North Beach Flood Control Zone District Economic Impacts Research Methodology

12-30-2024

Authored by Robin Vance

[Under contract to SBGH-Partners in Support of Grays Harbor County's "NBFCZD – Gaging and Governance Project" as executed and dated 10-18-2024]

Data Sources:

Data Axle Reference Solutions - Granular Business Data

Host Compliance® - Short Term Rental Data

Grays Harbor County Assessor's Office - Parcel Data

Grays Harbor County GIS Department - Parcel and District Shapefiles

---

Data Axle Reference Solutions

Note on Data Axle Reference Solutions

Data Axle Reference Solutions gathers U.S. business data from over 5,200 Yellow and White Page directories, utility connects and disconnects, annual reports, press releases, news feeds, and web research. The data is further verified and enhanced by a team of more than 350 database specialists who manually check each record for quality and completeness (Source: https://www.referenceusagov.com/Static/OurQuality?utm).

Their U.S. Business database contains a total of 103 million businesses, including 18 million verified businesses and 85 million unverified businesses that are updated weekly. This database is further enhanced by conducting more than 24 million phone calls annually, ensuring the most accurate and comprehensive data possible. This makes it a leading resource for reliable and up-to-date business information (Source: http://www.referenceusa.com/Home/Home).

Collecting the Data:

For the purposes of this research, Data Axle Reference Solutions' "U.S. Businesses" database was utilized to identify businesses within the North Beach Flood Control Zone District (NBFCZD) and the Ocean Beach Subzone. The database was queried using a user-defined geographic polygon that encompassed the boundaries of these areas. The query returned a list of relevant businesses along with associated details.

Due to a limitation in the download feature, which restricts search results to 250 entries per download, the data was downloaded as six separate CSV files. Each file contained a subset of the businesses within the defined geographic area. These CSV files were then combined into a single Excel workbook. The combined dataset served as the foundation for subsequent analysis and was saved for further processing and refinement.

Filtering the Data:

The Data Axle Reference Solutions dataset was imported into ArcGIS Pro as a table, where each business entry's address was geocoded to generate corresponding X and Y coordinates. To ensure that only businesses located within the NBFCZD were included in the analysis, businesses outside the spatial extent of the district were identified and removed from the layer and its attribute table. The resulting geocoded table, now limited to businesses within the NBFCZD, was then exported and included in the Excel Workbook file "All_NBFCZD_Businesses_Both_Geocoded_NotGeocoded" for further filtering and analysis. A few dozen entries could not be accurately geocoded within ArcGIS Pro due to lacking address information, so a separate *Non_Geocoded* sheet was created for this data. The main sheet is *All_NBFCZD_Businesses_Geocoded*.

The Data Axle Reference Solutions dataset categorizes businesses as "Verified," indicating they have been quality-checked by Data Axle, "Unverified," for entries not yet fully verified, and "Closed / Out of Business," representing dissolved businesses. All entries marked as "Closed / Out of Business" were removed prior to analysis.

To further refine the dataset, erroneous "Unverified" entries were filtered as thoroughly as possible. Following an initial filtering of duplicate business entries, all "Unverified" businesses were attended to. Due to the extensive scope of the dataset, which included over 300 unverified business entries, and the challenges in accurately determining the operational status of these businesses, manual verification was not conducted for every unverified entry.

To streamline the process, manual checks were prioritized for businesses whose last update in the Data Axle database predated the year 2020. This targeted approach allowed for efficient validation of older, potentially outdated entries, ensuring a more accurate dataset without requiring a full review of every unverified business. This process involved cross-referencing multiple external sources to verify the activity and presence of businesses within the NBFCZD. These sources included the Washington Secretary of State Corporations and Charities Filing System, which was used to check the registration status of businesses; the Grays Harbor County Assessor's Office "Taxsifter" parcel search, which provided additional location and property information; Google Maps, for identifying physical business locations; and Google Web Search, for gathering further evidence of operational status.

Job Count Calculation

The number of jobs in the North Beach Flood Control Zone District was extrapolated by summing employee data from the project's Excel Workbook, with adjustments to account for incomplete data and refine the results.

The direct summing of employee counts from the dataset was made and recorded in cell B2 under the *Calculations* sheet. Many entries within the Data Axle Reference Solutions dataset lacked employee number or range counts, resulting in "0" values where employee numbers were recorded in the Excel Workbook. A second calculation adjusted for missing data by assuming every business had at least one employee, ensuring businesses with zero-reported employees were included in the total recorded in cell B3 under the same sheet. Results were further refined by excluding businesses deemed irrelevant to employee counts, such as ATMs and gas stations (see *Excel Functions for Calculations* for a full list of exclusions), while still applying the minimum one-job assumption for businesses with missing or zero-reported data. Such were recorded in cell B4.

<u>Excel Functions for Job Count Calculations</u>

The following three formulas are designed to calculate sums across specific columns in two worksheets, All_NBFCZD_Businesses_Geocoded and Non_Geocoded, within the Excel workbook "All_NBFCZD_Businesses_Both_Geocoded_NotGeocoded". Each formula addresses different goals for summing the data.

