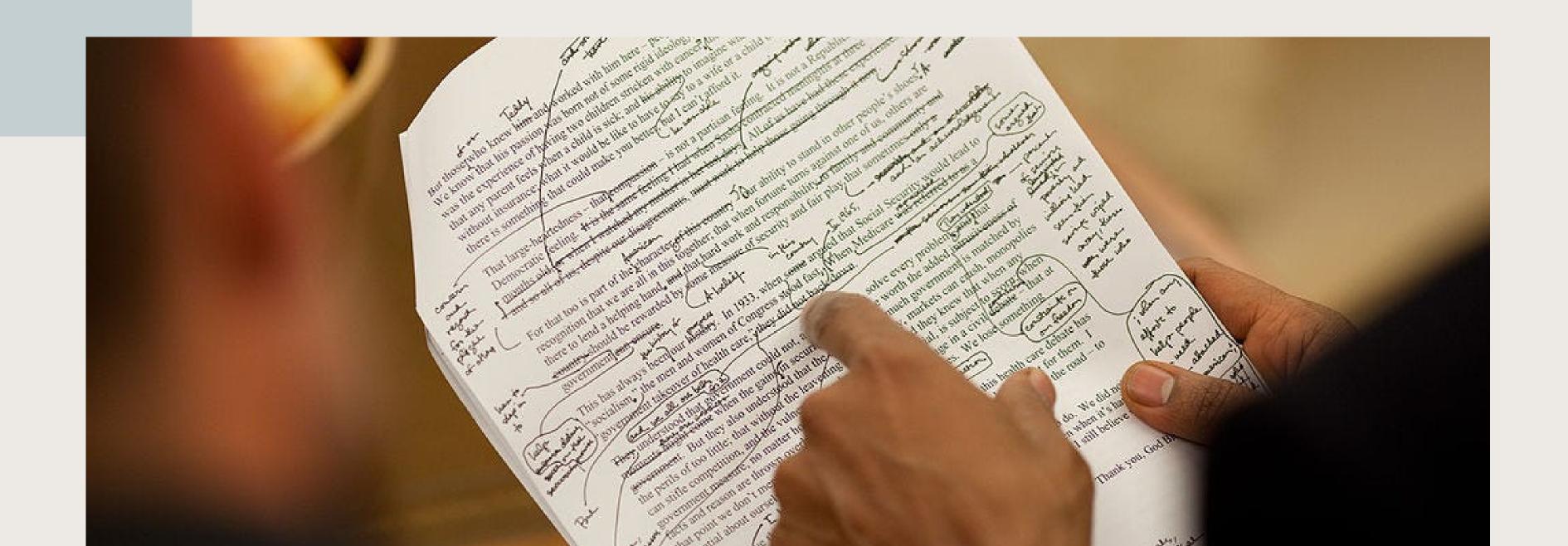
WHAT MAKES A GREAT SPEECH?

Capstone Sprint presentation

BY DREWE MACIVER



WHAT MAKES A GREAT SPEECH?

A SPEECH'S INFLUENCE

Speeches inform, motivate, and inspire change.

BUT WHAT MAKES A SPEECH MOST IMPACTFUL?

Who can benefit from our findings?

Business managers (town halls; meetings)

Students (presentations)

TED speakers

Speech givers (at toasts and family events)

Anyone looking to influence thru the power of speech.



OUR GOAL:

To quantify an answer to the question, "What makes a great speech?"

(Analyze TED Talks to identify what correlates with their most influential and well-liked talks)

COMMUNICATION IS COMPLICATED. A COMPUTER WILL FIND THE CORRELATIONS.

SPEECH CONTENT ATTRIBUTES:

Quantifying the elements of the speech itself.

These are it's replicable performance metrics.





Language (NLP of transcripts)



Topic (subject of talk)



Tone (voice inflection)

SITUATIONAL CORRELATIONS:

Measuring the impact of non-performance metrics.



