

Demo Day Story

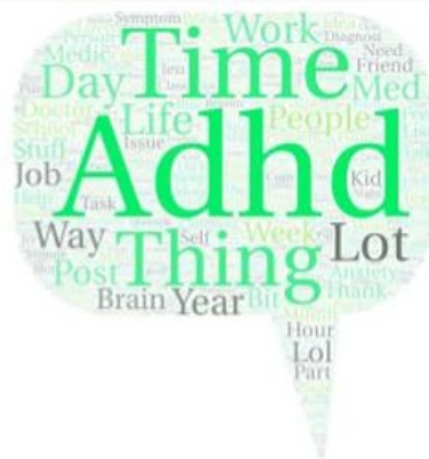
Title	Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA
-------	--------------	------------	----------	-------------------------------------	---------------------	--	--------------------------------------



NLP and the Challenges of Social Media

Topic modelling and Sentiment Analysis on a Reddit message board for women with ADHD.

By A. J. Hammond



reddit

Demo Day Story

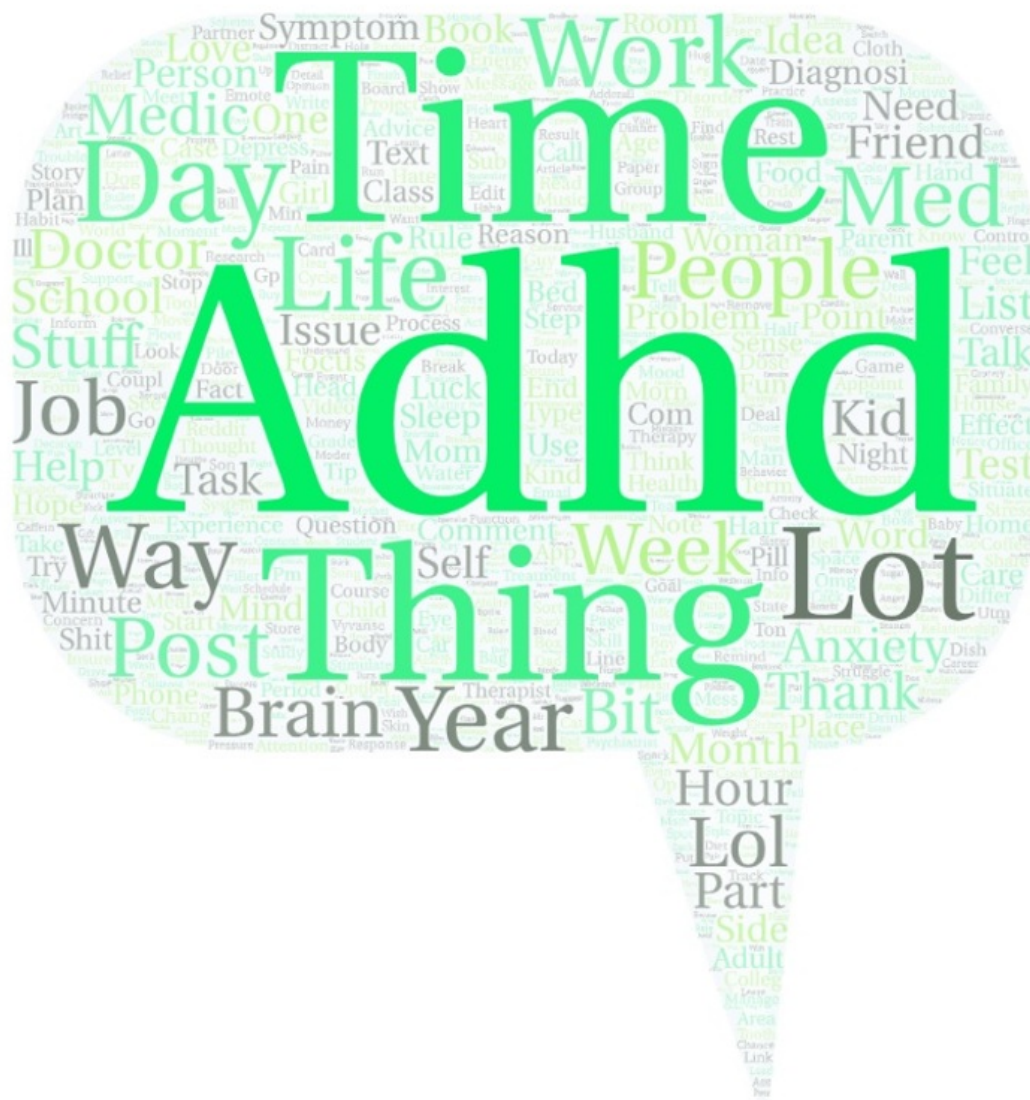
Title	Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA
-------	--------------	------------	----------	-------------------------------------	---------------------	--	--------------------------------------

