

Disney Movies Through the Years:

A Text-Mining Analysis of Language, Sentiment, and Themes in The Lion King, Up, and Zootopia

Yechen Cao, Xianya Fu, Qianping Wu

Table of contents

- 01 Introduction
- 02 Word Analysis
- 03 Sentiment Analysis
- 04 Latent Dirichlet Allocation: LDA
- 05 Conclusion & Limitations



01

Introduction

Motivation

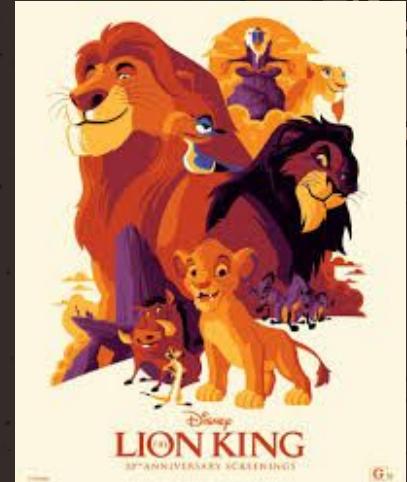
We grew up with Disney animated films and always have special feelings for Disney stories that are full of magic and wonder. As time passes, we can see noticeable changes in the characters, stage designs, and even songs of Disney movies. This evolution reflects Disney's commitment to staying relevant to contemporary audiences, adapting narratives to suit their tastes [1].

As Disney movie enthusiasts with text analysis skills, we want to know how Disney's content style has changed over time.



Data Introduction

- **Research question:** How have the language usage, sentiment expression, and thematic focus in Disney animated movies evolved over time?
- We chose to conduct a text-mining study of three productions from different periods: The Lion King (1994)[2], Up (2009)[3], and Zootopia (2016)[4]. We found the scripts of these three movies online, including environmental descriptions and third-person narration, which are about 100 pages for each. We also performed a data-cleaning process on these texts to prepare them for further analysis.



Data Cleaning

1. Simple Transforms
2. Conversion to Lowercase
3. Remove Numbers
4. Remove Punctuation
5. Remove English Stop Words
6. Remove Own Stop Words
(Including characters' names and special terms in scripts)
7. Strip Whitespace
8. Stemming

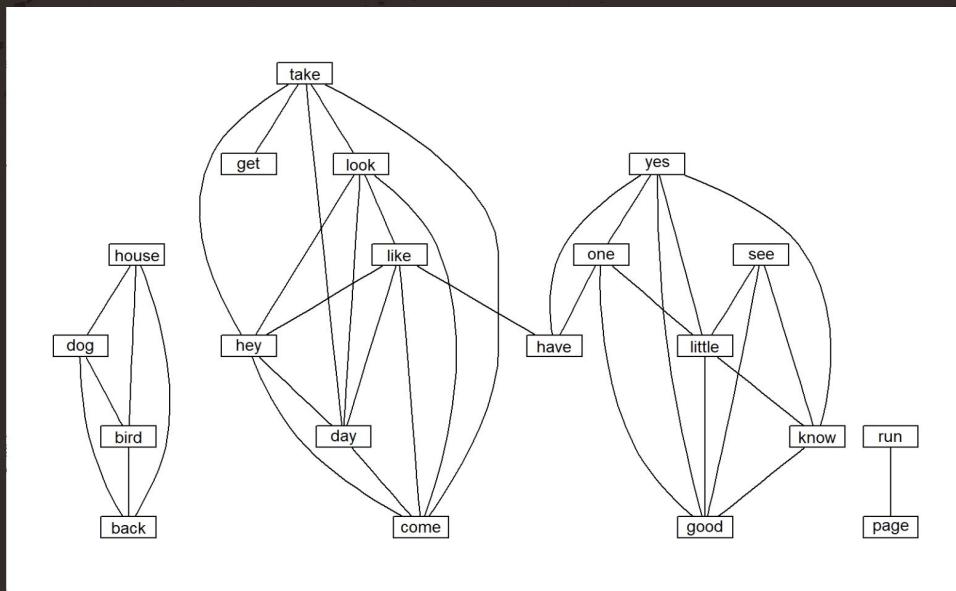
* We also noticed that movie scripts often use EXT. (Exterior) and INT. (Interior) to indicate the context for visualization. Each occurrence of EXT. or INT. signals the beginning of a new scene, which effectively marks the transition to a new segment of the film. Therefore, we used these markers to divide the three movies into chapters.

