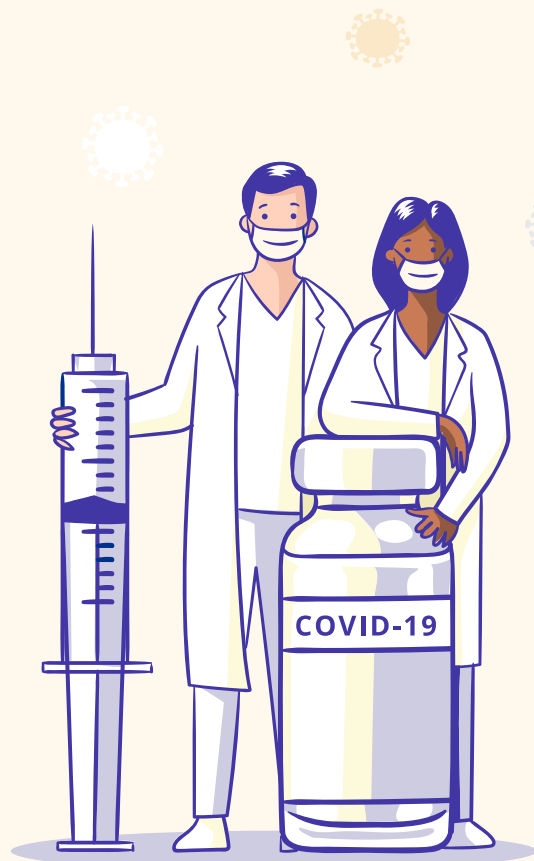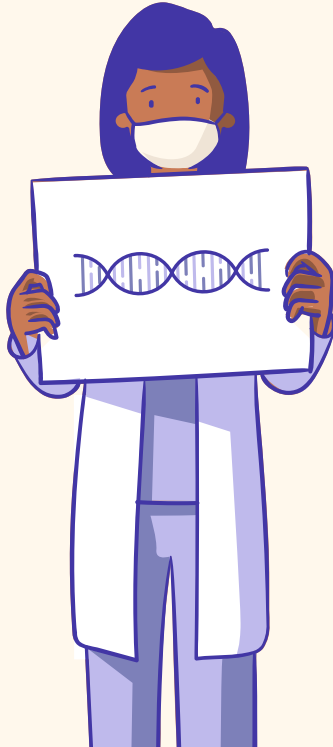# Discovering Popular Covid-19 Vaccine Myths

Using Unsupervised Learning and Natural Language Processing on Reddit Anti-Vaxx Threads

Motivation

How can we help?

NYC Health Department launched multicultural and multi-lingual vaccination ad campaigns

Goal: to win over skeptical Americans, curb the spread of the novel coronavirus

Nearly 70% of new cases are delta variant

Stats:
Fully vaccinated in NYC:
4.5 million – 53.5% of city's total population
At least one dose:
4.8 million – 58.1% of city's total population

Methodology

Data Acquisition

- Reddit Praw API
- 2834 Documents
- Post titles
- Post body
- Comments

EDA
- Pandas
- Numpy
- Matplotlib
- Plotly

Topic Modeling
- NMF (CountVectorizer, TFIDF)
- LSA (CountVectorizer, TFIDF)

