“Speaking Ill of the Dead”: A Statistical Analysis of Media Sentiment Before and After Celebrity Deaths

Titouan Dupleich

April 2024

Abstract

The Latin expression “De mortuis nil nisi bonum dicendum est”, which translates to “Of the dead, nothing but good should be said”, suggests that it is improper for the living to criticize those who have passed away, as they are unable to defend or explain themselves. This saying highlights the *death positivity bias*, a phenomenon that describes the human tendency to portray the deceased in a favourable light, often disregarding their past actions (Allison and Eylon 2005). Exploring the evolution of media sentiment surrounding the deaths of prominent figures offers a unique lens through which to examine societal attitudes and values, shedding light on the complex interplay between celebrity, public perception, and cultural norms. In this context, this thesis seeks to delve into the nuances of media sentiment related to celebrities around the time of their death. This project analyses the media sentiment associated with celebrities who passed away in the last 10 years. Media articles are collected using Event Registry’s News API and assigned a sentiment score. Two periods of interest are selected for comparison: pre-death and post-death. 38 celebrities are analyzed in different societal spheres. Those include politicians, musicians, and athletes among others (e.g., Italian politician Silvio Berlusconi, American basketball player Kobe Bryant). Statistical tests are computed to test for differences in sentiment between before and after death. Additional characteristics of the celebrities including gender, age at death, industry,

origin, or type of death are manually collected. These features are used to understand whether some celebrity characteristics can help explain variations in media sentiment after their death. Finally, Natural Language Processing (NLP) methods are used to show what themes are covered by the media pre- and postmortem, and attempt to understand the reasons behind the bias. The paper finds evidence of a small death positivity bias, especially for show business celebrities and those who died of suicide or natural causes. The themes detected before and after death are very relevant to the celebrities being studied but understand the reasons behind the bias would require additional research.

Contents

List of Figures

List of Tables

1 Introduction

Going back hundreds of years, death has always fascinated humanity, captivating imagination and influencing cultural, philosophical, and artistic expressions. This fascination is encapsulated in the Latin expression "De mortuis nil nisi bonum dicendum est", which translates to "Of the dead, nothing but good should be said". This saying highlights the human tendency to remember and portray the deceased in a favourable light, often disregarding their past actions, whether good or bad. Known as the *death positivity bias*, it "involves forming more favorable perceptions and appraisals of the dead than the living" (Allison and Eylon 2005, p. 6). This phenomenon manifests in obituaries, the media and survey studies, where the deceased are remembered with reverence and idealization (Alfano, Higgins, and Levernier 2018; Heynderickx and Dieltjens 2016; Rusu 2020). However, most studies overlook the public perception of individuals before their death and fail to capture whether post-death positivity is due to an actual change in narrative at death, or merely due to an already existing pre-death appreciation. Additionally, most of the literature on the topic makes a qualitative assessment of the discourse around the deceased. This thesis aims to bridge those gaps by uncovering patterns and quantifying the death positivity bias using Natural Language Processing (NLP) and statistical methods. In this context, this project analyses the media sentiment associated with celebrities who passed away in the last 10 years. 80 celebrities across the industries of music, film, academia, sports, and public affairs are selected (e.g., Italian politician Silvio Berlusconi, American basketball player Kobe Bryant). Two periods of interest are chosen, one pre-death and one post-death for each celebrity. News articles about the celebrities are fetched in both periods using Event Registry's API and a sentiment score is assigned to each article. After filtering out articles about stars with insufficient media coverage, 7600 about 38 celebrities remain for analysis.

The following research questions (RQ) are addressed:

- **RQ1:** Does the media speak more positively about celebrities once they have passed away? (Section 3.3)

- **RQ2:** Which attributes of celebrities can account for variations in media response? (Section 3.4)
- **RQ3:** What themes are highlighted by the media before and after celebrities' passing? Can they help explain the death positivity bias? (Section 3.5)

The paper proceeds as follows: Section 2 provides additional context to this work by discussing the relevant scientific literature. Section 3 covers the methodology used to access the data and answer each research question. Section 4 describes the chosen celebrities and articles, and checks for any inherent biases in the sources of articles before and after death. Section 5 presents the results of the experiments. Lastly, Section 6 concludes, addresses limitations of the study, and offers suggestions for future research.

2 Literature Review

2.1 Thanatology

Thanatology is the scientific study of death and dying, encompassing a broad spectrum of associated practices. It is a multidisciplinary field of research that delves into various aspects of mortality, biological, medical, forensic and psychological facets, while also considering the broader societal implications of death (Anderson 2016). While the field of thanatology as a science is relatively recent, human fascination with death has persisted throughout history. Ancient Greek philosophers already theorised death. Socrates and Plato considered death as a passage to another life where souls continue to live, Epicurius saw it as the cessation of sensation. More recently, Schopenhauer pondered death as the inevitable faith of life (Singh 2016), joined by Nietzsche, Sartre, Montaigne, and numerous other authors who have contemplated its significance. Religion has also played a major role in the way we approach death, shaping rituals and providing explanations. While different religions bring different approaches to death, Fonseca and Testoni (2012) argue that the recent detachment from religion in Western society has greatly contributed to the creation of thanatology as a scientific field of study. The study of death extends beyond the theoretical and clinical aspects explored in thanatology; it also encompasses the profound societal impact of individual deaths, particularly those of public figures which is the focus of the current thesis.

2.2 Celebrity death effects

The deaths of celebrities have various effects. Fans, mourners and the general public unite in support of the deceased. Many fans experience a sense of grief and loss. Peaks in social media activity related to those who passed away are reported as people express bereavement online (Burgess, Mitchell, and Münch 2018; Ueda et al. 2017). In some cases, such events bring interest to the topics associated with people's fights during their lifetime. For example, known for her key role in Star Wars, Carrie Fisher struggled with bipolar disorder and addiction and remained a dedicated advocate for mental health.

Through semantic network analysis, E. L. Cohen and C. Hoffner (2016) found that her death brought a significant wave of praise for her openness and lifelong advocacy of mental health issues on social media X (formerly Twitter) with the #InHonorOfCarrie hashtag. Similarly, the suicides of actor Robin Williams and DJ Avicii came as a shock and brought waves of posts related to suicide prevention, depression and addiction (Niederkrotenthaler, Till, and Garcia 2019; Park and C. A. Hoffner 2020). Such sharing of educational content is an example of prosocial behaviour that was observed in other survey studies (Bae, Brown, and Kang 2010; Myrick 2017; Myrick and Willoughby 2019).

Additionally, studies have shown that there is a positive celebrity death effect on stock returns, especially for celebrities who received greater media attention post-death and those who died unexpectedly (Jansen 2016; Lepori 2021). Such deaths also greatly boost the sales of products related to those who just passed away. While the type and context of death impact the market response, marketers adapt quickly to match the demand for such products (Radford and Bloch 2013). For instance, there were five times more Spotify streams of David Bowie's music in the 9 hours following his death announcement compared to the 48 hours that preceded them (Yanofsky 2016), Michael Jackson's passing increased his album sales to record levels (Kaufman 2009), and the manufacturer of Steve Jobs' iconic black mock turtleneck shirts faced difficulties keeping up with the demand following the CEO's death in 2011 (Yin 2011). As Forbes outlines in its yearly report of the highest-paid dead celebrities, this market effect persists throughout the years (Dellatto 2023). Although he died almost 50 years ago, Elvis Presley is estimated to still earn $100 million per year in product sales. Part of these earnings are prosocial behaviours in the form of donations to organisations associated with the celebrity, as observed in a 2019 survey following the death of American actress Mary Tyler Moore (Myrick and Willoughby 2019). A good example is the tragic and unexpected death of internationally renowned conservationist Steve Irwin whose donations to his organisation reached $2 million in the month following his death (*Irwin donations top two million* 2006). The intense reactions outlined above can be explained by the intimate one-way connection that fans feel towards celebrities during their lifetime. Their passing marks the end of a close relationship, a parasocial relationship, the meaning of which is explained in the following paragraph.

2.3 Parasocial relationships

Initially conceptualized by Horton and Richard Wohl (1956), parasocial relationships have since been extensively covered in the literature. They are defined as a "one-sided mediated relationship that people may experience with similar emotional strength to personal relationships" (Gach, Fiesler, and Brubaker 2017, p. 1). Fans cultivate an intimate relationship with celebrities, knowing them on a sometimes personal level without reciprocity. While they existed before the internet era, parasocial relationships have been largely explored with the rise of online news and social media. With this recent phenomenon, parasocial bonds between fans and celebrities are strengthened. Highly publicised deaths bring a large public response online which exemplifies the relationships as fans grieve the deceased. While there are disagreements and misunderstandings among online users as to why and how to mourn those who passed on (ibid.), there is strong evidence of parasocial responses to celebrities' deaths. Numerous studies cover the social reaction following the death of celebrities in various fields. There is evidence of parasocial breakups in the music industry with Michael Jackson (Sanderson and Hope Cheong 2010), David Bowie (Burgess, Mitchell, and Münch 2018; Gach, Fiesler, and Brubaker 2017; Van den Bulck and Larsson 2019), Alan Rickman, Prince (Gach, Fiesler, and Brubaker 2017), or Avicii (Niederkrotenthaler, Till, and Garcia 2019); but also in sports with the deaths of race car driver Dale Earnhardt (Radford and Bloch 2012), and basketball player Kobe Bryant (Bingaman 2022); in cinema with Indian actor Om Puri's unexpected passing (Rajan and Sarkar 2018), Robin William's suicide (E. L. Cohen and C. Hoffner 2016), Carrie Fisher's death (Park and C. A. Hoffner 2020); and even in more niche fields such as wildlife conservation with Steve Irwin's deadly accident (Brown 2010). Through surveys or social media analyses on Reddit, X and Facebook, these studies all outline an outpouring of grief and bereavement following the passing of celebrities. According to Terror Management Theory (TMT), a psychological theory exploring how humans deal with the existential fear that arises from the awareness of their mortality, much of human behaviour is driven by the desire to manage this fear of death (Greenberg, Pyszczynski, and Solomon 1986). Supported by over 500 studies to date, TMT suggests that the death of a celebrity is a reminder of the fragility of life and people's own mortality. Positive portrayals of deceased celebrities can help the public manage their existential anxiety. Commemorating the achievements and positive contributions of the deceased in the media highlights the lasting impact individuals can have, providing a sense of continuity and meaning beyond death to the living (Greenberg, Vail, and Pyszczynski 2014). In that sense, the news media take the role of "national healers" (Kitch 2000, p. 189).

Nevertheless, a recurring debate surfaces whenever the event concerns celebrities with a controversial background (Burgess, Mitchell, and Münch 2018). For instance, Bowie's death was followed by criticism of his alleged past illegal sexual relationships on social media X (Gach, Fiesler, and Brubaker 2017). Similarly, Michael Jackson faced numerous scandals throughout his life regarding allegations of child sexual abuse, his changing appearance due to extensive plastic surgery, and his unconventional lifestyle as the owner of exotic pets and a theme park at his home. These behaviours brought extensive media attention and criticism reappeared at the time of his death with tweets condemning him or even stating "being glad that [he] had died" (Sanderson and Hope Cheong 2010, p. 334). This comes in sharp contrast with the majority who perpetuate the "King of Pop"'s legacy in ways "that one would expect from a disciple of an actual religious figure" (ibid., p. 338). Other fans are faced with the dilemma of whether to continue listening to these artists knowing what they have done. The debate remains even years later with news articles covering past controversies of the deceased and contemplating whether the artist should be separated from the art (Dededer 2023; Freeman 2019; Jonze 2024; Kuo 2023). These comments spark debates on whether it is right to mourn stars who have engaged in morally questionable behaviour or committed crimes in the past. Some argue that separating the artist from their actions allows for appreciation of their contributions to society, and others believe that celebrating such individuals overlooks the harm they may have caused and sends the wrong message to society. Although some instances of negative reactions to the deaths of controversial celebrities were shown, it is likely that those commenters already thought negatively of the celebrities before they passed away. The question remains whether the passing of celebrities improved the discourse around them. Do our perceptions of the dead differ from those of the living?

2.4 The death positivity bias

From the Latin expression "De mortuis nil nisi bonum dicendum est", stems the "Do not speak ill of the dead" norm. This implies that one should speak favourably of the dead, disregarding their past behaviour, whether fruitful or controversial. To put this norm to the test, Heynderickx and Dieltjens (2016) qualitatively analysed 150 obituaries published in the staff magazines of five Flemish companies between 1940 and 2010. The study shows that all the texts are positive overall. Whenever rare occurrences of negativity appeared, those were contrasted by a stronger positive statement about the deceased. A more thorough analysis of obituaries was conducted through two studies by Alfano, Higgins, and Levernier (2018). Those were run on 930 and 74 obituaries collected from local newspapers and The New York Times Obituaries, respectively. The researchers manually assigned characteristic traits of the individual described in each obituary (e.g., "hard-working", "honest", "generous"). They then plotted co-occurrences of traits as an undirected graph with traits as vertices and edges weighted by the number of obituaries with co-occurrences of the connected traits. Colours were used to outline gender differences. While expressed differently depending on the gender of the departed and the obituary source, both studies show that positive terms overshadow negative ones. Though the two research papers presented here provide evidence that we do not speak ill of the dead, they overlook whether individuals were already praised before their passing, and therefore do not address whether death changes the discourse around them.

To test the latter, Allison, Eylon, et al. (2009) explore the concept of *death positivity bias* which claims that individuals tend to biasedly speak more favourably of those who passed away (Allison and Eylon 2005). To do so, they run five studies in different settings. Study 1 is an A/B test in which both groups ($n = 56$) are presented with a vignette describing a fictional leader's achievements. The control group is told that this leader is still alive while the treatment group is informed that he has passed away. Participants in both groups are then asked to fill out a questionnaire assessing their overall liking of the character. An analysis of variance (ANOVA) showed that the treatment group formed more favourable impressions overall, and considered him more inspiring and more competent. In a second study, the researchers show that appreciation also goes up as a group of 66 students answers the same questionnaire before and after learning about a fictional leader's death. While this is true for laypersons, similar results are found for real-world celebrities in a third study. The researchers collected and labelled as positive or negative a sample of 697 news articles about 8 deceased celebrities across politics, entertainment and athletics, 4 years prior to and 4 years after their deaths. Results from ANOVA and chi-squared tests suggest that media sources may also be subject to the death positivity bias. Hayes (2016) conducted research that provided additional evidence in favour of the death positivity bias. In a first study, 72 participants were randomly assigned to groups for A/B testing. Some were requested to picture a person they knew and was still alive while others were asked to imagine that the person in their mind had died. Similarly to previous research, each group was asked to answer open-ended questions and Likert-scale ratings about their relative. ANOVA tests showed that participants in the death group used more favourable descriptors of their chosen relative than those in the other group. These findings remain valid whether the target was distant or close to the surveyed individual. A second study on 101 individuals in a similar setting extended this conclusion to disliked relatives.

While most of the literature introduced so far supports the death positivity bias, some evidence suggests it does not always hold. Allison, Eylon, et al. (2009)'s Study 4 shows that a leader's pre-death moral standing impacts their post-death popularity. Indeed, when presented with the vignette of a deceased fictional leader who had behaved immorally, survey participants rated him more unfavourably compared to leaders having not faced controversy. It should be noted that this study is the only one that finds evidence against the death positivity bias by actually comparing the deceased against the living.

The remaining works presented below suggest that the bias might not hold but only looking at post-death writings. A qualitative analysis of 505 comments related to 142 suicides on MyDeathSpace.com showed that the majority of posts focused on "judgment, speculation, joking, disgust, stereotyping, and a general lack of sympathy" (Leonard and Toller 2012, p. 401) with varying degrees of violence. However, these findings should be interpreted cautiously since most communications on the platform are conducted anonymously. Another study on 63 Romanian celebrities who passed away between 2013 and 2016 finds some additional evidence against the theory of death positivity bias (Rusu 2020). 1148 articles were collected within one week after celebrity death from three news sources: a news agency–Mediafax, a broadsheet gazette–Gândul and a tabloid newspaper centred on celebrities' lives and gossip–Libertatea. The researchers assigned each article a label from one of five categories (negative-condemnatory, negative-critical, neutral-informative, positive-conventional, positive-eulogizing). They found that almost one in four articles portrays a negative evaluation of the deceased and 13.9% critically assess them. The latter goes up to 36% if selecting only celebrities that have a controversial reputation. The counts of articles were also compared against other variables by running chi-squared tests to test for differences in article frequencies. The researchers found differences in response depending on the type of news source. Evaluations were becoming more negative from news agency, through broadsheet gazette, to tabloid newspaper. This result was expected as news agencies tend to be more formal and neutral/positive in comparison to the tabloid newspaper which convey more sensationalist news and scandals, labelling it as "deathtertainment" (ibid., p. 588). The profession of the deceased also had an impact. Businesspersons received less obituary coverage and fewer eulogising assessments compared to stars of popular culture (18.5% vs. 39.6%).

2.5 Thesis objectives

As shown in this review, there has been extensive coverage of celebrity deaths in the literature. Most studies make a qualitative assessment of social media posts or media articles by manually grouping them and retrieving the key themes that stem from the passing of celebrities (Burgess, Mitchell, and Münch 2018; Gach, Fiesler, and Brubaker 2017; Rajan and Sarkar 2018; Sanderson and Hope Cheong 2010). Another paper takes a semi-quantitative approach by manually labelling media articles into sentiment groups and running chi-squared statistical tests to understand what celebrity characteristics can help explain variations in response (Rusu 2020). However, these studies all assess post-death texts without comparing them to their pre-death counterparts. Only two papers by Allison, Eylon, et al. (2009) and Hayes (2016) do make the comparison. Both studies are mostly survey-based and centred around fictional leaders or relatives of the participants. Nevertheless, only a section of Allison, Eylon, et al. (2009)'s paper extends the findings to real-world celebrities by fetching media articles and running ANOVA and chi-squared tests with some limitations that should be outlined. First, articles are manually labelled as either positive or negative which fails to capture the nuance in discourse. Second, sample sizes are small. Only 8 celebrities are assessed through 290 and 407 pre- and post-death articles respectively. Some individuals have as few as 10-15 pre-death articles. Results are therefore hardly generalisable to the wider celebrity sphere. Third, no covariates were used in the analysis. Those would make the analysis more robust by potentially capturing some of the observed effect. The current thesis proposes a novel approach to assessing changes in discourse around celebrity deaths by addressing the limitations of previous research. It does so taking a quantitative approach from start to end, in contrast with previous qualitative papers. A sentiment algorithm is used to compute a continuous sentiment scale and capture appreciation on a more granular level. Post-death assessments are done in comparison with pre-death to ensure that post-death positiveness is not merely due to an already existing pre-death appreciation, but rather to an actual change with the event of death. While previous studies covered at most 8 celebrities, 38 are analysed here through a total of 7600 articles. Lastly, the greater sample size allows for comparisons of effect size across groups of celebrities based on their individual characteristics. The objective of this thesis is, not only to uncover evidence of the death positivity bias, but also to understand whether the signal is stronger for some types of individuals, and to see if article content can help understand the shift in narrative.

3 Methodology

As shown previously, comparisons between pre-death and post-death sentiments around celebrities have not been covered extensively in the literature. The methods used here are novel, making this research preliminary.

3.1 Data collection

3.1.1 Targeted celebrities

The search for celebrity figures was conducted by reviewing lists of celebrity deaths in recent years. Those include but are not limited to, Wikipedia's *Lists of people by cause of death* (n.d.) or movie reviewer IMBd's *Advanced name search* (n.d.) and *In Memoriam: Deaths in the 2010s* (n.d.). To narrow down the search, celebrities were selected based on two criteria: (1) death date, and (2) Google web searches:

1. **Death date:** The API used to collect data does not go back further than January 1st, 2014 (see Section 3.1.3). To ensure at least 6 months of media coverage before death, the celebrity must have passed away after June 2014. Similarly, since data was collected in May 2024 and to ensure at least 6 months of media coverage after death, the individuals must have died no later than October 31st, 2023.

2. **Google web searches:** To ensure sufficient media coverage around those celebrities before and after their death, the interest in celebrities was measured using Google Trends[1]. The tool provides data on the search queries of users around the world. Thanks to Google's Named Entity Recognition (NER) feature, monthly interest over time can be downloaded for each celebrity, as shown in Figure 1. Celebrities are selected only if their Google Trends interest is consistently greater than 0% between January 2014 and the query date. This ensures that the celebrity attracted sufficient public interest to be covered in the media before and after their death. Some exceptions can be found if a young celebrity went viral for the first time after 2014. For instance, American rapper Juice Wrld saw sudden popularity following

[1]Google Trends: `https://trends.google.com/trends/`

the release of two songs that climbed the American Billboard Hot 100 chart in 2018 (*Juice WRLD* n.d.). He died of an overdose shortly after in 2019, at the age of 21.

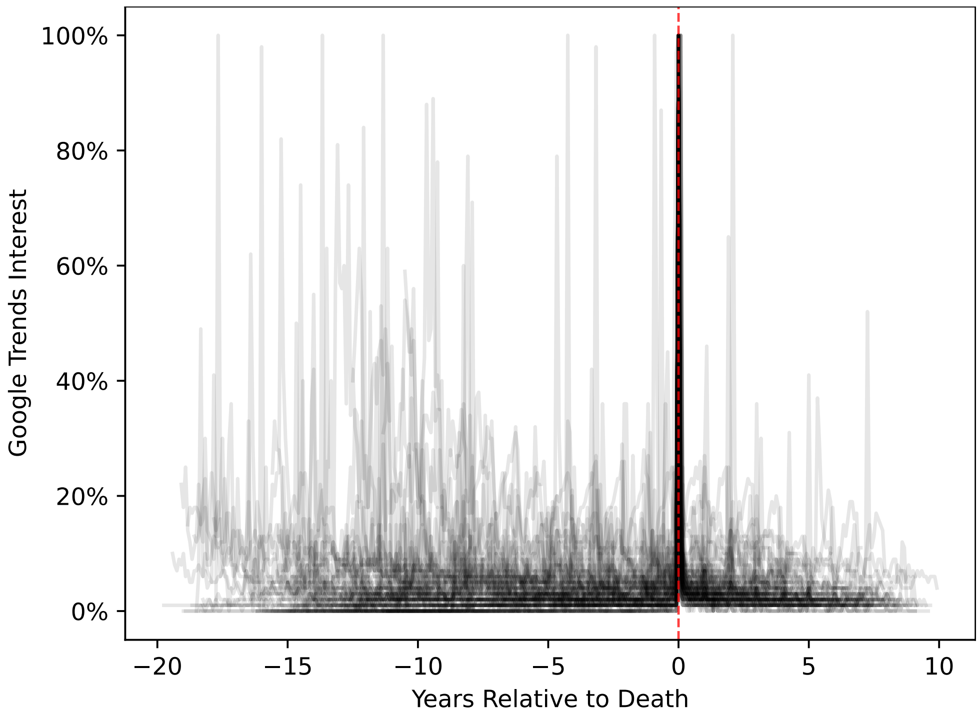

Figure 1: Time series of monthly Google Trends interest relative to death date for all the celebrities (N = 80)

Each line corresponds to one celebrity. The x-axis shows the number of years from the death date, at value 0. The y-axis shows the Google Trends monthly interest as a percentage of the maximum value for each celebrity.

