Name: Yunzhi Chen
Student ID: 32051018

Restaurants near me - Uber Eats

FIT5147 Data Visualisation Project
Tutor: Deep Mendha
Tutorial 10

Table of Contents

Introduction

In 2014, Uber launched an online food ordering and delivery platform which is called Uber Eats. Users can use a mobile app, or through a web browser, to read menus, view restaurant locations and ratings, order and pay for food from participating restaurants. Besides, with Uber Eats delivery, all people favorite foods are right at doors with just a tap of phone.

This report aims to illustrate the results found in Data Exploration Project, which uses R for data wrangling and R shiny to demonstrate visualization. The critical message shows the distribution of popular restaurants by price type and the geographical location of popular restaurants on the US map. In addition, a visual approach to expressing the flow route of people spending on food, rather than simply using tables.

The intended target audience for communicating the key information will be Uber Eats users, who will have a new perspective on U.S. restaurant data by communicating key information about nearby restaurants, choosing restaurants more informative and mutually beneficial for both restaurants and customers, thus contributing to the economic growth of the restaurant industry and even the takeout industry. The integration of restaurant data also makes it easier for people to understand the overall information of the nation's gourmet restaurants and provides a more comprehensive picture for data analysts.

Design

The Five Design Sheets approach (FdS) is a method for generating ideas for the final visualization of a project that shows the thought process of how visualization works and how to answer questions. FdS requires designers to create five design sheets. It relies on brainstorming to expand the design space of possibilities, centered on the project, and then three separate design sheets will record the three ideas from the initial brainstorming (Sheets 2, 3, and 4). Detailed information on how the visualization artifacts will work or will be created is then shown in Sheet 5. More details of each design sheet are shown in the Appendix part.

In my first sheet, I came up with 14 sheets to answer my intended questions, but I filtered four plots because they are in varying degrees not suitable. I categorized and combined the remaining visualization methods in conjunction with the research problem.

Secondly, I used a pie chart plus a select bar to interactively show the share of different price types of restaurants with different reputations, then a dot map to show the distribution of restaurants in the country, and finally a flow map to show the traveling

flow route of people for food. Drop-down menu to select different regions for display. But according to Munzner's (2014) suggestion, I found that neither a pie chart nor a dot map is a good choice, because the same data attributes encoded using different visual variables will not be perceived effectively.

And then, I chose to use a density plot with drop-down options, then a leaflet map to show the distribution of restaurants by price type across the country, and finally a Sankey diagram to show the route information. All things being equal, visual channels with as little interference as possible should be used. Visual variables like color and position are separable because there is little interference between them. So, I chose different colors to represent the density of restaurants with different price types. But the Sankey diagram is not a good choice as it only shows the starting and ending points, not the distance and the exact route.

For sheet4 I used a combination of a bar chart, heat map and flow map, and two select bars to be able to select different scores & rankings and different zip codes respectively. However, the heat map destroys the rule of discriminability, which means it cannot allow the reader to easily distinguish, compare data and identify specific values, which is why I did not choose it in the end.

Sheet5 is to combine sheet2 3&4, which to choose a bar chart, leaflet map, and flow map, using select bar interactively choose different information to present. The limited capacity of working memory means that we are better off encoding multiple attributes into a single visual object, rather than using a separate visual object for each attribute so that more information can be held in visual working memory. I decided to use a leaflet map to reflect the distribution of restaurants with different price types and different evaluations across the United States, and I can see the number of them, users can choose different price types, show the total number of restaurants in particular areas on the map, and then by clicking marker, they can get these restaurants information in detail.

Implementation

In this project, the tool used is R shiny for chart interaction and visualization. The libraries that I used are "shinythemes", "tidyverse", "lubridate", "plotly", "leaflet", "maps", "geosphere", "RColorBrewer". The library of "shinythemes" is used to change the basic theme of the shiny app. "lubridate" and "maps" are for data wrangling and cleaning, and the rest of libraries such as "plotly", "leaflet", "geosphere" and "RColorBrewer" are all as part of demonstrating the interactivity and visualization of the shiny app.

