

The Portrayal of Health Disparities in New Jersey's Local Newspapers: A Text Mining Approach

Mumtahina Obaid¹; Rifat Afrin^{1*}; Ahasan Harun²; Gayle Prybutok¹; Victor R Prybutok³

¹Department of Rehabilitation and Health Services, College of Health and Public Service, University of North Texas, Denton, TX, USA.

²Department of Information Systems, Robert C. Vackar College of Business & Entrepreneurship, University of Texas-Rio Grande Valley, Edinburg, TX, USA.

³Vice Provost for Graduate Education and Dean of the Toulouse Graduate School, Regents Professor of Decision Sciences, ITDS Dept., G. Brint Ryan College of Business, University of North Texas, Denton, TX, USA.

*Corresponding Author: [Rifat Afrin](#)

Email: rifatafrin@my.unt.edu

Abstract

Introduction: Health disparity ensues when racial and ethnic minorities experience limitations in their ability to access health services, resulting in poor health outcomes. The extent to which the health disparity issue is covered in the news impacts how individuals perceive the severity of the issue and individuals' attitudes toward the issue, which can influence policymakers' actions. Although a large body of academic literature focuses on health disparities, the perspective of news media's coverage of this issue has not received much attention. We address this gap by concentrating on how health disparities are portrayed in the local newspapers in New Jersey, USA.

Methods: We collected relevant news articles to investigate the framing of health disparity. We then employed a co-occurrence network to identify the words that mostly co-occur together and build a frame. In addition, we used rapid automatic keyword extraction (RAKE) to identify the most frequent phrases from news texts and sentiment analysis to detect the expressed emotions in the news frames.

Results: Health disparity coverage in New Jersey newspapers reveals several distinct frames. These highlight issues such as the effects of the pandemic on marginalized communities, racial disparities in health policies, and the timing of advocacy actions. RAKE-identified keywords in the news coverage emphasize health issues such as colorectal cancer, mental health, maternal mortality, and the positive impact of programs like needle exchange while underscoring the disproportionate effects of health disparities on the Black community and the need for healthcare reform.

Conclusion: This research broadens our understanding of health disparity issues in news media from a unique local perspective. These findings make it clear that the news media significantly influences how the public perceives health disparities. Our research emphasizes how crucial it is to watch media coverage to guarantee empathetic and educational conversations about health disparities.

Keywords: Health disparity; Newspapers; Health communication; Community health.

Citation: Obaid M, Afrin R, Harun A, Prybutok G, Prybutok VR. The Portrayal of Health Disparities in New Jersey's Local Newspapers: A Text Mining Approach. Med Discoveries. 2024; 3(1): 1108.

Introduction

Health disparities are the differences that socially disadvantaged groups encounter in the burden of diseases, injuries, violence, or opportunities to reach optimal access to health care services [1]. Examples of health disparities include variations in the prevalence of a particular illness or condition, how much a person is likely to get sufficient and high-quality care, or someone's access to doctors and medical insurance coverage. Populations of color — Hispanics, Blacks, Alaska Natives, and American Indians — are more likely to be uninsured [2]. They are also more likely to encounter financial care barriers and to accrue medical debt [2]. Communities of color have suffered significantly and negatively due to years of racism in the US [3]. Print media is a key channel for disseminating health-related information in society [4] and has the power to affect public perception and opinion on a specific subject [5]. In the US, newspapers are among the media outlets with the highest readership and influence [6]. Several scholarly research articles have shown how health topics are discussed in newspapers [7-9]. For this reason, this study focuses on the topic of health disparity coverage in newspapers.

Despite the improvement that has been made to better health care service delivery, ethnic and racial health disparities remain among the most difficult to correct [10]. Moreover, while evaluating health disparities, ethnicity and race become relevant considerations [11]. Newspapers are effective means for informing the public about societal concerns. Hence, it is important to analyze current news coverage of health disparities to determine how recently and how thoroughly the topic of health disparities has been discussed. We also need to identify what unique themes related to health disparities received more attention. Additionally, there is a need to focus on US local or state coverage. We first need to look at the circumstances in different states to get a perspective about the coverage of health disparities. Along with access to medical services and the overall standard of care, people's health statuses can differ significantly between and within states [2]. For this reason, the study's primary focus is on New Jersey newspapers' reporting about health disparities.