Introduction

- Natural Language Processing
- Sentiment Analysis
- Topic Modelling
- Reddit message boards
- (Misnamed) Attention Deficit Hyperactivity Disorder

Demo Day Story

Title	Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA
-------	--------------	------------	----------	-------------------------------------	---------------------	--	--------------------------------------



Demo Day Story

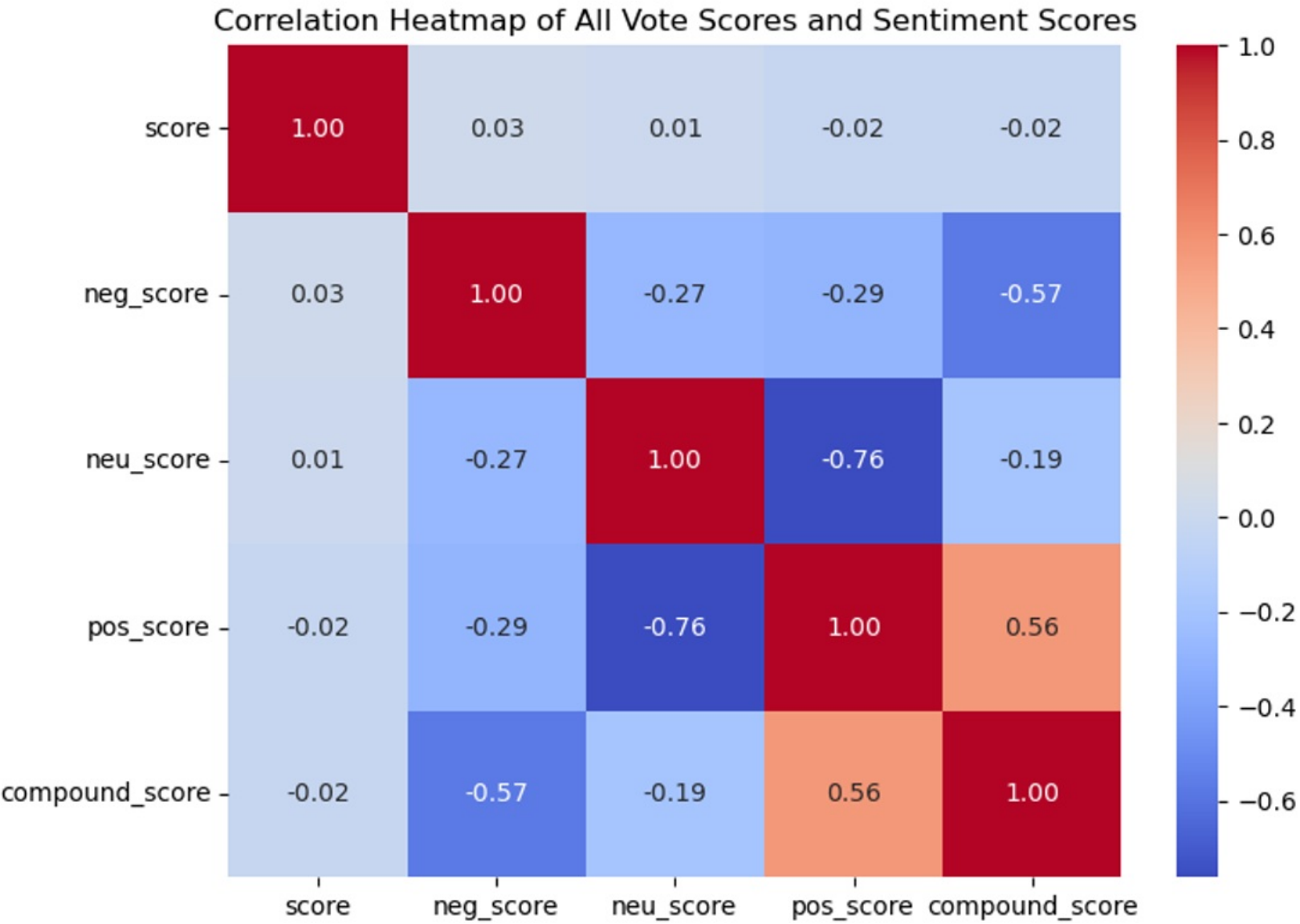
Title	Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA
-------	--------------	------------	----------	-------------------------------------	---------------------	--	--------------------------------------



