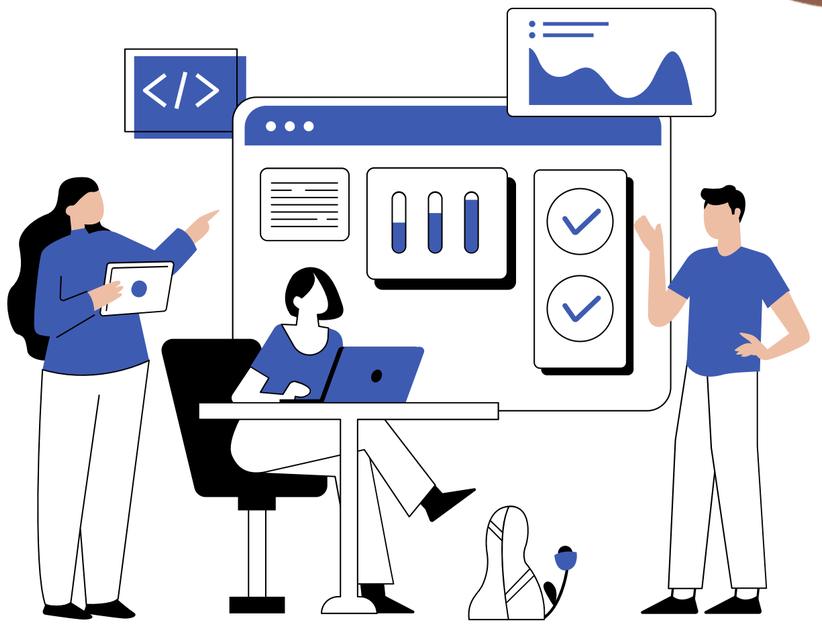
Exploring Culinary Landscapes:
An Interactive Analysis of Restaurant Data

