

An Exploratory Data Analysis of Uber Based on Twitter

Presented by Aksam Ahmad, Siyao Chang, Fengshi Niu, Chenyu Wang, and Xinyue Zhou



Introduction

In our analysis, we look at tweets about Uber as a way of understanding the public mindset about Uber and its services. Our exploratory analysis begins with answering certain questions, such as when, where and what people tweeted about Uber. Along the way, we do further analysis on the negative tweets. Finally, tweets pertaining to surge pricing and tweets pertaining to its new logo are decomposed further.

Data Collection and Mining

- English tweets containing #Uber from January 20, 2016 to February 13, 2016
- Subset tweets with a UTC-offset value and converted to local time
- Tokenized tweets and removed stop and auxiliary words

Most Frequently Tweeted Words With Uber

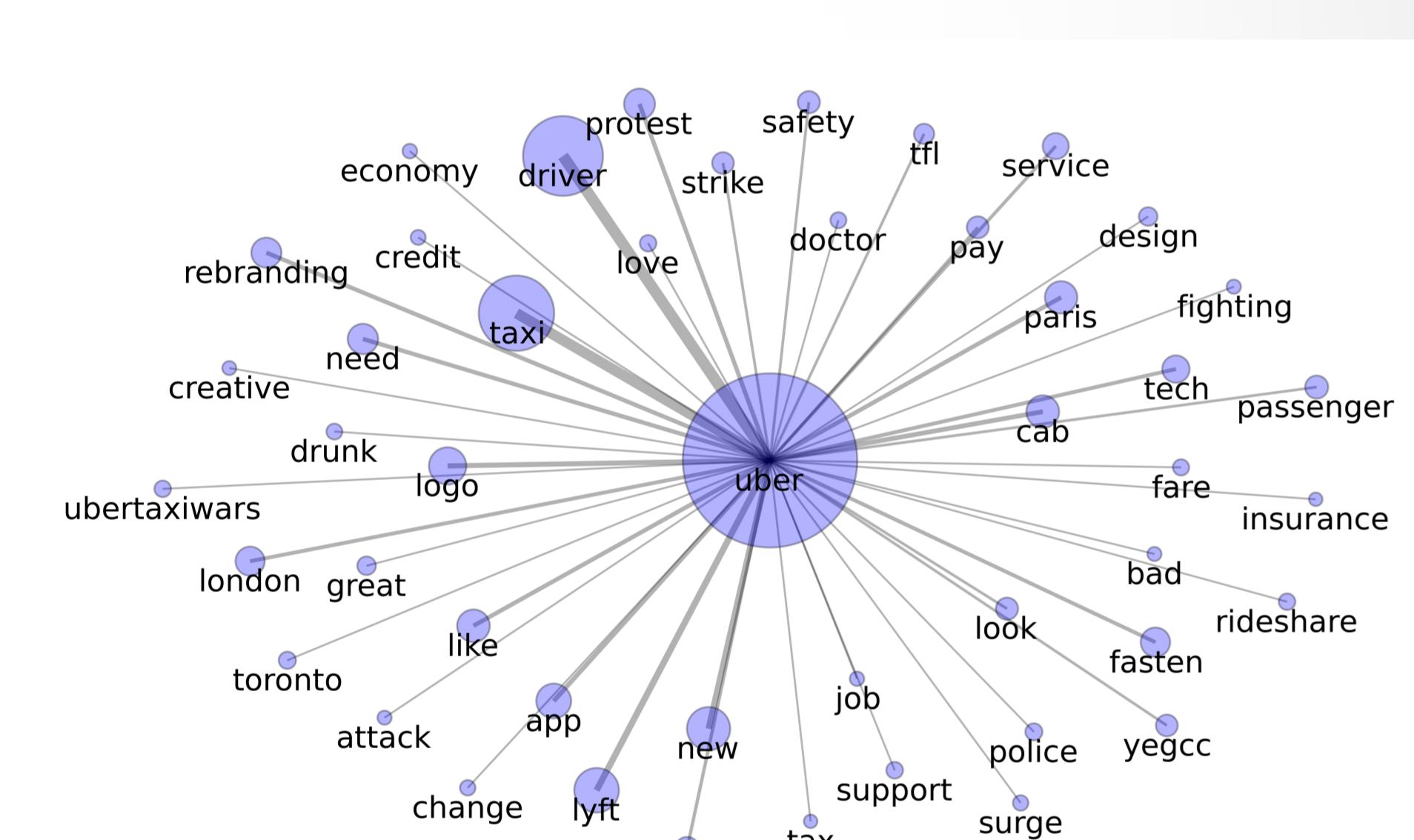


Figure 1: Top Frequency Words Network: Size of the nodes is proportional to their appearing frequencies.

