**Zomato** is a popular restaurants listing website in India (Similar to Yelp) and People are always interested in seeing how to download or scrape Zomato Restaurants data for Data Science and Visualizations.

In this post, We’ll learn how to scrape / download Zomato Restaurants (Buffets) data using R. Also, hope this post would serve as a basic web scraping framework / guide for any such task of building a new dataset from internet using web scraping.

**Steps**

* Loading required packages
* Getting web page content
* Extract relevant attributes / data from the content
* Building the final dataframe (to be written as csv) or for further analysis

*Note: This post also assumes you’re familiar with Browser Devtools and CSS Selectors*

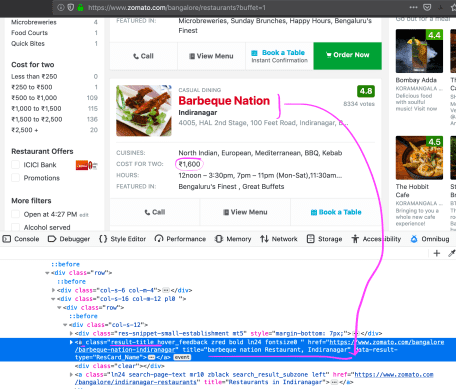
**Packages**

We’ll use the R-packages rvest for web scraping and tidyverse for Data Analysis and Visualization

**Loading the libraries**

library(rvest)

library(tidyverse)



zomato web scraping

**Getting Web Content from Zomato**

zom <- read\_html("<https://www.zomato.com/bangalore/restaurants?buffet=1>")

**Extracting relevant attributes**

Considering, It’s Restaurant listing – the columns that we can try to build are – Name of the Restaurant, Place / City where it’s, Average Price (or as Zomato says, Price for two)

**Name of the Restaurant**

This is how the html code for the name is placed:

[Barbeque Nation](https://www.zomato.com/bangalore/barbeque-nation-indiranagar)

So, what we need is for a tag with class value result-title, the value of attribute title.

zom %>% html\_nodes("a.result-title") %>%

html\_attr("title") %>%

stringr::str\_split(pattern = ',') -> listing

As a good thing for us, Zomato’s website is designed in such a way that the name and place of the Restaurant are within the same css selector a.result-title – so it’s one scraping. And it’s separated by a , so we can use str\_split() to split and the final output is now saved into listing which is a list.

**Converting List to Dataframe**

zom\_df <- do.call(rbind.data.frame, listing)

names(zom\_df) <- c("Name","Place")

In the above two lines, we’re trying to convert the listing list to a dataframe zom\_df and then rename the columns into Name and Place

**Extracting Price and Adding a New Price Column**

zom\_df$Price <- zom %>% html\_nodes("div.res-cost > span.pl0") %>%

html\_text() %>%

parse\_number()

Since the Price field is actually a combination of Indian Currency and Comma-separated Number (which is ultimately a character), we’ll use parse\_number() function remove the Indian currency unicode from the text and extract only the price value number.

**Dataset**

head(zom\_df)

## Name Place Price

## 1 abs absolute barbecues Restaurant Marathahalli 1600

## 2 big pitcher Restaurant [Old Airport Road 1800](https://www.google.com/maps/search/Old+Airport+Road++1800?entry=gmail&source=g)

## 3 pallet Restaurant Whitefield 1600

## 4 barbeque nation Restaurant Indiranagar 1600

## 5 black pearl Restaurant Marathahalli 1500

## 6 empire restaurant Restaurant Indiranagar 500

**Price Graph**

zom\_df %>%

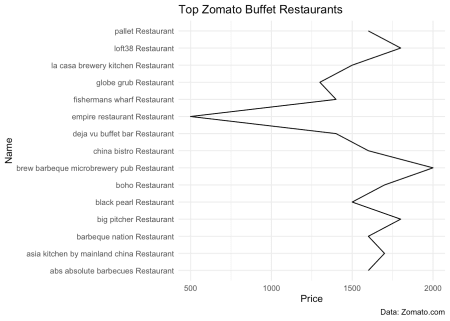
ggplot() + geom\_line(aes(Name,Price,group = 1)) +

theme\_minimal() +

coord\_flip() +

labs(title = "Top Zomato Buffet Restaurants",

caption = "Data: Zomato.com")



**Summary**

Thus, We’ve learnt how to build a new dataset by scraping web content and in this case, from Zomato to build a Price Graph.