

MATH513 Practical

Using Twitter data to gain business advantage in the electric
and fuel car industry

Group: 36

Members: 10649798, 10654115, 10775412, 10777441

December 06, 2022

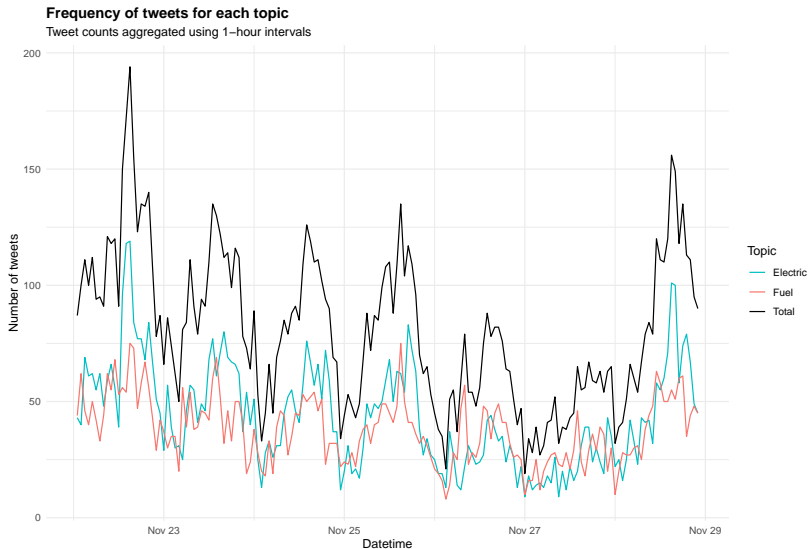
Introduction

- ▶ The innovation of electric cars over the years seems to gain more market recognition globally.
- ▶ There are so many advantages of electric cars as they try to curb environmental deterioration.
- ▶ Fuel Cars manufacturer seems to be overwhelmed with the fast pace of Electric Cars growth in the past few years.
- ▶ This presentation aims to compare and contrast the business insight of electric cars and fuel cars. We shall collect and analyse data from twitter to gain insight on the best marketing strategies to be deployed.

Methodology

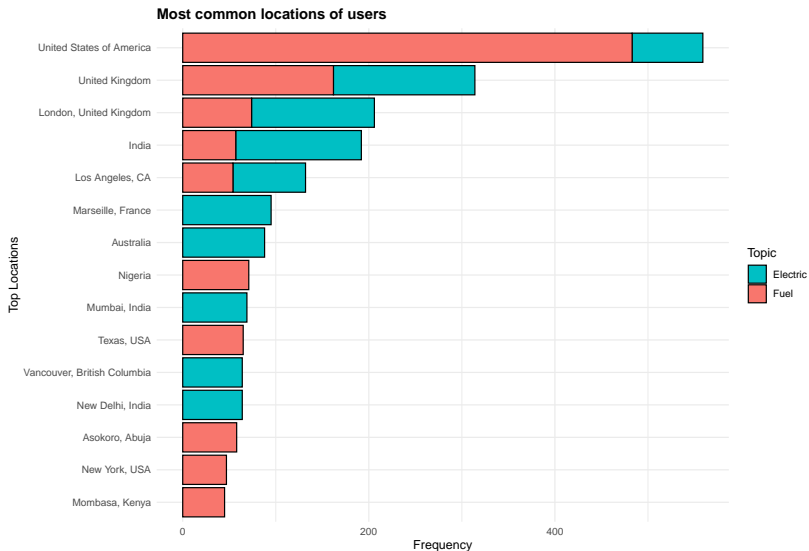
- ▶ This presentation was created using R markdown in a RStudio IDE, using R programming language.
- ▶ The process began by searching tweets about electric and fuel cars, using the twitter API. However, as the 'free' twitter API is limited to only getting the data from the past 7 days, the data were far from representing the whole customer base.
- ▶ Useful customer data was then extracted from the tweets, by making use of some basic R data manipulation techniques.
- ▶ The data was then visualized by creating graphs, to make the understanding of the data easier
- ▶ All the packages and tools that were used in the process are documented at the end of the presentation.

