Mechanical and Industrial Engineering
Department IE 6600: Computation and
Visualization.
Spring 2025

Food and Nutrition

Group 2

Members: Praneeta Mahentesh, Bhavya Kapadia, Soumya Gosala

Introduction:

Access to nutritious food is a fundamental component of public health, particularly in urban environments where food insecurity and service accessibility vary across neighborhoods. Washington, D.C., hosts a diverse array of food and nutrition support services provided by local nonprofits, community centers, and outreach programs. Understanding where these services are located, how they are distributed geographically, and which areas may be underserved is crucial for improving equity in food access.

This project aims to perform a geospatial analysis of food and nutrition services in Washington, D.C., using an official dataset of service providers. By leveraging Python and Plotly, we created interactive visualizations—ranging from maps and pie charts to cluster and trend analyses—to uncover patterns in service availability across ZIP codes, wards, and census tracts.

The results of this analysis provide actionable insights for city planners, nonprofit coordinators, and health policymakers who aim to optimize service delivery, identify coverage gaps, and strengthen food security initiatives in the region.

Data Processing and Analysis

1. Data Acquisition:

The dataset titled "Food_and_Nutrition.csv" was sourced from a government open data portal (data.gov) and provides detailed information about food and nutrition service providers in Washington, D.C. The dataset consists of 23 records and originally included 26 columns, capturing essential geographic, organizational, and administrative attributes.

Key Fields in the Dataset:
- CATEGORIES_OF_SERVICE: Specifies the type of service provided, such as "Food Delivery" or "Meal Services."
- ORGANIZATION: Name of the service provider or agency.
- FULL_ADDRESS: Physical address of the service location.
- EMAIL / PHONE / WEBSITE: Contact information for outreach or client support.
- ZIPCODE / MAR_WARD / MAR_CENSUS_TRACT: Geographic and administrative identifiers that enable location-based analysis.
- Latitude & Longitude: Provided under the columns MAR_LATITUDE and MAR_LONGITUDE, enabling accurate geospatial visualization.

Purpose of the Dataset:
This dataset serves as a foundational resource for examining how food and nutrition services are distributed across different parts of Washington, D.C. Through geospatial visualization and statistical exploration, this dataset enables us to:
- Identify geographic clusters of high or low service density.
- Highlight underserved communities with limited access to nutrition services.

- Support policy design and data-driven resource allocation by community organizations and city planners.

This dataset forms the core for generating interactive maps, density analyses, cluster models, and other geospatial visualizations that offer actionable insights into the city's food service infrastructure.

2. Data Inspection and Cleaning:

2.1 Initial Data Inspection

Before performing any analysis or visualization, the dataset was examined using Python functions such as df.info() and df.describe(). These commands helped assess the structure, completeness, and integrity of the data.

Key observations from initial inspection:

- The dataset contained 26 columns and 23 rows.
- Several columns were empty, offering no analytical value.
- Coordinate fields were available in both projected and geographic formats.
- Some fields contained missing or inconsistent values, particularly in contact information.

2.2 Cleaning Process

Dropping Redundant and Empty Columns:

Columns with 100% null values or no analytical utility were removed to streamline the dataset. The following fields were dropped:

- SERVICE_AVAILABLE_TO, ADDITIONAL_SERVICES_MAYINCLUDE: Fully null.
- MAR_ERROR, MAR_IGNORE: Contained only missing or uniform values without analytical purpose.

This step reduced the total column count from 26 to 22, retaining only fields necessary for geospatial and categorical analysis

2.3 Handling Missing Values

The dataset contained missing values in a few non-critical fields:

- The EMAIL field had 7 missing entries, which were filled with the placeholder "Not Provided" to preserve record completeness.
- Other fields crucial to mapping (e.g., Latitude, Longitude, Ward, etc.) were inspected and ensured to have complete values.

Rows with missing or invalid coordinates were dropped to ensure visualizations like maps and clustering remained accurate.

2.4 Renaming Columns for Clarity

To enhance readability and maintain naming consistency, select fields were renamed:

- MAR_LATITUDE → Latitude
- MAR_LONGITUDE → Longitude
- X, Y → Longitude_Projected, Latitude_Projected

These updates ensured clarity during visualization scripting and made the dataset more human-readable.

