

Relp

Taikgun Song

4/29/2017

Abstract

A popular webpage yelp.com was reconstructed into a `shiny` application. The application focuses on imitating two basic features of the webpage: providing user-generated reviews and ratings, and providing basic information of the local businesses. The imitation was done by first scraping the yelp.com webpage, and re-building the features with interactive visualization and data re-organization via R language and Shiny application.

Introduction

The name of the project re-construction of yelp.com using R. `Relp` (reads Re-lp) building my own web page.

Since the application educational purpose reviewing class materials from STAT585 course.

The application was focused on learning HTML language framework, implementing R output into

yelp.com

yelp.com (“Yelp” 2017) is an online and mobile application providing platforms for internet users to publish reviews and rate local businesses, as well as online delivery service and reservation service. For the project, only the user-generated reviews and local business information was taken into consideration.

Some of the basic features of this webpage was reconstructed using the `shiny`(Chang and others 2017) application in R(R Core Team 2017) framework.

`shiny`(Chang and others 2017), `tidyverse` (Wickham and RStudio 2017), `ggmap` (Kahle and Wickham 2016), `leaflet` (Cheng 2017), `DT` (Xie 2016).



Figure 1: Front page of the ‘yelp.com’



Figure 2: Main page of the ‘yelp.com’

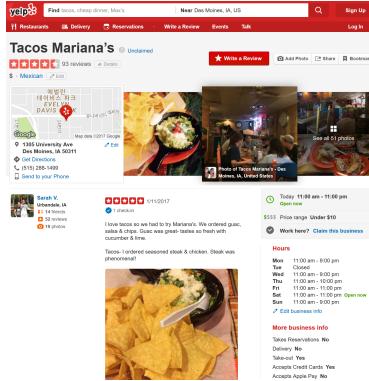


Figure 3: An example of a selected restaurant page on ‘yelp.com’

Relp

The website was initially scraped using the `rvest`(Wickham and RStudio 2016) package within the `tidyverse`(Wickham and RStudio 2017) package.

Front panel

The front page of the `Relp` tried to mimic the front page of the `yelp.com`. As illustrated in Figure 4, the page has an image of a delicious dish and a couple of widgets to enter inputs for the application. When the search button is clicked on the real `yelp` front page, the user is directed to a new web page with a list of businesses information. The realization of this feature was constructed on `Relp` by using the conditional panel with the “search”" button to initiate. That is, when the red search icon is clicked with appropriate “Find” and “Near” variable, the shiny application switches its panel to the main panel. The application starts scraping information from the `yelp.com` once the search button is pressed.



Figure 4: Front page of the Relp application

Main panel

The main panel of the `Relp` application consists of two different sections: the input section and the output section. The input section has multiple widgets and is located on the left column of the page. User of the `relp` application could change the “Find” and the “Near” variable, and also filter the price variable. New output could be generated by clicking the “refresh” button. The output section located on the right side of the page with the following three sub-panels: map, review, and image panels. An example of the main page is provided in Figure 5. The “Find” and the “Near” variable was changed from “Ames, restaurant” in Figure 5a to “San Francisco, Bars” in Figure 5b.



(a) Initial map panel of the Relp application

(b) Map panel of the Relp application with different variables

Figure 5: Example of the use of widget on the main page

Map

Location, price, and average rating are important factors when choosing the right business. These information could be obtain in the map panel. Visualization of the location information is reflected as a map using the `leaflet` package. Other information are organized into a data table by the `DT` package. Interaction between the data table and the map is also established. As illustrated in Figure 5b, when a row in the data table is highlighted, the corresponding marker on the map changes its color from blue to red.

Review

Crowd-sourced reviews and ratings often provide valuable indirect experiences of a business to the users reading the evaluation. The review panel displays phone number, hours of operation, and features of a particular business as well as the crowd-sourced reviews and ratings. An example of the review panel is given in Figure 6. When a row of the data table from the map panel is highlighted, the corresponding url of the business is retrieved. The web page of the url is then scraped and re-organized as a shiny web panel.

