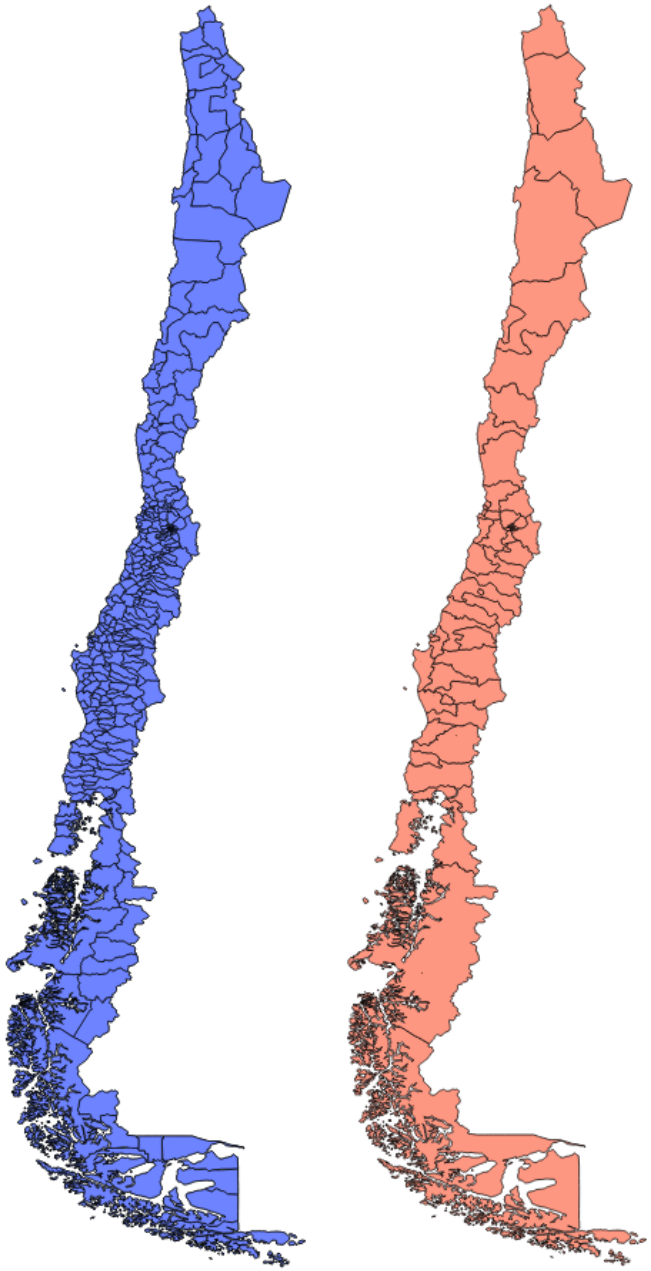
- In QGIS is also very easy to transform raster data to attributes of polygons using `Raster/Zonal Statistics`
- As an example, we will download a raster showing satellite night lights from outer space
  - For more information see "Measuring Economic Growth from Outer Space" (Henderson et al 2012 *American Economic Review*).
- This data serves as a way to measure local economic output in the absence of local statistics.

Figure: Satellite nights

Figure: Satellite nights and polygons

Shapefiles from Shapefiles

- As a final application, we can construct new shapefiles as a function of existing shapefiles using
- Example: electoral districts are polygons usually composed by the union of multiple municipalities
- For this we can use `Vector/Geoprocessing Tools/Dissolve`



The map on the left shows municipalities and the one on the right shows electoral districts, which are constructed as the union of municipalities.