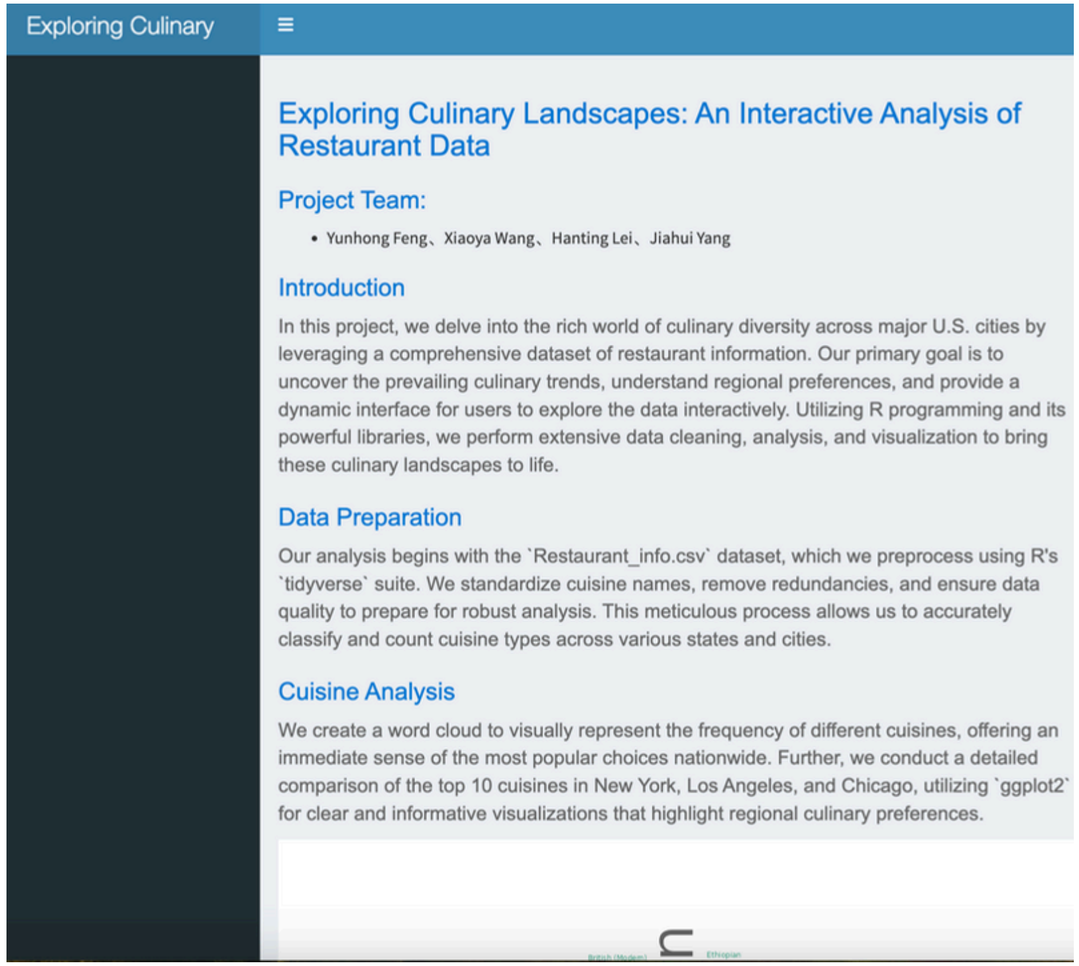


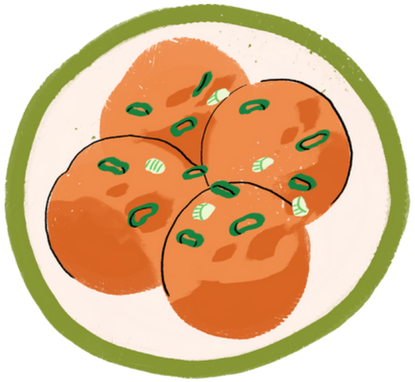
Team Members: Yunhong Feng, Xiaoya Wang, Hanting Lei, Jiahui Yang

Introduction

The primary goal of the project is to uncover and analyze the predominant culinary trends and regional food preferences, offering both insights and a practical tool for users to explore these trends interactively.

We aim to provide an in-depth understanding of how various cuisines distribute themselves across urban areas, highlighting not only popular food choices but also niche markets and emerging trends. The interactive features of our website encourage users to engage directly with the data, making the exploration of culinary data both informative and accessible.



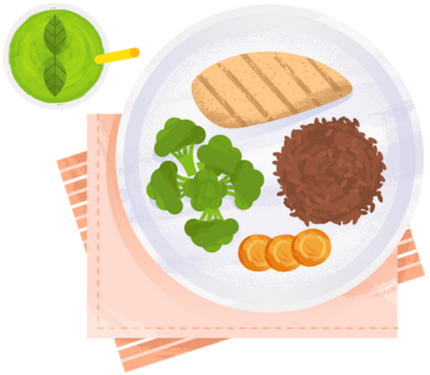


address	categories	primaryCategories	city	cl
1045 San Pablo Ave	Restaurant,Asian/Pacific,Cafe,Vegetarian / Vegan Restaurant,Vegetarian,Vegetarian Vegan Restaurant,Caf,Thai	Accommodation & Food Services	Albany	
1045 San Pablo Ave	Restaurant,Asian/Pacific,Cafe,Vegetarian / Vegan Restaurant,Vegetarian,Vegetarian Vegan Restaurant,Caf,Thai	Accommodation & Food Services	Albany	
1045 San Pablo Ave	Restaurant,Asian/Pacific,Cafe,Vegetarian / Vegan Restaurant,Vegetarian,Vegetarian Vegan Restaurant,Caf,Thai	Accommodation & Food Services	Albany	
1045 San Pablo Ave	Restaurant,Asian/Pacific,Cafe,Vegetarian / Vegan Restaurant,Vegetarian,Vegetarian Vegan Restaurant,Caf,Thai	Accommodation & Food Services	Albany	
1045 San Pablo Ave	Restaurant,Asian/Pacific,Cafe,Vegetarian / Vegan Restaurant,Vegetarian,Vegetarian Vegan Restaurant,Caf,Thai	Accommodation & Food Services	Albany	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	
16411 Bernardo Center Dr	Indian Restaurant,Restaurant,Caterers,Food Dining,Catering,Vegetarian Restaurants,Restaurants,Indian Restaurants,Food & Dining,Middle Eastern Restaurants	Accommodation & Food Services	San Diego	

Data Source

"Uber Eats USA Restaurants and Menus" dataset available on Kaggle:
This dataset has been extensively modified and merged to form the final dataset, **Restaurant_info.csv**, which incorporates various attributes relevant to our analysis. This includes restaurant names, addresses, cuisines, pricing information, and geolocations. The original data offers a snapshot of the diverse culinary offerings across multiple U.S. cities, capturing details that range from menu items to restaurant categories.

Data Processing



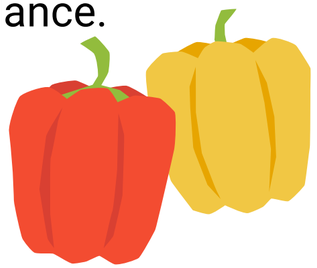
- **Data Cleaning and Transformation**
  - Separation and Trimming: We began by separating multiple cuisines listed together into distinct rows using the separate_rows function, ensuring each cuisine is analyzed individually. This was essential for accurately categorizing restaurants by their offered cuisines.
  - Normalization: Cuisine names were standardized by trimming spaces and removing redundant suffixes like "Restaurants," ensuring consistency across the dataset.
  - Categorization: We applied conditional logic to rename vague terms like "Veg" to "Vegetarian" to clarify the dataset and focus on dietary-specific trends.
- **Data Selection:**
  - We refined the dataset by selecting relevant columns such as name, city, address, cuisines, state, and pricing, which are crucial for our geographic and pricing analysis.
- **Geolocation Processing:**
  - For restaurants in key areas like New York and Brooklyn, we specifically selected and mapped their geolocations to visualize the distribution of popular cuisines, enhancing the geographic precision of our analysis.

