Eurographics Conference on Visualization (EuroVis) 2018
J. Heer, H. Leitte, and T. Ropinski
(Guest Editors)

Volume 37 (2018), Number 3

Towards Easy Comparison of Local Businesses Using Online Reviews

Yong Wang[1], Hammad Haleem[1], Conglei Shi[2], Yanhong Wu[3], Xun Zhao[1], Siwei Fu[1] and Huamin Qu[1]

[1]the Hong Kong University of Science and Technology, China
[2]Airbnb Inc., USA
[3]Visa Research, USA

Abstract

With the rapid development of e-commerce, there is an increasing number of online review websites, such as Yelp, to help customers make better purchase decisions. Viewing online reviews, including the rating score and text comments by other customers, and conducting a comparison between different businesses are the key to making an optimal decision. However, due to the massive amount of online reviews, the potential difference of user rating standards, and the significant variance of review time, length, details and quality, it is difficult for customers to achieve a quick and comprehensive comparison. In this paper, we present E-Comp, a carefully-designed visual analytics system based on online reviews, to help customers compare local businesses at different levels of details. More specifically, intuitive glyphs overlaid on maps are designed for quick candidate selection. Grouped Sankey diagram visualizing the rating difference by common customers is chosen for more reliable comparison of two businesses. Augmented word cloud showing adjective-noun word pairs, combined with a temporal view, is proposed to facilitate in-depth comparison of businesses in terms of different time periods, rating scores and features. The effectiveness and usability of E-Comp are demonstrated through a case study and in-depth user interviews.

CCS Concepts

•***Human-centered computing*** $\rightarrow$ *Information visualization; Visualization design and evaluation methods;*

1. Introduction

Numerous customer review websites or e-commerce platforms (e.g., Yelp, TripAdvisor and Airbnb) have been launched recently. All of them allow users to post reviews for restaurants, shops, hotels, etc., which are called *local businesses* in this paper. These reviews, usually consisting of both a numerical rating and a text description, have a strong influence on the purchase decision of customers [FM14, YLG09]. The online reviews have become a major resource to help customers make a purchase decision in many application scenarios, for example, finding a good restaurant for an important celebration [Nie], choosing a professional barbershop for a stylish haircut [Mor11] or booking a suitable hotel for a distant trip [GY08]. In most cases, there would be many candidates for a specific type of service or products, satisfying the customer requirements. Customers need to read online reviews to compare them and make an optimal choice.

However, it is not an easy task to use existing online reviews to make a quick comparison. Most review websites calculate an average rating for each business entity from all the customers' ratings. Such an average rating provides an indication about customers' overall assessment to some extent, but the average rating score does not necessarily reveal the true business quality due to its bimodal distribution [HPZ06]. Also, it lacks the temporal details of customer reviews [GS12, WH08] and other important information. For example, how do the customer evaluations for a selected business entity evolve with time (e.g., good evaluations in the past but bad ratings in the recent, or vice versa)? What is the main difference between two businesses with exactly the same rating? Customers often check the review text to get more details of previous customers' opinion. Unfortunately, there are usually many reviews for single local business and the large volume of review text leads to information overload [BAC12]. Moreover, customers have to compare various business features revealed in the review texts by memory. These factors make it difficult to conduct an efficient review summary and reliable comparison of local businesses.

Prior research work mainly uses review text to summarize customer opinions [YNTT11b, WZG*14, WWL*10]. They either use traditional word clouds showing single words with most frequency or abstract the review sentiments into an augmented scatterplot to show the review content. However, their major focus is to provide a review summary for individual business and lack detailed context (as will be discussed in Section 6.3). Thus, they are not suitable for detailed comparison between different businesses. Opinion Observer [LHC05] compares review sentiment of two businesses using the basic bar charts, but it is still unable to provide users with deep insights about the difference between them.

In this paper, we present *E-Comp*, a visual analytics system, to help customers quickly compare local businesses, where both the numerical rating and review text are considered. Based on our interview with end users, we divide the whole comparison of local businesses into two stages: *preliminary comparison* and *detailed comparison*. Preliminary comparison aims at quickly selecting the interested businesses from all the choices and detailed comparison supports revealing insightful differences between them. An intuitive glyph is designed to encode the most important features for preliminary comparison, including customer number, average rating, each rating percentage, business similarity and price. For detailed comparison, we propose a novel augmented word cloud showing grouped adjective-noun word pairs, to support an easy comparison of different aspects of local businesses. Considering that different customers may have significantly different criteria when rating the businesses, we propose using the rating difference by the *common customers* who rated both businesses to gain more reliable comparison through a grouped Sankey diagram. Interactive comparison of the temporal distribution of reviews is also enabled, helping users quickly understand the temporal evolution of the customer reviews. The major contributions of this paper can be summarized as follows:

- An interactive visualization system to help customers conduct preliminary and detailed comparisons of local businesses.
- An intuitive glyph design to support a fast selection of local business candidates and a novel augmented word cloud showing adjective-noun word pairs to enable detailed comparison.
- A case study and in-depth user interviews to demonstrate the effectiveness and usability of the proposed method.

2. Related Work

The related work of this paper can be categorized into three groups: visual comparison, review visualization and opinion extraction.

2.1. Visual Comparison

Visual comparison is a typical and fundamental visualization task [Gle18, KH13], which aims at understanding the similarity or difference between data. According to Gleicher et al. [GAW*11], visual comparison can generally be classified into three complementary categories: 1) spatial and temporal juxtaposition (e.g., side-by-side comparison), 2) superposition (i.e., overlay), and 3) explicit encoding of the difference (i.e., visually displaying the similarities and differences). These three basic methods and combinations of them have been widely used in many applications [SWL*14, WSA*16, SMDS14]. Juxtaposition places the objects side-by-side to support comparison between them [GLG*13, NSH*18, KPBG13, ZWC*18]. Superposition shows the objects to be compared in the same space to investigate the differences between data, which is especially useful when the spatial location is a key component of the comparison [WDSC07]. When the difference between objects can be explicitly computed, explicit encoding can be a good choice for visual comparison, which has been applied in many applications including TACO [NSH*18], GraphDiaries [BPF14] and AmbiguityVis [WSA*16]. Besides, interaction is often used to enhance visual comparisons [Gle18, HVW08, SMDS14].

In this paper, we use side-by-side comparison and enable interactions to compare different aspects of local businesses.

2.2. User Review Visualization

Visual analysis of user reviews can facilitate an overall understanding about the products or services and has been extensively researched in recent years. For example, visual analytics has been applied to analyzing user reviews of devices [LHC05, OHR*09], hotels [WWL*10], films [BML17]. They mainly extracted different features of the products or services from the reviews and further visualize those features to provide an overall understanding of the reviews. Other researchers often focused on visualizing review text, where word clouds combined with sentiment analysis [KPK17] are usually used. Yatani et al. [YNTT11b] visualized user reviews by showing a word cloud consisting of adjective-noun word pairs. Similarly, Wang et al. [WZG*14] arranged semantically similar words close to each other in the word cloud. *Compare Cloud* [DES*15] visualized words of text corpora to compare media frames. *Prefix tag clouds* [BLPW13] used a prefix tree to group different word forms and visualized them as a word cloud. The recently proposed SentenTree [HWS17] linked words based on their co-occurrence to provide more semantic meanings. These methods can give users a basic overview of the review content, but are difficult to support detailed comparison and provide deep insights about the reviews.

Inspired by [YNTT11b], we propose an augmented word cloud design to show grouped adjective-noun word pairs, enabling better comparison between local businesses through user reviews.

