**Five Food Stories from Five Cities**

Jing Ouyang Tong Wu and Jing Zhou[[1]](#footnote-1)

University of Southern California, Los Angeles CA 90007, USA

**Abstract.** In this paper, we use Yelp dataset to analyze the restaurant data and present recommendations for users with visual design and use HTML, D3, ArcMap and other tools to illustrate our points. We get inspirations from previous works and improve their designs with the method we learned from class.

**Keywords:** Yelp, restaurant, data visualization

**1 Introduction**

With the increasing number of immigrations and international visitors, the United States and Canada have more diversity on their food genre. People would love to enjoy not only traditional and local food, but also desire to explore different type of food with the help of recommendation applications, such as Yelp. For this final project, we would love to use Yelp’s data for analysis which provides information and suggestions for users. In order to achieve this goal, we would enhance data visualization methods such as graphs, maps and charts to illustrate our points, then we make webpages to present our result for users. We would talk about this project from four different aspects, including literature review, story introduction, graph design and conclusion.

**2 Literature Review**

Before we narrow our scope down to the Yelp dataset, we searched several datasets online for previous works on relative topics. There were several data analysis case studies and data visualization on Yelp dataset. For the majority of studies, they used the whole dataset which included restaurant rating, sightseeing rating, shopping center rating, etc. other than that, they applied ggplot from R to visualize the dataset with bubble plot, box plot, scatter plot and matplot. However, they were only present the distribution of location and the relationship between rating and location. Some other studies were focusing on review analysis over time, such as review trend over time for one specific restaurant or a category of restaurant.

Other than that, we found out that several scholars would like to use pie chart and bar chart to highlight the distribution of the content of our target dataset. Additionally, they made some improvements to those charts by reordered the categories and changed the color combination to enhance the contrast. They also used word cloud to analysis the review in order to get a summary for target restaurant. Unfortunately, the result was neither pleasant nor efficient since there were abundant reviews about services, locations and reservations rather than the food. Therefore, users could not get the information they wanted through the word cloud.

Except creating charts and graphs, previous researchers also used maps to offer a straight forward way to illustrate the feature of this dataset. By doing the literature review, we had a draft about how to apply those data visualization methods on our target dataset, combine information with users’ needs and avoid problems while making our own websites. The main problem of all previous work was the dataset, therefore, we decided to use only the restaurant data from Yelp dataset and focus on five cities which included Madison, Las Vegas, Phoenix, Cleveland, Charlotte, Montreal and Toronto.

**3 Story Introduction**

Our final project is aiming to present five stories about food depends on five different cities. We would love to show the feature of food in different cities and make recommendations for our users based on Yelp dataset. First, we would like to compare these five cities in a general level. Therefore, our users could see the big picture of those five cities food industry. Then, we would be willing to elaborate the story of each city based on their features, such as their traditional food, trending food, top rating food genres and recommendations. For this part of project, we plan to do some research about the most famous and iconic food for each city and visualize it with flat design and catching design. After that, we would provide top five recommendations based on Yelp rating and attach brief introductions of each restaurant.

Nowadays, people have the trend to try various type of food while enjoy the most authentic local cuisine, therefore, users might want to try something uncommon. So, we argue that presenting the high rating uncommon restaurant would match their needs. In order to visualize this part, we need to filter out the minority food genre in each city and then get the most popular and high rated restaurants. Finally, we could show both of the iconic food and uncommon food at the same time.

**4 Graph Design**

**5 Conclusion**

We started from another idea which was also interesting, however, we could not find enough data and previous researches to work with. Therefore, we changed our topic several times to adjust and tried to reach a balance between our interest and the reality. This final project did not only help us to apply all of the data visualization method that we learned from lecture, practice our website design coding skills outside of our required assignments. The most important thing about this final project was pushing us to learn by ourselves and trained us to get used to solve unexpected problems while making websites. Those problems and bugs were hard to forecast while planning, therefore, we had to get use to solve problems and change our original plan while still achieve our goal.

The whole team had many brilliant ideas while brain storming and everyone knew themselves better while they were finishing the final project. We all contributed to this project with our efforts and thought as a team to work together. If we have more time and skills, we would polish our project and make more content from the dataset.

**Reference:**

1. Kaggle Yelp dataset page, <https://nycdatascience.com/blog/student-works/project-1-exploratory-visualizations-of-yelp-academic-dataset-draft/>
2. Kaggle Yelp dataset page, <https://www.kaggle.com/vksbhandary/exploring-yelp-reviews-dataset>
3. Kaggle Yelp dataset page, <https://www.kaggle.com/ibadia/visualization-on-yelp-dataset>
4. Kaggle Yelp dataset page, <https://www.kaggle.com/ambarish/a-very-extensive-data-analysis-of-yelp>

1. [↑](#footnote-ref-1)