



# VAPING TRENDS

Apoorv Mehrotra | Arnav Deshwal | Tushar Gupta | Kumar Udbhav



## Vaping is Big \$\$\$, but..

2018

\$14.5B\*\*

No recorded statistics

2019

\$18B\*\* @20% CAGR  
(Projected)

**47\* to date**



\*\*<https://www.prnewswire.com/news-releases/vaping-market-to-reach-a-value-29-39-billion-at-a-cagr-of-20-3-from-2018-2022--a-report-from-tbrc-300904265.html>

\*[https://www.cdc.gov/tobacco/basic\\_information/e-cigarettes/severe-lung-disease.html](https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html)

# The future is at risk !!!

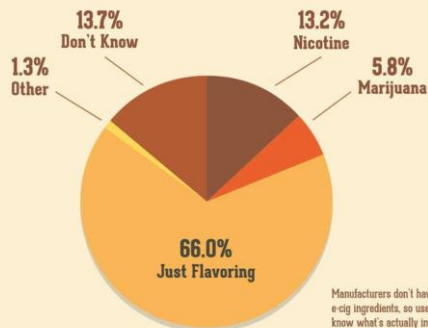
## TEEN E-CIG USERS ARE MORE LIKELY TO START SMOKING.\*<sup>2</sup>

Start Smoking Within 6 Months



\*Includes combustible tobacco products (cigarettes, cigars, and hookahs)

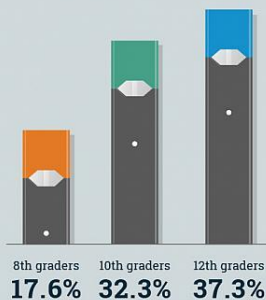
## WHAT DO TEENS SAY IS IN THEIR E-CIG?<sup>3</sup>



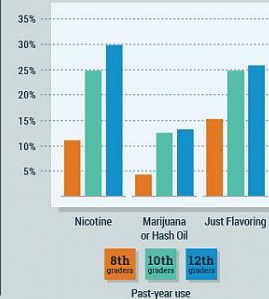
Manufacturers don't have to report e-cig ingredients, so users don't know what's actually in them.

## TEENS USING VAPING DEVICES IN RECORD NUMBERS

### PAST-YEAR VAPING



### WHAT DO TEENS SAY THEY ARE VAPING?



NEARLY 2 IN 5 STUDENTS IN 12TH GRADE REPORT PAST-YEAR VAPING, RAISING CONCERNS ABOUT THE IMPACT ON BRAIN HEALTH AND POTENTIAL FOR ADDICTION.