However, I made some changes to the final app based on the 5 design sheets to better represent the visualization. The difference with the 5 design sheets is that in part1 (Price situation of popular restaurants), I used the 'facet_wrap' function to separate the bar

chart by price type, the reason for this is that the number of restaurants in different price categories varies greatly, and without separation, it would be impossible to identify the number of smaller data. I then added the exact value of each bar to allow color-blind users to understand the data as well. I also added generalized text at the end of this section to summarize the distribution of price types of restaurants with good reviews.

In the part of distribution of popular restaurants, I added the price type to the marker, so the information is not only highlighted from select panel but also the marker. Users can click on the marker to get detailed information about the restaurant, including restaurant name, full address, price type, rating, etc. if they want to know.

One more change is in the traveling flow for delicious foods section. I changed the interactive section in the original design from a select drop-down menu to a checklist box, also changed the selection from zip code (show region) to end city, which is another interactive way to make my interactive section not monotonous.

My project also has the following challenge. Since my flow map is represented by a leaflet, I need to spend a lot of effort upfront to process the uber data, add the latitude and longitude for the starting and ending cities respectively and then integrate the data to form a form that can be used to display the flow eventually.

User guide

Firstly, to run the R shiny application, click on the "Run App" button on the RStudio page. A window will pop up then press on "Open in Browser" to show the full view of the webpage.

At the top of the app are the project name and four selectable sections (Figure1), namely "Price Type Situation", "Restaurants Distribution Map", "Traveling Flow Map" and "About". "Traveling Flow Map" and "About". Users can switch between the different sections by clicking on them.

Restaurants near me - Uber Eats $ Price Type Situation   🍴 Restaurants Distribution Map   🚗 Traveling Flow Map   💬 About

Figure1

On the "Price Type Situation" page (Figure2), users can switch between different ranking types to see the number of restaurants of different price types. Mouse over the bar chart can get detailed information interactively.
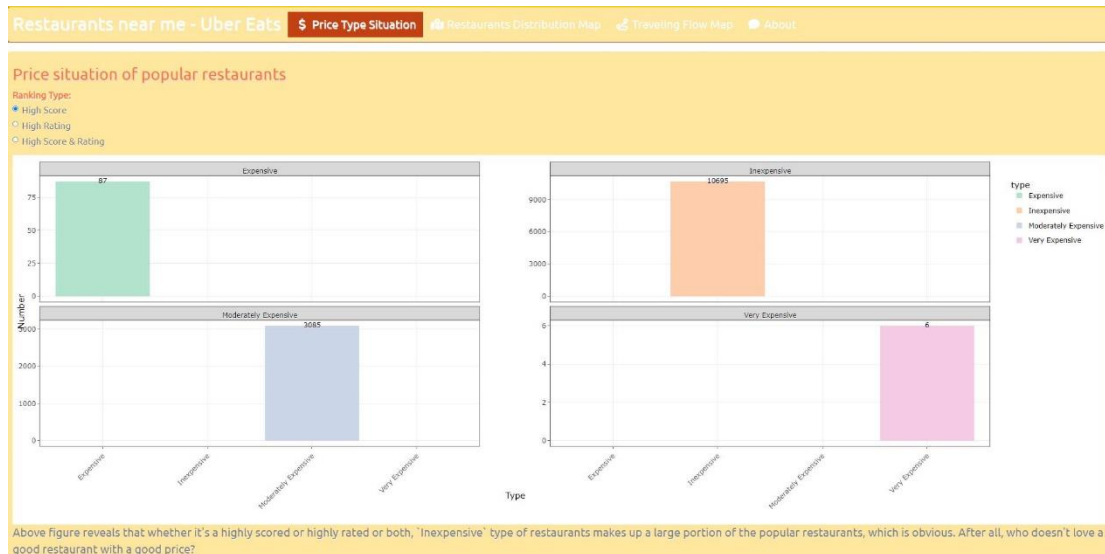
Figure2