2.5 Data Type Conversion
Geospatial fields (Latitude, Longitude) were explicitly **converted to numeric data types** using pd.to_numeric() with error coercion. This ensured compatibility with Plotly and other geospatial tools. Any rows where these conversions failed were dropped from the dataset.

3. Final Data Verification

Following the cleaning process, the dataset contained:
- 23 complete rows (unchanged from the original but cleaned).
- 22 meaningful columns, all verified to be relevant and properly typed.
- No null values in critical fields used for visualizations and clustering.

This refined dataset provided a solid foundation for building interactive maps, trend lines, cluster plots, and categorical breakdowns in the subsequent analysis.

Exploratory Data Analysis (EDA)
1. Interactive Map of Service Locations
A line plot of the average suicide rate over time (1950-2019) provides insight into the historical evolution of suicide rates across different decades.
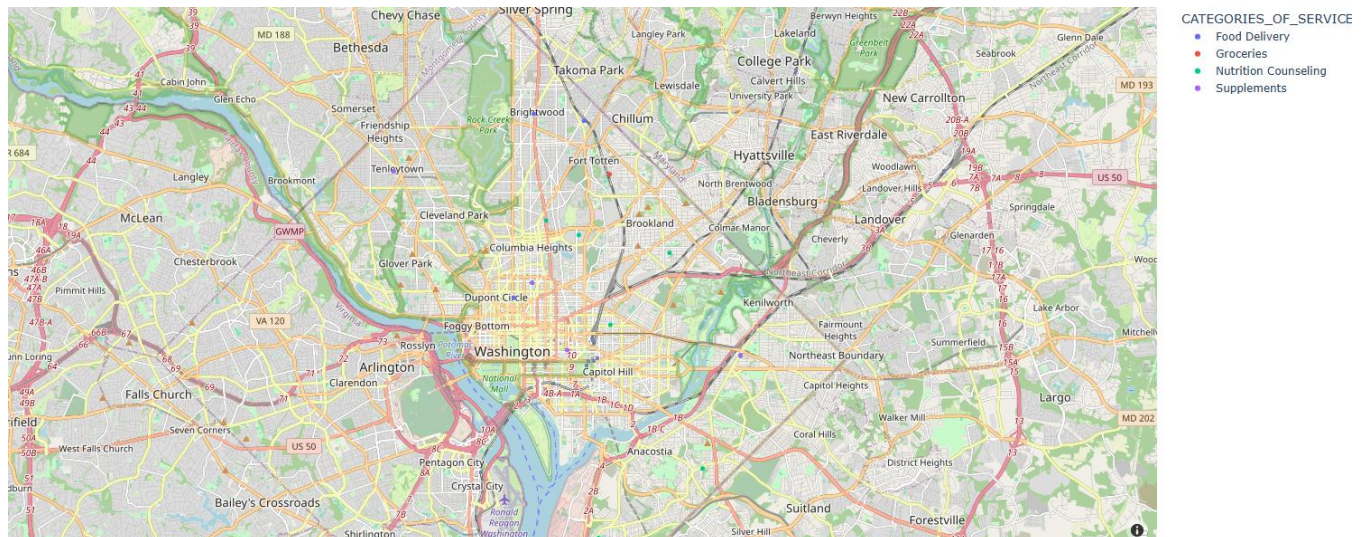
Key Observations
- Service Clusters in Central and Northeast D.C.: The map reveals a high concentration of services around areas like Columbia Heights, Fort Totten, and Capitol Hill, indicating strong coverage in densely populated neighborhoods.
- Ward 4 and Ward 7 as Major Service Hubs: Consistent with bar chart findings, Wards 4 and 7

display the greatest number of mapped services, suggesting targeted outreach in these regions.

- Food Delivery Dominates the Service Landscape: The blue markers indicate that food delivery services are the most widely available across the city, likely due to their flexible, at-home service model.
- Noticeable Gaps in Outlying Areas: Regions on the outskirts of Washington, D.C.—such as Forestville, Falls Church, and parts of Southeast—have few or no services, pointing to potential service deserts.
- Accessibility and Urban Placement: Most services are located in urban, transit-friendly zones, enhancing accessibility. However, this may limit access for individuals in low-density or less connected neighborhoods.

