

Sentiment Analysis of Customers Reviews on Restaurants

Yefei Fan, Spencer Glass, Tariq Haque, Liya Zhang

Kehu is an application that takes unique user-generated restaurant reviews and processes them with SQL and RShiny to gain insights on overall trends in the reviews. It uses real-world Yelp reviews as its data source. The brand name is derived from Chinese words 客户获取 (Kèhù huòqǔ) meaning customer acquisition.

Our mission statement:

“*to leverage user-generated data to develop analytical solutions that will allow clients to improve sales effectiveness*”

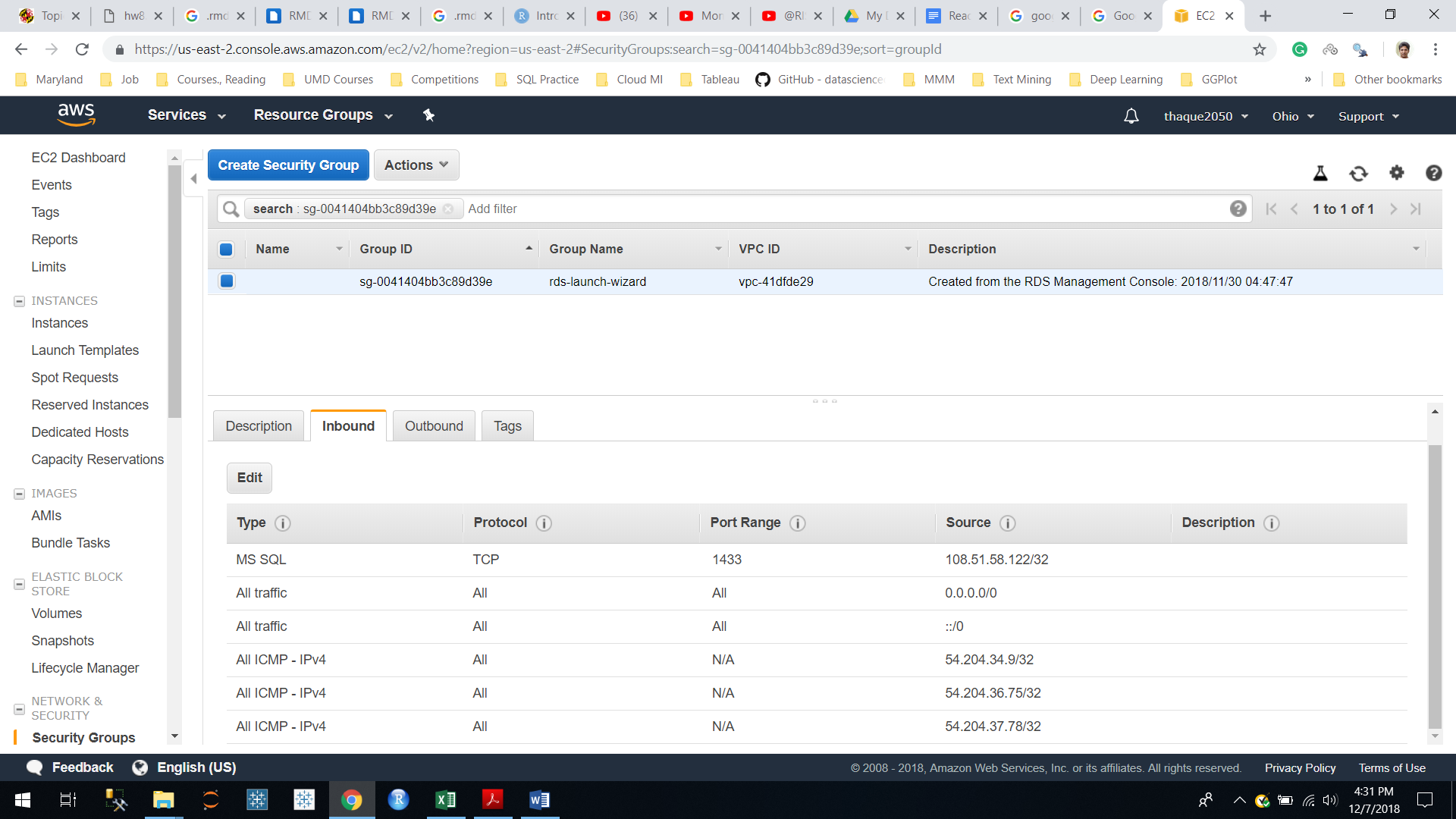
Kehu provides business intelligence tool to sales force to help them engage effectively with their customer during sales calls. To achieve this, Kehu allows a restaurant (a customer that has sought consulting services about their reviews) to sort and view analyzed reviews by many categories. Kehu also uses a sentiment analysis package to quantify sentiment trends in reviews. Kehu can also identify a business’ competitors by their similar attributes and then compare them via metrics like overall star ratings and sentiment ratings.

