

Calculating Zip Code Distances

Nathan Schiff

Updated September 18, 2015

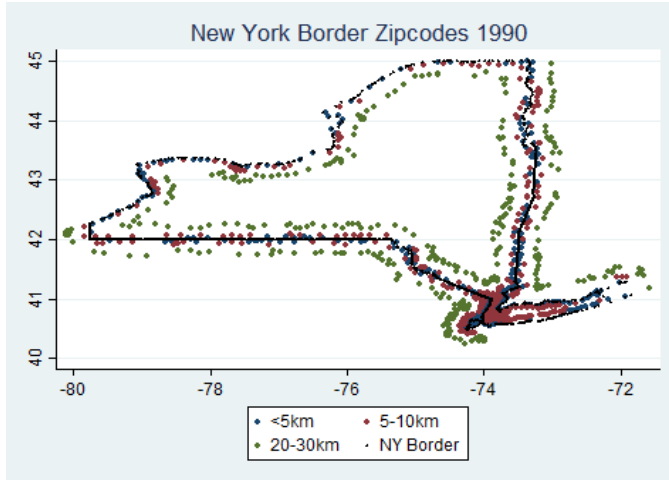
1 Calculation Notes

First I downloaded the ZCTA centroids from the Census Gazetteer webpage for 1990, 2000, 2010, 2012, 2013, 2014. Then I downloaded the TIGER Line US state border files for Census 2010 and the 2010 “cartographic boundary” files for 2010; the ZCTA centroids and both boundary files use the NAD 83 geographic coordinate system. The TIGER files are very precise boundaries that include water and coastal boundaries (ex: RI-NY ocean boundary line) while the cartographic boundaries are restricted to land areas and have a simpler shape (ex: coast lines are smoothed). For 2010 and 2012-2014 zipcodes the Census did not give us the state and so I used the precise TIGER files to assign states to zipcodes (point in polygon). To calculate distances I used the cartographic boundaries to try and exclude water boundaries. I brought these into R and created 2500 points along every state border. For a given Zip Code tabulation area (ZCTA) centroid I calculated the great circle distance to all 2500 points along a given state border. I selected the point along the border with the minimum distance to the ZCTA centroid as the closest point and used this distance as the distance between ZCTA and state. I repeated this process for all states and all ZCTAs.

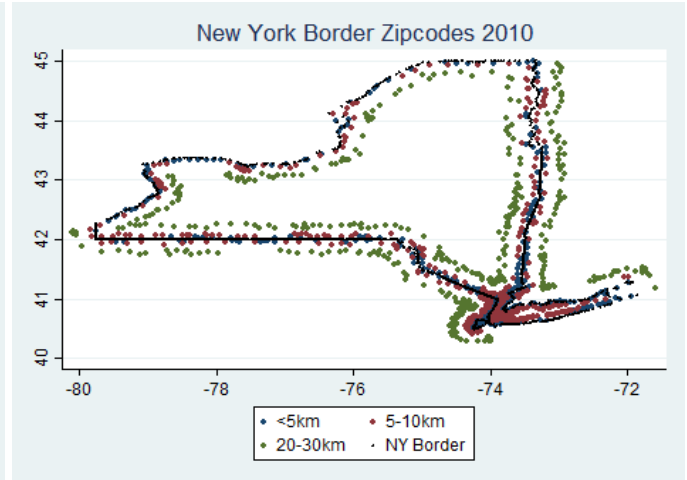
Notes:

1. In 2000 the Census had zipcodes with letters in them, such as “020HH.” For this reason I made the zip code field a string for all files.
2. The number of unique ZCTAs in each year is: (1990, n=29,191), (2000, n=32,705), and n=32,701 for years 2012, 2013, 2014.
3. The state assignment using GIS seems quite accurate. I compared the Census state to the GIS-assigned state for 1990 and there were only two differences out of 29,194 zipcodes (ZCTA centroids located right on the KY-IN and NM-TX borders). For 2000 there were 4 differences out of 32,760 zipcodes, also all border ZCTAs.
4. In a few (probably rare) instances, islands and water bodies create border distances that are deceptively close. For example, RI zipcode 02891 is only 14km from the New York State border because there is a New York island (Fisher Island) just off the southern coast of RI.