3.1.2 Feature selection

The following 7 features were collected for each celebrity:

- nationality
- gender
- industry
- age at death
- cause of death
- whether the death was unexpected/sudden (dummy variable: 1 if unexpected, 0 otherwise)

- whether the individual faced controversy during their lifetime (dummy variable: 1 if controversial, 0 otherwise)

An attempt was made to extract some of those features (i.e., industry, age at death) by pulling data from DBpedia[2], a database that provides Wikipedia data in a structured way using the SPARQL query language. Due to the inconsistent structure of Wikipedia pages across celebrities, this method led to numerous bugs. Accessing Wikipedia pages using the official API also turned out to be time-consuming and prone to errors. For faster processing, it was more convenient to manually collect and label data from Wikipedia directly. Manually labelling data comes with challenges. Justifications for the choices made and more detailed descriptions of the features are provided in Section 4.1.

3.1.3 Event Registry API

News articles about celebrities are collected by using an API provided by Event Registry[3]. Event Registry is a news intelligence platform that collects, organizes, and analyzes news articles from thousands of RSS feeds around the world in real time. It serves as a powerful tool for researchers, journalists, and businesses to monitor and understand global events, trends, and sentiments associated. The platform aggregates articles from various sources, including news websites, press releases, and blogs since 2014. They provide additional features such as NER and various NLP methods to group articles by topic, as well as advanced search capabilities through their API[4].

Event Registry's built-in NER feature associates each entity with a Wikipedia page. In this way, articles can be easily targeted by providing celebrities' Wikipedia pages to the queries. Even if a celebrity is mentioned only once in an article, the entity concept of that individual becomes associated with the article. The issue with this approach is that articles which mention a celebrity one time only are still retrieved, even though most of the content is unrelated. For instance, an article which covers films that stood

[2]DBpedia: `https://www.dbpedia.org/`
[3]Event Registry: `https://eventregistry.org/`
[4]Event Registry News API: `https://www.newsapi.ai/`

out at the Oscars[5] mentions basketball star Kobe Bryant once in its introduction. To prevent the retrieval of such an article, an additional condition is added to the query. The first and last names of celebrities are used as keywords to look for in the title of articles. This ensures that all articles retrieved are directly related to the individual of interest. Also, Event Registry's NER feature sometimes incorrectly identifies celebrities. For instance, using musician Prince's Wikipedia page as an entity retrieves articles related to princes of the British Royal Family[6]. In such confusing case, the Wikipedia page title typically contains a word between parentheses describing the star's occupation (e.g., "Prince (musician)"[7]). To limit such errors, whenever the Wikipedia page title contains parentheses, the celebrity's industry is used as an additional entity concept to detect in articles. In this case, "Music", associated with the Music Wikipedia page[8], is added as a condition to the query. Another example is "David Edwards (basketball)"[9] for whom the "Sport"[10] industry was added to the query. Details of all the API parameters can be found in Table 1. Table 1 also provides an example query to search articles about American singer Mac Miller.

[5]Example article unrelated to Kobe Bryant: https://www.cnnbrasil.com.br/esportes/de-rocky-a-nyad-10-filmes-de-esporte-que-se-destacaram-no-oscar/

[6]Example article unrelated to Prince (musician): https://www.foxnews.com/entertainment/king-charles-struggles-evict-prince-andrew-disgraced-royals-home-disrepair-experts

[7]Prince (musician) Wikipedia page: https://en.wikipedia.org/wiki/Prince_(musician)

[8]Music Wikipedia page: https://en.wikipedia.org/wiki/Music

[9]David Edwards (basketball) Wikipedia page: https://en.wikipedia.org/wiki/David_Edwards_(basketball)

[10]Sport Wikipedia page: https://en.wikipedia.org/wiki/Sport

Function	Parameter	Example value	Explanation
QueryArticlesIter	keywords	ER.QueryItems.OR([‘Mac’, ‘Miller’])	Look for either of the keywords.
	conceptUri	`https://en.wikipedia.org/wiki/Mac_Miller`	Use NER to look for articles associated with the concept of the celebrity’s Wikipedia page.
	categoryUri	er.getCategoryUri(‘music’) if ‘(’ in ‘Mac Miller’ else None	Look for a specific industry concept in case of parenthesis in the name (it does not apply here).
	lang	‘eng’	Find articles in English.
	dateStart	‘2017-01-01’	Start date of the search period.
	dateEnd	‘2017-12-31’	End date of the search period.
	keywordsLoc	‘title’	Look for the provided keywords in the title of the articles, rather than the body, to ensure that articles are solely about the celebrity of interest.
	isDuplicateFilter	‘skipDuplicates’	Skip the articles that are duplicates of other articles.
	dataType	[‘news’, ‘pr’, ‘blog’]	Search for all kinds of articles: news, press releases, and blogs.
execQuery	sortBy	‘rel’	Sort returned articles by relevance.
	articleInfo	ER.ArticleInfoFlags(sentiment=True)	Include the sentiment of articles.
	maxItems	100	Maximum number of articles to return.

Table 1: API parameters of the *ER.QueryArticlesIter* and *q.execQuery* functions

ER.QueryArticlesIter is a function in the *eventregistry* module imported as *ER*. It returns a query object *q*.
q.execQuery is a method called on the query object *q*. It returns the queried articles in a json format.
The example query returns the 100 (or less) most relevant English articles in 2017 associated with the concept of Mac Miller and mentioning either his first or last name in the title. These articles can be of type news, press release, or blog, and come with an associated dictionary sentiment.

3.1.4 Selection of the periods of interest

Articles were needed both before and after the death of each celebrity. Due to the limited access to API tokens, accessing all articles across all RSS feeds until today would have been too costly. For this reason, two periods of media articles were selected for each celebrity, one pre-death and another one post-death. The logic behind the period selection is explained in the following paragraphs and illustrated in Figure 2.

Pre-death period The pre-death period is selected based on the Google Trends interest surrounding celebrities during their lifetime. The year with the highest peak is selected between January 2014 (i.e., the start of Event Registry's search period) and the celebrity's death. This peak highlights an important event related to the person's career and is likely to become what the celebrity will be remembered for. A full year is selected for three reasons. First, this ensures that the peak is fully captured. Indeed, most viral phenomena do not happen from one day to the next and hence cannot be captured in the span of a few days or weeks. For instance, Leonard Cohen's release of his 13th album, Popular Problems, on September 24th 2014 brought a spike in web searches that lasted for several weeks. Second, one month is often insufficient to fetch enough articles about each individual. Third, selecting a full year ensures that other relevant events would also be covered. Those include the celebrity's birthday and any other yearly anniversary related to the individual. Event Registry's API's tokens are consumed for each search year. To save tokens, instead of selecting a fixed time window around the peak (e.g., $\pm$ 6 months), the selected year goes from January 1st to December 31st. It should also be noted that Google Trends provides a tool that tracks News Searches[11] rather than Web Searches[12]. While the former seems to be a good indicator of media attention, it only tracks searches made by users on the Google News search bar. The search volumes are therefore much lower and only a minority of news readers use the platform. This makes Google Web Searches more general-purpose and a better indicator of celebrity interest.

[11] Google News: `https://news.google.com/`
[12] Google Search: `https://google.com/`

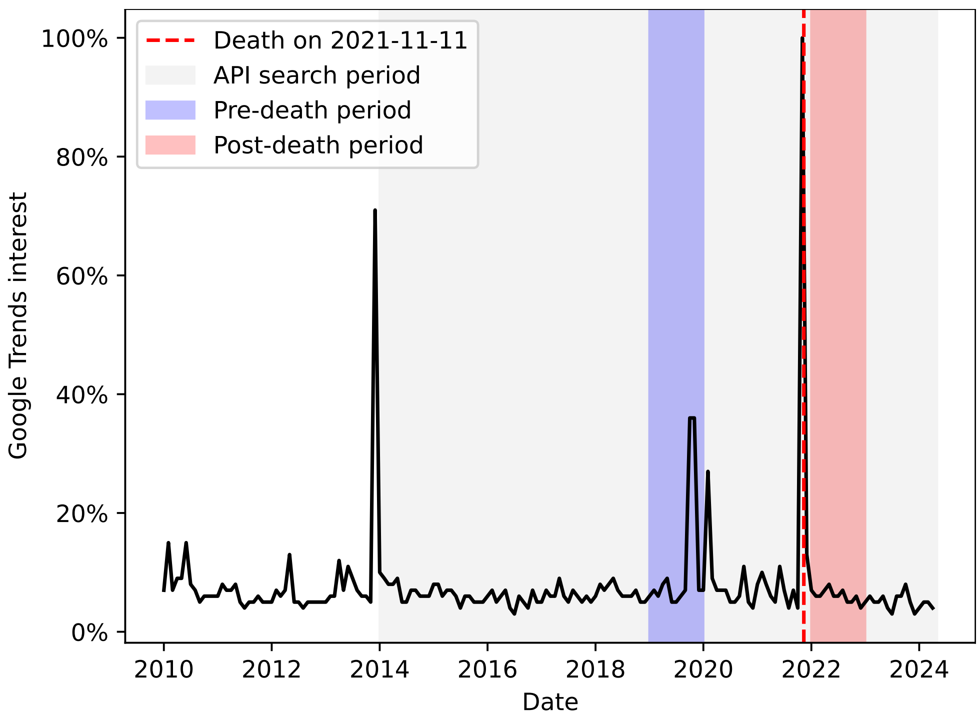

Figure 2: Example of pre-death and post-death period selection for F.W. de Klerk

The black line represents the Google Trends Interest related to South African politician F.W. de Klerk. The vertical red dotted line marks his date of death. The blue and red bands mark the selected pre- and post-death periods for that celebrity.

Post-death period Celebrity deaths systematically bring a huge spike in interest. As shown in Figure 1, the highest peak of Google Trends interest typically matches the person's death, except in some cases (e.g., Pope Benedict XVI's election in April 2005 and abdication in February 2013 brought more interest than his death in late 2022). On X alone, Swedish DJ Avicii's and rockstar David Bowie's deaths brought huge traction with over 300,000 and more than 1 million tweets per hour respectively in the immediate aftermath (Niederkrotenthaler, Till, and Garcia 2019; XUK 2016). While this is a period in which media attention is at its highest, news articles around this time use terminology mostly focused on death, funeral, bereavement and other negatively connoted lexicon. This phenomenon was also detected in Niederkrotenthaler, Till, and Garcia (2019)'s study where a peak in negative emotions was observed on X, primarily due to wording related to sadness. Although post-death articles typically retrace celebrities' achievements during their lifetime, sentiment models could fail to identify positive discourse in such a context. For this reason, the post-death period was postponed to one year after death. This period makes sense for two reasons. First, interest has likely moved away from the funeral and sorrow, to focus more on remembrance of the deceased. Second, this marks the first anniversary of the death and some media publish articles commemorating celebrities. Such articles focus on the celebrities' achievements and potential controversies throughout their lifetime, which is the focus of the current study. Finally, a full year from January 1st to December 31st is selected for the reasons outlined in the previous paragraph.

An API request was sent for each celebrity and time period. If less than 100 articles were available in either period, additional pre-death and post-death periods of 1 year were added before and after the initial ones to try and meet the minimum count goal. If there were still less than 100 articles per period after several rounds of queries in different years, the celebrity was dropped from the analysis. On the contrary, some celebrities receive high media attention (i.e., more than 100 articles available before or after death). To ensure popular stars were not overrepresented and limit token usage, the number of articles for each celebrity and each period was capped at 100 based on relevance (each article retrieved by the API is automatically assigned a relevance score based on the relevance of the match with the query). Only the 100 articles with maximum relevance are fetched.

3.2 Sentiment analysis

Event Registry automatically assigns a sentiment score to each fetched article using a dictionary-based approach (Hardeniya and Borikar 2016). This method uses a predefined lexicon of words associated with specific sentiment values. After preprocessing, each word of an article is matched against the sentiment lexicon: positive words contribute to a positive score, and negative words contribute to a negative score. The overall sentiment is determined by averaging the scores across all the words in the article. Although this method is simplistic and has some limitations, it is a convenient and fast way to extract sentiment from text. In this case, scores are computed on a continuous scale ranging from -1 for most negative to +1 for most positive.

3.3 Evidence of the death positivity bias (RQ1)

This section addresses RQ1: Does the media speak more positively about celebrities once they have passed away?

3.3.1 Mann-Whitney U-test

Statistical tests are used to compare the pre- and post-death sentiment distributions quantitatively. The t-test is the simplest option to compare means between two groups (Kim 2015). However, it should only be used when the data of the two groups are approximately normally distributed. While this assumption can be tested using the Shapiro-Wilk test (Shapiro and Wilk 1965), a more convenient approach is the Mann-Whitney U-test (Mann and Whitney 1947). A non-parametric test is preferred here as it is not assumed that the data follows a known distribution. Sometimes viewed as the non-parametric equivalent of the t-test, the U-test tests a slightly different null hypothesis where the distributions of both groups are identical. In other words, there is a 50% probability that an observation randomly drawn from a distribution X is greater than an observation randomly selected from a distribution Y. To do so, values in both groups are ranked from lowest to largest. The ranks are then averaged in both groups and the test computes a p-value that depends on the discrepancy between the mean ranks of the two groups. A greater discrepancy leads to a lower p-value.

Computation The U-statistic is given by:

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \tag{1}$$

where R_1 is the sum of the ranks for sample 1, and n_1 is the number of observations in sample 1. Note that it does not matter which of the two samples is used as sample 1. The formula can also be written as:

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} \tag{2}$$

The smaller value of U_1 and U_2 is used as the final U statistic to compute the p-value associated with the test.

Assumptions The following four assumptions must be met to run a Mann-Whitney test:

1. The dependent variable is continuous or ordinal. This assumption is respected since sentiment is measured on a continuous scale ranging from -1 to 1, and sentiment scores can be ordered.

2. The independent variable is categorical, it consists of two groups and each observation is assigned to exactly one of those groups. In the current case, articles are assigned to one of two groups: pre-death or post-death based on the date on which they were published relative to the death date.

3. Observations are independent from one another. There is no relationship among observations in each group, or between the groups themselves. While there is no formal way to test this assumption, the current study design suggests no such relationship.

4. The samples are randomly drawn from the population. As covered in Section 3.1.4, articles with the highest relevance measure were selected to maximise celebrity representativeness. In that sense, one might argue that article selection is not fully randomized. However, fully randomizing the search would have resulted in articles which are not representative of the individuals they portray.

Hypotheses A one-sided Mann-Whitney test is run to answer RQ1. It is summarised in the following set of hypotheses:

$$H_0: \quad P(Y > X) \leq P(X > Y)$$
$$H_1: \quad P(Y > X) > P(X > Y)$$

where X and Y are the pre- and the post-death sentiment distributions respectively. A small p-value would suggest that post-death and pre-death article sentiments are drawn from different distributions and that the former tends to be more positive. In other words, the media would speak more favourably about celebrities once they have passed.

In simple terms, the one-sided version of the test detects whether post-death sentiment is on average greater than pre-death sentiment. This version of the test directly answers RQ1. However, it does not detect celebrities for whom the relationship might be reversed. For completeness, a two-sided version of the test is also run. It can be summarised as follows:

$$\begin{aligned} H_0: &\quad P(Y > X) = P(X > Y) \\ H_1: &\quad P(Y > X) \neq P(X > Y) \end{aligned}$$

3.3.2 Effect size

Statistical tests are useful in understanding differences between groups. While a low p-value tells us whether an effect exists or if there is a significant difference between groups, they do not inform us about the practical importance of the difference. A tiny difference between groups could be statistically significant due to a large sample size, but might not be so important in a real-world context. Effect size measures can help us distinguish between statistically significant and practically important results. Sentiment analysis is very complex due to the inherent context, nuance, ambiguity and other challenges found in natural language. For this reason, the sentiment score assigned to articles is subject to some variability. For instance, on a scale from -1 to 1, there is arguably no difference between scores of 0.2 and 0.22. However, a statistical test might consider this difference statistically significant due to a large sample size. Measuring effect size is coherent in this context.

There are various parametric and non-parametric measures of effect size, the most famous being Cohen's d (J. Cohen 2013). Some measures are characteristic of specific statistical tests. When it comes to Mann-Whitney U analysis, the *rank biserial correlation* coefficient r is the most appropriate (Cureton 1956). This non-parametric measure is calculated based on the U-statistic and sample sizes. The correlation coefficient ranges from -1 for a perfect negative relationship (all the values of the first sample are smaller than all the values of the second sample), to 1 for a perfect positive relationship (all the values of the first sample are greater than all the values of the second sample), through 0 when there is no relationship. Since there are no pre-defined thresholds for r in the current field of study, J. Cohen (2016, Table 1, p. 157)'s are used. Those are .10, .30, and .50 for small, medium and large effect sizes respectively.

The rank biserial correlation coefficient is calculated as follows:

$$r = \frac{2U_1}{n_1 n_2} - 1 \tag{3}$$

or

$$r = 1 - \frac{2U_2}{n_1 n_2} \tag{4}$$

where U_1 and $U2$ are the Mann-Whitney U-statistics, and n_1 and n_2 are the sample sizes of group 1 and group 2 respectively. In this case, group 1 represents pre-death and group 2 is post-death.

3.3.3 Bootstrap of the test statistic and effect size

Bootstrapping is a statistical method used to strengthen the estimate of a statistic of interest by computing confidence intervals around it (Wasserman 2013, p. 107). It involves B iterations of random sampling with replacement to create B simulated samples. The statistic of interest is computed on each of these samples and its final estimate is computed by taking the mean statistic across the B samples. Due to the Central Limit Theorem, the sampling distribution obtained tends towards a normal distribution and a Normal confidence interval for the estimate can be computed using the z-score. In the current context, the method is used to make the estimate of effect size more robust. Confidence intervals are computed for the estimated biserial rank correlation coefficient r, as motivated by Banjanovic and Osborne (2019).

Computation The method is summarised in the following 6-step procedure:

1. Draw two samples of size $n_1 = 100$ and $n_2 = 100$ with replacement from the original pre-death and post-death distributions X and Y:

$$X_b^* = \{x_1^*, x_2^*, \ldots, x_{n_1}^*\} \sim X = \{x_1, x_2, \ldots, x_{n_1}\} \tag{5}$$

$$Y_b^* = \{y_1^*, y_2^*, \ldots, y_{n_2}^*\} \sim Y = \{y_1, y_2, \ldots, y_{n_2}\} \tag{6}$$

2. Calculate the statistic of interest (e.g., U-stat, r) with the bootstrap samples:

$$\hat{\theta}_b^* = f(X_b^*, Y_b^*) \tag{7}$$

3. Repeat steps 1 and 2 B times to obtain $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$:

4. Compute the estimate:

$$e_{boot} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_b^* \tag{8}$$

5. Compute the standard error of the estimate:

$$se_{boot} = \sqrt{\frac{1}{B} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - e_{boot} \right)^2} \tag{9}$$

6. Compute the Normal confidence interval around the estimate:

$$C_{boot} = e_{boot} \pm z_{\alpha/2} se_{boot} \tag{10}$$

The methods described in Section 3.3 are computed on the whole dataset of articles and then repeated at the celebrity-level (i.e., one test is run for each celebrity separately).

3.4 Variations in media response (RQ2)

This section addresses RQ2: Which attributes of celebrities can account for variations in media response?

3.4.1 Article sampling and pairing

The second research question explores how celebrity characteristics can help explain variations in media response following their death. In other words, is the difference between pre- and post-death sentiment greater for some groups of individuals? This is answered using an ANOVA, and Tukey's Honestly Significant Difference (HSD) tests (see Sections 3.4.2 and 3.4.3 respectively for details). The interest is on the effect size of the difference. However, three issues can be noted. First, there is only one value of effect size for each celebrity. While it is a summary statistic of the two distributions, it does not capture them in their entirety and some information is lost. Second, celebrities are grouped by the features introduced in Section 3.1.2. However, as shown in Section 4.1 on the celebrities dataset, some features are highly unbalanced. For instance, while the cinema industry group comprises 13 celebrities, sport and academia only contain 5 and 1 respectively. Similarly, 14 people died of illness whereas less than half of that died of accident, suicide, natural causes, and assassination combined. Testing differences in effect size for groups of less than 5 celebrities would not provide statistically convincing results. For these reasons, the ANOVA and HSD tests are run at article level rather than celebrity level. This is done by randomly pairing pre- and post-death article sentiments for each celebrity, and computing their difference as detailed in Algorithm 1.

It should be noted that the random samplings in lines 6 and 7 are done without replacement to ensure that articles are paired exactly once. The subsequent ANOVA and Tukey's HSD tests are run using the difference distribution D as the dependent variable. The total number of observations goes from n to $n \times 100$. This ensures that small groups of celebrities have large enough sample sizes (i.e., at least 100). Furthermore, more information is captured by the distribution in comparison with a celebrity-level analysis. Although this method greatly increases sample size, results should be analysed with care

Algorithm 1 Article sampling and pairing

$c \leftarrow$ list(celebrities) ▷ c: list of celebrities
$n \leftarrow length(c)$
$d \leftarrow$ array$(n, 100)$ ▷ d: empty array for sentiment differences
for $i \leftarrow 1$ to n **do**
  for $j \leftarrow 1$ to 100 **do**
    $x_{i,j} \leftarrow random\_sample(X_i)$ ▷ X: pre-death sentiment distribution
    $y_{i,j} \leftarrow random\_sample(Y_i)$ ▷ Y: post-death sentiment distribution
    $d_{i,j} \leftarrow y_{i,j} - x_{i,j}$
  end for
end for
$D \leftarrow flatten(d)$ ▷ D: Aggregate sentiment difference distribution

since small groups are still over-represented by only a few celebrities.

3.4.2 ANOVA test

An ANOVA test is run to differentiate article sentiment pre- and postmortem. The ANOVA test allows a comparison of more than two groups simultaneously to understand whether there exists a relationship between them. There are 3 types of ANOVA tests: sequential (Type I), hierarchical (Type II) and unique (Type III). Those vary in their approach to comparing the Sum of Squares of each factor j (SSF_j) to other factors. Each has its drawbacks. In Type I, the order in which features are added to the model matters when computing interactions. Type II assumes that there are no interactions among features. Type III compares SSF_j with a model containing all other effects through interaction terms but without the sequential constraint of Type I. In this context, Type I is inappropriate since there is no logical or temporal component to celebrities' factors, hence the order in which factors are added to the model should not matter. Type I is also inappropriate for unbalanced data, which is the case across several celebrity features (see Section 4.1). While Type II assumes no interactions among model features, Langsrud (2003) shows that there are clear limitations to Type III and that Type II also has more statistical power (i.e., the probability that the test correctly rejects the null hypothesis H_0 when the alternative hypothesis H_1 is true). For these reasons, type II is used in this experiment. The ANOVA test's significance is measured through an F-statistic for each feature, which is calculated according to the procedure described in the next paragraph.