Top Words Related to Uber							
Top Words	driver	taxi	lyft	new	logo	app	cab
Frequency	17828	15861	5572	5237	3909	3400	2992

Table 1: This table shows the most frequently tweeted words with Uber.

Top Words Related to Uber by Group	
Top Topic	Related words
taxi (19645)	taxi, cab, ubertaxiwiars
logo change (13832)	change, design, logo, look, rebranding, new
strike (6978)	paris, protest, strike
drunk doctor (2007)	doctor, drunk, attack
surge pricing (678)	surge

Table 2: This table gives context to what the most frequently tweeted words with Uber were related to.

The word “like” is mainly used in a negative context. (Detailed explanations later in the poster.) From the network diagram and these charts, the top topics related to Uber are **taxi**, **logo change**, **strikes**, **drunk doctor** and **surge pricing**.

Where Do People Tweet about Uber in the U.S.?

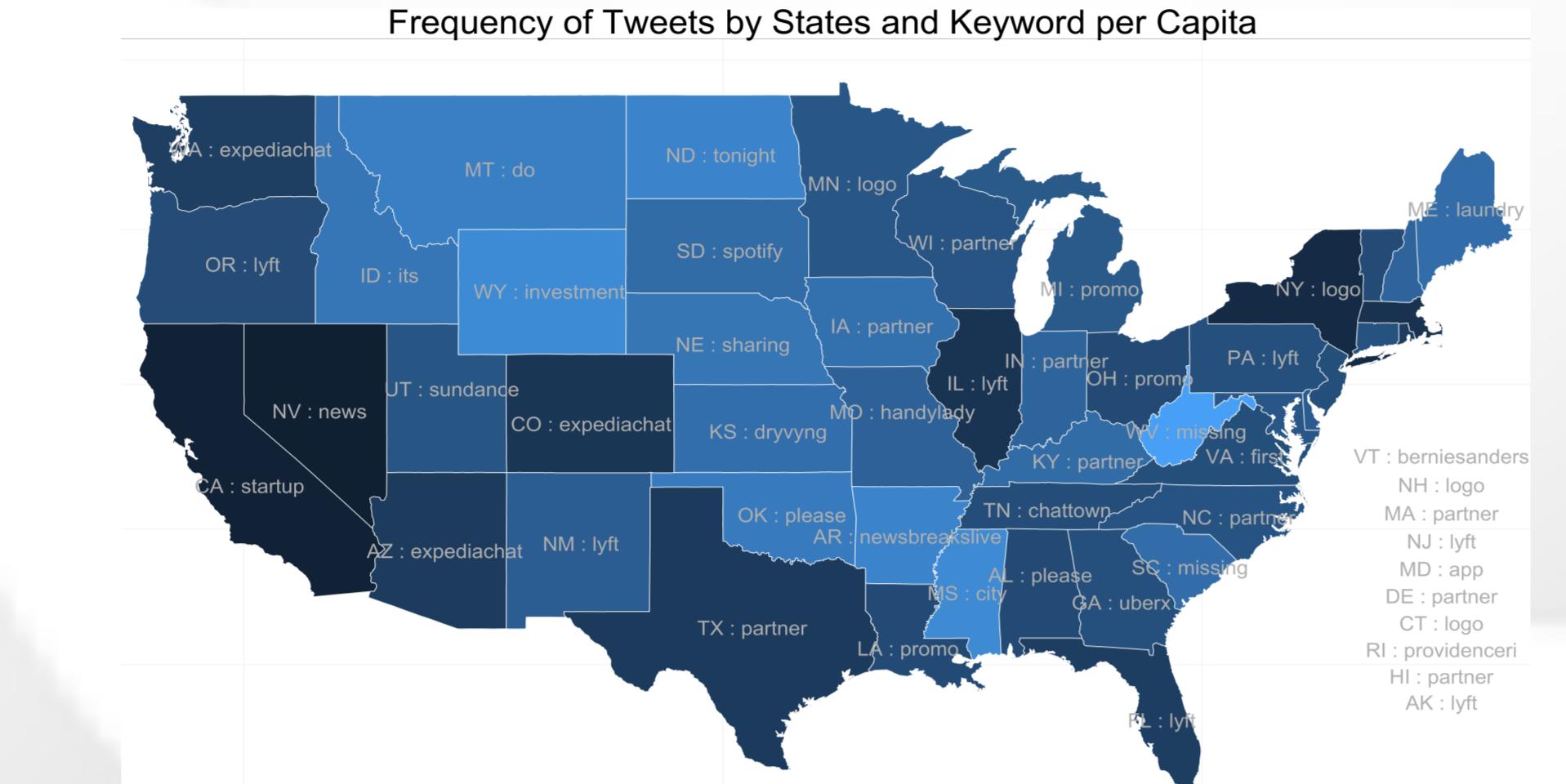


Figure 2: In this heatmap, darker color implies more tweets about Uber. The text in each state is the most frequently tweeted word for the state.

The frequency level in the **heatmap** is computed from the number of tweets about Uber in each state after normalizing by corresponding population. The map shows that California, Nevada and New York are the top three most active states in the U.S.

Top Used Words by State					
State	Keyword	Freq	State	Keyword	Freq
CA	startup	965	NV	news	69
TX	partner	221	FL	lyft	67
NY	logo	143	IL	lyft	58
HI	partner	106	MA	partner	54

Table 3: This table shows the most frequently used words in the top 8 most active states. The U.S. data are a subset of the larger global dataset.

- **startup:** Californians love to talk about startups with Uber, as Uber is one of the fastest growing startups worldwide.
- **logo:** New Yorkers seem to be very passionate about new Uber logo.
- **Lyft:** People in Florida and Illinois tend to juxtapose Lyft and Uber, as these companies have similar business and are very active in both states.
- **partner:** Uber has an expanding business in Texas and Massachusetts, resulting in a lot of advertisements about “how to be a Uber **partner**”. “Uber partner” has the same meaning as “Uber driver”.

