

The Truth about Secrets

Laura Chen

Agenda

- Introduction to Whisper
- Exploratory Data Analysis
- Sentiment Analysis
- Topic Modelling/Clustering
- App concept
- Next Steps

About Whisper

Whisper is a popular, anonymous platform for sharing secrets with the world.

- Stress relief through ranting or confessing without consequences
- Mental health & suicide prevention
- Voicing honest/unpopular opinions
- Citizen reporting and political opinions in censored countries



High level Overview

Whisper

250M

monthly users

187

countries

75%

usership between
the ages of 18 - 34

Secret data

34,241

secrets scraped

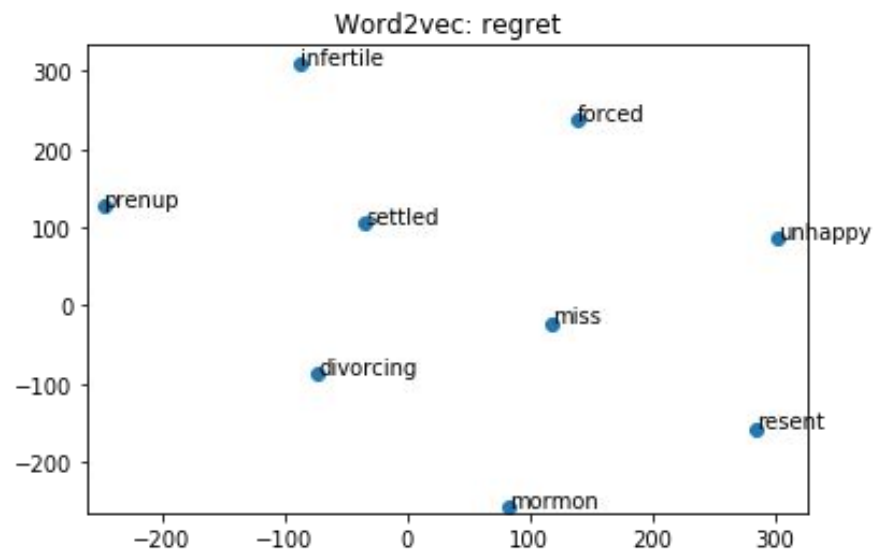
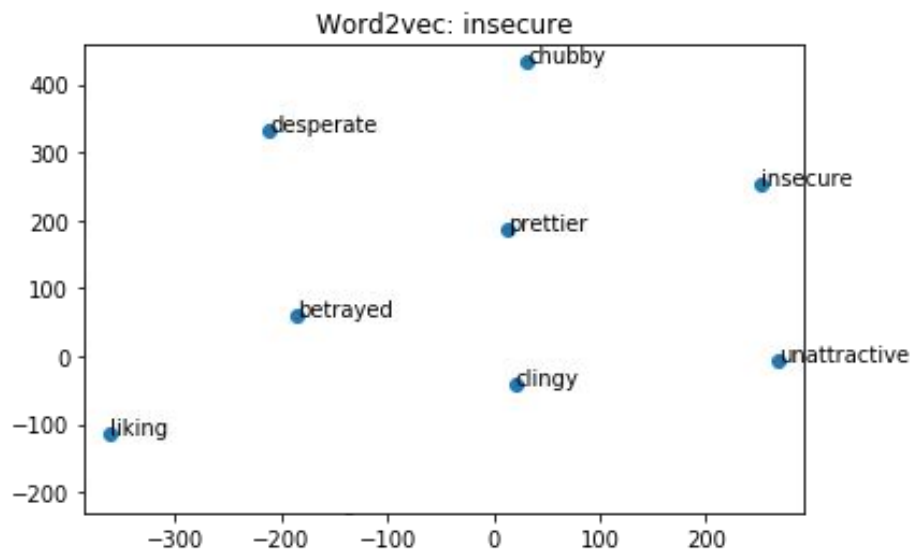
11

Avg words per secret
(no stopwords)

