

Sentimentanalyse von YouTube-Kommentaren

Video: *Opulenz - Contrapoints*

Thomas Haase

19. Januar 2025

Justus-Liebig-Universität Gießen

Matrikelnummer - 6033199

Blockseminar:

Prof. Dr. Dennis Klinkhammer - *Digitale Ungleichheits- und Geschlechterverhältnisse*
Wie Algorithmen exklusive Narrative erzeugen und bestärken können

Der folgende Text enthält 1987 Wörter

Inhaltsverzeichnis

1	Begründung der Materialauswahl	2
2	Zusammenfassung des Textmaterial	2
2.1	Natalie Wynn	2
2.2	Kanalname "Contrapoints"	2
2.3	Zusammenfassung Videos	3
2.4	Beispielhafter Einblick in Topkommentare	4
3	Sentimentanalyse	4
3.1	Plot 1	6
3.2	Plot 2	6
3.3	Plot 3 und 4	7
4	Ergebnisreflexion	8
5	Quellen	8
6	Anhang	9
6.1	Topkommentare	9
6.2	R-Code	10
6.3	Eigenständigkeitserklärung	11

1 Begründung der Materialauswahl

Im Folgenden werden die Voraussetzungen und die Auswahl des Materials erklärt, um einerseits die Motivation der kurzen Analyse zu skizzieren, und andererseits die persönliche Involviertheit transparent darzulegen, damit die Objektivität der Analyse nachvollzogen werden kann.

Mein Ziel bei der Auswahl des Materials war, dass der zu analysierende Inhalt thematisch auf das Thema des Seminartitels Bezug nimmt. Auch wenn das Tool von Prof. Dr. Klinkhammer für die Analyse längerer Texte ausgelegt ist habe ich mich für die Analyse von YouTube-Kommentaren entschieden, weil Sentimentanalyse besonders populär in der Auswertung von Nutzerdaten vieler Individuen ist. Der Nutzen einer Einübung der Analyse eines solchen Anwendungsfeldes in einem Seminarsetting ist daher plausibel und gegeben.

Ich habe mich für das Video mit dem Titel *Opulenz* von Contrapoints entschieden. Im Video *Opulenz* wird thematisiert, wie Machtverhältnisse durch Ästhetik und Geschmack manifestiert werden, diese Ästhetik die Kapitalistische Utopie verdeckt und unhinterfragt, genauer: "natürlich" erscheinen lässt. Zuletzt thematisiert die Sprecherin des Videos, wie *Opulenz* die Ästhetik des Goths erzeugt, die die von vielen ausgeblendete Vergänglichkeit des Überflusses darstellt.

Das Video befasst sich daher sehr direkt mit Ungleichheit in der Gesellschaft und erfüllt die zu Beginn selbst gestellte Anforderung.

2 Zusammenfassung des Textmaterial

In diesem Abschnitt wird kurz auf den Kanal *Contrapoints* eingegangen, der Inhalt des Videos zusammengefasst und anschließend werden die obersten 10 Topkommentare des Videos analysiert.

2.1 Natalie Wynn

Contrapoints wird von Natalie Wynn geführt. Wynn studierte Klavier am Berklee College of Music, hat einen Master in Philosophie und brach ihr PHD Programm ab um den YouTube-Kanal zu gründen. Wynn veröffentlicht auf ihrem Kanal Videoessays, in denen Sie Themen wie die neue Rechte, Rassismus, Ungleichheit und Radikalisierung (bspw. Incels) behandelt. Die Videos sind mit ca. 1-2 Millionen Views sehr populär und besitzen ein hohes Niveau an künstlerischer, gestalterischer und inhaltlicher Tiefe welche dem Inhalt oft eine neue Ebene hinzufügen. Um die Videos unterhaltsam zu gestalten kommen (oft auch sexuelle) Pointen und Sarkasmus zum Einsatz. Wynn ist Transgender und vertritt sozialdemokratische Werte.

2.2 Kanalname "Contrapoints"

Der Kanalname *Contrapoints* ist gleichzeitig Homonym wie Oxymoron, da das Wort zwei gegensätzliche Bedeutungen vereint.

In der Musik wird mit Kontrapunkt (lat.: punctus contra punctum = "Note gegen Note") eine Technik bezeichnet, durch die einer vorgegebenen Stimme eine zweite hinzugefügt wird. Dabei komme ein komplexes Regelsystem zur Anwendung. Ziel war es im Mittelalter (ca. 1450) Dissonanzen in der Musik zu vermeiden. Besonders bekannt ist Johann Sebastian Bach für seine außergewöhnlich geniale Anwendungen des Kontrapunktischen Prinzips.

Im Bezug auf Videoessays, die als Kommentare in einer gesamtgesellschaftlichen Diskussion verstanden werden können, bedeutet Contrapoints einen *"Gegenpunkt"* in einer Diskussion zu liefern.

Kurz: In der musikalischen Bedeutung werden mit "Contrapoints" Dissonanzen vermieden, wenn man eine Stimme hinzufügt. In der kommunikativen Bedeutung werden Dissonanzen erzeugt um eine aufklärerische Auseinandersetzung eines Themas zu ermöglichen. Harmonie durch Dissonanz zu erzeugen ist ein Ausdruck und eine Bezeugung zum Grundprinzip der Demokratie.