Remarks

- You can also modify existing shapefiles by simply changing polygons, lines, or points with the `edit` tool
- There are any other spatial operations that you can do with QGIS. Many of these operations are included by default when you download QGIS
- You can download additional operations by installing plugins. Go to the menu bar and click in `Plugins/Manage and Install Plugins...` in order to do this

Overview

1. Introduction to QGIS
2. Applications using QGIS
3. Other applications

Other Applications

1. ~~Buffers~~
2. Network Analysis
3. Dijkstra's algorithm

Network Analysis

Losing Your Dictator

- What is the impact of a political transition on firms? We exploit the timing of Chile's unexpected transition to democracy to answer this
- Augusto Pinochet was dictator of Chile between 1973 and 1990, and he called for elections to validate his regime in 1988. However, unexpectedly, he lost
- We classify all publicly traded firms by how connected they were to Augusto Pinochet using a network analysis

Figure: Network of firms during Pinochet's dictatorship

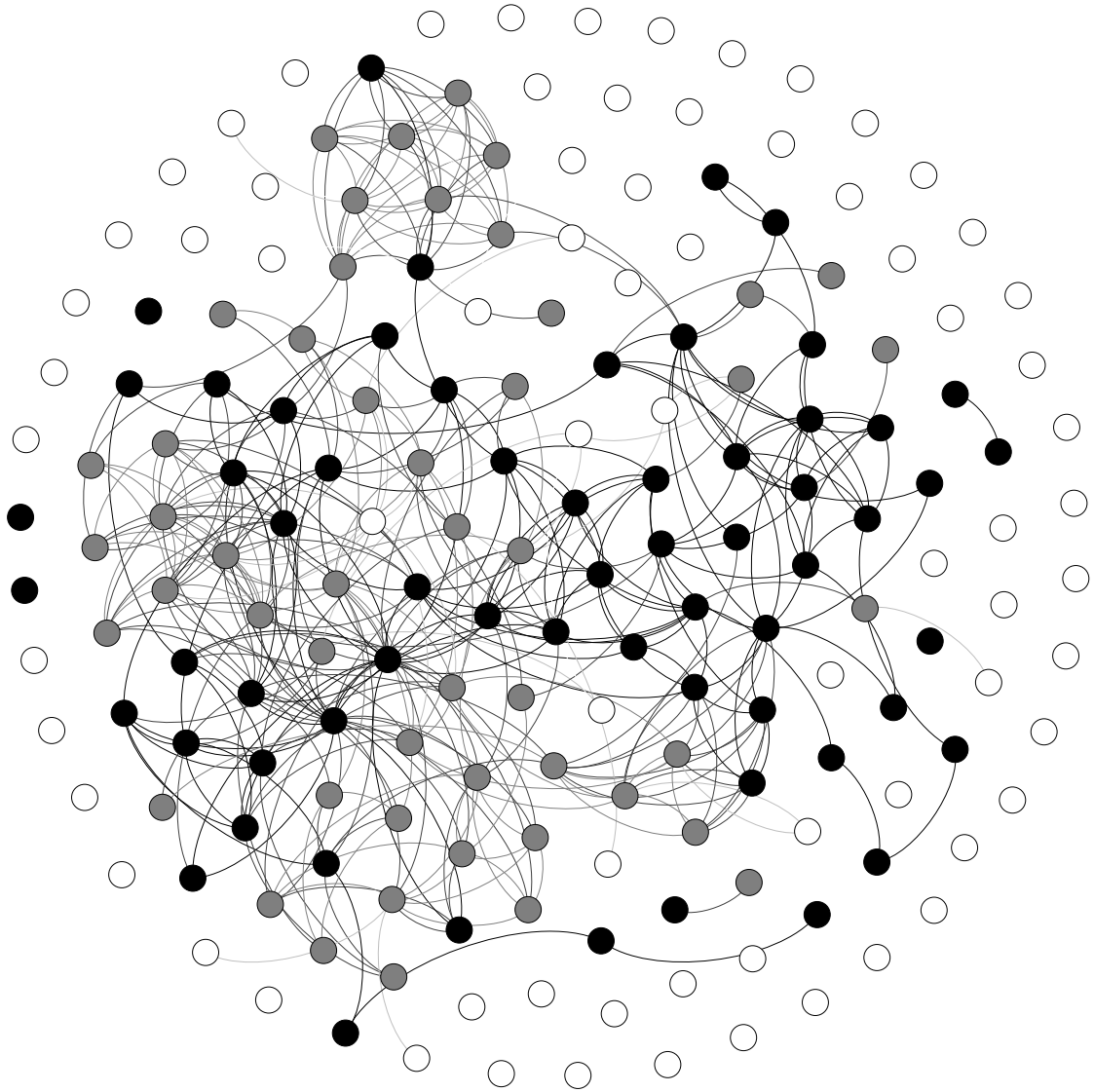


Figure: Stock price returns

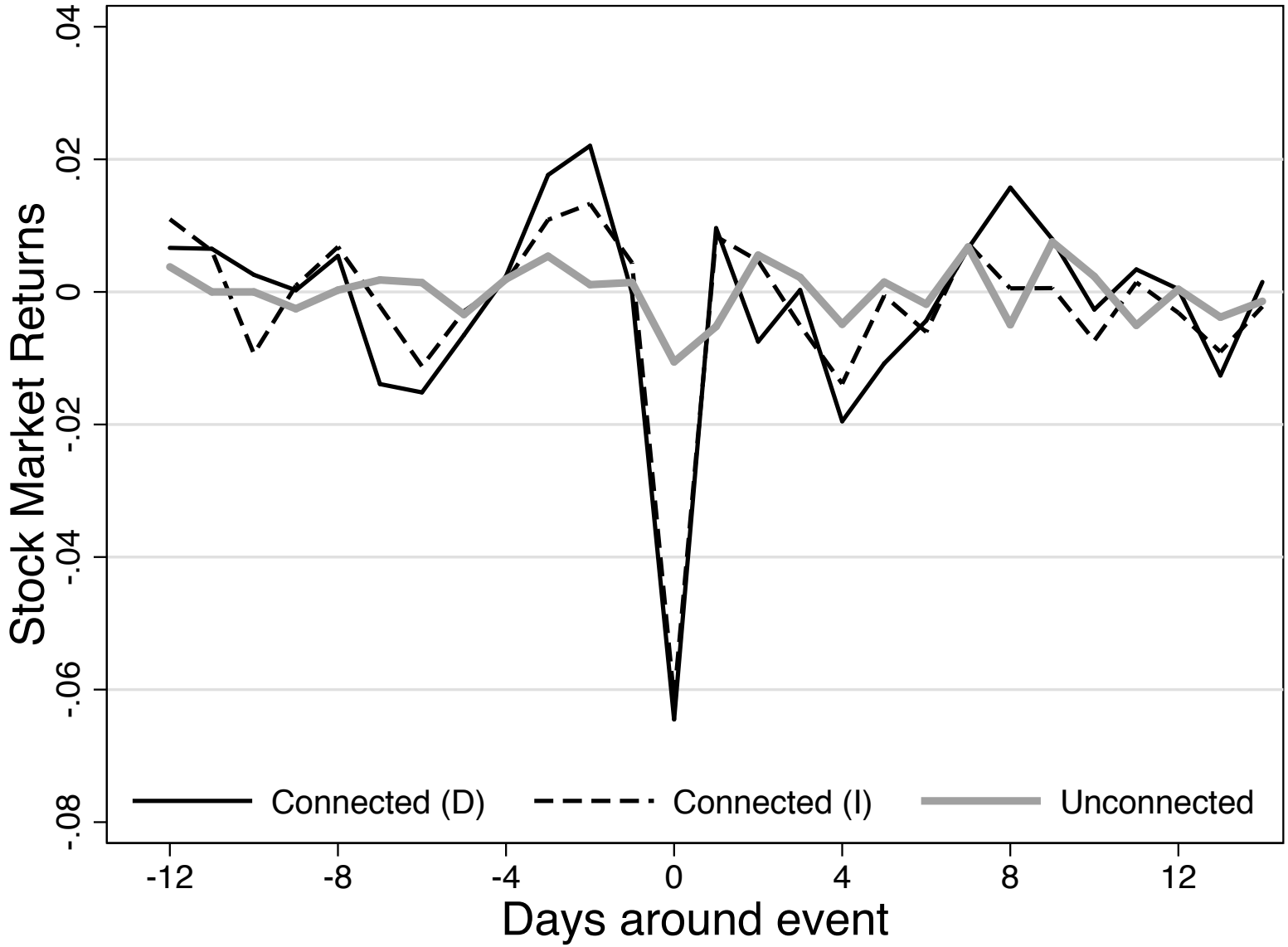


Figure: Investment before and after the Plebiscite

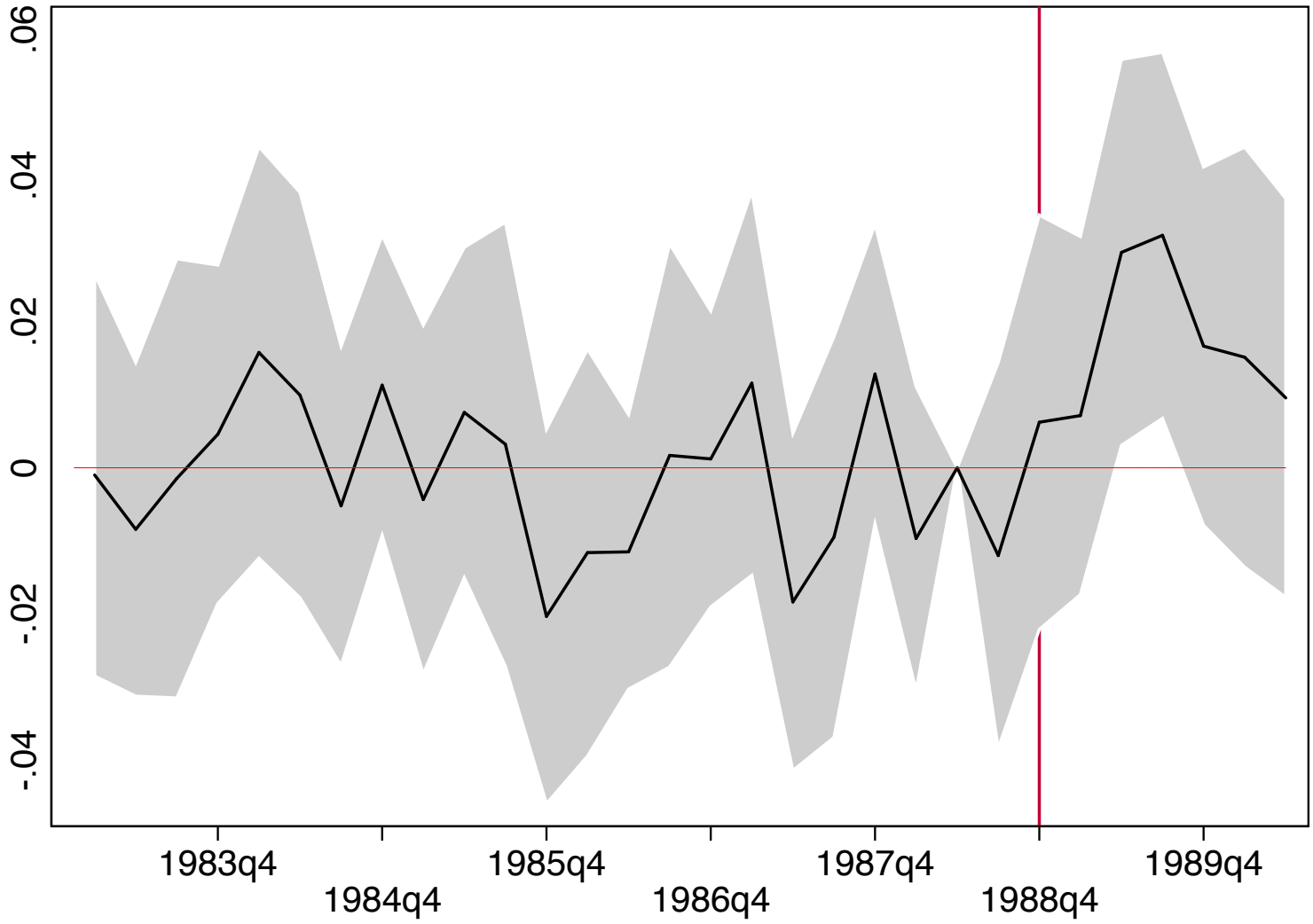
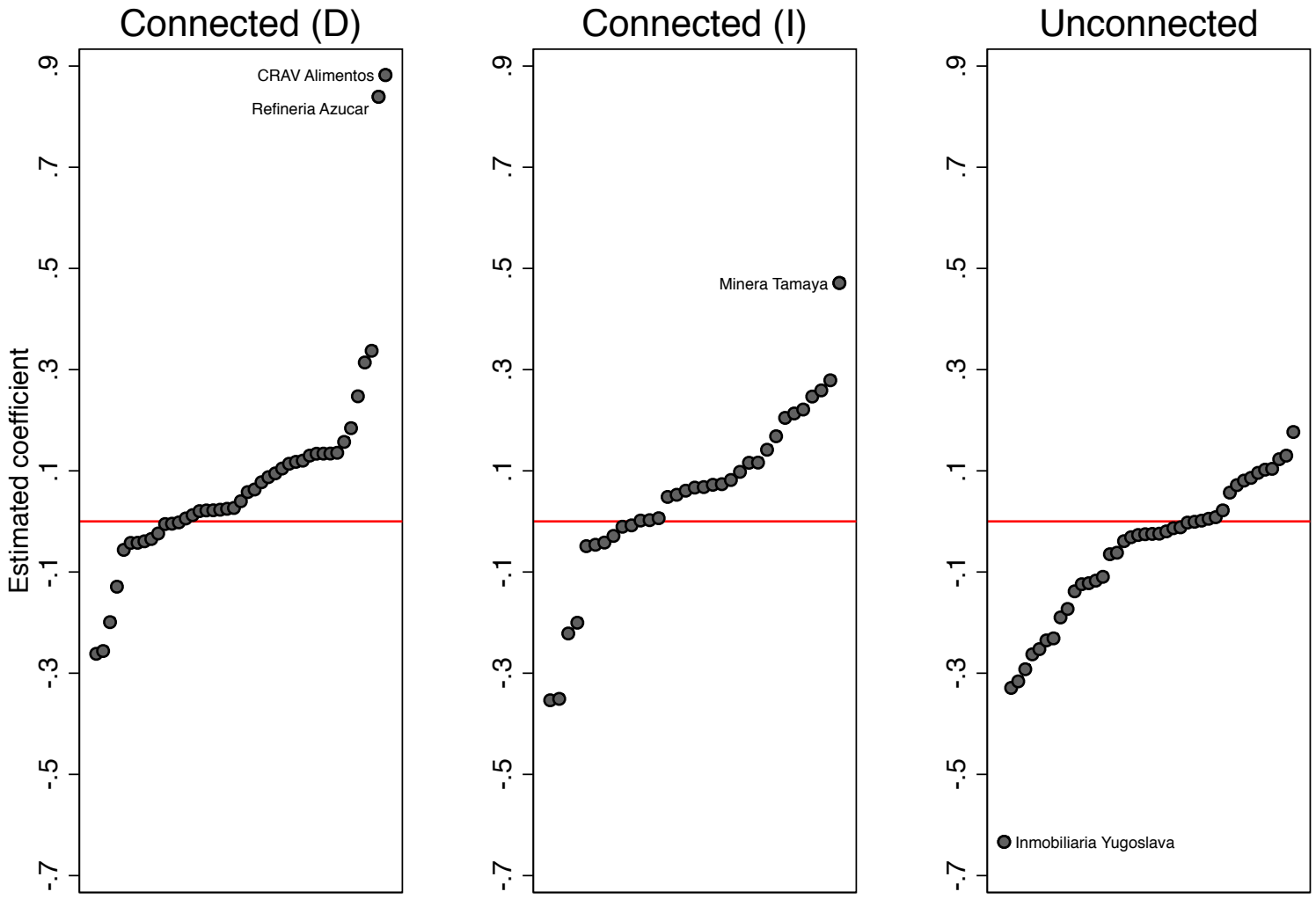