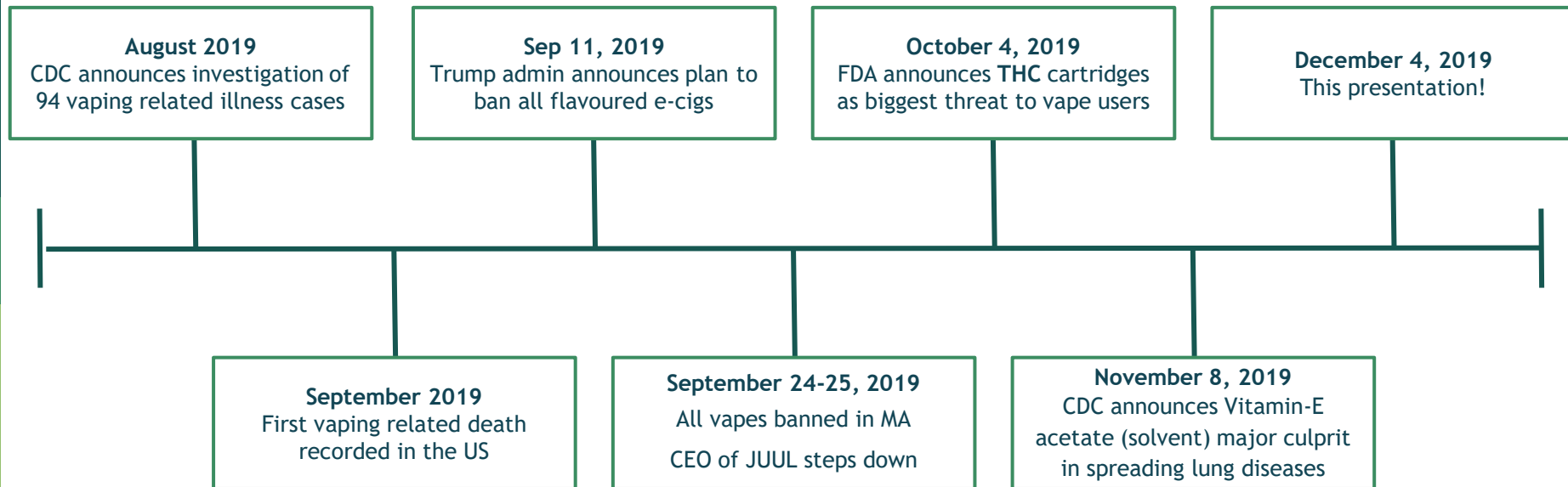


NIH National Institute on Drug Abuse

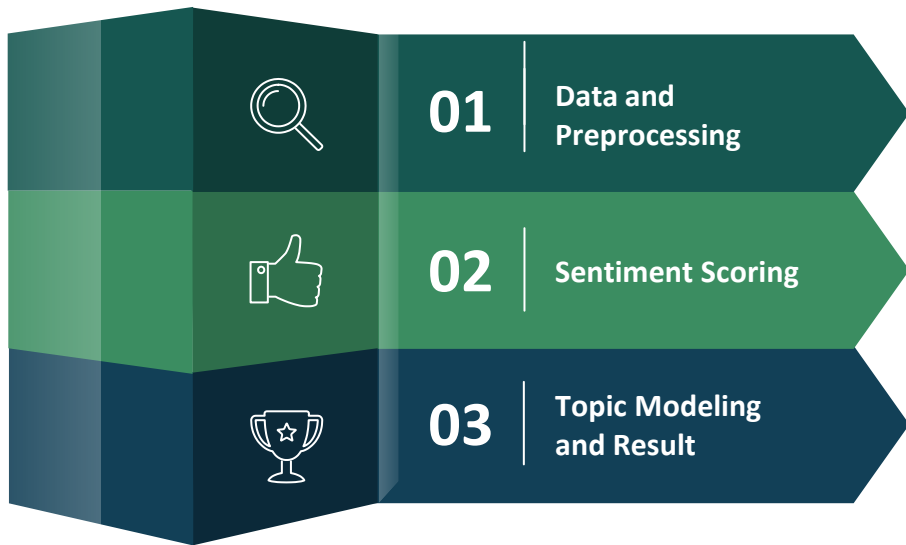
DRUGABUSE.GOV



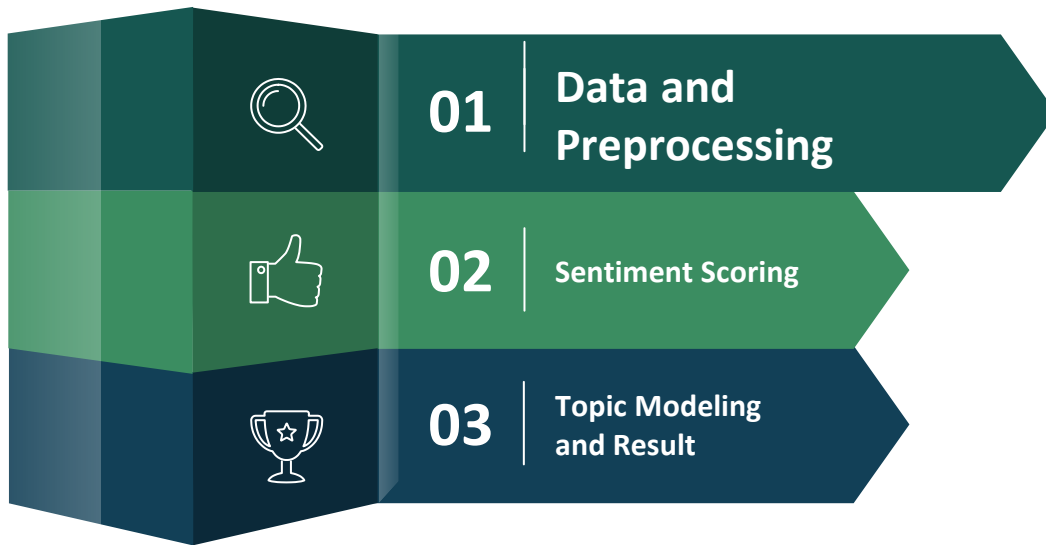
## Looking through the cloud of smoke



# Agenda



# Agenda





## Data extraction and cleaning process

### Data Scraping

Scraping of general discussions on vaping forum of vaportalk.com

### Data Cleaning

Words to lowercase,  
removal of punctuation,  
non-alphanumeric  
words, and stopwords

### Lemmatization

Extraction of root words  
from inflected words

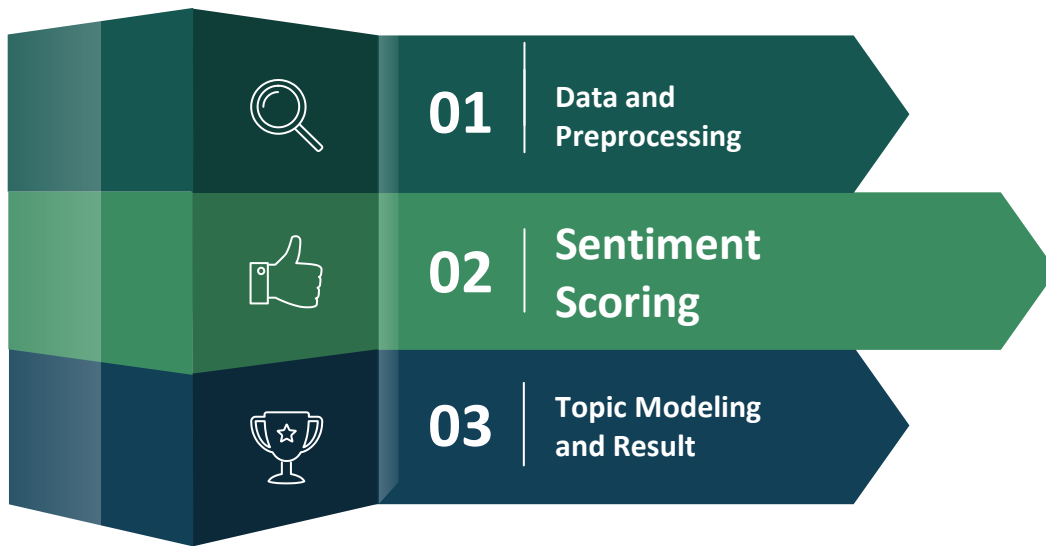


## Sample data

Thread Name	Thread Date	Thread User	Thread Replies	Comment User Name	Comment Date	Comment User Posts	Comment
need help with mixing liquid	2019-10-24T14:19:16Z	dbster	5	adamabbot	2019-11-19 22:14:36+00:00	5	thanks answer also interested good mix may recommend know mix interested taste mix
aspire breeze nxt 50vg50pv or 70vg30pg to maximise coil use	2019-11-23T13:26:16Z	sophie777	1	Tam	2019-11-23 17:13:16+00:00	4489	mean find one first number company use first second others use first number eliquid faster gunk coil
flavor ban	2019-09-20T15:40:51Z	Tyson	18	Bebop	2019-11-16 20:54:09+00:00	2127	one whole county area banned vape product sale