Negative Tweets About Uber

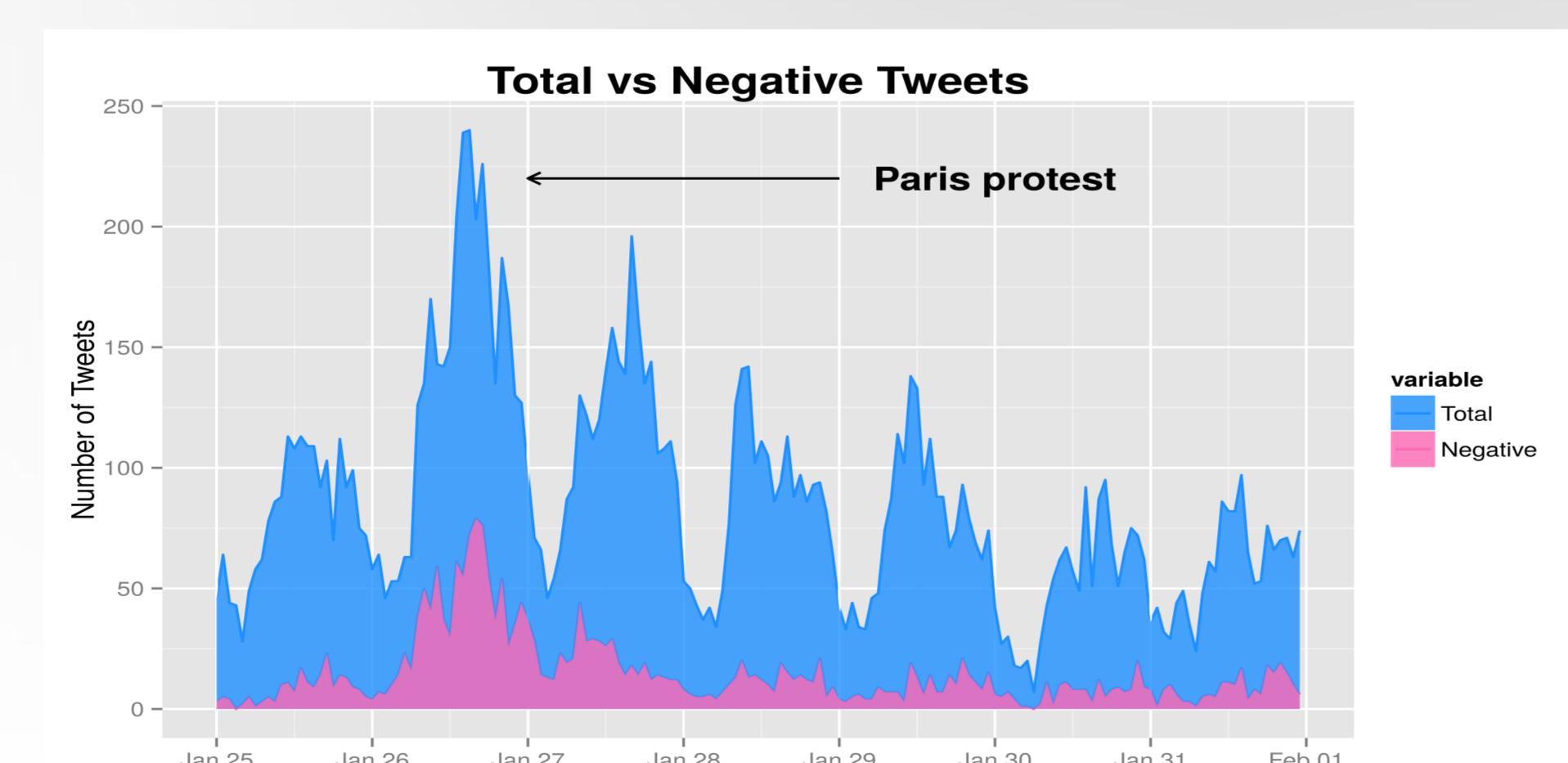


Figure 3: This is a one-week time series plot, presenting the time and the amount people tweet about Uber. The proportion of negative tweets is depicted in pink.

In general, people tweet less about Uber on weekends. Most tweets happen around rush hour or party times. The amount of negative tweets is generally proportional to the amount of total tweets.