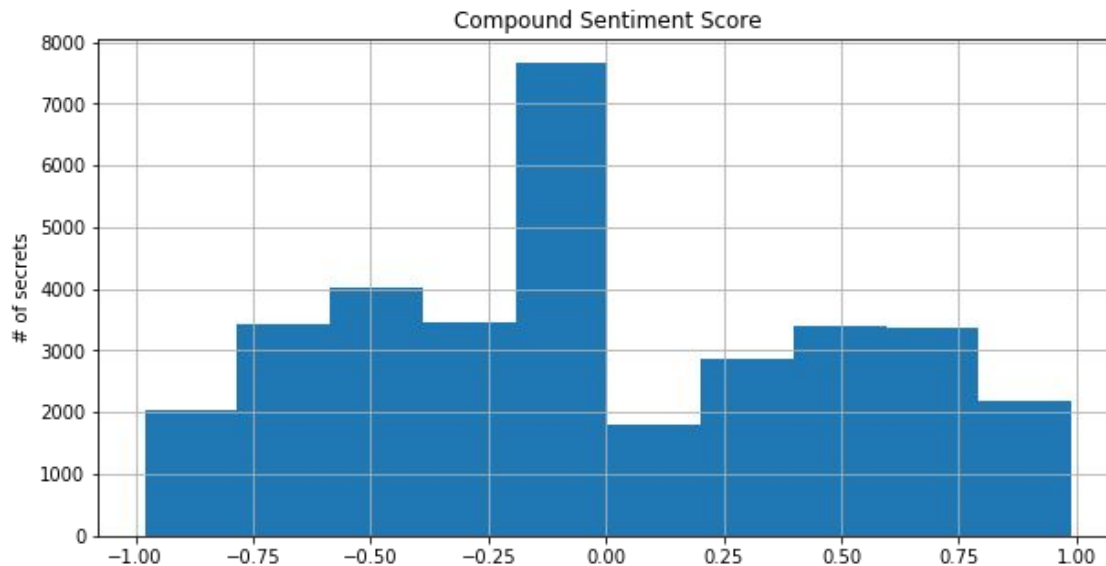
.07

Unique words/
Total words

Word2vec Exploration



Sentiment Analysis



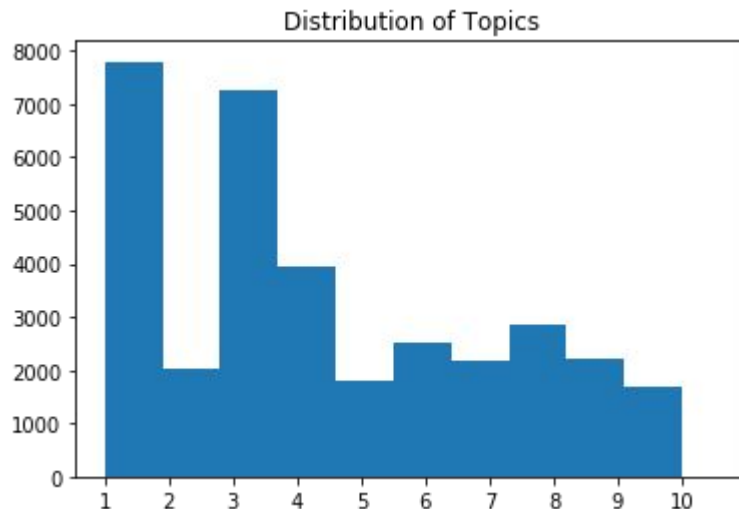
"i would feel guilty bringing a child into this world full of hate, pollution, crime, sexism, racism, drugs, and disgusting people."

"my twin sister and i have switched places for about a week, on multiple occasions."

"i love seeing pda because there isn't enough love in the world and it's beautiful and adorable."