Sentiment Analysis
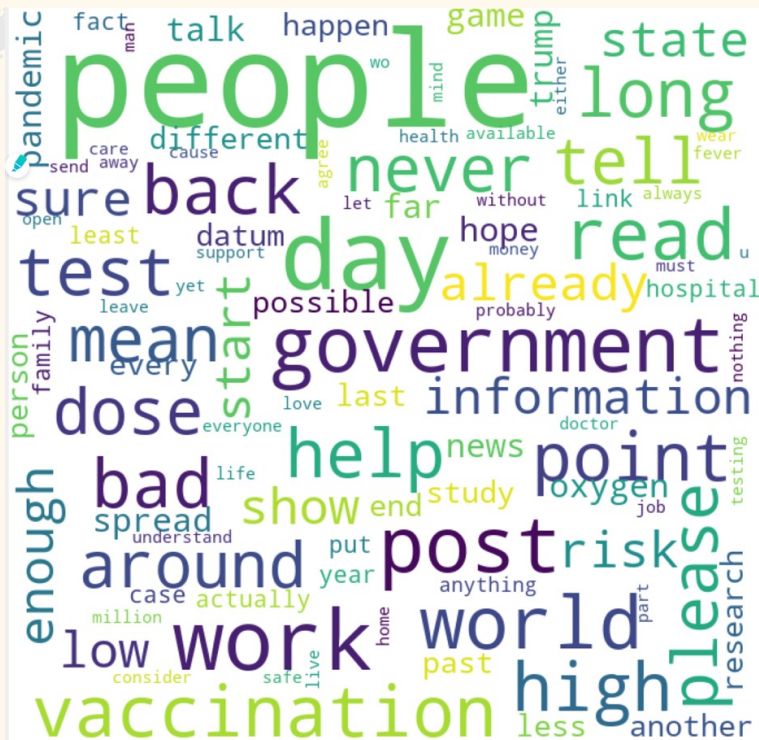- Textblob
- Polarity scores

Topic Interpretation:
- Ranked Topics
- Coherency
- Relevancy
- Popularity
- Uniqueness

1   2   3   4   5

Start   End

Topic Modeling:

Top Choice: NMF using CountVectorizer

Topic 1: Work, post pandemic

Topic 2: High Tech

Topic 3: Trump/Russia

Topic 4: Vaccine Side Effects

Topic 5: Science vs Conspiracy

Topic 5: Variant

Topic 6: Jobs

Topic 7: Tracing calls

Topic 8: Academic Research

Topic 9: Vaccine in Rural Areas

Topic 10: Government's prioritization of economy over lives

Topic Interpretation:

Most to Least Popular Topics

Topic 9: Vaccine in Rural Areas

Topic 0: Work- post pandemic

Topic 10: Government's prioritization of economy over lives

Topic 6: Tracing calls

Topic 3: Science vs Conspiracy

Topic 7: Tracing calls

Topic 8: Academic Research

Topic 2: Trump/Russia

Topic 5: Variant

Topic 1: High Tech

Topic 4: Science vs Conspiracy

Topic 9: Vaccine in Rural Areas

Topic 10: Government's prioritization of economy over lives

Topic 3: Vaccine Side Effects

Topic 0: Work, post pandemic

Topic 4: Science vs Conspiracy
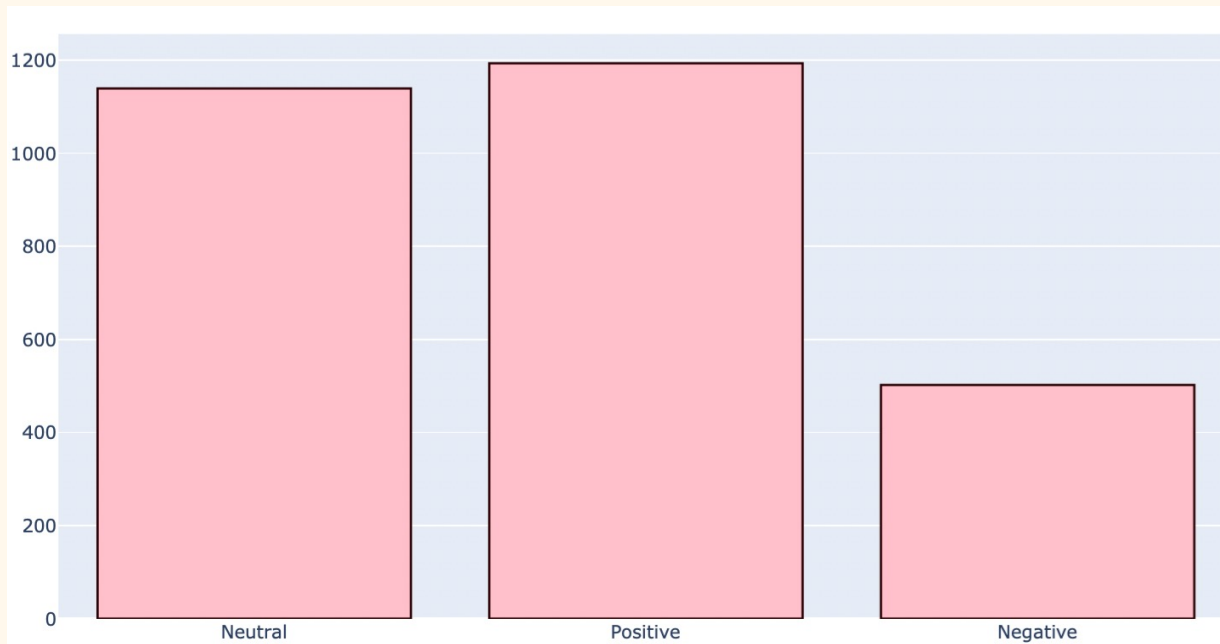
Topic 8: Academic Research