02

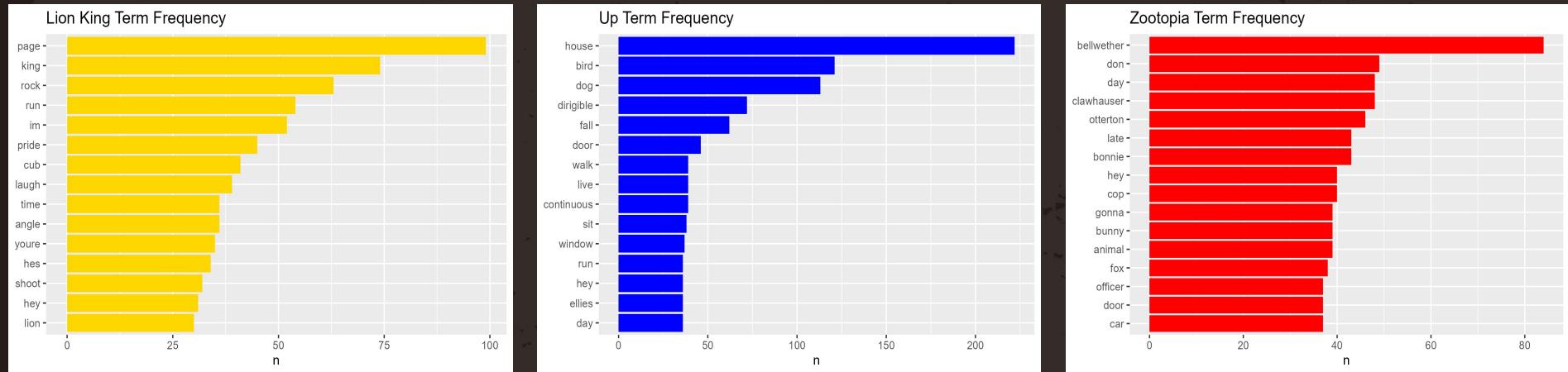
Word Analysis

Association between Terms

Lowest frequency = 100, correlation threshold = 0.9



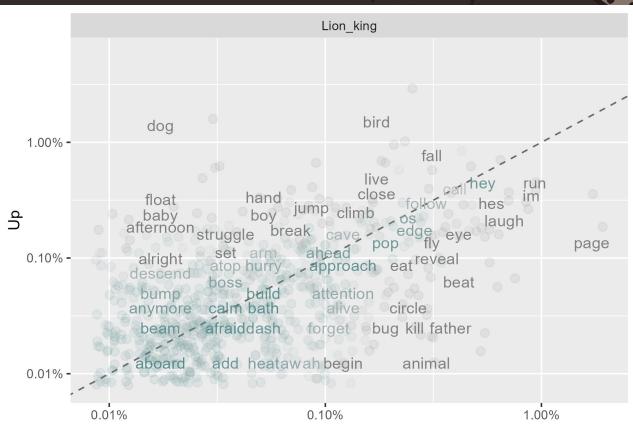
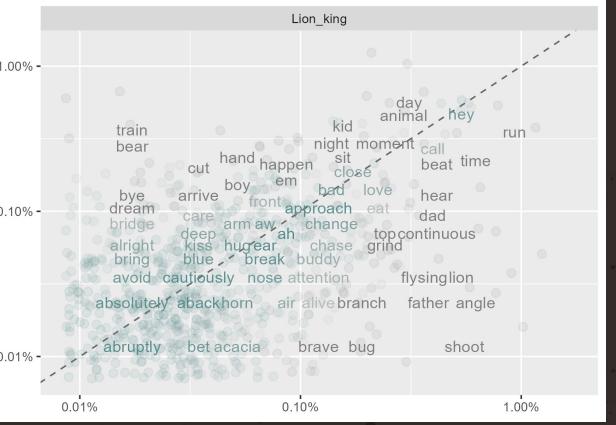
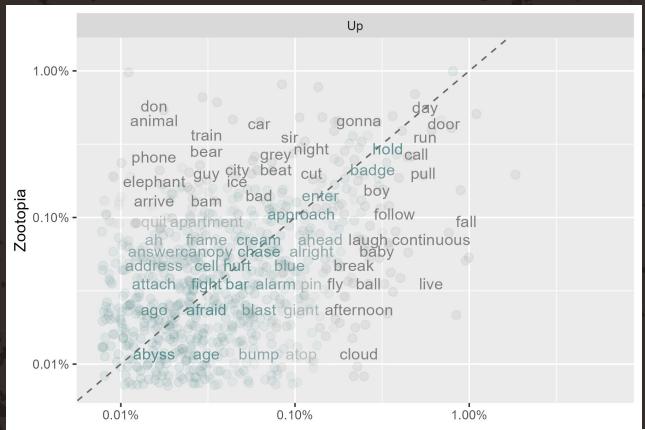
Term Frequency



TF-IDF

Movie	word	n	ft	idf	tf-idf
Up	house	222	0.034	1.099	0.038
Lion King	king	74	0.013	1.099	0.014
Zootopia	enter	84	0.012	1.099	0.013
Up	dirigible	72	0.011	1.099	0.012
Lion King	pride	45	0.008	1.099	0.009
Lion King	cub	41	0.007	1.099	0.008