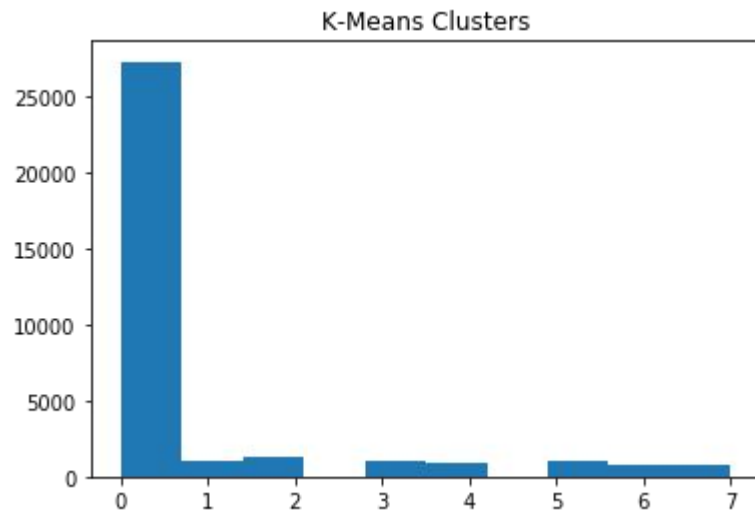
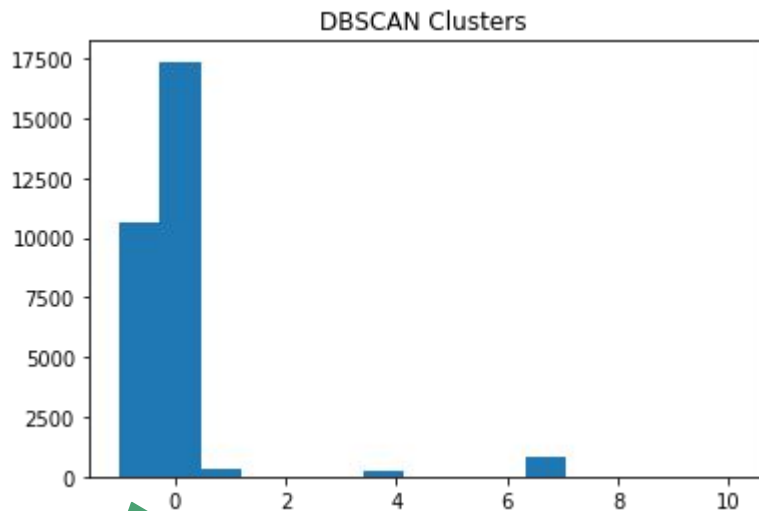
Topic Modeling with NMF

1. **Parents/family**
2. Sex
3. **Jobs/Money**
4. **Exes**
5. Hatred
6. Marriage
7. Kids
8. Pregnancy
9. High School/College
10. Dating



15% of secrets have a max topic coefficient of less than .00001

Clustering from NMF Topics



“It feels like a sin to wash bacon grease off a pan”

Web App Mockup

Shared Secrets

Type your secret below:

i hate my job and can't wait to quit

Find Similar Secrets

Similarity: 1.00

Secret: never in a million years would i have thought i would hate my dream job.

Similarity: 1.00

Secret: the manager at my job had a gambling addiction and he stole money from our tip pool so everyone was shorted every week for 2 years. hate that guy.

Similarity: 1.00

Secret: i loved working for a grocery store; it was the terrible management that made me hate the job.

Next Steps

- Classification of secrets in a higher level = regrets, wishes, insecurities
- Analyzing sentiment and topics of secrets from different countries
- Comparing models to non-anonymous social media sites such as Twitter, Facebook, and Instagram

Thank you!

Questions?

Appendix

Model Challenges

Examples of secrets with low NMF coefficients:

- Zero Context:
 - " it appears part 3 of my dastardly plan has been a success"
- Oddly specific:
 - "why does this girl have so many pics of her hamster? Freak."
- Random thoughts:
 - "it feels like a sin to wash bacon grease off a pan."
- Lighthearted confessions:
 - "i've caught 3 pokemon in less than 5 min at church."
- Unpopular opinions:
 - "i'm an immigrant and i strongly support trump and the building of a wall."

Exploring the data

Most popular Bigrams:

- i feel like i 'm
- i n't know i 'm
- i tell people i 'm
- i n't know i pregnant
- i 'm pregnant i 'm
- i 'm love best friend
- i know i 'm love
- i 'm afraid i 'll
- i ca n't help feel

