INF554 Project Report

Dongyu Li, Yiming Liu and Tianxin Zhao

Abstract— **The project is a Restaurant Market Analysis System wisely implemented via leading visualization and web development techniques. The interactive web application presents user-friendly charts and graphs with real-time and accurate data to help users better analyze restaurant markets.**

I. Introduction

Restaurants are an essential part of modern life. People can find places of business selling food, drinks and desserts of various types on every corner in metropolises or small towns. In the meantime, more and more people decide to start or invest a restaurant business. To help people wisely make decisions, the system provides information about existing restaurants and present the valuable information towards users via intelligently and reasonably visualizations.

The restaurant information presented in the system is derived from Yelp Fusion API, which provides a wealth of data from over 50 million businesses. The Apache Tomcat Server is deployed as back end to fetch data from Yelp Fusion API as well as p reprocess and format data. The React.JS framework is built as user interfaces to visualize information and present graphs and charts implemented via D3.JS and Google Map API.

The final system receives location and area information as input chosen by users, and automatically shows multiple corresponding charts with interactive effects. In addition, the system also generates comprehensive analysis graphs to help comparisons of restaurant markets among various areas. These helps to find out areas with higher dinning consuming level, areas with higher rating on restaurants of specific categories, areas with higher people density and so on for investment decision making.

II. System Design

A. Data Description

The data source of the project is Yelp Fusion API. The endpoint $/businesses/search$ of Search API could return up to 1000 restaurants data based on the provided search criteria. The returned data provides some basic information about the business, including restaurant name, address, geospatial location, category, price, rating, review count, and etc. Figure 1 shows the example restaurant information object after fetching from Yelp API and processing from back end.

B. System Architecture

1) Web Server and External APIs: The restaurant information is fetched from Yelp Fusion API, the data includes restaurant name, category, price, rating, review count, and so on. For the back end, the system uses Apache Tomcat Server to fetch, process and format data.

```
{
    "address": "\"124 Castro St\",\"Mountain View, CA 94041\"",
    "item_id": "chop-and-pub-mountain-view",
    "city": "Mountain View",
    "image_url": "https://s3-media2.fl.yelpcdn.com/bphoto/2EWbBqMcKUdhQwt3_9kjMA/o.jpg",
    "latitude": 37.39482,
    "rating": 4.5,
    "review_count": 11,
    "url": "https://www.yelp.com/biz/chop-and-pub-mountain-view?adjust_creative
        =ocDAWV6tkSxVGkKwE2dErw&utm_campaign=yelp_api_v3&utm_medium
        =api_v3_business_search&utm_source=ocDAWV6tkSxVGkKwE2dErw",
    "zipcode": "94041",
    "phone": "+16504473000",
    "price": 2,
    "name": "Chop & Pub",
    "state": "CA",
    "categories": [
        "Tapas/Small Plates",
        "Cocktail Bars"
    ],
    "longitude": -122.07871
},
```

Fig. 1. Example Restaurant Business Information Object

2) Front end Framework and Visualization Tools: For the front end, the system uses the React.JS framework as user interface to visualize information, and the plots and diagrams are implemented using D3.JS and Google Map API.

C. Visualization Methods

To better visualize formatted data, the project carefully choose visualization techniques with considerations of human vision and recognition characteristics as well as user experiences knowledge. The system takes advantage of geospatial, charts/plots, and graphs these three types of visualization methods in the end.

1) Visualization Considerations:

1) Color Scheme
   The choice of color scheme considers the following aspects:
   - Use distinctive color to represent categorical variables in implementing bubble chart, hierarchical bar chart, and sun burst graph.
   - Use black fonts in white background in interactively showing detailed information to provides the most possible contrast in achromatic channel.
   - Use black fonts in white background in interactively showing detailed information to provides the most possible contrast in achromatic channel.
   - Use consistent aesthetics color across multiple graphics.
   - Normalize the data for mapping to more distinctive colors.

2) Distinctiveness
   Use high degree of contrast of the objects in SVG with the backgrounds, and maximize differences of pre-attentive features to make pop-out effect more marked.