In the "Restaurants Distribution Map" section (Figure3), users can initially see the distribution and number of popular restaurants across the US in the "Inexpensive" and "Expensive" price categories. Users can add and remove price types from the drop-down menu to view the selection. Since there are no hot restaurants in the "Very Expensive" category, I removed that option. Zooming in on the map can show more detail about each restaurant's location. By clicking on the marker (Figure4), people can also get a lot of detailed information about the chosen restaurant.
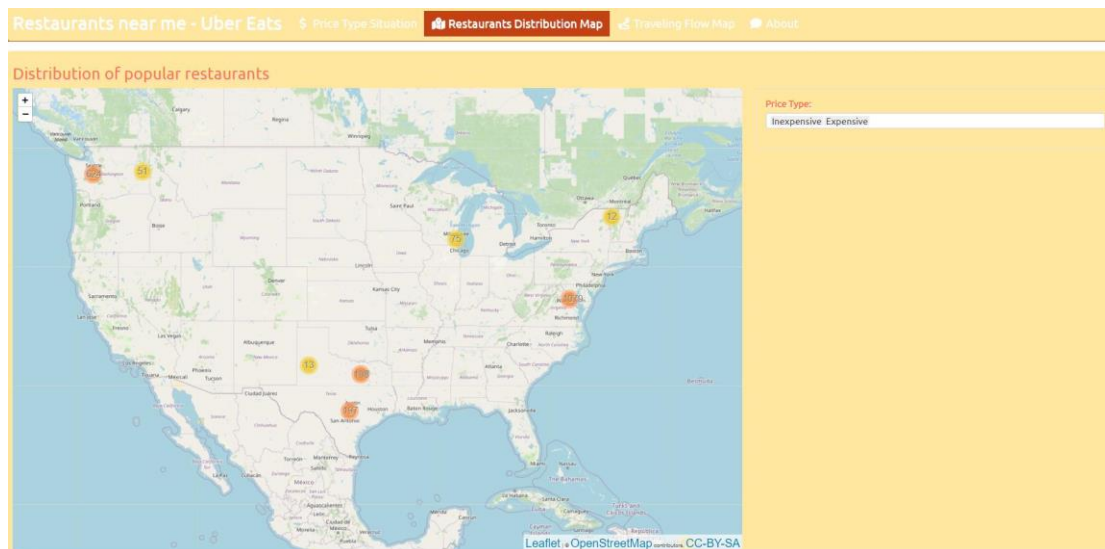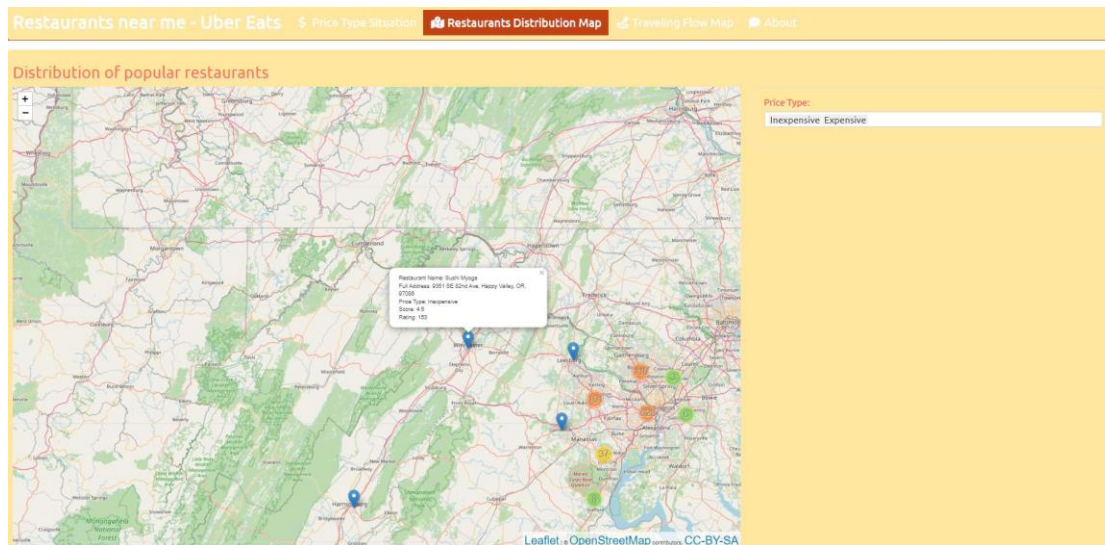