In New Jersey, areas with a high concentration of Black people continue to bear an unfair burden during the pandemic. Two New Jersey towns provided evidence of the COVID pandemic's ongoing racial divide. In New Jersey, the towns of Teaneck (dominated by Black and Hispanic neighborhoods) and Ridgewood (dominated by a white population) are a short distance apart. They were, however, poles apart in terms of illness and fatalities during the COVID pandemic, highlighting the ongoing racial disparity in the disease's toll. Teaneck's death rate is about twice that of neighboring Ridgewood [12]. Such discrepancies between people of color and white people are brought on by limited access to care, especially primary care, and a lack of health insurance [12]. The New Jersey Department of Health seeks to attain health equity and decrease premature death among New Jerseyans [13]. Therefore, it is necessary to explore how the local newspapers of New Jersey portray this current health disparity.

To do so, we explore the following research questions in our study: how is the topic of health disparity framed in New Jersey newspapers? And what sentiments do these news pieces express? To address these questions, our research has several unique objectives. The first objective was to help healthcare providers to redouble their efforts to address the underlying

health disparities and to determine whether any activities or behaviors on their part are increasing the identified health disparities. This will help healthcare planners and providers to effectively utilize available resources to obtain the most positive patient outcomes possible. The second objective was to use a co-occurrence network, RAKE, and sentiment analysis to deliver a more useful assessment. The co-occurrence network and RAKE help to identify the significant frames or themes in textual documents. Finally, we will attempt to deepen our understanding of the news items' attitudes and emotions through sentiment analysis.

Methodology

Data collection: Examining how health disparities are covered in New Jersey newspapers is the goal of this study. The NewsBank database was used to compile the New Jersey newspapers for this project. The chosen time frame, July 2021 to June 2022, provided access to recent news coverage on health disparities. The newspapers were searched containing the word "health" and the words either "gap", or "divide", or variations on the abbreviations inequ* or disparit* as well as terms that refer to particular groups of race or ethnicity or the more general terms "ethnic" or "race", or "racial" [14]. A total of 123 news articles were retained after duplicates and irrelevant articles were removed.

Analysis procedure: Initially, we used a co-occurrence network procedure. This network of word co-occurrences explains relationships between commonly used terms. A node in this diagram represents each word, and an edge represents each word's relationship to other words. Moreover, we used the RAKE (rapid automatic keyword extraction) method. This is an unsupervised, language and domain independent approach for retrieving words or phrases from textual documents [15]. Furthermore, we performed an automated sentiment analysis. The methodology utilized in this research is lexicon-based. The sentiment analysis approach that is based on lexicons is superior to/outperformed methods for machine learning [16]. Prior research has indicated that the lexicon-based technique for sentiment analysis is best. Hence, it is employed in this study. Positive and negative word lists make up the lexical resources that are polarity-oriented [17].

Findings: Figure 1 presents the co-occurrence network that detects the word clusters or news frames from health disparity coverage in New Jersey newspapers. We can see that there are a total of ten frames. In the most dominant frame or frame one, we find that health has connections with the pandemic, community, program, state, etc. As mentioned, some people suffer from health disparities because of their social environments, which constrains their opportunities in life. These news items corroborate that fact. During the Covid pandemic, these truths became more pronounced. People of color frequently expressed a lack of confidence in government assistance. The theme of racial disparity is more prominent in frame two as we can see the words disparity, racial, policy, etc. Frame three contains the words time, day, group, etc., and provides information regarding the timing of various events, like whether or not advocacy groups are taking action. This frame implies there is now a short time to take necessary actions because these actions have been neglected for a long time. People and month are among the words from frame four. This frame depicts why some people with COVID had few or no symptoms while others developed life-threatening conditions or passed away, as expressed in the news reports. Frame five contains the words year, family, child,

etc. Frame six contains public, school, and student. To prevent a national disaster amid rising juvenile mental health problems, more needs to be done, particularly as kids become older. From frame seven, we can see the words and symbol %, rate, week, etc., which gives us some idea about disparity statistics in New Jersey. New, home and support are among the words in frame eight. Homeless individuals are more likely to need more care when receiving medical attention, to require more hospitalizations overall, and to stay in hospitals longer. Frame nine provides us an idea about the impacts of the Covid situation on health disparity and policy changes by mentioning the patient and hospital. Frame ten mentions woman and mother, which implies maternal health and maternal mortality.