Figure 1: Border Zipcodes for New York: 1990 and 2010

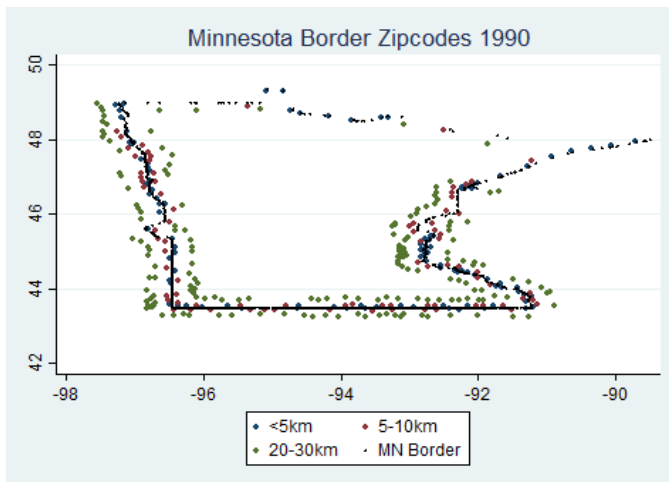


(a) NY1990

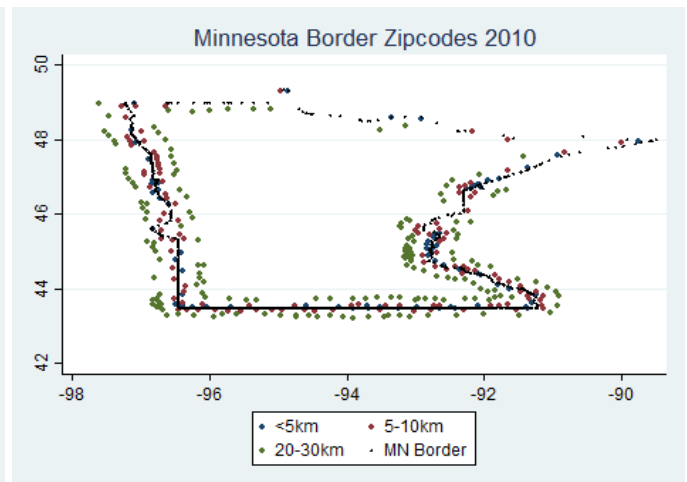


(b) NY2010

Figure 2: Border Zipcodes for MN: 1990 and 2010



(a) MN1990



(b) MN2010

Figure 3: Difference between TIGER (black) and Cartographic boundary files