Tweet frequency for each topic



Source: Data collected from Twitter's REST API via rtweet

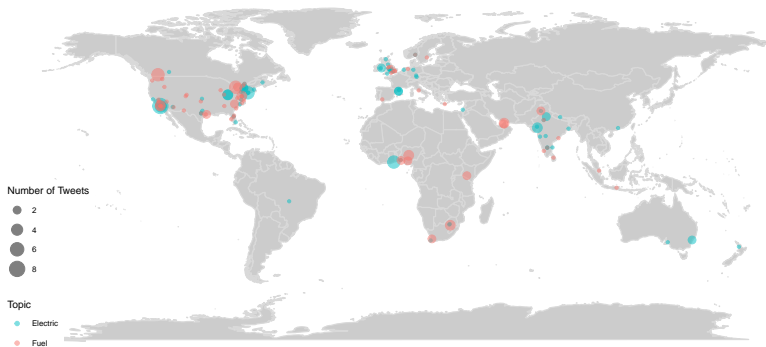
Location of users



Source: Data collected from Twitter's REST API via rtweet

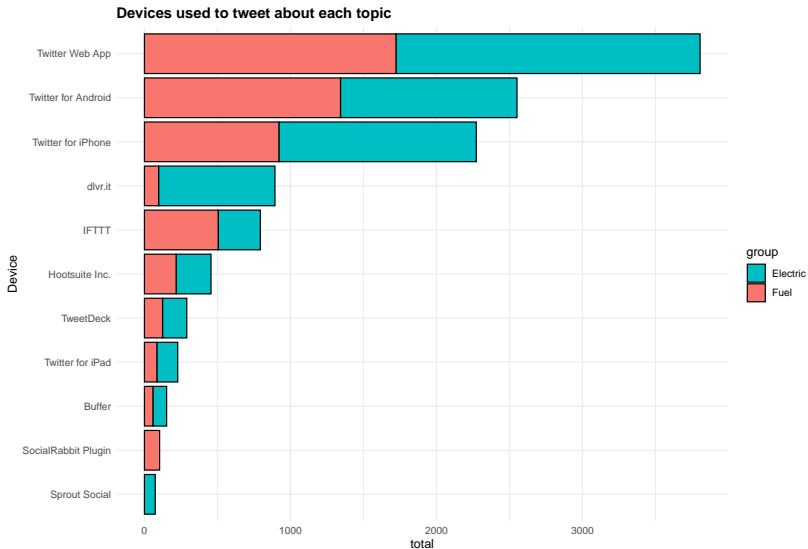
Where people are tweeting from

Locations of tweets for each topic



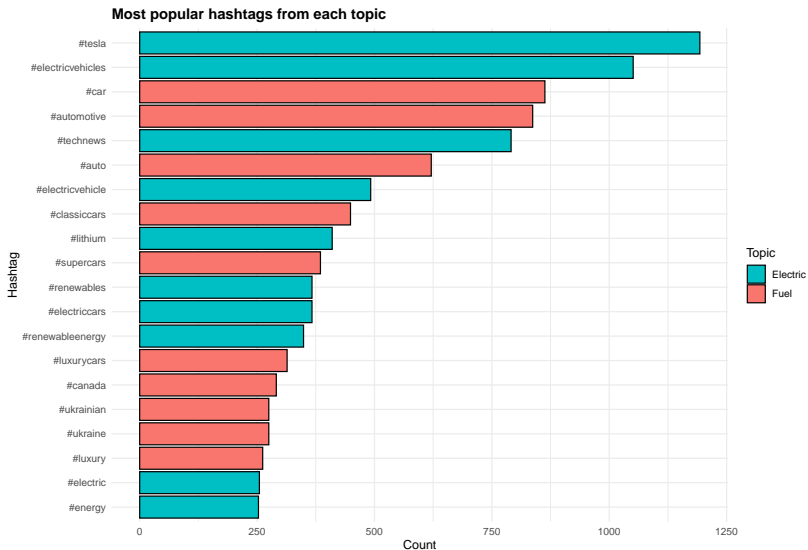
Source: Data collected from Twitter's REST API via rtweet

Devices used



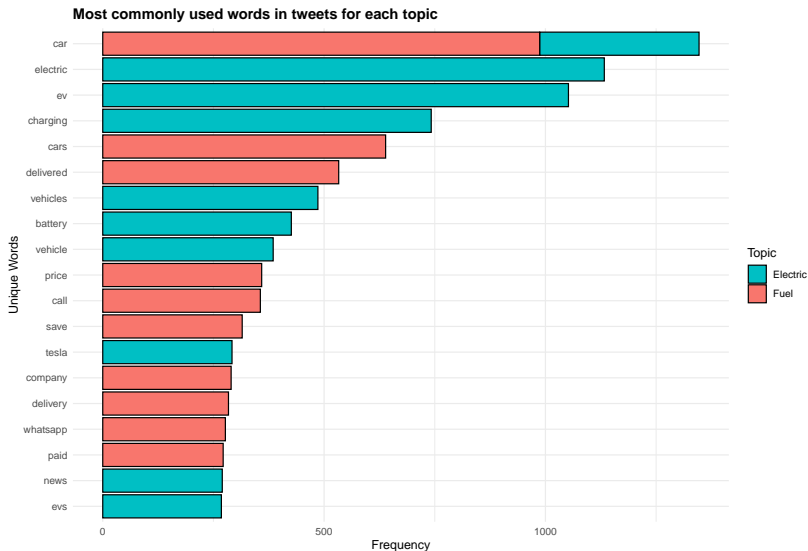
Source: Data collected from Twitter's REST API via rtweet