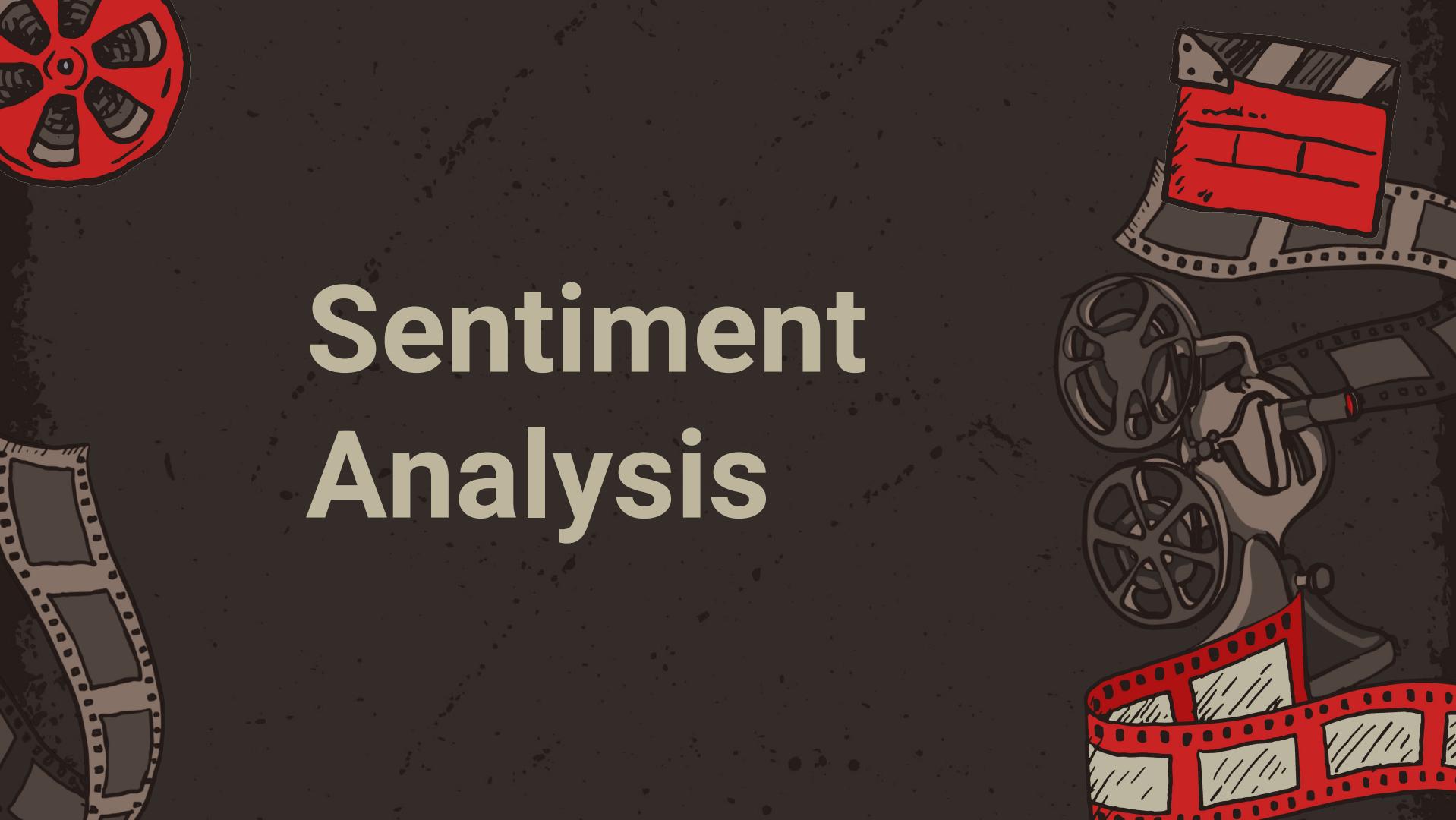
Unique Words



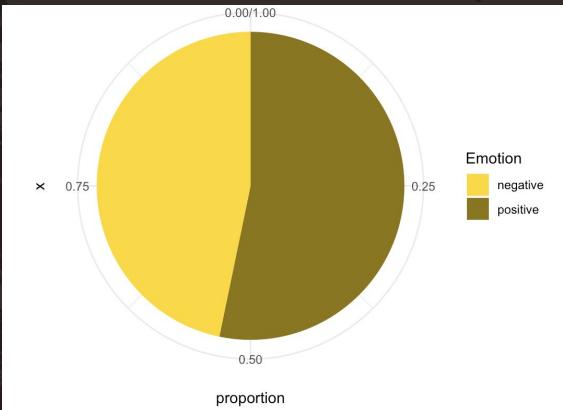
Correlation between 3 Movies

Comparison	p-value	95% CI	Estimated Correlation
Lion King vs. Up	< 2.2e-16	(0.3006982 0.4294261)	0.3668
Lion King vs. Zootopia	< 2.2e-16	(0.4329664 0.5426022)	0.4897
Up vs. Zootopia	< 2.2e-16	0.3943091 0.5079980)	0.4530