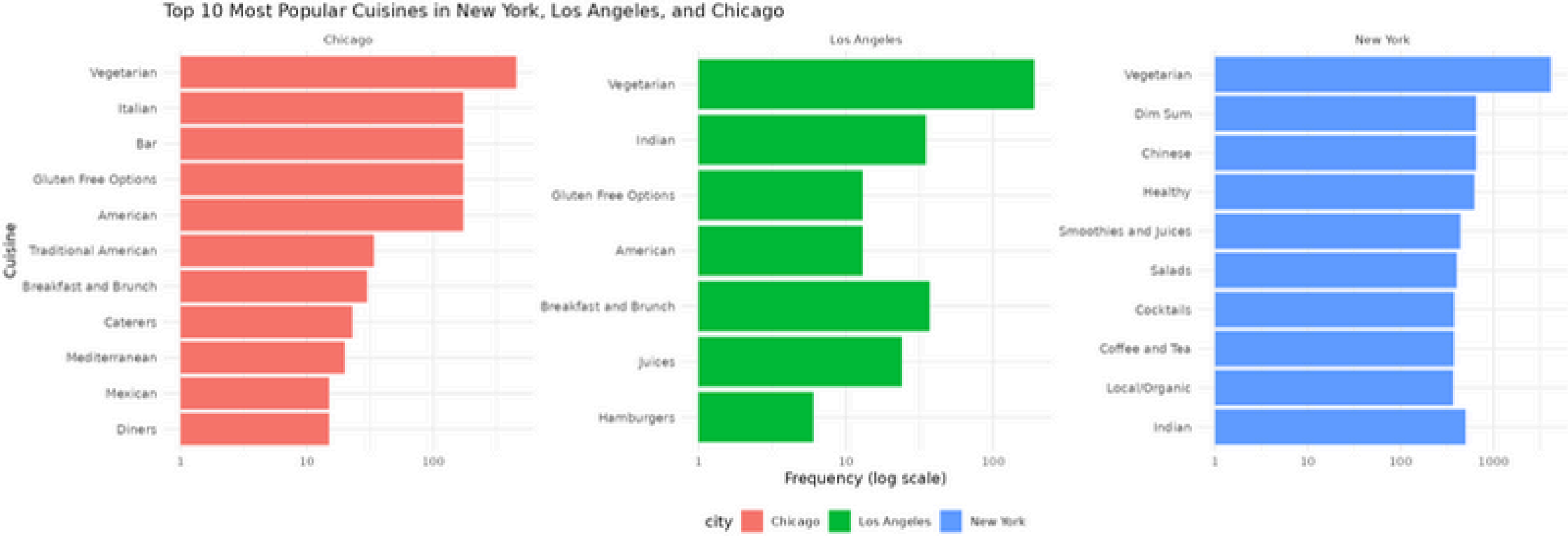
Visualization and Analysis Methods

Word Cloud of Cuisine Popularity

The word cloud serves as a dynamic gateway into the landscape of U.S. culinary tastes, graphically emphasizing the prevalence of various cuisines. By representing cuisines with different font sizes based on their frequency in the dataset, this visualization instantly communicates which flavors dominate the American culinary scene.