# Agenda





## Google to the rescue...

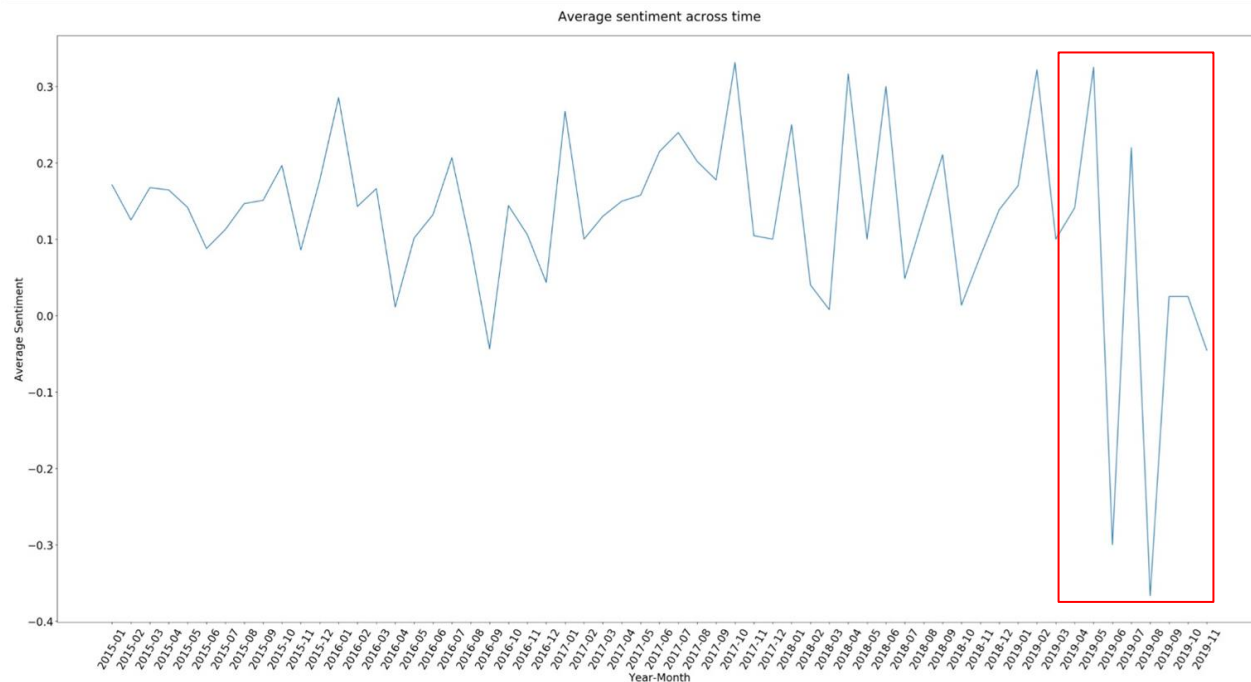
- Used Google's Natural Language API for sentiment scoring
- Received JSON response with sentiment score (on a scale of -1 to 1), magnitude of sentiment, and detected language
- Defined thresholds to identify sentiments  
**-1 to -0.3: Negative**   **-0.3 to 0.3: Neutral**   **0.3 to 1: Positive**

*Tricky part is to constraint the number of hits to the API to 600 per minute for the free tier*