From Figure 2, we can see the keywords identified by RAKE in New Jersey news coverage. Some keywords are blood pressure, mental health, maternal mortality, health insurance, needle exchange, infant mortality, immune system, colorectal cancer, etc. The Black community is disproportionately affected by colorectal cancer. Because of health care disparities, colorectal cancer represents a terrible disease for numerous families. However, early identification and routine examinations are the best way to prevent it. Moreover, programs for exchanging used needles have been demonstrated to benefit areas where opioid misuse is a problem. These initiatives assist lower drug use and overdose rates while lowering the prevalence of infectious diseases, including HIV and hepatitis. Hence, advocates want to protect the needle exchange program in Atlantic City of New Jersey. Furthermore, due to various medical and economic disparities, black women are, for instance, more likely than white women to die during childbirth. Contrary to white moms, Black and Hispanic mothers in New Jersey have more complications during childbirth, maintaining a long-standing pattern in the state.

Figure 3 shows us the sentiments that are present in New Jersey news coverage. In this figure, we can see eight emotional valances: trust, joy, surprise, anger, sadness, anticipation, fear, and disgust, along with polarity (positive or negative). We see that positive emotion is the most prevalent emotion in the newspaper frames. The good emotion of trust is also frequently seen in news reports. Figure 4 gives us an idea of the top words for each emotion. The most prevalent words in positive emotion are community, public, medical, justice, study, etc. On the other hand, pandemic, abortion, crisis, etc. are among the top words in negative emotions. Because news articles not only present the unfortunate aspect of health disparities but also offer advice and discuss if actions are being taken, positive emotion is frequently present in news reports.

Discussion

This study examined how health disparities are portrayed in New Jersey newspapers to understand better how the issue is portrayed. Despite prior research on health disparities in newspapers, none of it has a state-specific focus. Therefore, this study adds to the field of local health disparities in that regard. Health disparities demand policymakers' attention because they can significantly impact a person's quality of life and cause earlier deaths in certain situations. American policymakers need to understand how the general public perceives health disparities and whether that perception differs from residents of one state to other particular states or regions. Hence, local news media, which serves as a platform for informing the local public about health issues, deserves attention.

We can see from the findings that lack of opportunities in life or chances to lead a better life led to health disparities. If we look at some news articles it becomes prominent. For example, when the pandemic arrived, federal aid was disproportionately distributed to white farmers, bringing discrepancies to a peak [18]. Young Hispanic men in New Jersey are among the Covid pandemic's most severely affected populations [19]. The state medical examiner's office reports that, in 2022, 624 persons had passed away from suspected overdoses. Blacks accounted for 28% of the deaths, with a per-capita overdose rate nearly twice as high as whites [20]. This increase can be linked to lingering disparities in access to healthcare as well as the COVID-19 pandemic's widespread but uneven effects. The coronavirus not only severely affected communities of color but also played a role in the rise in substance misuse generally, fueled by social isolation, economic instability, and fresh societal fears [20].

The news articles prominently bring about racial health disparities. The COVID -19 death toll was far higher for persons of the race [21]. Black Americans have around 3.5 times a higher mortality rate following coronavirus exposure [21]. Moreover, the number of children who need mental health treatment is growing [22]. One of the most ethnically diverse states in the US, New Jersey, has sizable Latino and Black populations. Black and Latino pupils may not be able to cope to the same extent if counselors of those ethnicities are not placed in the schools of those areas [22]. Racial differences were noticeable among homeless people, especially among African Americans. Although they represent 13% of the state's population, they account for 24% of the poor and 49% of the homeless [23]. The notion is that a secure, long-term home is essential for better health. Homeless individuals are more likely to need more care when receiving medical attention, to require more hospitalizations overall, and to stay in hospitals longer. A law allowing hospitals to construct accommodations for homeless people was signed; proponents of the law claim it is a significant step toward better medical care [23].