Topic 5: Variant

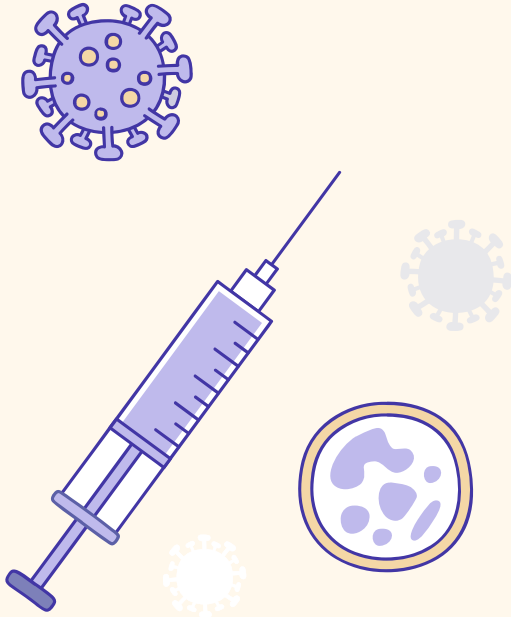Topic 2: Trump/Russia

Topic 6: Jobs

Topic 7: Tracing calls

Topic 1: High Tech

Sentiment Analysis: TextBlob

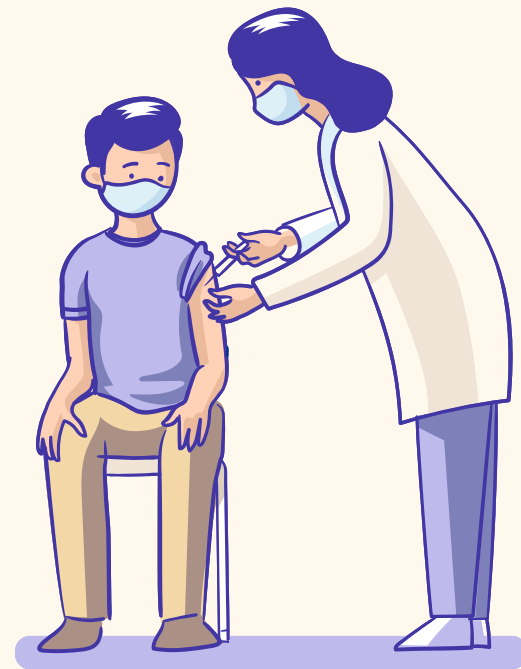# Final Recommendation to NYC Health:
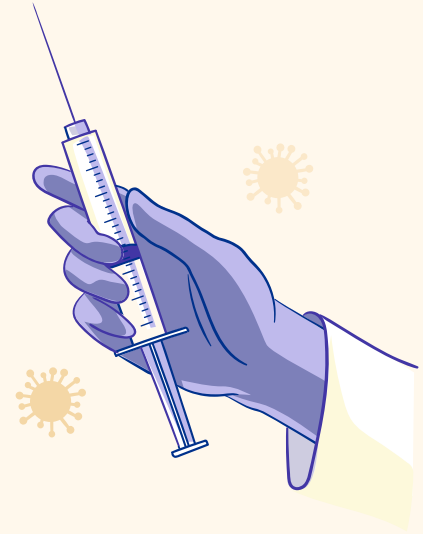
Focal Points for Next Ad:

Topics

10 & 3 !

Thanks!

Have you gotten the vaccine?

Visit nyc.health/vaccinefinder

for scheduling information

Appendix:

Polarity Distribution