Figure: Investment



Dijkstra's Algorithm

Drug trafficking organizations

- In our final application we will use Dijkstra's's Algorithm to calculate the least cost route between two points in a network
- After specifying a cost function to move over a network, the algorithm calculates the optimal path between two points
- We will identify least cost routes between Mexican municipalities where drug trafficking organizations operate and entry points to the U.S.
  - See "Trafficking Networks and the Mexican Drug War" for a nice application of this idea (Dell 2014)

Figure: Presence of a drug-trafficking organization in 2010

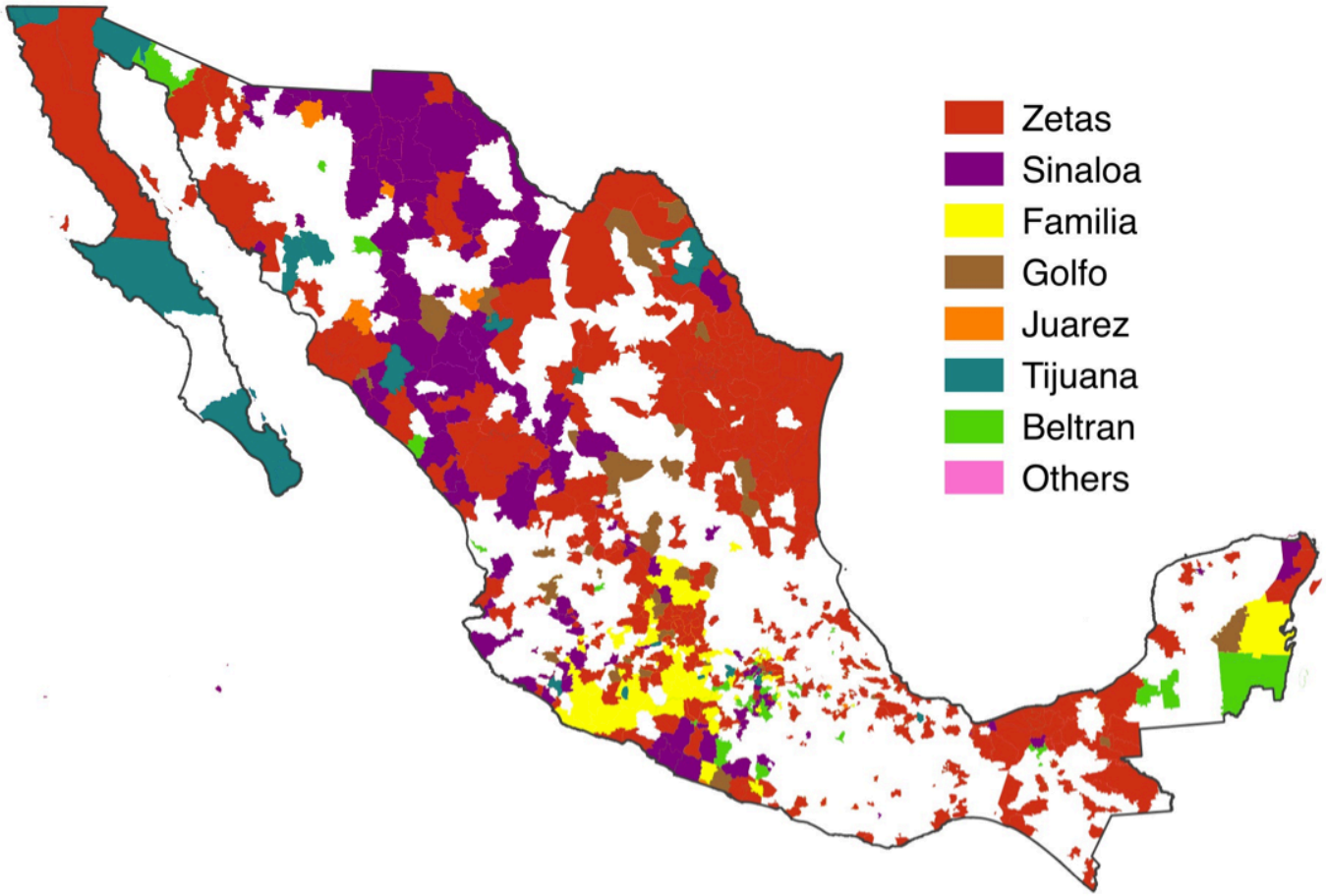
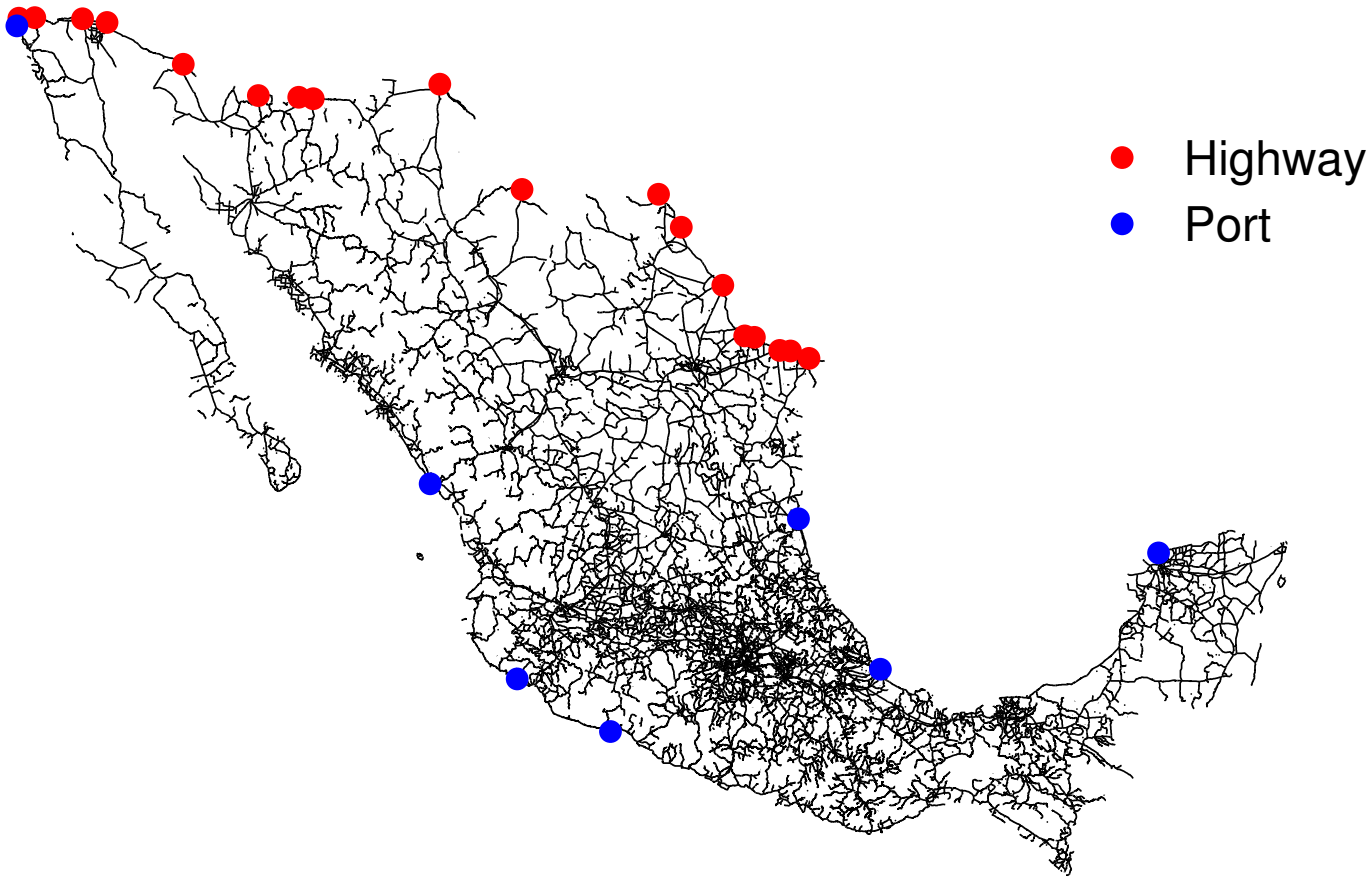