This study demonstrated the usefulness of news item analysis for tracking the development of local health disparities in recent times. Newspaper illustrations can show the most prevalent concepts and points of view on a subject and to what extent the subject has become critical to the broader public can be ascertained [24]. The value of strategic health communication is demonstrated by several research before [25-27]. Therefore, it is necessary to examine how health disparities are covered in newspapers in New Jersey and other US states so that policymakers can compare the results with those from this study. Because state officials can learn from one another's approaches to problems and use that knowledge to find an effective solution. Hence, this study would stimulate further research in other states or areas that can provide an overview of the US from the perspectives of policy analysis, geographic disparities, and other issues.

This research also brings out the importance of states in policymaking. Policy discussions frequently only cover effective strategies adopted in Washington, DC. However, for a substantial number of Americans, local decisions impact daily life that is at least as great as those made farther away. In crucial ways, the US still adheres to its primary practice of federalism: localities and states are heavily involved in establishing policy [28]. In specific ways, local and state government decision-making is more important than federal government decision-making. In

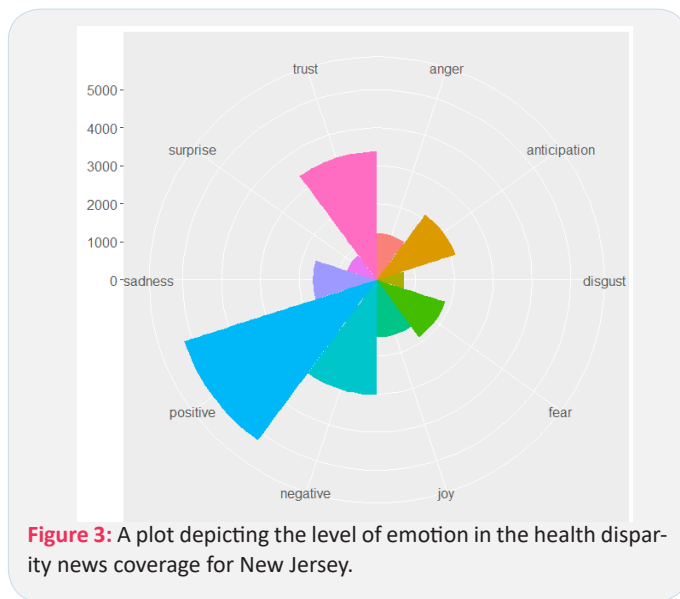
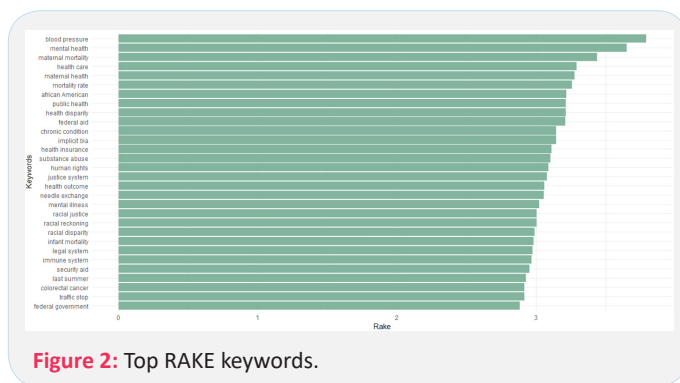
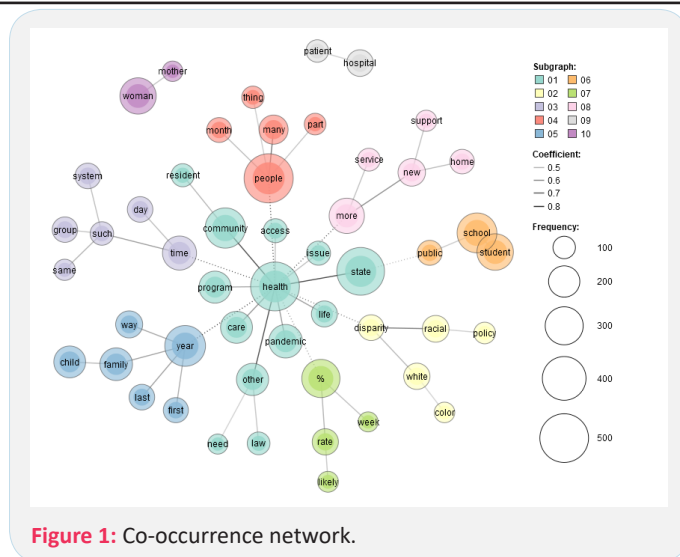
fact, crucial investment choices are made by state and local governments in various areas, including infrastructure, education, and numerous others, that affect how well the entire economic system will perform in the longer term [28]. Therefore, state authorities may be able to provide some remedies to lessen health disparities because health disparities become profound when certain people face restrictions in life prospects.