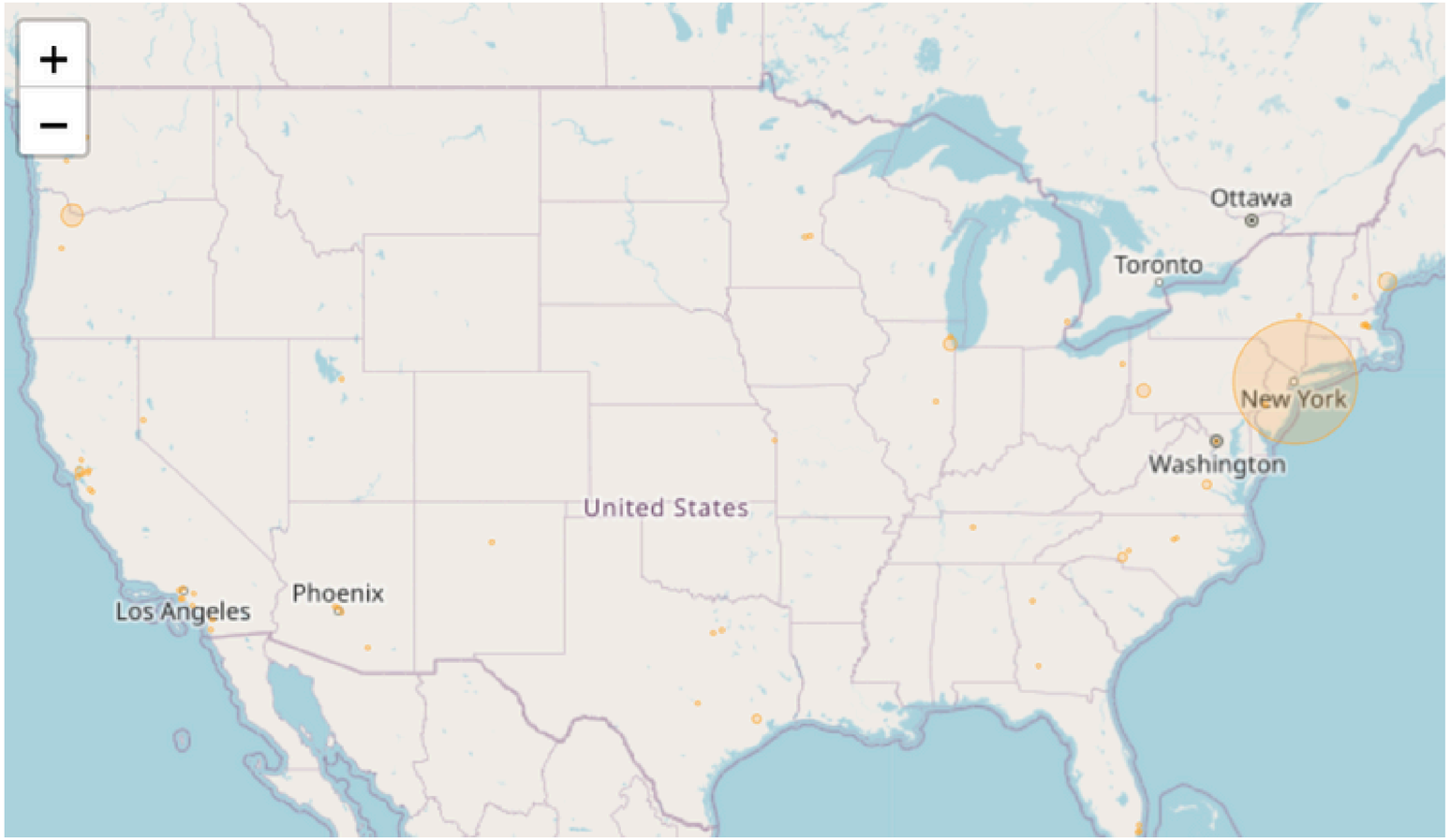
It captures visual interest and provides a quick, intuitive understanding of national food preferences, serving as a compelling summary for viewers to grasp the broader trends at a glance.





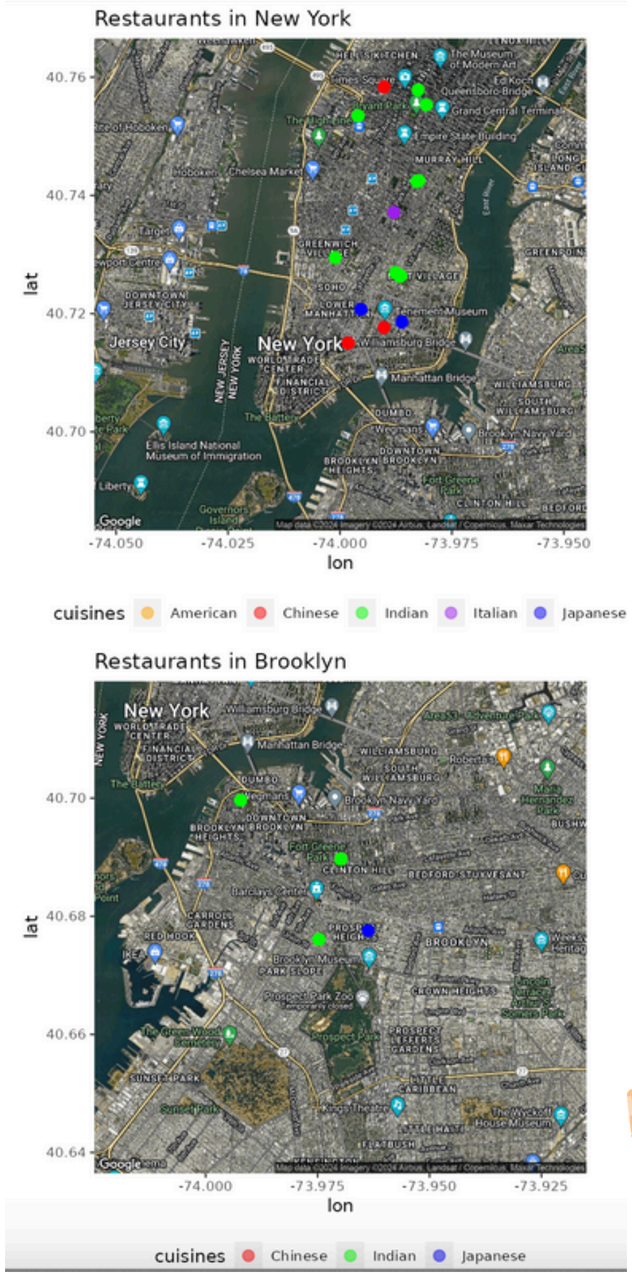
Top 10 Cuisines in Major Cities (New York, Los Angeles, Chicago)