Figure3

Figure4

In the last part of the presentation (Figure5), the detailed route can be obtained by clicking on the arriving city, and the user can select several cities or just a single city to get the route. By placing the mouse over the line, you can also see the start and end points of the route, as well as distance information. Note that since there are many cities with the same name in many states in the United States, the flow map will show multiple endpoints and overlapping routes even if only one endpoint city is selected.
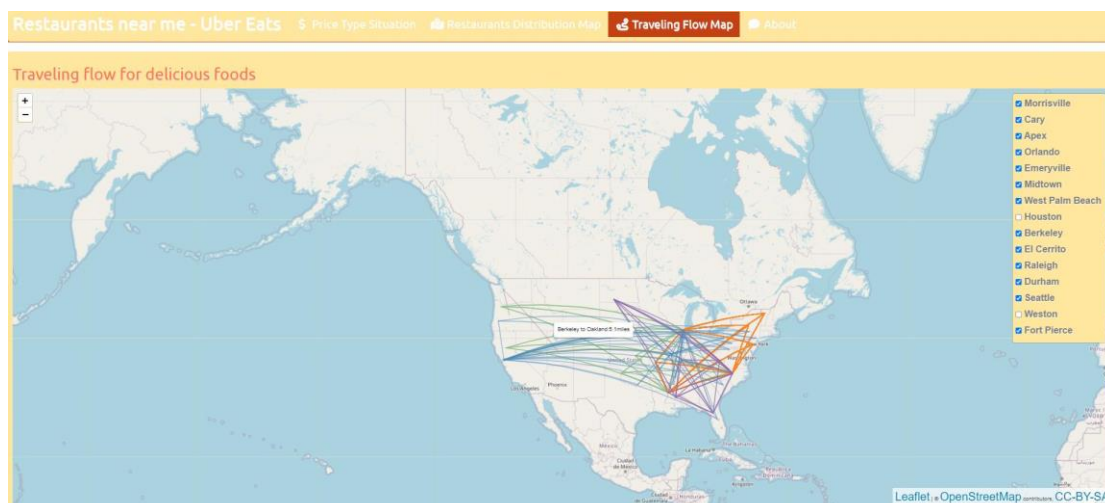

Figure5

Users can find out instructions about the author, the data, and the target audience on the "About" page (Figure6) at the end.

Figure6

Conclusion

Overall, this project visualizes the distribution of popular restaurants in the U.S. and the routes people take for food. Based on 5 design sheets, visualization and interaction were presented and implemented. The shiny app serves the purpose of creating user interaction and manipulation and displaying information about user choices.

However, the project still has limitations. The part that could be improved is the flow map could be tried in the form of animation, which would better show the mobility changes. Besides, although there are many cities with the same name, an attempt could be made to filter out the routes with the longest distance to draw the map to avoid the situation where multiple endpoints lead to a mix of routes.

Bibliography

Munzner, T. (2014). *Visualization analysis and design.* CRC Press.

Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert and Barbara Borges (2022). shiny: Web Application Framework for R. R package version 1.7.2. https://CRAN.R-project.org/package=shiny

Winston Chang (2021). shinythemes: Themes for Shiny. R package version 1.2.0. https://CRAN.R-project.org/package=shinythemes

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." *Journal of Open Source Software,* *4*(43), 1686. doi:10.21105/joss.01686 https://doi.org/10.21105/joss.01686.

Garrett Grolemund, Hadley Wickham (2011). Dates and Times Made Easy with lubridate. Journal of Statistical Software, 40(3), 1-25. URL https://www.jstatsoft.org/v40/i03/.