2.3. Opinion Extraction

Opinion extraction, also called opinion mining or sentiment analysis, aims at summarizing the opinions expressed in the text, which has attracted lots of attention in the text mining field. Several comprehensive surveys about opinion extraction and summary can be found in [KGSZ11, BHML16]. More specifically, Hu and Liu [HL04] summarized opinions from review text by extracting the product features and further identifying opinions. Yu et al. [YHSZ16] proposed a phrase-based approach for summarizing reviews. Yatani et al. [YNTT11a] obtained the adjective-noun word pairs to show the opinions in review text. Other researchers tried to extract opinions through sentiment analysis [PL08]. In addition, many excellent natural language processing (NLP) libraries are published (e.g., NLTK [nat], TextBlob [Lor17]), allowing general users to process text data and extract opinions.

In this paper, we use the NLP libraries to analyze review text and extract adjective-noun word pairs. But different from [YNTT11a], we further classify the word pairs into several semantically meaningful categories and group the word pairs with common noun word to facilitate the comparison between local businesses.

3. Requirement Analysis and Data Model

In this section, we summarize the design requirements and basic tasks. The dataset used in this paper is also introduced.

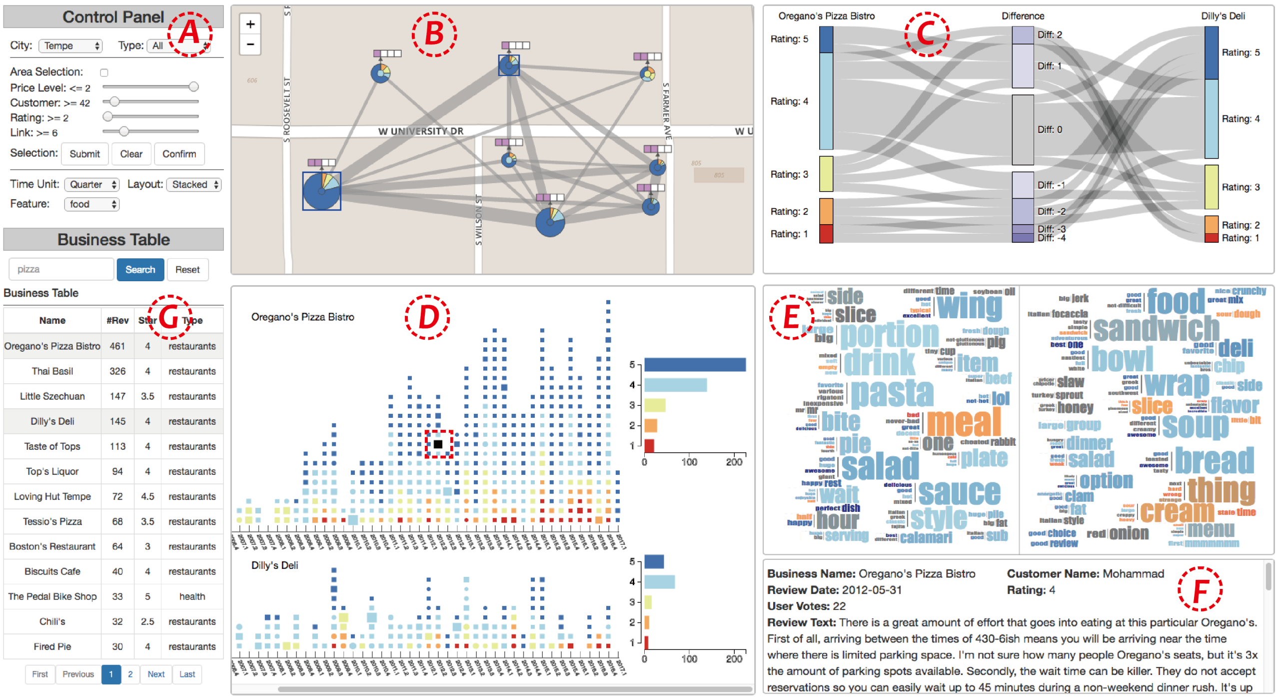

Figure 1: *The user interface of E-Comp. (A) Control panel allows users to filter the data and change other views. (B) Map view shows glyphs encoding the basic attributes of local businesses, where the selected businesses are marked with a blue rectangle. (C) Common customer comparison view shows the rating difference of common customers. (D) Temporal view visualizes the temporal trend of reviews and encodes the helpfulness of individual reviews through the rectangle or circle size. (E) Augmented word cloud view compares the adjective-noun word pairs with high frequency for one feature dimension in the review text. (F) Detailed review text view shows the details of the selected review (the dotted red rectangle) in Temporal View. (G) Business table lists the local businesses in the selected region.*

3.1. Design Requirements and Tasks

To better understand the user requirements when reading online reviews to compare local businesses, we conducted user interviews with four participants (1 female, age: 22 to 29). They are postgraduate students or research staff in a university and have an experience of at least four years in using online reviews for making purchase decisions. We asked them about the major procedures and important review features they will use to compare businesses. We also encouraged them to report their expected requirements for a review-based business comparison tool. According to their feedback, when using online reviews for purchase decision-making, their exploration of using online reviews for purchase decision-making usually includes two stages: ***preliminary comparison*** to select candidates and ***detailed comparison*** to make a final decision. In the preliminary comparison, they mainly use the basic information of businesses, e.g., average rating score, percentage of each level of ratings, location and price, to select a few candidates. For the detailed comparison, all the participants mainly rely on the detailed ratings and review texts. However, the overwhelming volume of user reviews hinders them from gaining a quick understanding of the overall reviews, which is consistent with prior studies [BAC12]. The majority of the participants instead choose to read only several latest reviews, though it is probably biased and unreliable. Thus, they would highly appreciate it if a tool supporting reliable and fast comparison of local businesses is available. Besides the user feedback, we also investigated prior studies on online reviews. Based on these, we compiled a list of design requirements:

R1 Quick overview for filtering potential candidates. As suggested by participants, the first step of finding an optimal local business is to select several potential candidates satisfying customers' basic requirements by using features like price, location and other customers' review rating. For the review rating, they mainly use the average rating score, the ratio of good and bad ratings, and the total review number (indicating the popularity of a local business) for quick filtering. Prior studies [LHC05, DFAG13] also confirm the importance of quick summary of online reviews. However, in most existing online review platforms, users need to browse a long list of businesses and sometimes switch between the basic information pages of the interested businesses, which is not convenient for quick selection of candidates.

R2 Reliable comparison between businesses. Due to the large volume of online reviews, the participants said that they usually are not able to read enough reviews quickly before making a purchase, which makes their decisions probably biased. Also, different customers may have quite different rating criteria and preference, making them not as reliable as expected. Therefore, a comprehensive and reliable comparison of online reviews facilitates more accurate decision making [MS10].

R3 Temporal analysis of user reviews. Prior work [GS12, WH08] has studied the temporal evolution of online reviews. The temporal analysis supports a deep understanding of the dynamic changes such as the overall trend of user rating [WH08] and the dynamic influence of online reviews on user purchase behaviors [Cad15]. Our interview participants also showed interest in similar questions: Is the business receiving more and more good reviews or just the opposite? What is the difference between recent reviews and older reviews? Answering these questions requires a temporal analysis of reviews.

R4 Insightful details of important features. Our participants reported that they are usually interested in the important features of local businesses. For example, for a restaurant, they may care about the price, food taste, service and environment. Extracting the important features from online reviews is also emphasized in prior studies [DFAG13, HL04]. Manually reading all the reviews to gain insights is time-consuming. Automatically providing insightful details of important features would greatly help the purchase decision making.

R5 Detailed review exploration on demand. The participants said that they usually read some reviews carefully to gain more details, especially those reviews with high helpfulness. It helps confirm the exploration findings by them.