Dieser Abschnitt diene dem Zweck die verschiedenen Zuschreibungen im zuvor genannten Abschnitt mit einem Beispiel zu hinterlegen.

2.3 Zusammenfassung Videos

Aus Simplitätsgründen werden folgende Videoparameter nur aufgelistet, nicht weiter erläutert und helfen sich kurz einen Überblick zu verschaffen.

Länge	49:06 Minuten
Uploaddatum	12.10.2019
Aufrufe	3.248.191
Likes	147.872
Kommentare	13.156

Tabelle 1: Deskriptive Werte des YouTubevideos

Wynn beginnt ihr Video "Opulenz" mit der These, dass als Kind Ästhetik durch absolute Freude und Lust (zum Beispiel an glitzernden Steinen) entsteht, mit der fortschreitenden Sozialisation wird uns allerdings ein Geschmack antrainiert.

Opulenz gelte als Zeichen von Erfolg und sei ein Symbol dafür, dass man den American Dream erreicht hat. Das ist allerdings unmöglich, weil der American Dream eine Utopie ist und keine Leistungs- oder Chancengerechtigkeit gibt, es handelt sich lediglich um Ideale. Trotzdem existiert dieser Gedanke weiter in der Gesellschaft, denn die Vorstellung viel zu besitzen ist schön, selbst wenn sie nur als *"Ballreichtum"* ("ich sehe reich aus, also könnte ich reich sein") gespielt werde. Durch solches Verhalten wird allerdings das ausbeuterische Moment der Kulturindustrie noch weiter verstärkt (Glücksspiel, Wille nach mehr). Wynn bringt es im Video folgendermaßen auf den Punkt:

"For the wealthy and powerful opulence is a flex, but for the poverished opulence is a sepulcrum of the power they have been denied" - Natalie Wynn

Anschließend widmet sich Wynn der Marxistischen Vorstellung von Klassen. Diese sei zu einfach für die heutzutage existierende Gesellschaft. Sie fährt mit einer Erläuterung der Begriffe Habitus, kulturellem Kapital und Klasse fort. Auch die Vorstellung von Ästhetik spiegelt sich in den verschiedenen Gesellschaftsschichten wieder. Ästhetik ist nach Kant und Platon nicht-materiell und transzendent - man kann sie also nicht kaufen. Geschmack ist ein Indikator für Macht und Kultiviertheit.

Die Mächtigen der Gesellschaft kontrollieren durch Kunstaussstellungen und Museen was geschmackvoll ist. Zum Beispiel wurden koloniale Strukturen durch das ausschließen schwarzer Kunst gefestigt oder in der Renaissance die weiße Farbe von Statuen einfach angenommen, obwohl sie früher gefärbt waren.

Im vorletzten Teil des Videos behandelt Wynn den Begriff "Glamour" (von *grammery* = magic/ illusion/ charm). Ursprünglich sei dieser Ausdruck mit fremdem und heiligem verknüpft gewesen. Heutzutage ist der Begriff stark mit Reichtum verknüpft. Glamour ist distanziert statt integrativ, individuell statt gesellschaftlich und im Zwiespalt zwischen Identifizierung und Un-erreichbarkeit.

Zum Schluss widmet sich Wynn der Goth-Ästhetik. Bei dieser spielen Vampire eine wichtige Rolle. Diese sind reiche, aristokratische, gehobene Verführer die eine untergehende Weltordnung repräsentieren. In dieser werden Menschen ausgebeutet/gesaugt. Vampire wohnen in den Medien klassischerweise in verfallenen Burgen oder Mansions aus genau dieser untergehenden Weltordnung. Auch heutzutage sind diese Motive im American Horror wiederzuerkennen (lost places wie alte shoppingmalls, etc.) und konfrontieren die Menschen auch heutzutage mit ihrer Vergänglichkeit und damit der Entwertung der Opulenz. Goth ist der Tod und Verfall der Opulenz.

Wynn beendet das Video mit den Worten: *"I accept my fate..."*

2.4 Beispielhafter Einblick in Topkommentare

Die an dieser Stelle knapp analysierten Topkommentare befinden sich mit ihren induktiv gebildeten Codes im Anhang. Analyseeinheit sind die einzelnen Kommentare.

Auffällig ist, dass unter diesen Kommentaren rege Diskussionen über das Thema des Videos stattfanden. Der Topkommentar hat zum Beispiel 44 Antworten. Aus den Kommentaren lässt sich durch sehr grobes, skizzierendes, Induktives Codierung folgendes erkennen:

Code	Anteil der 10 obersten Topkommentare zum 9.5.23
positive Interaktion	10/10
Inhaltliche Auseinandersetzung	4/10
Transgenderbezug	3/10

Tabelle 2: Ergebnisse des knappen, induktiven Codings

Anzumerken ist auch, dass der letzte der 10 Kommentare mit 2201 Zeichen mehr Zeichen als alle anderen Kommentare in Summe besitzt. Ohne diesen liegt die Durchschnittliche Zeichenzahl eines der Top 10 Kommentare also bei $\frac{1645}{9} = 182,7$, mit dem letzten Kommentar $\frac{3846}{10} = 384$ Zeichen.