This research also highlights the significance of policies in governing society and reducing health disparities. People who experience health disparities endure social disadvantages. They frequently exhibit traits distinct from those of the majority due to the many forms of discrimination they frequently encounter. Policies and initiatives intended to improve minority health or reduce health inequities must consider these factors [29]. Authorities and other essential stakeholders frequently want to know whether a specific initiative saves costs. For this situation, a cost-benefit analysis of the initiatives to address health disparities is required, with expenditures and achievements expressed in monetary terms. For example, if we look at Delaware, the state adopted a health disparity intervention program for African Americans. This program costs approximately one million dollars annually and aims to increase colorectal cancer screening rates. The outcome estimated that \$8.5 million in treatment expenses were saved yearly [29]. The most significant impact of this program is the reduction in health disparities resulting from improved health outcomes. This indicates that preventative or interventional programs work well in reducing health disparities. States can employ evidence-based policymaking to create effective laws by creating procedures and instruments that utilize evidence to guide legislation and budgetary choices in all the areas under consideration. In other words, states might use previous policy initiatives that reduced health disparities as steppingstones for developing new policies.

This research also highlights the importance of local newspapers in providing health disparity reports. How will people be made aware of the incidence of health disparities in their state, or how will people know if appropriate actions are taken if those newspapers do not report on how health disparities are occurring in New Jersey? Local news is currently under crisis in the US because of closures of news outlets. The pandemic, which caused such widespread devastation, hurt the sector, resulting in employment losses and the shutdown of local media outlets [30]. Such a situation has stripped neighborhoods of an organization necessary for uncovering wrongdoings and promoting civic involvement, depriving many Americans of a crucial local source of information [31]. The crisis has sparked a widespread turn away from local participatory democracy. The number of candidates running for local office in areas that have lost journalists has decreased as Americans have turned away from local news. This development has also decreased state and local election turnout [31]. Local journalism must thrive and provide investigative reporting regarding health disparities to inform the public about local issues and secure a sustainable future.

Conclusion

Several considerations limit the findings of this investigation. The sampling technique for news articles collection, while thorough and containing important phrases, may have overlooked potentially pertinent news pieces, which is one of the study’s limitations. In addition, a year’s worth of news coverage is only considered. Nevertheless, this study advances our knowledge of the variety and extent of the content that was shared during this duration. Scholars can develop this project’s concept for



forthcoming studies. Our research only examines how health disparities are covered in New Jersey newspapers. It is critical to look into how electronic news sources address health disparities. Moreover, it is crucial to keep up with news from other US states. We can drive the research in a new direction by comparing how New Jersey newspapers cover health disparities with those of other US states and foreign news outlets cover this important topic and using various computational methods. We can find whether the frames will be the same if topic modeling is employed in the study. These topics and methods need to be dealt with in the future to confront this public health risk factor and for the scholarly development of the field of health disparities research.