R6 Intuitive visual designs. The target users of the proposed system are general customers with no background knowledge of visualization. Therefore, intuitive visual designs would be more desirable than complex designs that probably confuse users.

3.2. Data Abstraction

We mainly use the Yelp dataset [yel]. It contains 4.1 million reviews for 144 thousand local businesses in 11 cities across 4 countries, where we extracted the review data in three cities: Toronto, Las Vegas and Tempe. Each review contains both a numeric rating ranging from 1 to 5 and review text. Other information about the review, including the actual time, the corresponding information of businesses and customers and the helpfulness votes, are also provided. The detailed business information includes its name, location, average rating, city, etc. The proposed system is mainly tested by comparing Yelp restaurants using online reviews, but it is not limited to restaurants and can also applied to other applications such as different shops.

4. System Overview

We name the proposed visual analytical system *E-Comp*, since its design goal is to support ***E***asy ***Comp***arison of businesses on ***E***-commerce platforms. We pre-processed the source review data offline. The most important features for comparing local business, including statistical features, common customers between local businesses, temporal features and frequent word pairs, are first extracted and stored in the database. Figure 1 shows the major user interface of *E-Comp*. It has four carefully-designed visualization views: *Map View*, *Common-customer Comparison view*, *Temporal View* and *Augmented Word Cloud View*. *Map View* shows intuitive glyphs encoding statistical features of local business to help users conduct a *preliminary comparison*, i.e., finding the potential candidates. The remaining other three views supports *detailed comparison* between local businesses: *Common Customer Comparison View* provides a more reliable comparison between local businesses by using the ratings from common customers, *Temporal View* visualizes and compares the temporal distributions of customer ratings, and *augmented word cloud view* is proposed to help customers quickly understand the major characteristics of local businesses.

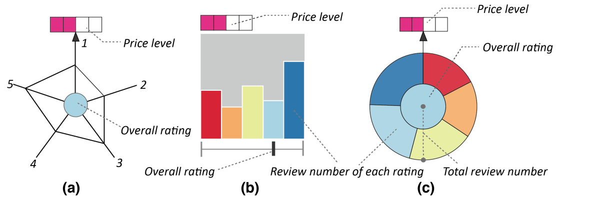

Figure 2: *Visual design for the glyph overlaid on the map. (a) A radar-diagram based glyph design, (b) a bar-chart based glyph design, (c) a pie-chart based glyph design.*

5. Preliminary Comparison

This section introduces the visual designs for preliminary comparison.

5.1. Glyph-based Visualizations

According to **R1**, one fundamental task of using online reviews to make a purchase decision is to quickly select the potential local businesses. Customers usually do filtering based on the factors like average ratings, percentages of good/bad ratings, popularity indicated by the number of existing reviews and the price level, which are also used for preliminary comparison in this paper. Considering the advantages of glyphs in conveying the overall information of multiple attributes [War08], we chose to design intuitive glyphs to show these essential features (**R6**). We proposed a pie-chart based glyph, as shown in Figure 2c. Considering that area is an effective visual channel in pie chart [SK16, BEW16], we use the outer circle area to indicate the total number of customer reviews. Each sector of the pie chart encodes the review number of each rating level. The inner circle has fixed area and its color represents the average rating of all the customers. The upper bar shows the price level of a business, which can be gained from the Yelp review dataset.

Alternative design: Before finalizing the glyph design, we also considered the radar-diagram based glyph (Figure 2a) and bar chart based glyph (Figure 2b). Bar charts are commonly seen and can facilitate the direct estimation of the absolute number of reviews at each rating level. However, the bar chart based glyph is not able to clearly visualize the percentage of good/bad ratings and the total number of reviews, which are the major features that customers use for quick filtering of business candidates (as shown in Section 3.1). The radar-diagram based glyph suffers from the same problems with the bar chart based glyph. Moreover, it has been shown that the shape characteristics of radar diagram affect the accurate interpretation of the underlying data [KHW09]. On the contrary, a prior study [SL91] has demonstrated that pie chart has an advantage over bar chart in displaying percentages and it has been successfully applied to showing the percentages in real applications [CLLT15]. The total number of the reviews is also directly revealed by the circle area of a pie chart. Taking all these factors into account, we finally chose the pie-chart based glyph (Figure 2c).

Color encoding: Initially, a sequential color scheme (bad rating:

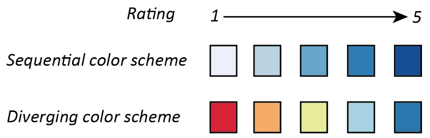

Figure 3: *Color scheme for the rating. The diverging color scheme is finally chosen to emphasize both the higher and lower ratings.*

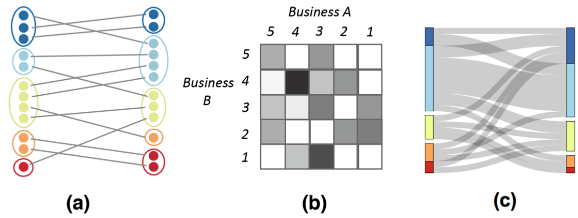

Figure 4: *Alternative designs for common customer comparison: (a) bipartite graph, (b) matrix, (c) traditional Sankey diagram.*

light blue, good rating: dark blue) is chosen to show the number of reviews in each rating level. However, after checking with the target users in the prototype stage, we found they usually regard a three-star rating as a neutral review and are more interested in the percentage of positive and negative ratings than the neutral rating, which is also confirmed in [MS10, PS80]. However, a sequential color scheme tends to emphasize only the positive ratings. Therefore, a diverging color scheme is finally chosen (Figure 3).

Location information: The location information is also essential for preliminary comparison of local businesses, which is usually shown by drawing a label icon on the map in existing online shopping platforms. Instead, we proposed overlaying meaningful glyphs on the map, making the preliminary comparison more convenient.

5.2. Encoding of Common Customers

The number of common customers between local businesses has been widely used in the recommendation systems for E-commerce applications [LSY03], as it indicates users' common interest in several business items. Therefore, to help users easily find local businesses of their interest (**R1**), we add links between local businesses with the link width encoding the number of common customers. Interactive filtering of the links is also supported to reduce the possible visual clutter when too many links are shown.

6. Detailed Side-by-side Comparison

The detailed comparison aims at further comparing the initially filtered candidates to select the optimal one as the final choice. To provide users with a reliable and comprehensive comparison between local businesses using online reviews, we designed three major views to achieve the design requirements discussed in Section 3.1, including common customer comparison view, temporal view and augmented word cloud view.

Due to the limited screen size, we need to make a trade-off between the details to show and the number of businesses to compare. Considering the preliminary comparison has already enabled the quick filtering and selection of local businesses and it is important to show reliable and comprehensive details in the detailed comparison stage, we finally chose the side-by-side comparison to compare two candidates, instead of comparing more businesses simultaneously.

6.1. Common Customer Comparison View

The design goal of common customer comparison view is to provide users with a more reliable and deeper comparison between two local businesses (**R2**). Using all the customer reviews can reveal the overall differences between local businesses to some extent. However, one possible risk of using all the customer reviews is the inconsistency across user reviews [MS10, BAC12], as different customers may have quite different personal preferences and rating standards. Therefore, we choose the review ratings by common customers to achieve a more insightful comparison, since the rating standards of individual customers are relatively stable and the rating difference by the same customer is able to reveal more details of the two local businesses.

Considering the expressiveness of Sankey diagram in conveying the data relationship [KBH06], we propose a Sankey-diagram based design to visualize the difference of ratings by common customers (Figure 1C). The left and right rectangles encode the five-level ratings of two local businesses. The middle rectangles explicitly summarize the difference of ratings by each common reviewer, i.e., the left rating subtracted by the right one. To further facilitate comparison of reviews by the same customers, we group the ratings with a similar overall sentiment, where a review rating of three in the five-level ratings usually indicates a neutral evaluation, while the rating above or below three can be regarded as positive or negative feedback, respectively. We use the same diverging color scheme discussed in Section 5.1 for showing the review ratings and a sequential color scheme for showing the rating difference.

Alternative designs: Several alternative designs are also considered in the prototype stage. The bipartite graph (Figure 4a) is the initial option for comparing two businesses, where the numerical ratings by common customers are directly shown. However, it has serious scalability issues with the increase of reviews. Also, it is difficult to quickly understand the overall rating difference by common customers. Matrix (Figure 4b) is another possible visual design here, with the color opacity of each cell encoding the number of corresponding common reviewers. But it is difficult to gain an accurate understanding on the overall rating difference by common customers. The traditional Sankey diagram (Figure 4c) has better scalability than the bipartite graph, but viewers still need to manually compare each level of ratings side-by-side, which is not efficient especially when both businesses have similar rating distributions. Figure 1C shows exactly the same data with Figure 4c. But with the explicit encoding of rating difference in the middle column of Figure 1C, users can more easily know which is relatively better when compared with using Figure 4c.

6.2. Temporal View

Temporal information of reviews is also critical for comparing local businesses (**R3**), which, however, is not well conveyed by existing review visualization techniques. In this paper, we propose a temporal view to explicitly visualize the review variation along with time, as shown in Figure 1d. Each review is encoded as a rectangle or circle, where circles represent the reviews by the common users of the selected businesses under comparison. To visualize the overall trend of reviews and evolution of reviews in each rating level, both ***stacked layout*** and ***layered layout*** are provided. Users can specify the layout mode through interaction. For the time unit, users can also interactively specify it as either one quarter or one month, depending on their needs.

It is of great importance to inform users of the helpfulness of each review, as many businesses have an overwhelming number of reviews, making it time-consuming to read all the reviews. When users are able to know which reviews are more helpful, they can focus on those important reviews to accelerate the review exploration. According to the comprehensive study by Mudambi and Schuff [MS10], there are four major factors that affect the perceived helpfulness of reviews: rating extremity, review depth (i.e., the word count of a review), helpfulness votes by other customers and product type. Product type mainly moderates the effect of review extremity on the review helpfulness. They further claimed that reviews with extreme ratings are less helpful than reviews with middle ratings for the so-called *experience goods*, which are defined as products that are difficult to obtain information on product quality prior to purchase [MS10]. The local businesses in Yelp dataset belong to this category. Review depth and helpfulness votes have a positive effect on review helpfulness. Based on their study, we propose a metric to show the helpfulness of reviews:

$$H = \alpha \cdot R_e + \beta \cdot R_d + \gamma \cdot R_v \tag{1}$$

where R_e, R_d and R_v are rating extremity, review depth and helpfulness votes, respectively. α, β and γ are the weights ($\alpha+\beta+\gamma = 1$). Review depth and the number helpfulness votes can be directly obtained from the original data. Rating extremity in a five-level rating scheme is calculated as the closeness to the neutral rating:

$$R_e = 3 - |r - 3| \tag{2}$$

where r refers to an individual review rating. All the three metrics are normalized to $[0,1]$ by the maximum metric value for the selected local businesses. We empirically set $\alpha = 0.25$, $\beta = 0.25$ and $\gamma = 0.5$. We encode the helpfulness of reviews using the size of each rectangle or circle.

To support detailed review exploration on demand (**R5**), we also include a detailed review text view to show the details of an individual review including business name, customer name, review date, rating, helpfulness votes by other customers and detailed review text, as shown in Figure 1f. Combining the detailed review text view with the temporal view, where the review helpfulness is explicitly encoded as the size of rectangles and circles, the users can easily check the details of helpful reviews and gain a deep understanding of local businesses efficiently.

6.3. Augmented Word Cloud View

Review text represents the detailed evaluation of customers on a business. To compare local businesses through review text more efficiently and accurately, we propose a novel augmented word cloud to support insightful review summary and effective comparison between local businesses (**R4**).

Traditional word clouds focus on visualizing the frequency and sentiment of single words. But in most cases, single words do not convey any contexts, which makes it difficult to know the opinions of the reviews and further compare the businesses. For example, when viewing a word cloud consisting of only single words generated from the restaurant reviews, we are not able to know whether the word *"good"* means the restaurant has *good service, good food or good ambiance*. For the single word *"pizza"*, we are also not able to know whether the pizza is delicious or not. Inspired by [YNTT11b], we show adjective-noun word pairs in a word cloud to provide meaningful context. But instead of randomly positioning word pairs [YNTT11b], we group word pairs with the same noun word into a cluster and carefully place the words to support more effective summary and comparison of reviews. Moreover, we classify the word pairs into different categories based on the aspects they are describing about the local businesses. Users can interactively select the interesting category of word pairs to show, further benefiting an easy comparison of reviews.

The generation of the augmented word cloud view consists of *word pairs extraction and classification* and *word pairs layout*.

Word pairs extraction and classification: We use the part-of-speech (POS) tagger developed in NLTK [nat], a popular language processing library, to label the part of speech of each word in the review text. Then, similar to [YNTT11b], we extract adjective-noun word pairs from each sentence by keeping the noun and the corresponding adjective that modifies it. To accurately reflect customer opinions, we handled the negative expressions specially. For example, when a review says "*the food is not delicious*", we extracted "not delicious food" rather than "delicious food" to preserve the original opinions. A heuristic rule-based method is used here: we check if there are negative adverbs (e.g., *"not", "never", "no"*) between the noun and adjective words in a single sentence with a linking verb (e.g., *"is", "are"*). If yes, then keep the negative adverb in the extraction result. Common abbreviations like *"isn't", "aren't", "can't"* in the reviews are also expanded to guarantee the accurate detection of negative expressions. Furthermore, we conduct sentiment analysis for each word pair by using NLTK.

To classify the adjective-noun word pairs to several categories following the interests of customers, we manually label a set of representative words for each category, then calculate the similarity between the input word pairs and labeled words of each category by using word2vec [MSC*13]. Finally, we assign the word pairs to the category with the highest similarity. For the restaurant reviews in Yelp dataset, we classify the word pairs into four categories: food, price, service and ambiance, which are the four detailed aspects of a restaurant that customers are generally interested in [DFAG13].

Word pairs layout: We propose to render the word pairs of different categories separately to provide a semantically meaningful summary and quick comparison. Users can interact with *E-Comp* to choose the category of interest. In addition, we group word pairs describing the same object together to further enhance the effectiveness of visual summary and comparison of review text. As shown in Figure 8, the adjective words describing the same noun word are clustered and aligned vertically and sorted by their frequency. The color of adjective words represents their word sentiment. Most of the noun words have a neutral sentiment. To provide users with more information about the user reviews, we calculate the sentiment value of the noun words as a weighted average of the sentiment values of all the adjective words modifying this noun word. Blue, black and red are adopted to indicate the positive, neutral and negative sentiments, respectively. The word size of noun words indicates their overall word frequency. The adjective words modifying the same noun word share the same size to make the less frequent adjective words can also be viewed more clearly. The sum of their sizes is equal to the size of the corresponding noun word, reflecting their word frequency to some extent.

