**VDSS\_ZIP\_RESIDENTIAL\_STATUS.R - README**

**Author: Dylan Craig**

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

**File Name: VDSS\_ZIP\_RESIDENTIAL\_STATUS.R**

**Purpose**

This script classifies ZIP codes in Virginia as Residential, Non-Residential, or Unknown based on IRS 2020 population estimates. It merges USPS ZIP-County data with IRS ZIP code data and assigns residential status based on population values.

**Overview**

The script processes ZIP code data to:

1. Load USPS ZIP-County and IRS population data.
2. Merge datasets based on ZIP codes.
3. Assign residential status to each ZIP code based on population values.
4. Save the final dataset as an Excel file for further use.

**Datasets and Tools**

**Input Files:**

1. **USPS\_Zip\_County.xlsx:** Contains ZIP code and county mappings.
2. **zip\_code\_IRS\_2020.xls:** Provides IRS 2020 population estimates for each ZIP code.

**Output File:**

1. **Zip\_Residential\_Status.xlsx:** Final dataset containing residential classifications for ZIP codes.

**Dependencies:**

* **R Libraries:**
  + dplyr: For data manipulation.
  + readxl: For reading Excel files.
  + writexl: For writing data to Excel files.

**Script Functionality**

**1. Loading Data:**

* Reads USPS ZIP-County and IRS ZIP code data.
* Removes the FIPS column from USPS data and retains unique ZIP codes.

**2. Merging Datasets:**

* Merges USPS ZIP-County data with IRS population data based on ZIP codes.

**3. Assigning Residential Status:**

* Assigns residential status based on the POPULATION column:
  + **Residential:** Population > 0.
  + **Non-Residential:** Population = 0.
  + **Unknown:** Missing population data.

**4. Saving the Output:**

* Exports the final dataset as an Excel file.

**File Output**

**Zip\_Residential\_Status.xlsx:**

* Contains ZIP codes with assigned residential statuses (Residential, Non-Residential, or Unknown).

**Usage**

1. Place the input files in the specified directories.
2. Install the required R libraries.
3. Execute the script in an R project environment.
4. Access the output file in the Data Outputs/VDSS\_Zip\_Residential\_Status/ directory.

**Notes**

* The accuracy of the classification relies on the completeness of the IRS 2020 population data.
* Any discrepancies or missing data in the inputs may affect the results.