Valence Aware Dictionary sEntiment Reasoner

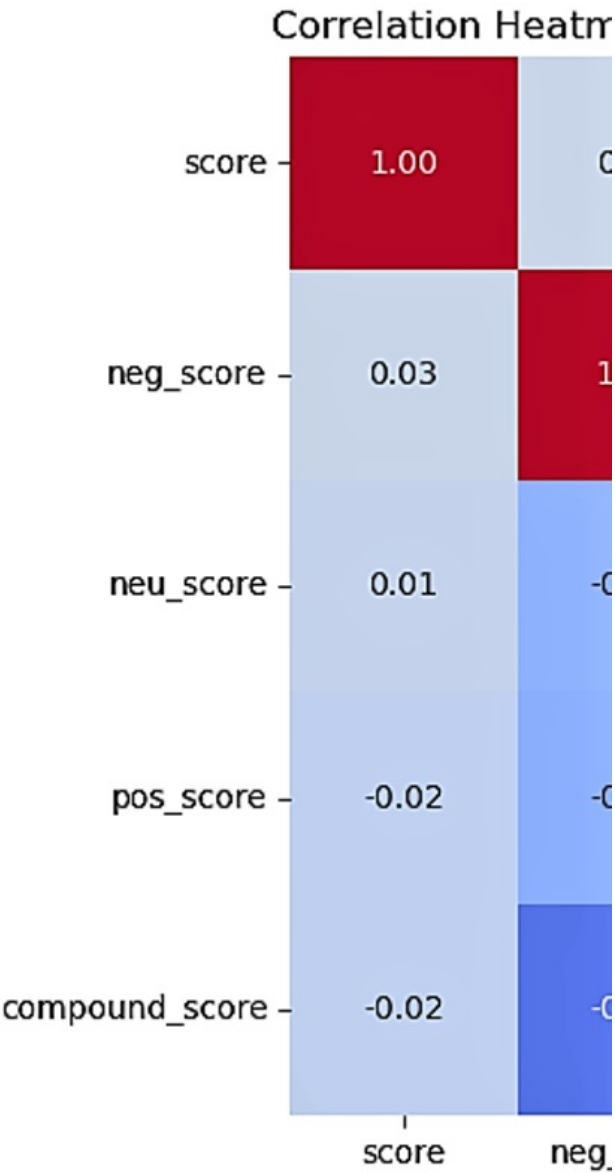
Demo Day Story

Title	Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA
-------	--------------	------------	----------	-------------------------------------	---------------------	--	--------------------------------------	-----



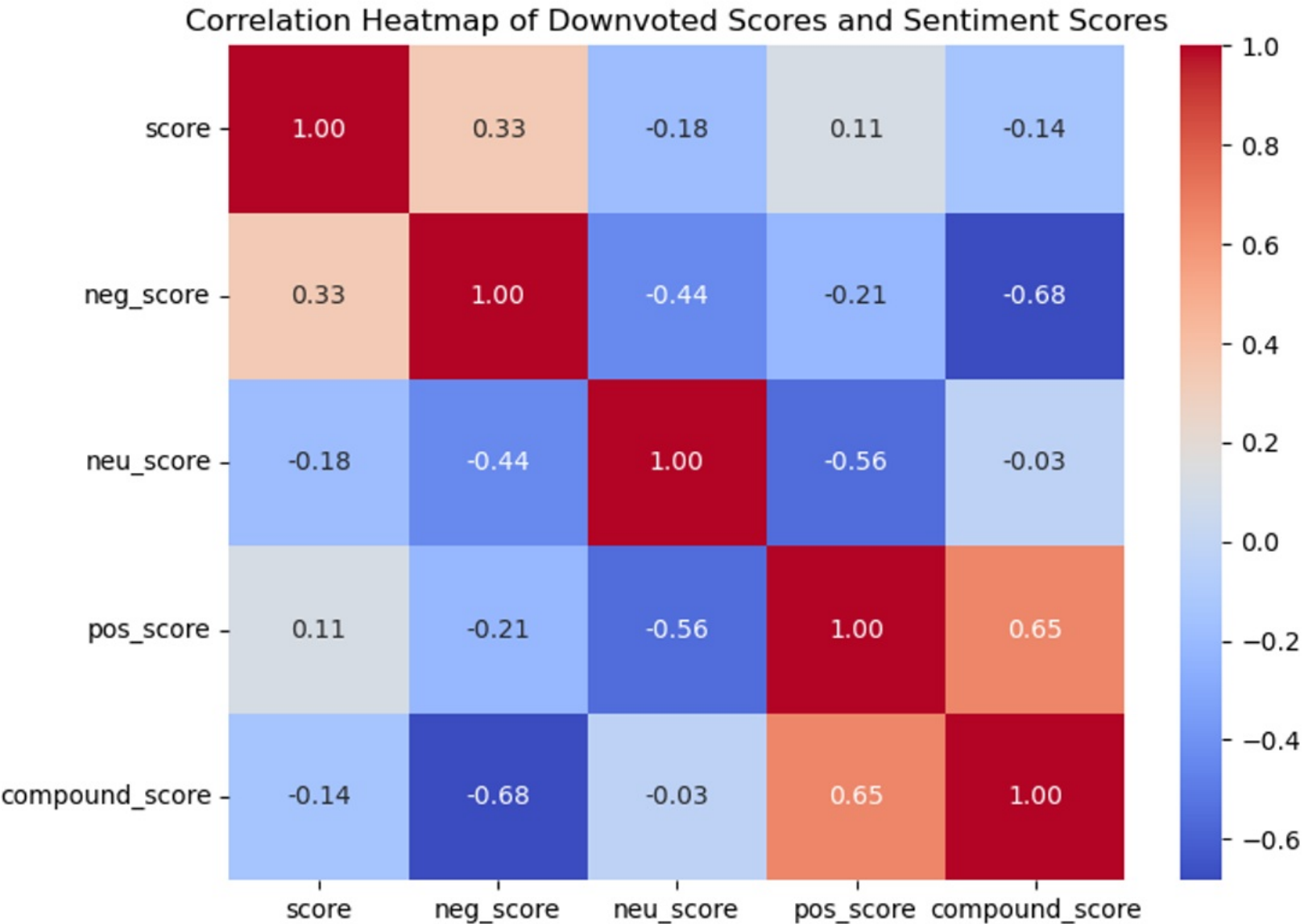
Demo Day Story

Introduction	Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA	9-topic pyLDavis
--------------	------------	----------	----------------------------------	---------------------	-------------------------------------	-----------------------------------	-----	------------------



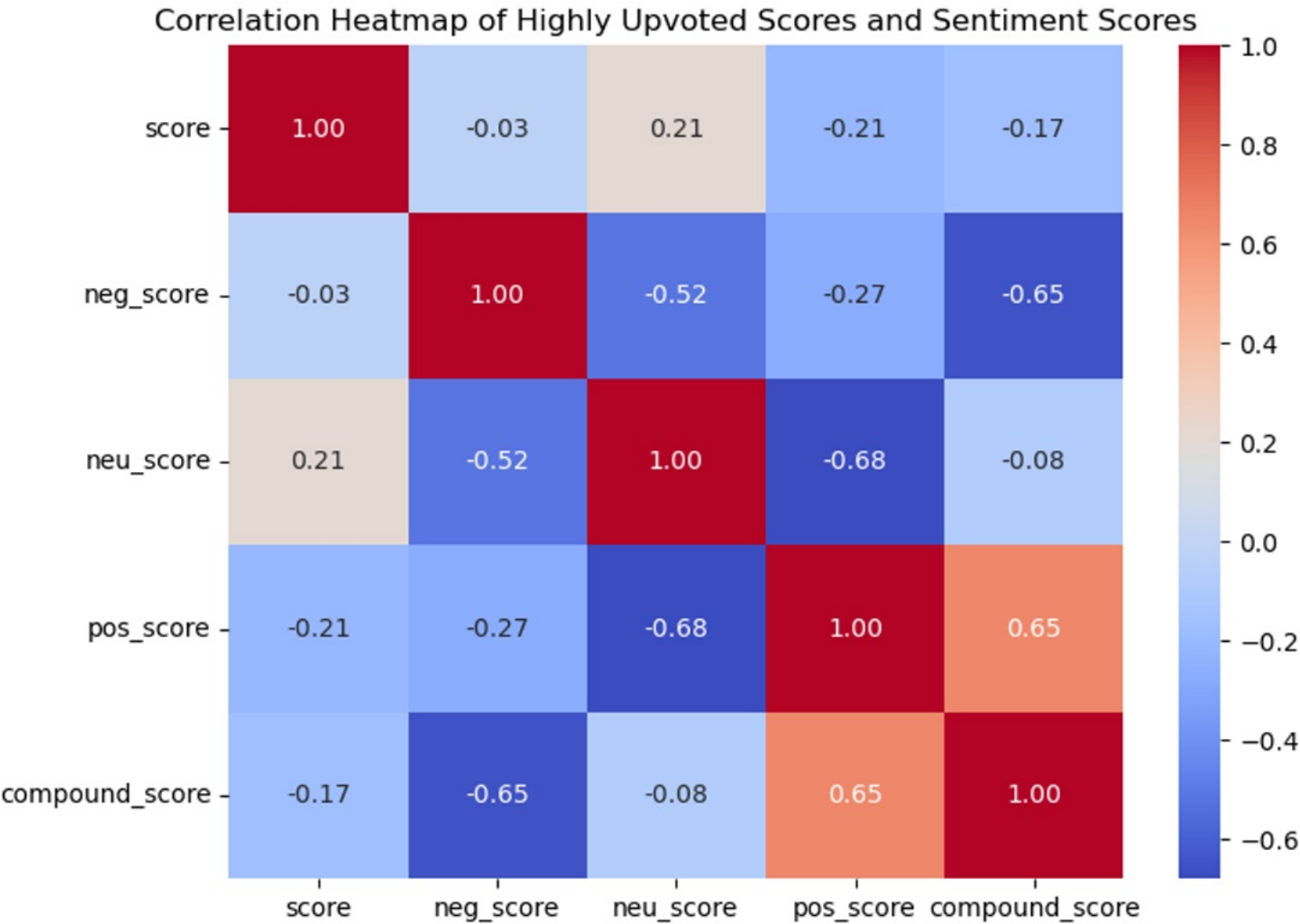
Demo Day Story

Word Cloud	VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA	9-topic pyLDavis	4-topic pyLDavis
------------	----------	----------------------------------	---------------------	-------------------------------------	-----------------------------------	-----	------------------	------------------



Demo Day Story

VADER SA	Correlations: RedditVotes and SA	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA	9-topic pyLDavis	4-topic pyLDavis	Topic pyLDavis ..
-------------	-------------------------------------	---------------------	--	--------------------------------------	-----	------------------	------------------	----------------------



Demo Day Story

Correlations: R..	Correlation Zoom-In	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA	9-topic pyLDAvis	4-topic pyLDAvis	Topic pyLDAvis Comparison	Key Take 1: Reddit ra..
-------------------	---------------------	-------------------------------------	-----------------------------------	-----	------------------	------------------	---------------------------	-------------------------

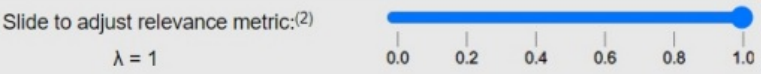


Latent Dirichlet Allocation

Demo Day Story

Correlat ion Zoo..	Correlations: Most Downvoted and SA	Correlations: Most Upvoted and SA	LDA	9-topic pyLDavis	4-topic pyLDavis	Topic pyLDavis Comparison	Key Take 1: Reddit range	Key Take 2: No correla..
-----------------------	--	--------------------------------------	-----	------------------	------------------	------------------------------	-----------------------------	-----------------------------

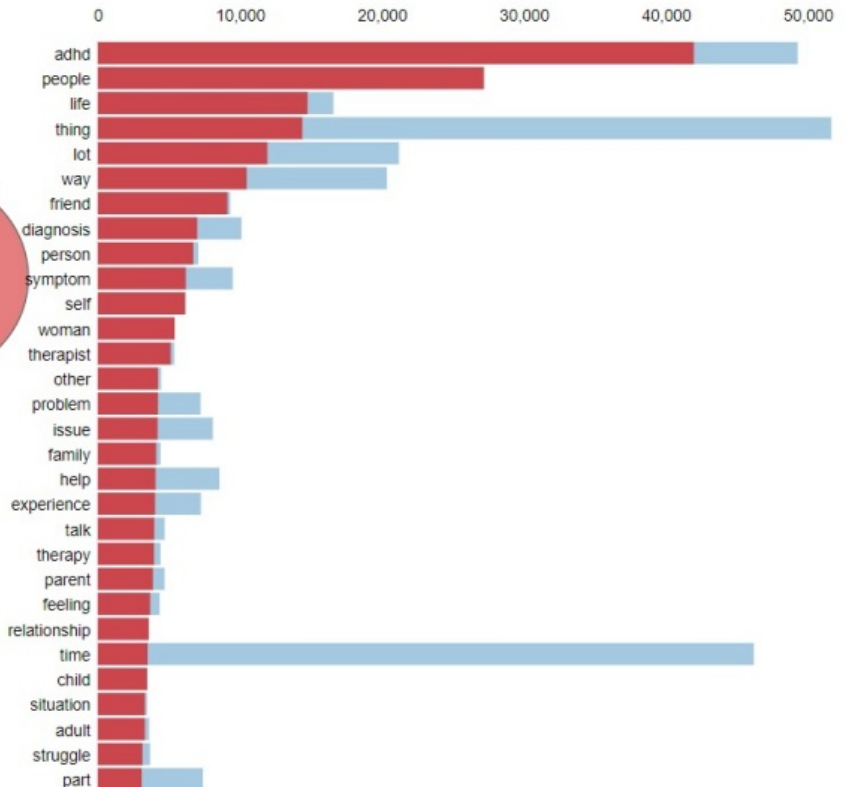
Selected Topic:



Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 1 (20% of tokens)



Demo Day Story

Correlations: M..	Correlations: Most Upvoted and SA	LDA	9-topic pyLDAvis	4-topic pyLDAvis	Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitativ..
-------------------	-----------------------------------	-----	------------------	------------------	---------------------------	--------------------------	----------------------------	--------------------------

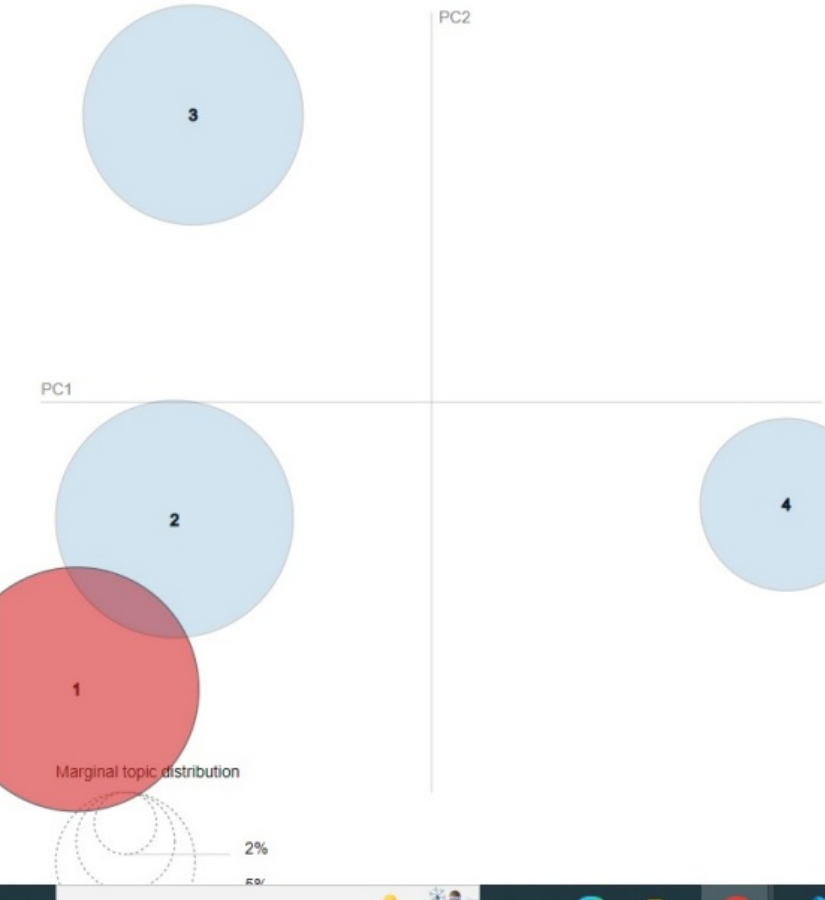
Selected Topic:

Slide to adjust relevance metric:(2)

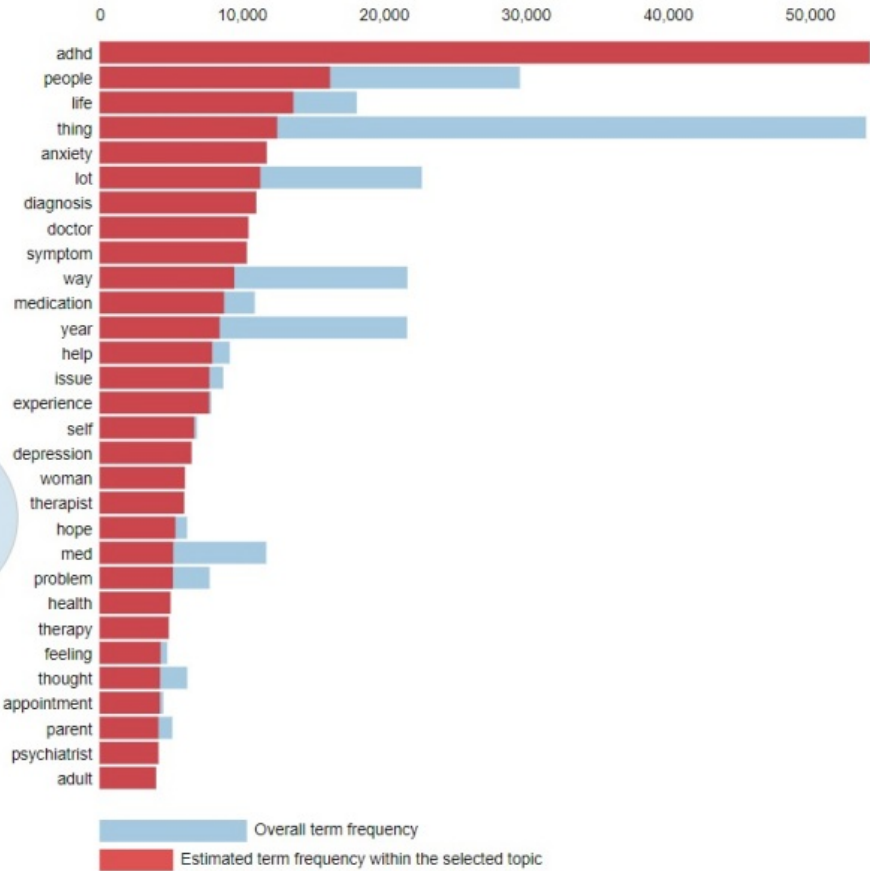
$\lambda = 1$

0.0 0.2 0.4 0.6 0.8 1.0

Intertopic Distance Map (via multidimensional scaling)



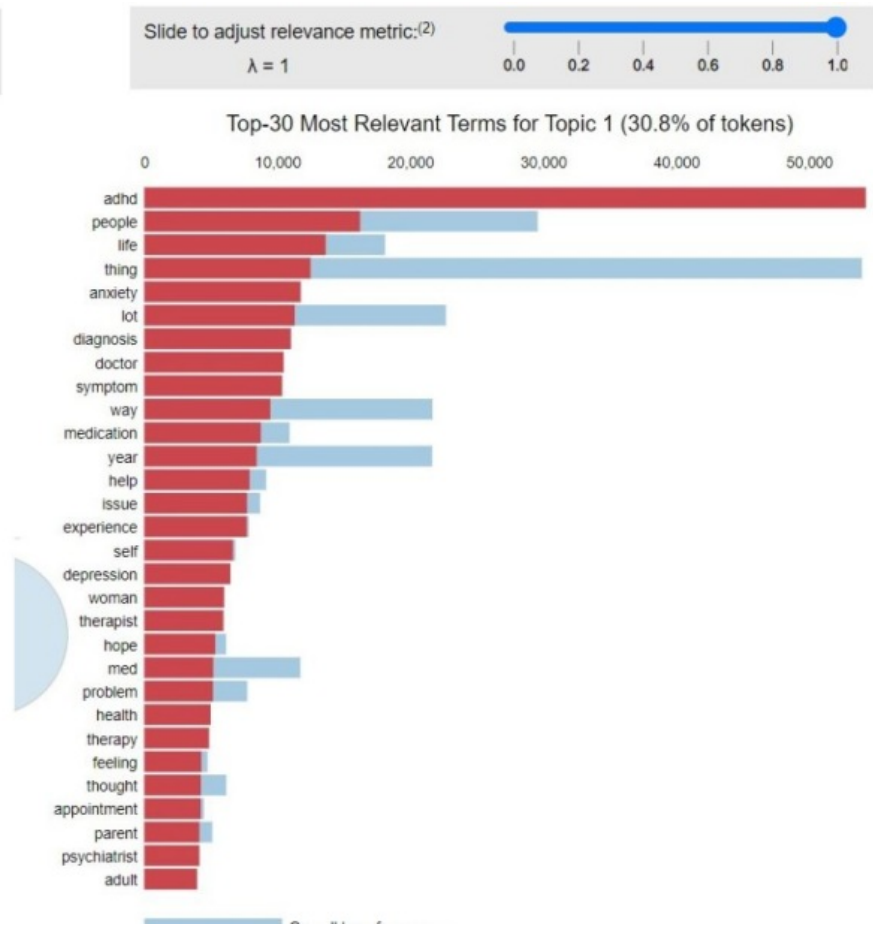
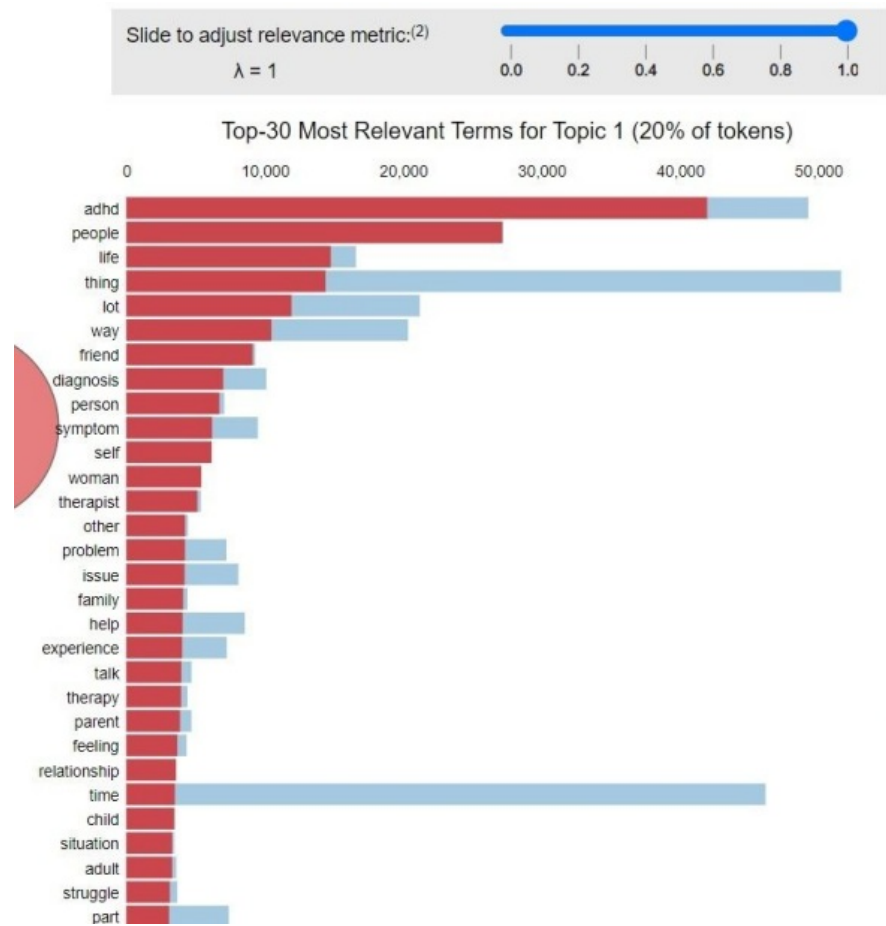
Top-30 Most Relevant Terms for Topic 1 (30.8% of tokens)



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)

Demo Day Story

Correlations: M..	LDA	9-topic pyLDavis	4-topic pyLDavis	Topic pyLDavis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-M..
-------------------	-----	------------------	------------------	---------------------------	--------------------------	----------------------------	------------------------------------	-------------------------



Demo Day Story

LDA	9-topic pyLDAvis	4-topic pyLDAvis	Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work
-----	------------------	------------------	---------------------------	--------------------------	----------------------------	------------------------------------	------------------------------------	-------------

Key Takeaways

- Redditors discuss a range of challenges and coping strategies in living with ADHD.

Demo Day Story

9-topic pyLDavis	4-topic pyLDavis	Topic pyLDavis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide
------------------	------------------	---------------------------	--------------------------	-----------------------------------	------------------------------------	------------------------------------	-------------	-----------

Key Takeaways

- Redditors discuss a range of challenges and coping strategies in living with ADHD.
- Positive or negative sentiment in a post is not correlated with how much readers will upvote or downvote it.

Demo Day Story

4-topic p yLDavis	Topic pyLDavis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide	Dashboard
----------------------	------------------------------	-----------------------------	-------------------------------	---------------------------------------	---------------------------------------	-------------	-----------	-----------

Key Takeaways

- Redditors discuss a range of challenges and coping strategies in living with ADHD.
- Positive or negative sentiment in a post is not correlated with how much readers will upvote or downvote it.
- Evaluating the results of topic modelling and sentiment analysis has qualitative, subjective aspects.

Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide	Dashboard
---------------------------	--------------------------	----------------------------	------------------------------------	------------------------------------	-------------	-----------	-----------

Key Takeaways

- Redditors discuss a range of challenges and coping strategies in living with ADHD.
- Positive or negative sentiment in a post is not correlated with how much readers will upvote or downvote it.
- Evaluating the results of topic modelling and sentiment analysis has qualitative, subjective aspects.
- Subject-matter expertise is integral to the evaluation.

Demo Day Story

Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide	Dashboard
---------------------------	--------------------------	----------------------------	------------------------------------	------------------------------------	--------------------	-----------	-----------

Future Work

- Domain-specific dictionaries for both topic modelling and sentiment analysis.
- Qualitative analysis of what changes with hyperparameter tuning.

Demo Day Story

Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide	Dashboard
---------------------------	--------------------------	----------------------------	------------------------------------	------------------------------------	-------------	-----------	-----------



Thank
you!



Demo Day Story

Topic pyLDAvis Comparison	Key Take 1: Reddit range	Key Take 2: No correlation	Key Take 3: Qualitative Evaluation	Key Take 4: Subject-Matter Exper..	Future Work	End Slide	Dashboard
---------------------------	--------------------------	----------------------------	------------------------------------	------------------------------------	-------------	-----------	-----------