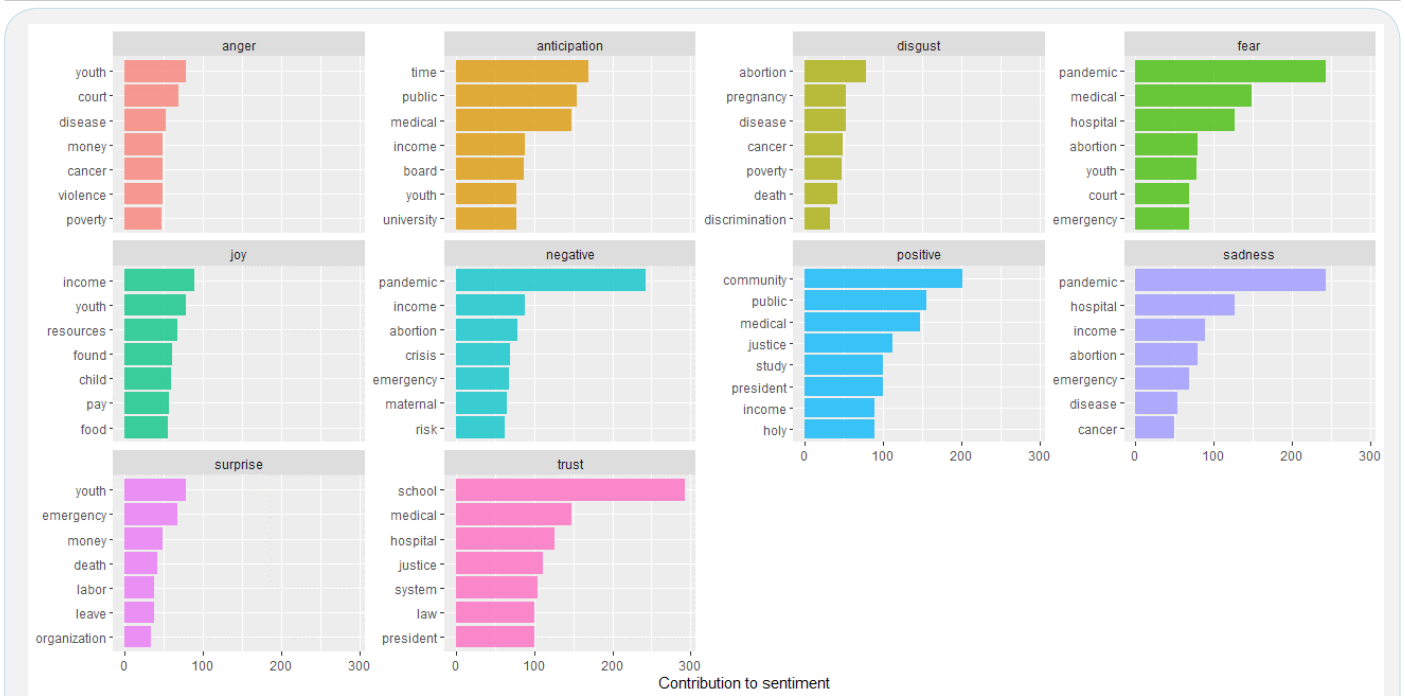


Figure 4: Diagram of each emotional valence.

References

1. CDC. Health Disparities. Centers for Disease Control and Prevention. 2020. <https://www.cdc.gov/healthyyouth/disparities/index.htm>.
2. Radley DC, Baumgartner JC, Collins SR, Zephyrin L, Schneider EC. Achieving racial and ethnic equity in US health care. The Commonwealth Fund. 2021.
3. CDC. Racism and Health. Centers for Disease Control and Prevention. 2021. <https://www.cdc.gov/healthequity/racism-disparities/index.html#:~:text=The%20data%20show%20that%20racial>
4. Rooke C, Amos A. News media representations of electronic cigarettes: an analysis of newspaper coverage in the UK and Scotland. *Tobacco Control*. 2014; 23(6): 507-12.
5. Burnett E, Johnston B, Corlett J, Kearney N. Constructing identities in the media: newspaper coverage analysis of a major UK *Clostridium difficile* outbreak. *Journal of Advanced Nursing*. 2014; 70(7): 1542-52.
6. Al Bawab AQ, AlQahtani F, McElnay J. Health care apps reported in newspapers: Content analysis. *JMIR mHealth and uHealth*. 2018; 6(10): 10237.
7. Afrin R, Mumtahina O, Prybutok G, Harun A, Prybutok VR. Exploring the Coverage of Cyberchondria Addiction in Newspapers: A Perspective from the USA. *J Addict Recovery*. 2023; 6(2): 1043.
8. Hart PS, Chinn S, Soroka S. Politicization and polarization in COVID-19 news coverage. *Science communication*. 2020; 42(5): 679-97.
9. Nimegeer A, Patterson C, Hilton S. Media framing of childhood obesity: A content analysis of UK newspapers from 1996 to 2014. *BMJ open*. 2019; 9(4): 025646.
10. Baciu A, Negussie Y, Geller A, Weinstein JN, National Academies of Sciences, Engineering, and Medicine. The state of health disparities in the United States. In *Communities in action: Pathways to health equity*. National Academies Press (US). 2017.
11. Bell J, Lee MM. Why place and race matter: impacting health through a focus on race and place. *Policylink*. 2011.
12. