



Group 25

Bogdan Ene
Dusko Trajkov
Gloria Boekhouder
Nicola Pieruzzini
Thodoris Zois



RESEARCH

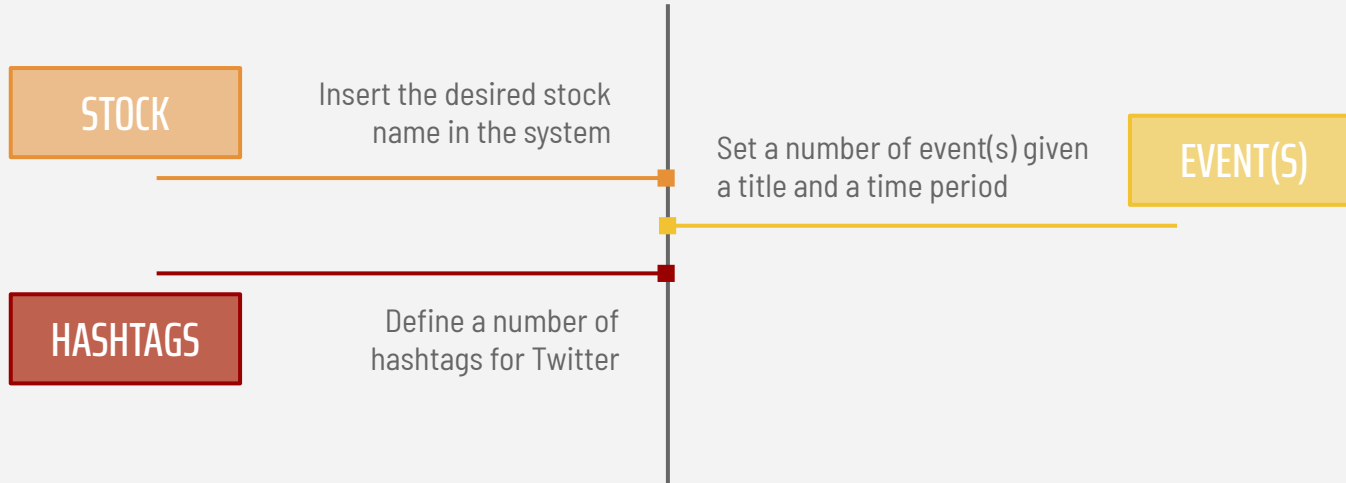
- What is the impact that a certain brand-related event has on the brand itself in terms of consumers' opinions and which is the correlation between these latter and the stock prices?

- How can investors investigate whether the public opinion affects a company that they are interested to invest in?



WEB-BASED ANALYTICS TOOL

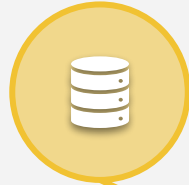
INPUT



PIPELINE

1. DATA GATHERING

Twitter data
Stock prices



2. DATA CLEANING

No location
Unknown location
Duplicate tweets (same user)



3. MACHINE LEARNING

Sentiment Analysis
Topic Analysis (LDA)



4. ANALYTICS APP

Interactive graphs



00

OVERALL STATISTICS

01

CORRELATION

02

GLOBAL SENTIMENTS

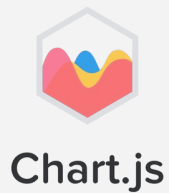
03

TOPIC ANALYSIS

APP STRUCTURE

TECHNOLOGIES

FRONTEND



BACKEND



CONTAINER



SERVER



USE CASE



TESLA

- » Unveils its new Cybertruck
- » Highest stock price till date of \$962.86
- » Worst day of the year after brutal earnings report and loss of CTO

00 OVERALL STATISTICS

Data collection and processing information



OVERALL STATISTICS



Event 1 Event 2 Event 3  GitHub

Tesla unveils its new Cybertruck

Data collection information

Event start: [2019-11-21](#)

Event end: [2019-11-23](#)

