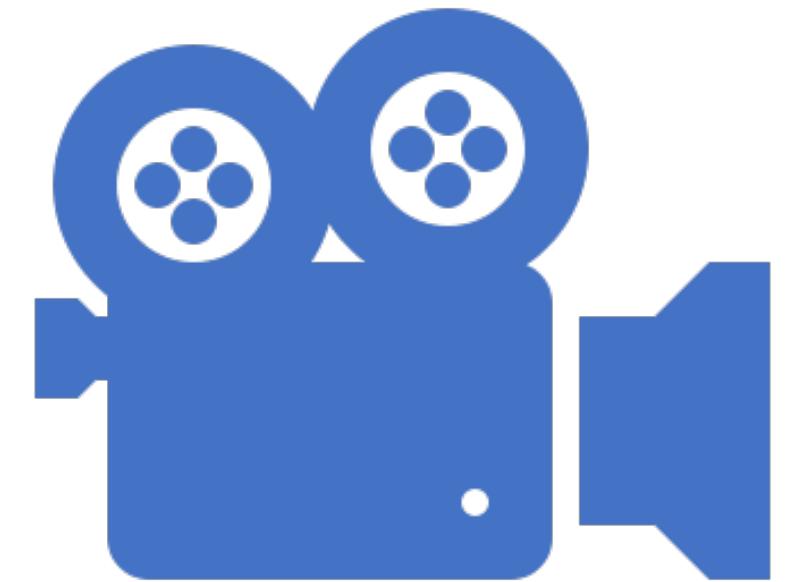


Movie Scripts

Marshal Will, Mikolaj Wieczorek



Introduction

Motivation: interesting insights of words common to specific movies

- Rarely/Commonly used words
- Clustering of Movies based on words
- Association between certain words
- Frequent Types of Emotions in Specific Movies
- Emotional Variance Over Time in a Movie

Database Source

- Source: Internet Movie Script Database (IMSDb)
- Pulled Movies based on specific Genre like Horror or Drama
- Categories Chosen: Action, Adventure, Comedy, Crime, Drama, Family, Fantasy, Horror, Romance, Sci-Fi, Thriller, and War



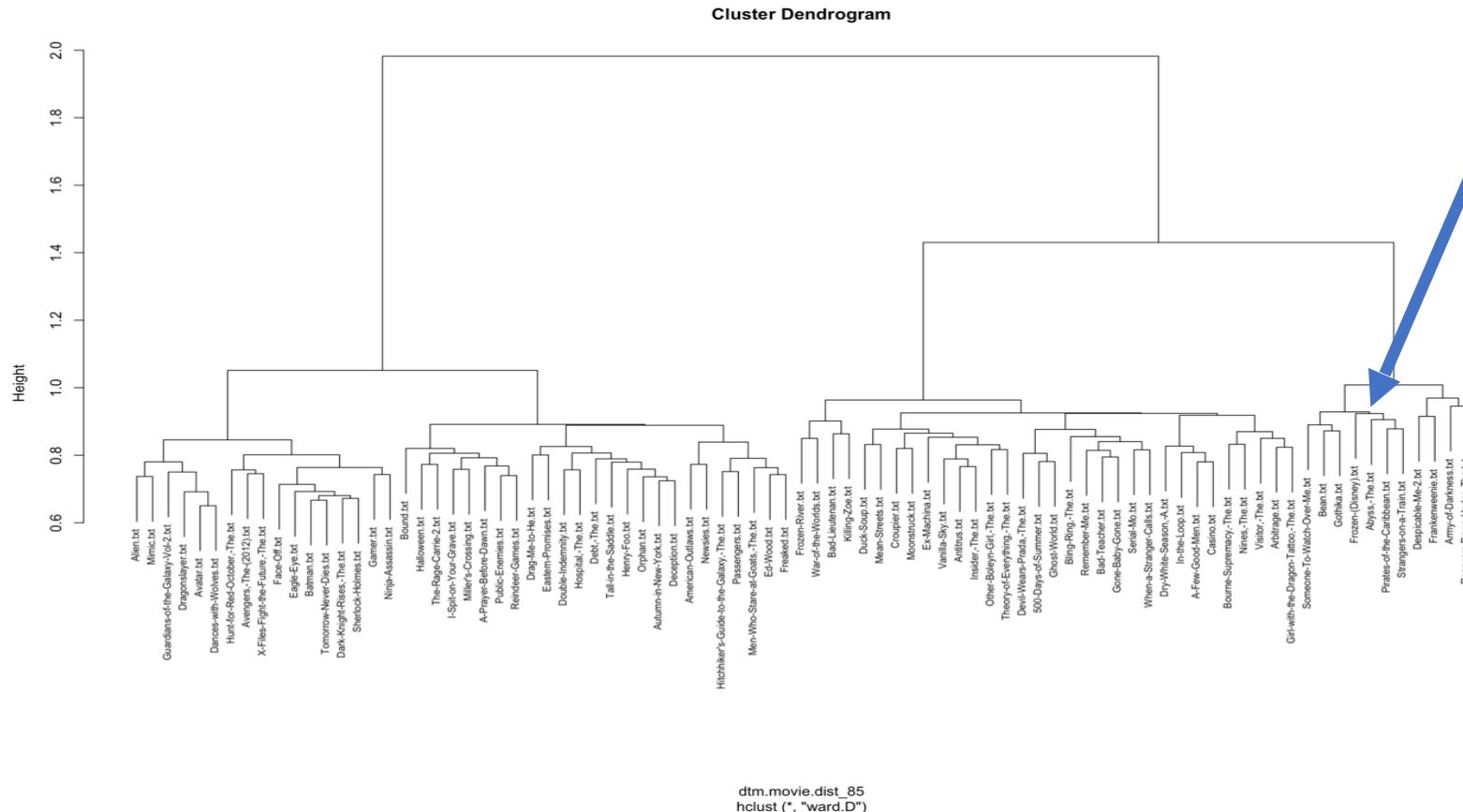
The web's largest
movie script resource!

