**VDSS\_DISTANCES\_2012\_ZCTA.R - README**

**Author: Dylan Craig**

**Date Last Updated: 12/29/2024**

**File Name: VDSS\_DISTANCES\_2012\_ZCTA.R**

**Purpose**

This script calculates multiple geospatial distances and travel times, including Haversine, driving, and transit distances, between ZCTA (ZIP Code Tabulation Area) midpoints and LDSS (Local Department of Social Services) office midpoints for October 2012. The calculations incorporate datasets for geolocation and ZIP-FIPS mappings to provide comprehensive metrics.

**Overview**

The script computes:

1. **Haversine distances** between ZCTA midpoints and LDSS midpoints.
2. **Driving distances and times** using the Google Maps API.
3. **Transit distances and times** using the Google Maps API.

These results are integrated into a final dataset for analysis, providing insights into geospatial relationships and travel accessibility.

**Datasets and Tools**

**Input Files:**

1. **HRSA\_ZCTA.xlsx:** Contains the crosswalk between ZIP codes and ZCTAs.
2. **VDSS\_Offices\_2012.xlsx:** Location data for VDSS offices as of October 2012.
3. **USPS\_Zip\_County.xlsx:** Includes ZIP code to county FIPS mappings.
4. **TIGRIS Shapefiles:** Used to obtain 2010 ZCTA shape data for Virginia.

**Output Files:**

1. **VDSS\_Office\_ZCTA\_GeoCode\_2012\_Distances.xlsx:** Final dataset containing Haversine, driving, and transit distances and times.

**Dependencies:**

* **R Libraries:**
  + dplyr: Data manipulation.
  + readxl: Reading Excel files.
  + tigris: Accessing TIGER/Line shapefiles.
  + sf: Handling spatial data.
  + geosphere: Calculating Haversine distances.
  + openxlsx: Exporting data to Excel files.
  + gmapsdistance: Interfacing with the Google Maps API.
  + progress: Displaying progress bars.

**Script Functionality**

**1. Data Loading and Preprocessing:**

* Reads ZCTA crosswalk, VDSS office data, and ZIP-FIPS mappings.
* Obtains ZCTA shapefiles for Virginia from TIGRIS and extracts centroids (midpoints).

**2. Merging Datasets:**

* Combines ZCTA centroid data with ZCTA crosswalk information.
* Merges ZIP-FIPS mappings and VDSS office data based on county FIPS codes.

**3. Distance and Time Calculations:**

* **Haversine Distances:** Calculates straight-line distances between ZCTA midpoints and LDSS midpoints.
* **Driving Distances and Times:**
  + Utilizes the Google Maps API to calculate road travel distances and times.
  + Tracks progress with a progress bar.
* **Transit Distances and Times:**
  + Utilizes the Google Maps API to calculate public transit distances and times.
  + Tracks progress with a progress bar.

**4. Unit Conversion:**

* Converts distances from meters to miles.
* Converts times from seconds to minutes.

**5. Data Export:**

* Saves the final dataset with calculated distances and times as an Excel file.

**File Output**

**VDSS\_Office\_ZCTA\_GeoCode\_2012\_Distances.xlsx:**

* Includes calculated Haversine, driving, and transit distances and times, along with geospatial metadata.

**Usage**

1. Place input files in the appropriate directories.
2. Install required R libraries and set up the Google API key.
3. Run the script in an R project environment.
4. The output file will be saved in the Data Outputs/VDSS\_ZCTA\_Office\_Distances/ directory.

**Notes**

* Ensure TIGRIS shapefiles are accessible and correctly downloaded for the specified year.
* Verify the Google API key is active and has sufficient quota for API calls.
* Review input datasets for completeness and accuracy to ensure correct distance calculations.
* Rows with missing distance metrics are flagged for further inspection.