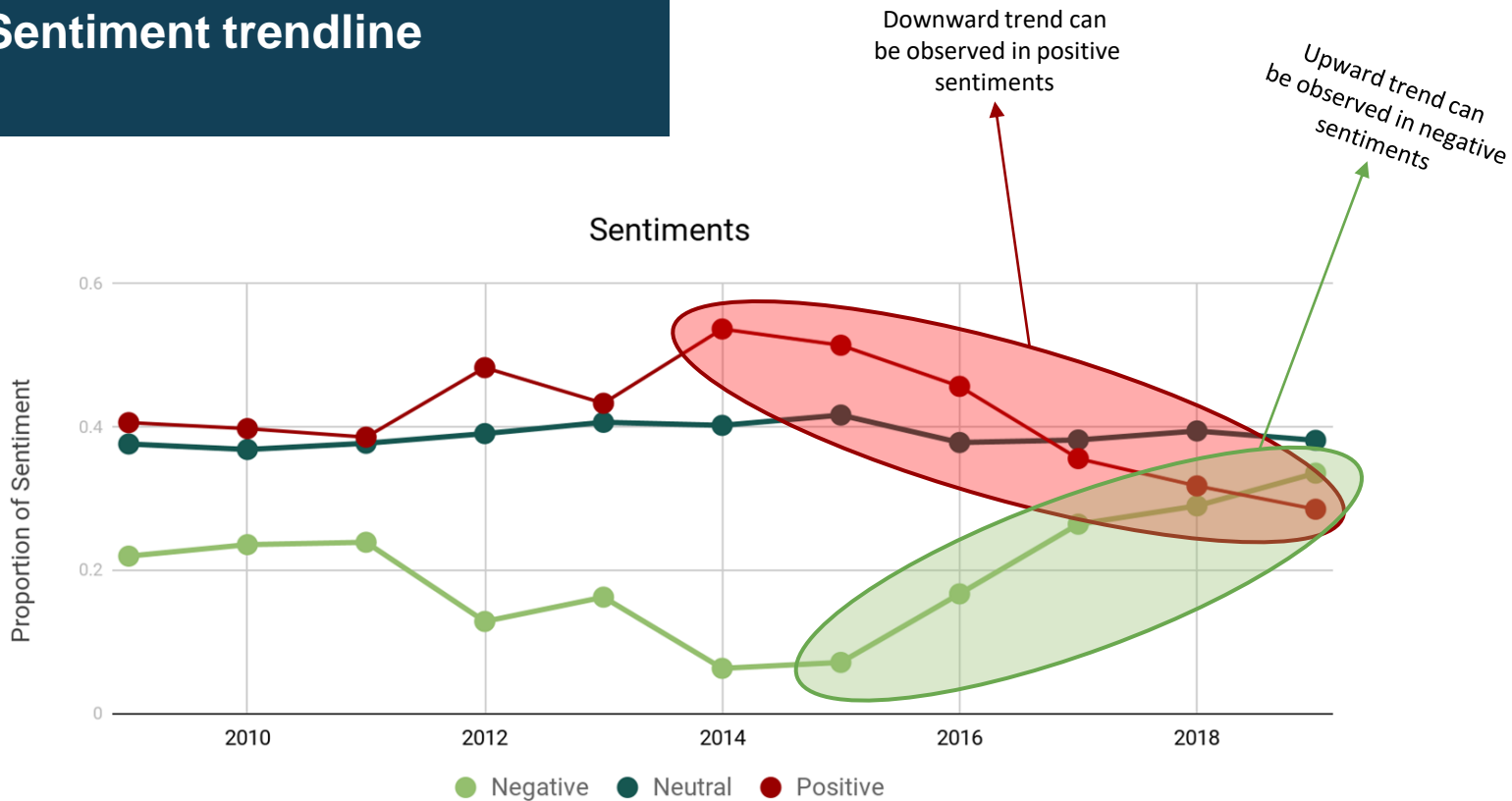
## Overall sentiments

- Overall sentiments have dropped in coherence with the national events related to vapes