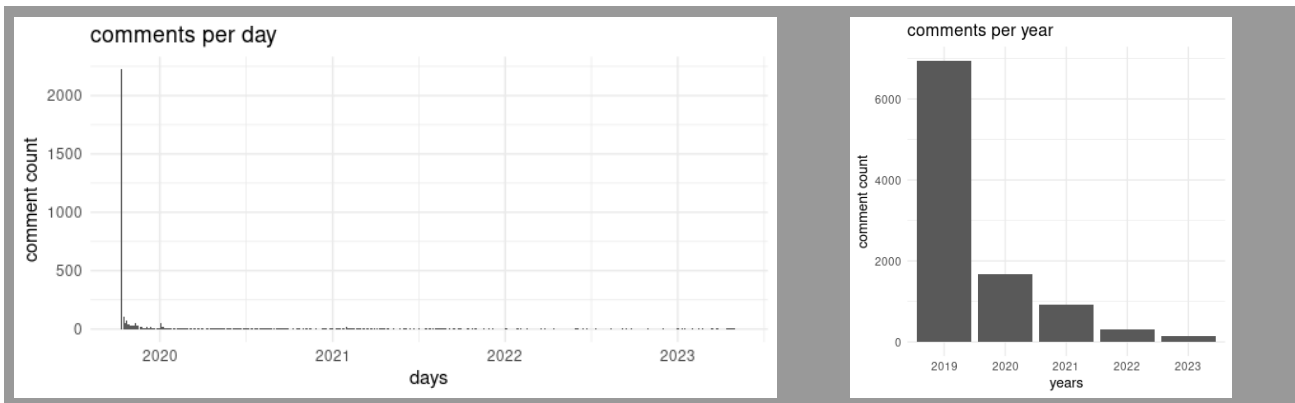
Diese Zahl verwundert nicht. Wynns Videos haben einen für YouTube überdurchschnittlichen Grad an analytischer Tiefe. Zuschauer die an solchen Inhalten interessiert sind möchten sich auch mit diesen auseinandersetzen. Wer dies nicht will sieht sich kein Video an, in welchem nahezu eine Stunde gesellschaftliche und künstlerische Aspekte ausgehandelt werden. Der Bezug auf Wynn als Person erschließt sich auch intuitiv. Wynn betreibt (vor allem online) Aktivismus und steht für kosmopolitische Werte. Über Soziale Medien sind Personenkulte kaum zu vermeiden, wenn dann diese Person mehreren Minderheiten angehört, steht sie als Vorbild für viele Personen in der Gesellschaft.

3 Sentimentanalyse

Im folgenden wird die Analyse der Kommentare mit Hilfsmitteln der quantitativen Statistik erfolgen. Zentral wird hierbei das R-Skript von Prof. Dr. Dennis Klinkhammer aus dem Seminar

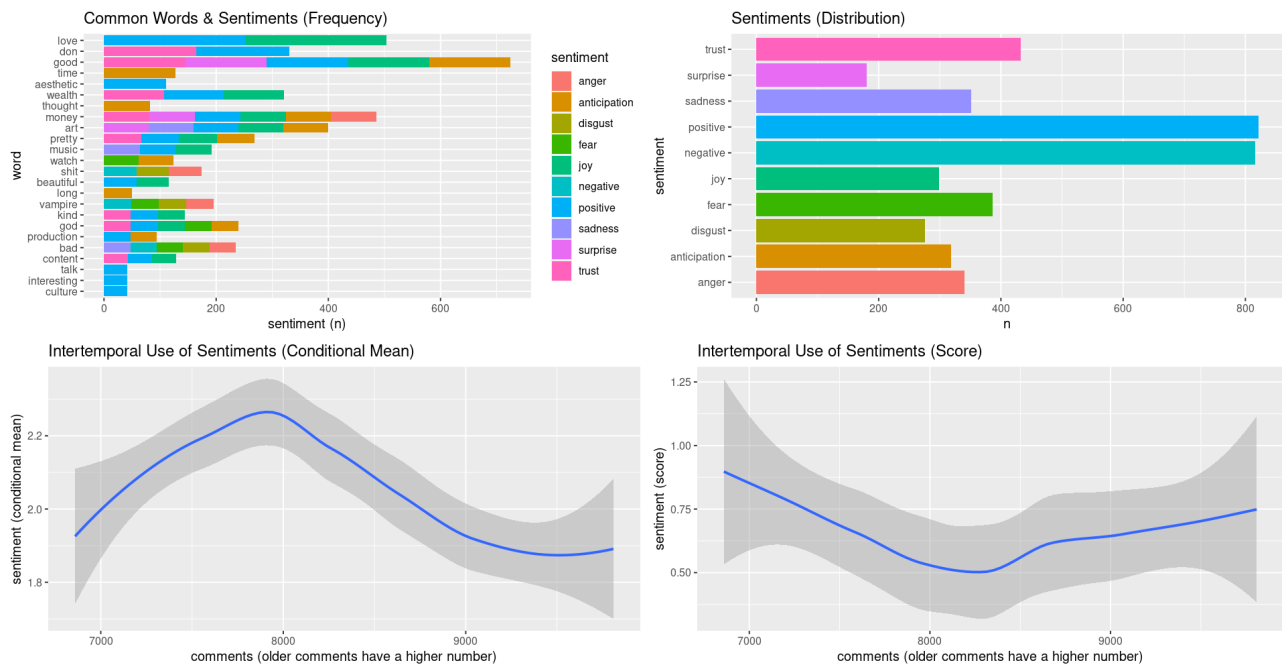
zum Einsatz kommen.