Scrapping and Cleaning Processes

- Loop had to cycle through each HTML link of IMSDb (10,000+ movies!)
- Chose 900 movies and scraped from mentioned genres
- Spoken word script clean-up: extensive process
 - Removing “(CONTINUED)” or “Contd”
 - Removing descriptive parts of the script
 - Removing punctuation, etc.

bridge
mask
york
truck
bill
dances
lieutenant
trees
kicking
nurse
same
gold
code
river
horse
prison
restaurant
deck
outer
shape
charlie
stage
progress
video
parking
wine
fist
dawn
general
bird
beer
king
summer
mary
pause
frank
agent
peter
sister
guards
cabin
standard
court
boat
agent

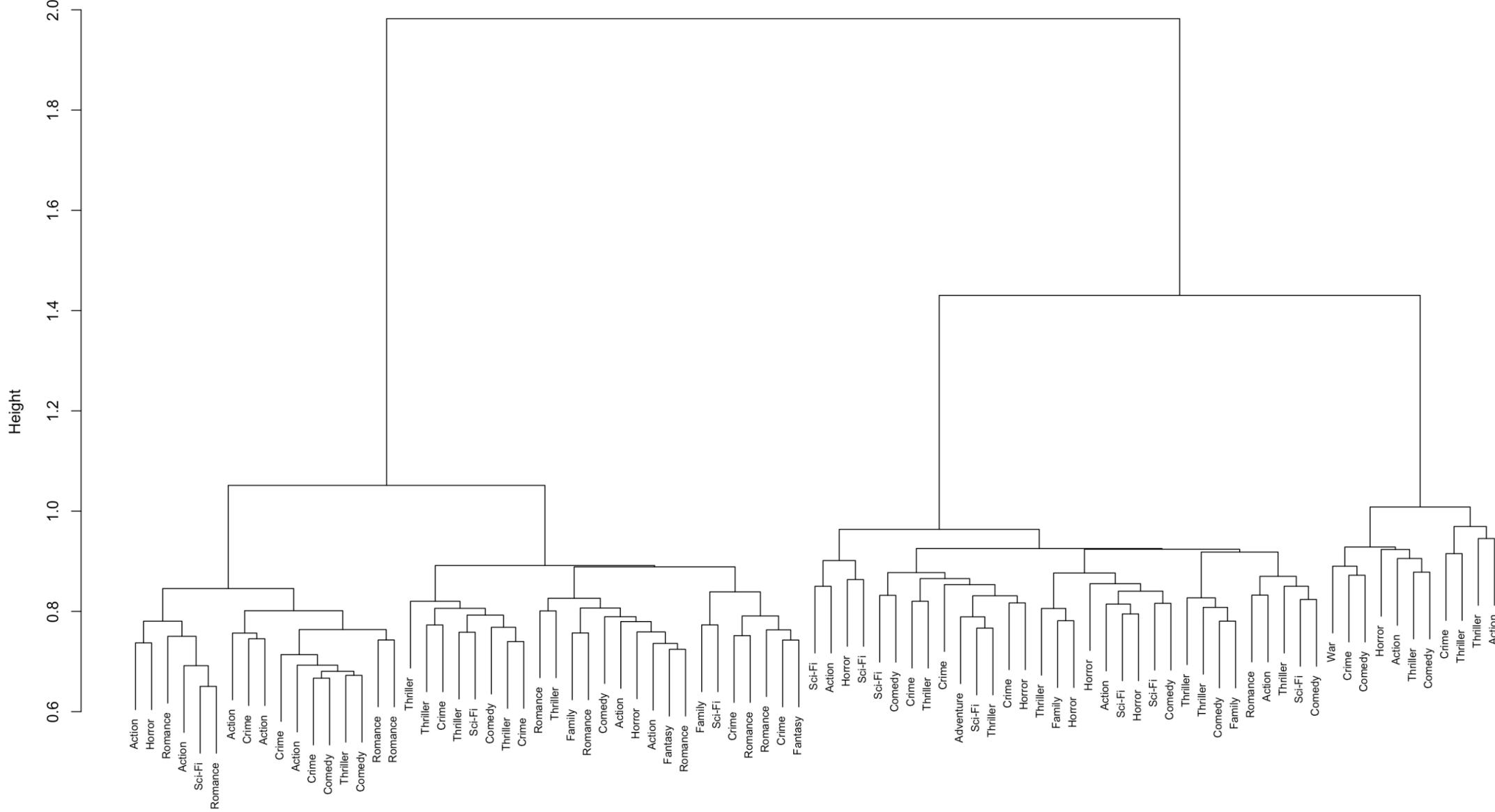
Dendrogram Movie Clusters



Interesting Cluster Disney Movies

- Despicable Me
- Frozen
- Pirates of the Caribbean

Cluster Dendrogram

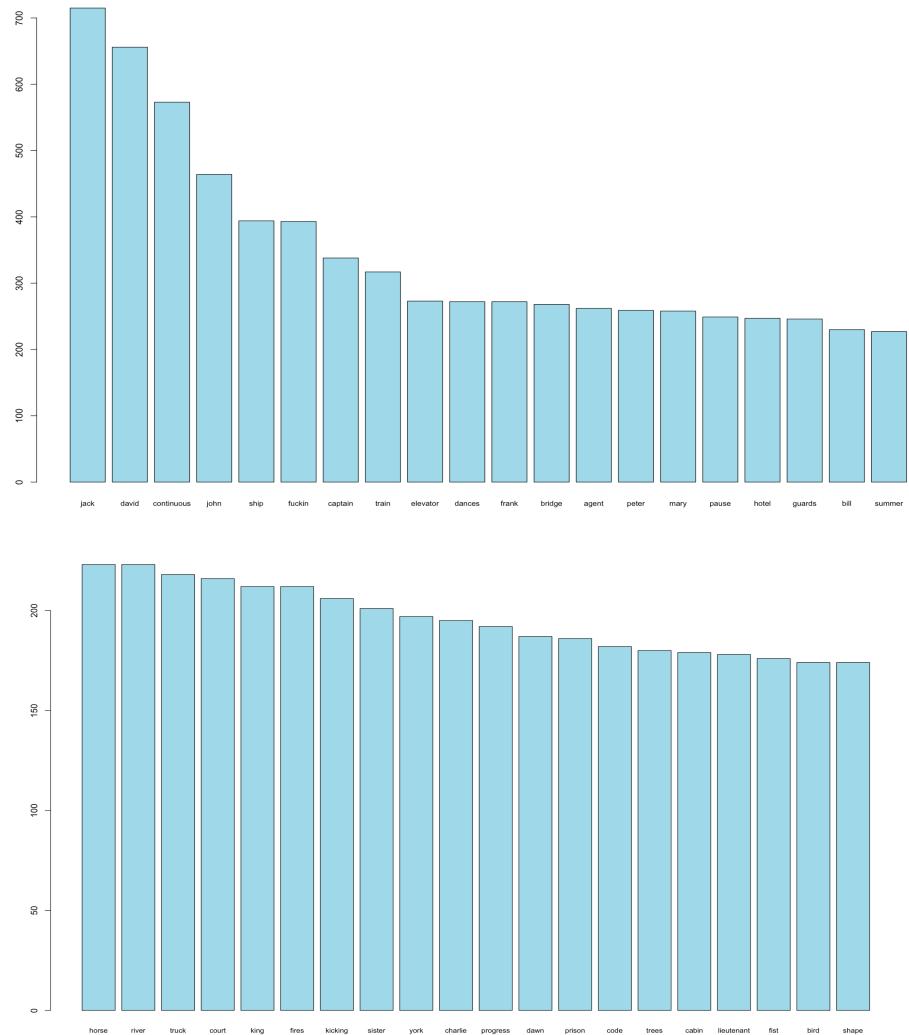