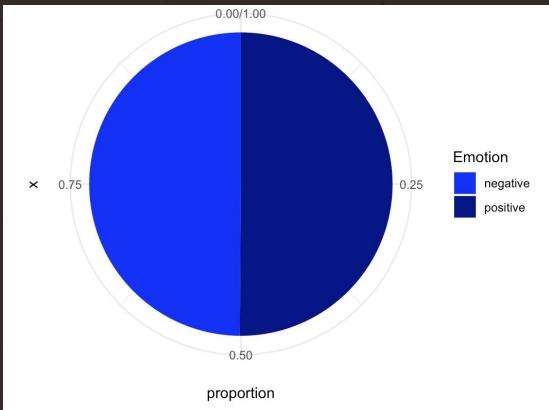
Sentiment Analysis



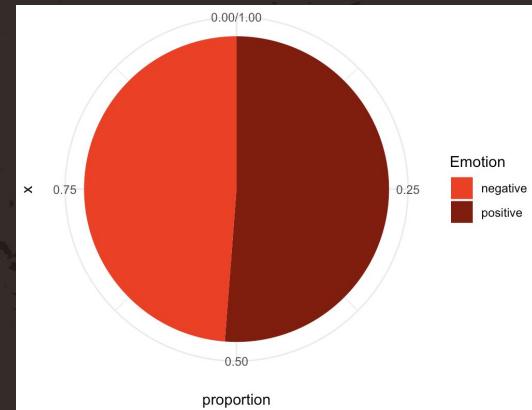
Proportion of Positive/Negative Sentiment



Lion King

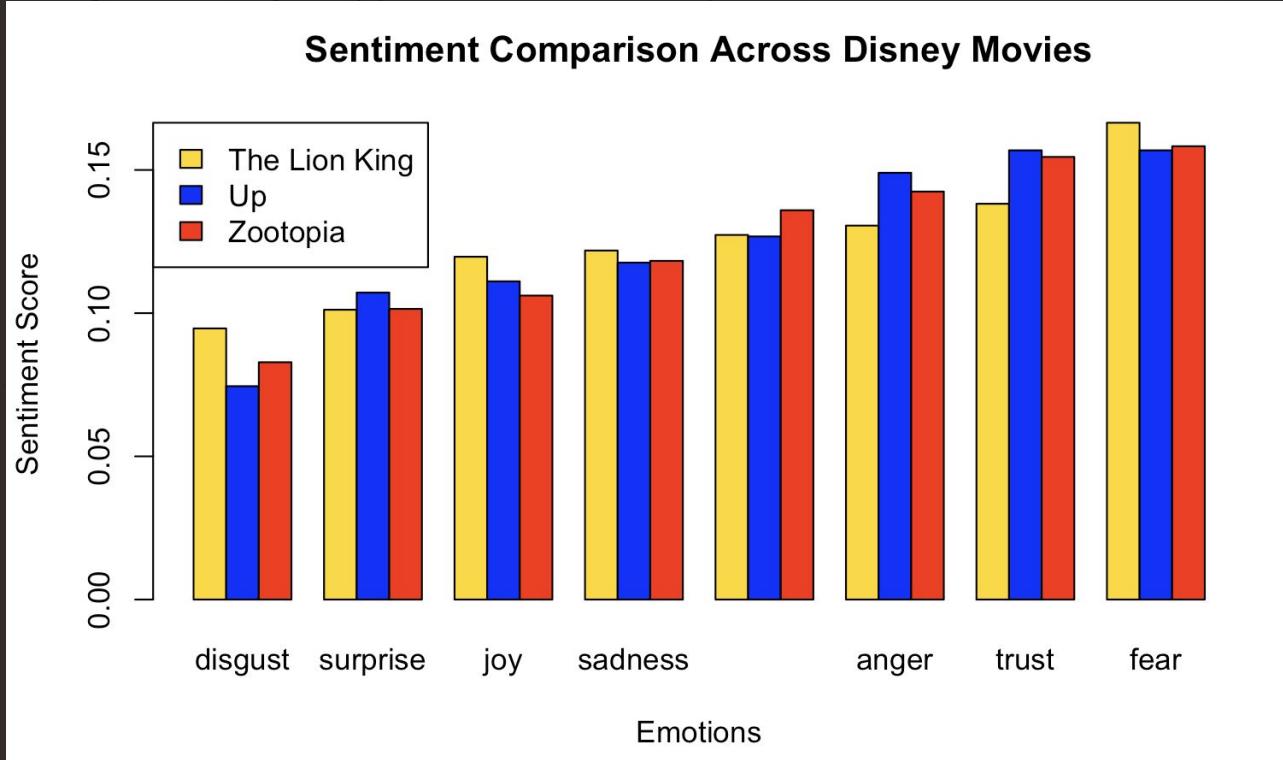


UP



Zootopia

Barchart for Sentiments



Wordcloud



prepare elephant
whisper frame time
doesn't remember
shenzi catch branch
sire understand banzus
grind day leap
hakuna grind branch
cub ed day leap
run pride land
lionhead son bird kill sound eye
stop rise hill fly huge father his
circle hold seq exit
life sun im death animal tree
sky fall continue stick
reveal free hyena scream
edge eat leave matata moment
move life sing call what's your rear stand
edge sing call what's your roar stand
move life sing call what's your roar stand
edge sing call what's your roar stand
shoot hear love shadow
wildebeest shadow simbas
paw angle laugh beat
continuous scar walk
jungle

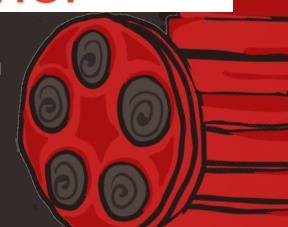
Lion King



master dirigible balloon russells fall chocolate
adventure imvoice hall
talk laugh imvoice hall
explorer leave
front alpha carls hand leave
wait foot collar grab
beta door grind paradise
you're watch climb hear day kid ball
baby morning timesit jump tie
steer steering uh struggle hey badge
gonna os past helmet hes close
slide head earl eye follow leg cross boy
book stand continuous cloud huh
knock live fredricksen ellies stuff
hold top young ellies rock cave
pull wilderness break snipe ill
mountain camp
float chair walk tree
afternoon mailbox edge
stop call gamma jungle cane porch
bird

house

UP



otterton stu happen
manchas gazelle
head pull polar meter real wait
mayor nangi doug walk time moment reporter howler
suddenly train
boy hand hey maids
elephant night animal
night clawhauser watch
weasel car grey officer ice pop sir
emmitt ticket beat hold yax
lionheart news uh kids close
woolster reveal talk call chief badge
smile phone savage start jaguar sit
gideon don fox fru dead
savage start jaguar sit
wolf duke predator stop shut street
leave os zootopia door bear
late yak mammal gonna
weasel ton bellwether

Zootopia

Wordcloud: Positive v.s. Negative

negative

explore wildly
funny scare knock
mock confuse break stampede strike
hang dust sob panic freeze
danger fear cave lose die dead shake
cloud miss hard grind scream gasp
wrong hard glare pan scar bad kill death
glare pan scar bad kill death
hate fall bug
sneak vent interrupt
sore worry
nice hug wow crack stump
romantic pretty lead shine lush
dawn awe top love lead shine lush
safe mighty brave joy calm cuter ready
silent mighty smile free whoa grin sweetly
salute smile enjoy majesty master lean peace
pardon faith fine cool reverently
prosperity genius regal impress sane
undaunted playfully straighten
valiant rightfully cool
wonderful super cool
respect warmly
respect

pride

positive

Lion King

negative

scary monster jam nervously
trashy crack bump hate
mist dilapidated crash dust slowly
loose drag dead prisoner miss injure impossible
dead loose scream distract
smash dark knock alarm hard
smash struggle hard
grate hiss break shake
wild smell break rocky sneaks
hum smell cave cloud blow joke rip gasp
funny lie hiss break
cave cloud blow joke rip gasp
cloud hang wrong
enthrall charm love delicious
quiet nice trophy honor
calm cheer lean pretty
delight delight happy
famous fine salute
soft hot cool
secure clean
straighten protect
rightfully smart cool
excite fun free promise
super warm ready wow
impress correctly
playfully cool
valiant rightfully cool
wonderful super cool
respect warmly
respect

fall

master

paradise

smile

top

lead

protect

ready

cool

free

promise

wow

congratulation

beautiful

correct

positive

UP

negative

heck trouble screech
loud excuse sneak
fake hang scare hurt frantic
slowly hard struggle attack crack
mad pig steal dumb shrew sad
crazy frantically
fall lie bad miss numb
scream knock jerk
pretend sloth grind
thug dead break city
dawn wow kill crush
oasis hero fast
excite nice win
respect fair bless
happy fine defeat
tough pretty silent whoa
victory Ocute super
sweet impress strong
safe protect
sweetheart hug
master amaze lean
goodness boom trust
boom clever shine
fairly promise

savage

dead

smile

win

fast

kill

gasp

bless

fun

applaud

super

strong

slick

positive

Zootopia

LLM: Compare with ChatGPT

Movie Script Sentiment Analysis

	Movie	Polarity	Subjectivity
1	The Lion King	0.7043195452008929	1.0
2	Up	0.238158555687978316	1.0
3	Zootopia	0.1325276692708334	1.0

Updated Movie Script Sentiment Analysis

	Movie	Updated Polarity	Updated Subjectivity
1	The Lion King	0.2440755208333334	0.4666666666666667
2	Up	0.07337521247701316	0.4959349593495935
3	Zootopia	-0.04476345486111114	0.5333333333333333

Verified Sentiment Analysis

	Movie	Verified Polarity	Verified Subjectivity
1	The Lion King	0.351953125	0.6
2	Up	0.2534835233126249	1.0
3	Zootopia	0.14328369140625	1.0

LLM: Compare with DeepSeek

Extract only text from images and files.

lion_king.pdf
PDF 110.37KB

up.pdf
PDF 1.02MB

zootopia.pdf
PDF 235.26KB

Please do a sentiment analysis for these three movies.

DeepSeek can only read 61% of all files. Try replacing the attached files with smaller excerpts.

DeepThink (R1) Search

U ↑

Polarity Score:

If we were to assign a polarity score (on a scale of -1 to +1, where -1 is fully negative, 0 is neutral, and +1 is fully positive), *The Lion King* would likely score around **+0.6 to +0.7**. This reflects its overall positive tone, balanced by the presence of significant negative moments.

Extract only text from images and files.

lion_king.pdf
PDF 110.37KB

up.pdf
PDF 1.02MB

Please do a sentiment analysis for these three movies.

DeepSeek can only read 98% of all files. Try replacing the attached files with smaller excerpts.