Die YouTube-Kommentare wurden aus Gründen der Einfachheit nicht durch die YouTube-API sondern Netlytic heruntergeladen. Dadurch ließen sich nur 10.000 der insgesamt 13.156 Kommentare herunterladen. Das Fehlen der Kommentare schränkt die Analyse nur bedingt ein wie ich im folgenden erläutern werde. Im folgenden sind die Veröffentlichungszeitpunkte der Kommentare geplottet. Wie zu sehen ist sind die Kommentare seit Videoveröffentlichungszeitpunkt im Datensatz enthalten, lediglich die 3000 aktuellsten Kommentare fehlen. Ein kurzer Blick in den Datensatz bestätigt, dass der älteste Kommentar am Datum der Videoveröffentlichung verfasst wurde.



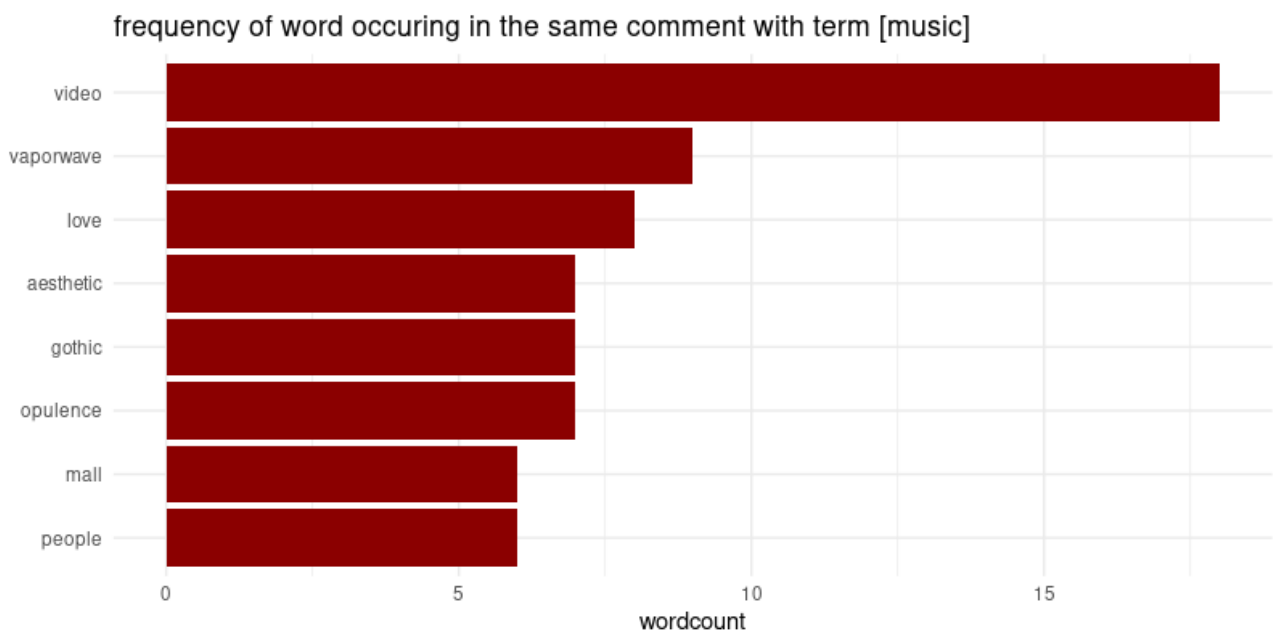
In dem *comments per year*-Plot ist zu erkennen, dass die Kommentaranzahl pro Jahr stark gesunken ist, die wichtigsten $\sim 75\%$ der Kommentare sind demnach im Datensatz enthalten. lediglich die ca. 3000 neusten Kommentare sind nicht abgebildet.

Aus der Sentimentanalyse ergeben sich folgende Plots:



3.1 Plot 1

In dem ersten Plot sind die häufigsten Wörter der Kommentare mit ihren jeweiligen Sentiments aufgelistet. Es fällt direkt auf, dass in den dargestellten 24 häufigsten Wörtern nur 7 Wörter (*time, thought, watch, shit, long, vampire, bad*) keine positives Sentiment aufweisen. Diese Beobachtung bestätigt die zuvor aus den Beispielen erkannte positive Interaktion mit dem Video. Außerdem ist erkennbar, dass viele der häufigsten Wörtern direkt in den Themenbereichen des Videos relevant sind. Zum Beispiel *aesthetic, wealth, money, art, vampire, god, culture*. Dadurch lässt sich die im Beispiel benannte inhaltliche Auseinandersetzung wiedererkennen (Auch, dass das 7. häufigste Wort "*thought*" ist, die Zuschauer denken mit!). Die hohe positive Resonanz lässt sich sogar mit den folgenden Wörtern vermuten: *aesthetic, music, watch, beautiful, production, content*. Das Video ist in der tat ungewöhnlich professionell und künstlerisch produziert. Interessant ist, dass *music* hier scheinbar oft als Thema kommentiert wird. Im folgenden sind die häufigsten Wörter aufgelistet, die in den Kommentaren, in denen *music* erwähnt wird, vorkommen.