Computation Assume there are n observations and p factors denoted as $F_1, F_2, \ldots, F_p$ split among $k_1, k_2, \ldots, k_p$ levels respectively.

1. Total Sum of Squares:

$$SST = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \tag{11}$$

where Y_i is the value of Y for observation i, and $\bar{Y}$ is the grand mean of all observations.

2. Sum of Squares for each factor j:

$$SSF_j = \sum_{i=1}^{k_j} n_{ij} (\bar{Y}_{ij} - \bar{Y})^2 \tag{12}$$

where $\bar{Y}_{ij}$ is the mean of Y for level i of factor j, and n_{ij} is the number of observations for level i of factor j.

3. Error Sum of Squares:

$$SSE = SST - \sum_{j=1}^{p} SSF_j \tag{13}$$

4. Degrees of Freedom for each factor j:

$$dfF_j = k_j - 1 \tag{14}$$

5. Degrees of Freedom for error:

$$dfE = n - \sum_{j=1}^{p} k_j + (p - 1) \tag{15}$$

6. Mean Squares for each factor j:

$$MSF_j = \frac{SSF_j}{dfF_j} \tag{16}$$

7. Mean Squares for error:

$$MSE = \frac{SSE}{dfE} \tag{17}$$

8. F-statistic for each factor j:

$$F_j = \frac{MSF_j}{MSE} \tag{18}$$

Each F-statistic follows an F-distribution with dfF_j and dfE degrees of freedom. p-values can be computed from the F-statistics. A low p-value indicates a high variability between the k_j groups of that factor F_j and a low variability within them. In other words, there is evidence to reject the null hypothesis of no difference between the k groups. It suggests that at least one group mean is significantly different from the others at the chosen confidence level.

To measure the contribution of each factor j to the model, an effect size, also known as *partial eta squared*, is calculated for each factor j:

$$\eta_j^2 = \frac{SSF_j}{SSF_j + SSE} \tag{19}$$

η_j^2 measures the proportion of variance in the dependent variable associated with feature j in the ANOVA model. It can be thought of as the correlation between effect j and the dependent variable. By construction, $0 \leq \eta_j^2 \leq 1$. J. Cohen (2013)'s thresholds are used for interpreting the strength of association. Those are .01, .06, and 0.14 for small, medium and large effect sizes respectively.

Assumptions On top of the assumption of no interaction for Type II ANOVA, the following four assumptions must be met:

1. Normality: The data within each group should be normally distributed. However, it was shown that ANOVA tests are not very sensitive to moderate deviations from normality, the false positive rate (i.e., also known as Type I error, it is the proba-

bility that H_0 is rejected when it is actually true in the population) remains rather unchanged (Glass, Peckham, and Sanders 1972; Harwell et al. 1992; Lix, J. C. Keselman, and H. J. Keselman 1996).

2. Homogeneity of variance: The variance of the data within each group should be equal. The variances of each group are computed in Table 3.

3. Independence: The observations within each group should be independent. Again here, while there is no formal way to test for independence, the current study design suggests no such relationship.

4. Random sampling: Observations are sampled randomly. The sampling procedure described in Section 3.4.1 ensures randomisation.

Hypotheses The ANOVA test can be resumed in the below set of hypotheses for each factor F:

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_2 = \cdots = \mu_k \\ H_1 &: \quad \exists\, i, j \text{ such that } \mu_i \neq \mu_j \end{aligned}$$

where:

- k is the number of groups for factor F
- i and j are indices representing different groups, with $1 \leq i, j \leq k$
- μ_i, μ_j are the population means of the i-th group, and the j-th group respectively

In this case, the dependent variable is the difference distribution D obtained from the article pairing procedure described in Section 3.4.1. The ANOVA test tests for differences in D. Groups of celebrities are based on the celebrity features detailed in Section 4.1. An F-statistic close to 0 provides evidence to reject the null hypothesis that all groups of individuals in this factor have the same population mean change in sentiment after their passing.

3.4.3 Tukey's HSD test

ANOVA tests for differences in means across all k groups of a factor. The F-statistic tests for differences in all groups simultaneously, but it does not tell where those differences lie exactly. It is common practice to run post hoc tests to find out which specific pairs of group means differ in cases where features have more than 2 groups. Tukey's Honestly Significant Difference (HSD) test is commonly used as a follow-up to ANOVA where the F-statistic has revealed the existence of a significant difference between some of the tested groups. Tukey's HSD test assesses the significance of differences between all pairs of group means.

Computation The test statistic is given by:

$$q = \frac{|\bar{Y}_i - \bar{Y}_j|}{SE} \tag{20}$$

where $\bar{Y}_i$ and $\bar{Y}_j$ are the means of the two groups i and j being compared, and SE is the standard error for the sum of the means.

A p-value can be computed from the resulting q-statistic using a q-distribution.

Assumptions The test follows the same assumptions as the ANOVA test as described in Section 3.4.2.

Hypotheses Tukey's HSD test tests the below set of hypotheses for each pair of means of each factor F:

$$H_0: \quad \mu_i = \mu_j$$
$$H_1: \quad \mu_i \neq \mu_j$$

where:

- i and j are indices representing two groups of factor F, with $1 \leq i, j \leq k$, and k the number of groups for factor F

- μ_i, μ_j are the population means of the i-th group, and the j-th group respectively

As in the ANOVA test, the dependent variable is the difference distribution D obtained from the article pairing procedure described in Section 3.4.1. Tukey's HSD tests are run for all features that have more than two groups and for which the F-statistic is lower than 0.05 (i.e., there is evidence of a difference in means between at least two of the groups).

3.5 Article content analysis (RQ3)

This section addresses RQ3: What themes are highlighted by the media before and after celebrities' passing? Can they help explain the death positivity bias?

3.5.1 Text pre-processing

To answer RQ3, article contents are analysed by leveraging NLP techniques. Text pre-processing is an essential step in NLP. It involves transforming raw text data into a clean and usable format for machine learning models. The goal is to normalize the text, remove noise, and prepare it for feature extraction or model training. A common preprocessing pipeline is run to clean the articles' text. Text is converted to lowercase, and digits, punctuation, and special characters are removed. The text is then tokenized, stop words are removed, and the remaining tokens are lemmatized with the help of Part-Of-Speech (POS) tagging. POS tagging involves assigning a part of speech tag (e.g., noun, verb, adjective) to each word using their surrounding words as context. Such tagging helps the lemmatizer to do its job. Lemmatization is the process of reducing each token/word to its lemma. By doing so, words with the same root can be grouped together and analysed as a single term (e.g., the words "changing", "changed" and "change" are all converted to their common root: "change").

Additionally, since the articles target celebrities, celebrity names and organisations associated with them are often mentioned. Those entities are however not useful in understanding the content of text. NER is run to filter out such entities. NER identifies words or series of words that refer to a pre-determined category. For instance, in the sentences: "The Harry Potter star, 69, tied the knot with Rima Horton, who he has been dating since they met as teenagers in the 1960s, in an intimate ceremony in New York. He announced the news in an interview with German newspaper Bild, which was published on Thursday (23rd April 2015).", the following entities are identified and removed from the final preprocessed text:

- Harry Potter: PERSON
- Rima Horton: PERSON
- 1960s: DATE
- New York: GPE (geopolitical entity)
- German: NORP (nationalities or religious or political groups)
- Bild: ORG (organisation)
- 23rd April 2015: DATE

Though NER is quite reliable, the algorithm may miss some occurrences of celebrities (PERSON). To ensure no names appear in the preprocessed texts, an additional check drops tokens matching the first or last name still present after NER. After all the preprocessing steps described above, the example sentences become: "star tie knot date since meet teenager intimate ceremony announce news interview newspaper publish thursday".

3.5.2 Word frequency using TF-IDF

For machine learning models to run, words must be vectorized. Term Frequency-Inverse Document Frequency (TF-IDF) is a statistical method used in NLP to reflect the importance of a word in a document relative to a collection of documents (i.e., the corpus of articles). The $TF\text{-}IDF(t, d)$ score of term t in document d is calculated as follows:

1. Term Frequency (TF):

$$TF(t, d) = \frac{\text{Number of occurrences of } t \text{ in } d}{\text{Total number of terms in } d} \tag{21}$$

2. Inverse Document Frequency (IDF):

$$IDF(t) = \log \frac{D}{\text{Number of documents containing } t} \tag{22}$$

where D is the total number of documents in the corpus.

3. Term Frequency-Inverse Document Frequency (TF-IDF):

$$TF\text{-}IDF(t, d) = TF(t, d) \times IDF(t) \tag{23}$$

TF-IDF assigns higher importance to terms that are frequent in a document but less common across the whole corpus. A higher TF-IDF score indicates greater significance of the term t in document d. On the contrary, a term t that is either not very present in document d (low TF), and/or is repeated many times across all documents (low IDF), would receive a lower TF-IDF score since it is not representative of document d. Common words (e.g., “the”, “is”, “and”) appear frequently across all documents and do not carry much meaning. TF-IDF helps reduce the weight of such common words, thereby focusing more on unique and meaningful terms. In that sense, TF-IDF helps reduce noise. The output of TF-IDF is a two-dimensional matrix $\mathbf{V}(D \times T)$ where D is the total number of documents and T is the number of unique terms in the corpus. This matrix is used as input in machine learning models where D is the number of observations and T is the number of features. In the current analysis, D articles make the observations, and unigrams and bigrams make up the T features.

To understand which terms are most representative of either period, the article period is encoded as a dummy variable (1: pre-death, 0: post-death). A logistic regression is trained and tested on the $\mathbf{V}$ matrix with the period as target. By construction, the regression coefficients indicate the importance of each term to help predict whether an article belongs to the pre- or post-death category. Terms with a positive coefficient are most relevant to the pre-death period and those with a negative coefficient are representative of the post-death period. Ranking articles by their coefficient shows what terms are prevalent in one period and not the other.

3.5.3 Topic Modelling

Topic modelling is run to understand the themes highlighted by articles before and after death. Topic modelling is a collection of unsupervised machine learning methods to cluster similar words in text. Non-Negative Matrix Factorization (NMF) is an example of such a method. It is a dimensionality reduction technique whose goal is to identify k latent topics in a corpus. NMF uses factorisation to approximate a non-negative input matrix $\mathbf{V}(D \times T)$ into two non-negative matrices, $\mathbf{W}(D \times k)$ and $\mathbf{H}(k \times T)$ such that $\mathbf{V} \approx \mathbf{WH}$. The non-negativity constraints make the results more interpretable in the context of term frequencies. In this case:

- $\mathbf{V}(D \times T)$ is the input matrix of TF-IDF scores where V_{dt} is the TF-IDF score of term t in document d.
- $\mathbf{W}(D \times k)$ is the basis (or document-topic) matrix where W_{di} is the prevalence of topic i in document d.
- $\mathbf{H}(k \times T)$ is the coefficient (or topic-term) matrix where H_{it} is the contribution of term t to topic i.

The approximation of $\mathbf{V} \approx \mathbf{WH}$ is achieved through the following objective function (O'callaghan et al. 2015):

$$\min_{W,H} \sum_{d=1}^{D} \sum_{t=1}^{T} (V_{dt} - (WH)_{dt})^2 = \min_{W,H} \|V - WH\|_F^2 \quad \text{subject to:} \quad W \geq 0, \; H \geq 0 \tag{24}$$

where $\| \cdot \|_F$ denotes the Frobenius norm.

Selecting the optimal number of topics k is crucial for obtaining meaning and interpretable results. The coherence score is a measure used to help approximate a good value for k. It measures how coherent terms are within a single topic by assessing the degree of semantic similarity among high-scoring terms of that topic. Coherence is based on the idea that a topic consists of terms that frequently co-occur within documents. There are various methods to calculate coherence. In this case, the C_v coherence metric is used as introduced by Röder, Both, and Hinneburg (2015) due to its proven overperformance compared to other metrics, when it comes to correlation with human ratings. The method is detailed in Syed and Spruit (2017, Section II-B, pp. 167-168). It is worth noting that though coherence is a good indicator to pick k, human intervention is often required to assess the relevance of topic modelling results (O'callaghan et al. 2015). It is sometimes worth losing a little on the coherence score for a lower value of k and less redundancy across topics. For this reason, the number of topics may be adjusted based on human judgment.

To understand the distribution of topic importance for each celebrity separately, the mean W_{di} coefficient is calculated across the 200 articles corresponding to each celebrity c. This results in a topic distribution matrix $\mathbf{M}(C \times k)$ for each celebrity where C is the number of celebrities and k is the number of topics. Each element of $\mathbf{M}$ is calculated as follows:

$$M_{ci} = \frac{1}{200} \sum_{d=1}^{200} W_{di} \quad \text{where } d \text{ is an article about celebrity } c \tag{25}$$

To understand the distribution of topic importance before or after death for each celebrity separately, two separate $C \times k$ matrices are computed for pre- and post-death articles separately, $\mathbf{M}^{(pre)}$ and $\mathbf{M}^{(post)}$. Each element of either matrix is calculated using the above formula applied to the 100 pre- and the 100 post-death articles separately as follows:

$$M_{ci}^{(pre)} = \frac{1}{100} \sum_{d=1}^{100} W_{di} \quad \text{where } d \text{ is a pre-death article about celebrity } c \tag{26}$$

$$M_{ci}^{(post)} = \frac{1}{100} \sum_{d=1}^{100} W_{di} \quad \text{where } d \text{ is a post-death article about celebrity } c \tag{27}$$

A $C \times k$ difference in topic distribution matrix $\mathbf{M}^{(diff)}$ for each celebrity is then computed by taking the element-wise difference between $\mathbf{M}^{(post)}$ and $\mathbf{M}^{(pre)}$. Each element is calculated as follows:

$$M_{ci}^{(diff)} = M_{ci}^{(post)} - M_{ci}^{(post)} \tag{28}$$

- If $M_{ci}^{(diff)} > 0$, then topic i is more prevalent after death for celebrity c.
- If $M_{ci}^{(diff)} < 0$, then topic i is more prevalent before death for celebrity c.
- If $M_{ci}^{(diff)} = 0$, then topic i is as prevalent before and after death for celebrity c.

4 Data

4.1 Celebrities

An initial dataset of 80 celebrities is manually compiled using the methods outlined in Section 3.1. Filtering out the celebrities for whom there were not enough articles (at least 100 both before and after death) resulted in a final sample of 38 stars. Features are summarised in Figure 3.

Nationality (English-speaking country) The nationality feature is heavily biased towards Americans. Indeed, there are 43 Americans, 9 French, 6 British, 2 Canadian-Americans, 2 British-Americans, 2 Indians, 2 Russians, and numerous other underrepresented nationalities in the original dataset. Similarly, the filtered dataset comprises 21 Americans, 6 British, 2 Canadian-Americans, and 9 other nationalities. To partly address this bias and simplify the analysis later on, the nationality feature was split into two groups: English-speaking and non-English-speaking countries. The two categories contained 61/80 (76%) and 30/38 (79%) celebrities respectively. The increased share between the original and filtered datasets can be explained by the fact that only English articles are fetched. Likely, celebrities outside the English-speaking world do not benefit from as much exposure in English media.

Gender Both datasets have approximately the same gender distribution. Male individuals are greatly overrepresented (82%). The strong gender discrimination, already in the original dataset (76%), comes as a direct consequence of poorer media coverage of female celebrities. This is consistent with previous research on obituaries that were heavily biased towards male individuals. Rusu (2020)'s study contains only 13% of female obituaries. This study outlines other papers with similar findings. R. Kastenbaum, Peyton, and B. Kastenbaum (1977) found that men are attributed 4 times more obituaries than women. 20 years later, the trend had worsened. Men were found to be 6 times more likely to receive an obituary after their death (Moremen and Cradduck 1999).

Industry The studied individuals come from six industries. Cinema includes movie directors, actors, and producers. Music encompasses singers and musicians in general. Public affairs comprises politicians and other state officials such as Queen Elizabeth II and Pope Benedict XVI. Sport contains sportspeople from various disciplines (e.g., Basketball, Sailing, Boxing). The crime category contains 6 American death row inmates who were executed in the period of interest, and Ayman al-Zawahiri, an important figure of the Islamist organisation al-Qaeda, who was killed in a drone strike led by the U.S. Central Intelligence Agency. While analysing the public response to serious criminals would have been valuable, the minimum count of articles was not met for any of the convicted. Although the selected criminals showed some pre-death Google Trends interest (see Section 3.1.4), the media coverage before their passing is insufficient. Criminals are therefore all dropped for the final analysis. Finally, the academia sector initially comprised scientists and Nobel Prize nominees. The filtered dataset filtered out all nominees to keep only one figure, the famous physicist Stephen Hawking. Also, it should be noted that only one label was assigned to each celebrity. For example, Marcia Strassman, known as an actress and a singer, was labelled with "cinema" since that is what she is most renowned for.

Cause of death The proportions of cause of death remain similar before and after filtering. As expected, a majority of deaths are due to illness. The cardiac category comprises celebrities who died of cardiac arrest, heart failure, septic shock, heart attack, or aortic dissection. Accidents are of various types, such as car or helicopter crashes. Sentenced individuals are the 7 criminals outlined in the previous paragraph. The remaining labels are self-explanatory. It is worth noting that annotating the cause of death is sometimes challenging. For instance, James Bond actor Sean Connery died of respiratory failure due to pneumonia, atrial fibrillation and old age. In this case, his death was labelled as heart failure, under the cardiac category but another annotator could have decided differently. American Singer Natalie Cole makes for another interesting case. She was diagnosed with hepatitis C, had kidney failure not long after and died of congestive heart failure as a complication of idiopathic pulmonary arterial hypertension, potentially due to hepatitis. In the dataset, her cause of death was manually labelled as "illness" due to the longstanding disease but one could argue that heart failure would be more appropriate. Similarly, American comedian Joan Rivers experienced serious complications during a minor throat intervention procedure. She died a few days later from brain damage caused by a lack of oxygen never having awakened from a medically-induced coma. After investigations, it was found that the clinic made numerous mistakes before and during the intervention. While her death is due to medical reasons, it was labelled as "accident" considering the context.

Age at death Age at the time of death is grouped into decades for visualisation purposes. Groups will be kept as such to perform the statistical tests described in Section 3.4. As one would expect, the majority of deaths are concentrated around 50+ year-old individuals. While there is only one death in the 30-49 range after filtering, 5 early and unexpected deaths remain. Those are the collision of American football player Dwayne Haskins with a truck, the suicide of Swedish DJ Avicii, the overdoses of American rappers Mac Miller and Juice Wrld, and the shooting of American rapper XXXTentacion.

Unexpected death After filtering, half of the remaining celebrities passed on unexpectedly. This is reflected by the 5 cases outlined in the above paragraph. Other unexpected deaths include Kobe Bryant's fatal helicopter accident, American musician Chris Cornell's suicide, American actor Matthew Perry's ketamine overdose, American rally driver Ken Block's snowmobile accident, American singer Prince's fentanyl overdose, or Joan Rivers' controversial death as outlined previously. Similarly to the cause of death, some subjective assessment is required to understand whether death was sudden and unexpected. Though lethal accidents and overdoses are evidently unexpected, other kinds of "slower" deaths such as illness can still come as a surprise to the public. For instance, though Bowie suffered from liver cancer, he had not made his condition public and his death at 69 felt sudden to his fans. His death is therefore labelled as "unexpected" in the dataset. In that regard, a death is considered sudden/unexpected depending on whether or not the condition that caused the death was previously known to the public.

Controversy After filtering, the controversy variable was also evenly balanced with half the figures having faced controversy in their lifetime. The dummy variable measuring controversy is also subject to subjective evaluation since there are varying degrees of controversy. Celebrities are considered controversial whenever they faced criticism for their action or acted in a way that generated a polarised debate in the media at least once in their lifetime. There are numerous reasons for controversy, making some figures highly contentious. Those include Pope Benedict XVI for his handling of the church's sex abuse scandal (Povoledo 2022), British rock musician Lemmy for owning and wearing Nazi uniforms (Michaels 2008), Argentinian football player Diego Maradona for his drug abuse (Bernardi 2022), Italian politician Silvio Berlusconi for his series of sex scandals (Yuhas 2023), American rapper XXXTentacion for his various criminal charges including stabbing events, domestic violence towards a pregnant woman, and false imprisonment (Coscarelli 2018).