Hashtags: [#cybertruck](#) [#Cybertruck](#) [#tesla](#) [#Tesla](#) [#teslamotors](#)
[#TeslaMotors](#) [#ElonMusk](#) [#elonmusk](#)

Processing information

Total tweets: [8725](#)

Tweets with location: [6166](#)

Tweets with unknown location: [1174](#)

Duplicate tweets: [242](#)

Tweets analyzed: [4750](#)

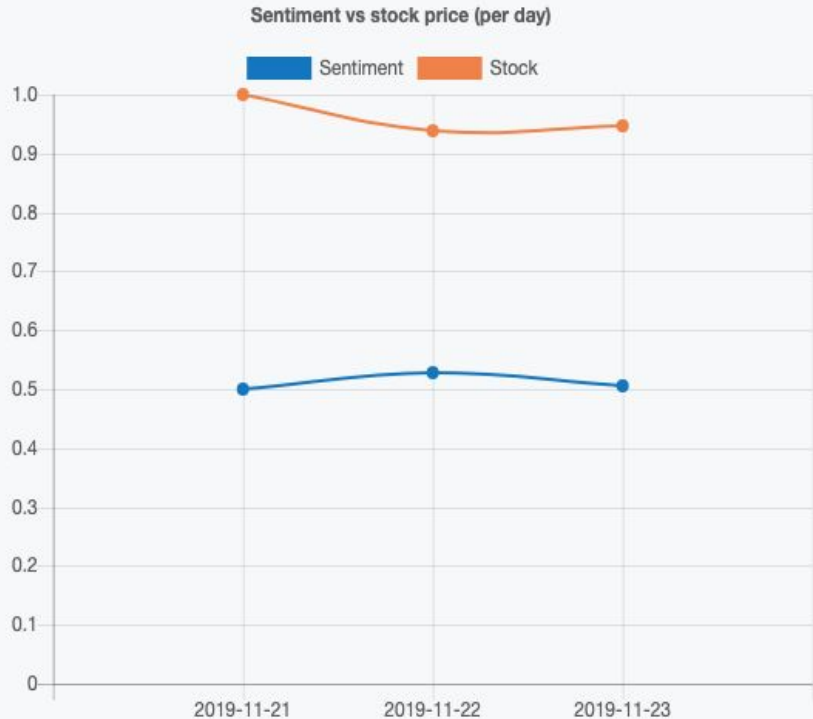


CORRELATION

01

Compare the trends of stock prices and users' sentiments

CORRELATION

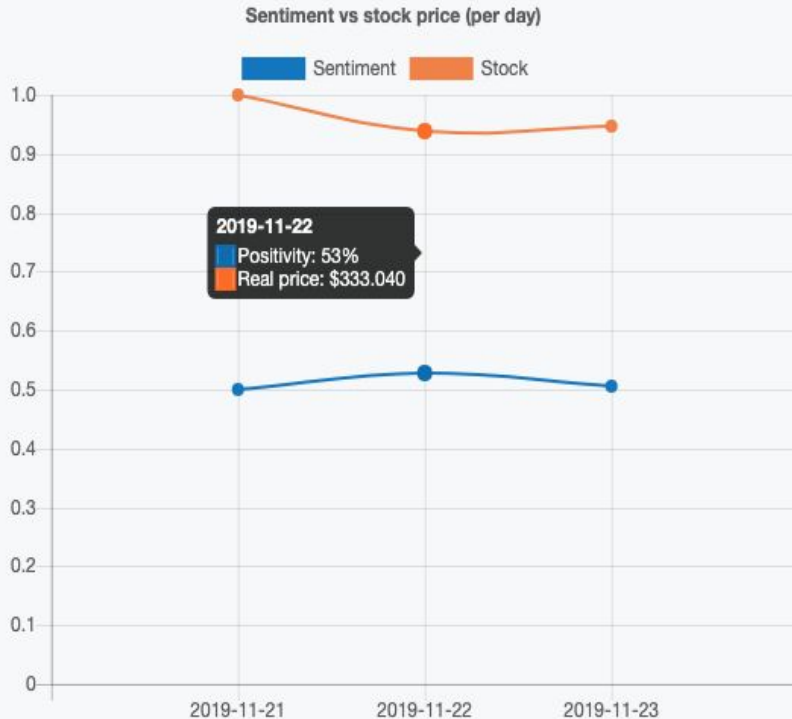


$$\rho(X,Y) = \text{cov}(X,Y) / \sigma(X)\sigma(Y)$$

Correlation Coefficient

-0.78565

CORRELATION



$$\rho(X,Y) = \text{cov}(X,Y) / \sigma(X)\sigma(Y)$$

Correlation Coefficient

-0.78565

02

GLOBAL SENTIMENTS

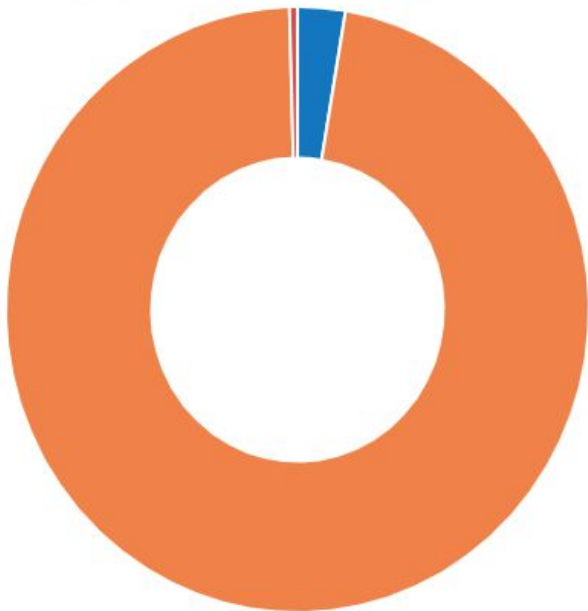
Distribution of the
sentiments across the world



