**Memorandum**

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

To: Derek Wu, Neil Cholli

From: Dylan Craig

Date Created: 12/26/2024

Subject: Data Analysis Memo on Bad ZipCounty Rate by Zip Code Type and Border Status

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

**Purpose:** To provide descriptive statistics on the percent of residents reporting to the incorrect VDSS local office based on their county of residence by zip code type and zip code border status.

**Data Collection:**

Data was at the Zip-FIPS level and included the following:

* Zip Code Type
  + Standard, Unique, or PO Box
* Zip Code Border Status
  + Interior, bordering, overlapping, no shapefile (meaning there wasn’t a corresponding ZCTA)
* Bad Zip-County Rate
  + Percent of zip code residents going to the “incorrect” VDSS local office

**Data Analysis: Zip Code Border Status**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **BORDER\_STATUS** | **avg\_N** | **avg\_bad\_zipcounty** | **weighted\_bad\_zipcounty** | **count\_N** | **count\_bad\_zipcounty** |
| Bordering | 13872.23324 | 0.027812743 | 0.013105442 | 343 | 343 |
| Interior | 7901.391459 | 0.040331811 | 0.013334288 | 281 | 281 |
| No Shapefile | 151.2745098 | 0.301188618 | 0.140419098 | 153 | 153 |
| Overlapping | 18215.28854 | 0.011529199 | 0.008322939 | 253 | 253 |
|  |  |  |  | 0 | 0 |

The weighted bad zipcoding rate accounts for population differences across zip codes. Zip codes without a ZCTA shapefile show the highest incorrect LDSS reporting rate at around 14%. However, their small population size means this statistic may not be particularly meaningful.

Looking at the other categories, Interior zip codes have a rate of 1.33%, followed closely by Bordering at 1.31%, with Overlapping showing the lowest at 0.8%. This ordering is unexpected - we would predict that zip codes fully within a county's boundaries would have the lowest rate of incorrect LDSS office visits. While the difference between bordering and overlapping rates isn't statistically significant, both differ from interior zip codes significantly at the 95% confidence level.

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

**Data Analysis: Zip Code Types**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TYPE** | **avg\_N** | **avg\_bad\_zipcounty** | **weighted\_bad\_zipcounty** | **count\_N** | **count\_bad\_zipcounty** |
| PO Box | 293.7956989 | 0.241869515 | 0.089045859 | 186 | 186 |
| Standard | 13788.98926 | 0.023727954 | 0.011128003 | 838 | 838 |
| Unique | 12 | 1 | 1 | 1 | 1 |
|  | 49.8 | 0.8 | 0.425702811 | 5 | 5 |

For zip code types, Unique zip codes show the highest weighted bad zipcounty rate at 100%, but this is based on a single zip code with just 12 residents. PO Box zip codes follow with an 8.9% weighted average, while standard zip codes have a much lower rate at 1.1%. This notable difference may be explained by PO Box zip codes typically being located in more rural areas than standard zip codes.