**DATA SOURCE**

Yelp provides the data for open challenges however the data was in NDJSON format. We used jsonlite package of R to parse data and store into csv formats. The code for parsing and storing as csv file is also included with the suite.

**DATABASE DESIGN**

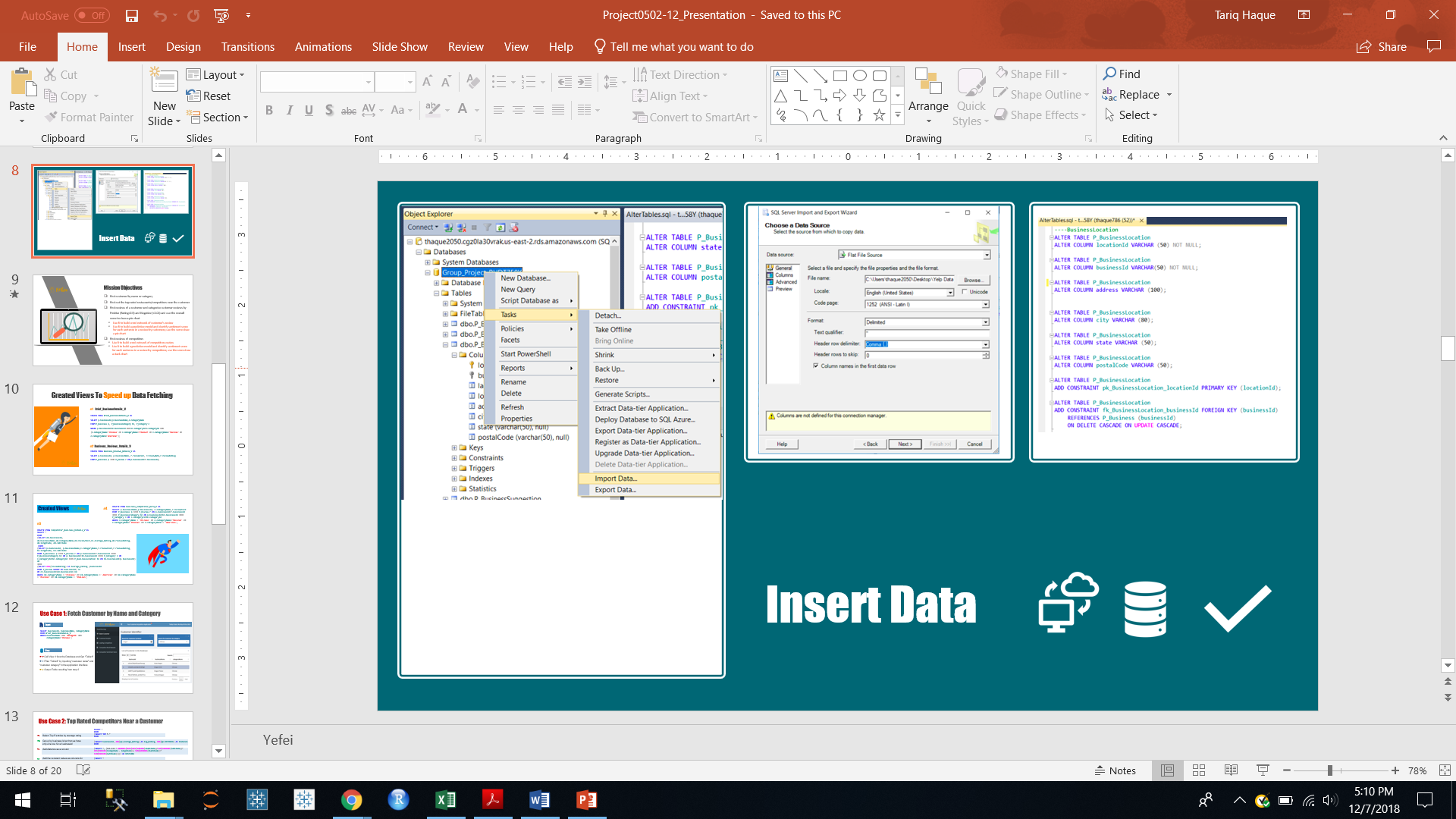
Since the on-premise server space provided by the business school was behind firewall, we used public cloud offered by amazon web services to host our data. AWS gave us choice to enable database access to any inbound traffic.