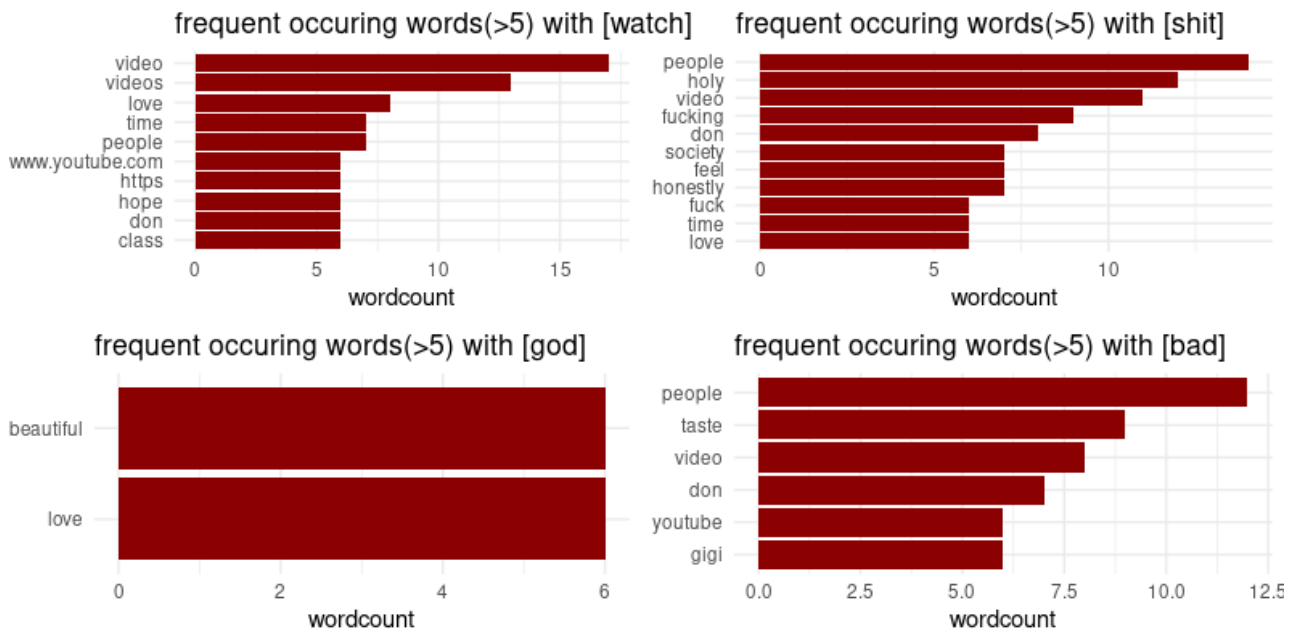


Die folgenden Wörter im Plot: *aesthetic, gothic, opulence, mall* sprechen erneut für die Inhaltliche Auseinandersetzung mit dem Video, Musik wird also oft in Verknüpfung mit inhaltlichem gesehen und *vaporwave* und *waporwave, love* sprechen für positive Resonanz der Zuschauer mit der Musik des Videos.

3.2 Plot 2

Im zweiten Plot des Sentimentplots aus dem Seminar [3] lassen sich nicht nur "gute" Sentiments ablesen ("*trust, positive, joy, anticipation*"), sondern auch eine große Anzahl an negativen Sentiments ("*negative, fear, sadness, anger, disgust*") erkennen. Diese sind aber nicht auf Videokritiken zurückzuführen, ansonsten wären viele nicht-im Video erscheinende Begriffe in Plot 1 abgebildet. Stattdessen werden im Video Themen angesprochen, die im "NRC word-emotion association lexicon" als negativ eingestuft werden.

Im Plot 1 lassen sich einige dafür ursächliche Begriffe erkennen: *watch*, *shit*, *god*, *bad* aber auch *art*, *money* werden mit genau diesen negativen Sentiments gelabelt. Die letzteren beiden Begriffe sind Kernthemen des Videos und damit zu erklären, die anderen Begriffe gilt es im folgenden genauer zu analysieren. Hierzu werden wieder die häufigsten Wörter aufgelistet, die zusammen mit den zu erklärenden Wörtern in den jeweiligen Kommentaren auftreten.



Der Begriff "*watch*" ist nicht als negativ einzuordnen, weil Zuschauer ihn entweder mit neutralen Wörtern, positiven Wörtern wie "*love*" in Verbindung bringen, oder Videos geteilt haben ("*watch*" ist auch Teil einer spezifischen YouTube-URL).