=SUM(All_NBFCZD_Businesses_Geocoded!AF2:AF346) + SUM(Non_Geocoded!Y2:Y15), performs a sum of values in column AF from the first sheet and column Y from the second sheet, **equating to the summation of employees for all businesses listed across sheets.**

=B2+COUNTIF(All_NBFCZD_Businesses_Geocoded!A2:AF346, 0) + COUNTIF(Non_Geocoded!Y2:Y15, 0) adds a layer of logic to the calculation. For both columns, it replaces any 0 values with 1 and adds these values to the original employee total, **equating to the summation of employees for all businesses across sheets where a value of '0' is converted into '1' (i.e. all businesses are associated with at least 1 employee).**

=SUM(IF(ISNA(MATCH(All_NBFCZD_Businesses_Geocoded!Y2:Y346, {"Atm-Automated Teller Machines","Mailing & Shipping Kiosks","Electric Charging Station","Gas-Propane-Refilling Stations","Propane Tank Kiosks","Service Stations-Gasoline & Oil"}, 0)), IF(All_NBFCZD_Businesses_Geocoded!AF2:AF346=0, 1, All_NBFCZD_Businesses_Geocoded!AF2:AF346), 0)) +SUM(IF(Non_Geocoded!Y2:Y15=0, 1, Non_Geocoded!Y2:Y15)) introduces additional filtering criteria. This formula refines the results by excluding businesses categorized as:

- Atm-Automated Teller Machines
- Mailing & Shipping Kiosks (e.g., USPS collection boxes)
- Electric Charging Station
- Gas-Propane-Refilling Stations
- Propane Tank Kiosks
- Service Stations-Gasoline & Oil

For the remaining rows, it replaces any `0` values with `1` to ensure every included business contributes at least one job. Rows that match the predefined excluded categories are assigned a value of `0`, ensuring they are omitted from the summation. **This equates to the summation of employees for all businesses across sheets where a value of '0' is converted into '1' with certain businesses excluded.**

***Reasoning for exclusions:** These types of businesses were considered irrelevant to employee counts in the context of calculating jobs because they typically have minimal or no on-site staffing requirements. For instance, ATMs and propane tank kiosks are primarily self-service and do not require dedicated employees. Similarly, mailing and shipping kiosks, as well as electric charging stations, are often automated or require only occasional maintenance rather than consistent staffing. Gas-propane-refilling stations and service stations for gasoline and oil may employ staff, but their roles are usually limited and highly specific, which may not align with the broader employment or economic impact goals being analyzed. Excluding these businesses allows the calculation to focus on establishments more likely to contribute significantly to the district's workforce, providing a clearer and more meaningful picture of employment.

Business Count Calculation

An initial count of NBFCZD businesses was conducted by tallying all business entries within the dataset, which was recorded in cell B6 of the *Calculations* sheet. Similarly to the job count calculation, the dataset was filtered to exclude businesses deemed irrelevant to the analysis, though only the "Mailing & Shipping Kiosks" category was filtered out. This exclusion ensured that only relevant businesses were included in the final count, which was recorded in cell B7 under the same sheet.

Excel Functions for Business Count Calculations

The following formulas were used to calculate the total number of businesses in the NBFCZD using data from two worksheets, `All_NBFCZD_Businesses_Geocoded` and `Non_Geocoded`, within the Excel workbook "All_NBFCZD_Businesses_Both_Geocoded_NotGeocoded." Each formula addresses a specific goal for counting businesses relevant to the district.

`=COUNTA(All_NBFCZD_Businesses_Geocoded!A2:A346) + COUNTA(Non_Geocoded!A2:A15)` calculates the total number of business entries by counting all non-empty rows in column `A` from both the geocoded and non-geocoded datasets. This formula assumes that any row with data in column `A` represents a business and provides a baseline total count for all businesses included in the datasets.

`=SUMPRODUCT((All_NBFCZD_Businesses_Geocoded!A2:A346<>"")*(All_NBFCZD_Businesses_Geocoded!Y2:Y346<>"Mailing & Shipping Kiosks")) + COUNTA(Non_Geocoded!A2:A15)` introduces filtering criteria to refine the count. In this formula, rows from the geocoded dataset are excluded if column `Y` indicates the business is categorized as "Mailing & Shipping Kiosks." These businesses were deemed irrelevant to the analysis due to their minimal staffing requirements or automated nature.

Host Compliance®

A dataset of short-term rental (STR) listings within the North Beach Flood Control Zone District (NBFCZD) was gathered using Host Compliance® Software and integrated into the *All_NBFCZD_Businesses_Both_Geocoded_NotGeocoded* Excel Workbook under a new sheet titled *STR*. To ensure the data was limited to the NBFCZD, a polygon layer of the district was used to spatially filter the listings, retaining only those within the district's boundaries.

The dataset was further refined to include only entries that met Host Compliance's definition of a short-term rental. Listings that did not meet this definition, such as hotels and motels (many of which were already present within the business entries of the Excel Workbook), were excluded from the *STR* sheet. To calculate the total number of short-term rentals that are likely to be active in the NBFCZD, all short-term rental entries currently advertised or advertised within the past year were summed and reported in cell B9. This total was incorporated into the total number of jobs calculations, as reported in cell B10 of the *Calculations* sheet, as well as the total number of businesses calculations, as reported in cell B11 of the same sheet.

<u>Excel Functions for Short Term Rental Calculations:</u>

The total number of short-term rentals was calculated using the formula `=SUM(STR!BF2:BF208)`, which sums all valid entries in column `BF` of the *STR* sheet. These entries represent short-term rentals currently advertised or advertised/updated within the past year. To integrate short-term rentals with the broader business count while excluding specific categories such as "Mailing & Shipping Kiosks" (e.g., USPS collection boxes), the formula `=B9+B7` was used. Here, `B7` represents the total short-term rental count, and `B9` represents the count of businesses excluding USPS collection boxes. Lastly, the total of short-term rentals and employee counts with exclusions (adjusted for non-zero values) was calculated using the formula `=B9+B4`. This combines the previously refined business count (`B9`) with the total employee count, where any zero-reported values were replaced with one to ensure all businesses contributed to the total.