C. Sievert. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida, 2020.

Joe Cheng, Bhaskar Karambelkar and Yihui Xie (2022). leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 2.1.1. https://CRAN.R-project.org/package=leaflet

Original S code by Richard A. Becker, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka and Alex Deckmyn. (2021). maps: Draw Geographical Maps. R package version 3.4.0. https://CRAN.R-project.org/package=maps
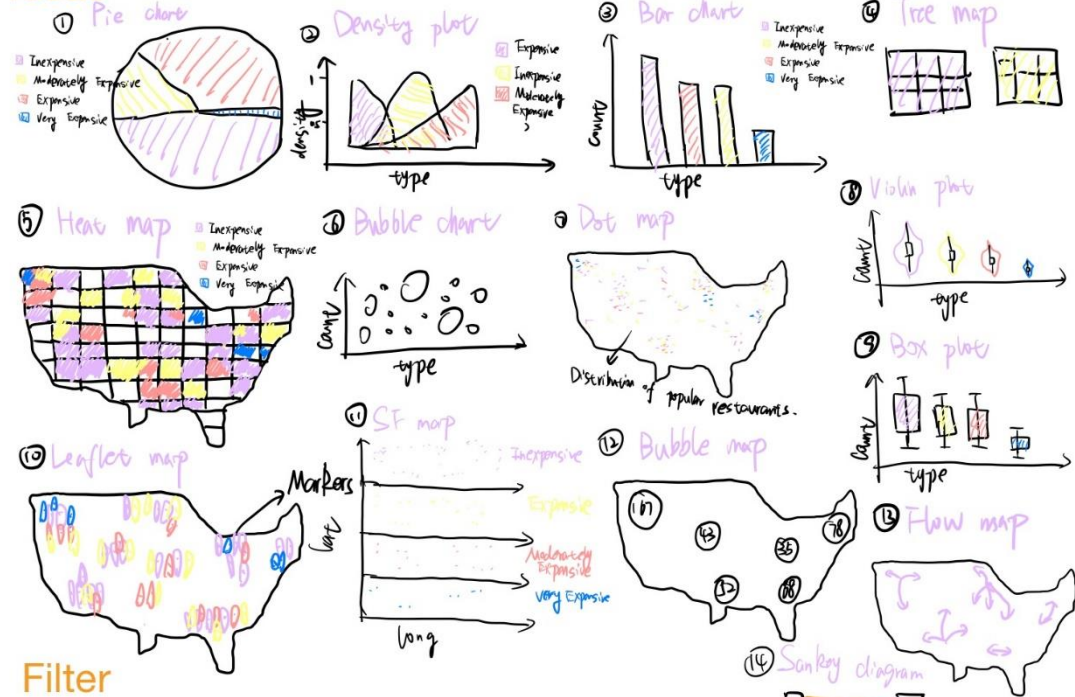
Robert J. Hijmans (2021). geosphere: Spherical Trigonometry. R package version 1.5-14. https://CRAN.R-project.org/package=geosphere

Erich Neuwirth (2022). RColorBrewer: ColorBrewer Palettes. R package version 1.1-3. https://CRAN.R-project.org/package=RColorBrewer

Appendix

Sheet1

Ideas

① Pie chart
Inexpensive
Moderately Expensive
Expensive
Very Expensive

② Density plot
density
type

③ Bar chart
Inexpensive
Moderately Expensive
Expensive
Very Expensive
Count
type

④ Tree map

⑤ Heat map
Inexpensive
Moderately Expensive
Expensive
Very Expensive

⑥ Bubble chart
Count
type

⑦ Dot map
Distribution of popular restaurants.

⑧ Violin plot
Count
type

⑨ Box plot
Count
type

⑩ Leaflet map
Markers
lat

⑪ SF map
Inexpensive
Expensive
Moderately Expensive
Very Expensive
lat
long

⑫ Bubble map

⑬ Flow map

⑭ Sankey diagram

Filter

④ Tree map   ⑥ Bubble chart
⑤ Heat map   ⑨ Box plot

Categories

① ② ③ ⑧   ⑦ ⑩ ⑪ ⑫   ⑬ ⑭

Combine and Refine

① ② ③ ⑧:

Those plots can be used to show the situation of the price type of popular restaurants. Pie chart and density plot is straight-forward to show the proportion or density of each type, meanwhile the violin plot can show the distribution
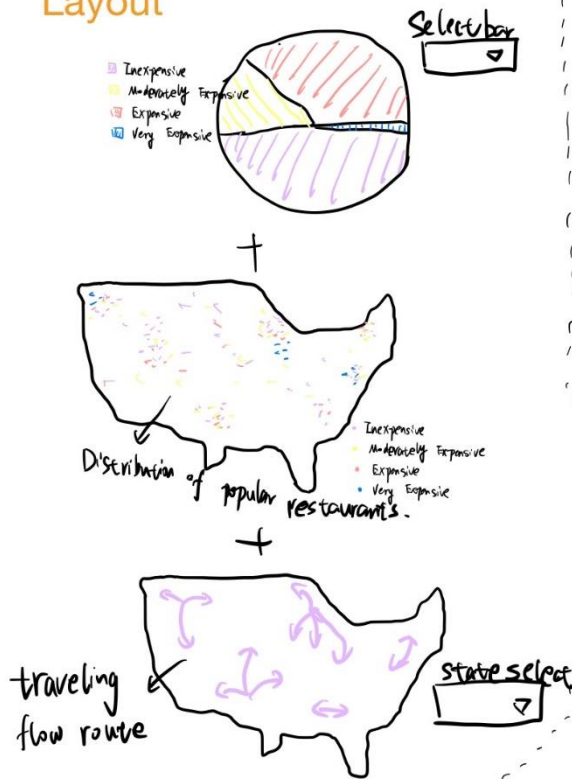
⑦ ⑩ ⑪ ⑫:

As for the distribution of restaurants in US, the ⑦ ⑩ ⑪ ⑫ can be good way to visualize. We can use the leaflet map combine the bubble map to show the location of restaurants directly. ⑬, ⑭ can used to show the customers traveling route for delicious restaurants.

Markers
Inexpensive
Moderately Expensive
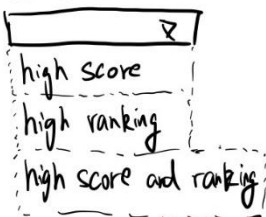Expensive
Very Expensive
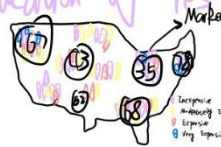
Questions

1. How is the price situation of those popular restaurants?
2. What is the distribution of those popular restaurants?
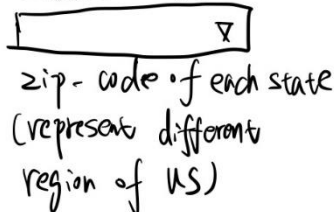3. What is the traveling trend of customers to get to the gourmant restaurants?

Sheet2

Layout

Inexpensive
Moderately Expensive
Expensive
Very Expensive

Selectbar

+

Distribution of popular restaurants.

Inexpensive
Moderately Expensive
Expensive
Very Expensive

+

traveling flow route

State select

Focus

Select bar

high score
high ranking
high score and ranking

State select

zip-code of each state (represent different region of US)

Operation

1. Pie chart: Has a select panel to choose the choice of different judgment of popular restaurants, in order to output the proportion of different price type.

2. Dot map: Using ggplotly to have a interactive plot of distribution of restaurants in US

3. Flow chart: By choosing different zip-code (which represents variance region), interactively showing the traveling flow in different area of US.
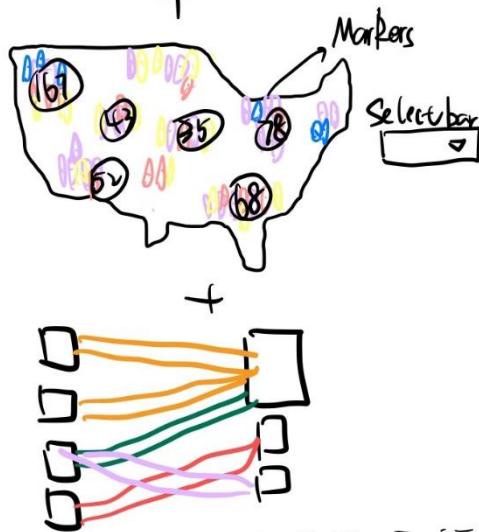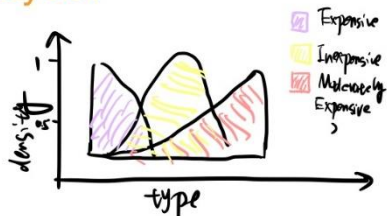
Discuss