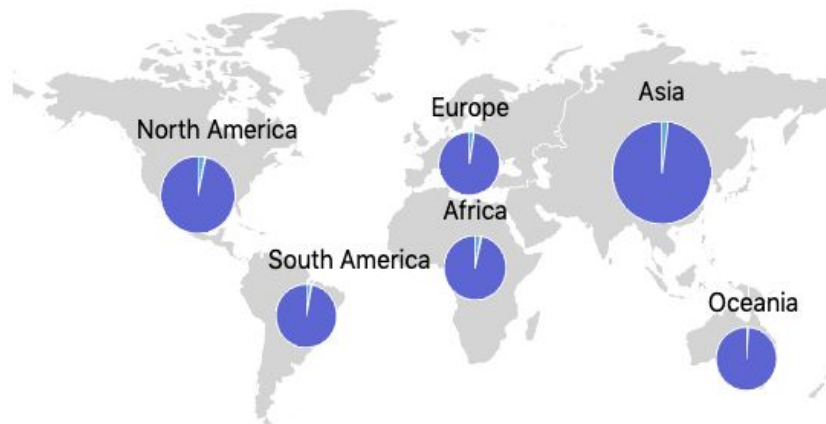
GLOBAL SENTIMENTS

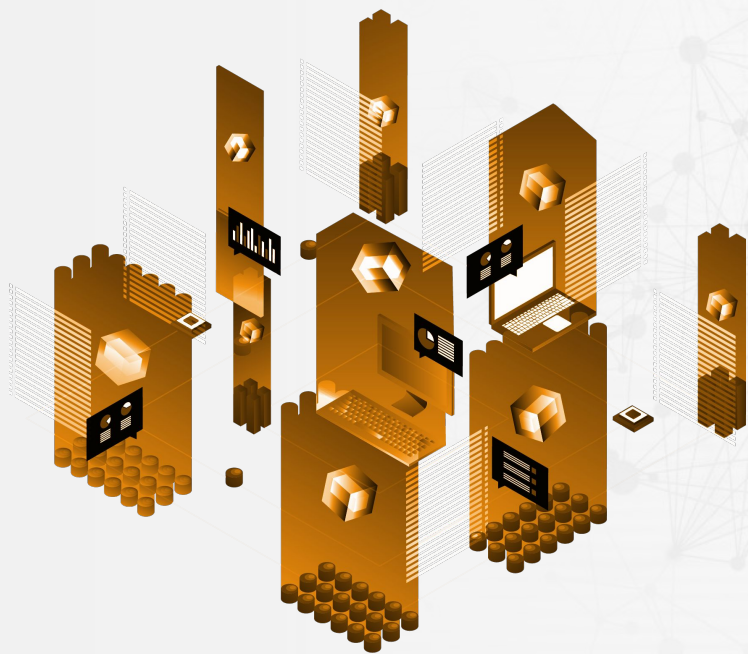
Overall sentiment for the event (percentage)

Positive Neutral Negative



Sentiment per continent (percentage)



