\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Memorandum**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

To: Derek Wu, Neil Cholli

From: Dylan Craig

Date Created: 12/26/2024

Subject: Data Management Memo on Zip Code Borders

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Purpose:** To demonstrate the data cleaning and management process for deriving which zip codes belong to several counties based on geography and population share (2012).

**\*\*Read READ\_ME(s) as well as Script Directory for technical breakdown including file names**

**Step 1: Derive list of bordering counties based on shapefile**

Utilizing an R script, the TIGRIS package, and the VA Counties Tiger shapefile, a list of bordering counties were created for each county. A bordering county is designated if it touches another county.

Note: Some counties, such as those on Virginia’s peninsula, were considered bordering based on non-land borders (i.e., rivers, lakes, bays).

**Step 2: Derives list of Zip codes and 1) what FIPS codes they belong to based on population share and 2) which FIPS codes border the FIPS codes they belong to**

Utilizing another R script as well as the 2012 Q4 [HUD USPS Zip Code Crosswalk Files](https://www.huduser.gov/portal/datasets/usps_crosswalk.html), population shares (ratio of county addresses as a share of the zip codes total addresses). If a county makes up more than 1% of a zip codes addresses, that county can be considered associated. Then, remaining FIPS codes from Step 1 with population shares less than 1% are considered bordering only.

**Step 3: Derives list of bordering zip codes based on ZCTA shapefiles and population**

Utilizing ZCTA TIGER shapefiles via the R TIGRIS package as well as the prior “associated FIPS” codes (based on population) to create the following designations:

* Interior: The zip code only has one associated FIPS code and only intersects with one county
* Bordering: The zip code has one associated FIPS code but intersects with several county shapefiles
* Overlapping: The zip code has multiple associated FIPS codes and intersects with several county shapefiles