2. Distribution of Service Categories

Key Observations

- Nutrition Counseling is the Most Common Service: Nearly 48% of all services fall under Nutrition Counseling, indicating a strong emphasis on educational and preventive care in the region.
- Food Delivery Services Hold a Significant Share: Food Delivery makes up 26.1%, reflecting high demand for direct-to-home food access, likely supporting seniors, individuals with disabilities, or low-mobility residents.
- Supplement Programs Are Moderately Represented: Supplements account for 21.7% of the services, suggesting a growing focus on micronutrient support and targeted health interventions.
- Grocery Services Are Minimal: Only 4.35% of providers focus on Groceries, pointing to a potential gap in direct food access options like food pantries or fresh produce distribution.
- Balanced Mix Indicates a Holistic Service Approach: While Nutrition Counseling dominates, the presence of multiple service types suggests an effort to address food insecurity through both education and direct aid.

3. Service category per Ward

Service Count by Ward – Food Delivery
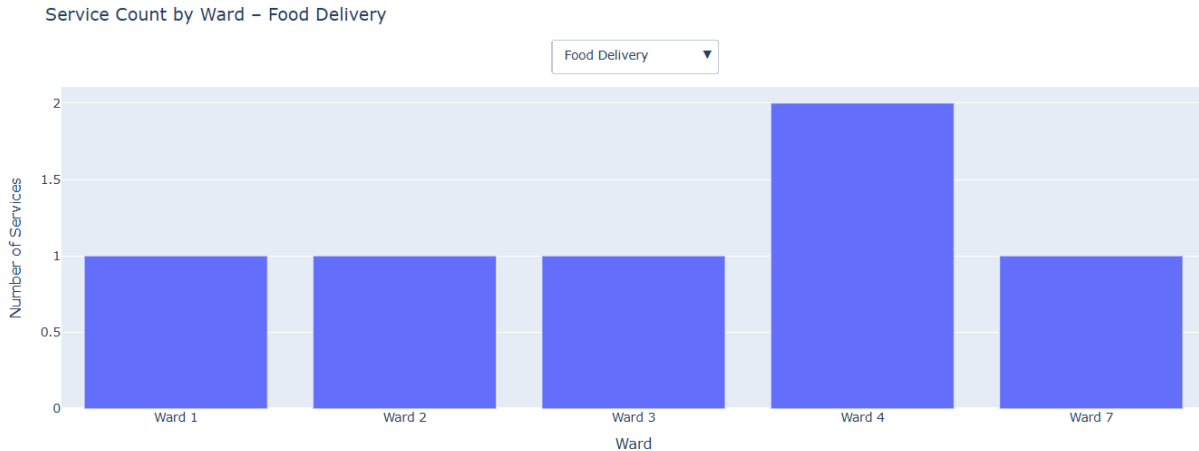
Key Observations

- Ward 7 Leads in Service Count: With 6 services, Ward 7 has the highest number of food and nutrition programs, indicating it may be a focal point for community outreach and high-demand populations.
- Wards 3 and 1 Show Strong Service Presence: Both Ward 3 and Ward 1 have 4 services each, suggesting moderate but consistent service coverage in these central areas of the city.
- Wards 5 and 2 Have Mid-Level Access: Each of these wards hosts 2–3 services, reflecting balanced but potentially improvable access to nutrition support.
- Wards 8 and 6 Appear Underserved: With only 1 service each, Wards 6 and 8 may be considered service deserts, potentially requiring targeted policy or nonprofit intervention.
- Inequity in Distribution Across Wards: The chart highlights a disparity in service availability, which could be tied to socioeconomic factors, demographic demand, or historical funding decisions.