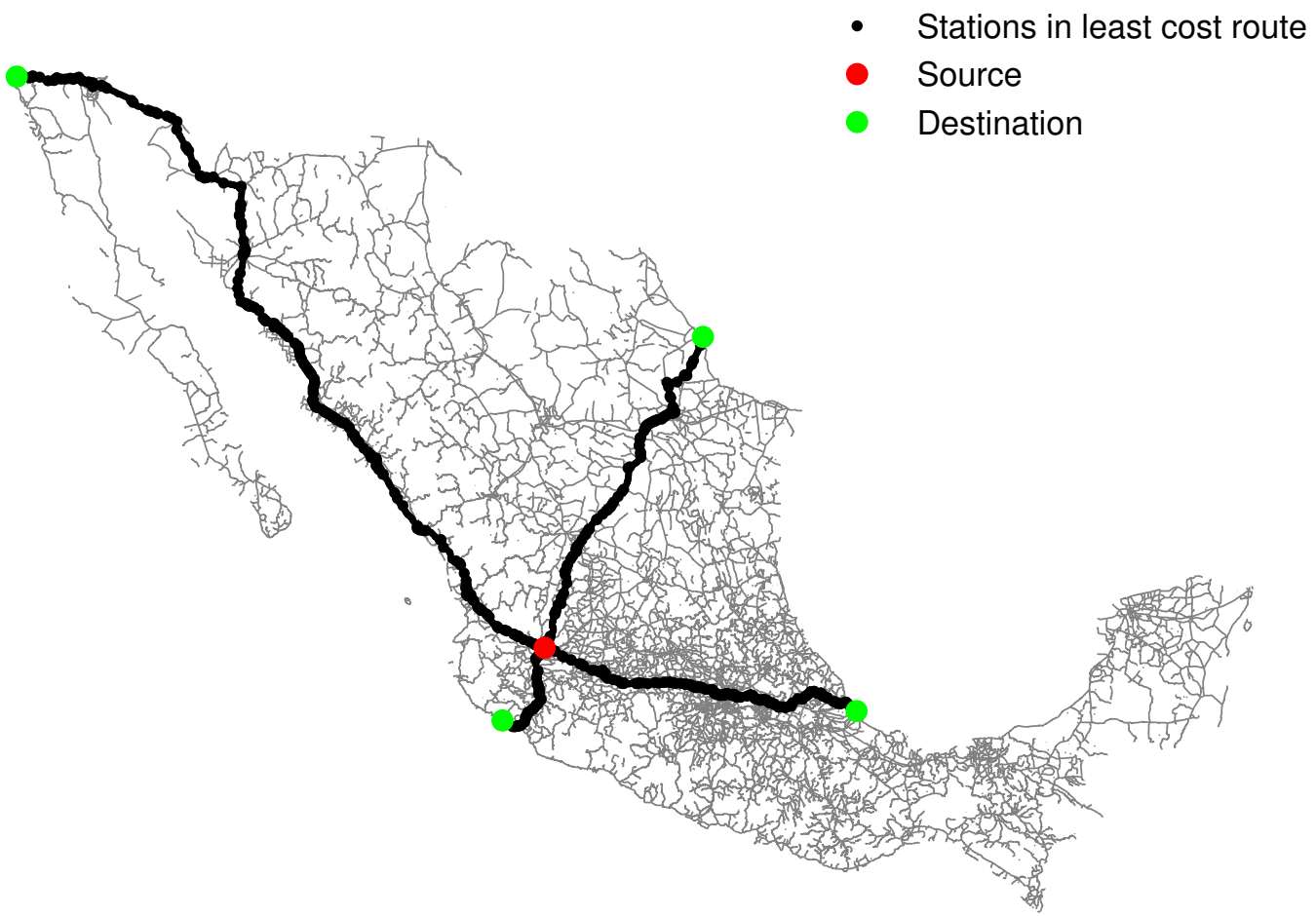
Figure: Road network and entry points to the U.S.



Algorithm

- Make use of the resources available in the internet. You don't have to code this algorithm, several people have already did it
- However, you do have to be able to read their code to make sure there are no mistakes and to learn how to properly use it
- Matlab central is a great place to look for resources, i.e. functions coded by other people to make your work easier

Figure: Least cost route using Dijkstra's algorithm



Thanks