The database can be accessed using the details as follows from the Microsoft Management Studio 2012:

|  |  |
| --- | --- |
| Image result for aws | Server=thaque2050.cgz0la30vrak.us-east-2.rds.amazonaws.com; 1433  Database=Group\_Project\_BUDT758Y  UID=thaque786  PASSWRD=tariq214H |

For data insertion on database, we have used csv files to upload the document on server and then used alter table to set primary and foreign keys. The following screen shot provided the required details. We have included the .SQL file for alter table with the suite.



**FINAL PRODUCT**

Kehu is hosted online, at https://kehu2050.shinyapps.io/finalversion22/

**INSTALLATION/EXECUTION**

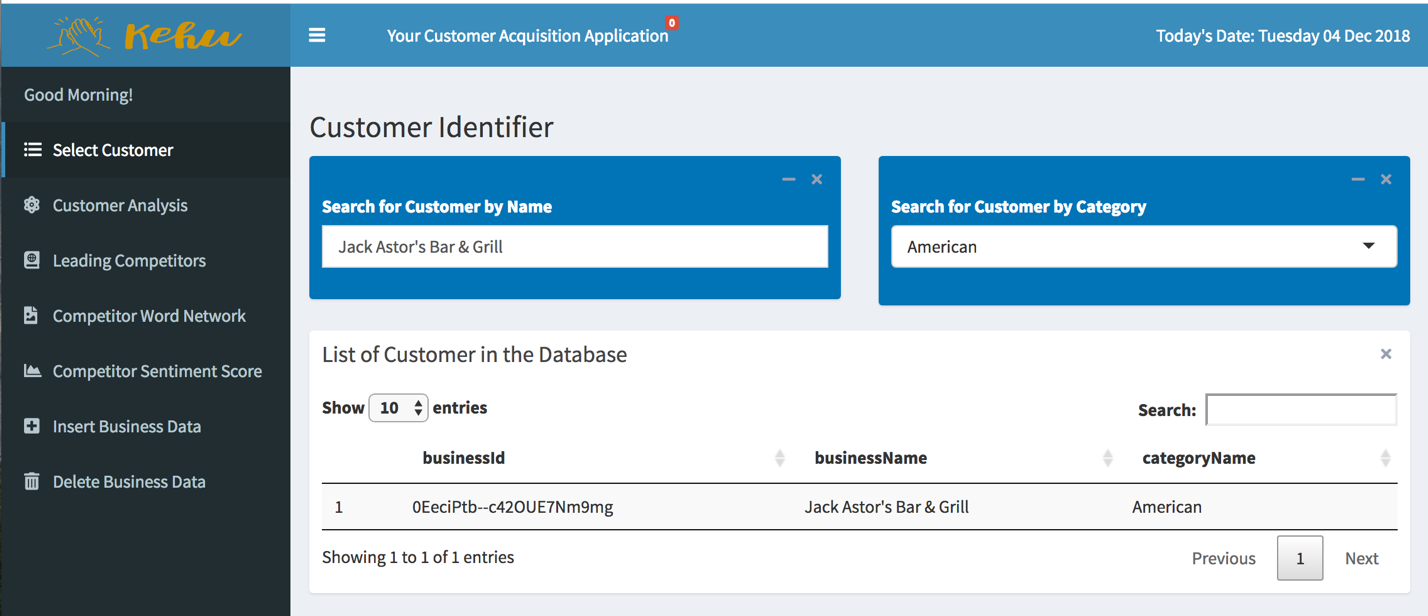
|  |  |
| --- | --- |
| **Install on Local Machine** | 1. Open file “**FinalVersion\_Local Machine**” in RStudio and run the app 2. The app is ready to use |
| **Publish on shinyapps.io** | 1. Create account on shinyapps.io and get token and secret key as follows. Copy this entire text and paste in the location described in step 3.      1. Install and load **rsconnect** package in Rstudio 2. Get Rstudio ready to host apps on shinyapps.io using the following steps:    1. Go to “Manage Accounts” as seen in the screenshot below      * 1. Click on the shinyapps.io as shown below and a new window will appear asking for token and secret key      * 1. Paste the toke and secret key obtained in step 1 in the box as shown below and click connect account. The Rstudio is ready to host apps on shinyapps.io.      1. Open file “**FinalVersion\_ShinyApps\_Hosting**” in Rstudio    1. Use the button below to publish application      * 1. Enter title of your choice as seen in the screenshot below and then click publish      1. The app will now appear in a browser and is ready to use |

**APPLICATION USAGE**

All features of Kehu can be found in the left-most column. There are 7 functions available here.