4. Treemap of Service Categories by ZIP Code

Treemap of Service Categories by ZIP Code
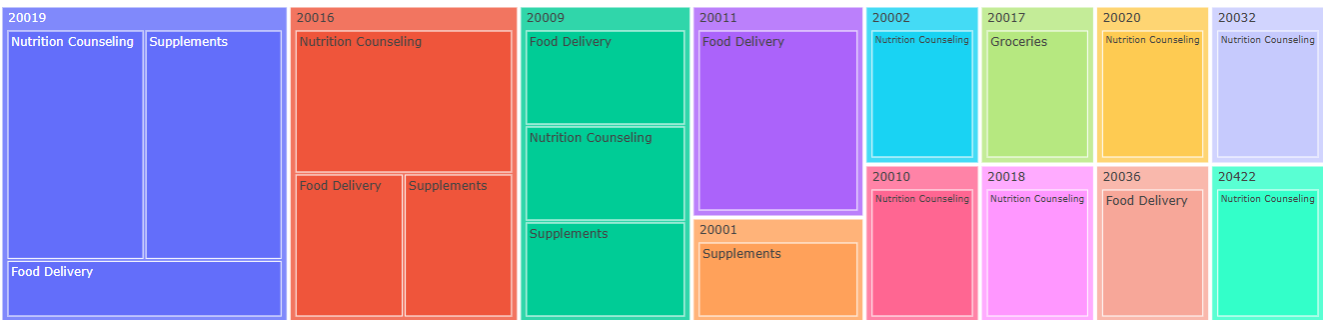
Key Observations

- ZIP Code 20019 is the Most Diverse and Densely Served Area: With prominent blocks for Nutrition Counseling, Supplements, and Food Delivery, ZIP 20019 emerges as a highly serviced area offering a broad spectrum of nutrition support.
- ZIP 20016 Offers a Balanced Mix of Services: This ZIP code shows nearly equal representation of three key services, indicating strong community investment and potentially high demand in that region.
- ZIP 20009 Has Strong Coverage with Three Core Services: Nutrition Counseling, Food Delivery, and Supplements are well represented, showcasing it as another multi-service neighborhood with holistic outreach.
- ZIP Codes with Single-Service Categories May Indicate Limited Access or Specialized Programs: Areas like 20020, 20032, 20036, and 20422 each show only one service category, mostly Nutrition Counseling or Food Delivery—suggesting narrower program availability in those ZIPs.
- Service Diversity Is Not Evenly Distributed Across ZIPs: Some ZIP codes have multiple stacked services (e.g., 20009, 20019), while others are limited to a single category. This points to geographic disparities in service availability and the need for more even distribution of holistic care.

5. Service Distribution Across Census Tracts


Service Distribution Across Census Tracts

Key Observations

- Uneven Distribution of Services Across Census Tracts: The line plot shows clear peaks and valleys, indicating that services are concentrated in a few census tracts, while many others have only one or no services at all.
- Census Tract in the 10,000 Range Has Highest Service Count: One census tract near 10,000 stands out with 5 services, making it a significant hub for food and nutrition access—possibly overlapping with highly populated or underserved regions.
- Low-Service Tracts Are Common: Several tracts have only 1 service, highlighting widespread underrepresentation in those zones and suggesting a potential need for outreach or redistribution.

- Isolated Spikes Reflect Targeted Service Allocation: The few spikes (e.g., near tracts 2,000, 4,500, and 10,000) suggest that service delivery may be influenced by local nonprofit initiatives or municipal funding tied to specific neighborhoods.
- Line Plot as a Trend Proxy for Geographic Spread: Although not temporal, the plot mimics a "trend" across space, revealing how food access ebbs and flows geographically through D.C.'s neighborhoods.