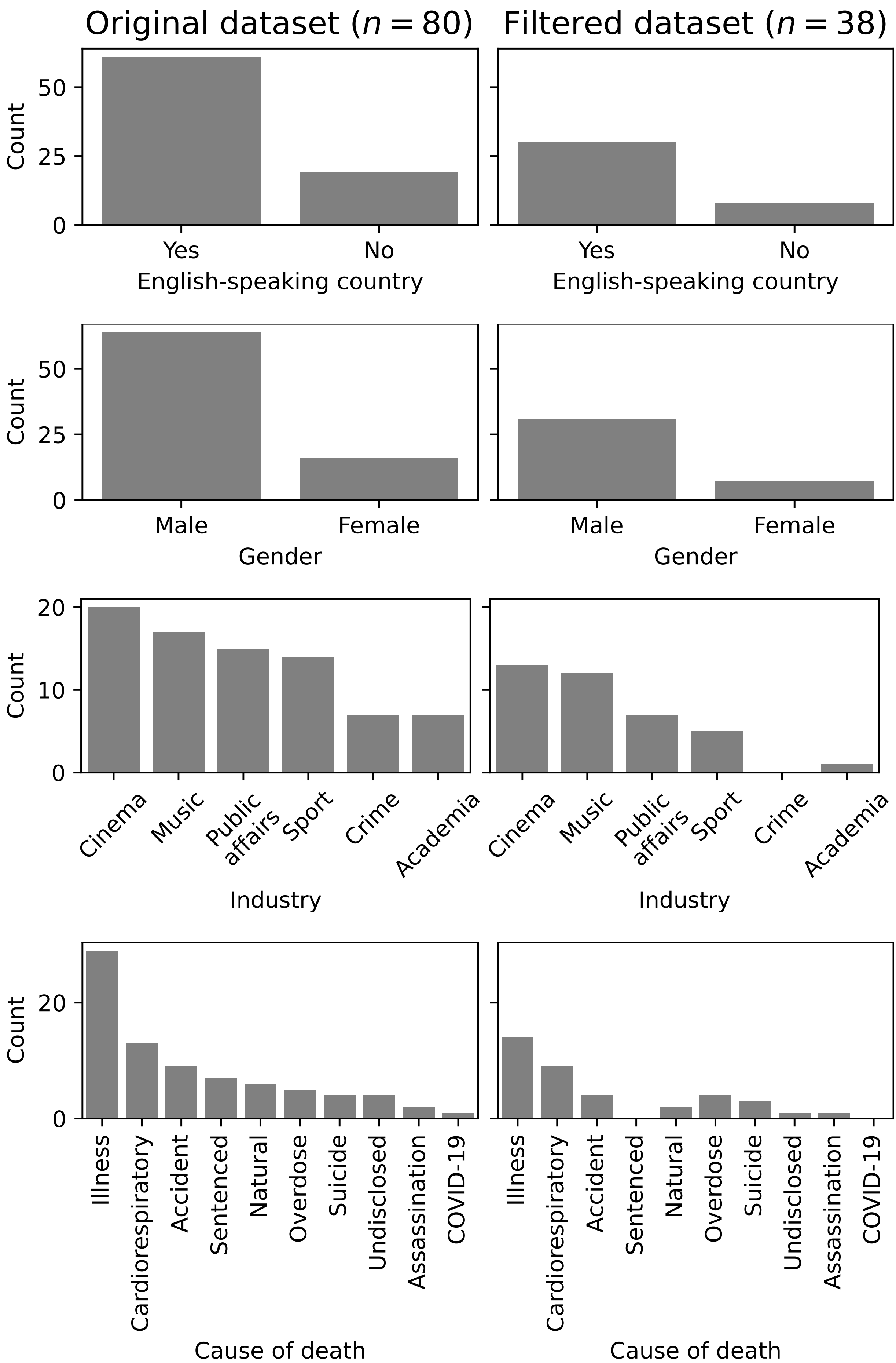

Figure 3: Summary of features before and after filtering

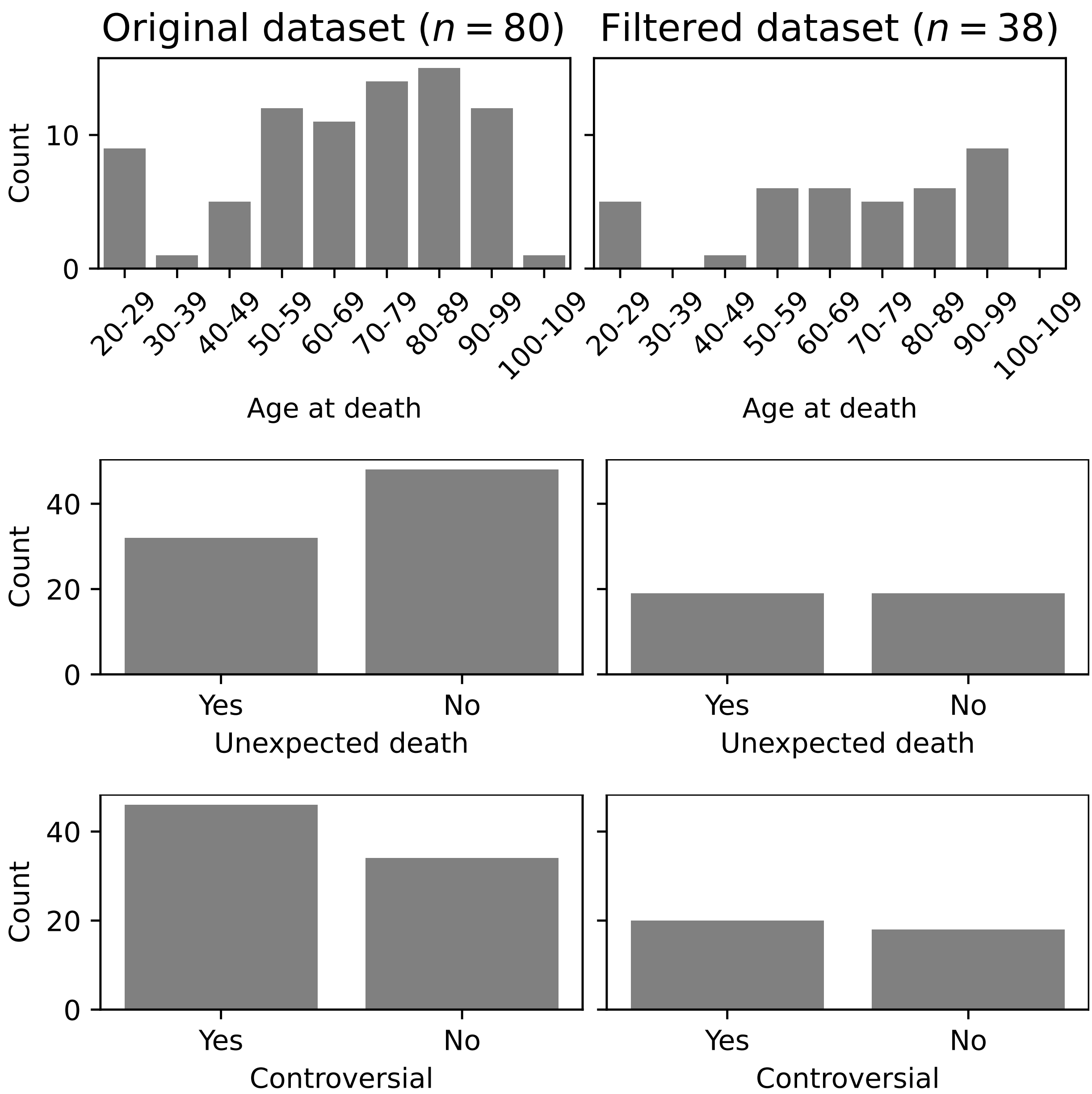

Figure 3: Summary of features before and after filtering (cont.)

The filtered dataset contains celebrities for whom at least 100 articles were fetched both before and after death.

4.2 Article sentiment

A total of 32,668 articles was fetched using the API. Duplicate articles and those with a missing sentiment score were dropped and only the top 100 pre- and post-death articles were kept for each celebrity (Section 3.1.4). 7600 articles (= 100 articles × 2 periods × 38 celebrities) remained after filtering. This section gives a preliminary description of pre- and post-death sentiment distributions of the remaining articles.

4.2.1 All articles

Figure 4 shows the pre- and post-death sentiment distributions across all articles. Table 2 contains descriptive statistics of the two distributions. The distributions have almost the same mode. While the post-death curve has a slight negative skew, the pre-death one is bimodal with the second highest peak in the negative sentiment range. This second peak drives the mean and median lower than their post-death counterparts. These early observations suggest that there might be some evidence of the death positivity bias. In other words, the media would speak more positively about celebrities once they have passed.

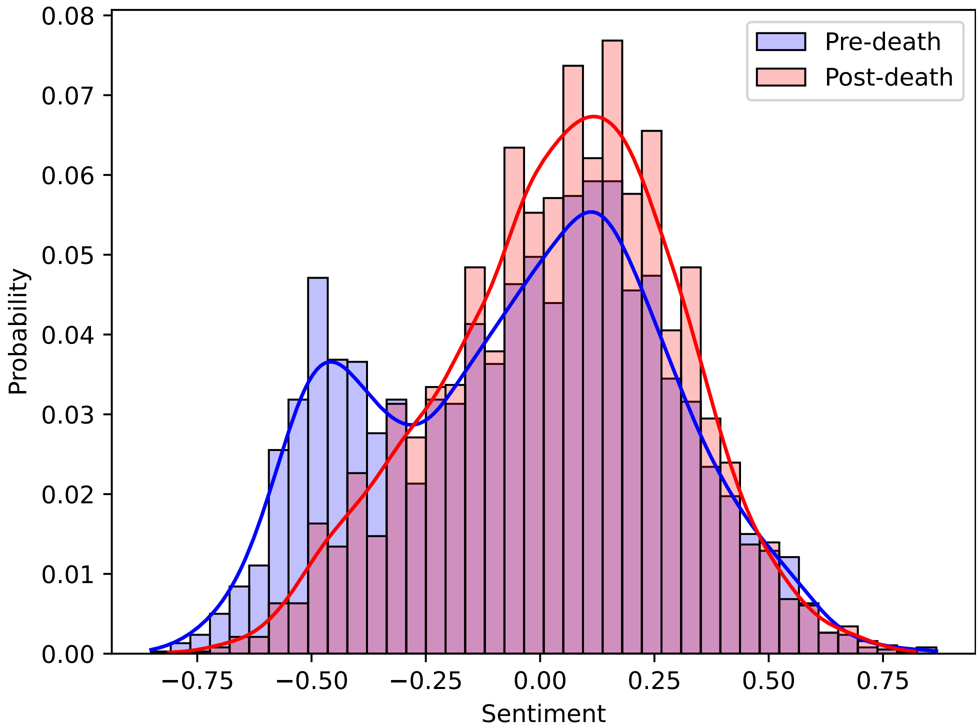

Figure 4: Probability distributions of pre-death and post-death sentiment for all celebrities

Table 2: Sentiment descriptive statistics

Period	**Mean**	**Median**	**Mode**	**Std**	**Var**	**Skew**	**N**
Pre-death	-0.0471	-0.0118	0.1216	0.3167	0.1003	-0.0955	3800
Post-death	0.0435	0.0667	0.098	0.257	0.066	-0.2393	3800

4.2.2 Groups of articles

Figure 5 shows the sentiment distributions of pre-death and post-death articles for each feature group as box plots. The following paragraphs describe the data. A more thorough analysis of sentiment scores is covered in Section 5.2 following statistical tests.

Nationality (English-speaking country) There seems to be some evidence of death positivity bias in countries of the English-speaking world, whereas other countries do not exhibit the effect.

Gender Although in the minority, articles about female celebrities have a more widespread range of pre-death sentiment values. Both groups portray similar post-death distributions.

Industry The academia and cinema industries show some increase in sentiment post-death. No signal is detected for the remaining industries.

Cause of death Assassination, illness and suicide portray a death positivity bias. It is worth noting that the undisclosed group shows an overall negative postmortem response.

Age at death Younger age groups, 20-29 and 40-49 display a more negative sentiment post-death in comparison with pre-death. The remaining age groups exhibit a positive signal for the death positivity bias.

Unexpected death Whether death is sudden or not does not seem to have a strong impact on the change in sentiment following celebrities' passing. Both groups exhibit a death positivity bias.

Controversy The death of celebrities who faced controversy during their lifetime does not seem to impact the media sentiment towards them. In contrast, non-controversial individuals show a rather strong positive shift in sentiment following their death.

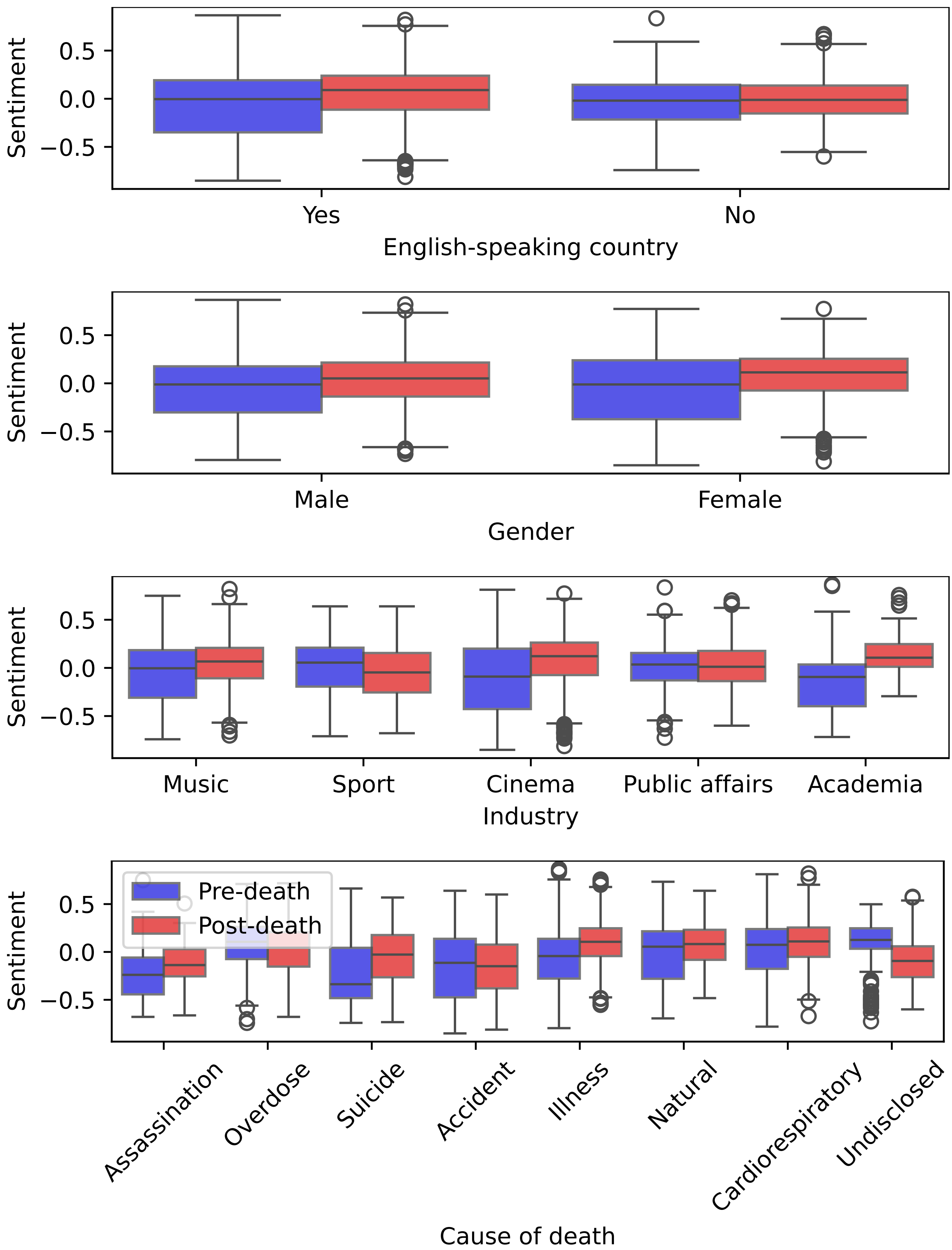

Figure 5: Box plots of sentiment distributions pre-death and post-death for each feature

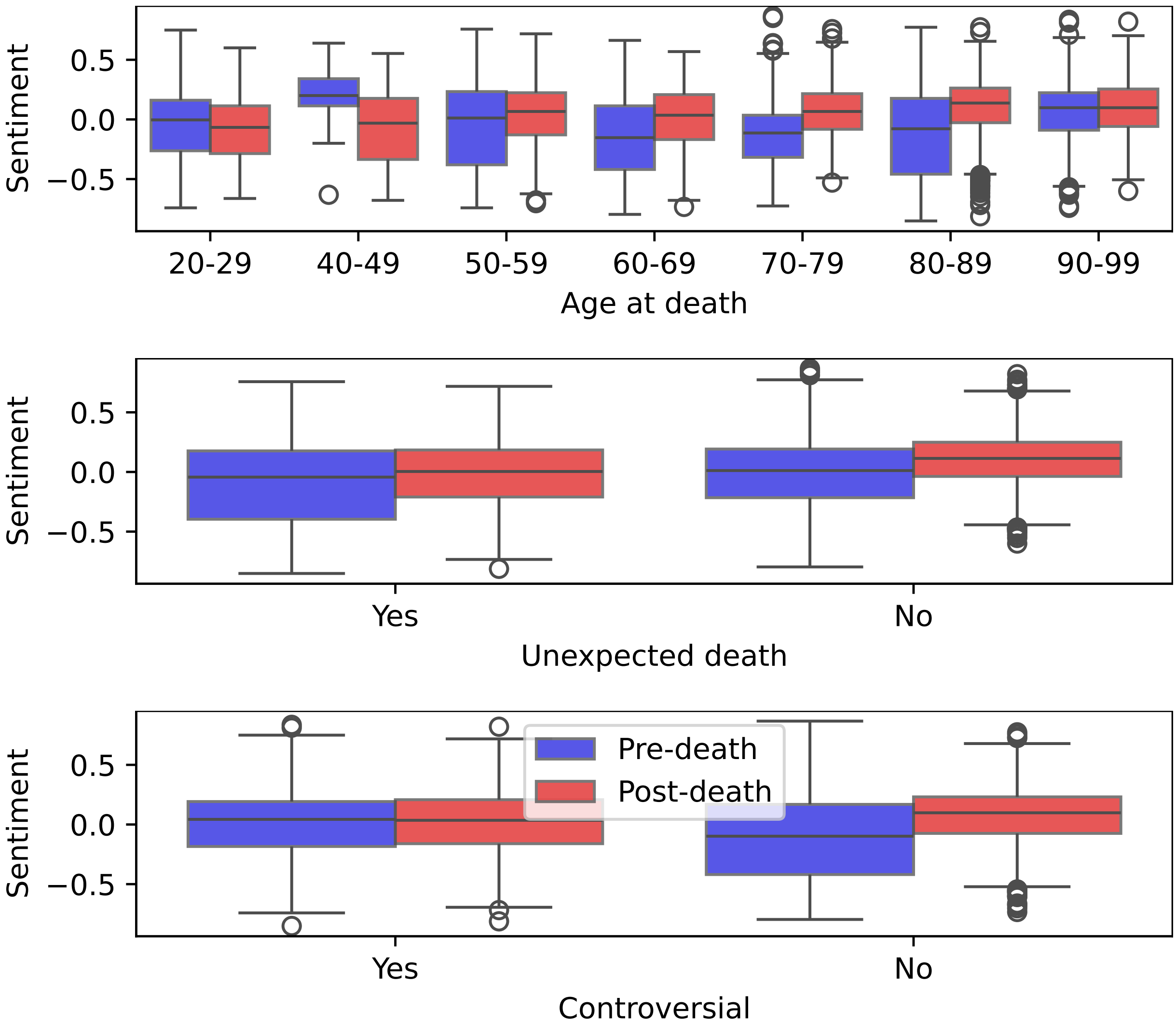

Figure 5: Box plots of sentiment distributions pre-death and post-death for each feature (cont.)

4.3 Article sources

This section ensures that there are no biases in the dataset. Indeed, varying article sources between the pre-death and post-death groups might impact the analysis. As mentioned in the literature review, Rusu (2020) found differences in the evaluation of deceased leaders depending on the type of news source (See Section 2.4). Sentiment tends to be more negative in tabloid newspapers than in more formal news agencies. Since Event Registry's API collects articles from all kinds of sources, it is necessary to check whether article sources are consistent between the pre-death and post-death samples. The API provides information on the sources of the news articles. Those include the url, the location of the news source and information on its rank. More information is provided in the below paragraphs.

Source URL Articles come from 1904 different sources. The distribution of pre- and post-death article counts for the top 20 sources are plotted in Figure 6.

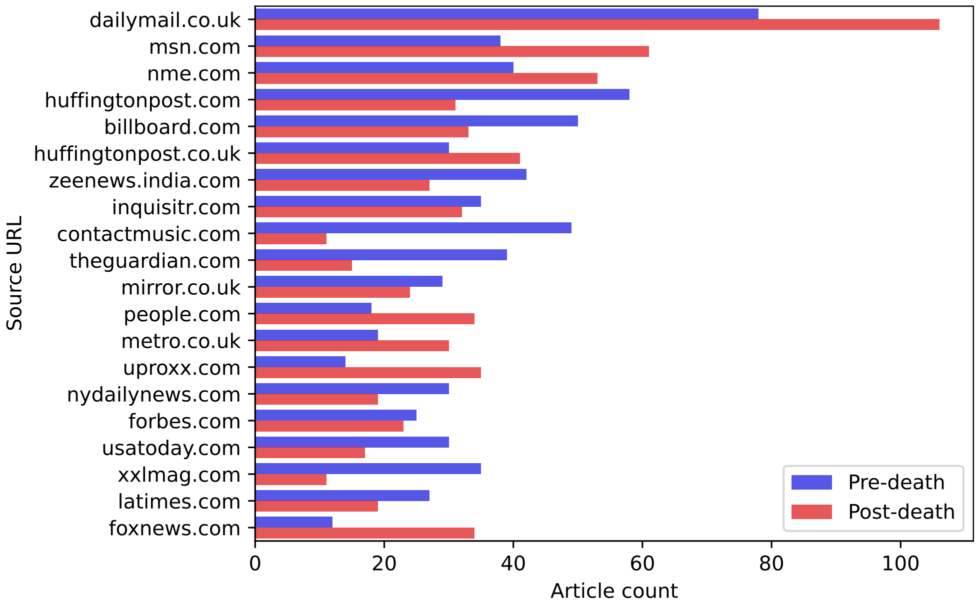

Figure 6: Counts of pre-death and post-death articles for top 20 sources

The British newspaper Daily Mail Online is the most represented source with 138 articles (2.42%). Other important sources include Microsoft news provider MSN (Microsoft Network, 1.30%), British music website NME (New Musical Express, 1.22%), American progressive news website HuffPost (1.17%), American music and entertainment magazine Billboard (1.09%), and British version of the HuffPost (0.93%). The remaining sources represent less than 1% of all sources. There is some variation in source counts across periods. The Spearman correlation coefficient indicates no relationship, $\rho(1791) = -.01, p = .63$. Furthermore, 49.86% of the sources have only one article and therefore appear in only one of the two periods. While there is a large pool of article sources overall, it is difficult to assess whether the writing tone is similarly distributed in the two periods.

Source location The location returned by the API comes at varying levels of granularity. Sources are located from the country level (e.g., "Canada"), through region level (e.g., "Wellington, New Zealand"), to the city level (e.g, "Charlotte, North Carolina, United States"), or even neighbourhood level (e.g., "Playa Vista, Los Angeles, California, United States"). 360 articles have an unknown location. To simplify comparison, locations were mapped to their respective country levels. Articles come from 86 different countries in total. Figure 7 shows the distribution of pre- and post-death article counts for the top 20 countries.

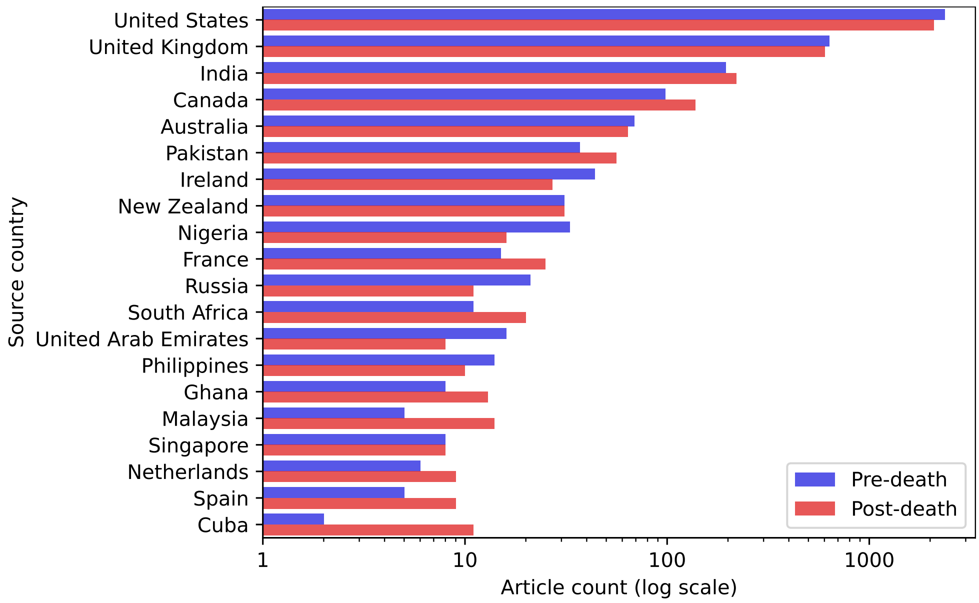

Figure 7: Counts of pre-death and post-death articles for top 20 source countries

The article count on the x-axis is measured on a $\log_{10}$ scale

Most articles were written by news sources located in the United States (61%). Other important sources are the United Kingdom (17%), India (6%), Canada (3%), Australia (2%), and Pakistan (1%). The remaining countries represent less than 1% of all source countries. Figure 7 shows that the top 9 countries have English as an official language. Those countries represent 93% of all articles, which reflects the fact that only English articles were fetched. The distribution of source country counts is balanced across the two periods with a strong positive Spearman correlation, $\rho(84) = .58, p < .001$. Both groups of articles are well-balanced in terms of source location. The results of the subsequent analysis are not due to a bias in the country selection.