**1. “Select Customer” Tab**

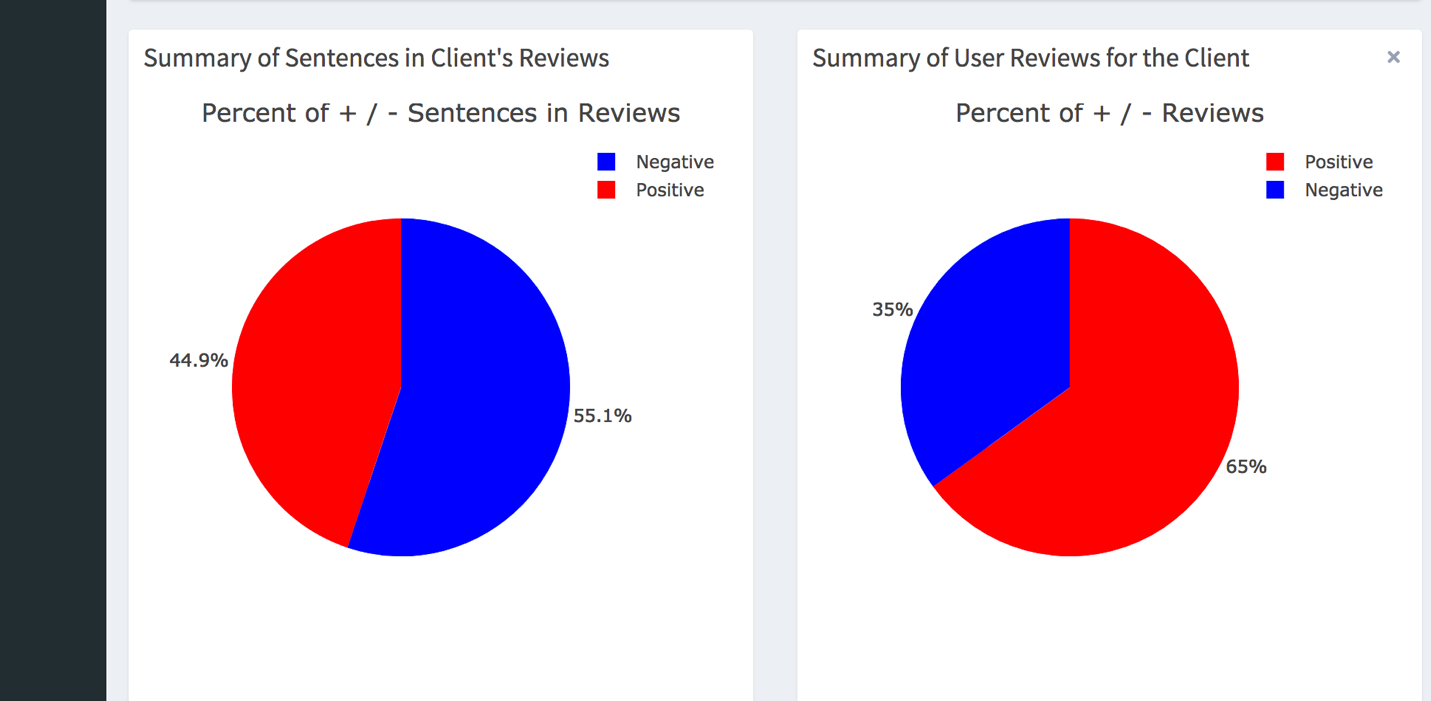
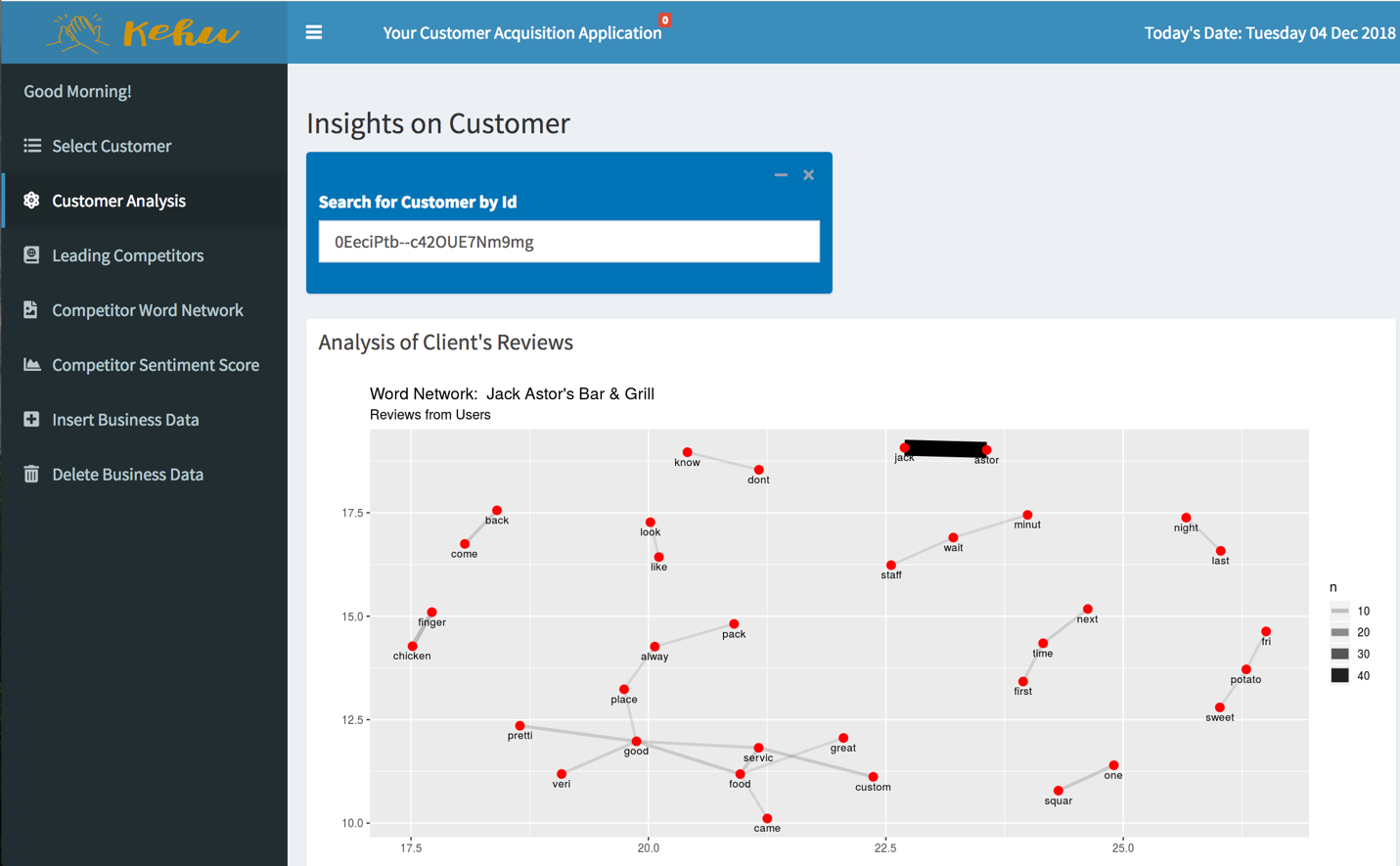
Once this tab is selected, the user can enter a restaurant name and/or restaurant category in the two boxes shown. Then, they can search the database and view all results, which include the businessId, businessName, and categoryName of the restaurants that are returned by the search.



