**VDSS\_HUD\_USPS\_2011\_2012\_Type\_Comparison.R - README**

**Author: Dylan Craig**

**Date Last Updated: 12/28/2024**

**File Name: VDSS\_HUD\_USPS\_2011\_2012\_Type\_Comparison.R**

**Purpose**

This script analyzes ZIP code data for Q4 2011 and Q1 2012, focusing on the distribution of TYPE and NUMBER\_TYPES across the specified quarters.

**Overview**

The script, **VDSS\_HUD\_USPS\_2011\_2012\_Type\_Comparison.R**, filters and processes historical ZIP code data to examine and summarize distributions for ZIP TYPE and NUMBER\_TYPES during Q4 2011 and Q1 2012. It outputs the results in an Excel file for further analysis.

**Datasets and Tools**

**Input Files:**

1. **VA\_HUD\_USPS\_Zip\_Data\_Collapsed.xlsx**: Contains historical ZIP code data, including classifications and metrics across multiple quarters.

**Output File:**

* **VA\_HUD\_USPS\_2011\_2012\_TYPE\_COMPARISON.xlsx**: Includes the distribution summaries for TYPE and NUMBER\_TYPES for Q4 2011 and Q1 2012.

**Dependencies:**

* **R Libraries:**
  + dplyr: Data manipulation
  + readxl: Reading Excel files
  + writexl: Writing Excel files

**Script Functionality**

**1. Loading Data:**

* Reads the collapsed HUD USPS ZIP data from the input file.

**2. Filtering Data:**

* Filters rows for the specified quarters: Q4 2011 and Q1 2012.

**3. Summarizing Distributions:**

* **TYPE Distribution:** Summarizes the count of ZIP codes by TYPE for each quarter.
* **NUMBER\_TYPES Distribution:** Summarizes the count of ZIP codes by NUMBER\_TYPES for each quarter.

**4. Exporting Results:**

* Combines the TYPE and NUMBER\_TYPES distributions into a list and exports them as an Excel file.

**File Output**

**VA\_HUD\_USPS\_2011\_2012\_TYPE\_COMPARISON.xlsx**:

* **Sheets:**
  1. TYPE Distribution: Distribution of ZIP codes by TYPE for Q4 2011 and Q1 2012.
  2. NUMBER\_TYPES Distribution: Distribution of ZIP codes by NUMBER\_TYPES for Q4 2011 and Q1 2012.

**Usage**

1. Place the input dataset in the specified directory.
2. Install required R libraries.
3. Run the script within the designated R project environment.
4. The output will be saved automatically to the specified folder.

**Notes**

* Ensure input files are correctly formatted and located in the appropriate directory.
* The analysis focuses solely on Q4 2011 and Q1 2012. For additional quarters, modify the YEAR\_QUARTER filter in the script.
* The accuracy of the results depends on the completeness and quality of the input dataset.