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

**Memorandum**

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

To: Derek Wu, Neil Cholli

From: Dylan Craig

Date Created: 12/15/2024

Date Updated: 12/20/2024

Subject: VDSS Office Distances README

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

**Purpose:** To calculate distances between VDSS office coordinates and midpoints of associated zip codes

**Step 1:** Gathering VDSS Office data

* Utilized the wayback machine and VDSS office location data to track office locations by date
* Separated datasets into 1) VDSS office locations existing in October 2010 and 2) VDSS office locations for all dates
* Files produced
  + **"VA\_CommonHelp\Raw Data\VDSS\_Office\_Distances\VDSS\_Offices\_2012.xlsx"**
  + **"VA\_CommonHelp\Raw Data\VDSS\_Office\_Distances\VDSS\_Offices\_All.xlsx"**

**Step 2:** Geocoding VDSS and Zip Code Coordinates

* Manually geocoded coordinates for each VDSS office using Google Maps
* Ran R Scripts ("**VA\_CommonHelp\Scripts\VDSS\_Office\_Distances\VDSS\_ZIP\_GGMAP\_GEOCODING\_2012.R**"; **"VA\_CommonHelp\Scripts\VDSS\_Office\_Distances\VDSS\_ZIP\_GGMAP\_GEOCODING\_ALL.R"**) assigning zip codes to VDSS offices using USPS Zip-to-FIPS database (**"VA\_CommonHelp\Raw Data\VDSS\_Office\_Distances\USPS\_Zip\_County.xlsx"**) and then, using GoogleAPI and R Package “Geosphere”, added midpoint coordinates for each zip code
  + Due to errors in some zip code coordinates (<100), manually corrected incorrect zip code midpoints (see all\_zip\_corrections .xlsx file)
* Files produced
  + **"VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\VDSS\_Office\_Zip\_GeoCode\_2012.xlsx"**
  + **“VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\VDSS\_Office\_Zip\_GeoCode\_All.xlsx"**
  + **"VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\VDSS\_Office\_Geocode\_All\_Corrected.xlsx"**

**Step 3:** Calculating Distances

* Ran R scripts calculating the Haversine distance, distance in miles/transit, travel time in driving/transit between each VDSS office and the midpoints of its associated zip codes (**VA\_CommonHelp\Scripts\VDSS\_Office\_Distances\VDSS\_HAVERSINE\_DISTANCES\_2012.R"; “VA\_CommonHelp\Scripts\VDSS\_Office\_Distances\VDSS\_DISTANCES\_ALL.R"**)
* Files produced
  + **"VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\VDSS\_Office\_Zip\_GeoCode\_2012\_Distances.xlsx"**
  + **"VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\VDSS\_Office\_Zip\_GeoCode\_All\_Distance.xlsx"**

**Step 4:** Analyzing Distances Calculated for VDSS\_All

* Ran R script producing summary statistics, histograms, correlation table, etc. for different measurements of distance generated for all years ("**VA\_CommonHelp\Scripts\VDSS\_Zip\_Office\_Distances\VDSS\_Distances\_Analysis\_All.R**")
* Files produced
  + Plots and Tables can be found in **“VA\_CommonHelp\Plots\VDSS\_Zip\_Office\_Distances**"

**Step 5:** Assigning Treatment Status to VDSS Office / Zip Codes

* Assigned levels of treatment (e.g., above median distance compared to all other haversine distances)
* File produced
  + "**VA\_CommonHelp\Data Outputs\VDSS\_Office\_Distances\zip\_treat.dta**"