**Memorandum**

**To:** Derek Wu, Neil Cholli  
**From:** Dylan Craig  
**Date Created:** [Insert Date]  
**Subject:** Data Management Memo for VDSS ZCTA ACS Characteristics

**Purpose:**  
To explain how the VDSS ZCTA ACS dataset was cleaned and prepared, focusing on population, employment, education, and poverty indicators.

**Step 1: Load Datasets**

* Loaded raw data using a relative project path.
* Used dplyr, tidyverse, and readr for data processing.

**Step 2: Population Data**

* Extracted total population estimates for ZCTAs and years.
* Renamed variables to improve clarity.

**Step 3: Employment and Industry**

* Separated employment data into two periods (2010-2012 and 2013-2022) to address variable changes.
* Included industry-specific data to analyze economic activity.
* Calculated employment, unemployment, and not-in-labor-force rates.

**Step 4: Age Groups**

* Aggregated age groups from sex-specific data to get population by age range.
* Created a new child population variable (0-17 years).

**Step 5: Median Income**

* Extracted and renamed median household income (2011 inflation-adjusted dollars).

**Step 6: Race and Ethnicity**

* Combined race and ethnicity data, including Hispanic, White non-Hispanic, Black non-Hispanic, and others.
* Created an "Other non-Hispanic" category for less common groups.

**Step 7: Poverty**

* Calculated poverty counts and rates, including a child poverty variable.
* Derived poverty rates for total and child populations.

**Step 8: Education**

* Categorized education into "No High School," "High School Graduate," "Some College," and "College Graduate."
* Summarized totals by gender and education level.

**Step 9: Marital Status**

* Aggregated marital status into "Never Married," "Married (Spouse Present/Absent)," "Widowed," and "Divorced."

**Step 10: Merge Datasets**

* Merged all datasets by ZCTA and year using left\_join.
* Checked for consistency in totals (e.g., age, sex, race).

**Step 11: Weighted Averages**

* Calculated weighted averages for key indicators (e.g., poverty, employment) using population weights.

**Step 12: Save Outputs**

* Saved the final dataset as CSV and Excel files for further use.

**Conclusion:**  
The cleaned dataset consolidates demographic, economic, and social indicators for ZCTAs, ensuring consistency and usability. Each step focused on accuracy and reproducibility using modular R scripts.