dtm.movie.dist_85
hclust (*, "ward.D")

Clustering: K-means

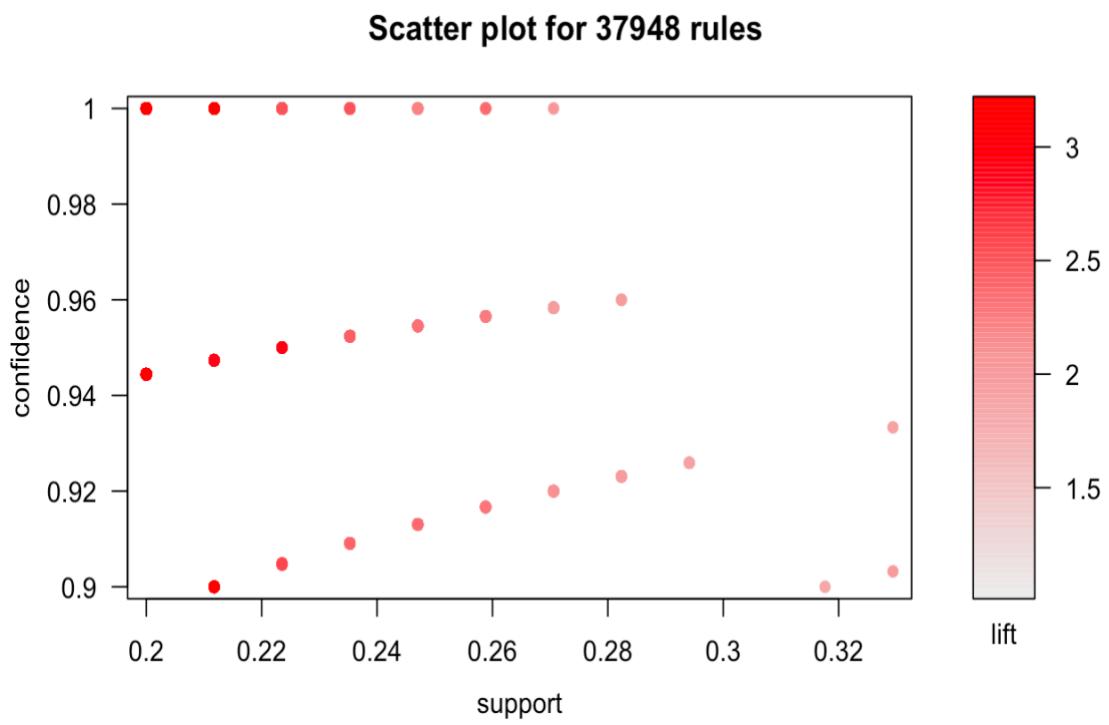


	Action	Adventure	Comedy	Crime	Family	Fantasy	Horror	Romance	Sci-Fi	Thriller	War	
1	3		1	4	3	2	0	5	1	7	6	0
2	7		0	4	7	2	2	2	9	3	6	0
3	2		0	2	2	0	0	0	1	0	0	3



Word	Frequency
jack	715
david	656
continuous	573
john	464
ship	394
fuckin	393
captain	338
train	317
elevator	273
frank	272
bridge	268
agent	262
peter	259
mary	258
pause	249
hotel	246
guards	246
bill	230
horse	223
river	223
truck	218
court	216
fires	212
kicking	206
sister	201
york	197
charlie	195
progress	192
dawn	187
prison	186
code	182
cabin	179
lieutenant	178
fist	176
nurse	173
video	171
restaurant	170
bird	174
deck	169
gate	167
boat	166
parking	166
tony	165
general	163

Association Rules



- A lot of words were associated with “fuckin”