3) Avoid Over-plotting
   In order to avoid over-plotting, the system presents the
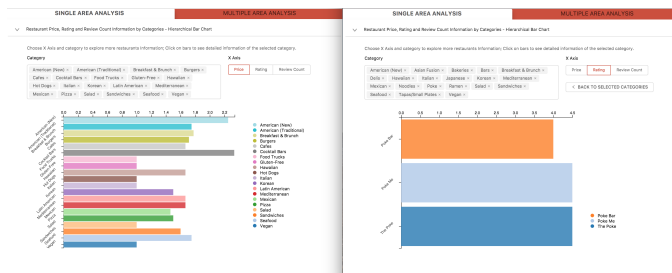
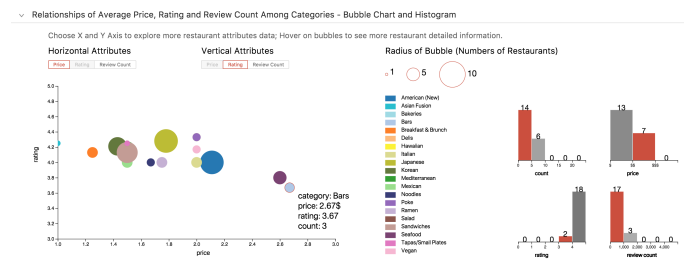Fig. 2. Hierarchical Bar Chart (Left: All categories; Right: Selected category)

Fig. 3. Bubble Chart (Left) and Histogram (Right)

data by calculating average information among each category and use hierarchical bar chart to hide detailed information into second layer.

4) Comprehensible Visualizations

To make the system more acknowledgeable as well as considered the attention bottleneck of human being, the system chooses relatively familiarized visualization techniques and offers graphs on summarized data. The system also provides text hints and legends to help user better understand contents.

5) Smooth Transition

The system implements smooth transitions to avoid change blindness.

2) Visualization Techniques:

1) Hierarchical Bar Chart

The hierarchical bar chart is composed of two layers. The first layer displays the price, rating or review count information on specific kind of restaurant category selected by the user. The second layer shows up after clicking on the bar of one category, and the detailed information within the category would display. Figure 2 (left) shows the overall situation in the area, and the second Figure 2 (right) shows the price information of each restaurant belongs to the category.

2) Bubble Chart and Histogram

The bubble chart displays three dimensional information. The radius represents the number of restaurants of each category. The x axis and y axis could be changing among price, rating and review count to show relationships between categories on every category in the selected area. See Figure 3.

The histogram displays the distribution of restaurants numbers, average price, average rating, average review counts of the whole selected area as shown in Figure 3.

3) Thematics Map - Dot Map and Choropleth Map

The Thematics Map used Google Maps as a geodetic system for mapping the data to specific locations by coordinates, and used Dot Map with dots in yellow to red to represent data of restaurants, and Choropleth Map with polygons in yellow to red for average value of each area. The Dot Map is for both single area analysis and multiple areas' comparison. And for each area, a polygon will be generated on the Choropleth Map.
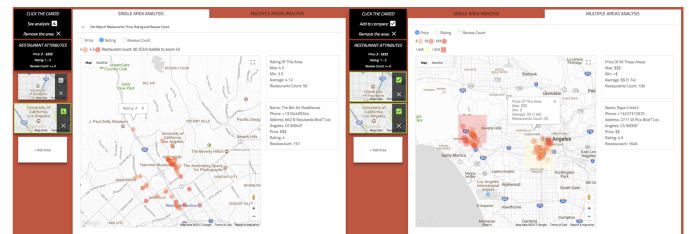

Fig. 4. Dot Map and Choropleth Map (Left: Dot Map fot One Area; Right: Dot Map and Choropleth Map fot Multiple Areas)

Both the dots and polygons has the color proportional to the value of prices, ratings or review counts (See Figure 4) - bigger values with deeper colors more close to red.

Combination of the Dot Map and Choropleth Map is an innovation. Most of the Dot Maps include tons of dots covering most of the areas where a Choropleth cannot be applied together. Here the restaurants are not that many so that a combination of Dot Map and Choropleth Map can clearly shows the value differences for both single restaurants and average level of a whole area.