A detailed description about rendering the augmented word cloud can be summarized as follows:

1. Group all the word pairs with the same noun word and sort the adjective words based on the word frequency. When there are more than four adjective words modifying a noun word, only the top four adjective words are kept.
2. For all the clusters of adjective-noun word pairs gained from Step 1, sort them by using the noun word frequency.
3. For the sorted list of word pair clusters, render the clusters one by one in descending order of the frequency of the noun word, following a radial layout. The cluster with the highest frequency of the noun word is positioned near the center.
4. Collision detection is performed to check whether the current word pair cluster has spatial overlap with previously placed clusters. When collision exists, the current word pair cluster is moved to the next position following the Archimedean spiral [Fei10] until there is no collision.
5. Repeat Steps 3 and 4 until either all the word pair clusters are rendered or the predefined maximum cluster is reached.

7. Interactions

To allow users smoothly compare local businesses by leveraging the user reviews, *E-Comp* provides a set of intuitive interactions.

Filtering. The control panel supports the filtering of local businesses with a certain price, rating level and the total number of customers. Users can brush on the map view to select interested regions and the local businesses of this region will be shown as glyphs for preliminary comparison. Users can click a glyph to select a local business for detailed comparison or double click to delete an uninterested business. The links with a small width indicating less common customers in the map view can also be interactively removed.

Details on demand. Users can drag, zoom in/out in the map view. They can switch the temporal view between stacked layout and layered layout to check the temporal details of customer ratings. Tooltips are supported for most visual elements of *E-Comp* to show the encoding details on demand.

Linked exploration. Linked analysis across multiple views are supported. When users select other interested businesses in the map view. All the views for detailed comparison will be updated simultaneously. Brushing on the temporal view triggers the selected reviews to be shown in the augmented word cloud view.

8. Evaluations

We evaluate the effectiveness and usability of *E-Comp* through both a case study and in-depth user interviews.

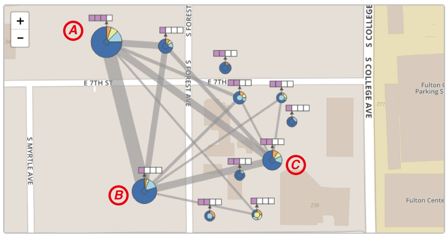

Figure 5: *The map view showing the restaurants filtered by the user in the local region around Arizona State University. Restaurants A,B,C receive more reviews and share many common customers.*

8.1. Case Study on Restaurant Comparison

In this section, we report a real case of restaurant comparison that was conducted by one of our interview participants in Section 8.2. This participant had a travel plan to Tempe and would like to visit his friends at Arizona State University (ASU), so he used *E-Comp* to compare restaurants near ASU.

Preliminary comparison: The participant first navigated to the region around ASU. After viewing the overall distribution of restaurants, he brushed on an interesting region with more restaurants. Then all the restaurants will be visualized as glyphs and the links between them indicate the number of common customers, as shown in Figure 5. He may also use the sliders to filter out the restaurants that obviously do not satisfy his requirements. From Figure 5, the participant easily noticed that Restaurant A receives good ratings from customers and is very popular, indicated by the high percentage of dark blue and big size of the glyph. However, its price bar shows that Restaurant A is a bit expensive for him. The participants quickly noticed that Restaurants B and C shares a large number of common customers with Restaurant A, but are much cheaper than Restaurant A. In addition, both of them also have good ratings and popularity, which makes them perfect restaurant candidates around that region. By hovering or clicking on the glyphs, the participants can quickly know that Restaurants B and C are called "The Chuck Box" and "Original Chopshop", respectively.

Detailed comparison: The selected two restaurants are very similar regarding overall rating, popularity and price, making it difficult to know which one is better. By clicking on the corresponding glyphs, the participant further checked their detailed differences.

From the common comparison view (Figure 6), the participant could quickly know the rating difference between the common customers. For example, the leftmost and rightmost grouped bars indicate that *Chuck Box* and *ChopShop* received a similar number of 4-star and 5-star ratings, but *Chuck Box* has more 1-star and 2-star ratings than *ChopShop*, indicating that *ChopShop* is relatively better than *Chuck Box*. The grouped bars in the middle show more details: about half of the common customers gave the same rating for both restaurants and there also exist a similar number of common customers who prefer one restaurant to the other (the red dotted rectangles in Figure 6). But there are more common customers who believe *ChopShop* is *significantly* better than *Chuck Box*, further confirming that *ChopShop* is relatively better than *Chuck Box*. This is out of the participant's initial expectation, since *Chuck Box*

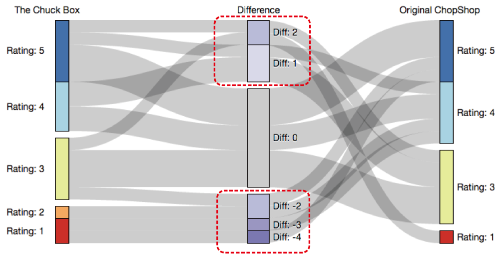

Figure 6: *The common customer comparison view shows the difference of the ratings by the common customers of Chuck Box and ChopShop. The top and bottom grouped bars marked in red dotted rectangles represent the common customers who give a better rating to Chuck Box and ChopShop, respectively.*

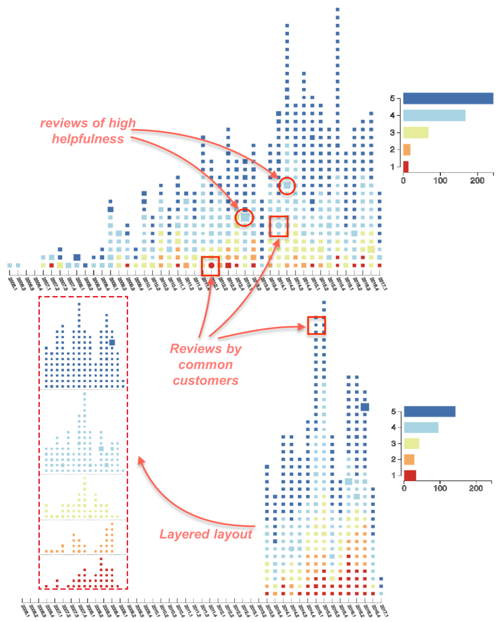

Figure 7: *The temporal view shows the temporal trend of reviews. Top: Chuck Box, bottom: ChopShop. The stacked layout can be switched to a layered layout. The circles represent reviews by common customers and the size of rectangle and circle encodes the helpfulness of reviews.*

is a bit more popular in terms of the number of total customers and has a higher percentage of 5-star ratings.

Figure 7 provides the participant with a clear impression about the temporal trending of user ratings. He immediately knew that *Chuck Box* has received a continuously increasing number of customers and *ChopShop* started its business about 7 years later than *Chuck Box*. Taking into account that both restaurants have a similar number of total customers, it further confirms the relative advantage of *ChopShop*. By switching to layered layout, the participant was able to easily know the review trend of each rating level. The participant could also brush a recent period to view the latest review details.

To glean useful information from the review text, the participant further used the augmented word cloud view to compare different features of the selected restaurants. He compared the food features of the two restaurants first, as shown in Figure 8. At first glance, most of the word pairs for both restaurants are shown in blue color, indicating the overall positive feedback on the food for both restaurants. When looking at the cluster of adjective-noun word pairs with high frequency, the participant quickly noticed that the most representative food enjoyed by customers: *Chuck Box* has “good burger”, “best hamburger”, “classic cheeseburger”, “awesome bacon” and “great beer”, etc., while *ChopShop* may be a perfect restaurant for vegetarians, since its previous customers enjoy its “sweet potato”, “nice rice”, “good salad”, “delicious juice”, etc. When focusing on the words in reddish color, the participant also gained a quick understanding of the major negative comments on the food. For instance, the beer at *Chuck Box* may be disappointing (“disappointed beer”) and the sort of the food on the menu may be weird or confusing (“weird sort”). The salad at *ChopShop* may be expensive (“expensive salad”) and the waitress may give you the wrong food order (“wrong order”). All these information is clearly shown to support an easy comparison. The participant further compared other features such as price, service and ambiance. By brushing on the temporal view, he also compared their reviews posted recently and the reviews of different rating levels.