By highlighting differences and similarities in culinary tastes across these cities, the visualizations not only offer insights into localized food cultures but also reflect how diverse cultural influences shape regional food scenes. For businesses and culinary enthusiasts, understanding these patterns can inform decisions about menu offerings and marketing strategies.



Vegetarian Landscape Map

This visualization is particularly useful for identifying regions with a higher density of vegetarian cuisine, which can help diners make informed choices and restaurant owners to spot potential areas for business expansion. It also reflects growing trends in health and dietary preferences, highlighting how urban centers adapt to dietary needs.



Diverse Cuisines of New York City

This detailed mapping of five popular cuisines in New York City goes beyond mere location plotting. Our analysis dives into five popular cuisines - Indian, Japanese, Chinese, Italian, and American - showcasing their presence across different neighborhoods. It provides a snapshot of the city's vast culinary diversity, breaking down the distribution by neighborhoods and showing how culinary cultures permeate different parts of the city. This visualization aids potential diners in locating their cuisine of choice and illustrates the mosaic of food options available.

Technologies and Tools

01

R and Shiny: The core of our application is built using R, combined with Shiny for creating interactive web applications. Shiny allows us to build user-interactive elements easily, integrating live data processing with user input.

02

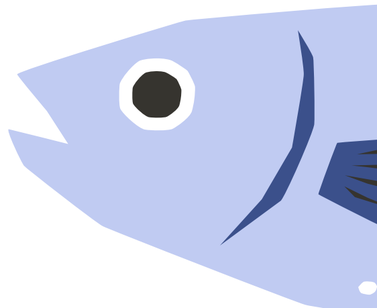
ggplot2 for Visualization: We employ ggplot2 to build bar charts and maps.

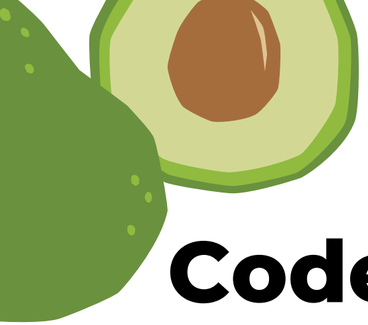
03

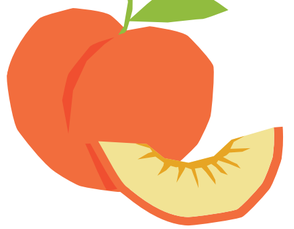
Leaflet for Mapping: Geographic data visualization is handled using the Leaflet package, which enables interactive map interfaces within R. This tool is crucial for implementing the dynamic maps that display various data points based on user interaction.

04

Tidyverse for Data Manipulation: The suite of data manipulation tools within Tidyverse (including dplyr and tidyr) is essential for preparing our dataset. These tools help us clean, organize, and transform data efficiently.







Code Implementation and Functionality

01

- **Word Cloud Generation**

```
wordcloud(words = cuisine_df$Cuisine, freq = cuisine_df$Frequency, min.freq = 1, max.words = 200, random.order = FALSE, rot.per = 0.35, colors = brewer.pal(8, "Dark2"))
```

This snippet generates a word cloud from our cuisine data, highlighting the popularity of different cuisines with varying sizes. It's a visually engaging way to start our analysis, giving a quick overview of popular trends.

02

- **Top Cuisine Analysis in Major Cities:**

```
ggplot(data = combined_cuisines, aes(x = reorder(cuisines, n), y = n, fill = city)) + geom_col() + coord_flip() + labs(title = "Top Cuisines")
```

ggplot2 is used to create a comparative bar chart of the top cuisines in New York, Los Angeles, and Chicago, showcasing regional preferences distinctly.