Figure 4: Zipcodes centroids in 1990

Zip Code Tract Area centroids in 1990

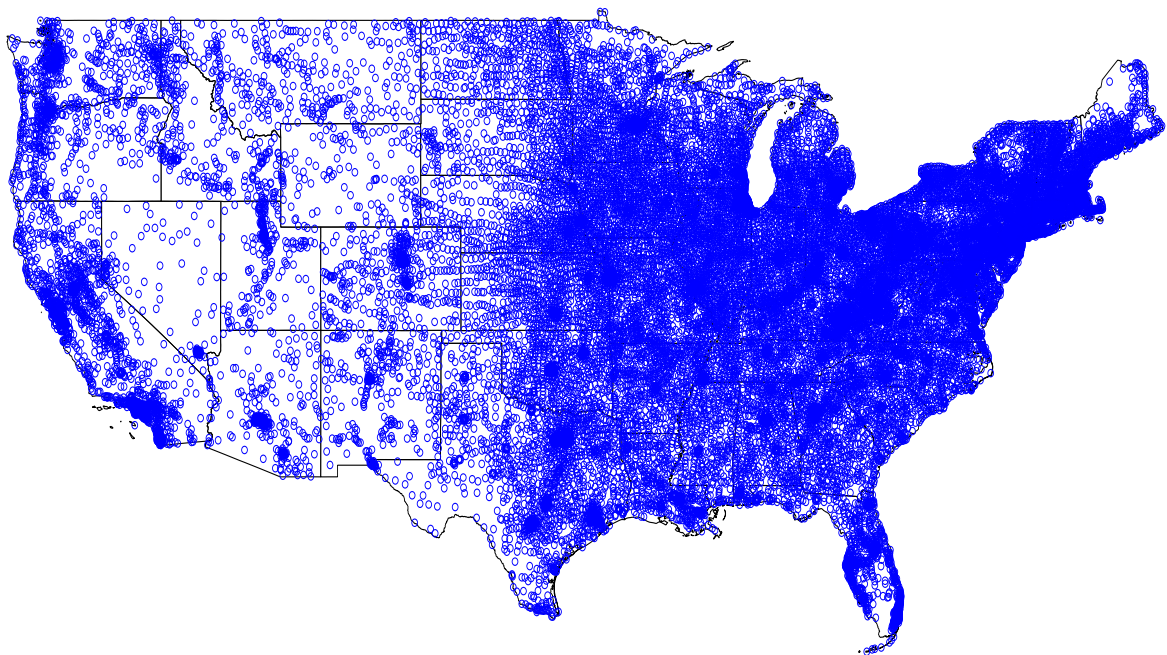


Figure 5: Zipcodes centroids in 2000

Zip Code Tract Area centroids in 2000

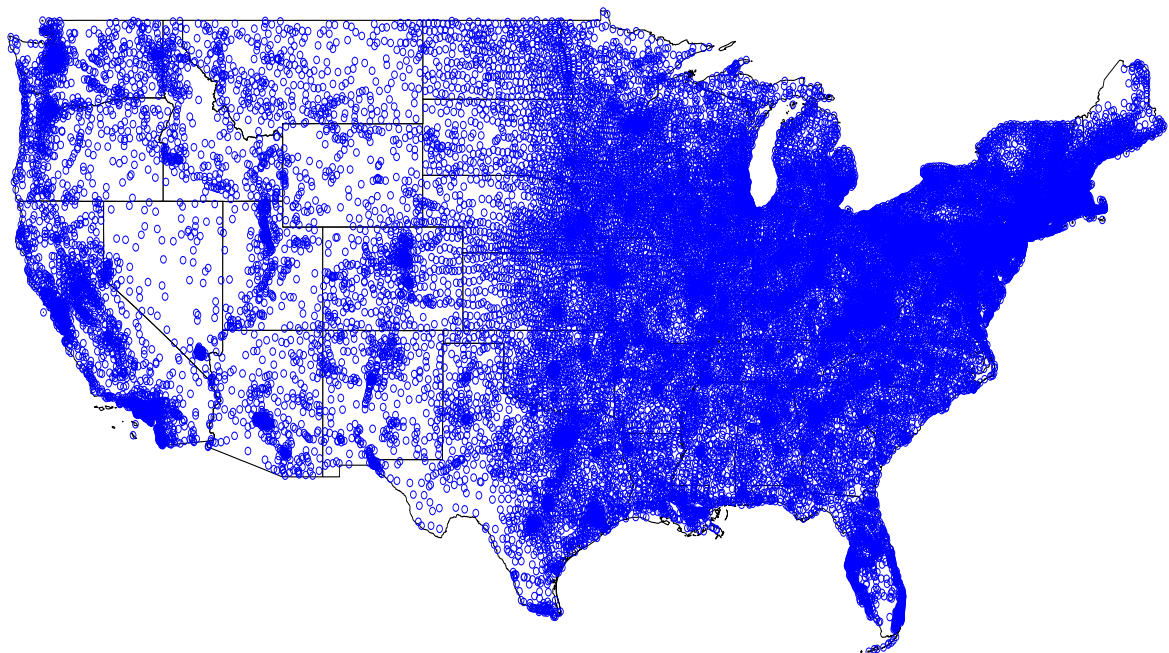


Figure 6: Zipcodes centroids in 2014

Zip Code Tract Area centroids in 2014

