zomato-mini-project.R

akhil

2021-08-31

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5 v purrr 0.3.4  
## v tibble 3.1.2 v dplyr 1.0.7  
## v tidyr 1.1.3 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(ggplot2)  
library(ggridges)  
library(plotly)

##   
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':  
##   
## last\_plot

## The following object is masked from 'package:stats':  
##   
## filter

## The following object is masked from 'package:graphics':  
##   
## layout

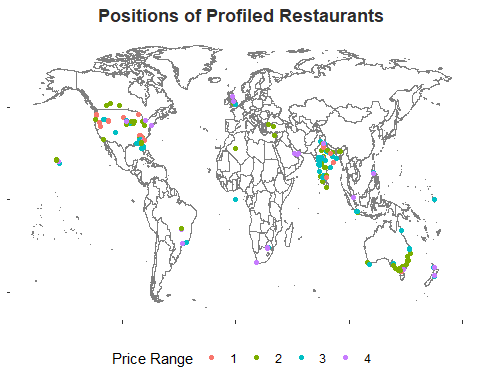
library(tidytext)  
  
  
zomato <- read.csv('zomato.csv')  
  
  
  
  
zomato %>%  
 str()

## 'data.frame': 9551 obs. of 21 variables:  
## $ Restaurant.ID : int 6317637 6304287 6300002 6318506 6314302 18189371 6300781 6301290 6300010 6314987 ...  
## $ Restaurant.Name : chr "Le Petit Souffle" "Izakaya Kikufuji" "Heat - Edsa Shangri-La" "Ooma" ...  
## $ Country.Code : int 162 162 162 162 162 162 162 162 162 162 ...  
## $ City : chr "Makati City" "Makati City" "Mandaluyong City" "Mandaluyong City" ...  
## $ Address : chr "Third Floor, Century City Mall, Kalayaan Avenue, Poblacion, Makati City" "Little Tokyo, 2277 Chino Roces Avenue, Legaspi Village, Makati City" "Edsa Shangri-La, 1 Garden Way, Ortigas, Mandaluyong City" "Third Floor, Mega Fashion Hall, SM Megamall, Ortigas, Mandaluyong City" ...  
## $ Locality : chr "Century City Mall, Poblacion, Makati City" "Little Tokyo, Legaspi Village, Makati City" "Edsa Shangri-La, Ortigas, Mandaluyong City" "SM Megamall, Ortigas, Mandaluyong City" ...  
## $ Locality.Verbose : chr "Century City Mall, Poblacion, Makati City, Makati City" "Little Tokyo, Legaspi Village, Makati City, Makati City" "Edsa Shangri-La, Ortigas, Mandaluyong City, Mandaluyong City" "SM Megamall, Ortigas, Mandaluyong City, Mandaluyong City" ...  
## $ Longitude : num 121 121 121 121 121 ...  
## $ Latitude : num 14.6 14.6 14.6 14.6 14.6 ...  
## $ Cuisines : chr "French, Japanese, Desserts" "Japanese" "Seafood, Asian, Filipino, Indian" "Japanese, Sushi" ...  
## $ Average.Cost.for.two: int 1100 1200 4000 1500 1500 1000 2000 2000 6000 1100 ...  
## $ Currency : chr "Botswana Pula(P)" "Botswana Pula(P)" "Botswana Pula(P)" "Botswana Pula(P)" ...  
## $ Has.Table.booking : chr "Yes" "Yes" "Yes" "No" ...  
## $ Has.Online.delivery : chr "No" "No" "No" "No" ...  
## $ Is.delivering.now : chr "No" "No" "No" "No" ...  
## $ Switch.to.order.menu: chr "No" "No" "No" "No" ...  
## $ Price.range : int 3 3 4 4 4 3 4 4 4 3 ...  
## $ Aggregate.rating : num 4.8 4.5 4.4 4.9 4.8 4.4 4 4.2 4.9 4.8 ...  
## $ Rating.color : chr "Dark Green" "Dark Green" "Green" "Dark Green" ...  
## $ Rating.text : chr "Excellent" "Excellent" "Very Good" "Excellent" ...  
## $ Votes : int 314 591 270 365 229 336 520 677 621 532 ...

WorldData <- map\_data('world')  
  
  
WorldData %>% filter(region != "Antarctica") -> WorldData  
WorldData <- fortify(WorldData)  
  
zomato\_world <- zomato %>% select(Latitude,Longitude,Price.range) %>% mutate(Price.range=as.factor(Price.range))  
  
  
p <- ggplot()  
p <- p + geom\_map(data=WorldData, map=WorldData,  
 aes(x=long, y=lat, group=group, map\_id=region),  
 fill="White", colour="#7f7f7f", size=0.5)