Der Begriff "*shit*" wird als Ausruf genutzt (mit *holy*, *fucking*, *feel*, *fuck*) oder wird im Kontext mit abstrakten Konstrukten wie "Gesellschaft" oder "Personen", aber nicht dem Video an sich genutzt.

Der Begriff "*god*" wird nur in positiven Kontexten verwendet.

Der Begriff "*bad*" bezieht sich wie "*shit*" auf "Personen" oder auf Themen im Video ("*bad taste*" ist ein Thema des Videos und gigi eine YouTubepersönlichkeit die im Video als Beispiel genutzt wird).

3.3 Plot 3 und 4

Die beiden letzten Plots können gut zusammen interpretiert werden. In Plot 4 ist zu sehen, dass bei ca. Kommentar Nummer 8200 ein Minimum entstanden ist. Das bedeutet es gab in dem Zeitraum, in dem die Kommentare um Kommentar 8200 verfasst wurden, besonders negative Kommentare. In Plot 3 ist für dieselben Kommentare eine Zunahme an Sentimentäußerungen zu erkennen. Es handelt sich um ein typisches Muster eines Shitstorms.

Durch eine Kurze Recherche auf Reddit[2], Twitter und Wikipedia lässt sich diese Theorie bestätigen. Auf Wikipedia ist der Grund des Shitstorms prägnant zusammengefasst:

”Shortly after, Wynn’s video ”Opulence” featured a quote from John Waters read by transsexual pornographic actor Buck Angel,[35] whose views on transgender and non-binary people have attracted criticism, including by some who see Angel’s views as being transmedicalist.[34][35] She was criticized for featuring Angel, including by journalist Ana Valens. Wynn, as well as other YouTubers associated with her channel, were widely harassed.[34][35]” -[1].

4 Ergebnisreflexion

Ich habe bereits in meiner Analyse ausreichend dargelegt, weshalb an welchen Stellen dem Sentiment-Tool zugestimmt werden kann und mit weiteren Einblicken in Analysen von Begriffskorrelationen gezeigt wo die Grenzen des Tools bzw. von Sentimentanalysen liegen. Sentimentanalysen liefern nur einen groben Richtwert, können allerdings auffällige Texte/Textpassagen gut hervorheben, die im Anschluss genauer analysiert werden müssen. Sentimentanalyse ist nur ein mögliches Tool im Werkzeugkasten eines Forschers, der quantitativ aber auch qualitativ Texte analysieren möchte. Wie jede Methode sollte sich auch hier die Methode am Forschungsgegenstand orientieren. Im Falle von YouTube-Kommentaren eignete sich die Sentimentanalyse hervorragend um den Ideen, die die Zuschauer teilten, aber auch die zeitliche Entwicklung der Resonanz mit dem Video zu beschreiben.

5 Quellen

Literatur

[1] Wikipedia contrapoints. online on Wikipedia.

[2] dogmaticmink. What is the twitter meltdown thing. r/LindsayEllis:, 2019-2020.

6 Anhang

6.1 Topkommentare

Comment	Likes	Antworten	Interpretation	Codierung
I saw this working in a hotel. The super wealthy treated us coldly but well, and tipped well. The poor who were there for special events were grateful, even if they didnt tip. The middle class were monsters. Demanded, screamed, tried to get you fired for any perceived slight.	6651	44	unterstützendes Beispiel	positive Interaktion/ Inhaltliche Auseinandersetzung
Opulence has really drowned in the cancellation debacle. Can we take a minute to appreciate what a true piece of mastery this really is?	2516	19	Lob	positive Interaktion
It is SO cool to hear her talk about a prediction of "Decaying Mall Gothic" aestheticism in 2019, when here in 2021 we're already seeing a huge rise of that in Weirdcore and Liminal Space aesthetics. Maybe I just linger in obscure "horror content" corners of the internet- but I'm definitely seeing a huge increase in people taking 80s decay for their gothic aesthetics. She's right on the money.	1576	17	unterstützendes Beispiel	positive Interaktion/ Inhaltliche Auseinandersetzung
my mom saw me watching this and said "shes so beautiful she dresses up to give you facts"	8496	21	Humorvoll	positive Interaktion/ ggf. Transgenderbezug
Just an offering to the algorithm deities. Who else watches Natalie's work over and over again?	654	16	Lob	positive Interaktion
The irony that the framing device of this video is Natalie playing someone who's about to be beheaded by a bloodthirsty mob, and people cancelled her in a bloodthirsty online mob as a result of three seconds of this video	185	8	Humorvoll, unterstützendes Beispiel	positive Interaktion/ ggf. Transgenderbezug
this video called me poor and i'm thanking it	18868	47	Lob	positive Interaktion
Now that Abigail Thorn is out of the closet between Abby and Natalie a new genre of youtube video has arrived: "Leftist transwoman in glamorous makeup and gorgeous costume pretend to be in their natural environment in fantastical places and wax philisophical." And as a leftist neurodivergent transwoman ex-theatre kid who likes facts and shiny things I have no choice but to stan hard.	626	10	Lob	positive Interaktion/ Transgenderbezug
An interesting thing I've observed is in contrast to the "Trump aesthetic" you describe - old rich who act like new rich, money without class - there is the direct opposite of that: "class" without money. I was raised in what was basically a poor branch of a distant, larger rich clan. My mom grew up middle class, but her taste and style were heavily informed by rich relatives in her childhood, and then the rich clients she went on to work for as an adult banker. She would scoff and roll her eyes at "nouveau riche" aesthetics: anything showy, ostentatious, excessive, "tacky." She loved pointing out when newly upper class people were insecure of their wealth and status, explained how you could tell they were lesser than the old rich, inferior compared to the people bred with class and etiquette as priority. She could always smugly point out which people were "poors getting a little taste of the good life" and subsequently turn her nose up at them. But even as she looked down on everyone who wasn't as assured and muted and "classy" as the old rich, she didn't venerate these supposedly superior people either. To her, being classy was never about the money, but the mindset: assurance, quiet confidence, simple but clean aesthetics in everything. She used to tell me, "The way you dress with class is not to wear things that make you look like you're rich, but to always have pristine hygiene for your body and clothes, and never be seen wearing or using damaged things." This was how she still considered herself classy and continued gladly ragging on the poor, middle class, and the new rich - even after three decades of retiring from her bank job and being stuck as the housewife of a family living in a slum. She taught me all the possible tricks to wash and maintain the few hand-me-down clothes I had so they'd always look new, how to haggle with smugglers of foreign overruns so I could buy original products at the price of fakes, how to cook like mothers in provincial haciendas and never like cheap urban canteens. She's a complicated woman I haven't talked to in a long time, and when I was living with her we always clashed, but this aspect of her has always stood out to me.	1260	13	unterstützendes Beispiel	positive Interaktion/ Inhaltliche Auseinandersetzung