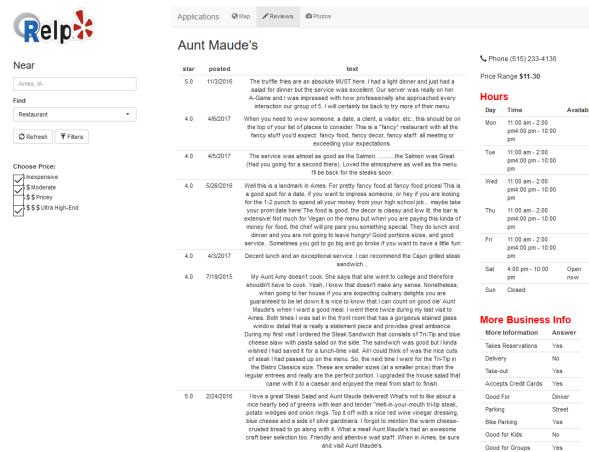


Figure 6: Review panel of the Relp application

Images

Pictures of a business is also an important factor when making decisions. The algorithm for generating the image panel is similar to that of the review panel. Image files of a business are first download in a temporary folder. Two action buttons are used to navigate the images. When a new business is selected, files in the temporary folder are deleted and new images are downloaded. Display of the image panel is portraited in Figure 7.

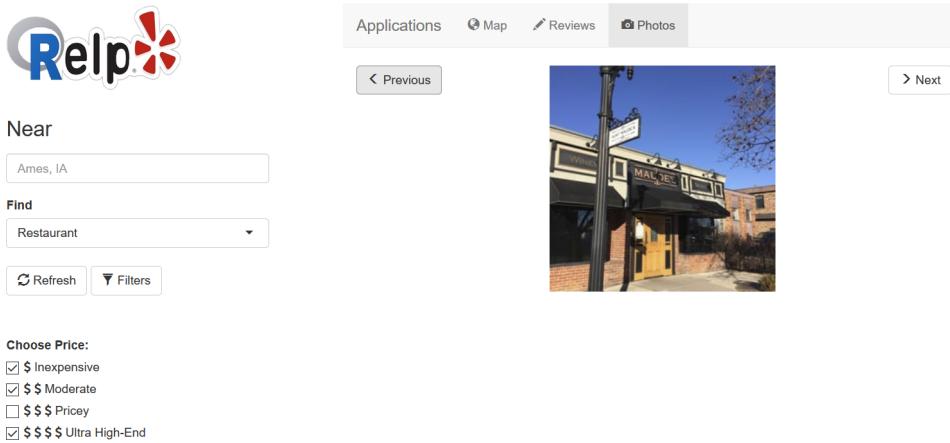


Figure 7: Image panel of the Relp application

Future work

Although much effort was put into the application to imitate that yelp.com, the application is not perfect. There are a few known bugs to be fixed, some yelp.com features are omitted, and the cosmetic part may be improved for better display.

Fix bugs

First, the application will crash when the Review and Image panel is selected with more than one business. This is because the script is written for only one selected business url. A simple statement (e.g. the `if` statement) may be applied to check if more than two business urls are chosen as an input, and generate a waring message when the condition is not met.

Second, a script should be updated to un-mark the selected marker on the leaflet output when clicked twice. Current script only marks the chosen input, and no code is written to reverse the selection. A leaflet proxy code could be applied to remove the marker when the business is clicked again.

Third, an error occurs when publishing the application onto the shinyapps.io. On the local machine, the application runs smooth without any issues regarding the `leaflet` package. However, uploading the application yields the following message:
`ERROR: could not find function "leafletOutput".` Apparently, similar issues were brought up a couple years ago (2015). One of the suggestion was to load the `leaflet` package on both `ui` and `server`, however, this is not a solution for the case. Other possible source of errors maybe due to the lack of version control or errors generated from `leafletProxy` and `htmlwidgets`. Further investigation should take place to fix this problem.

Add more features

Improve speed

Beautification

Conclusion

References

- Chang, Winston, and others. 2017. *Web Application Framework for R*. <https://cran.r-project.org/web/packages/shiny/shiny.pdf>.
- Cheng, Joe. 2017. *Create Interactive Web Maps with the Javascript 'Leaflet' Library*. <https://cran.r-project.org/web/packages/leaflet/leaflet.pdf>.
- Kahle, David, and Hadley Wickham. 2016. *Spatial Visualization with Ggplot2*. <https://cran.r-project.org/web/packages/ggmap/ggmap.pdf>.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Wickham, Hadley, and RStudio. 2016. *Easily Harvest (Scrape) Web Pages*. <https://cran.r-project.org/web/packages/rvest/rvest.pdf>.
- . 2017. *Easily Install and Load 'Tidyverse' Package*. <https://cran.r-project.org/web/packages/tidyverse/tidyverse.pdf>.
- Xie, Yihui. 2016. *A Wrapper of the Javascript Library 'Datatable'*. <https://cran.r-project.org/web/packages/DT/DT.pdf>.
- “Yelp.” 2017. <https://www.yelp.com/>.