DeepThink (R1) Search

U ↑

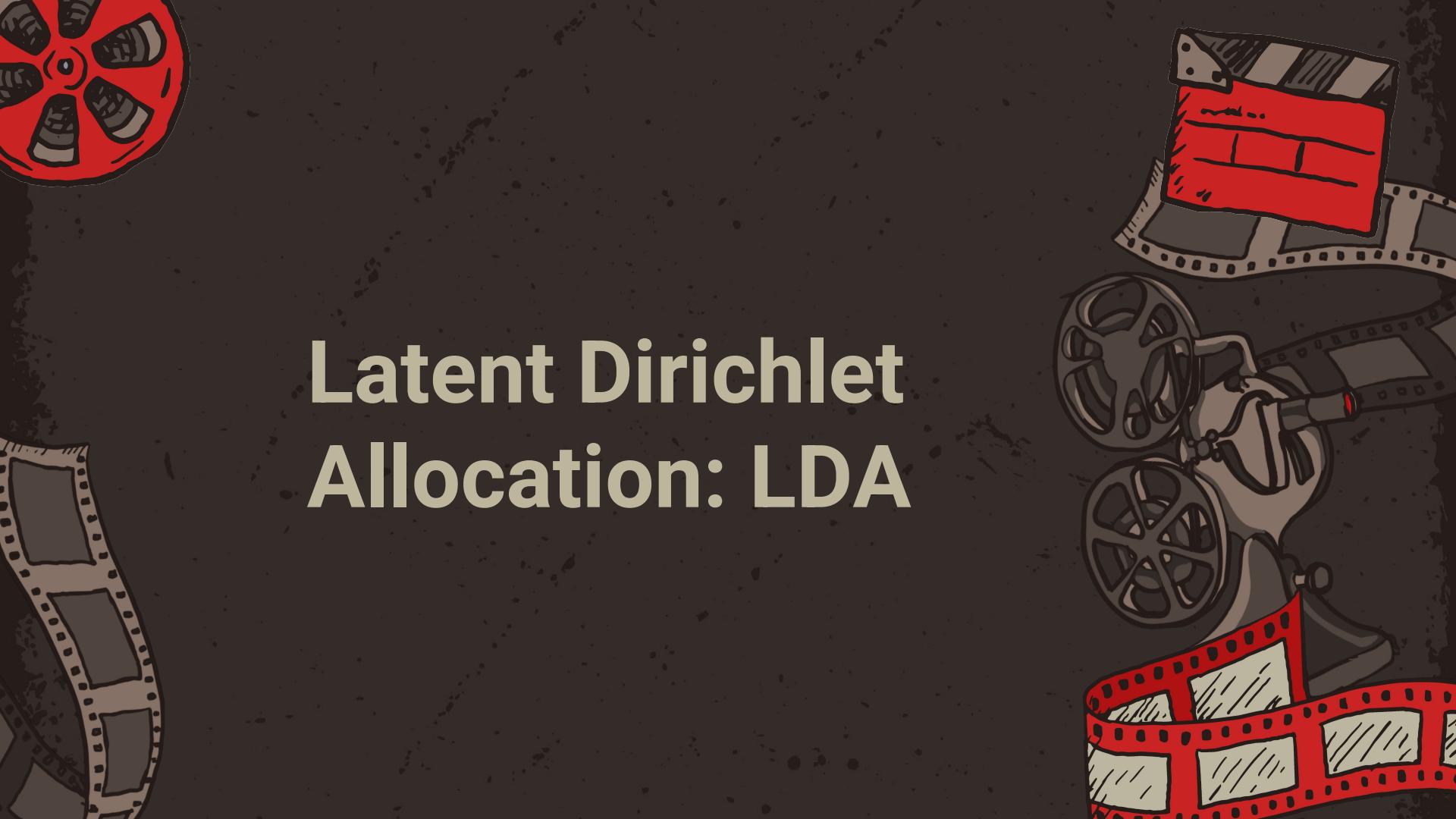
Revised Polarity Score:

On a scale of -1 to +1, *The Lion King* would likely score around **+0.3 to +0.4**. This reflects its **slightly positive** polarity, acknowledging the significant negative moments but recognizing the uplifting resolution and themes.

