Detecting food deserts by analyzing food delivery data

Analysis of food delivery demographics

Sitara Anand
University of Pittsburgh
Pittsburgh, PA
sra56@piit.edu

ABSTRACT

Food delivery services have emerged as a critical component of contemporary society, significantly influencing the accessibility and distribution of prepared meals. Prominent platforms such as Uber Eats, DoorDash, and Grubhub have facilitated the expansion of traditional restaurants' takeout capabilities, enabling these establishments to reach a broader customer base without the need to develop proprietary delivery infrastructure. The operational parameters of these services, including delivery time and geographic radius, are subject to fluctuations based on demand peaks. To optimize efficiency and maintain service quality, delivery companies use advanced algorithms that dynamically calculate delivery times and routes.

CCS CONCEPTS

Artificial Intelligence → Maching learning;

KEYWORDS

Uber Eats restaurants → delivery information

ACM Reference format:

1 Introduction

The primary objective of this research is to systematically evaluate food delivery services within a defined region, incorporating critical factors such as pricing and clustering. This study seeks to:
Identify and assess the distribution of food delivery services by analyzing spatial patterns across urban and suburban environments, with particular attention to the emergence of food delivery deserts—areas with limited or no access to delivery platforms.
Investigate equity in service availability by determining whether food delivery options are fairly distributed among diverse geographic and demographic areas.
Pinpoint underserved communities that lack adequate access to food delivery platforms, thereby highlighting potential gaps in service provision.
Analyze trends in food delivery, including categorization by cuisine or food type, to uncover broader patterns and shifts within the industry.
Visualize and interpret collected data to reveal correlations among variables such as cost, service availability, and restaurant types.
Examine cost structures and service options to understand better the economic landscape of food delivery and its implications for consumers.
Examine cost structures and service options to understand better the economic landscape of food delivery and its implications for consumers.
Assess the relationship between income levels and access to food delivery services, and whether lower-income areas are disproportionately affected by limited-service availability.
Conduct fundamental demographic analysis, such as evaluating population size and density, to contextualize findings and support more nuanced interpretations of service distribution and equity.
By addressing these objectives, the research aims to provide actionable insights into the accessibility, equity, and trends of food delivery services, supporting both academic inquiry and practical decision-making in the field.

Rows after cleaning:	2814292
Median Income range	$30907.00 - $112376.00
Number of unique zip codes:	1646
Number of unique restaurants:	30175
Number of unique menu items:	484593

Table 1: Raw data analyzed for this paper

Merged Data Description.

1	id	float64
2	restaurant_name	object
3	score	object
4	ratings	object
5	cuisine	object
6	price_range	object
7	full_address	object
8	zip_code	float64

Detecting food deserts by analyzing food delivery data

9	lat	float64
10	lng	float64
11	restaurant_id	float64
12	menu_category	object
13	menu_item	menu_item
14	description	object
15	price	float64
16	city	object
17	state	object
18	county	object
19	zip_lat	float64
20	zip_lng	float64
21	population	float64
22	median_income	float64

The features we use were sourced from the Uber Eats dataset with additional information from the USDA median incomes for metro areas.

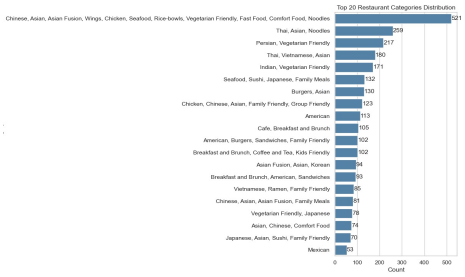
1.1 Finding existing correlations in the data

Before undertaking the primary analysis to identify food delivery deserts, it is essential to conduct a preliminary examination of the dataset's fundamental correlations.

1.1.2 Categorization of menu items

Analysis: Restaurants are systematically classified according to the cuisine they offer. This categorization facilitates subsequent analyses of food delivery patterns, enabling the identification of the most frequently ordered cuisines and their distribution across different neighborhoods.