Taking all the factors above into consideration, the participant said that he could easily obtain an insightful difference between the two restaurants and came to a conclusion that he may choose *ChopShop*, because it receives more support from common customers, becomes similarly popular with *Chuck Box* in a much short time and it also has his favorite vegetable food.

8.2. In-depth User Interview

We also conducted in-depth user interviews to further demonstrate the effectiveness and usability of *E-Comp*.

8.2.1. Study Design

We carefully designed the in-depth user interviews through recruiting representative participants and proposing essential exploration tasks and interview questions.

Participants: We recruited 12 participants (2 female, age: 20 to 34 (mean:25.3)) from our university to take part in our in-depth interviews. They are either students or research staff with a research background in business or engineering. All of them are familiar with E-commerce and have at least three years’ online shopping experience. To guarantee that the findings from the in-depth user interviews are general for common users, none of the participants has a background of visualization or HCI.

Apparatus and Procedures: We run *E-Comp* on a laptop connected to a 23-inch display with a 1920 $\times$ 1080 pixels resolution and 60 Hz refresh rate. By using this, we conducted all the user interviews with each participant one by one. Each participant interview took about 50 to 70 minutes. We first briefly introduced *E-Comp*. Then we went through an example to explain the functions and visual encoding of *E-Comp*. After that, we asked the participants to freely explore the system and use it to compare businesses they are interested in. Considering the key steps of making purchase decisions, we selected the following tasks for participants to finish in their free explorations:

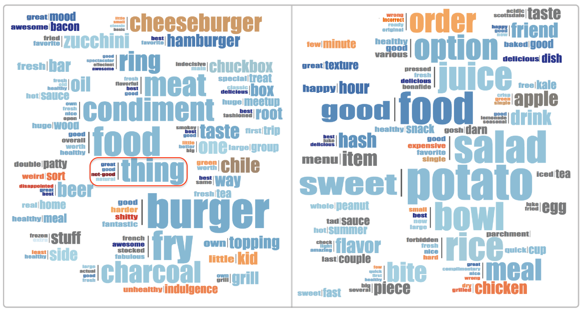

Figure 8: *The augmented word cloud view shows the clustered adjective-noun word pairs of food feature. Left: Chuck Box, right: ChopShop. An example of negative expression (i.e., "not-good") is shown in the cluster highlighted in a red rounded rectangle.*

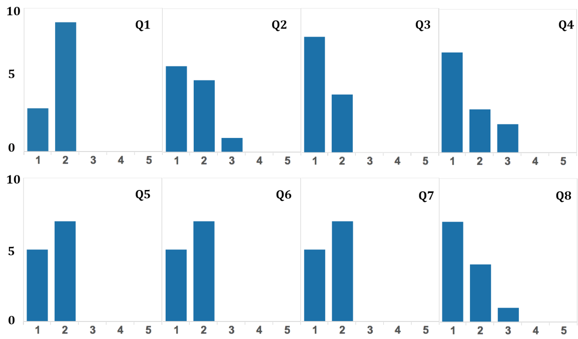

Figure 9: *The results of the post-study. User responded to Questions 1-8 on a 5-point Likert scale, where 1 to 5 represent the difficulty levels from "very easy" to "very difficult".*

T1 Find interesting local businesses with the preferable rating, price, location and popularity from the map view.
T2 Find the difference of review ratings by common customers.
T3 Compare the temporal trend of review amount of the selected business regarding both overall and different levels of ratings.
T4 Find the differences of the selected venues in terms of food, service, price and ambiance.
T5 Explore the major reasons that customers like or complain about a business from the high and low rating reviews.

We asked the participants to comment verbally during their exploration to collect more detailed user feedback. We kept observing the exploration process of each participant and wrote down their major operations and comments. After participants finished the free exploration, we interviewed each participant about their general comments on *E-Comp*, including the advantages and possible limitations, and the helpfulness of *E-Comp* in facilitating local business comparison. A post-study questionnaire was also conducted to collect detailed feedback on the effectiveness and usability of *E-Comp*, where eight questions were asked (Table 1). The participants responded to each question on a 5-point Likert scale (1=very easy, 5=very difficult) and provided the main reasons for their response.

It is also interesting to further compare *E-Comp* with existing review platforms. However, since the existing review platforms (e.g., the Yelp website) usually do not support an explicit visual comparison of local businesses, it is probably unfair and inappropriate to conduct a strict comparison between them and *E-Comp*. Therefore, we simply asked participants to try *E-Comp* and *Yelp* again and collected their general comments about the usability and effectiveness comparison between both systems for comparing local businesses.

8.2.2. Result

Overall, our in-depth user interviews demonstrate that *E-Comp* can effectively help users conduct an easy comparison between local businesses and it has good usability even for users who have no background knowledge in visualization or HCI.

Effectiveness: Figure 9 shows that most participants can use *E-Comp* without difficulty to quickly compare different local business, providing support for the effectiveness of *E-Comp*. For preliminary comparison, all the participants agree that the designed glyphs and links between them facilitate a quick comparison between local businesses (Q1), because the most important factors for selecting businesses, such as the percentage of good/bad ratings, average rating, popularity (i.e., the total number of reviews), price and location, are intuitively encoded. The links between glyphs provide participants with explicit hints about the common customers between businesses, accelerating the preliminary selection.

After preliminary comparison, the participants usually selected two local businesses with good ratings, popularity and similar prices. The participants commented that they appreciate the detailed comparison functions of *E-Comp*, which are very helpful for uncovering the detailed differences behind the overall similarity of the local businesses. From common customer comparison view, 11/12 participants agreed that they could discover some differences between businesses that may be hidden in the overall ratings, as shown in Q2 of Figure 9. For example, Participant 2 (P2) selected two sushi restaurants with exactly the same average rating and a similar number of total customers, but the common customer view indicates that lots of their common customers gave a better rating to one restaurant than the other. P2 commented that such information is usually unable to gain by reading the original reviews.

Other participants' feedback also confirmed that common customer view provides a quick way to gain insights into the subtle differences between local businesses. All the participants could easily understand the temporal trend of both overall and each level of ratings, as shown in Figure 9 Q3. The interactions in temporal view, especially the brushing and switching between stacked and layered layout, are also enjoyed by many participants (e.g., P1, P5, P6, P12), as they facilitate exploring reviews within the period and rating level of interest. With the augmented word cloud view, most participants (10/12) said that they could gain a quick summary about the major characteristics of each business in different feature dimensions, as shown in Q4 of Figure 9. P5 and P7 explicitly pointed out that the grouped adjective words, combined with the color-encoded sentiment, provide them with enough context to quickly understand the review opinions in the original review text. All the participants (Q5 of Figure 9) confirmed that the helpfulness encoded by the size of rectangles and circles of temporal view is useful for them to focus on the reviews with more helpful information. *E-Comp* provides them with a convenient way to check the review details by simply clicking on the rectangle or circle. Taking

Table 1: *The post-study questions. Q1-6 are focusing on assessing the effectiveness of E-Comp in facilitating an easy comparison of local businesses and Q7-8 evaluate its usability.*

Q1	Is it easy for you to do preliminary comparison and quick selection of business candidates by viewing the glyphs and links on map view?
Q2	Is it easy for you to recognize the differences of local businesses that may be hidden in the overall rating?
Q3	Is it easy for you to quickly understand the differences of temporal trend of overall and individual ratings?
Q4	By using the augmented word cloud view, is it easy for you to compare the basic features of two businesses?
Q5	Is it easy to for you to find out the relatively more helpful reviews and check the details of them?
Q6	Overall, is it easy for you to compare different local businesses, especially when compared with reading the original review text?
Q7	Is it easy for you to learn to use *E-Comp*?
Q8	Is it easy for you to understand the overall visual designs of *E-Comp*?

all these factors into consideration, all the participants evaluated *E-Comp* as an excellent tool for facilitating the comparison between different businesses, especially when compared with the way of reading the original reviews to do a comparison (Q6 of Figure 9).

Usability: Q7-8 of Figure 9 indicate that there is no difficulty for the general users to learn to use *E-Comp* and most of the participants can easily understand the visual designs (**R6**). The user feedback also shows the interactions integrated in *E-Comp* are generally smooth and helpful for the comparison. To sum up, these feedback demonstrates the good usability of *E-Comp*.

E-Comp vs. Yelp: Overall, the participants' feedback shows that *E-Comp* has an advantage over Yelp in supporting an easy comparison of local businesses. When using the Yelp website, they can only rely on the summary features like price, average rating and the total review number to filter candidates. For the detailed comparison, they need to read the review text and check back and forth to do a comparison, which is not efficient. The feedback further confirms the necessity of a review comparison tool and our design requirements in Section 3.1. Contrarily, *E-Comp* explicitly supports both preliminary and detailed comparisons, facilitating the comparison of local businesses. The advantage of Yelp, as mentioned by many participants (e.g., P2, P3 and P7), is that Yelp also shows other information like food images, restaurant images and user profiles.

Limitations and Suggestions: Despite all the positive feedback above, our participants also pointed out several limitations of *E-Comp* and offered suggestions. For example, P5 suggested supporting collision detection between glyphs in the map view to avoid occlusion of glyphs, which has been implemented in the latest prototype of *E-Comp*. P10 suggested that *E-Comp* can also benefit business owners, as they can use *E-Comp* to analyze reviews and improve their own business. P3 and P7 mentioned that it would be interesting if the posted images by customers can also be analyzed in *E-Comp*, which have been left for future work.

9. Discussion and Conclusion

In this paper, we have presented *E-Comp*, a visual analytics system to facilitate the comparison between local businesses by using online reviews. We designed novel glyphs, which are overlaid on the map, to support better preliminary comparison. Detailed comparison is achieved through carefully-designed visual encoding from the perspective of rating differences by common customers, temporal evolution of reviews and detailed features in the review text. *E-Comp* also supports rich interactions, allowing for flexible visual exploration. The case study and in-depth user interviews using Yelp dataset demonstrate the great effectiveness and usability of *E-Comp* for supporting easy comparison of local businesses. However, there are still several aspects that need further discussion.

Other applications and users. We use the restaurant reviews to demonstrate the effectiveness and usability of *E-Comp*. But it can also be applied in other local businesses like shops, fashion retailers and local services (e.g., barbershops, health clinics and gyms), as the metadata is exactly the same. Moreover, *E-Comp* can be easily extended to the review comparison of products in online shopping, where the only difference from local businesses is that the location information (the map of Figure 1B) is not necessary. Apart from the general customers, *E-Comp* can also be used for other users like restaurant managers, shop owners, etc., when they want to compare their establishment with other business competitors.

Scalability. E-Comp focuses on the preliminary comparison between tens of local businesses and detailed comparison between two candidates. This design choice conforms to the reported general exploration procedures of customers in Section 3.1 and strikes a balance between the details to be shown and the potential visual clutter resulting from the limited screen space. Therefore, *E-Comp* is not suitable for comparing hundreds of local businesses or even all the local businesses in a whole city due to the visual clutter.

Desktop system vs. mobile program. Currently, *E-Comp* is designed as a desktop-based system to help customers easily make a purchase decision in many situations, such as finding a good restaurant for an important celebration or booking a suitable hotel for a distant trip. However, with the wide use of smart mobile phones, a mobile version of *E-Comp* will also benefit the general customers.

In future work, we will try to further improve *E-Comp*, for example, building a mobile version of *E-Comp* and enhancing the word cloud view by considering the perceptual biases in font size [ACS*17]. In addition, we plan to further analyze the images contained in the online reviews along with review text and ratings to compare local businesses comprehensively. Moreover, we will apply *E-Comp* to more real datasets to further examine the effectiveness and usability of *E-Comp*.

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful comments and suggestions. This research was supported in part by HK RGC GRF 16208514.

References

[ACS*17] ALEXANDER E. C., CHANG C.-C., SHIMABUKURO M., FRANCONERI S., COLLINS C., GLEICHER M.: Perceptual biases in font size as a data encoding. *IEEE Transactions on Visualization and Computer Graphics PP*, 99 (2017), 1–1. 10

[BAC12] BAEK H., AHN J., CHOI Y.: Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce 17*, 2 (2012), 99–126. 1, 3, 5

[BEW16] BERTINI E., ELMQVIST N., WISCHGOLL T.: Judgment error in pie chart variations. In *Proceedings of the Eurographics / IEEE VGTC Conference on Visualization: Short Papers* (2016), pp. 91–95. 4

[BHML16] BEIGI G., HU X., MACIEJEWSKI R., LIU H.: An overview of sentiment analysis in social media and its applications in disaster relief. In *Sentiment Analysis and Ontology Engineering*, Pedrycz W., Chen S.-M., (Eds.). Springer, Cham, 2016, pp. 313–340. 2

[BLPW13] BURCH M., LOHMANN S., POMPE D., WEISKOPF D.: Prefix tag clouds. In *17th International Conference on Information Visualisation Information Visualisation* (2013), IEEE, pp. 45–50. 2

[BML17] BADER N., MOKRYN O., LANIR J.: Exploring emotions in online movie reviews for online browsing. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion* (2017), ACM, pp. 35–38. 2

[BPF14] BACH B., PIETRIGA E., FEKETE J.-D.: Graphdiaries: Animated transitions andtemporal navigation for dynamic networks. *IEEE Transactions on Visualization and Computer Graphics 20*, 5 (2014), 740–754. 2

[Cad15] CADARIO R.: The impact of online word-of-mouth on television show viewership: An inverted U-shaped temporal dynamic. *Marketing Letters 26*, 4 (2015), 411–422. 4

[CLLT15] CAO N., LIN Y.-R., LI L., TONG H.: g-miner: Interactive visual group mining on multivariate graphs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), ACM, pp. 279–288. 4

[DES*15] DIAKOPOULOS N., ELGESEM D., SALWAY A., ZHANG A., HOFLAND K.: Compare clouds: Visualizing text corpora to compare media frames. In *Proceedings of IUI Workshop on Visual Text Analytics* (2015), pp. 193–202. 2

[DFAG13] DI FABBRIZIO G., AKER A., GAIZAUSKAS R.: Summarizing online reviews using aspect rating distributions and language modeling. *IEEE Intelligent Systems 28*, 3 (2013), 28–37. 3, 4, 6

[Fei10] FEINBERG J.: Wordle. In *Beautiful Visualization: Looking at Data through the Eyes of Experts*, Steele J., Iliinsky N., (Eds.). O'Reilly Media, 2010, ch. 3, pp. 37–58. 7

[FM14] FILIERI R., MCLEAY F.: E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of Travel Research 53*, 1 (2014), 44–57. 1

[GAW*11] GLEICHER M., ALBERS D., WALKER R., JUSUFI I., HANSEN C. D., ROBERTS J. C.: Visual comparison for information visualization. *Information Visualization 10*, 4 (2011), 289–309. 2

[Gle18] GLEICHER M.: Considerations for visualizing comparison. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 413–423. 2

[GLG*13] GRATZL S., LEX A., GEHLENBORG N., PFISTER H., STREIT M.: Lineup: Visual analysis of multi-attribute rankings. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2277–2286. 2

[GS12] GODES D., SILVA J. C.: Sequential and temporal dynamics of online opinion. *Marketing Science 31*, 3 (2012), 448–473. 1, 3

[GY08] GRETZEL U., YOO K. H.: Use and impact of online travel reviews. *Information and Communication Technologies in Tourism* (2008), 35–46. 1

[HL04] HU M., LIU B.: Mining and summarizing customer reviews. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2004), ACM, pp. 168–177. 2, 4

[HPZ06] HU N., PAVLOU P. A., ZHANG J.: Can online reviews reveal a product's true quality? empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM conference on Electronic Commerce* (2006), ACM, pp. 324–330. 1

[HVW08] HOLTEN D., VAN WIJK J. J.: Visual comparison of hierarchically organized data. *Computer Graphics Forum 27*, 3 (2008), 759–766. 2

[HWS17] HU M., WONGSUPHASAWAT K., STASKO J.: Visualizing social media content with sententree. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 621–630. 2

[KBH06] KOSARA R., BENDIX F., HAUSER H.: Parallel sets: Interactive exploration and visual analysis of categorical data. *IEEE Transactions on Visualization and Computer Graphics 12*, 4 (2006), 558–568. 5

[KGSZ11] KIM H. D., GANESAN K., SONDHI P., ZHAI C.: *Comprehensive review of opinion summarization*. Tech. rep., 2011. 2

[KH13] KEHRER J., HAUSER H.: Visualization and visual analysis of multifaceted scientific data: A survey. *IEEE Transactions on Visualization and Computer Graphics 19*, 3 (2013), 495–513. 2

[KHW09] KLIPPEL A., HARDISTY F., WEAVER C.: Star plots: How shape characteristics influence classification tasks. *Cartography and Geographic Information Science 36*, 2 (2009), 149–163. 4

[KPBG13] KEHRER J., PIRINGER H., BERGER W., GRÖLLER M. E.: A model for structure-based comparison of many categories in small-multiple displays. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2287–2296. 2

[KPK17] KUCHER K., PARADIS C., KERREN A.: The state of the art in sentiment visualization. *Computer Graphics Forum 37*, 1 (2017), 71–96. 2

[LHC05] LIU B., HU M., CHENG J.: Opinion observer: Analyzing and comparing opinions on the web. In *Proceedings of the 14th International Conference on World Wide Web* (2005), ACM, pp. 342–351. 1, 2, 3

[Lor17] LORIA S.: TextBlob, python library for processing textual data., 2017. URL: https://web.archive.org/web/20161222050214/http://textblob.readthedocs.io/en/dev/. 2

[LSY03] LINDEN G., SMITH B., YORK J.: Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing 7*, 1 (2003), 76–80. 5

[Mor11] MORTON J.: Online business reviews and the public figure doctrine: An advertising-based standard. *Hastings Communications and Entertainment Law Journal 34* (2011), 403–424. 1

[MS10] MUDAMBI S. M., SCHUFF D.: What makes a helpful review? A study of customer reviews on Amazon.com. *Management Information Systems Quarterly 34*, 1 (2010), 185–200. 3, 5, 6

[MSC*13] MIKOLOV T., SUTSKEVER I., CHEN K., CORRADO G., DEAN J.: Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems* (2013), pp. 3111–3119. 6

[nat] Natural language toolkit. http://www.nltk.org/. Accessed: 2017-03-10. 2, 6

[Nie] NIELSEN: The reviews are in: Yelp users are four-star consumers. http://www.nielsen.com/us/en/insights/news/2013/the-reviews-are-in--yelp-users-are-four-star-consumers.html. Accessed: 2018-03-01. 1

[NSH*18] NIEDERER C., STITZ H., HOURIEH R., GRASSINGER F., AIGNER W., STREIT M.: Taco: Visualizing changes in tables over time. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 677–686. 2

[OHR*09] Oelke D., Hao M., Rohrdantz C., Keim D. A., Dayal U., Haug L.-E., Janetzko H.: Visual opinion analysis of customer feedback data. In *IEEE Symposium on Visual Analytics Science and Technology* (2009), IEEE, pp. 187–194. 2

[PL08] Pang B., Lee L.: Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval 2*, 1–2 (2008), 1–135. 2

[PS80] Presser S., Schuman H.: The measurement of a middle position in attitude surveys. *Public Opinion Quarterly 44*, 1 (1980), 70–85. 5

[SK16] Skau D., Kosara R.: Arcs, angles, or areas: Individual data encodings in pie and donut charts. *Computer Graphics Forum 35*, 3 (2016), 121–130. 4

[SL91] Spence I., Lewandowsky S.: Displaying proportions and percentages. *Applied Cognitive Psychology 5*, 1 (1991), 61–77. 4

[SMDS14] Sadana R., Major T., Dove A., Stasko J.: Onset: A visualization technique for large-scale binary set data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1993–2002. 2

[SWL*14] Shi C., Wu Y., Liu S., Zhou H., Qu H.: Loyaltracker: Visualizing loyalty dynamics in search engines. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1733–1742. 2

[War08] Ward M. O.: Multivariate data glyphs: Principles and practice. In *Handbook of Data Visualization*, Chen C., Härdle W. K., Unwin A., (Eds.). Springer, 2008, pp. 179–198. 4

[WDSC07] Wood J., Dykes J., Slingsby A., Clarke K.: Interactive visual exploration of a large spatio-temporal dataset: Reflections on a geovisualization mashup. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1176–1183. 2

[WH08] Wu F., Huberman B. A.: How public opinion forms. In *International Workshop on Internet and Network Economics* (2008), Springer, pp. 334–341. 1, 3, 4

[WSA*16] Wang Y., Shen Q., Archambault D., Zhou Z., Zhu M., Yang S., Qu H.: Ambiguityvis: Visualization of ambiguity in graph layouts. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 359–368. 2

[WWL*10] Wu Y., Wei F., Liu S., Au N., Cui W., Zhou H., Qu H.: Opinionseer: Interactive visualization of hotel customer feedback. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1109–1118. 1, 2

[WZG*14] Wang J., Zhao J., Guo S., North C., Ramakrishnan N.: Recloud: Semantics-based word cloud visualization of user reviews. In *Proceedings of Graphics Interface* (2014), Canadian Information Processing Society, pp. 151–158. 1, 2

[yel] Yelp dataset challenge. https://www.yelp.com/dataset_challenge/. Accessed: 2017-03-10. 4

[YHSZ16] Yu N., Huang M., Shi Y., Zhu X.: Product Review Summarization by Exploiting Phrase Properties. In *Proceedings of the 26th International Conference on Computational Linguistics* (2016), pp. 1113–1124. 2

[YLG09] Ye Q., Law R., Gu B.: The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management 28*, 1 (2009), 180–182. 1

[YNTT11a] Yatani K., Novati M., Trusty A., Truong K. N.: Analysis of adjective-noun word pair extraction methods for online review summarization. In *IJCAI International Joint Conference on Artificial Intelligence* (2011), pp. 2771–2776. 2

[YNTT11b] Yatani K., Novati M., Trusty A., Truong K. N.: Review spotlight: A user interface for summarizing user-generated reviews using adjective-noun word pairs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), ACM, pp. 1541–1550. 1, 2, 6

[ZWC*18] Zhao X., Wu Y., Cui W., Du X., Chen Y., Wang Y., Lee D. L., Qu H.: Skylens: Visual analysis of skyline on multi-dimensional data. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018), 246–255. 2