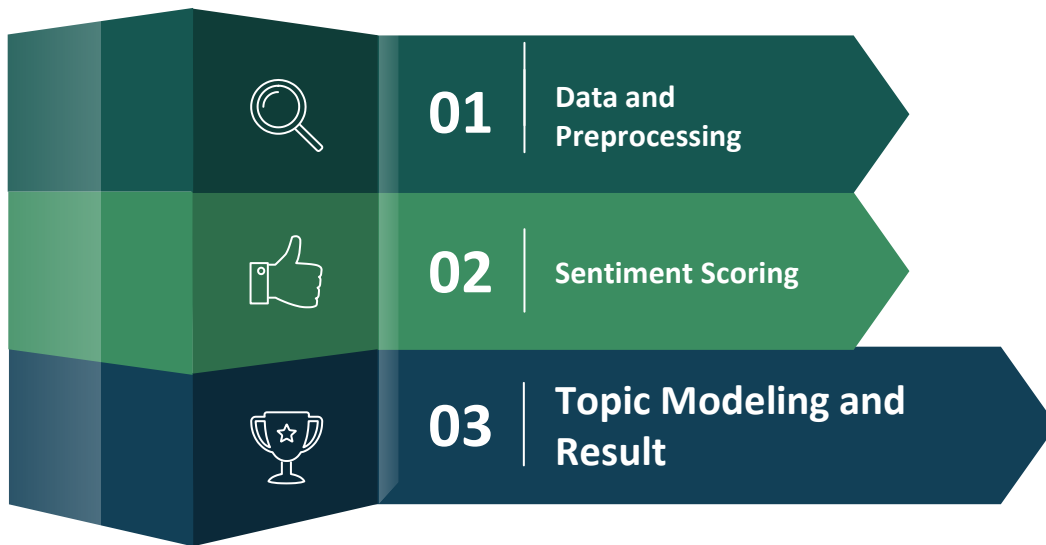




## Sentiment trendline



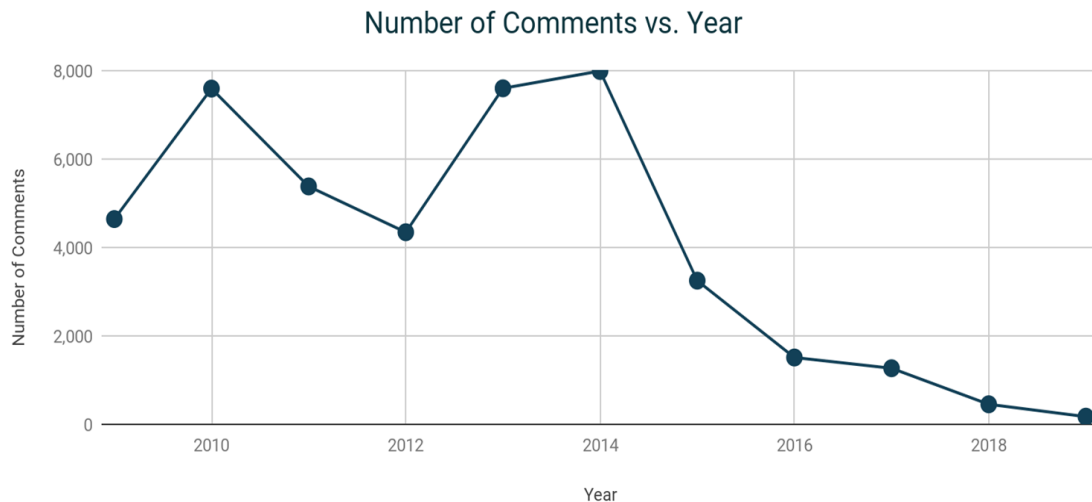
# Agenda





## Drop in discussion

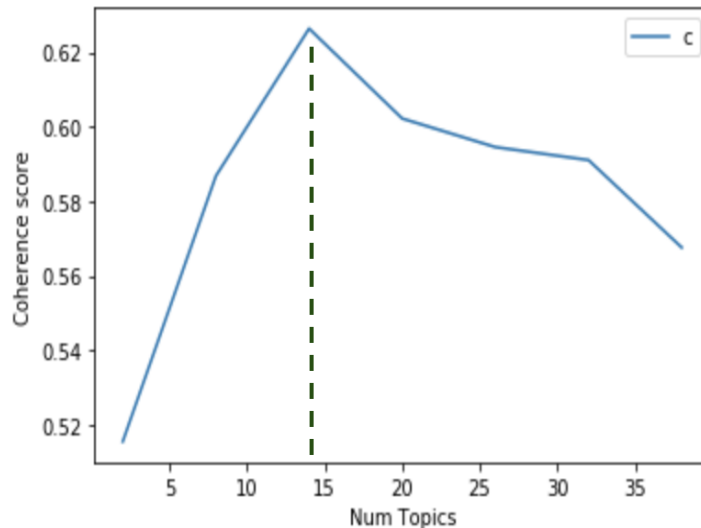
There is an overall **drop** in the volume of discussion across years





## Scree plot

Scree plot for coherence score across number of topics suggest **14 topics** across the discussions on the forum



Num Topics = 2 has Coherence Value of 0.5155

Num Topics = 8 has Coherence Value of 0.5868

Num Topics = 14 has Coherence Value of 0.6264

Num Topics = 20 has Coherence Value of 0.6022

Num Topics = 26 has Coherence Value of 0.5946

Num Topics = 32 has Coherence Value of 0.5911

Num Topics = 38 has Coherence Value of 0.5676