Google Maps' Data Layer and Event Listener API were used for this part. As many restaurants can be located together in a small area, the values was mapping to different colors rather than different sizes. And the map will be auto zoomed out when a dot is clicked - to avoid overlapping of the dots. For the comparison of different areas, squares are used to represent the selected area, with the boundaries of the squares are exactly the area that be selected. And the colors are used for representing values of properties.

Although the dots' and squares' colors are for the same attribute, separated scale legends are used because all the scale legends for Thematics Map instances are generated dynamically after normalizing and redistributed the values for current area, for a better visualization with more distinctive colors.

4) Sunburst

The zoom able Sunburst is similar to the tree map, except it uses a radial layout. The root node of the tree is at the center which represents the collection of all selected categories, while leaves are on the circumference. The inner layer displays each cate-
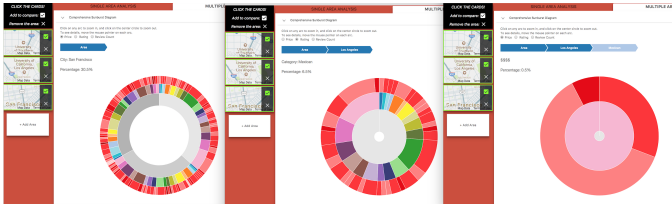
Fig. 5. Sunburst (1.Restaurant categories and price information for 3 areas ; 2 :Proportion of categories in San Fransisco;3 :Price proportion of one category.)
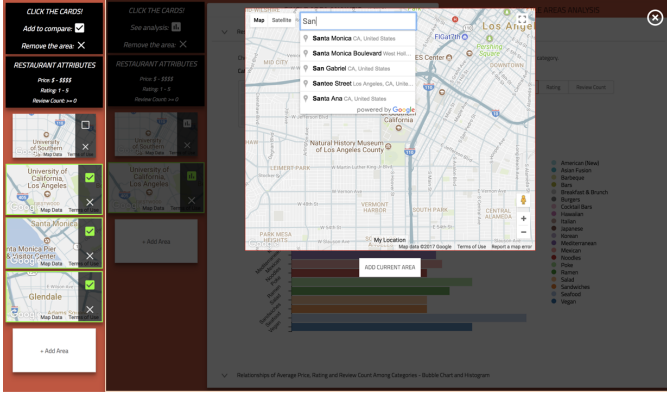

Fig. 6. Areas Management (Left) and New Area Adding (Right)

gory's name and proportion. The outer layer shows the restaurant parameter's overall proportion. In addition, the proportion of the measures is clear in each category. The parameters, price, rating, and review count, can be selected by radio buttons. Click on any arc to zoom in, and click on the center circle to zoom out.

3) Interactions:

1) Map Area Selection

   Areas to analyze are selected through Google Maps (See Figure 6). There is a "My Location" button on the bottom to get the device's current location and re-center the map. Or with the search bar on the left top corner, the map can quickly re-centered to a specified place. Zooming or dragging on the map provides high flexibility on area selection.

   The left navigation bar holds all the selected areas and providing area management(See Figure 6). The selected area of are in larger size. Clicking on the map card can toggle the selection status of the corresponding area. In the comparison tab, multiple areas can be selected, while in the single area tab selection of one area will force canceling all other selections. The "x" button on the lower right corner of each card is for totally deleting the area.

2) Plot Interactions

   - Hierarchical Bar Chart
     The hierarchical bar chart generated automatically according to users' choice of restaurant categories(like American, Asian, Burgers, ...) and the second layer shows up when clicking on the bars.

   - Bubble Chart and Histogram
     The bubble chart generated according to users' choice of comparison types(price, rating or review count) and the overall information of the whole category will show up when hovering on the bubble. Also, by hovering on the bubble, the bins of the category belongs to in the histogram will interactively turn to red.

   - Dot Map and Dot Map with Choropleth Map
     For a single area, the conclusion of this area is in the first information window on the right hand, displaying the maximum, minimum and average value of the selected attribute. The second information window is for the details of the corresponding restaurant. It will be activated when any dot is hovered. Also the value of current selected attribute will be displayed above the dot with the mouse over it. The clicking action on a dot will force the map re-centering to the dot and zooming in(See Figure 4).

     For comparison of multiple areas, the conclusion of all the restaurants of these areas is in the first information window on the right hand, displaying the maximum, minimum and average value of the selected attribute. Interactions on dots are the same. Hovering on a square will trigger a float window with the details of the corresponding area.(See Figure 4).