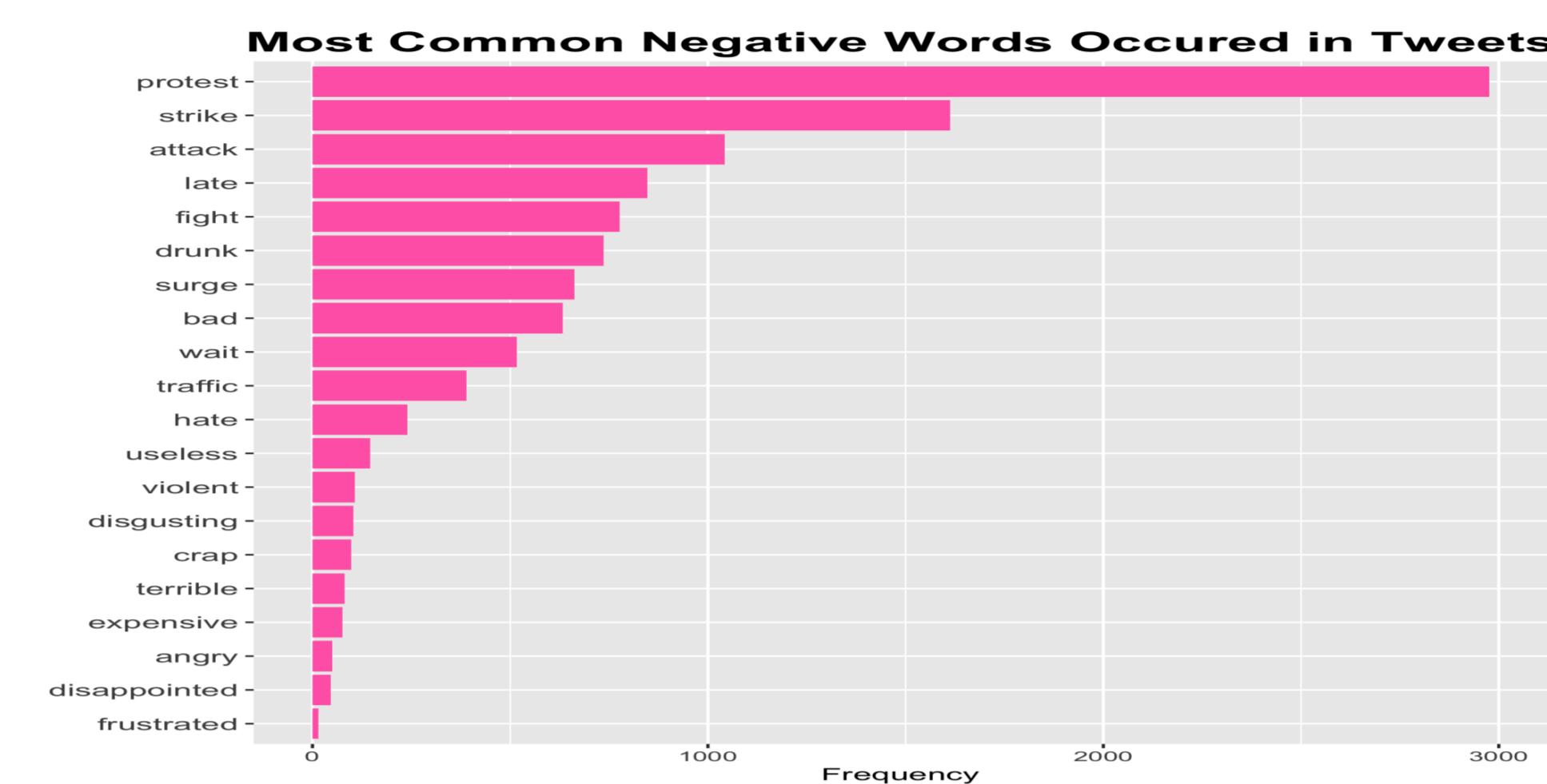


Figure 4: This shows the most common negative words occurred in tweets. X-axis represents the true value of frequency. As shown above, protest and strike have significant high frequency.

When does Surge Pricing Occur?