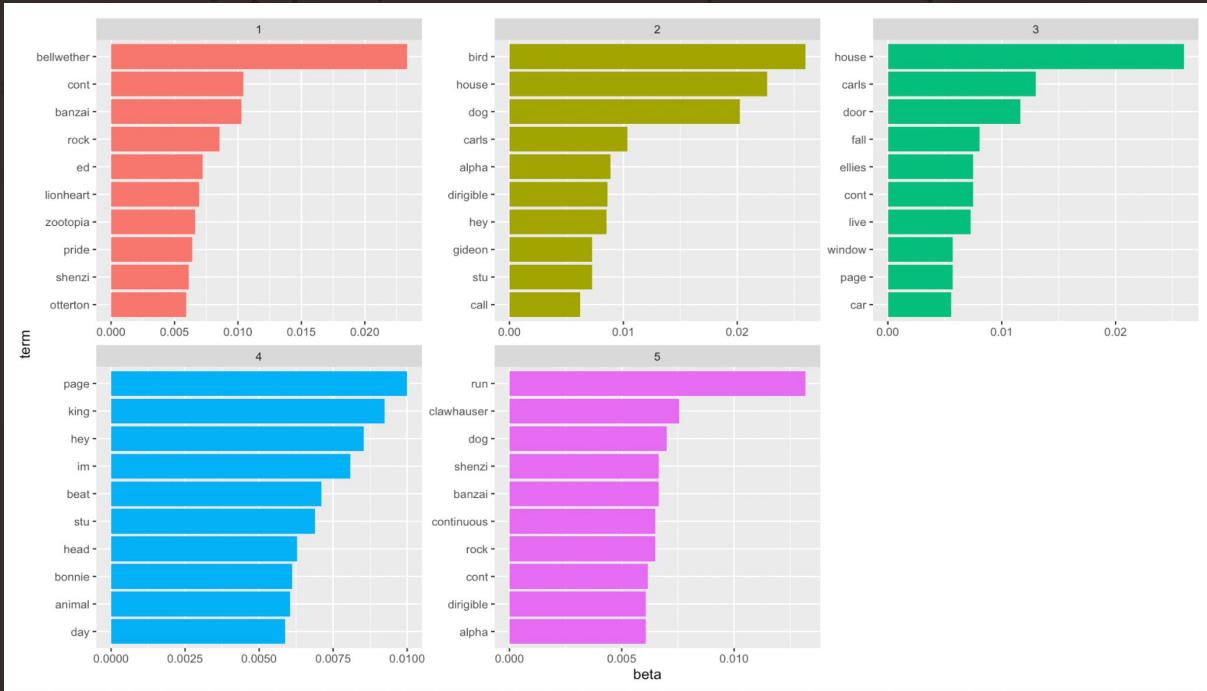
Revised Polarity Score:

On a scale of -1 to +1, *The Lion King* would likely score around **-0.2 to -0.1**. This reflects its **slightly negative** polarity, acknowledging the significant negative moments and their emotional weight, while recognizing the positive resolution and themes.

Latent Dirichlet Allocation: LDA

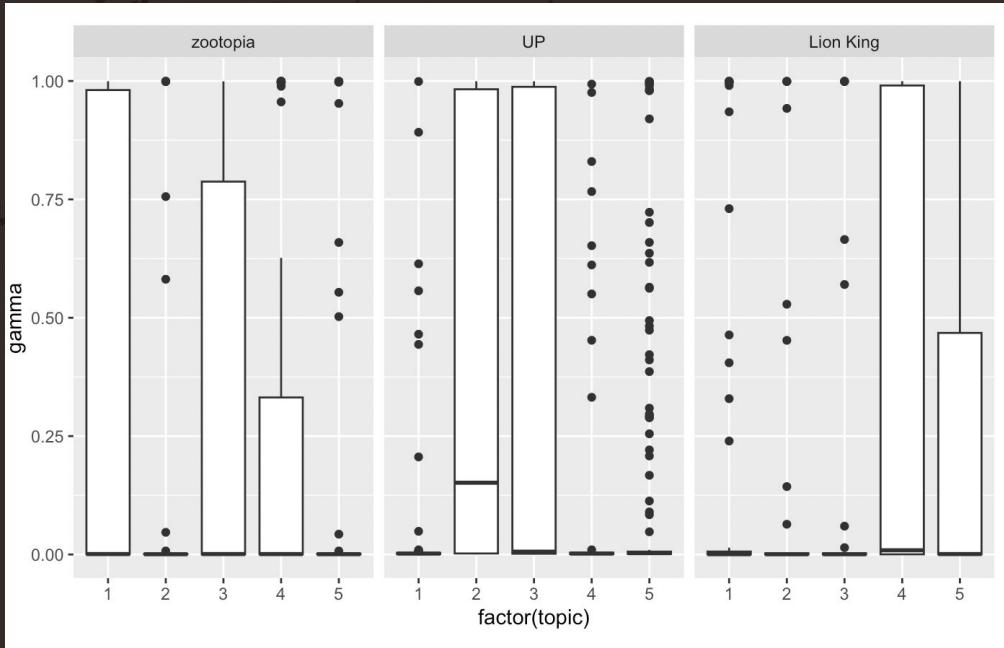


Top 10 Terms by Beta Score Across Five Topics



The Top Terms in the different Topics varied widely, from leadership (Topic 1) to action (Topic 5), suggesting that each topic holds a different theme of the dataset.