Popular hashtags



Source: Data collected from Twitter's REST API via rtweet

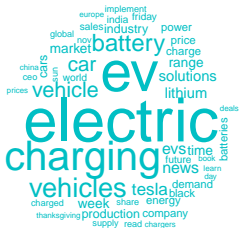
Most commonly used words



Source: Data collected from Twitter's REST API via rtweet

Wordclouds

Wordcloud for Electric cars



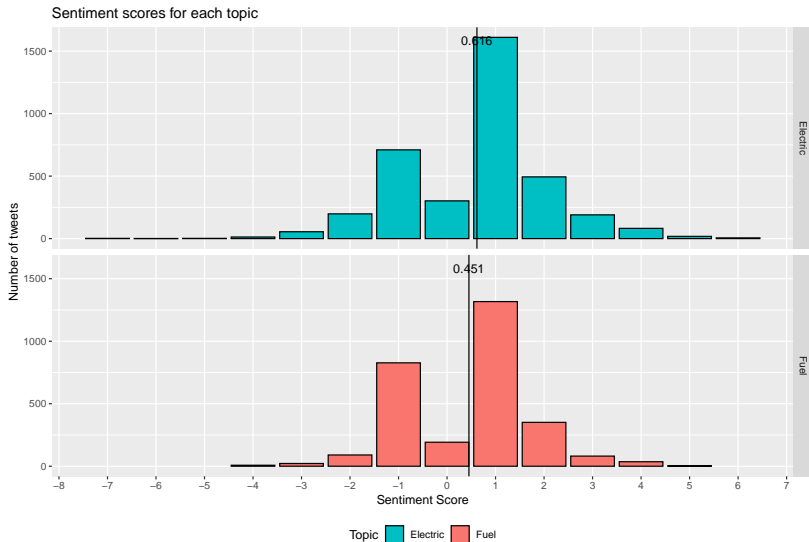
Source: Data collected from Twitter's REST API via rtweet

Wordcloud for Fuel cars



Source: Data collected from Twitter's REST API via rtweet

Sentiment scores



Source: Data collected from Twitter's REST API via rtweet

Statistical test

- ▶ To see whether there is a significant difference between the sentiment scores, we can conduct a statistical test.
- ▶ To check whether we can use a t-test we firstly need to check if the data results produce normal, bell-shaped distribution. We can do that by conducting a normality test, such as Shapiro-Wilk test.
- ▶ The test results of Shapiro-Wilk test are as follows:
 - ▶ Electric cars p-value: $7.5183629 \times 10^{-40}$
 - ▶ Fuel cars p-value: $1.8516033 \times 10^{-41}$
- ▶ Since the p-value is under 0.05, a t-test wouldn't be appropriate. Therefore we will use Mann-Whitney-Wilcoxon test. which yielded the following results:
 - ▶ p-value: 7.5701715×10^{-7}
- ▶ As the p-value is under 0.05, we can say that there is a significant difference between the sentiment scores.

Conclusion

- ▶ The various outcomes of our data analysis, data visualization and sentiment analysis provides us with useful insight on how to further expand the business of electric cars globally.
- ▶ The highest frequency of tweets occurred at 2PM UTC indicating the best time to reach our customers on the twitter platform.
- ▶ The location of users from our data visualization provided us with useful information on how Electric cars manufacturers should expand its business in the United States of America and in the United Kingdom.

Conclusion - continued

- ▶ The electric car manufacturers could reach out to their customers mostly by using Twitter Web App.
- ▶ The tweets which used popular hashtags and words gain more impressions, therefore very useful for digital marketing.
- ▶ From the sentiment score, tweets about electric cars are more positive than fuel cars. This indicates that the electric cars are well accepted from English speaking countries.
- ▶ Based on our analysis, we advise the manufacturers of fuel cars to fully integrate the technology and innovation of electric cars as positive sentiment scores favored electric cars.

Citations

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. (R version 4.2.1)

Kearney, M. W. (2019). rtweet: Collecting and analyzing Twitter data, Journal of Open Source Software, 4, 42. 1829. doi:10.21105/joss.01829 (R package version 0.7.0)

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wickham H, François R, Henry L, Müller K (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.10, <https://CRAN.R-project.org/package=dplyr>.

Arnold J (2021). ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. R package version 4.2.4, <https://CRAN.R-project.org/package=ggthemes>.

Jeroen Ooms (2014). The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. arXiv:1403.2805 [stat.CO] URL <https://arxiv.org/abs/1403.2805>.

Silge J, Robinson D (2016). "tidytext: Text Mining and Analysis Using Tidy Data Principles in R." JOSS, 1(3). doi:10.21105/joss.00037 <https://doi.org/10.21105/joss.00037>, <http://dx.doi.org/10.21105/joss.00037>.

Fellows I (2018). wordcloud: Word Clouds. R package version 2.6, <https://CRAN.R-project.org/package=wordcloud>.

Wickham H, Girlich M (2022). tidyr: Tidy Messy Data. R package version 1.2.1, <https://CRAN.R-project.org/package=tidyr>.

Cheng J, Chang W (2022). httpuv: HTTP and WebSocket Server Library. R package version 1.6.6, <https://CRAN.R-project.org/package=httpuv>.

Wickham H (2022). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.1, <https://CRAN.R-project.org/package=stringr>.

Wickham H, Hester J, Bryan J (2022). readr: Read Rectangular Text Data. R package version 2.1.3, <https://CRAN.R-project.org/package=readr>.