+VE

• Flow map can vividly expresses the trip routes that people in different states spend on food.
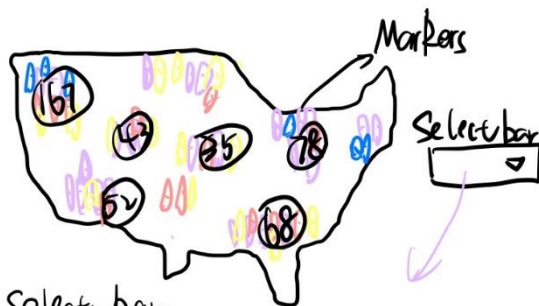
-VE

• Pie chart can't show information in detail, and people don't really understand the size of circles and angles.

• The information brought by dot map is intuitively flat and single, which is not enough to have visual impact.

Sheet3

Layout

Operation

1. Desity plot: Has a select panel to choose the choice of different judgment of popular restaurants, in order to output the density of different price type.

2. Leaflet plot: Using leaflet function to display the count of this areas' restaurants, with marker to represent each place, by selecting various price type to get diverce map.

3. The overall trip diagram is obtained by popularting the origin and destination in the sankey diagram.

Focus

Select bar

Inexpensive
Expensive
Moderately Expensive
Very Expensive

Select bar is used to choose different type of price to show the geographical distribution of hot places.
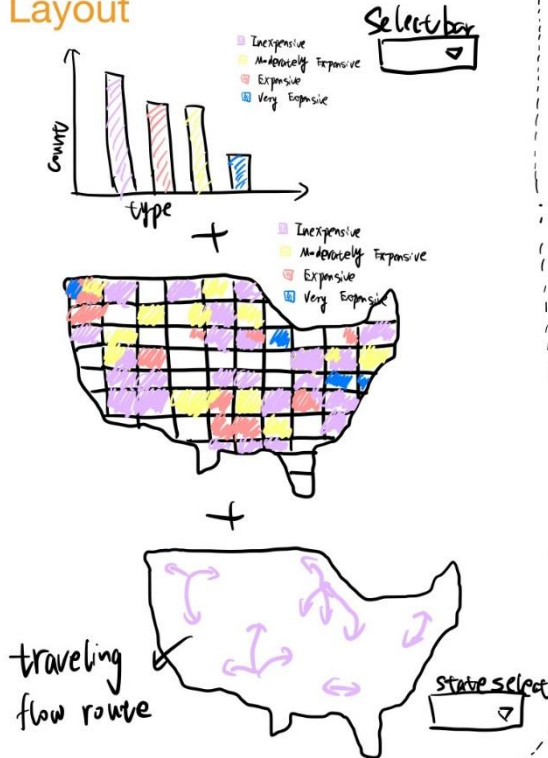
Discuss

+VE

- Leaflet is a very useful visual tool, which can display data information using makers and popups.

−VE

- Sankey map can only simply show where to start and end, not how long the path is and whether the trip crosses state line.

Sheet4

Layout

traveling flow route

Focus

Select bar

- high score
- high ranking
- high score and ranking

State select

zip-code of each state (represent different region of US)

Operation

1. Bar chart: Has a select panel to choose the choice of different judgment of popular restaurants, in order to output the count of different price type.

2. Heat map: By mapping the hot spots of restaurants with diverse price levels, we can get a visualization of the price type of popular restaurants and their geographical locations.