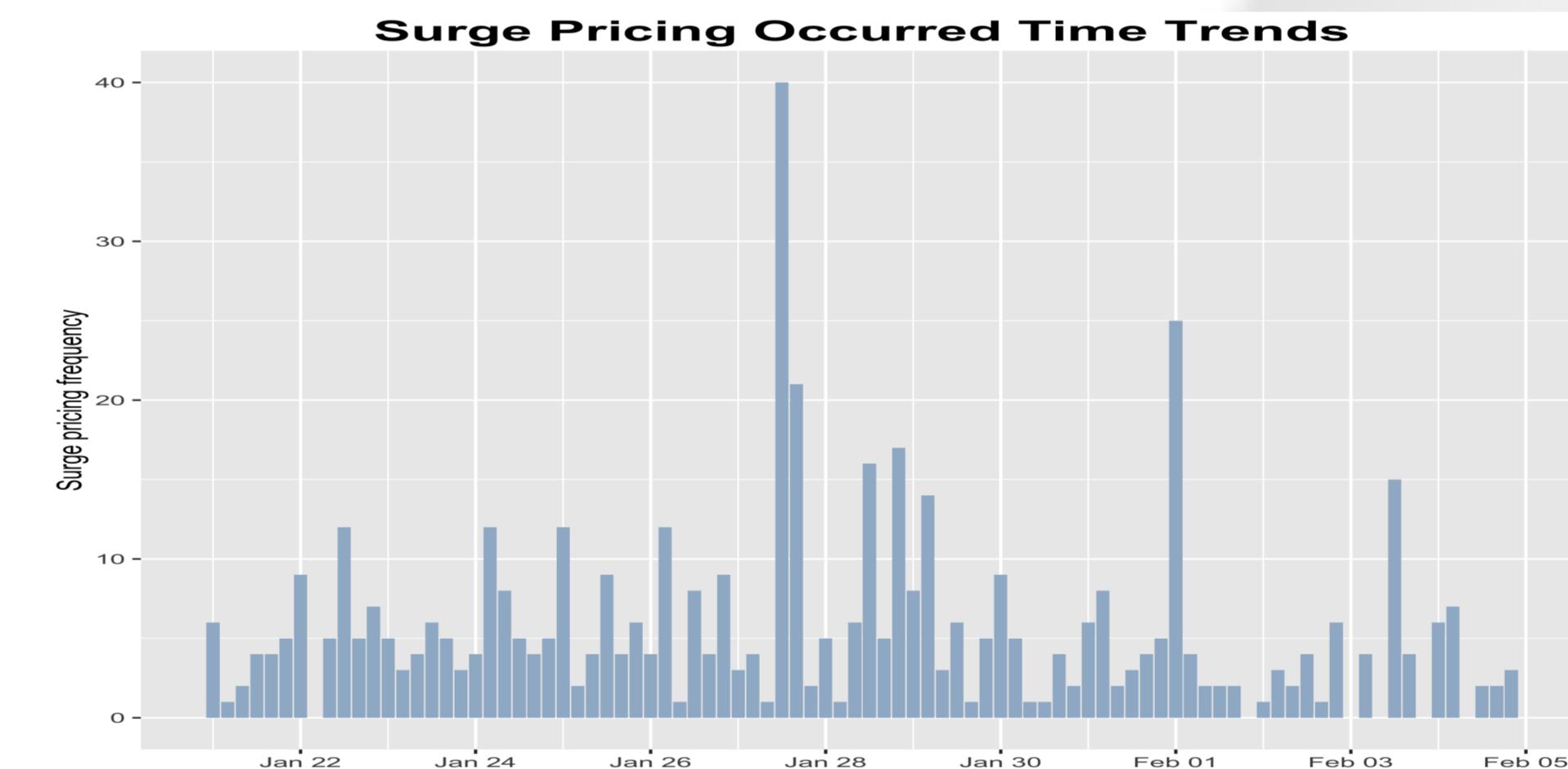


Figure 5: This is a time-series plot for two weeks from Jan 21st to Feb 4th. Each bar represents a 4-hour period. It shows the frequency of surge pricing occurred per bin.

Although there are certain spikes in the surge-pricing, we do not see any discernible patterns. We further explore the patterns of surge-pricing in the the following plot.