TED Event



Speaker Occupation



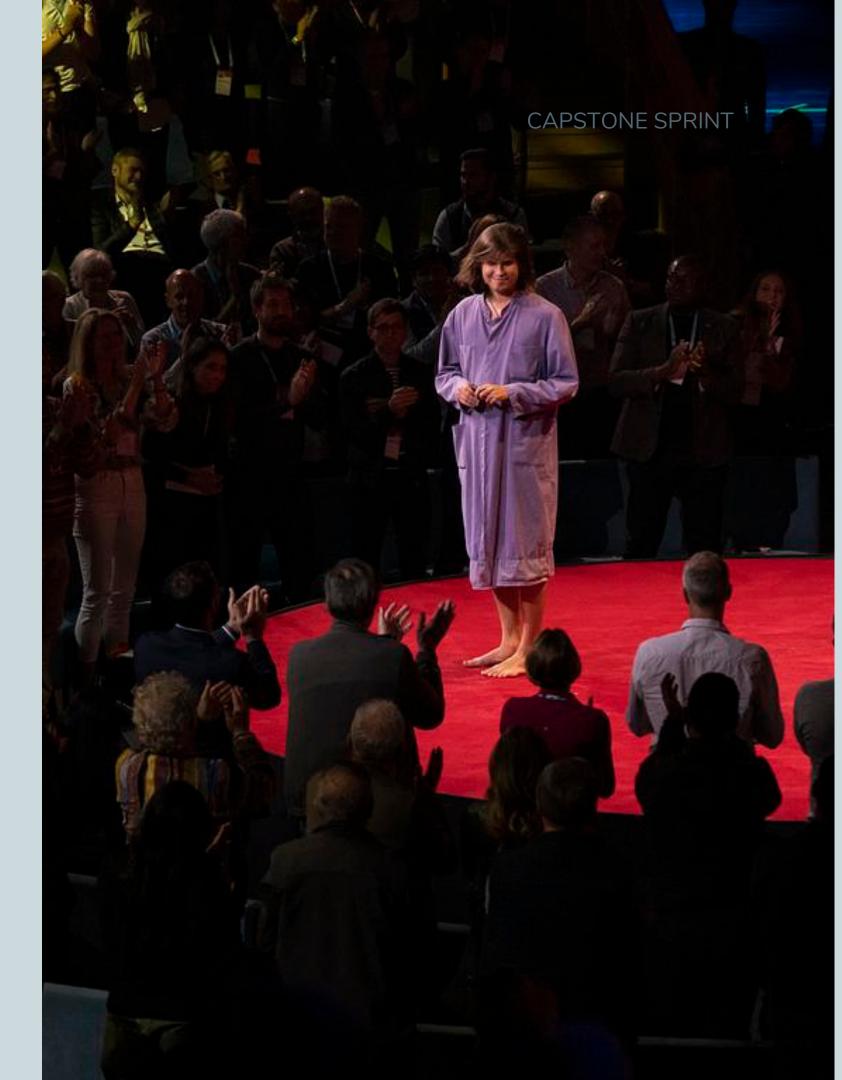
Date Recorded



Date Published (on TED website)

EMPIRICALLY IMPACTFUL SPEECHES

Combining empirically successful attributes to maximize a speeches impact. Language Words used. Cadence Tone Talking speed. Voice inflection. Maximum impact.



THE DATASET

Source: Kaggle.com; TEDLIUM 3

4,003 TED Talks from TED.com

Available Data:

Transcripts, # of Views, # of Likes, Audio Recordings, Speaker Information, etc...