## 14 topics identified

Closer topics on an **'Intertopic Distance Map'** are grouped into broader topic segments based on the commonality of discussion in those topics





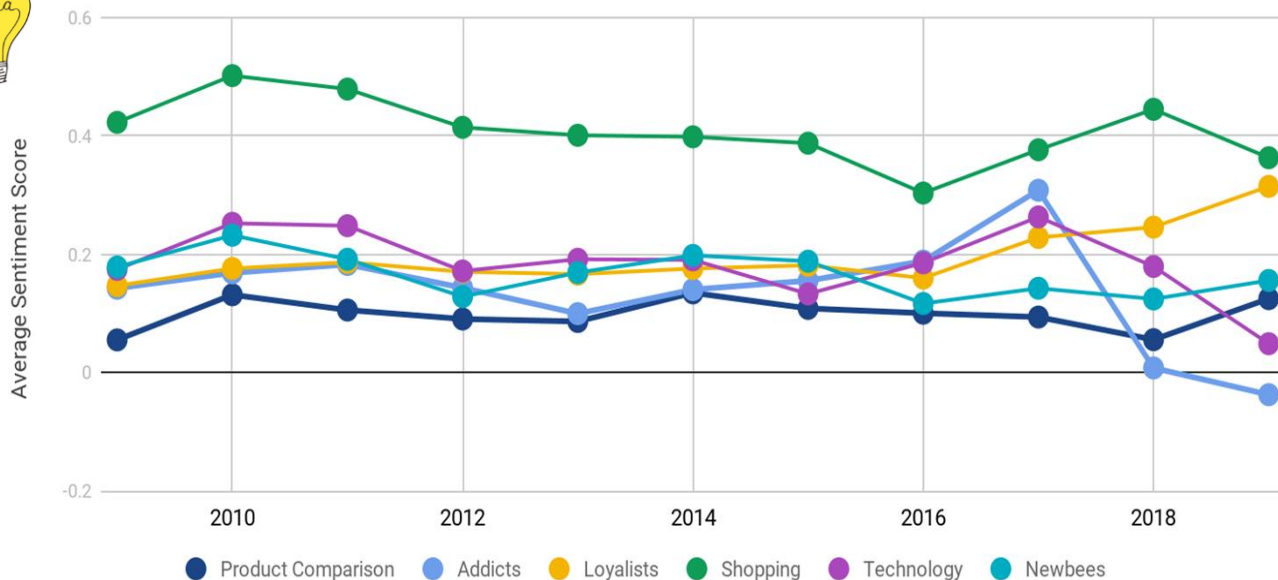


## Sentiment score by group

- Vape addiction topics show a **drop** in average sentiments after 2017
- Vape Loyalists are still fighting the battle
- Sentiment score across discussion topics like vape buying, product comparisons etc is fairly constant



Sentiment Score by Group



A young man with short brown hair is blowing a large, dense cloud of white vapor from a vape pen. He is wearing a red jacket. The background is dark. The word "Questions?" is overlaid in a large, green, sans-serif font across the center of the image.