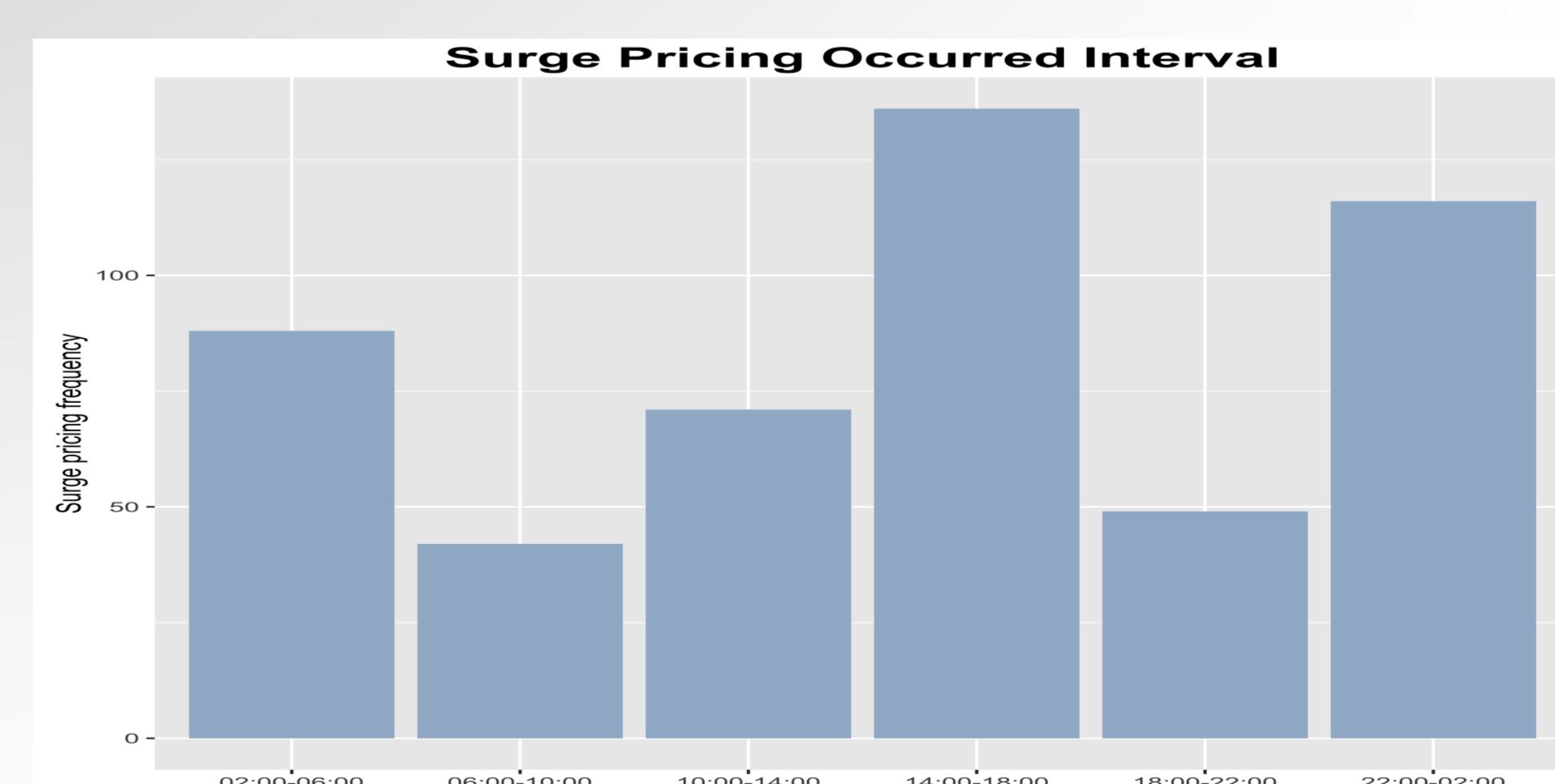


Figure 6: Each bar in this plot represents the frequency of surge-pricing occurring in each 4-hour period. The frequency is calculated based on the previous plot by summing all the same four-hour periods in two weeks together. It indicates that surge pricing tends to occur during 14:00-18:00 (rush hour) and 22:00-2:00 (party times).