## Warning: Ignoring unknown aesthetics: x, y

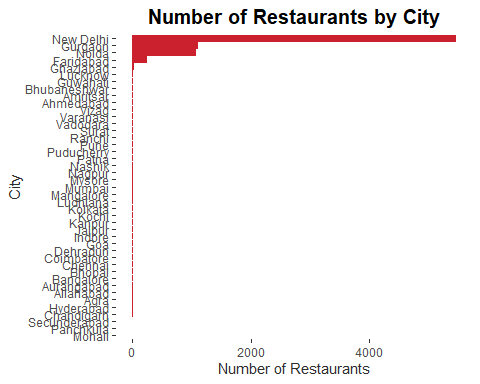
p <- p + geom\_point(data=zomato\_world,aes(x=Longitude,y=Latitude,color=Price.range))  
  
p+theme(panel.background = element\_blank(),axis.title = element\_blank(),  
 axis.text=element\_blank(),legend.position = 'bottom',legend.background = element\_blank(),  
 plot.title = element\_text(hjust=0.5,face='bold',color='#2d2d2d'),  
 legend.key = element\_rect(fill='white',color = 'transparent'))+labs(title='Positions of Profiled Restaurants',  
 color='Price Range')



zomato %>%  
 filter(Country.Code==1) %>%  
 select(Restaurant.ID) %>%  
 unique() %>%  
 nrow()

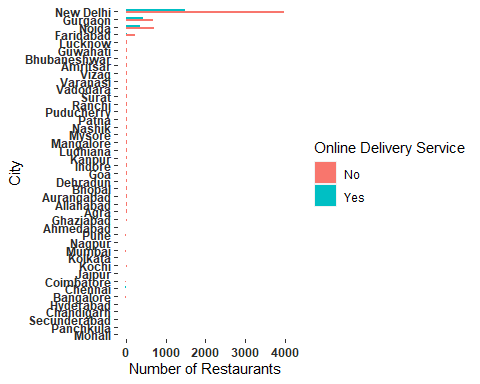
## [1] 8652

zomato %>% filter(Country.Code==1) %>%  
 select(Restaurant.ID,City) %>%  
 unique() %>%  
 group\_by(City) %>%  
 summarise(n=n()) %>%  
 ggplot(aes(x=reorder(City,n),y=n))+geom\_bar(stat = 'identity',fill='#cb202d')+  
 coord\_flip()+  
 theme(panel.background = element\_blank(),  
 strip.background = element\_blank(),  
 axis.title = element\_text(color = '#2d2d2d'),  
 strip.text.x = element\_text(color='#2d2d2d',face='bold',size=10),  
 plot.title = element\_text(hjust=0.5,face='bold',size=15))+  
 labs(x='City',y='Number of Restaurants',title="Number of Restaurants by City")



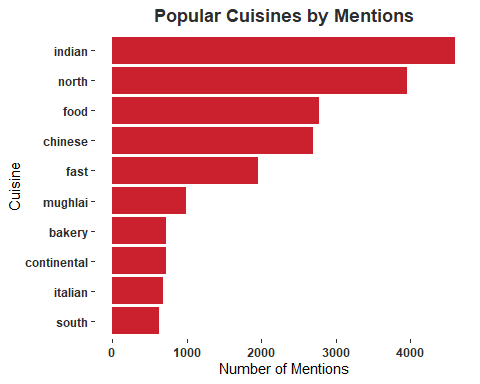
zomato %>%  
 filter(Country.Code==1) %>%  
 select(Restaurant.ID,Has.Online.delivery,City) %>%  
 unique() %>%  
 group\_by(City,Has.Online.delivery) %>%  
 summarise(n=n()) %>%  
 ungroup() %>%  
 rename(`Online Delivery Service`=Has.Online.delivery) %>%  
 ggplot(aes(x=reorder(City,n),y=n,fill=`Online Delivery Service`))+  
 geom\_bar(stat='identity',position = 'dodge',width = 0.5)+  
 labs(x='City',y='Number of Restaurants')+coord\_flip()+  
 theme(plot.title = element\_text(hjust=0.5,face='bold',color='#2d2d2d'),panel.background = element\_blank(),  
 strip.text = element\_text(face='bold',color='#2d2d2d'),axis.text.x = element\_text(face='bold',color='#2d2d2d'),  
 axis.text = element\_text(face='bold',color='#2d2d2d'))

## `summarise()` has grouped output by 'City'. You can override using the `.groups` argument.



library(Rcpp)  
  
zomato %>%  
 filter(Country.Code==1) %>%  
 select(Restaurant.ID,Cuisines,Average.Cost.for.two) %>%  
 unique() %>%  
 mutate(Cuisines=as.character(Cuisines)) %>%  
 unnest\_tokens(ngram,Cuisines,token='ngrams',n=1) %>%  
 group\_by(ngram) %>%  
 summarise(n=n()) %>%  
 arrange(desc(n)) %>%  
 top\_n(10) %>%  
 ggplot(aes(x=reorder(ngram,n),y=n))+geom\_bar(stat='identity',fill='#cb202d')+  
 theme(plot.title = element\_text(hjust=0.5,face='bold',color='#2d2d2d'),panel.background = element\_blank(),  
 strip.text = element\_text(face='bold',color='#2d2d2d'),axis.text.x = element\_text(face='bold',color='#2d2d2d'),  
 axis.text = element\_text(face='bold',color='#2d2d2d'))+  
 coord\_flip()+labs(x='Cuisine',y='Number of Mentions',title='Popular Cuisines by Mentions')

## Selecting by n



#Conclustion  
#A majority of restaurants are located in India  
#A majority of Indian restaurants are located in New Delhi  
#A majority of Indian restaurants do not have online delivery