   - Sunburst
     The zoom-able sunburst generated automatically according to users' comparison types(price, rating or review count) and the selected categories. The proportion and amount will show up when hovering on the arcs. When users click on any arc, the chart will zoom in, or click on the center circle to zoom out. The zoom function provides the possibility of viewing in details.((See Figure 5).

4) Responsiveness: The main user interface is responsive for devices in most size, including mobile devices and PCs. CSS styling helps for this part. For each chart, as the data visualization emphasize the intuition and clear relations of data sets, the complete and clear charts are only available for bigger screens like PCs. Some charts can be viewed by scrolling on small screens while some may mess up. For the maps, Google Maps helps to handle the responsiveness problems.

III. EVALUATION

A. Evaluate by Visualization Wheel

In refers to the Alberto Cairo's visualization wheel, the visualization could be evaluated in 6 dimensions. The website provides bar charts, bubble charts, histogram, sunburst and the dot map with functions like searching any area the user concerns and dynamic comparisons between real-time data, so it is considered as a more abstractive, functional,
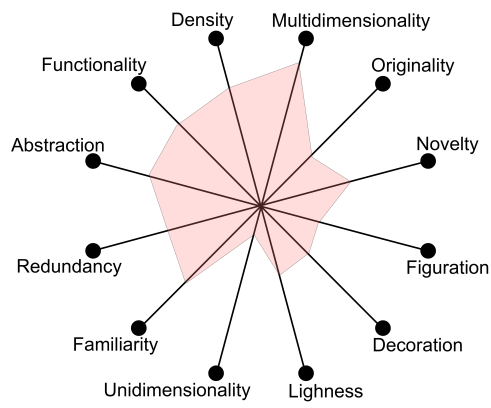
Fig. 7.   Visualization wheel

informative tool other than figurative, decorative and redundant pages. The analytic charts and maps display in both multidimensional and uni-dimensional ways. In addition, the website owns more familiarity than originality, because the plots and diagrams are combinations of simple components. ((See Figure 7).

B. Improvements

After collecting evaluations by Professor Luciano and other classmates, improvements for better understanding their functions should be added. One of the major logistic problems to deal with is to reorganize the navigation buttons which control the single area analysis and the comparison between areas. Also, for the hierarchical bar chart, some default categories need be chosen at first and the legend of bubble size is required in the bubble chart. For the analytic results in map views, the dot map should be applied instead of the proportional symbol map to present the location precisely while different colors of the dots give the attributes of restaurants. In addition, legends and explanation words need to be added to remove the confusions. Due to the time limitation, some of the problems are still in process.

IV. RELATED WORK

Preparation work includes study on React.JS, D3.JS, and reading documents of Google Maps API and Yelp fusion API. As the Yelp fusion API does not allow the direct access from client side, a Java server is used for back-end support.

According to researches, many charts analytics are based on states or countries restaurants. However, there is lack of integrated tools combining map searching and analytic visualizations. Besides, the implementation of Google Map API and Yelp API provides the possibility of comparisons all around the world. Comprehensive views of multiple restaurant and analysis of an area not only a single restaurant are provided. Current applications like Google Maps, Yelp only provides pure map version or the info of single restaurant.

V. CONCLUSION

All visualizations are created following the foundational patterns. In the ideation phase, descriptive analysis is used

to examine the data profile, found out suitable variables for restaurant business decision making and showing Yelp restaurants insights. The interactive web application with real-time data and reasonable diagrams and plots has been implemented via web development techniques. Therefore, this visualization is meant to serve as a starting point for restaurant investors or market analysts use as a guide and build on.

REFERENCES

[1] Murray, S. (2017). Interactive data visualization for the Web: an introduction to designing with D3. Beijing: OReilly.
[2] Cairo, Alberto. The Functional Art: an Introduction to Information Graphics and Visualization. New Riders, 2013.
[3] d3. d3/d3 [Internet]. [cited 2017Nov30]. Available from: https://github.com/d3/d3/blob/master/API.md
[4] Sidelnikov G. React.js Tutorials [Internet]. React JS Tutorials. [cited 2017Nov30]. Available from: http://www.reactjstutorial.net/