Waitress → Fuckin, Friday → Fuckin,

Society → Fucked, Guilty → Fucked

- Other common associations

Assistant, Pockets → Elevator

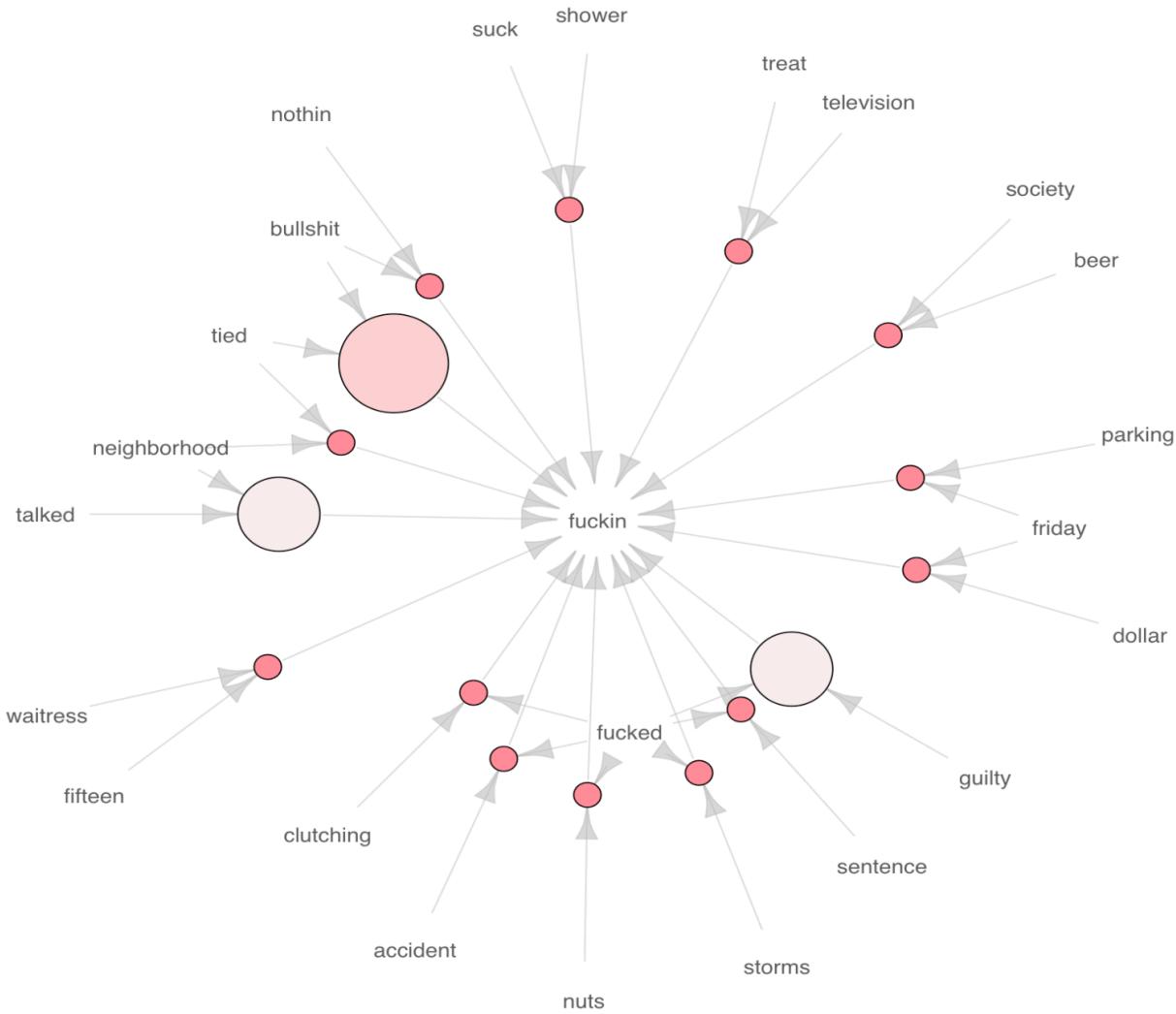
Yelling → Train

Sides, Awkwardly → Dancing

rules	support	confidence	lift
{gear,shifts,dozen,cracks} => {tunnel}	0.18	0.94	3.32
{reports,route} => {captain}	0.12	0.91	3.09
{floating,surface} => {ship}	0.14	0.80	3.09
{directly,floating} => {ship}	0.14	0.80	3.09
{waitress,fifteen} => {fuckin}	0.12	0.83	2.95
{familiar,agents} => {agent}	0.15	0.87	2.73
{elevator,bullet,struggle} => {agent}	0.15	0.87	2.73

Graph for 16 rules

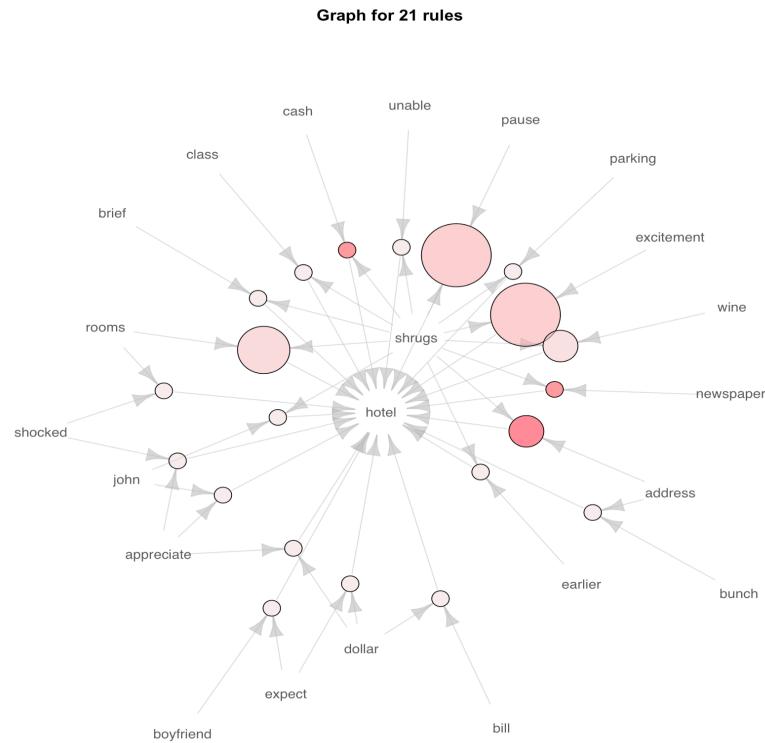
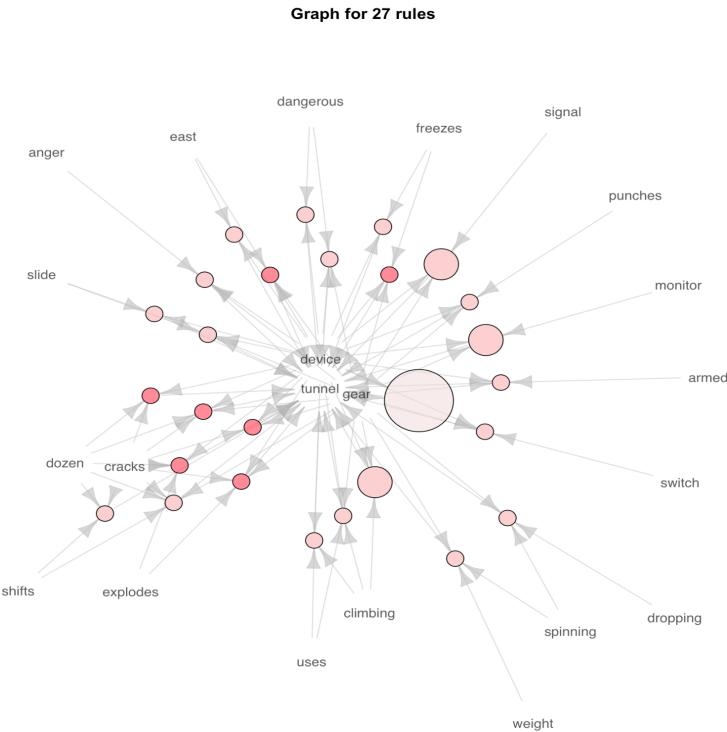
size: support (0.118 - 0.153)
color: lift (2.833 - 2.951)