Analysis Categorization of data



1.1.3 Geographic clustering of restaurants

The simple graph shows all 30175 restaurants in the dataset to review their geographic distribution. This will provide the basis for geographic clustering.

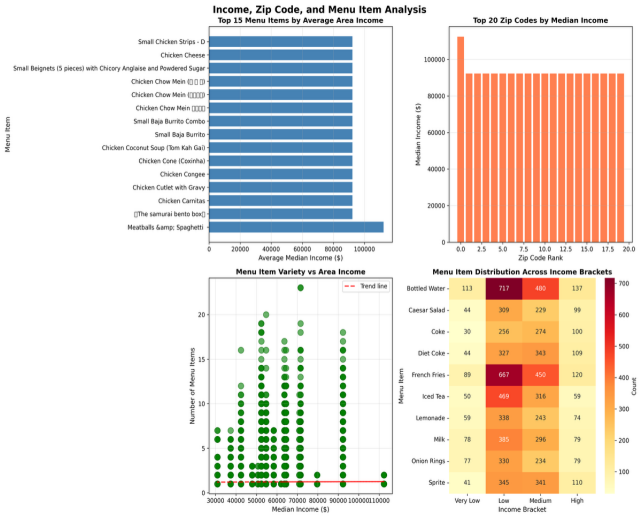
1.1.4 Correlation of income, zip, and Menu items ordered

Conducting an analysis of the correlation between median income, zip code, and menu item popularity enables the identification of patterns in consumer preferences across different geographic and socioeconomic areas.

Conclusion: Analysis of the top 20 menu items reveals an equitable **distribution** across income levels, indicating that consumer preferences for these items are consistent regardless of socioeconomic status

The second analysis, which examines the number of items ordered, reveals a positive correlation between income level and order volume; as income increases, so does the number of items ordered. Notably, the highest concentration of orders is observed among households situated within the middle of the median income range.

Detecting food deserts by analyzing Delivery data



1.1.5 Identification of food delivery deserts

This is the paper's primary analysis. Identifying food delivery deserts is the main objective of the analysis.

A food delivery desert is an area where residents have limited or no access to food delivery services, such as those provided by platforms like Uber Eats, DoorDash, or Grubhub. These areas are characterized by a lack of available restaurants participating in delivery platforms, restricted delivery radii, or operational limitations that prevent food delivery companies from serving specific neighborhoods. Food delivery deserts often overlap with traditional food deserts—regions where access to healthy, affordable food is limited due to socioeconomic factors, geographic isolation, or inadequate transportation.

Major food delivery platforms, including Uber Eats, DoorDash, and Grubhub, typically operate within a delivery radius of approximately ten miles from participating restaurants. The acceptance of delivery requests is further influenced by periods of peak demand, with companies employing advanced algorithmic systems to dynamically adjust delivery availability and optimize service efficiency.

As a result of conducting this analysis, we can identify the zip codes that are provided in the dataset and identify the areas that are covered.

The main algorithm uses a well-known formula to calculate the distance between 2 points.

a = sin(dlat / 2) ** 2 + cos(lat1) * cos(lat2) * sin(dlon / 2) ** 2
c = 2 * atan2(sqrt(a), sqrt(1 - a))
distance = R * c

Explanation of the formula

This formula calculates the straight-line distance between two points on Earth's surface (like the distance "as the crow flies").

Here's what each part does:

R = 3959 - This is Earth's radius in miles, which we need for the final calculation.

Converting to radians: GPS coordinates are usually in degrees (like 40.5° N), but trigonometric functions need radians, so we convert all four coordinates (two latitudes and two longitudes).

dlat and dlon - These are just the differences between the two latitudes and the two longitudes.