03

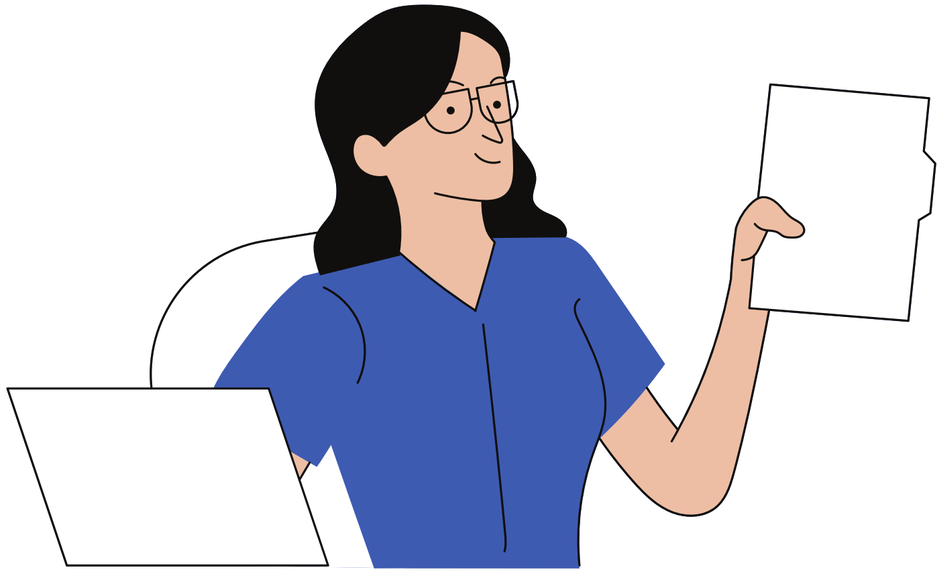
- **Interactive Maps for Vegetarian Distributions:**

```
leaflet(data = city_map_data) %>% addTiles() %>% addCircles(lng = ~long, lat = ~lat, radius = ~n * 50, color = '#ffa500')
```

This code utilizes Leaflet to plot circles on a map indicating the concentration of vegetarian restaurants, with circle size proportional to the number of options in each city.

Introduction to Interactive Features

We ensure that our interactive features not only serve analytical purposes but also engage and educate users by providing intuitive and dynamic tools. Our application utilizes R Shiny to deliver these capabilities.



01

Dynamic Filters: Users can interactively adjust their views by filtering data based on cuisine types, price ranges, or cities.

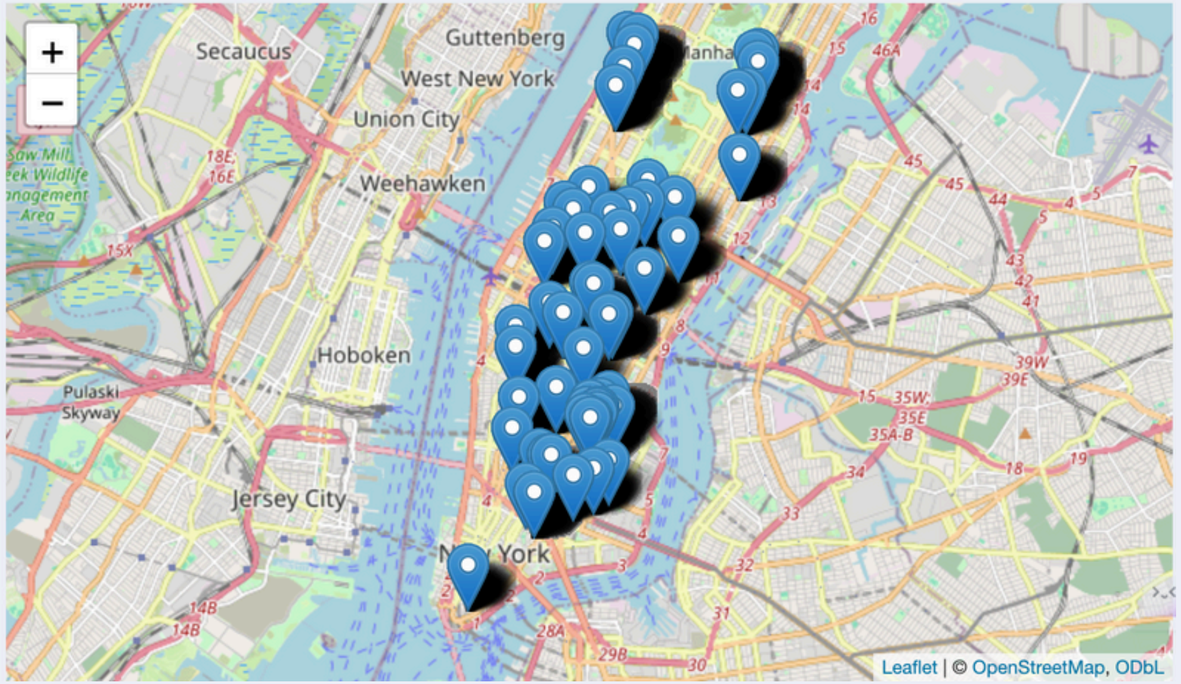
02

Interactive Maps: Utilizing the leaflet library, we offer maps that detail the geographical distribution of restaurants. Users can zoom in/out and click on markers to get information about each restaurant, enhancing their ability to explore data spatially.