6.2 R-Code

Dies ist der Code für alle Informationen, die in der Abgabe genutzt wurden. den Gelb hinterlegten Code habe ich hinzugefügt

```
library(dplyr) # Datenverarbeitung
library(stringr) # Textverarbeitung
library(tidytext) # Textverarbeitung
library(ggplot2) # Visualisierung
library(gridExtra) # Visualisierung

# Textbibliotheken einmalig installieren
library(textdata) # Textverarbeitung
get_sentiments("nrc") # NRC Bibliothek
# get_sentiment("bing") # BING Bibliothek

# Text in R einlesen
text_df <- read.delim("beispiel_nato.txt", header=F, sep="\t")
path <- "Opulenz.csv"
comments <- read.csv(path)

# Text als Tibble
text_df <- tibble(line=1:dim(comments)[1], text=comments$description)
text_df

# Text als Tidy
text_tidy <- text_df %>%
  unnest_tokens(word, text)

# Woerter ohne interpretativen Gehalt herausfiltern
# data(stop_words)
# text_tidy <- text_tidy %>%
# anti_join(stop_words)
# filtering words with theoretical background of Zipfs law:
text_tidy <- text_tidy %>%
  filter(line > quantile(line, 0.7)) %>%
  filter(line < quantile(line, 0.95))

#####
# H a e U F I G S T E W o e R T E R U N D S E N T I M E N T #
#####

# Positive und negative Woerter auflisten
nrc_word_counts <- text_tidy %>%
  inner_join(get_sentiments("nrc")) %>%
  count(word, sentiment, sort = TRUE) %>%
  ungroup()
# nrc_word_counts
# Haeufigste Woerter und deren Sentiment
plot1 <- nrc_word_counts %>%
  filter(n > 40) %>%
  mutate(word = reorder(word, n)) %>%
  ggplot(aes(word, n, fill = sentiment)) +
  geom_col() +
  coord_flip() +
  labs(y = "sentiment (n)") +
  ggtitle("Common Words & Sentiments (Frequency)")

#####
# A N Z A H L A N S E N T I M E N T S #
#####

# Absolute Sentiments auflisten
plot2 <- nrc_word_counts %>%
  inner_join(get_sentiments("nrc")) %>%
  count(word, sentiment) %>%
  ggplot(aes(sentiment, n, fill = sentiment)) +
  geom_col(show.legend = FALSE) +
  ggtitle("Sentiments (Distribution)") +
  coord_flip()

#####
# S E N T I M E N T E I N S A T Z I M V E R L A U F #
#####

# Verlauf des Sentiments auswerten
nrc_word_counts <- text_tidy %>%
  inner_join(get_sentiments("nrc")) %>%
  count(line, sentiment, sort = TRUE) %>%
  ungroup()
nrc_word_counts

# Verlauf des Sentiments darstellen
plot3 <- ggplot(data = nrc_word_counts, mapping = aes(x = line, y = n)) +
  geom_smooth(method="loess", formula="y~x") +
  xlab("Comments (older comments have a higher number)") +
  ylab("sentiment (conditional mean)") +
  ggtitle("Intertemporal Use of Sentiments (Conditional Mean)")

#####
# S E N T I M E N T V E R L A U F #
#####

# Absolutes Sentiment auswerten
bing_word_counts <- bind_rows(
  text_tidy %>%
    inner_join(get_sentiments("bing")) %>%
    mutate(method = "Bing et al.))

bing_word_counts

# Absolutes Sentiment in Summenscore ueberfuehren
sentiment_sum <- ifelse(bing_word_counts$sentiment == "positive", 1, -1)
sentiment_sum_df <- cbind(bing_word_counts$line, sentiment_sum)
colnames(sentiment_sum_df) <- c('var1', 'var2')
sentiment_sum_df <- as.data.frame(sentiment_sum_df)
sentiment_sum_df <- aggregate(sentiment_sum_df$var2,
  by=list(line=sentiment_sum_df$var1), FUN=sum)

# Verlauf des Sentiments darstellen
plot4 <- ggplot(data = sentiment_sum_df, mapping = aes(x = line, y = x)) +
  geom_smooth(method="loess", formula="y~x") +
  xlab("Comments (older comments have a higher number)") +
  ylab("sentiment (score)") +
  ggtitle("Intertemporal Use of Sentiments (Score)")

#####
# G E S A M T A U S W E R T U N G #
#####

grid.arrange(plot1, plot2, plot3, plot4)

# Mein hinzugefuegter Code

library(lubridate)
library(tidyverse)

comments <- comments %>%
  mutate(date = as_date(pubdate))

comments_per_day <- ggplot(comments, mapping = aes(x = date, fill =
  date)) +
  geom_bar() +
  labs(x = "days", y = "comment count", title = "comments per day") +
  theme_minimal()
comments_per_day

comments_per_year <-
  ggplot(as.data.frame(table(year(comments$date))), aes(x = Var1, y =
  Freq)) +
  geom_bar(stat = "identity") +
  labs(x = "years", y = "comment count", title = "comments per year") +
  theme_minimal()
comments_per_year

min(comments$date)

library(widyr)
library(RColorBrewer)
library(forcats)

# count words co-occurring within comment
cor_music <- text_tidy %>%
  anti_join(stop_words) %>% #filter stopwords
  pairwise_count(word, line, sort = TRUE) %>% #count words that occur in
  same line/comment
  filter(item1 == "music") %>% #filter all comments with music
  filter(n > 5) %>% # filter words occurring more often then 5 times
  arrange(n) %>% mutate(item2=factor(item2, levels=item2)) %>% # order
  columns in plot
  ggplot(aes(n, item2, fill = item2)) +
  geom_col(fill="darkred") +
  theme_minimal() +
  theme(legend.position = "none") +
  labs(x = "wordcount", y = NULL, title = "frequency of word occurring in the
  same comment with term [music]")
cor_music

# hier der Plot von cor_music nur als Funktion um die naechsten Plots zu
  beschleunigen

my_cor <- function(MYWORD,number){
  cor_bad_words <- text_tidy %>%
  anti_join(stop_words) %>% #filter stopwords
  pairwise_count(word, line, sort = TRUE) %>% #count words that occur in
  same line/comment
  filter(item1 == MYWORD) %>% #filter all comments with music
  filter(n > number) %>% # filter words occurring more often then 5 times
  arrange(n) %>% mutate(item2=factor(item2, levels=item2)) %>% # order
  columns in plot
  ggplot(aes(n, item2, fill = item2)) +
  geom_col(fill="darkred") +
  theme_minimal() +
  theme(legend.position = "none") +
  labs(x = "wordcount", y = NULL, title = paste("frequent occurring
  words(> ",number,") with [",MYWORD,"]",sep = ""))
  return(cor_bad_words)
}

grid.arrange(my_cor("watch",5),my_cor("shit",5),my_cor("god",5),my_cor("bad",5))
my_cor("bad video",1)
```

6.3 Eigenständigkeitserklärung

Schriftliche Erklärung*

Hiermit erkläre ich,

Name:	Haase
-------	-------

Vorname:	Thomas
----------	--------

geboren am:	05.09.2000
-------------	------------

gegenüber dem Institut für:	Soziologie oder Politikwissenschaft, je nachdem wo Herr Klinkhammer angestellt ist
-----------------------------	--

der Justus-Liebig-Universität Gießen, dass die vorliegende schriftliche Leistung mit dem Thema:	Sentimentanalyse von YouTube-Kommentaren - Video: Opulenz - Contrapoints
---	--

Name des / der Lehrenden der Veranstaltung:	Prof. Dr. Dennis Klinkhammer
---	------------------------------

Titel der Lehrveranstaltung:	Digitale Ungleichheits- und Geschlechterverhältnisse
------------------------------	--

im Semester / in der Veranstaltung:	T13 - Seminar 1
-------------------------------------	-----------------

selbständig und nach den Regeln guter wissenschaftlicher Praxis der Justus-Liebig-Universität Gießen (<http://www.uni-giessen.de/cms/studium/lehre/plagiate>) angefertigt wurde. Sämtliche benutzten Quellen und Hilfsmittel habe ich in der Arbeit angegeben. Ich dulde die Überprüfung mittels Anti-Plagiatssoftware und habe die Leistung auch in digitaler Form (durchsuchbar) eingereicht.

Ort und Datum	12.05.2023
---------------	------------



eigenhändige Unterschrift

* Der Text dieser Erklärung kann von der Homepage des Instituts abgerufen werden. Diese Erklärung ist der eigenständig erstellten Arbeit als Anhang beizufügen. Arbeiten ohne diese Erklärung werden nicht angenommen.