**2. “Customer Analysis” Tab**

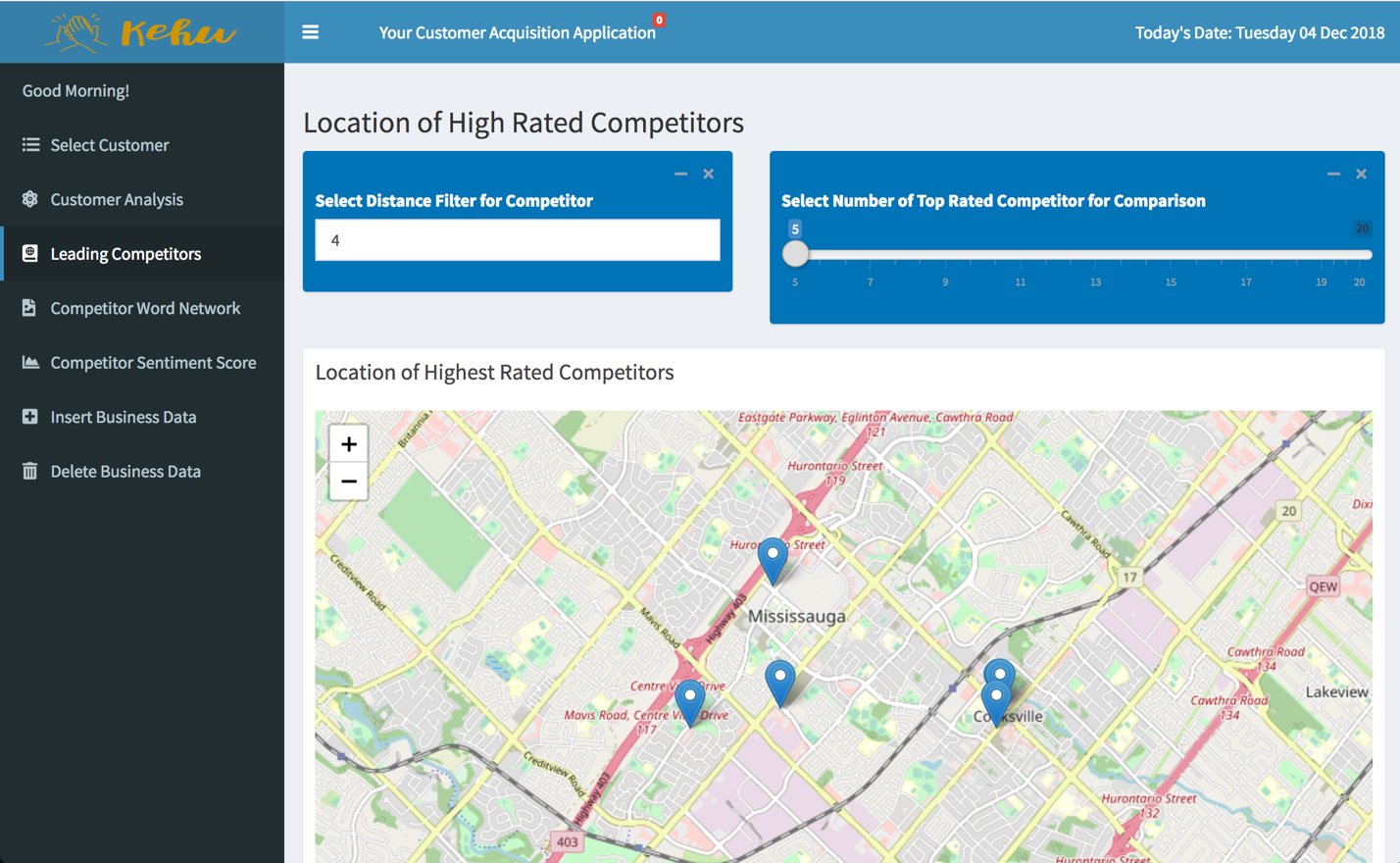
Once this tab is selected, the user can search for a restaurant (identified in the first tab) by using its businessId. Then, all the text of all reviews for that restaurant are analyzed with three visuals.

* The first is a word network of all the texts of all reviews and allows user to gather intelligence from the reviews
* Second is a pie chart showing the percent breakdown of positive and negative sentences (as determined by an R sentiment analysis package) in all the reviews.
* Third is another pie chart showing the percent breakdown of overall positive and overall negative reviews as provided by the user (>2.9 is positive and <2.01 is considered negative)