6. Service Location Density Map
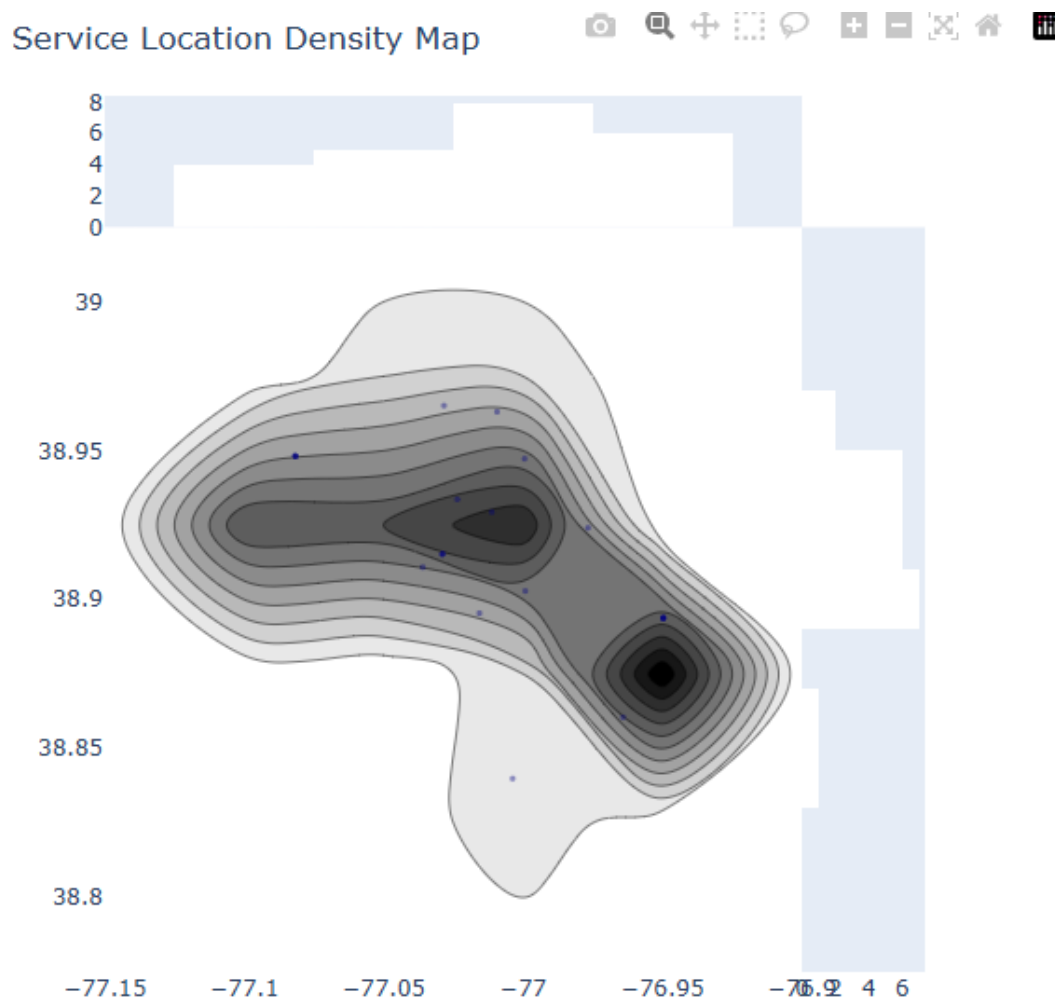
Key Observations
- Two Primary Service Hotspots Identified: The darkest regions on the map represent highest service concentration zones, with one cluster around longitude -76.95 and another near longitude -77.05, likely corresponding to neighborhoods in Northeast and Northwest Washington, D.C.
- Core Urban Area Receives Most Coverage: The central contours indicate that urban neighborhoods are more heavily serviced, which aligns with population density and accessibility to resources.

- Sparse Coverage in Peripheral Areas: The lighter outer contours and absence of blue dots around the edges highlight lower service density in suburban or fringe areas, suggesting potential outreach gaps.
- Visual Correlation with Zip and Ward Distribution: This heat-based density map visually reinforces earlier findings from the ward and ZIP code bar charts, where Ward 7 and ZIP 20019 also showed high service counts.
- Useful for Planning Resource Allocation: Density visualization like this is valuable for targeting underserved areas and optimizing the geographic deployment of new food or nutrition services.

Conclusion

This project provided a comprehensive geospatial analysis of food and nutrition service providers across Washington, D.C., using interactive visualizations to uncover critical insights into service availability, distribution, and accessibility.

By examining multiple dimensions—such as ZIP codes, wards, census tracts, and service categories—we observed that food and nutrition services are unevenly distributed, with clear clusters in central and northeastern wards like Ward 7 and ZIP 20019. The most commonly offered service was Nutrition Counseling, followed by Food Delivery, highlighting a dual focus on education and direct assistance.

Visual tools like scatter maps, density contours, treemaps, and bar charts revealed both highly served areas and underserved neighborhoods, enabling data-driven identification of potential service gaps. While certain wards and tracts demonstrate strong infrastructure and organizational involvement, others lack representation and may benefit from strategic expansion.

Overall, this analysis not only maps the current landscape but also serves as a decision-support tool for policymakers, nonprofits, and city planners seeking to enhance food security, optimize resource allocation, and promote equitable access to nutritional support across all communities in the District of Columbia.