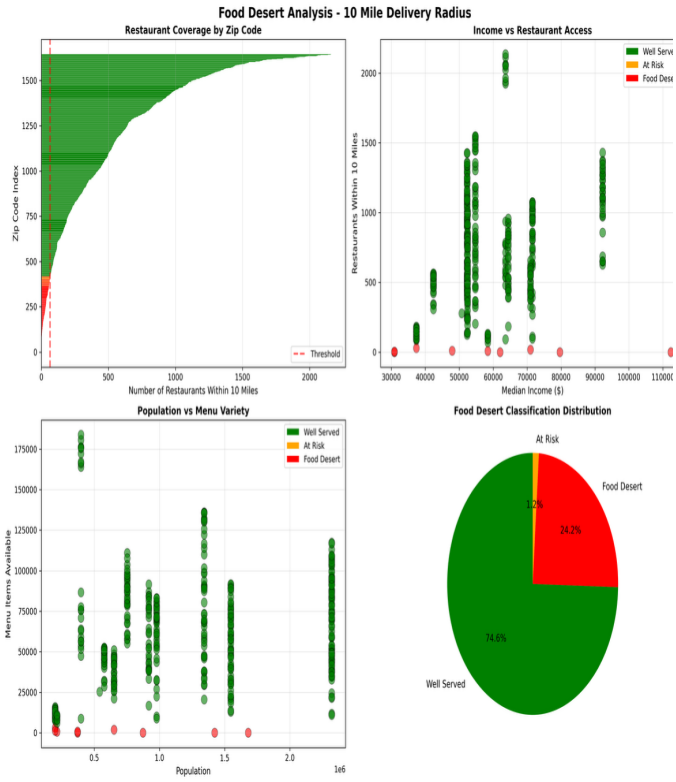
The "a" calculation - This is the heart of the formula. It uses sine and cosine functions to account for Earth's curvature. Think of it as figuring out how much of a "chord" you'd need to connect the two points through Earth's sphere.

The "c" calculation - This converts the chord measurement into an actual angle (in radians) from Earth's center.

distance = R * c - Finally, we multiply that angle by Earth's radius to get the actual distance in miles.

The formula was used based on the latitude and longitude to calculate 10-mile radius for each row. This provided the basis for calculating the food deserts.

Detecting food deserts by analyzing food delivery data



Conclusion

Analysis 1. Identify the restaurant coverage by zip code.

Preliminary analysis indicates that restaurants are disproportionately concentrated within specific zip codes. Even a basic spatial assessment reveals that specific neighborhoods lack food delivery coverage within a ten-mile radius. Despite the limitations of the available dataset, it is evident that food delivery deserts exist, with some areas experiencing significant gaps in access to delivery services.

Analysis 2. **Correlation between median income and food delivery deserts.**
The analysis revealed no significant correlation between median income and the extent of food delivery coverage.

Analysis 3. **Population vs Menu Variety**
The analysis revealed no significant correlation between population menu items and the extent of food delivery coverage.

Analysis 4 : **Food Delivery Deserts**

The algorithm was developed to identify food delivery deserts by determining whether zip codes in the dataset fell outside a ten-mile radius from participating restaurants. Using this approach, the analysis concluded that approximately 24.2% of restaurants were located in areas classified as delivery deserts. Given that the dataset was limited to Uber Eats data, this identified food delivery deserts are more accurately described as "Uber Eats deserts".

AKNOWLEDGMENTS

REFERENCES

Primary Dataset: Uber Eats USA Restaurant Menus
Source: Kaggle
Citation: Sharma, A. (2021). Uber Eats USA Restaurant Menus [Data set]. Kaggle. https://www.kaggle.com/datasets/ahmedshahriarsakib/uber-eats-usa-restaurants-menus

Content: Contains comprehensive restaurant information, including: Geographic coordinates (latitude/longitude),
Restaurant categories and cuisines
Pricing information
Ratings and reviews
Delivery availability and estimated delivery times
Coverage across multiple major U.S. cities

Other sources: Median income data from the us census data, population
https://data.census.gov/map?q=Income+by+Zip+code+tabulation+area&layer=VT_2022_860_Z2_PY_D1&loc=43.3751,-113.1138,z2.6270