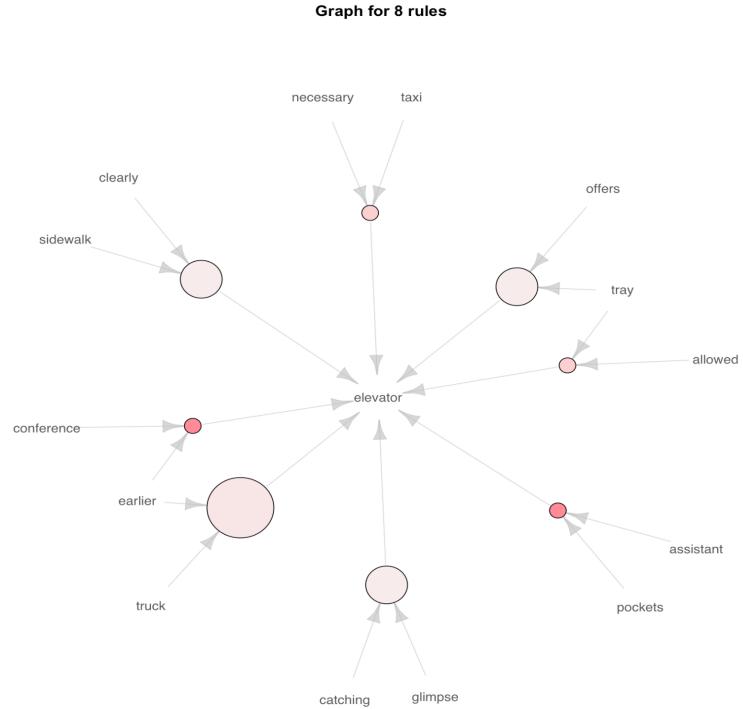
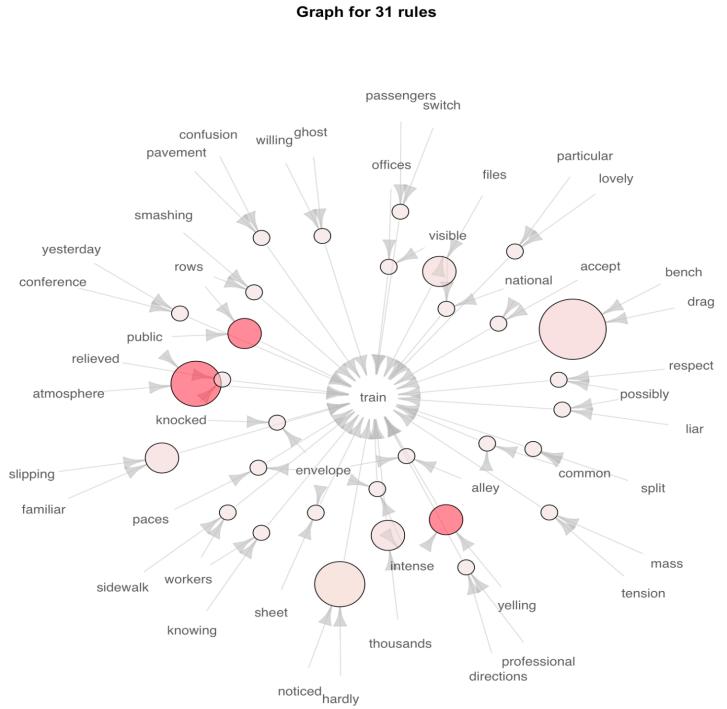


Graph for 10 rules

size: support (0.141 - 0.165)
color: lift (3.091 - 3.566)



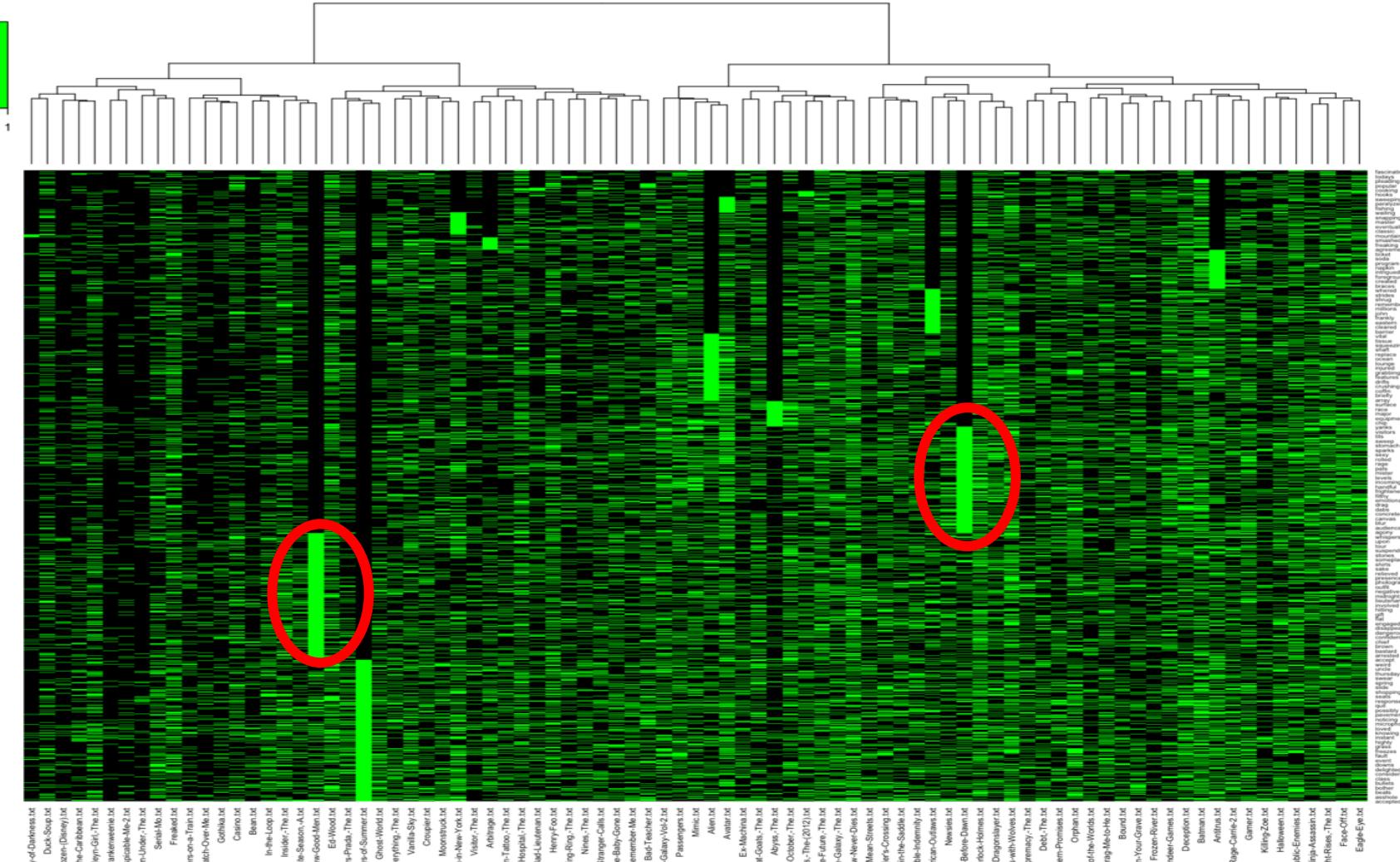




More Findings: Associations

- Lots of words are around Hotel like
 - Boyfriend → Hotel
 - Shocked → Hotel
 - Lunch → Hotel
- Hotel and Shrugs were strongly related
 - Newspaper → Hotel
 - Shrugs, Excitement → Newspaper

Heatmap: Only clustering movies



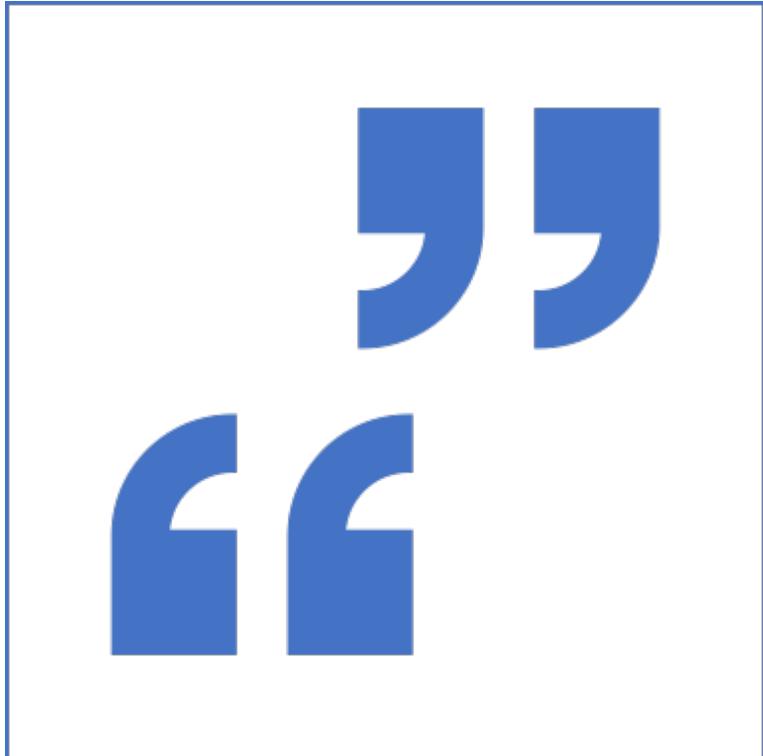
Chose these two movies for further sentiment analysis:

A Few Good Men

&

A Prayer Before Dawn

Short Story Summary



- A Few Good Men
 - The film revolves around the court-martial of two U.S. Marines charged with the murder of a fellow Marine and the tribulations of their lawyers as they prepare a case to defend their clients.
 - Fun movie quote “You can’t handle the truth!”
- A Prayer Before Dawn
 - The true story of an English boxer incarcerated in one of Thailand's most notorious prisons as he fights in Muay Thai tournaments to earn his freedom.

their prisoners
prison cigarettes watches
chains tries boxers shouts
watching around moment moore
through other heroin
doesn't takes after empty
again watching walks
watching smiles empty
trustees hands
stares alone there
before boxing floor
music another behind
diesel blood keeps
stands guards
guard

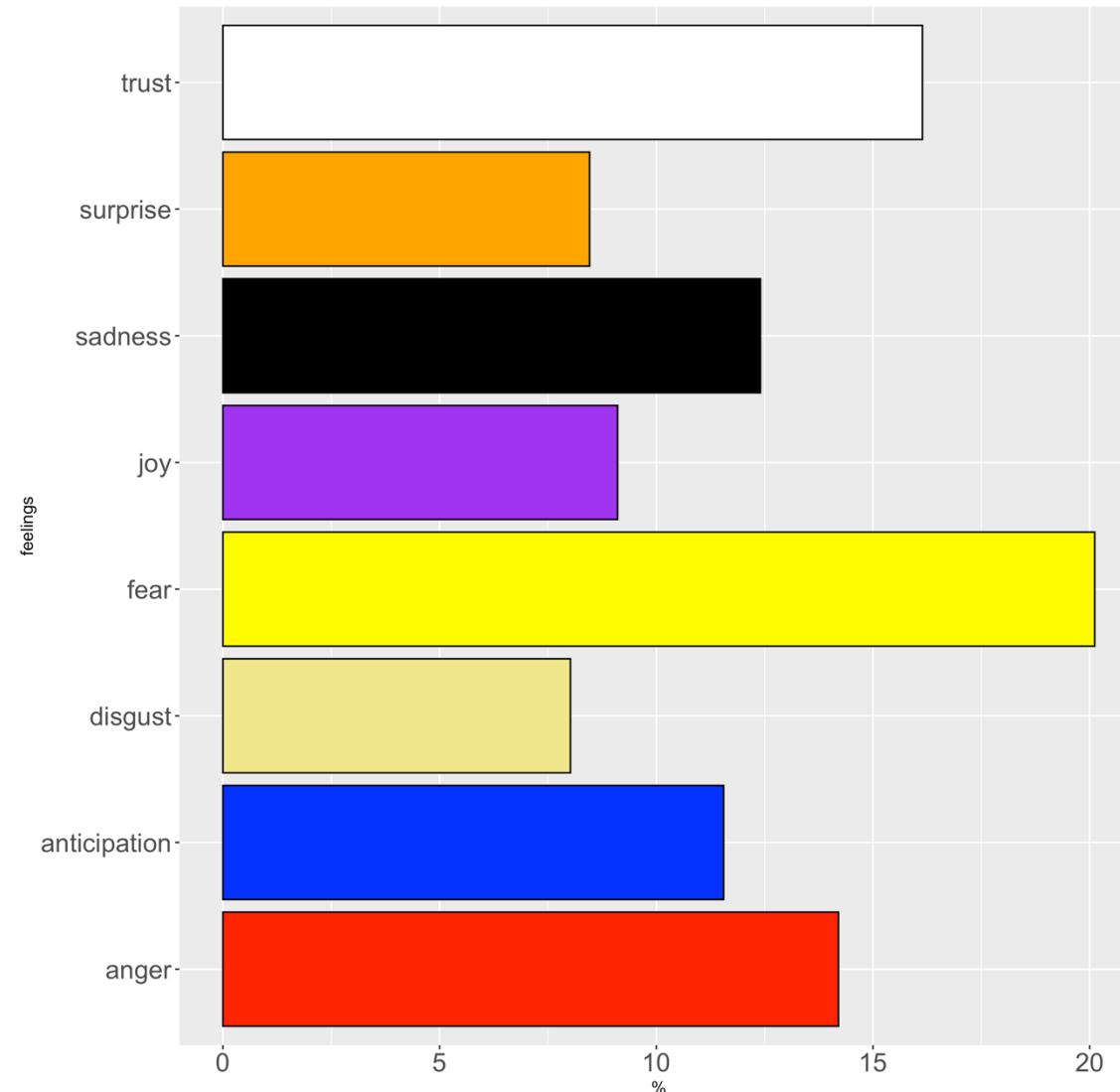
A Prayer Before Dawn

order defense because commander
private courtroom where
think thank before captain
colonel something please
Kaffees going wasnt stand
doesnt around morning
corps barracks three jessep
markinson santiagos other
their transfer being lieutenant
court first witness didnt lawyer
corporal downey would night
right kendrick moment
marine thats
dawson you're