Questions?

# Appendix

# Summary of topics

- Topic 9,12 : welcom comments for new members joining the forum
- Topic 8,10,11 : Vape technicalities - discussions around technical details of a vape - battery, resistance, tank, coil etc
- Topic 13, 14: addiction, comarison to nicotine addiction etc.
- Topic 6 : Battery discussions mostly
- Topic 7 : vape look and feel, product related discussions
- Topic 5 : vape supplies, ordering, time taken to deliver etc
- Topic 1,2,3,4 : Vape supporters, switching from cig to vape, vape flavors, juice, taste etc.

# Gensim's inbuilt version

```
(0, '0.016*"day" + 0.016*"vaping" + 0.015*"time" + 0.014*"one"')
(1, '0.025*"vapor" + 0.022*"get" + 0.021*"taste" + 0.019*"flavor"')
(2, '0.026*"know" + 0.025*"thanks" + 0.020*"would" + 0.018*"guy"')
(3, '0.068*"battery" + 0.031*"ego" + 0.018*"mod" + 0.018*"one"')
(4, '0.046*"cartos" + 0.034*"vapor" + 0.027*"talk" + 0.019*"water"')
(5, '0.032*"juice" + 0.026*"flavor" + 0.022*"like" + 0.020*"order"')
(6, '0.024*"tank" + 0.023*"cart" + 0.022*"atomizer" + 0.020*"drip"')
(7, '0.065*"nice" + 0.042*"cool" + 0.040*"glad" + 0.035*"sound"')
(8, '0.037*"post" + 0.031*"video" + 0.030*"forum" + 0.028*"review"')
(9, '0.066*"ecigs" + 0.059*"awesome" + 0.045*"wow" + 0.040*"http"')
```

# Mallet's version ( using optimal number of topics)

```
(0, '0.074*"tank" + 0.056*"coil" + 0.028*"hole" + 0.025*"mini"')
(1, '0.068*"drip" + 0.051*"tip" + 0.036*"cart" + 0.025*"water"')
(2, '0.039*"mouth" + 0.030*"welcome" + 0.025*"congratulation" + 0.025*"computer"')
(3, '0.027*"order" + 0.021*"attys" + 0.016*"vapor" + 0.015*"supplier"')
(4, '0.060*"nicotine" + 0.056*"nic" + 0.028*"manual" + 0.026*"pack"')
(5, '0.035*"vaping" + 0.023*"day" + 0.020*"smoking" + 0.019*"smoke"')
(6, '0.050*"volt" + 0.037*"http" + 0.033*"ohm" + 0.027*"voltage"')
(7, '0.037*"one" + 0.034*"got" + 0.018*"day" + 0.017*"week"')
(8, '0.073*"thanks" + 0.044*"good" + 0.039*"guy" + 0.034*"lol"')
(9, '0.027*"flavor" + 0.027*"juice" + 0.025*"like" + 0.021*"atty"')
(10, '0.146*"battery" + 0.056*"ego" + 0.028*"charger" + 0.023*"use"')
(11, '0.030*"know" + 0.028*"would" + 0.025*"like" + 0.022*"one"')
(12, '0.056*"mod" + 0.038*"look" + 0.036*"video" + 0.031*"like"')
(13, '0.046*"hey" + 0.039*"congrats" + 0.033*"forum" + 0.029*"spade"')
```