Available Data

n = 4003

Title, Speaker, Occupation About Speaker, # of Views, Recorded Date, Published Date, TED Event, # of Comments, Duration, Topics, Related Talks, URL, Description, Transcript

Supplemental Datasets:

- # of Likes (via Kaggle) (n=5007)
- Audio Recordings (via TEDLIUM 3) (n=2351)

Data Quality Concerns

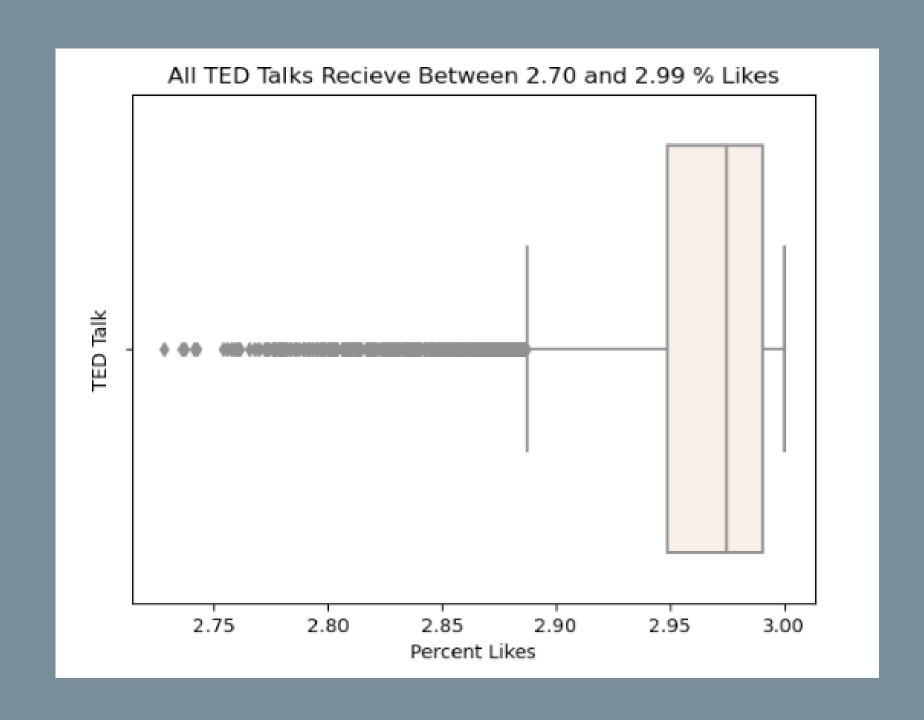
Limited Quantity:

There are only 4,003 talks datapoints in the dataset.

Similar Quality:

TED is a high-quality speech catalogue.

Determining "good" speeches from a catalogue of good speeches may be difficult.



NEXT STEPS:

Processing 1

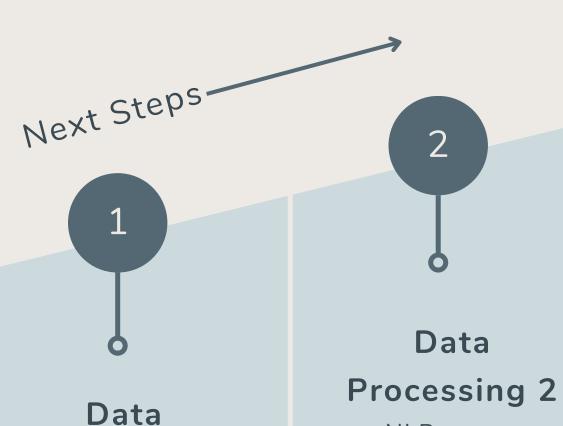
Merge datasets for

likes and views.

Possible incorporation

of additional talks.

DATA PROCESSING, FEATURE ENGINEERING, AND BASELINE MODELLING



NLP run on transcripts.

Feature Engineering 1

Exploratory considerations of additional feature variables.



Feature Engineering 2

Dummy variables encoded, initial correlation values determined.



Baseline Modelling

Iterative regression models, optimizing feature selection (using highlycorrelated values to start).



Audio File Incorporation

From a sample of the "best" speeches, discover which tonal attributes separate them from a random sample of other speeches.

What Makes A Great Speech?

Capstone Sprint

"NO ONE EVER COMPLAINS ABOUT A SPEECH BEING TOO SHORT!"

- IRA HAYES

CONTACT US

Drewe.Maciver@gmail.com +1.778.995.7801