Uber Changes Its Logo

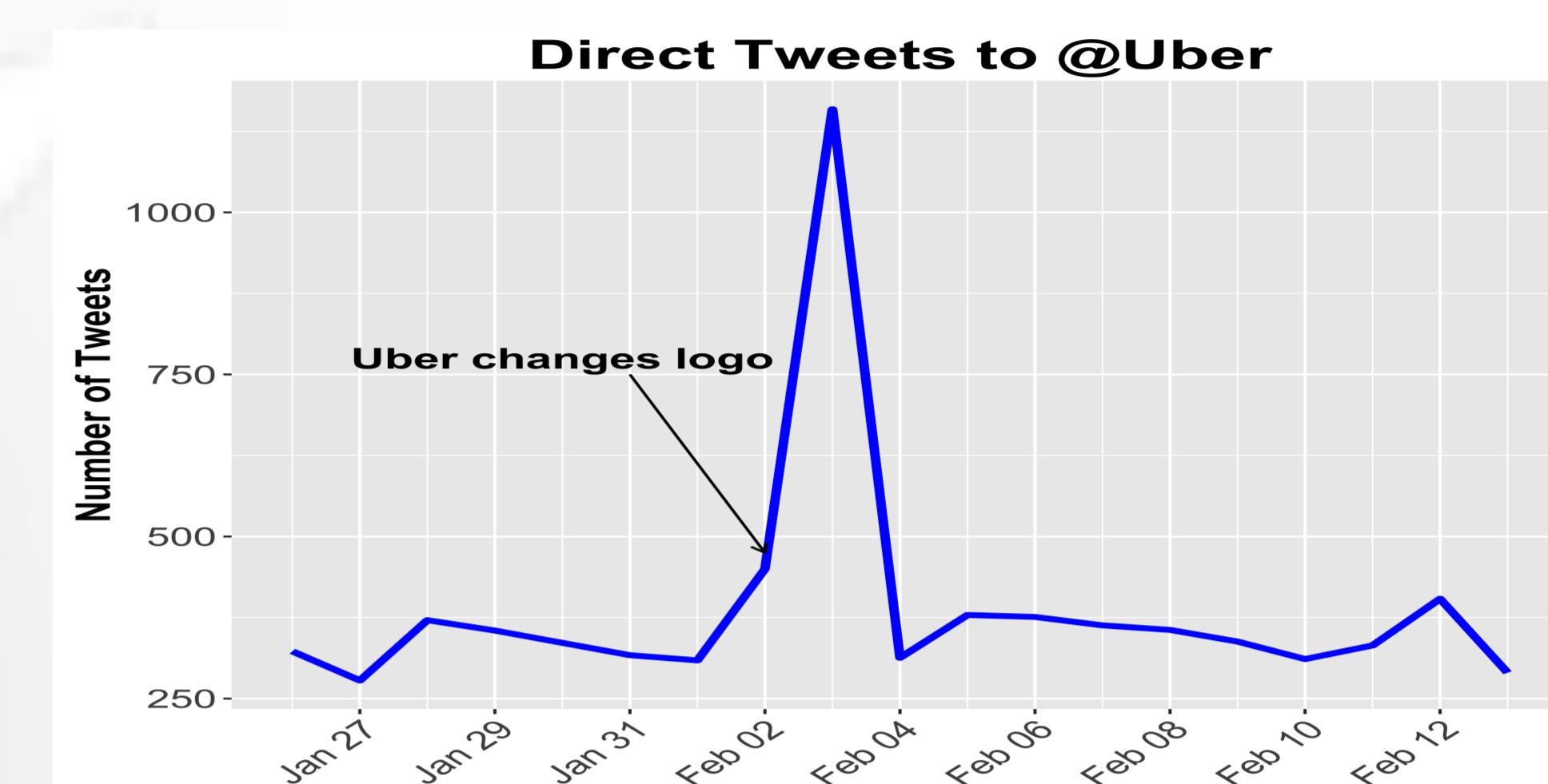


Figure 7: Upon announcement of Uber's new logo on February 2nd, 2016, the number of tweets directed to @Uber rose dramatically.



Figure 8: Most used word is ‘like’.

tall new sure is a bit bizarre i like older logo it was more distincti rolled out the new logo does n't like making logos he just likes to st d take credit for making it look like crap open letter to uber the new ld have thought not even drivers like the new uber logo the logo and t need to you get the uber logo i like upgrades i like new logos but ne the uber logo i like upgrades i like new logos but new logo is primit a toilet seat the new logo looks like an ancient chinese coin does thi logo is so ugly it 's making me like cabs more the new logo is so ug logo is so ugly it 's making me like cabs more just completely change logo is so ugly it 's making me like cabs more the new logo is so ug logo is so ugly it 's making me like cabs more this is a downgrade it ore this is a downgrade it looks like a microchip manufacturer the new logo is so ugly it 's making me like cabs more rolls out new look and s still an app icon which do you like most which do you like better th

The most frequently used word in tweets pertaining to the logo change was “like”, as shown above. Contrary to intuition, “like” was used in a negative context most of the time, as evidenced by the sample tweets above.

Conclusions

- California, New York, and Nevada have the most tweets per person.
- The most frequent negative words are protest, strike, and attack. These are driven by Paris Uber protest.
- Negative tweets represent a consistent proportion of total tweets.
- The public had a generally negative but short-term reaction to Uber's new logo.

Github: <https://github.com/Chenyu-Renee/UberAnalysis>