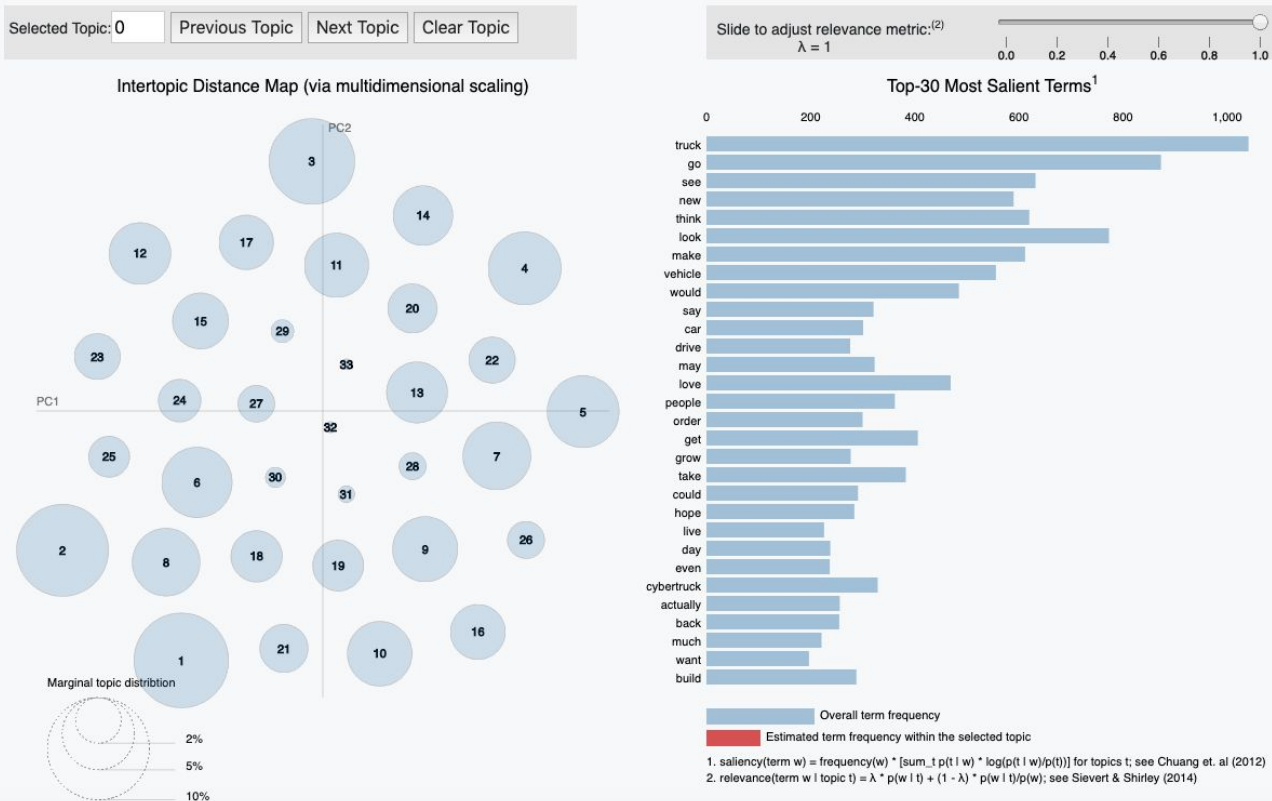


TOPIC ANALYSIS

Visualizations on most
relevant topics

03

TOPIC ANALYSIS



TOPIC ANALYSIS

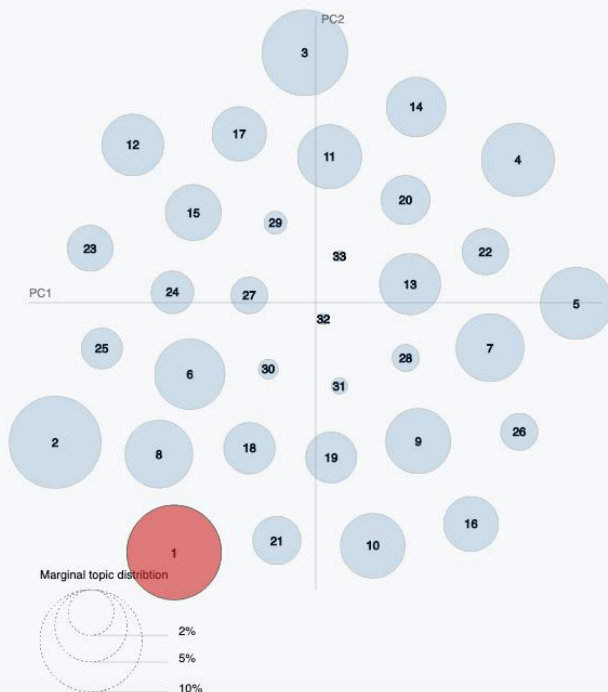
Selected Topic: 1 Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:⁽²⁾

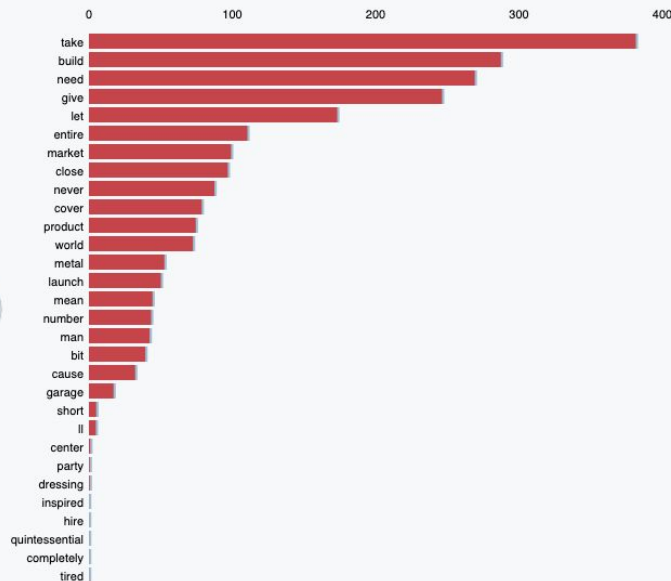
$\lambda = 1$

0.0 0.2 0.4 0.6 0.8 1.0

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 1 (8.6% of tokens)



Overall term frequency

Estimated term frequency within the selected topic

1. $saliency(\text{term } w) = \text{frequency}(w) * [\sum_t p(t | w) * \log(p(t | w)/p(t))]$ for topics t ; see Chuang et. al (2012)

2. $relevance(\text{term } w | \text{topic } t) = \lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

THANKS



Try it out at: <https://thzois.com/swanalytics>

 **GitHub** https://github.com/thzois/sw_2020