Source rank Each source is assigned an Alexa global rank score that is returned by the API. Alexa rank is a measure of a website's popularity relative to other sites. Websites across the internet are ranked using a combination of average daily visitors and pageviews over the previous 3 months. A lower Alexa rank indicates a more popular website. The Amazon-owned ranking system was shut down in May 2022 but Event Registry's API still provides previous rankings for news sources (Kehoe 2022).

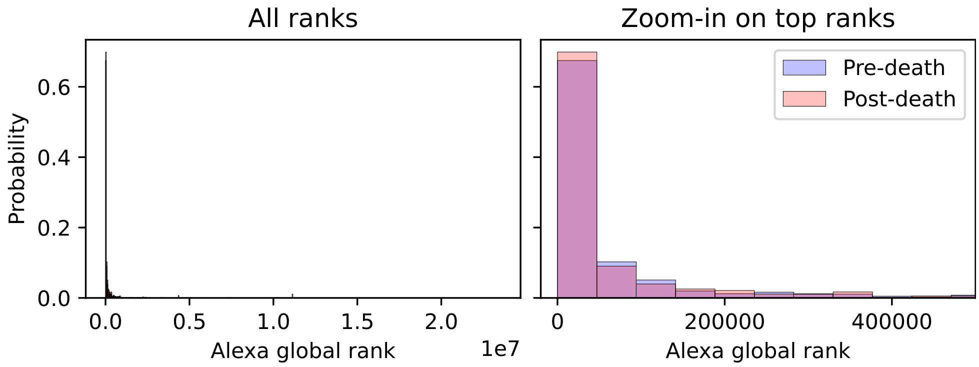

Figure 8: Probability distributions of pre-death and post-death Alexa global rank

The left figure shows the distribution of the ranks of all article sources. The right figure is a zoom-in on the top 2000 ranks.

Figure 8 suggests the distributions of pre-death and post-death Alexa global ranks are similar. A two-sided Mann-Whiney U-test finds evidence to reject the null hypothesis of no difference at the 10% level, but not at the 5% level of significance, $U = 6929383.5, p = .0593$. This detected difference is mostly due to the large sample size as the rank biserial correlation used as effect size remains negligible, $r = -.0403$. This shows that both groups of sources have similar distributions of website popularity.

5 Results

5.1 Evidence of the death positivity bias (RQ1)

Mann-Whitney U-tests and effect sizes were computed on all articles at once and on articles about each star individually. Results are found in Table 8 (Appendix A.3), and summarised in Figure 9. The bootstrap distribution of the rank biserial correlation coefficient r across all articles can be found in Figure 16 (Appendix A.2). Overall, there is evidence to reject the null hypothesis of no difference in sentiment between the two periods, $U_1 = 6044571, p < .01$. The rank order of article sentiment differs depending on whether the article was published pre- or postmortem. In other words, articles published after death tend to be more positive than those before. However, the effect size (rank biserial correlation) was found to be $r = -0.1628$, indicating a small effect (J. Cohen 2016, p. 157). To provide a robust estimate of this effect size, a 95% bootstrap confidence interval was computed using $B = 1000$ resamples. The bootstrap confidence interval ranges from -0.1880 to -0.1377, suggesting that the true effect size falls within this range with 95% confidence.

When diving into individual tests, Figure 9 provides evidence to reject the null hypothesis at the 10% level for 21 celebrities out of the 38. Out of those, 7 celebrities exhibit a large effect size ($|r| > .5$), 10 a medium effect ($.3 < |r| < .5$), and 4 a small effect ($.1 < |r| < .3$). For completeness, the two-tailed version of the Mann-Whitney U-test was also run. Detailed results are found in Figure 17 and Table 9 (Appendix A.4). The two-sided version of the test shows that although the null hypothesis of no difference is rejected for 31 celebrities, 11 of them exhibit a reversed effect ($r > 0$). In other words, 11 celebrities are spoken more negatively about once they have passed. These opposing results help explain the small effect size detected overall.

In summary, there is evidence of a death positivity bias. In other words, the media tends to speak more positively about celebrities once they have passed away. It should be noted that, although the signal is strong for some celebrities, some stars exhibit no effect or even a reverse relationship where death brings more negative media attention.

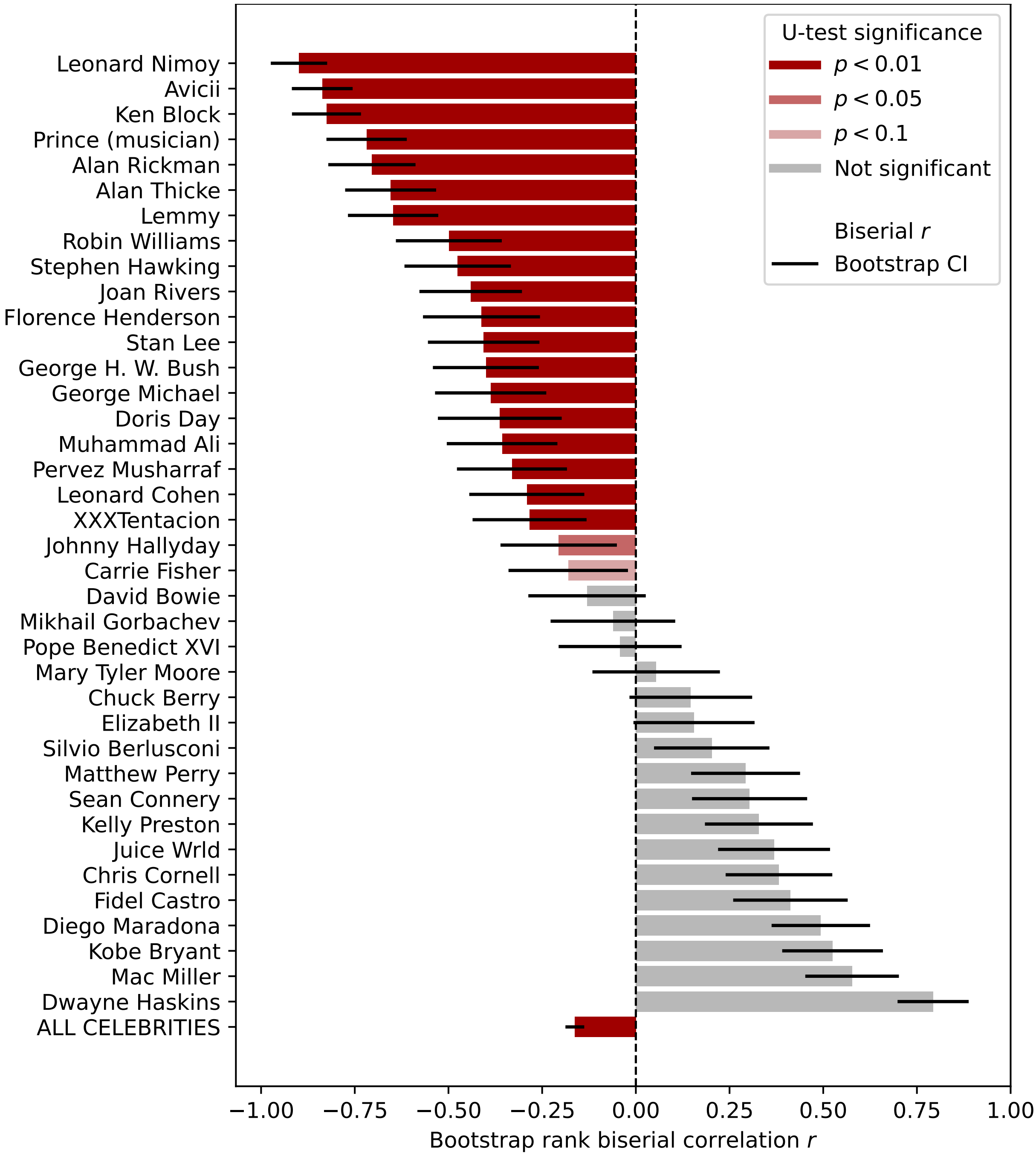

Figure 9: One-sided Mann-Whitney U-tests and effect sizes

The x-axis shows the bootstrapped effect size r of the difference between post-death and pre-death sentiment for each celebrity on the y-axis. The black bars match the bootstrap normal confidence intervals of r. A negative r indicates that post-death sentiment tends to be more positive than its pre-death counterpart. The bar colours represent the level of significance of the one-sided Mann-Whitney U-test formulated in Section 3.3.1. The values used to compute this plot are found in Table 8 (Appendix A.3).

5.2 Variations in media response (RQ2)

Results of RQ1 provide evidence of a small death positivity bias (Section 5.1). While the effect is large for some celebrities, others exhibit a smaller effect and the relationship is even reversed for some individuals. The current section explores whether such varying responses can be explained by individual characteristics of the celebrities and the context of their deaths.

Articles were randomly paired and differences between post- and pre-death sentiment scores were computed as described in Section 3.4.1. The sentiment difference distribution D can be found in Figure 18 (Appendix B.1). Descriptive statistics of each group are found in Table 3.

Table 3: Descriptive statistics of each group of the sentiment difference distribution D

Feature	Group	Mean ($\bar{D}$)	Std	Var	N
english	No	0.0394	0.3694	0.1365	800
english	Yes	0.1043	0.4068	0.1655	3000
gender	Female	0.0987	0.4003	0.1602	700
gender	Male	0.0888	0.4001	0.16	3100
industry	Academia	0.2405	0.3824	0.1463	100
industry	Cinema	0.1686	0.4138	0.1713	1300
industry	Music	0.0981	0.3946	0.1557	1200
industry	Public affairs	0.0071	0.3246	0.1054	700
industry	Sport	-0.0429	0.4169	0.1738	500
cause	Accident	0.016	0.4522	0.2045	400
cause	Assassination	0.1129	0.3307	0.1093	100
cause	Cardiorespiratory	0.0752	0.3687	0.1359	900
cause	Illness	0.1581	0.3726	0.1388	1400
cause	Natural	0.0765	0.394	0.1553	200
cause	Overdose	-0.0475	0.4016	0.1613	400
cause	Suicide	0.1881	0.4616	0.2131	300
cause	Undisclosed	-0.1497	0.3907	0.1527	100
age	20-29	-0.0273	0.4261	0.1815	500
age	40-49	-0.2824	0.3448	0.1189	100
age	50-59	0.0888	0.4291	0.1842	600
age	60-69	0.1428	0.3897	0.1519	600
age	70-79	0.1884	0.341	0.1163	500
age	80-89	0.2117	0.41	0.1681	600
age	90-99	0.029	0.3437	0.1181	900
unexpected	No	0.1202	0.3754	0.1409	1900
unexpected	Yes	0.061	0.4213	0.1775	1900
controversial	No	0.1663	0.4192	0.1757	1800
controversial	Yes	0.0225	0.3691	0.1362	2000

Dividing the number of articles N by 100 gives the number of celebrities in each group

An ANOVA test was run, complemented by post hoc Tukey's HSD tests for significant features with more than 2 groups. The results are found in Table 4 and Figure 10.

Nationality (English-speaking country) There is no evidence of a difference in response between celebrities whose countries have English as an official language and those who do not, $F(1, 4480) = .0471, p = .8281, \eta^2 = .0000$.

Gender There is a significant difference in sentiment response between male and female celebrities, $F(1, 4480) = 40.3958, p < .01$. On average, the media sentiment increases by

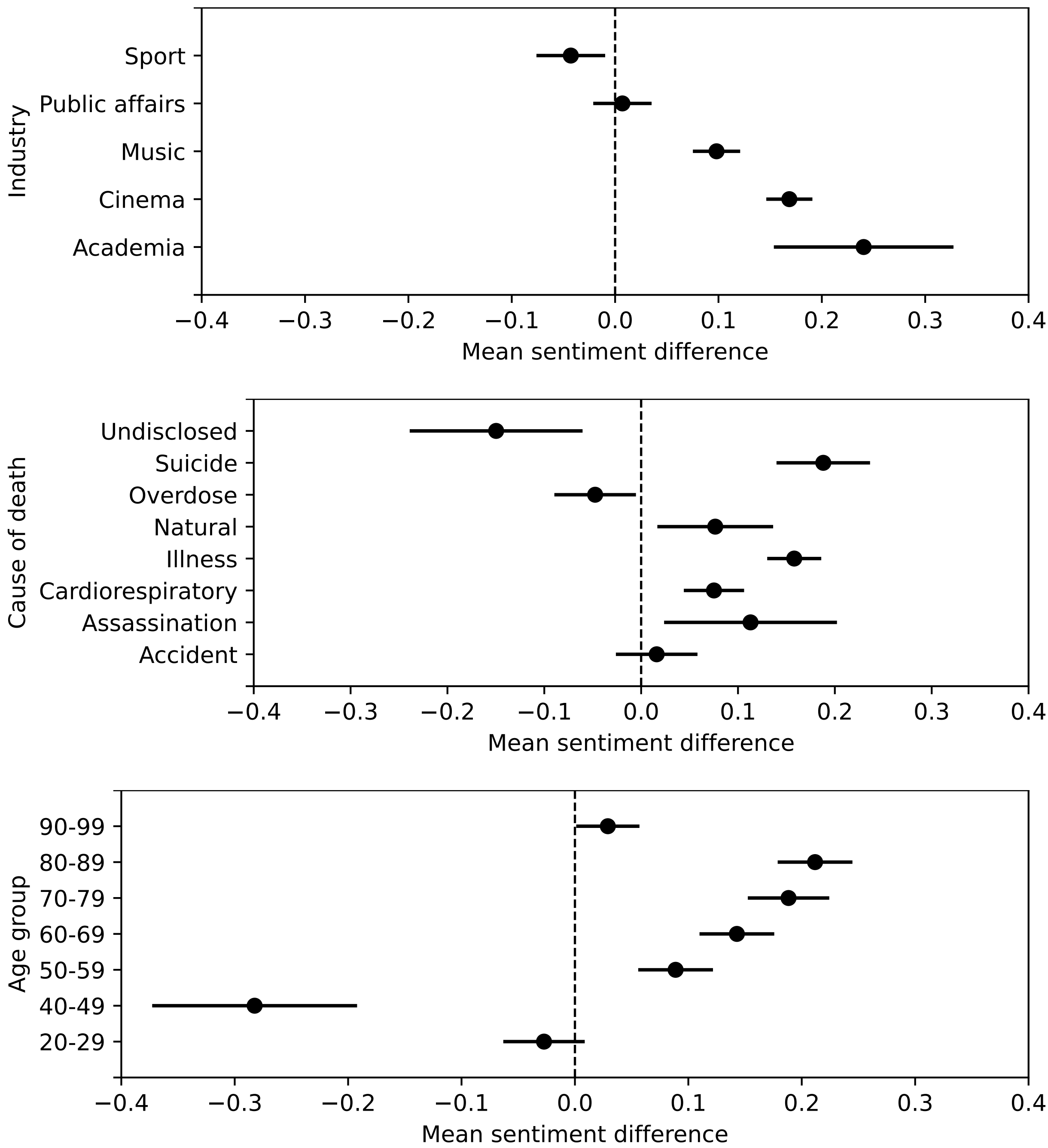

Figure 10: Tukey's HSD test results

The x-axis shows the mean sentiment difference for each group on the y-axis.
The means of each group are found in Table 3.
The horizontal black lines around the estimated means are 95% confidence intervals.
Two overlapping black lines indicate that there is not sufficient evidence to reject the null hypothesis of no difference between the two groups (i.e., $p > .05$).
Groups with means lower than 0 or whose confidence interval overlaps with 0 show no evidence of death positivity bias.
Details of the tests are found in Table 10 (Appendix B.2).

Table 4: ANOVA test results

Feature	**SS**	**df**	**F-stat**	**p-value**	$\boldsymbol{\eta^2}$
english	0.0066	1	0.0471	0.8281	0.0
gender	5.6634	1	40.3958	0.0***	0.0106
industry	8.3977	4	14.9748	0.0***	0.0156
cause	20.64	7	21.0317	0.0***	0.0375
age	22.1909	6	26.3806	0.0***	0.0402
unexpected	0.018	1	0.1281	0.7204	0.0
controversial	2.1439	1	15.2919	0.0001***	0.004
Residual	529.6648	3778	-	-	-

*Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$*

0.0987 for females and 0.0888 for males (Table 3). However, the detected difference is mostly due to the large sample sizes in both groups (700 and 3700) as the effect size remains small, $\eta^2 = 0.0106$ (J. Cohen 2013). The magnitude of the difference is small, as a sentiment increase of 0.0987 compared to 0.0888 is arguably the same in practice.

Industry The ANOVA test results provide evidence of a difference between different industries with a small effect size $F(4, 4480) = 14.9748, p < .01, \eta^2 = .0156$. Since the results are significant and there are more than two groups, Tukey's HSD tests were also run to understand the response for each group (Figure 10). Music, Cinema, and Academia have a death positivity bias signal. Academia and cinema exhibit the strongest signal (0.2405 and 0.1686 respectively). The confidence interval is large for academia since it is represented by only one individual, Stephen Hawking. In contrast, celebrities from public affairs and sport experience no change in media sentiment at their death.

Cause of death Similarly to industry, there is evidence of differences in means between different causes of death with a small effect size, $F(7, 4480) = 21.0317, p < .01, \eta^2 = .0375$. Tukey's HSD tests provide evidence of a death positivity bias for celebrities who died of suicide, natural causes, illness, cardiorespiratory failure, or assassination. Suicide and death by illness have the highest increases in media sentiment at death (0.1881 and 0.1581 respectively). The undisclosed group shows the opposite response, media sentiment becomes more negative after death. A possible explanation would be speculation on the actual cause of death.

Age at death Similarly to industry and cause of death, the ANOVA test shows evidence of mean differences among age groups with a small effect size, $F(6, 4480) = 26.3806, p < .01, \eta^2 = 0.0402$. According to Tukey's HSD test results, celebrities who passed away between 50 and 89 years old are subject to the death positivity bias. Within that age range, the greater the age, the greater the bias for a maximum of 0.2117 for the 80-89 group. Interestingly, there is no evidence of a bias for higher and lower age groups. No conclusion can be drawn for the 40-49 age group which contains only one individual (basketball player Kobe Bryant), and for the 30-39 group which is empty.

Unexpected death There is no evidence of a difference in response between celebrities who died of an unexpected, sudden death and those who did not, $F(1, 4480) = .1281, p = .1281, \eta^2 = .0000$.

Controversy Although the ANOVA model provides evidence to reject the null hypothesis of no difference between controversial and non-controversial individuals, $F(1, 4480) = 15.2919, p < .01$, the detected difference is mostly due to the large sample size since effect size is negligible, $\eta^2 < 0.01$. In other words, the independent variable explains only a tiny proportion of the variance in media sentiment. The practical impact of the difference is too small to justify an actual difference in response between controversial and non-controversial celebrities.

In summary, some individual characteristics of celebrities and the context of their deaths help explain variations in media response after death. In particular, the death positivity bias is the strongest for celebrities of the entertainment industry (music and cinema). No evidence was found for celebrities of other industries. In terms of cause of death, celebrities who died of suicide exhibit the strongest bias, followed by illness, assassination, natural, and cardiorespiratory failure. Undisclosed deaths and those by accident or natural causes do not show any bias. Those who died between the ages of 50 and 89 exhibit a bias, that gets stronger as age increases; though no signal is found for celebrities who died at 90+ or before 50 years old. Considering the sensitivity of sentiment models, there is weak to no evidence that nationality, gender, controversy levels or whether death was unexpected, contribute to variations in media sentiment response after death.

5.3 Article content analysis (RQ3)

After showing the existence of a death positivity bias in Section 5.1 on RQ1, results of RQ2 provide evidence of variations in such response based on celebrity characteristics (Section 5.2). The current section dives into the content of articles to understand what terms and topics are covered in the media.

5.3.1 Word frequency results

Article texts were split into training (75%) and testing sets (25%). The training set was preprocessed and vectorized using TF-IDF. Unigrams and bigrams appearing in at least 1% of the training articles were kept. This resulted in 2753 n-grams. Those were used as features of a logistic regression trained to classify articles as pre- or post-death (see Section 3.5.2). Performance metrics of the logistic regression on the held-out test set of 1900 articles are detailed in Table 5. The classifier has an accuracy and an F1 score both reaching 0.88. Precision and recall range between 0.86 and 0.90 for both periods. These results are to be taken with care since there might be multicollinearity among features. Some n-grams likely appear in similar contexts.

Table 5: Logistic regression classification report on test set

	Precision	**Recall**	**F1 score**	**N**
Post-death	0.8894	0.8632	0.8761	950
Pre-death	0.8671	0.8926	0.8797	950
accuracy			0.8779	1900
macro avg	0.8782	0.8779	0.8779	1900
weighted avg	0.8782	0.8779	0.8779	1900

Output of scikit-learn's classification_report function

The classifier's high performance suggests a clear distinction in article content between the two periods. As detailed in Section 3.5.2, the top positive and negative coefficients show the relative importance of each n-gram to the pre- and post-death periods respectively. Figure 11 shows the most representative n-grams of each period. As expected, there is a clear distinction between the terms used before and after death. The pre-death group contains a mix of positive and negative n-grams on new releases in showbiz ("new", "album", "recent"), celebrity entertainment news ("couple", "marry", "wife"), health-related topics ("condition", "hospital", "health"), crime or justice-related events ("jail", "rehab"). In contrast, n-grams typical of the post-death group exhibit much more negative terms with the prominence of death-related lexicon ("death", "late", "die", "pass", "pass away", "suicide", "funeral"). Though death paints a rather gloomy picture, there is also a strong focus on remembering and paying tribute to the deceased ("tribute", "estate", "legacy", "honor", "documentary", "anniversary", "remember", "memorial", "exhibition", "posthumous"). Family-related n-grams are also common ("father", "mother", "brother", "widow"). It is worth noting that by constructions, the n-grams highlighted here are the most representative of each period. They are the terms that make each period distinctive from its counterpart. As the next section will show using NMF, some themes are common to both periods.

5.3.2 Topic modelling results

As detailed in Section 3.5.3, NMF was trained on article texts to understand what topics are covered before and after death. The number of topics k was chosen based on coherence score C_v and human intervention. The evolution of coherence based on k is plotted in Figure 12. $k = 11$ topics yielded the highest coherence score ($C_v = 0.5250$) but led to some overlapping themes. After experimenting with other high-coherence values, the final number of topics was set to $k = 7$ for a comparable coherence score ($C_v = 0.5087$).

(a) Pre-death

(b) Post-death

Figure 11: Wordcloud of the top unigrams and bigrams for each period

The size of n-grams is proportional to the absolute value of their logistic regression coefficient.

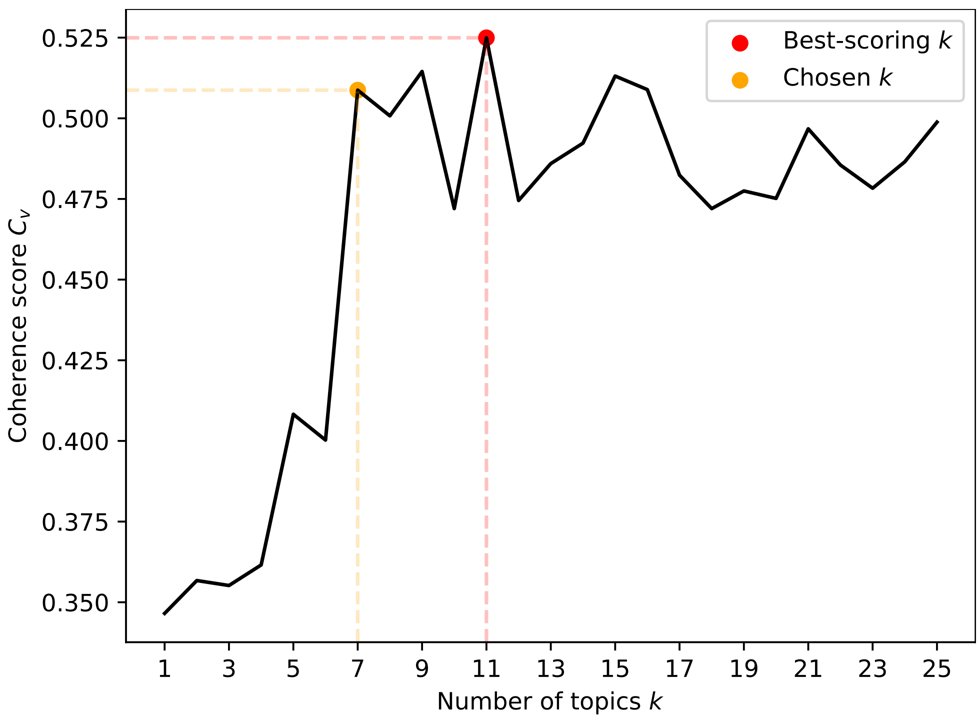

Figure 12: Plot of the coherence score based on the number of topics for NMF

The top 20 n-grams were extracted from the topic-term matrix $\mathbf{H}(k \times T)$ for each topic. The n-grams were then used to label each topic with a main theme as shown in Table 6. It is worth noting that industry-related topics are highlighted here. Apart from the academia industry which only contains one individual; cinema, public affairs, music, and sport are clearly reflected by topics 1, 2, 3, and 7 respectively. The remaining 3 topics are related to family/tribute (topic 4), justice/treason (topic 5), and crime/drugs (topic 6).

Table 6: NMF topics

Topic	Top 20 n-grams	Main theme
1	film, say, like, get, think, actor, movie, really, know, character, one, play, make, would, time, show, thing, role, want, people	cinema
2	president, say, leader, country, former, pope, people, minister, political, government, party, year, former president, prime minister, prime, would, election, war, world, visit	public affairs, leadership
3	album, song, music, record, release, new, rock, band, track, artist, lyric, tour, feature, studio, singer, new album, single, work, year, roll	music
4	family, love, mother, son, die, wife, daughter, share, death, tribute, child, life, year, post, day, say, pass, birthday, late, marry	family, tribute
5	court, treason, special court, sentence, case, verdict, high treason, special, justice, high, military, appeal, constitution, former, chief, treason case, death sentence, death, chief justice, trial	justice, treason
6	drug, lawsuit, death, claim, truck, report, allege, rapper, file, charge, accord, say, cause, estate, prescribe, attorney, kill, rehab, county, dump	crime, drugs
7	game, team, player, season, quarterback, win, football, car, play, one, video, league, coach, race, club, world, take, career, second, make	sport

Top 20 n-grams: 20 unigrams and bigrams with the highest coefficients (highest to lowest).
Main theme: Manually-labelled theme based on the top 20 n-grams.

The document-topic matrix $\mathbf{W}(D \times k)$ gives the relationship between articles and topics. Each row vector contains $k = 7$ coefficients representing the prevalence of each of the 7 topics in one article d. Each article was assigned its most prevalent topic as the one with the maximum coefficient among the 7. This allows to plot the count distributions of topics before and after death as shown in Figure 13. For a more granular analysis of the relationship between celebrities and topics, matrices $\mathbf{M}$ and $\mathbf{M}^{(diff)}$ are computed as detailed in Section 3.5.3. Their respective heatmaps are shown in Figures 14 and 15. The following paragraphs go over findings related to each of the 7 topics.

Topic 1 - Cinema As shown in Figure 14, topic 1 is covered in articles related to movie stars (e.g., Alan Rickman, Sean Connery, Leonard Nimoy, Matthew Perry, Carrie Fisher).

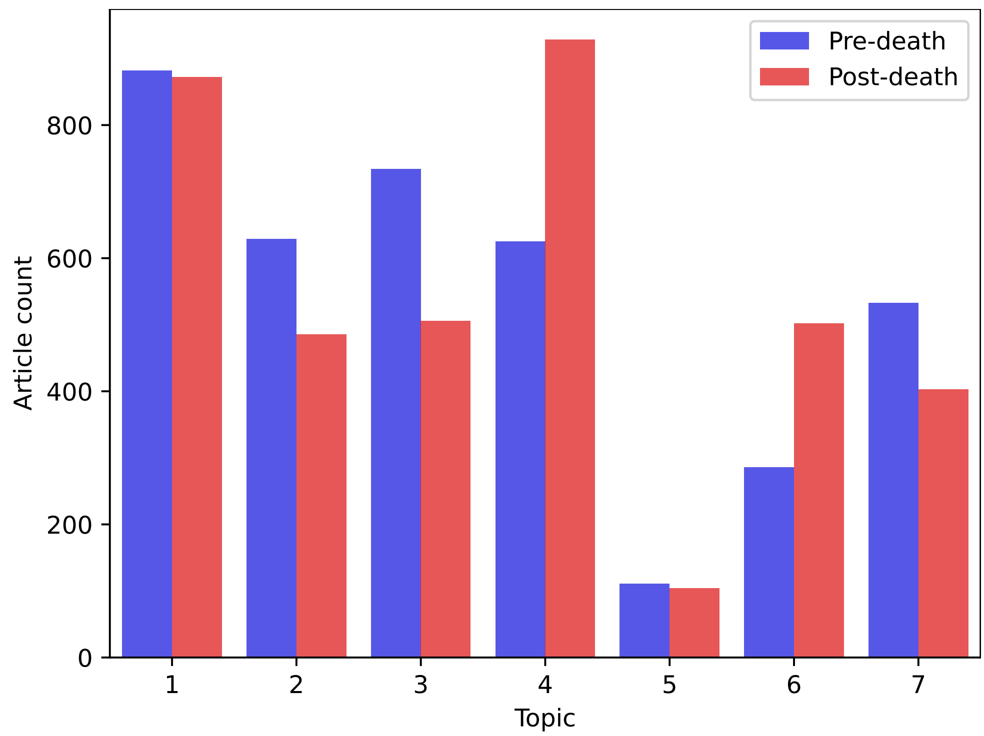

Figure 13: Pre- and post-death distributions of articles for each NMF topic

As expected, articles related to artists who had an acting career (e.g., David Bowie), or were the subject of movies (e.g., Lemmy, George Michael) also reflect topic 1. Figure 13 shows that topic 1 is covered as much before and after death which is explained by post-death articles that go over the achievements of celebrities during their lifetime.

Topic 2 - Public affairs, leadership Similarly to topic 1 with cinema figures, topic 2 is mostly represented by celebrities of its kind, public affairs. Those include Fidel Castro, Mikhail Gorbachev, Silvio Berlusconi, George H. W. Bush, and Pope Benedikt XVI among others (Figure 14). Although Muhammad Ali was characterised as a sports figure in the current work, his political activism is also reflected in media articles. Figure 15 shows that his political engagement was mostly covered in the media before he passed, in comparison with topic 7 related to sports. The fact that topic 2 is covered more extensively post-death is true of most celebrities and also reflected in Figure 13. Celebrities' political and leadership achievements seem to be of lesser interest to the media after death. The narrative would shift towards other topics for most individuals.

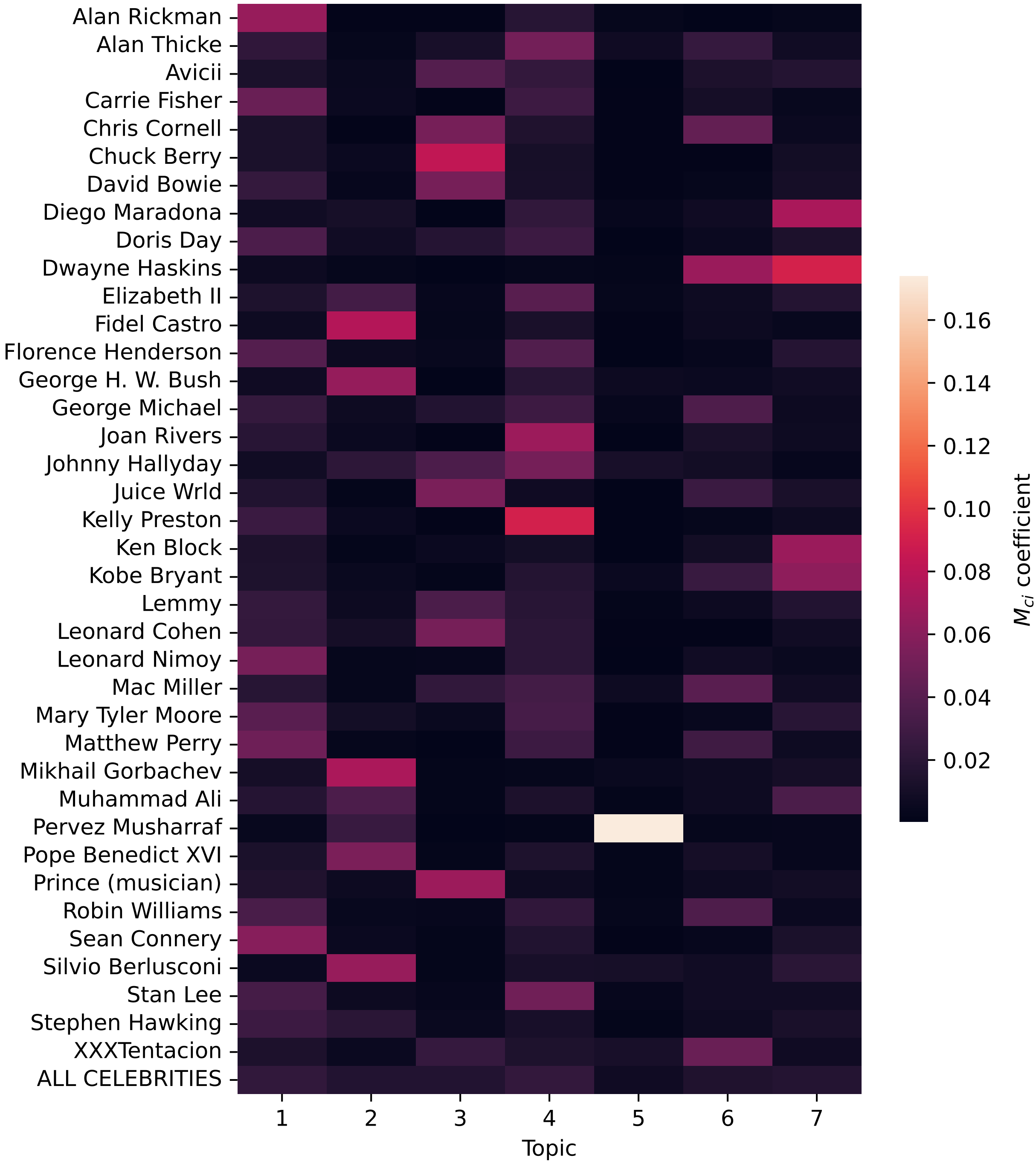

Figure 14: Heatmap of topic distribution matrix $\mathbf{M}$ for each celebrity

Matrix $\mathbf{M}$ is calculated as described in Section 3.5.3.

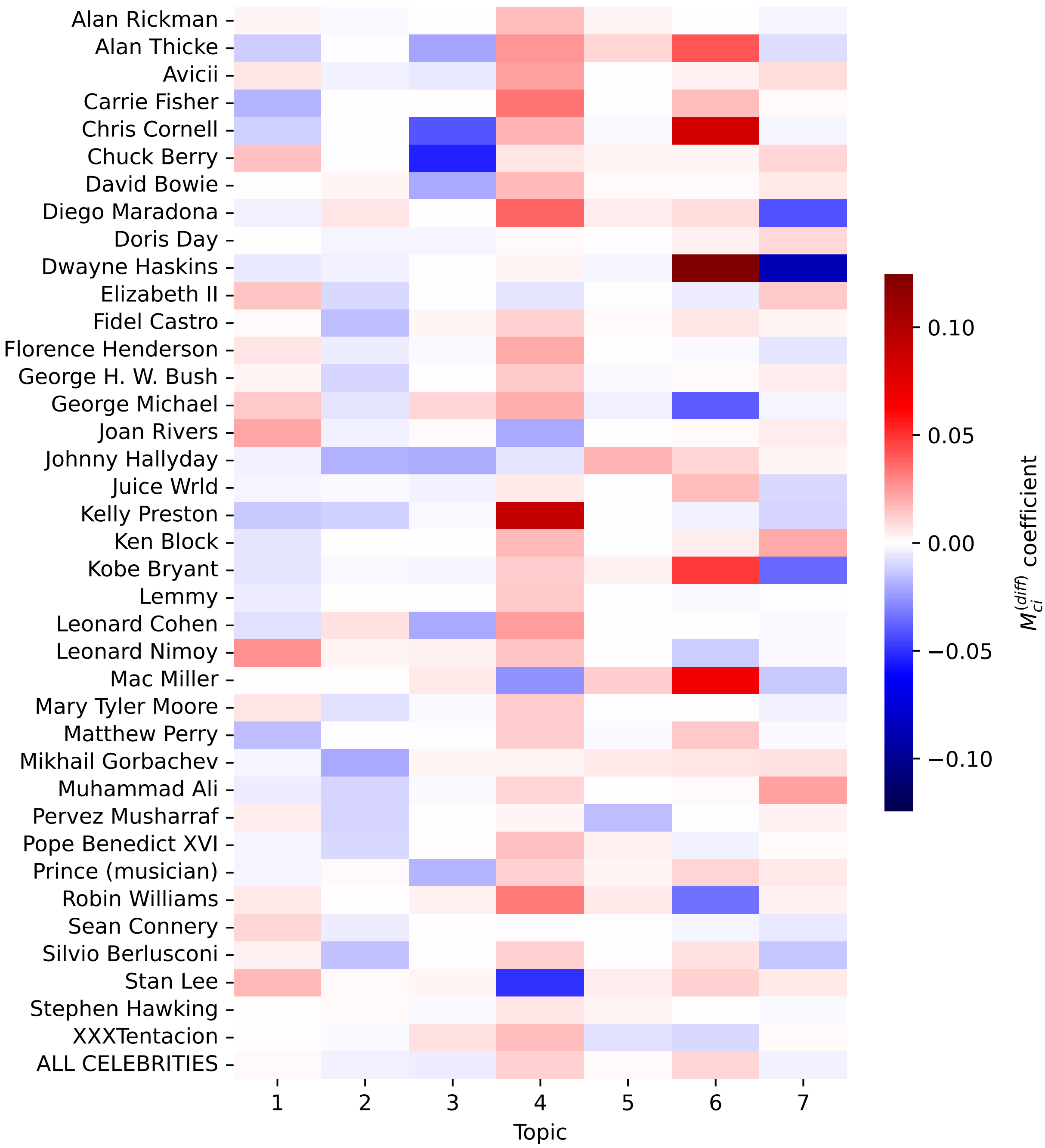

Figure 15: Heatmap of difference in topic distribution matrix $\mathbf{M}^{(diff)}$ for each celebrity

Matrix $\mathbf{M}^{(diff)}$ is calculated as described in Section 3.5.3.
Negative coefficients (blues) indicate that the topic is more prominent before death.
Positive coefficients (reds) indicate that the topic is more prominent after death.
Zero coefficients (white) indicate that the topic is as prominent before and after death.

Topic 3 - Music Again here, topic 3 covers musicians and singers including Chuck Berry, Prince, Juice Wrld, David Bowie, and Chris Cornell among others. Also known as a singer, actress Doris Day is also represented here (Figure 15). Figure 13 shows that similarly to topic 2 on cinema, topic 3 is more present before than after death. This is explained by the top n-grams on this topic which are related to new song and album releases, tours organised or studio recordings (Table 6).

Topic 4 - Family, tribute Topic 4 highlights the family and remembrance themes already detected in Figure 11b. This shows the role of the media in commemorating and paying tribute to those who passed. The focus on family is likely related to family members' actions and/or comments following the passing of their loved one. This theme is covered in articles related to almost all celebrities (Figure 14), though it mostly shows after death (Figure 13). Figure 15 shows that this holds for almost every celebrity, except for a few exceptions (e.g., Stan Lee, Mac Miller, Joan Rivers). The case of Joan Rivers could be explained by the writing of her book "From Mother to Daughter: Thoughts and Advice on Life, Love, and Marriage", where she offers life advice to her daughter and family lexicon is used.

Topic 5 - Justice, treason Topic 5 does not bring much information on the contents of the articles as it is mostly related to one individual. The former president of Pakistan, Pervez Musharraf received high media attention, especially before his death, for leading the 1999 coup d'état in Pakistan (Figures 14 and 15). The politician also faced controversy by declaring a state of emergency and suspending the constitution in 2007. The Pakistani government brought high treason charges against Musharraf who was sentenced to death for his 2007 actions, though the verdict was later annulled. It is safe to say that the 215 articles on topic 5 (Figure 13) are mostly related to Musharraf. Although those represent less than 3% of articles, the case is so unique that a dedicated topic was detected by NMF. The topic also appeared for other levels of k.

Topic 6 - Crime, drugs Topic 6 is closely related to the conditions of death of celebrities, more specifically those who died of drug overdose or were killed in an accident. The term "truck" comes from Dwayne Haskins' deadly accident where he was hit by a dump truck (Figure 14). Vocabulary of the lawsuit for a potential crime that followed is also found in the top n-grams (Table 6). XXXTentacion, Kobe Bryant and Chris Cornell are also related to such vocabulary due to the lawsuits that followed their respective assassination, suicide, and accident (Frisaro 2023; Kreps 2021; Matza 2023). Figure 14 implies that the drug-related lexicon is found in articles related to Mac Miller, Mathew Perry, and Juice Wrld who all died of a drug overdose. As shown in Figure 13, this topic is most prominent after the death of celebrities. Some exceptions in which the topic is more covered before death are found in Figure 15. The most notable examples are George Michael, Robin Williams, Leonard Nimoy and XXXTentacion who all suffered from drug addiction (Aitkenhead 2010; Engelking 2015; Hattenstone 2009; Moore 2018).

Topic 7 - Sport Similarly to topics 1, 2 and 3, topic 7 is representative of one industry (Table 6). Figure 14 shows that it is greatly related to sports figures (e.g., American football quarterback Dwayne Haskins, football player Diego Maradona, rally driver Ken Block, basketball player Kobe Bryant, boxer Muhammad Ali), but also to politician Silvio Berlusconi for his influence as Italian football club owner (Gates 2023). Figure 13 shows that topic 7 is more spoken about before death, though the evidence is weak since the trend is reversed for a number of celebrities remembered as prominent figures of their sport (e.g., Muhammad Ali, and Ken Block) (Figure 15).

In summary, 7 topics were detected (1: cinema; 2: public affairs, leadership; 3: music; 4: family, tribute; 5: justice, treason; 6: crime, drugs; 7: sport). Topics 1, 2, 3 and 7 reflect the 4 main industries covered in the current work. Academia is not reflected since it only consists of one individual. Except for the cinema industry (topic 1), these topics are more prominent before death. Although stars' achievements in their respective field are still relevant after death, much of the focus is shifted towards topics 4 and 6. The former is family-oriented and focuses on remembrance and paying tributes to the deceased. The latter identifies celebrities who were subject to drug addictions, and/or died of an overdose, or whose death was followed by a lawsuit. Topic 5 is mostly relevant to one individual and does bring additional insights. Although NMF performed well by identifying topics relevant to the celebrities being studied before and after death, they do not hep explain the reasons behind the death positivity bias.

6 Discussion

The current work aimed to identify whether the media is subject to a death positivity bias when celebrities pass away (RQ1). It also explored whether some characteristics of the deceased can help explain the strength of the signal (RQ2). Finally, NLP methods were leveraged to identify what themes are highlighted before and after death, and see whether those can help explain the bias (RQ3). This section discusses the main findings and critically assesses the methods used.

6.1 Evidence of the death positivity bias (RQ1)

The research finds evidence of a small death positivity bias, aligning with the societal norm to "not speak ill of the dead" and with previous studies by Allison, Eylon, et al. (2009) and Hayes (2016). As mentioned in Section 2.4, Terror Management Theory (TMT) claims that when a celebrity dies, the public is reminded of their own mortality (Greenberg, Pyszczynski, and Solomon 1986; Greenberg, Vail, and Pyszczynski 2014). On top of this, fans experience parasocial grief similar to losing a close acquaintance (Section 2.3). Such reactions prompt the news providers to align their coverage with the readers' need for more positive narratives. The media take on their role of "national healers" (Kitch 2000, p. 189) by providing eulogising statements of the deceased in response to the tragic events.

6.2 Variations in media response (RQ2)

The fact that the media tends to speak more positively about celebrities once they have passed is especially true for show business celebrities, aligning with Rusu (2020). Movies, music, and television reach people of all ages, creating widespread emotional connections through performances, stories, and personas. This leads to stronger feelings of nostalgia and loss, prompting more positive media coverage. In contrast, sports achievements are tied to specific teams or events, political figures are polarizing, and academic achievements are less visible and more specialized. The nature of these careers may not evoke the same level of broad emotional connection as showbiz celebrities. The death positivity bias is also stronger for those who died of suicide or natural causes. The former can be explained by its tragic nature that might trigger sympathy for the deceased, reinterpretation of their life struggles, and romanticize them as tragic heroes. The latter, natural death, is seen as a peaceful, less controversial end, reducing the potential for negative or sensationalistic coverage, also known as "deathtertainment" (Arnold et al. 2017; Rusu 2020). The absence of tragic or dramatic circumstances allows the media to focus on the individual's achievements and legacy rather than the circumstances of their death. This likely leads to more positive media sentiment. Additionally, the older the individual at the time of death, the stronger the signal; except for people who die at 90+ years old. One possible explanation is that older celebrities have more extensive careers to look back on. A greater share of articles is therefore allocated to their life story, rather than the negative connotations associated with death circumstances. The fact that the trend reverses from 90 years old is rather surprising. Would those who die at 90+ not receive the same level of intense coverage due to perceived lesser newsworthiness?

6.3 Article content analysis (RQ3)

7 themes are identified in the articles. Four of those reflect the industries analysed in this work (i.e., topic 1: cinema; topic 2: public affairs; topic 3: music; topic 7: sport). Those prevail before death with new releases of albums, movies, political updates, or upcoming tournaments. As expected, although industry-related topics are more prominent before death, post-death articles also cover them as they go over the milestones achieved by the deceased. 2 topics are most prominent after death (topic 4: family, tribute; topic 6: crime, drugs). The former focuses on family-related matters and paying tribute to the dead. This is consistent with results from Alfano, Higgins, and Levernier (2018) who found mentions of family in the obituaries of local newspapers, and studies on obituaries in general which consistently identify the theme of paying tribute to the deceased (Alfano, Higgins, and Levernier 2018; Heynderickx and Dieltjens 2016; Rusu 2020). The latter, related to crime and drugs, is mostly connected to celebrities' death circumstances, making

its post-death prominence understandable. Although the detected topics are relevant to the subjects being studied, they do not help understand why post-death sentiment tends to be more positive. Some identified topics are composed of rather neutral keywords, making it difficult to associate them with a sentiment. Other topics are composed of a mix of positive and negative keywords. For instance, topic 4 (family, tribute) contains positive terms such as “love”, “birthday”, and “marry”; but the tribute component is more negative with terms like “death”, “tribute”, “pass” or “late”. This might be because sentiment analysis and topic modelling focus use different algorithms and approaches. The former is trained on labelled data, whereas the latter is an unsupervised learning method that focuses on uncovering latent themes through term co-occurrences, without considering sentiment.

6.4 General comments and future research

Although this thesis finds promising results that align with previous studies on the death positivity bias, some comments and limitations should be outlined. First, the results of this study are dependent on the selected celebrities and sampled article sources. The current work mostly covers articles from American media about American males. Another set of individuals could yield different results. Second, API cost limitations prevented the search over longer pre- and post-death periods. It is assumed that the selected 1-year periods accurately capture the collective sentiment regarding celebrities. Future research could make use of the current methods and apply them to more celebrities, longer periods of interest, and more articles. Third, Event Registry’s sentiment scores are assigned using a dictionary-based approach which comes with difficulties when faced with subjectivity, irony, negation, comparisons, or sarcasm (Hardeniya and Borikar 2016). It would be worth reproducing the research with more advanced language models that can better deal with such challenges. Fourth, linking the death positivity bias to the themes uncovered by topic modelling is challenging. Using recent approaches that bridge this gap by performing sentiment analysis at the topic level would be beneficial (Pathak, Pandey, and Rautaray 2021).

Finally, similar research could be done on social media to check whether this thesis' findings can be extended to other communication channels. By analysing suicide response on social media X, Ueda et al. (2017) show that there is a low correlation between traditional media coverage and social media activity. Also, suicides by young entertainers tend to bring more interest on X than in traditional media sources which focus more on older individuals and public officials (ibid.). Furthermore, studying the reactions of individuals rather than traditional media would allow for testing for differences in responses depending on the author's age and gender. These attributes are crucial factors influencing people's interactions with celebrities (Brown, Basil, and Bocarnea 2003). As Bandura et al. (1986)'s Social Learning Theory suggests, individuals are inclined to identify more strongly with those who share similar age and gender characteristics. Studying celebrity death responses on social media would therefore uncover additional insights on the death positivity bias.

References

Advanced name search (n.d.). IMDb. URL: https://www.imdb.com/search/name/?death_date=2014-06-01,2023-10-31.

Aitkenhead, Decca (Sept. 2010). *Robin Williams: 'I was shameful, did stuff that caused disgust – that's hard to recover from'*. The Guardian. URL: https://www.theguardian.com/film/2010/sep/20/robin-williams-worlds-greatest-dad-alcohol-drugs.

Alfano, Mark, Andrew Higgins, and Jacob Levernier (2018). "Identifying virtues and values through obituary data-mining". In: *The Journal of Value Inquiry* 52, pp. 59–79.

Allison, Scott T and Dafna Eylon (2005). "The demise of leadership: Death positivity biases in posthumous impressions of leaders". In: *The psychology of leadership: New perspectives and research* 295.

Allison, Scott T, Dafna Eylon, et al. (2009). "The demise of leadership: Positivity and negativity biases in evaluations of dead leaders". In: *The Leadership Quarterly* 20.2, pp. 115–129.

Anderson, James R (2016). "Comparative thanatology". In: *Current Biology* 26.13, R553–R556.

Arnold, Michael et al. (2017). *Death and digital media*. Routledge.

Bae, Hyuhn-Suhck, William J Brown, and Seok Kang (2010). "Social influence of a religious hero: The late Cardinal Stephen Kim Sou-hwan's effect on cornea donation and volunteerism". In: *Journal of Health Communication* 16.1, pp. 62–78.

Bandura, Albert et al. (1986). "Social foundations of thought and action". In: *Englewood Cliffs, NJ* 1986.23-28, p. 2.

Banjanovic, Erin S and Jason W Osborne (2019). "Confidence intervals for effect sizes: Applying bootstrap resampling". In: *Practical Assessment, Research, and Evaluation* 21.1, p. 5.

Bernardi, Bruno (Nov. 2022). *Maradona revisited: on his drugs ban, Berlusconi and 'the suffocating love of Naples'*. The Guardian. URL: https://www.theguardian.com/football/2022/nov/20/maradona-revisited-on-his-drugs-ban-berlusconi-and-the-suffocating-love-of-naples.

Bingaman, James (2022). ""Dude I've never felt this way towards a celebrity death": Parasocial grieving and the collective mourning of Kobe Bryant on Reddit". In: *OMEGA-Journal of death and dying* 86.2, pp. 364–381.

Brown, William J (2010). "Steve Irwin's influence on wildlife conservation". In: *Journal of Communication* 60.1, pp. 73–93.

Brown, William J, Michael D Basil, and Mihai C Bocarnea (2003). "Social influence of an international celebrity: Responses to the death of Princess Diana". In: *Journal of communication* 53.4, pp. 587–605.

Burgess, Jean, Peta Mitchell, and Felix Victor Münch (2018). "Social media rituals: The uses of celebrity death in digital culture". In: *A networked self and birth, life, death.* Routledge, pp. 224–239.

Cohen, Elizabeth L and Cynthia Hoffner (2016). "Finding meaning in a celebrity's death: The relationship between parasocial attachment, grief, and sharing educational health information related to Robin Williams on social network sites". In: *Computers in Human Behavior* 65, pp. 643–650.

Cohen, Jacob (2013). *Statistical power analysis for the behavioral sciences.* routledge.

— (2016). "A power primer." In.

Coscarelli, Joe (Oct. 2018). *XXXTentacion Discusses Abuse and Stabbings on Tape Released by Prosecutors*. The New York Times. URL: https://www.nytimes.com/2018/10/24/arts/music/xxxtentacion-abuse-recording.html.

Cureton, Edward E (1956). "Rank-biserial correlation". In: *Psychometrika* 21.3, pp. 287–290.

Dededer, Claire (May 2023). *'Can I still listen to David Bowie?' A superfan's dilemma.* The Guardian. URL: https://www.theguardian.com/books/2023/may/06/can-i-still-listen-to-david-bowie-a-superfans-dilemma.

Dellatto, Marisa (Oct. 2023). *The Highest-Paid Dead Celebrities Of 2023*. Forbes. URL: https://www.forbes.com/sites/marisadellatto/2023/10/30/highest-paid-dead-celebrities-2023-michael-jackson-elvis-presley-whitney-houston/?sh=3d498e4e504b.

Engelking, Carl (Mar. 2015). *A Promising Drug for Disease That Took Leonard Nimoy's Life*. Discover. URL: https://www.discovermagazine.com/health/a-promising-drug-for-disease-that-took-leonard-nimoys-life.

Fonseca, Luciana Mascarenhas and Ines Testoni (2012). "The emergence of thanatology and current practice in death education". In: *OMEGA-Journal of Death and Dying* 64.2, pp. 157–169.

Freeman, Hadley (Mar. 2019). *The Michael Jackson accusers: 'The abuse didn't feel strange, because he was like a god'*. The Guardian. URL: https://www.theguardian.com/tv-and-radio/2019/mar/04/the-michael-jackson-accusers-the-abuse-didnt-feel-strange-because-he-was-like-a-god.

Frisaro, Freida (Mar. 2023). *Jury convicts 3 of murder in death of rapper XXXTentacion.* AP News. URL: https://apnews.com/article/xxxtentacion-rapper-killed-trial-verdict-56ba2e060ff65d639d0a87ffb04e63fa.

Gach, Katie Z, Casey Fiesler, and Jed R Brubaker (2017). ""Control your emotions, Potter" An analysis of grief policing on Facebook in response to celebrity death". In: *Proceedings of the ACM on Human-Computer Interaction* 1.CSCW, pp. 1–18.

Gates, Emmet (June 2023). *Silvio Berlusconi Was Arguably Football's Most Influential Club Owner*. Forbes. URL: https://www.forbes.com/sites/emmetgates/2023/06/15/silvio-berlusconi-was-arguably-footballs-most-influential-club-owner/.

Glass, Gene V, Percy D Peckham, and James R Sanders (1972). "Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance". In: *Review of educational research* 42.3, pp. 237–288.

Greenberg, Jeff, Tom Pyszczynski, and Sheldon Solomon (1986). "The causes and consequences of a need for self-esteem: A terror management theory". In: *Public self and private self*. Springer, pp. 189–212.

Greenberg, Jeff, Kenneth Vail, and Tom Pyszczynski (2014). "Terror management theory and research: How the desire for death transcendence drives our strivings for meaning and significance". In: *Advances in motivation science*. Vol. 1. Elsevier, pp. 85–134.

Hardeniya, Tanvi and Dilipkumar A Borikar (2016). "Dictionary based approach to sentiment analysis-a review". In: *International Journal of Advanced Engineering, Management and Science* 2.5, p. 239438.

Harwell, Michael R et al. (1992). "Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases". In: *Journal of educational statistics* 17.4, pp. 315–339.

Hattenstone, Simon (Dec. 2009). *George Michael: 'I'm surprised I've survived my own dysfunction'*. The Guardian. URL: https://www.theguardian.com/music/2009/dec/05/george-michael-interview-music-sex-drugs.

Hayes, Joseph (2016). "Praising the dead: On the motivational tendency and psychological function of eulogizing the deceased". In: *Motivation and Emotion* 40, pp. 375–388.

Heynderickx, Priscilla C and Sylvain M Dieltjens (2016). "An analysis of obituaries in staff magazines". In: *Death studies* 40.1, pp. 11–21.

Horton, Donald and R Richard Wohl (1956). "Mass communication and para-social interaction: Observations on intimacy at a distance". In: *psychiatry* 19.3, pp. 215–229.

In Memoriam: Deaths in the 2010s (n.d.). IMDb. URL: https://m.imdb.com/list/ls093128325/.

Irwin donations top two million (Oct. 2006). The Sydney Morning Herald. URL: https://www.smh.com.au/national/irwin-donations-top-two-million-20061015-gdolmn.html.

Jansen, Harm (2016). "Celebrity deaths and social mood". MA thesis. Radboud University Nijmegen.

Jonze, Tim (Apr. 2024). *Who's bad? From Michael Jackson to David Bowie, why are some stars uncancellable?* The Guardian. URL: https://www.theguardian.com/music/2024/apr/01/michael-jackson-david-bowie-uncancellable-cancel-culture-mj-sexual-allegations-musical.

Juice WRLD (n.d.). Billboard. URL: https://www.billboard.com/artist/juice-wrld/.

Kastenbaum, Robert, Sara Peyton, and Beatrice Kastenbaum (1977). "Sex discrimination after death". In: *OMEGA-Journal of Death and Dying* 7.4, pp. 351–359.

Kaufman, Gil (July 2009). *Michael Jackson Dominates 'Billboard' Charts*. URL: https://www.mtv.com/news/c8ojrs/michael-jackson-dominates-billboard-charts.

Kehoe, Adelle (2022). *Amazon Shut Down Alexa, Similarweb Ramped Up*. Similarweb. URL: https://www.similarweb.com/blog/updates/announcements/alexa-shutting-down-similarweb-ramping-up/.

Kim, Tae Kyun (2015). "T test as a parametric statistic". In: *Korean journal of anesthesiology* 68.6, p. 540.

Kitch, Carolyn (2000). "'A news of feeling as well as fact' Mourning and memorial in American newsmagazines". In: *Journalism* 1.2, pp. 171–195.

Kreps, Daniel (May 2021). *Chris Cornell's Family Settles With Doctor Who Prescribed Singer Drugs*. Rolling Stone. URL: https://www.rollingstone.com/music/music-news/chris-cornell-family-settlement-doctor-1166269/.

Kuo, Christopher (Aug. 2023). *Sexual Abuse Suits Against Michael Jackson's Companies Are Revived*. The New York Times. URL: https://www.nytimes.com/2023/08/18/arts/music/michael-jackson-sexual-abuse-lawsuits.html.

Langsrud, Øyvind (2003). "ANOVA for unbalanced data: Use Type II instead of Type III sums of squares". In: *Statistics and computing* 13.2, pp. 163–167.

Leonard, Lynnette G and Paige Toller (2012). "Speaking ill of the dead: Anonymity and communication about suicide on MyDeathSpace. com". In: *Communication Studies* 63.4, pp. 387–404.

Lepori, Gabriele M (2021). "A nonrandom walk down Hollywood boulevard: Celebrity deaths and investor sentiment". In: *Financial Review* 56.3, pp. 591–613.

Lists of people by cause of death (n.d.). Wikipedia. URL: https://en.wikipedia.org/wiki/Lists_of_people_by_cause_of_death.

Lix, Lisa M, Joanne C Keselman, and Harvey J Keselman (1996). "Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test". In: *Review of educational research* 66.4, pp. 579–619.

Mann, Henry B and Donald R Whitney (1947). "On a test of whether one of two random variables is stochastically larger than the other". In: *The annals of mathematical statistics*, pp. 50–60.

Matza, Max (Mar. 2023). *Kobe Bryant's widow gets almost $29m settlement from LA*. BBC. URL: https://www.bbc.com/news/world-us-canada-64806905.

Michaels, Sean (July 2008). *Motörhead's Lemmy in Nazi photoshoot scandal*. The Guardian. URL: https://www.theguardian.com/music/2008/jul/11/news.culture.

Moore, Suzanne (June 2018). *XXXTentacion's brutal life points to the problem with UK drug policy*. The Guardian. URL: https://www.theguardian.com/commentisfree/2018/jun/20/xxxtentacion-brutal-life-problem-uk-drug-policy.

Moremen, Robin D and Cathy Cradduck (1999). ""How will you be remembered after you die?" Gender discrimination after death twenty years later". In: *OMEGA-Journal of Death and Dying* 38.4, pp. 241–254.

Myrick, Jessica Gall (2017). "Identification and emotions experienced after a celebrity cancer death shape information sharing and prosocial behavior". In: *Journal of Health Communication* 22.6, pp. 515–522.

Myrick, Jessica Gall and Jessica Fitts Willoughby (2019). "The role of media-induced nostalgia after a celebrity death in shaping audiences' social sharing and prosocial behavior". In: *Journal of Health Communication* 24.5, pp. 461–468.

Niederkrotenthaler, Thomas, Benedikt Till, and David Garcia (2019). "Celebrity suicide on Twitter: Activity, content and network analysis related to the death of Swedish DJ Tim Bergling alias Avicii". In: *Journal of affective disorders* 245, pp. 848–855.

O'callaghan, Derek et al. (2015). "An analysis of the coherence of descriptors in topic modeling". In: *Expert Systems with Applications* 42.13, pp. 5645–5657.

Park, Sejung and Cynthia A Hoffner (2020). "Tweeting about mental health to honor Carrie Fisher: How# InHonorOfCarrie reinforced the social influence of celebrity advocacy". In: *Computers in human behavior* 110, p. 106353.

Pathak, Ajeet Ram, Manjusha Pandey, and Siddharth Rautaray (2021). "Topic-level sentiment analysis of social media data using deep learning". In: *Applied Soft Computing* 108, p. 107440.

Povoledo, Elisabetta (Dec. 2022). *Benedict was criticized for his handling of the church's sex abuse scandal.* The New York Times. URL: https://www.nytimes.com/2022/12/31/world/europe/pope-emeritus-benedict-xvi-sex-abuse-scandal.html.

Radford, Scott K and Peter H Bloch (2012). "Grief, commiseration, and consumption following the death of a celebrity". In: *Journal of consumer culture* 12.2, pp. 137–155.

— (2013). "Consumers' online responses to the death of a celebrity". In: *Marketing Letters* 24, pp. 43–55.

Rajan, Benson and Sahana Sarkar (2018). "Analysing grief on twitter: a study of digital expressions on Om Puri's death". In: *Funes-Journal of narratives and social sciences* 2, pp. 136–152.

Röder, Michael, Andreas Both, and Alexander Hinneburg (2015). "Exploring the space of topic coherence measures". In: *Proceedings of the eighth ACM international conference on Web search and data mining*, pp. 399–408.

Rusu, Mihai S (2020). "Celebrities' memorial afterlives: Obituaries, tributes, and posthumous gossip in the Romanian media deathscape". In: *OMEGA-Journal of Death and Dying* 80.4, pp. 568–591.

Sanderson, Jimmy and Pauline Hope Cheong (2010). "Tweeting prayers and communicating grief over Michael Jackson online". In: *Bulletin of science, technology & society* 30.5, pp. 328–340.

Shapiro, Samuel Sanford and Martin B Wilk (1965). "An analysis of variance test for normality (complete samples)". In: *Biometrika* 52.3-4, pp. 591–611.

Singh, R Raj (2016). *Death, contemplation and Schopenhauer.* Routledge.

Syed, Shaheen and Marco Spruit (2017). "Full-text or abstract? examining topic coherence scores using latent dirichlet allocation". In: *2017 IEEE International conference on data science and advanced analytics (DSAA).* Ieee, pp. 165–174.

Ueda, Michiko et al. (2017). "Tweeting celebrity suicides: Users' reaction to prominent suicide deaths on Twitter and subsequent increases in actual suicides". In: *Social Science & Medicine* 189, pp. 158–166.

Van den Bulck, Hilde and Anders Olof Larsson (2019). "'There'sa Starman waiting in the sky': Mourning David# Bowie on Twitter". In: *Convergence* 25.2, pp. 307–323.

Wasserman, Larry (2013). *All of statistics: a concise course in statistical inference.* Springer Science & Business Media.

XUK (2016). *Fans pay their respects to music legend David Bowie with over 2.3 million tweets, peaking at 20k tweets per min.* URL: https://twitter.com/XUK/status/686504368456253440.

Yanofsky, David (Jan. 2016). *This is what David Bowie's death looked like to Spotify*. URL: https://qz.com/591803/this-is-what-david-bowies-death-looked-like-to-spotify.

Yin, Sara (Oct. 2011). *Sales of Steve Jobs's Black Turtleneck Soar*. URL: https://uk.pcmag.com/news/113404/sales-of-steve-jobss-black-turtleneck-soar.

Yuhas, Alan (June 2023). *The Many Twists, Quotes and Scandals of Silvio Berlusconi*. The New York Times. URL: https://www.nytimes.com/2023/06/12/world/europe/italy-berlusconi-career.html.

A RQ1 Appendices

A.1 Filtered dataset of celebrities

Table 7: Filtered dataset of 38 celebrities

Celebrity	English	Gender	Industry	Cause of death	Age	Unexpected	Controversial
Alan Rickman	Yes	Male	Cinema	Illness	60-69	No	No
Alan Thicke	Yes	Male	Cinema	Cardiorespiratory	60-69	Yes	No
Avicii	No	Male	Music	Suicide	20-29	Yes	No
Carrie Fisher	Yes	Female	Cinema	Cardiorespiratory	60-69	Yes	No
Chris Cornell	Yes	Male	Music	Suicide	50-59	Yes	No
Chuck Berry	Yes	Male	Music	Cardiorespiratory	90-99	No	Yes
David Bowie	Yes	Male	Music	Illness	60-69	Yes	Yes
Diego Maradona	No	Male	Sport	Cardiorespiratory	60-69	Yes	Yes
Doris Day	Yes	Female	Cinema	Illness	90-99	No	No
Dwayne Haskins	Yes	Male	Sport	Accident	20-29	Yes	Yes
Elizabeth II	Yes	Female	Public affairs	Natural	90-99	No	Yes
Fidel Castro	No	Male	Public affairs	Undisclosed	90-99	No	Yes
Florence Henderson	Yes	Female	Cinema	Cardiorespiratory	80-89	No	No
George H. W. Bush	Yes	Male	Public affairs	Illness	90-99	No	Yes
George Michael	Yes	Male	Music	Natural	50-59	Yes	Yes
Joan Rivers	Yes	Female	Cinema	Accident	80-89	Yes	Yes
Johnny Hallyday	No	Male	Music	Illness	70-79	No	No
Juice Wrld	Yes	Male	Music	Overdose	20-29	Yes	Yes
Kelly Preston	Yes	Female	Cinema	Illness	50-59	Yes	No
Ken Block	Yes	Male	Sport	Accident	50-59	Yes	No
Kobe Bryant	Yes	Male	Sport	Accident	40-49	Yes	Yes
Lemmy	Yes	Male	Music	Illness	70-79	Yes	Yes
Leonard Cohen	Yes	Male	Music	Illness	80-89	No	No
Leonard Nimoy	Yes	Male	Cinema	Illness	80-89	No	No
Mac Miller	Yes	Male	Music	Overdose	20-29	Yes	No
Mary Tyler Moore	Yes	Female	Cinema	Cardiorespiratory	80-89	No	No

Table 7: Filtered dataset of 38 celebrities (cont.)

Celebrity	**English**	**Gender**	**Industry**	**Cause of death**	**Age**	**Unexpected**	**Controversial**
Matthew Perry	Yes	Male	Cinema	Overdose	50-59	Yes	Yes
Mikhail Gorbachev	No	Male	Public affairs	Illness	90-99	No	Yes
Muhammad Ali	Yes	Male	Sport	Cardiorespiratory	70-79	No	No
Pervez Musharraf	No	Male	Public affairs	Illness	70-79	No	Yes
Pope Benedict XVI	No	Male	Public affairs	Illness	90-99	No	Yes
Prince (musician)	Yes	Male	Music	Overdose	50-59	Yes	Yes
Robin Williams	Yes	Male	Cinema	Suicide	60-69	Yes	No
Sean Connery	Yes	Male	Cinema	Cardiorespiratory	90-99	No	No
Silvio Berlusconi	No	Male	Public affairs	Illness	80-89	No	Yes
Stan Lee	Yes	Male	Cinema	Cardiorespiratory	90-99	No	Yes
Stephen Hawking	Yes	Male	Academia	Illness	70-79	No	No
XXXTentacion	Yes	Male	Music	Assassination	20-29	Yes	Yes

A.2 Bootstrap of effect size

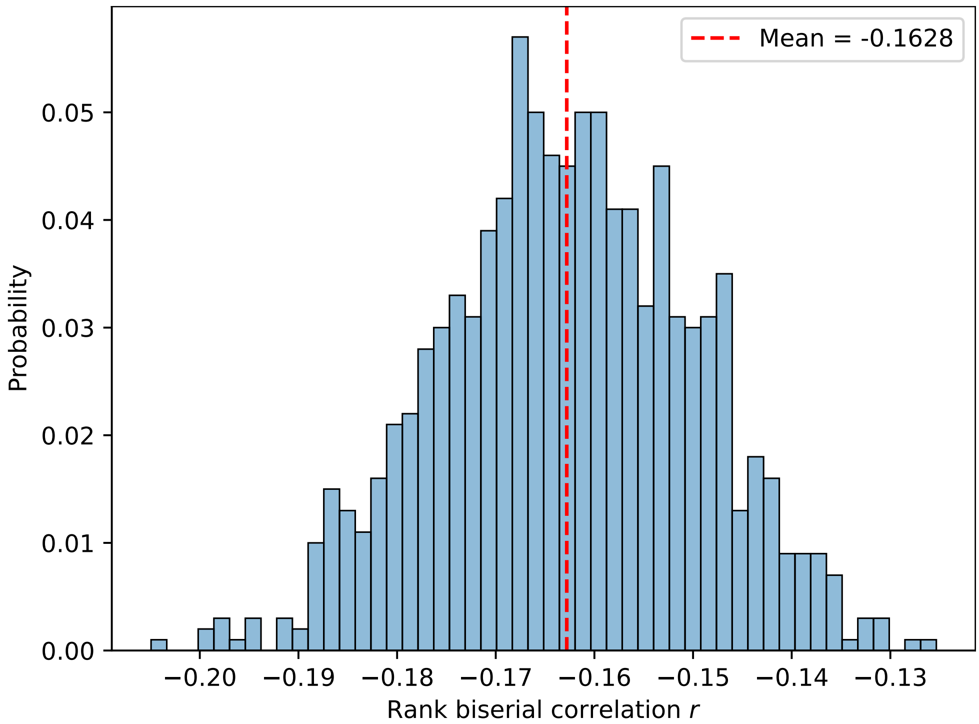

Figure 16: Bootstrap distribution of the rank biserial correlation coefficient r for all celebrities

A.3 One-sided Wilcoxon-Mann-Whitney U-tests and effect sizes

Table 8: Bootstrapped one-sided Wilcoxon-Mann-Whitney U-tests and rank biserial correlation coefficients r

Name	U-stat	p-value	r	r 95% CI
Leonard Nimoy	505.3665	0.0***	-0.8989	[-0.9741 -0.8237]
Avicii	815.767	0.0***	-0.8368	[-0.918 -0.7557]
Ken Block	872.7325	0.0***	-0.8255	[-0.9175 -0.7334]
Prince (musician)	1407.8355	0.0***	-0.7184	[-0.8257 -0.6112]
Alan Rickman	1478.494	0.0***	-0.7043	[-0.8207 -0.5879]
Alan Thicke	1728.681	0.0***	-0.6543	[-0.7757 -0.5329]
Lemmy	1761.3605	0.0***	-0.6477	[-0.7683 -0.5272]
Robin Williams	2504.9715	0.0***	-0.499	[-0.6403 -0.3577]
Stephen Hawking	2623.14	0.0***	-0.4754	[-0.6173 -0.3334]
Joan Rivers	2796.5275	0.0***	-0.4407	[-0.5775 -0.3039]
Florence Henderson	2940.3475	0.0002***	-0.4119	[-0.5681 -0.2558]
Stan Lee	2969.9885	0.0002***	-0.406	[-0.5548 -0.2572]
George H. W. Bush	2999.144	0.0001***	-0.4002	[-0.5415 -0.2588]
George Michael	3062.9955	0.0004***	-0.3874	[-0.5358 -0.239]
Doris Day	3185.158	0.0014***	-0.363	[-0.5281 -0.1978]
Muhammad Ali	3215.673	0.0011***	-0.3569	[-0.5044 -0.2093]
Pervez Musharraf	3347.0405	0.0015***	-0.3306	[-0.4774 -0.1838]
Leonard Cohen	3544.811	0.0047***	-0.291	[-0.4446 -0.1375]
XXXTentacion	3582.114	0.006***	-0.2836	[-0.4357 -0.1315]
Johnny Hallyday	3970.3505	0.0366**	-0.2059	[-0.3612 -0.0506]
Carrie Fisher	4098.138	0.0593*	-0.1804	[-0.3397 -0.021]
David Bowie	4348.9035	0.1282	-0.1302	[-0.287 0.0265]
Mikhail Gorbachev	4694.2325	0.301	-0.0612	[-0.2274 0.1051]
Pope Benedict XVI	4789.8075	0.3599	-0.042	[-0.2062 0.1221]
Mary Tyler Moore	5270.7955	0.675	0.0542	[-0.116 0.2243]
Chuck Berry	5733.697	0.8956	0.1467	[-0.0169 0.3104]
Elizabeth II	5775.9795	0.909	0.1552	[-0.0064 0.3168]
Silvio Berlusconi	6012.5195	0.9632	0.2025	[0.0485 0.3565]
Matthew Perry	6463.751	0.9956	0.2928	[0.1473 0.4382]
Sean Connery	6517.5595	0.9957	0.3035	[0.15 0.4571]
Kelly Preston	6642.246	0.9985	0.3284	[0.1844 0.4725]
Juice Wrld	6844.053	0.9993	0.3688	[0.2195 0.5182]
Chris Cornell	6909.3555	0.9997	0.3819	[0.2395 0.5242]
Fidel Castro	7062.49	0.9998	0.4125	[0.2597 0.5653]
Diego Maradona	7468.171	1.0	0.4936	[0.3623 0.625]
Kobe Bryant	7624.3215	1.0	0.5249	[0.3904 0.6593]
Mac Miller	7884.935	1.0	0.577	[0.4521 0.7019]
Dwayne Haskins	8965.9525	1.0	0.7932	[0.6983 0.8881]
ALL CELEBRITIES	6044570.695	0.0***	-0.1628	[-0.188 -0.1377]

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.4 Two-sided Wilcoxon-Mann-Whitney U-tests and effect sizes

Table 9: Bootstrapped two-sided Wilcoxon-Mann-Whitney U-tests and rank biserial correlation coefficients r

Name	**U-stat**	**p-value**	r	r **95% CI**
Leonard Nimoy	505.3665	0.0***	-0.8989	[-0.9741 -0.8237]
Avicii	815.767	0.0***	-0.8368	[-0.918 -0.7557]
Ken Block	872.7325	0.0***	-0.8255	[-0.9175 -0.7334]
Prince (musician)	1407.8355	0.0***	-0.7184	[-0.8257 -0.6112]
Alan Rickman	1478.494	0.0***	-0.7043	[-0.8207 -0.5879]
Alan Thicke	1728.681	0.0***	-0.6543	[-0.7757 -0.5329]
Lemmy	1761.3605	0.0***	-0.6477	[-0.7683 -0.5272]
Robin Williams	2504.9715	0.0***	-0.499	[-0.6403 -0.3577]
Stephen Hawking	2623.14	0.0***	-0.4754	[-0.6173 -0.3334]
Joan Rivers	2796.5275	0.0001***	-0.4407	[-0.5775 -0.3039]
Florence Henderson	2940.3475	0.0004***	-0.4119	[-0.5681 -0.2558]
Stan Lee	2969.9885	0.0003***	-0.406	[-0.5548 -0.2572]
George H. W. Bush	2999.144	0.0003***	-0.4002	[-0.5415 -0.2588]
George Michael	3062.9955	0.0008***	-0.3874	[-0.5358 -0.239]
Doris Day	3185.158	0.0027***	-0.363	[-0.5281 -0.1978]
Muhammad Ali	3215.673	0.0022***	-0.3569	[-0.5044 -0.2093]
Pervez Musharraf	3347.0405	0.003***	-0.3306	[-0.4774 -0.1838]
Leonard Cohen	3544.811	0.0094***	-0.291	[-0.4446 -0.1375]
XXXTentacion	3582.114	0.012**	-0.2836	[-0.4357 -0.1315]
Johnny Hallyday	3970.3505	0.0706*	-0.2059	[-0.3612 -0.0506]
Carrie Fisher	4098.138	0.1125	-0.1804	[-0.3397 -0.021]
David Bowie	4348.9035	0.2287	-0.1302	[-0.287 0.0265]
Mikhail Gorbachev	4694.2325	0.4051	-0.0612	[-0.2274 0.1051]
Pope Benedict XVI	4789.8075	0.45	-0.042	[-0.2062 0.1221]
Mary Tyler Moore	5270.7955	0.4313	0.0542	[-0.116 0.2243]
Chuck Berry	5733.697	0.1789	0.1467	[-0.0169 0.3104]
Elizabeth II	5775.9795	0.1651	0.1552	[-0.0064 0.3168]
Silvio Berlusconi	6012.5195	0.0704*	0.2025	[0.0485 0.3565]
Matthew Perry	6463.751	0.0088***	0.2928	[0.1473 0.4382]
Sean Connery	6517.5595	0.0086***	0.3035	[0.15 0.4571]
Kelly Preston	6642.246	0.003***	0.3284	[0.1844 0.4725]
Juice Wrld	6844.053	0.0014***	0.3688	[0.2195 0.5182]
Chris Cornell	6909.3555	0.0006***	0.3819	[0.2395 0.5242]
Fidel Castro	7062.49	0.0003***	0.4125	[0.2597 0.5653]
Diego Maradona	7468.171	0.0***	0.4936	[0.3623 0.625]
Kobe Bryant	7624.3215	0.0***	0.5249	[0.3904 0.6593]
Mac Miller	7884.935	0.0***	0.577	[0.4521 0.7019]
Dwayne Haskins	8965.9525	0.0***	0.7932	[0.6983 0.8881]
ALL CELEBRITIES	6044570.695	0.0***	-0.1628	[-0.188 -0.1377]

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

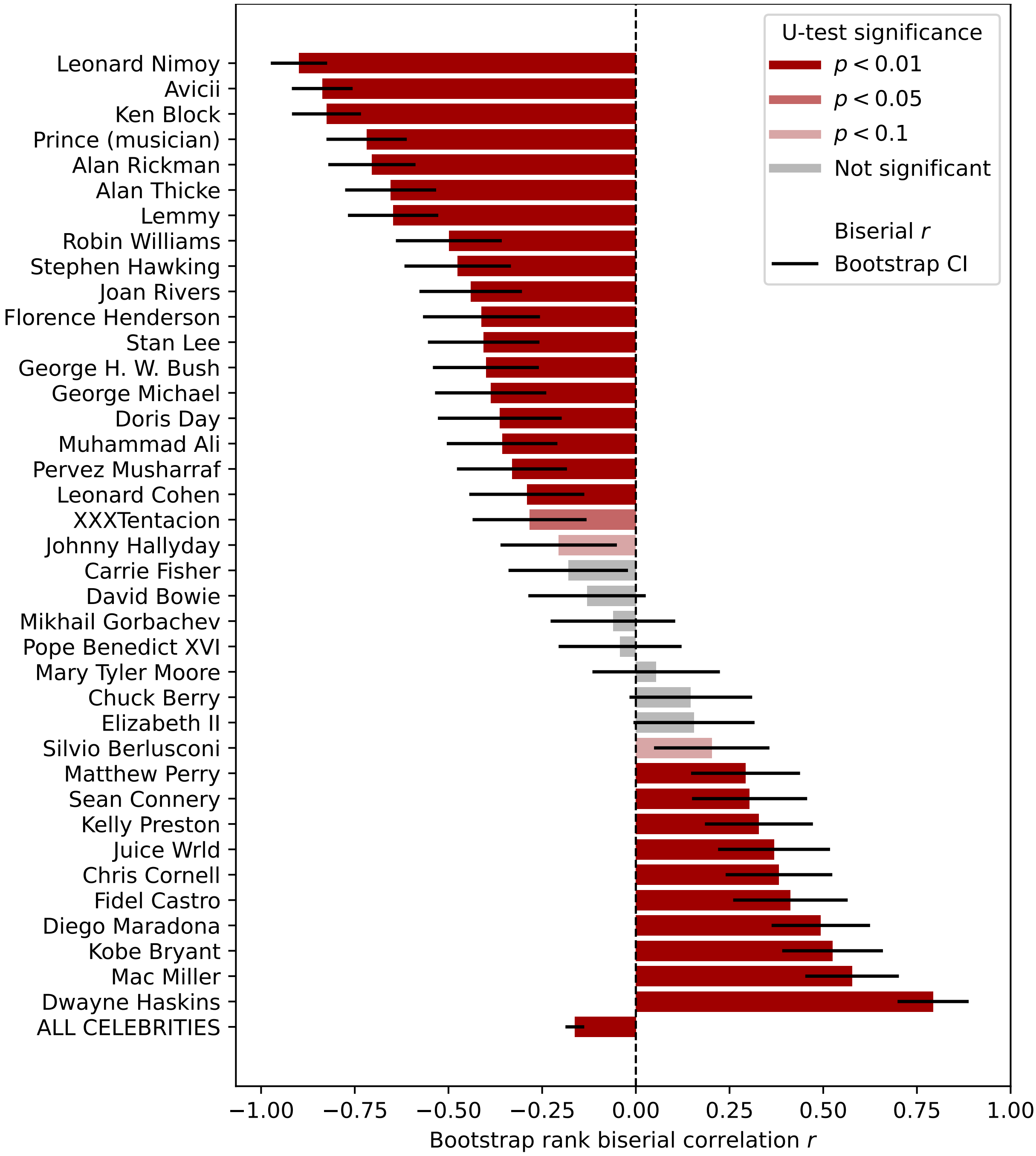

Figure 17: Two-sided Mann-Whitney U-tests and effect sizes

See the description of Figure 9. Here, the bar colours represent the level of significance of the two-sided Mann-Whitney U-test formulated in Section 3.3.1. The values used to compute this plot can be found in Table 9 (Appendix 9).

B RQ2 Appendices

B.1 Sentiment difference distribution

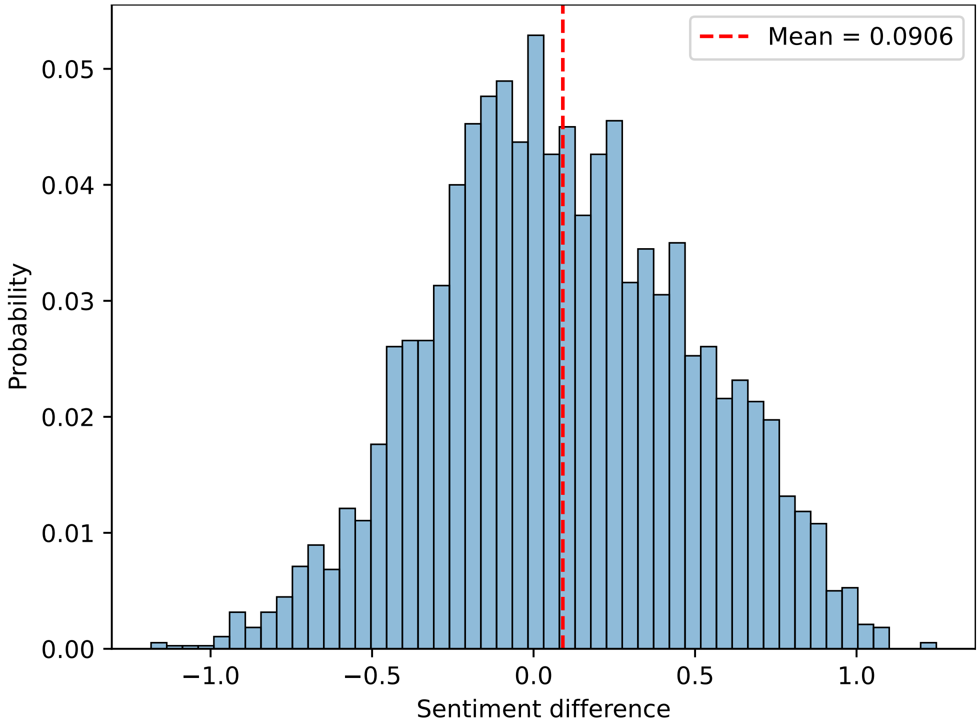

Figure 18: Distribution of the sentiment differences D for all celebrities

B.2 Tukey's Honestly Significant Difference tests

Table 10: Tukey's HSD tests

Feature	Group 1	Group 2	$\bar{D}_1 - \bar{D}_2$	95% CI	p-value
industry	Academia	Cinema	-0.0719	[-0.183 0.0392]	0.3936
industry	Academia	Music	-0.1424	[-0.2538 -0.0309]	0.0045***
industry	Academia	Public affairs	-0.2334	[-0.3479 -0.119]	0.0***
industry	Academia	Sport	-0.2833	[-0.4006 -0.1661]	0.0***
industry	Cinema	Music	-0.0705	[-0.1133 -0.0276]	0.0001***
industry	Cinema	Public affairs	-0.1615	[-0.2117 -0.1113]	0.0***
industry	Cinema	Sport	-0.2114	[-0.2678 -0.1551]	0.0***
industry	Music	Public affairs	-0.0911	[-0.142 -0.0401]	0.0***
industry	Music	Sport	-0.141	[-0.198 -0.084]	0.0***
industry	Public affairs	Sport	-0.0499	[-0.1126 0.0128]	0.1899
cause	Accident	Assassination	0.0969	[-0.036 0.2298]	0.3454
cause	Accident	Cardiorespiratory	0.0592	[-0.0123 0.1306]	0.191
cause	Accident	Illness	0.142	[0.0746 0.2094]	0.0***
cause	Accident	Natural	0.0605	[-0.0425 0.1634]	0.6332
cause	Accident	Overdose	-0.0635	[-0.1476 0.0205]	0.2984
cause	Accident	Suicide	0.172	[0.0812 0.2628]	0.0***
cause	Accident	Undisclosed	-0.1658	[-0.2987 -0.0328]	0.0039***
cause	Assassination	Cardiorespiratory	-0.0377	[-0.1631 0.0876]	0.9849
cause	Assassination	Illness	0.0451	[-0.078 0.1682]	0.9544
cause	Assassination	Natural	-0.0364	[-0.1821 0.1092]	0.9951
cause	Assassination	Overdose	-0.1604	[-0.2934 -0.0275]	0.0062***
cause	Assassination	Suicide	0.0751	[-0.0622 0.2124]	0.7138
cause	Assassination	Undisclosed	-0.2627	[-0.4308 -0.0945]	0.0001***
cause	Cardiorespiratory	Illness	0.0829	[0.0321 0.1337]	0.0***
cause	Cardiorespiratory	Natural	0.0013	[-0.0916 0.0943]	1.0
cause	Cardiorespiratory	Overdose	-0.1227	[-0.1941 -0.0512]	0.0***
cause	Cardiorespiratory	Suicide	0.1129	[0.0336 0.1921]	0.0004***
cause	Cardiorespiratory	Undisclosed	-0.2249	[-0.3502 -0.0996]	0.0***
cause	Illness	Natural	-0.0816	[-0.1714 0.0083]	0.1079
cause	Illness	Overdose	-0.2056	[-0.273 -0.1381]	0.0***
cause	Illness	Suicide	0.03	[-0.0457 0.1056]	0.9315
cause	Illness	Undisclosed	-0.3078	[-0.4309 -0.1847]	0.0***
cause	Natural	Overdose	-0.124	[-0.227 -0.021]	0.0064***
cause	Natural	Suicide	0.1115	[0.003 0.2201]	0.0389**
cause	Natural	Undisclosed	-0.2262	[-0.3719 -0.0806]	0.0001***
cause	Overdose	Suicide	0.2355	[0.1447 0.3264]	0.0***
cause	Overdose	Undisclosed	-0.1022	[-0.2352 0.0307]	0.2762
cause	Suicide	Undisclosed	-0.3378	[-0.4751 -0.2005]	0.0***
age	20-29	40-49	-0.2552	[-0.3803 -0.1301]	0.0***
age	20-29	50-59	0.116	[0.0469 0.1852]	0.0***

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The 4th column is subtracting the mean difference in sentiment $\bar{D}$ of group 2 from the one of group 1. The mean difference values $\bar{D}$ are detailed in table 3.

Table 10: Tukey's HSD tests (cont.)

Feature	Group 1	Group 2	$\bar{D}_1 - \bar{D}_2$	95% CI	p-value
age	20-29	60-69	0.1701	[0.1009 0.2392]	0.0***
age	20-29	70-79	0.2156	[0.1434 0.2879]	0.0***
age	20-29	80-89	0.239	[0.1698 0.3082]	0.0***
age	20-29	90-99	0.0563	[-0.0074 0.12]	0.1242
age	40-49	50-59	0.3712	[0.2478 0.4946]	0.0***
age	40-49	60-69	0.4252	[0.3019 0.5486]	0.0***
age	40-49	70-79	0.4708	[0.3457 0.5959]	0.0***
age	40-49	80-89	0.4942	[0.3708 0.6175]	0.0***
age	40-49	90-99	0.3115	[0.1911 0.4318]	0.0***
age	50-59	60-69	0.054	[-0.0119 0.12]	0.1913
age	50-59	70-79	0.0996	[0.0304 0.1687]	0.0004***
age	50-59	80-89	0.123	[0.057 0.1889]	0.0***
age	50-59	90-99	-0.0598	[-0.1199 0.0004]	0.0532*
age	60-69	70-79	0.0456	[-0.0236 0.1147]	0.4518
age	60-69	80-89	0.0689	[0.003 0.1349]	0.0337**
age	60-69	90-99	-0.1138	[-0.174 -0.0536]	0.0***
age	70-79	80-89	0.0234	[-0.0458 0.0925]	0.9547
age	70-79	90-99	-0.1593	[-0.223 -0.0956]	0.0***
age	80-89	90-99	-0.1827	[-0.2429 -0.1225]	0.0***

Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The 4$^{\text{th}}$ column is subtracting the mean difference in sentiment $\bar{D}$ of group 2 from the one of group 1. The mean difference values $\bar{D}$ are detailed in table 3.