3. Flow chart: By choosing different zip-code (which represents variance region), interactively showing the traveling flow in different area of US.
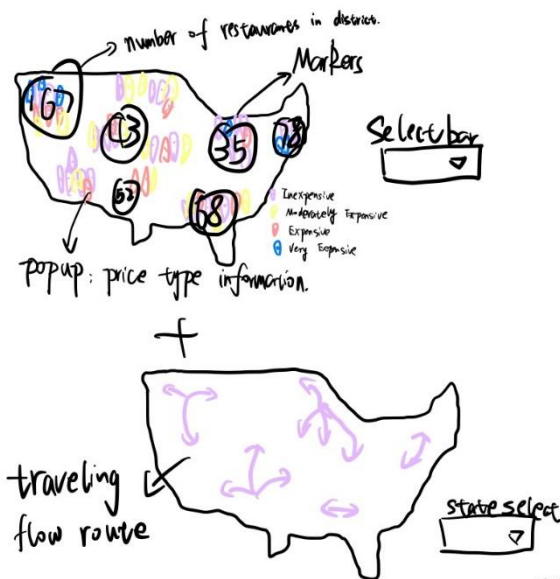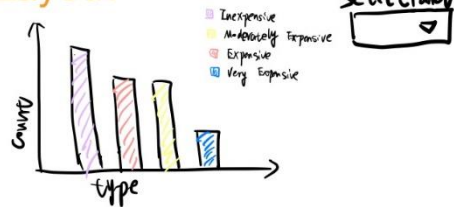
Discuss

+VE

- Flow map can vividly expresses the trip routes that people in different states spend on food
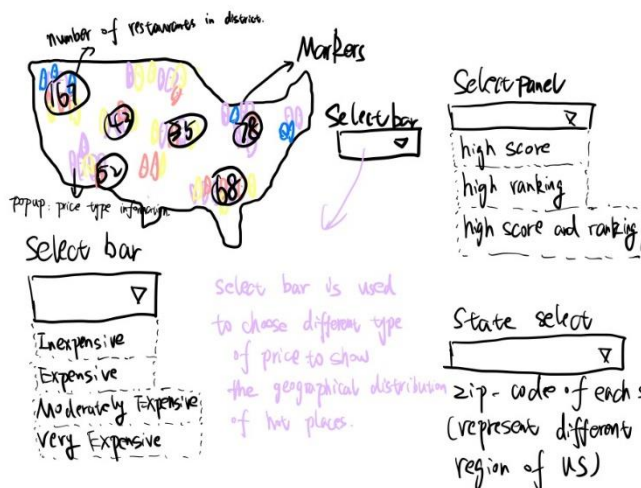- Bar chart is clear and straight-foward.

-VE

- Heat map doesn't accurately show precise message and can't easily identify specific values.

Sheet5

Layout

Inexpensive
Moderately Expensive
Expensive
Very Expensive

Select Panel

Count / type

number of restaurants in district.
Markers
Select bar

16, 13, 35, 28, 52, 68

popup: price type information.

Inexpensive
Moderately Expensive
Expensive
Very Expensive

+

traveling flow route

State select

Focus

number of restaurants in district.
Markers

16, 35, 28, 52, 68

popup: price type information.

Select bar

Inexpensive
Expensive
Moderately Expensive
Very Expensive

select bar is used to choose different type of price to show the geographical distribution of two places.

Select Panel

high score
high ranking
high score and ranking

State select

zip-code of each state (represent different region of US)

Operation

1. Bar chart: Has a select panel to choose the choice of different judgment of popular restaurants, in order to output the count of different price type.

2. Leaflet plot: Using leaflet function to display the count of this areas' restaurants with marker to represent each place, by selecting various price type to get diverse map. With the popups showing the price type information.

3. Flow chart: By choosing different zip-code (which represents variance region), interactively showing the traveling flow in different area of US.

Details
Dataset

Restaurant data has dimension of 40228 rows × 11 columns, has the information of rating, price, and geographic location.

Uber data has 1156 rows and 7 columns, which is based on Uber driver's trips, and contains variables such as trip start/end times, departure and arrival locations.

Dependencies

R Shiny

Estimate

Before 20/10/2022:
Shiny structure construction and graphs completion

Before 27/10/2022:
User Interaction: select

30/10/2022: Overall design and report ready.