Shubhendu Vimal – 11915067

Anmol More - 11915043

Sentiment words analysis

if (!require(tidytext)) {install.packages("tidytext")}

## Loading required package: tidytext

if (!require(ggplot2)) {install.packages("ggplot2")}

## Loading required package: ggplot2

if (!require(dplyr)) {install.packages("dplyr")}

## Loading required package: dplyr

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

read zomato and swiggy files

zomato <- readLines('data/zomato.txt')  
swiggy <- readLines("data/swiggy.txt")

function to return bing sentiment words

get\_positive\_negative\_word\_counts <- function(corpus) {  
 df = data\_frame(text = zomato) #create dataframe from corpus  
 bing\_sentiments = get\_sentiments("bing")  
   
 word\_counts <- df %>%  
 unnest\_tokens(word, text) %>%  
 inner\_join(bing\_sentiments) %>%  
 count(word, sentiment, sort = TRUE) %>%  
 ungroup()  
   
 return (word\_counts)  
}

bing sentiments add positive and negative scores Based on those score contribution

swiggy\_word\_counts <- get\_positive\_negative\_word\_counts(swiggy)

## Warning: `data\_frame()` is deprecated, use `tibble()`.  
## This warning is displayed once per session.

## Joining, by = "word"

swiggy\_word\_counts %>%  
 filter(n > 100) %>%  
 mutate(n = ifelse(sentiment == "negative", -n, n)) %>%  
 mutate(word = reorder(word, n)) %>%  
 ggplot(aes(word, n, fill = sentiment)) +  
 geom\_bar(stat = "identity") +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 xlab("Words") +  
 ylab("Count of words") +  
 ggtitle("Words contributing towards Swiggy sentiment")

zomato\_word\_counts <- get\_positive\_negative\_word\_counts(zomato)

## Joining, by = "word"

zomato\_word\_counts %>%  
 filter(n > 100) %>%  
 mutate(n = ifelse(sentiment == "negative", -n, n)) %>%  
 mutate(word = reorder(word, n)) %>%  
 ggplot(aes(word, n, fill = sentiment)) +  
 geom\_bar(stat = "identity") +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 xlab("Words") +  
 ylab("Count of words") +  
 ggtitle("Words contributing towards Zomato sentiment")