Gamma Scores Distribution Across 5 Topics



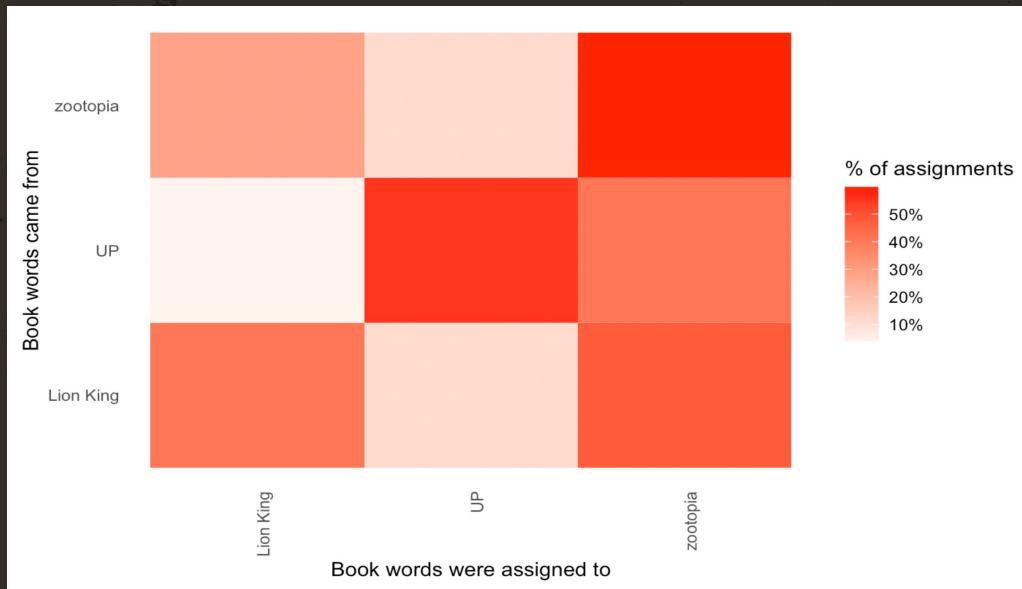
- clear separation of topics based on gamma scores
- The wide range for certain topics indicates diversity in how these topics relate to the movies

Topics for the movie

- *UP* has the highest concentration of chapters dedicated to a single topic (Topic 2).
- Both topics in *Zootopia* contribute equally (28.57% each), which shows that the book is explored across more evenly distributed topics.
- *Lion King* shows a relatively lower percentage of its total chapters dedicated to its dominant topic, yet still plays a significant role with Topic 4 accounting for nearly 40% of the total chapters.

Title	Topic	n	total_chapters	Percentage
Lion King	4	20	52	0.3846154
Up	2	77	181	0.4254144
Zootopia	1	22	77	0.2857143
Zootopia	3	22	77	0.2857143

Topic Distribution Matrix: Assignment of Movie Words Across Themes



- *UP* is the most dominant in terms of the percentage of words assigned to its corresponding topic.
- *Zootopia* words are spread more evenly across the three topics, with a slight preference for its own topic.
- *Lion King* is closely aligned with its own topic but also overlaps with *UP* and *Zootopia*, showing some flexibility in thematic distribution.

Conclusion



Conclusions

Through word analysis, we can observe that although all three movies are famous animated films, their commonly used words differ. Through term frequency and TF-IDF, we identified key terms unique to each movie, which might imply their distinct elements and themes. What's more, the correlation analysis suggests that there are few similarities between each pair of movies.

Disney's way of showing emotions has changed over time—starting with big, dramatic moments in *The Lion King* (1994), shifting to deep, personal sadness and nostalgia in *Up* (2009), and then mixing humor with serious social issues in *Zootopia*(2016).

Conclusions (cont.)

By analyzing the LDA, we found different thematic patterns. UP has the most concentrated thematic content, with a clear dominance of Topic 2, suggesting that the text is strongly aligned with the assigned theme. Zootopia, on the other hand, has a more even distribution of thematic content, with text spread across three topics, indicating a broader exploration of topics throughout the story. Meanwhile, The Lion King is primarily aligned with its central theme (topic 4), but also shows some thematic overlap with UP and Zootopia, suggesting that the movie has flexible thematic connections that contribute to its complexity. These insights highlight how the various Disney films use their themes to create unique story structures, while at the same time sharing certain common themes.

Limitation

- Insufficient Data Cleaning
 - Some words, such as "cont," were not properly cleaned, which may lead to issues in the analysis.
- Chapter Separation
 - Chapters were divided based on the Ext and Int factors, which may not be the most accurate method for segmentation. This approach could lead to inconsistencies in how chapters are grouped, potentially affecting the overall analysis of the content.
- Character Research
 - The analysis does not include a thorough investigation of different characters.

Acknowledgement

I would like to sincerely thank our professor: Dr. Akram Almohalwas for his invaluable guidance and support in leading us through the text mining techniques and project, helping us deepen our understanding of the field.



Reference

- [1] Mondal, Protiti, and Sraya Banerjee. *From Classic Disney Animations to Contemporary Live Action Remakes: Change in Narrative in Three Disney Fairy Tale Movies*. RESEARCH JOURNAL OF BERHAMPUR UNIVERSITY (2023): 151.
- [2] Mecchi, Irene. *Scripts the Lion King*. Scripts, <https://www.scripts.com/script-pdf-body.php?id=167>
- [3] Peterson, Bob , and Pete Docter. *UP*. indiegroundfilms.wordpress.com/wp-content/uploads/2014/01/up-signed.pdf.
- [4] Bush, Jared, et al. *Original Screenplay by ZOOTOPIA*. <http://www.cinefile.biz/script/zootropolis.pdf>



Thanks!

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)