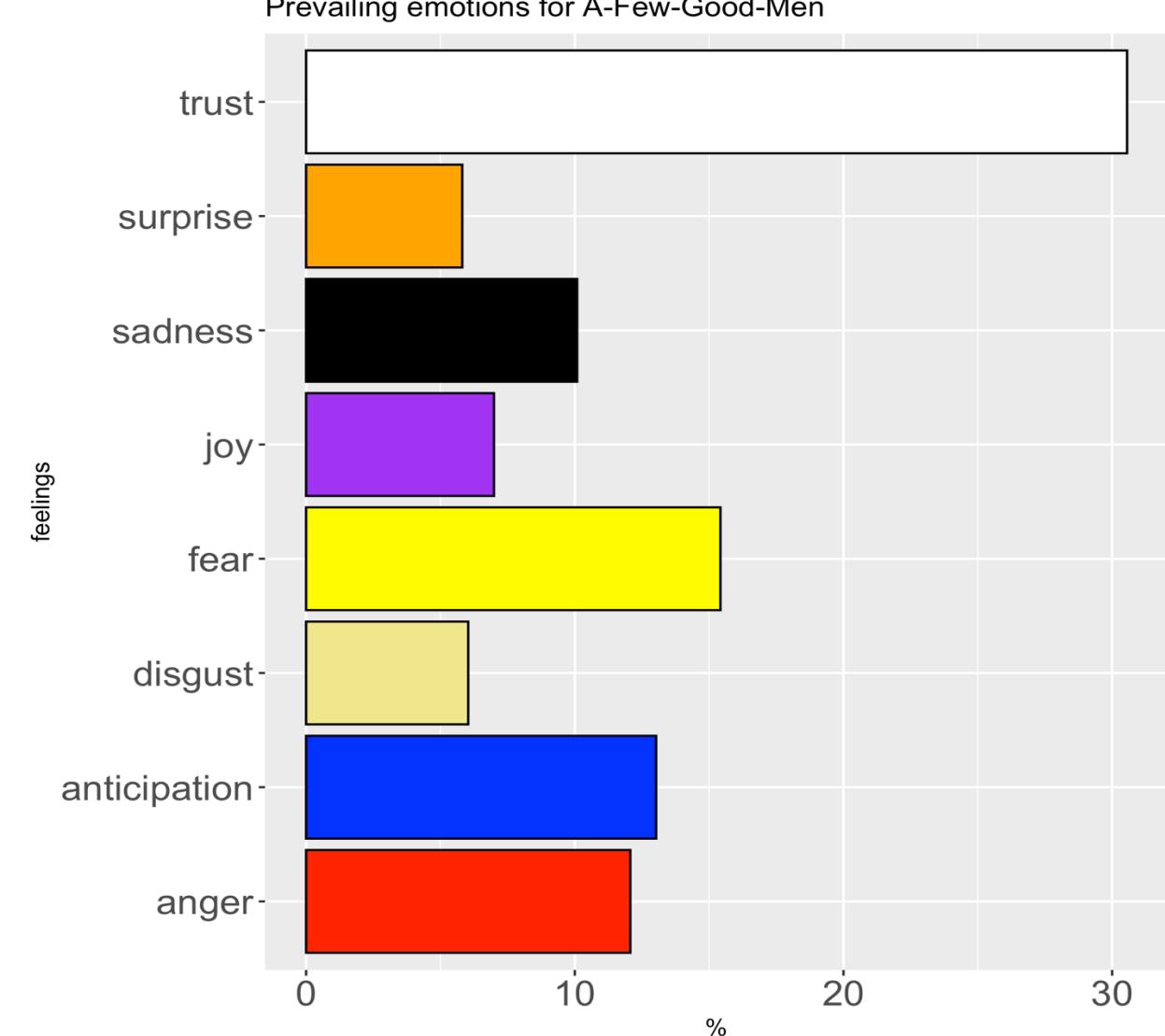
A Few Good Men

Comparing Emotional Sentiment

Prevailing emotions for A-Prayer-Before-Dawn

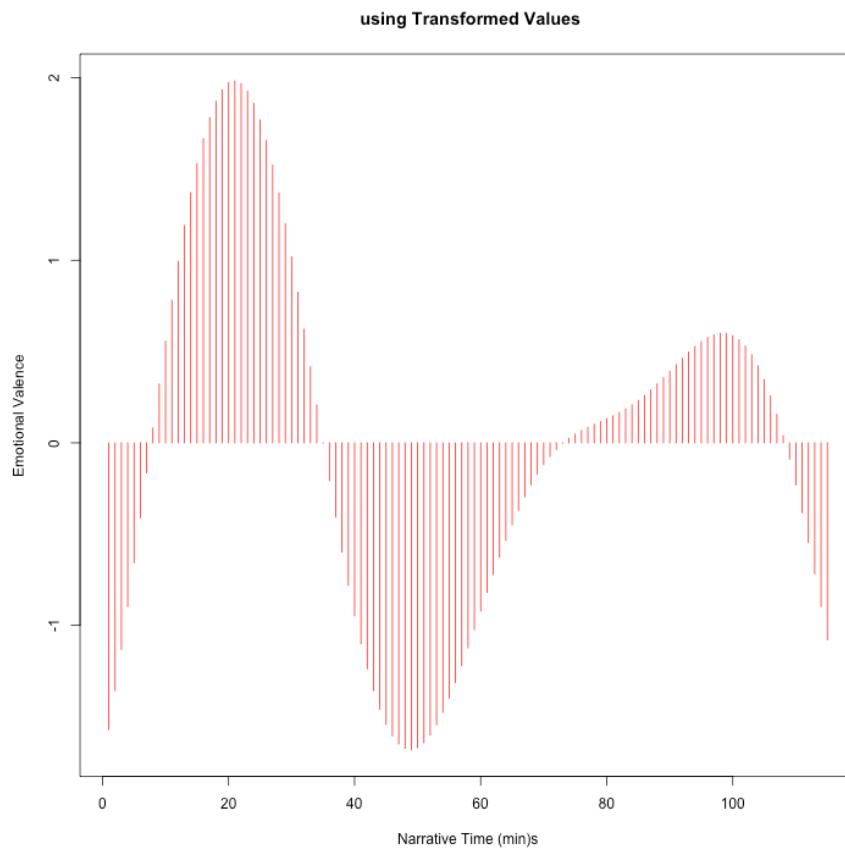


Prevailing emotions for A-Few-Good-Men

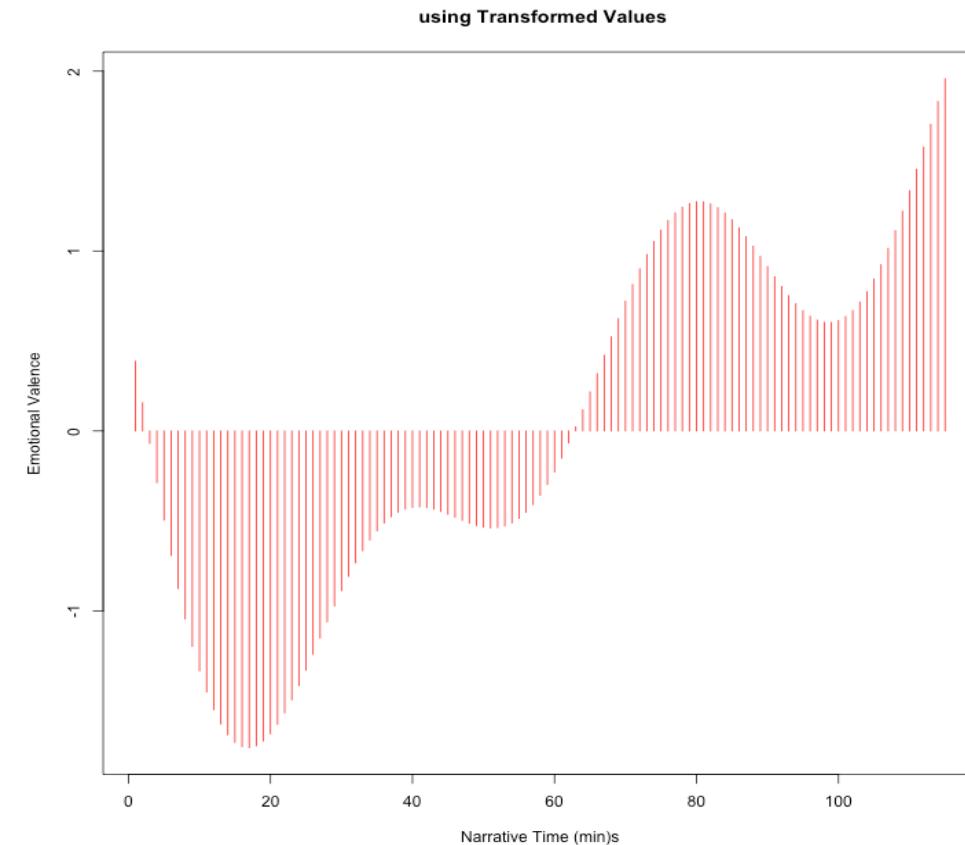


Emotional Trajectory

Above 0 → Positive Emotions
Below 0 → Negative Emotions



A Few Good Men



A Prayer Before Dawn



Questions?