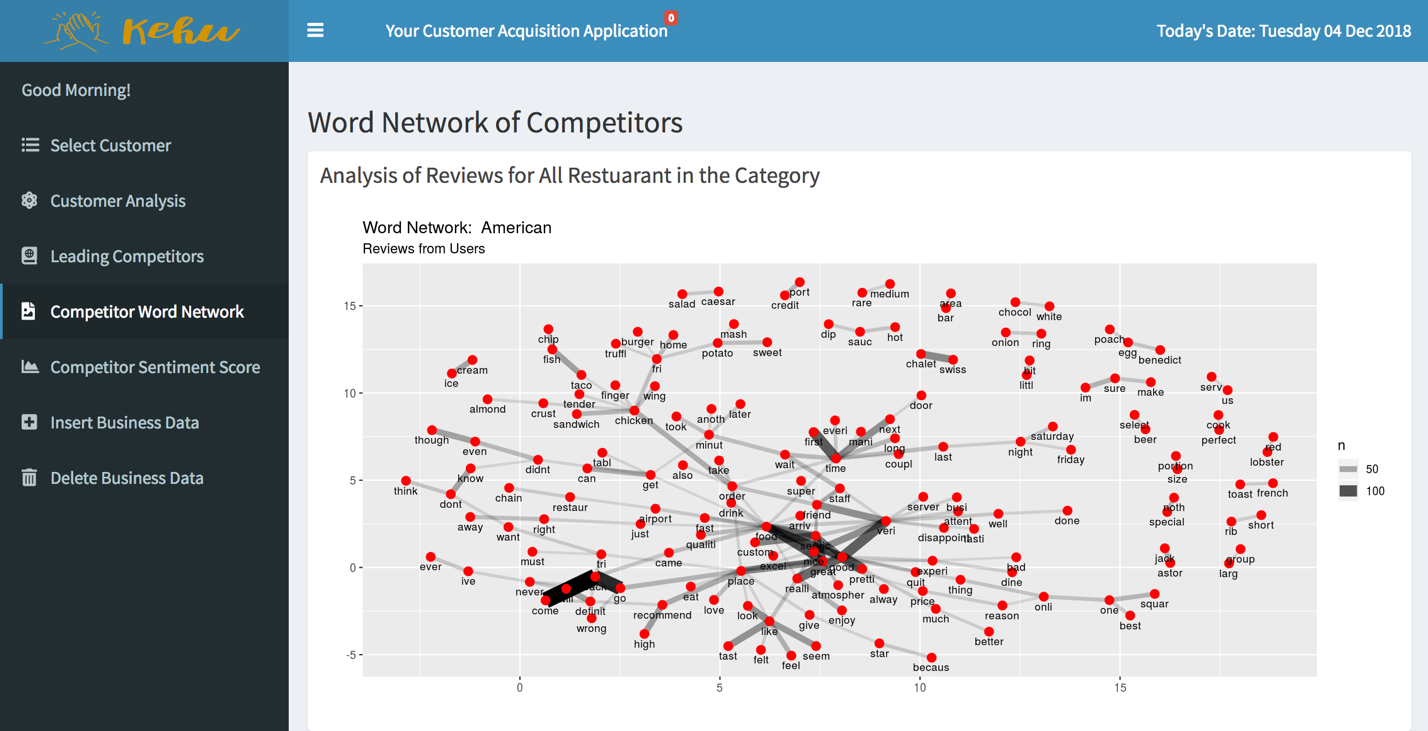
**3. “Leading Competitors” Tab**

Once this tab is selected, the user can view competing restaurants (restaurants in the same category) relative to the restaurant identified in the first tab. They can select how many miles of proximity they want to view results for, as well as how many results to return. These locations are then visualized on a map using R’s leaflet package. Customer select happens in the previous tab.