Interactive Maps

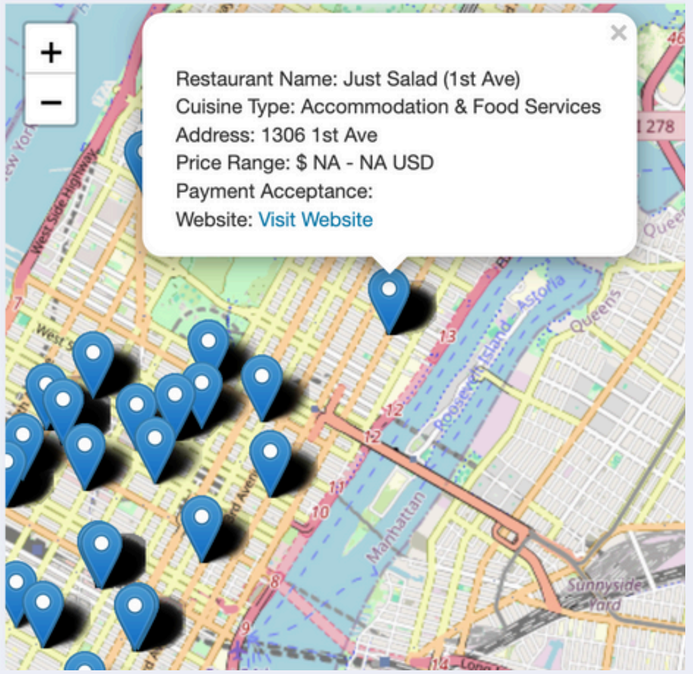
Choose a city:

New York



Choose a city:

New York

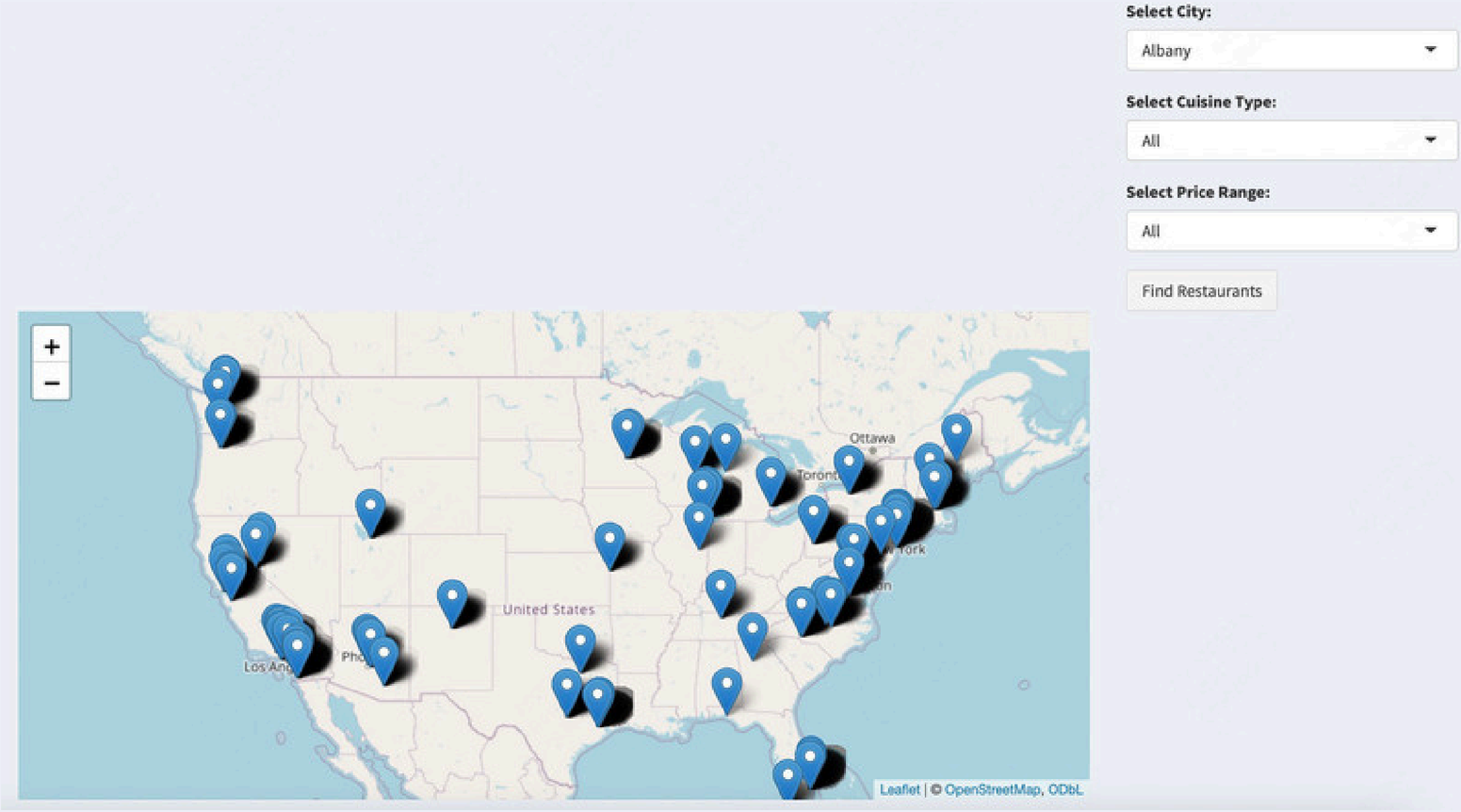


01 In our interactive map feature, users can effortlessly navigate the rich culinary tapestry of different places.

02 The map is dotted with location markers that represent individual restaurants, such as 'Just Salad' on 1st Ave, providing users with a bird's-eye view of dining options.

02 Upon selecting a marker, a popup window reveals the restaurant's key details, including name, cuisine type, address, and an external link to the restaurant's website for further exploration.

Interactive Maps



01 In our interactive map feature, users can effortlessly navigate the rich culinary tapestry of different places.

02 The map is dotted with location markers that represent individual restaurants, such as 'Just Salad' on 1st Ave, providing users with a bird's-eye view of dining options.

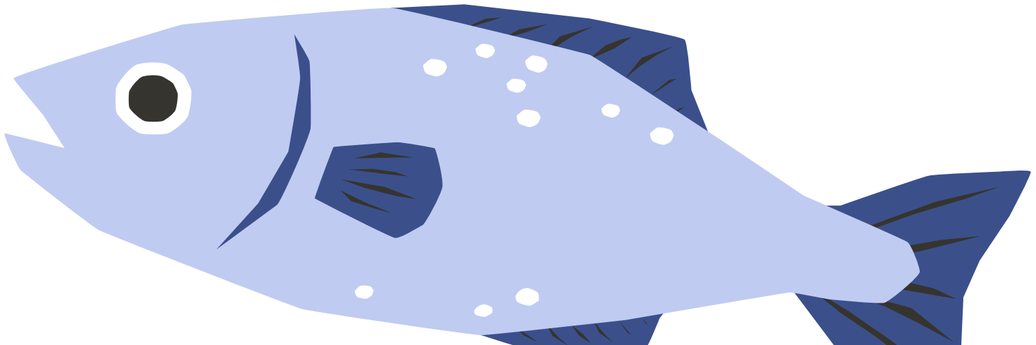
02 Upon selecting a marker, a popup window reveals the restaurant's key details, including name, cuisine type, address, and an external link to the restaurant's website for further exploration.

Project Challenges

01 **Data Complexity and Volume:** Handling and processing a large dataset with varying data quality required robust cleaning and standardization procedures.

02 **Performance Optimization:** Ensuring the app performs efficiently with real-time data updates and user interactions.

03 **User Experience Design:** Creating an intuitive interface that accommodates complex data visualizations and interactive elements.



Solutions

01 Utilized **tidyverse** for efficient data manipulation, enhancing our ability to clean and preprocess data effectively.

02 Optimized back-end processing and implemented reactive programming in Shiny to handle user inputs dynamically without performance lags.

03 Designed the UI with **shinydashboard** for a structured layout, focusing on usability testing to refine user interactions and visual design."

Outcomes and Reflections

01

Key successes:

- Effective data visualization that aids in understanding complex data easily.
- High user engagement levels demonstrated through interactive elements like maps and dynamic filters.

02

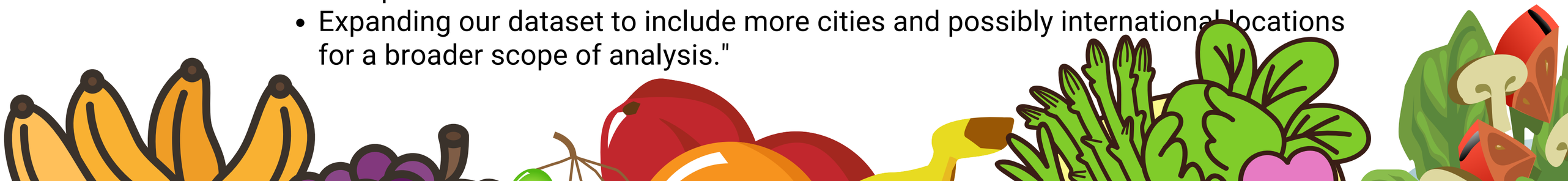
Areas for Improvement:

- Some users reported challenges in navigating the full scope of data available. More intuitive navigational aids could enhance usability.
- Data updates and maintenance could be automated to ensure the most current information is available.

03

Future Directions:

- Incorporating machine learning to predict trends and suggest restaurants based on user preferences.
- Expanding our dataset to include more cities and possibly international locations for a broader scope of analysis."



Any questions?