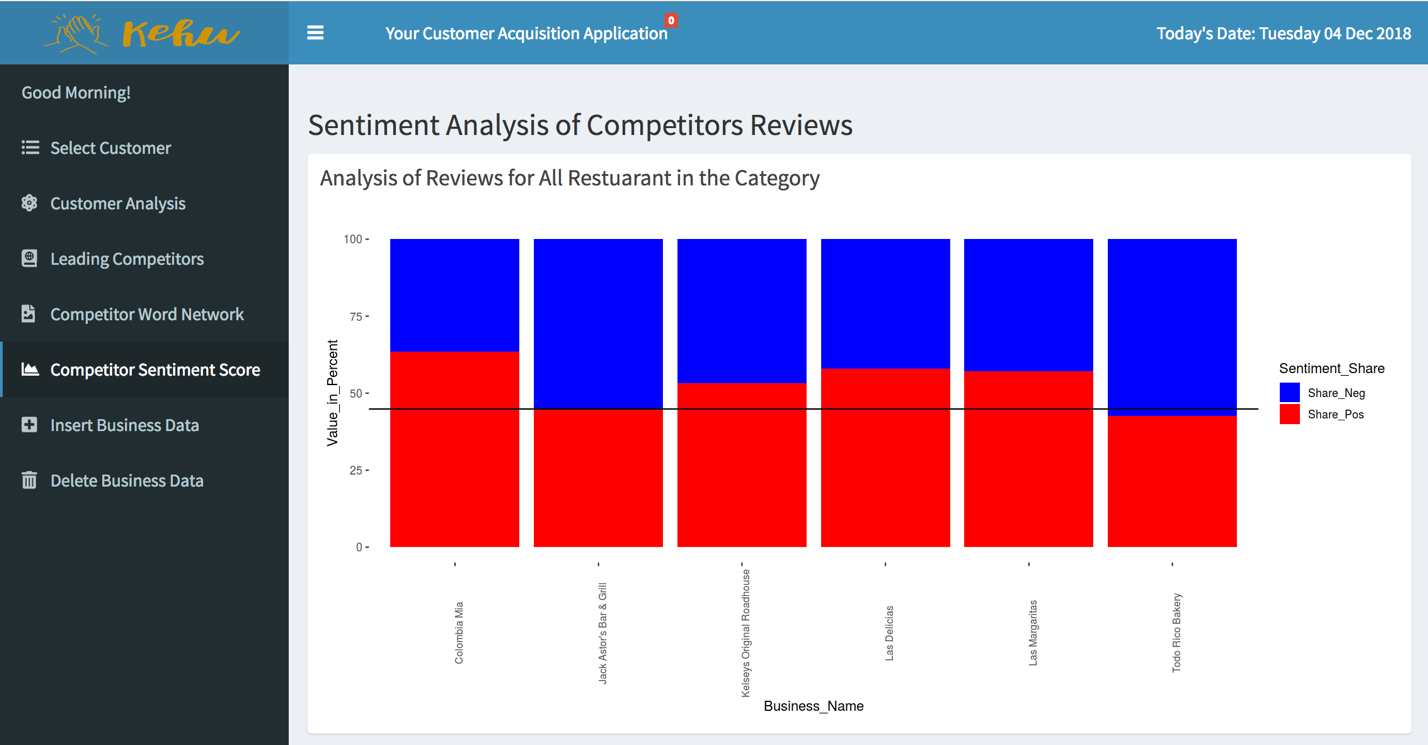
**4. “Competitor Word Network” Tab**

Once this tab is selected, the user can view a word network of all restaurants in the same category as the selected restaurant.



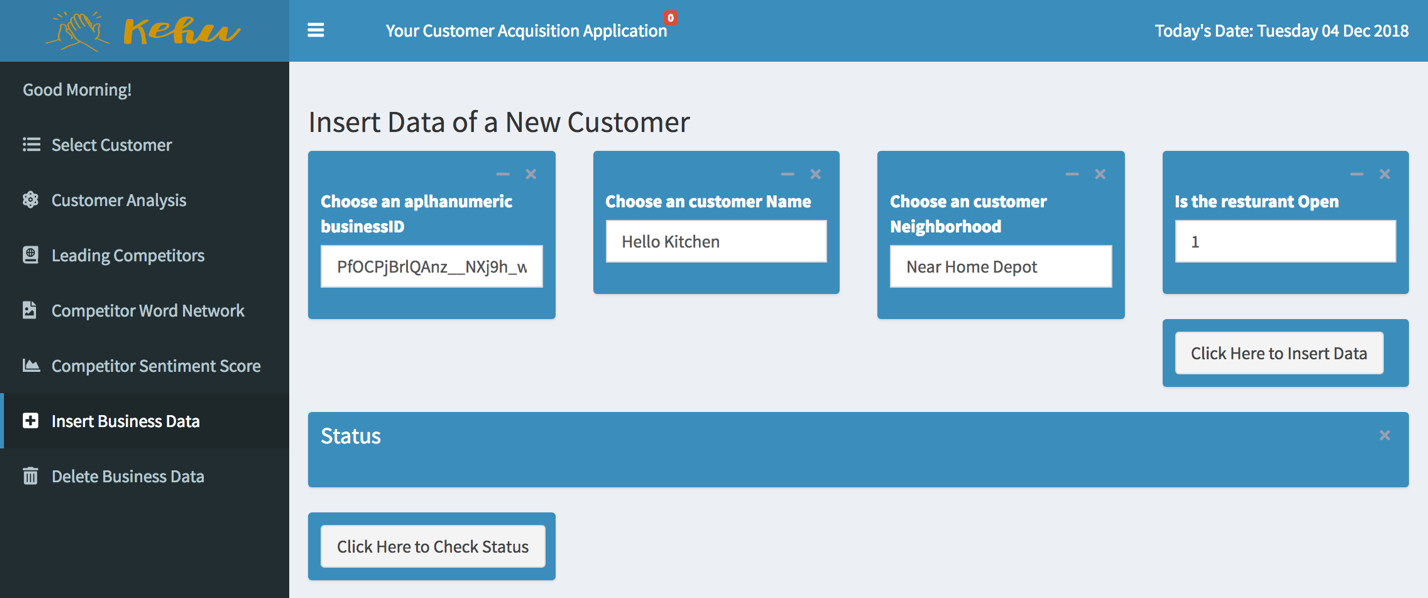
**5. “Competitor Sentiment Score” Tab**

Once this tab is selected, the percent breakdown of sentiment scores of reviews for competing restaurants are shown.



**6. “Insert Business Data” Tab**

Once this tab is selected, a user can add in any data about attributes in the Business Table into the database that they’d like. The functionality can be enhanced further to allow users to insert entries into more tables.



**7. “Delete Business Data” Tab**

Once this tab is selected, a user can delete all data about any restaurant that they’d like by entering the businessId. The functionality can be enhanced further to allow users to delete entries from more tables.

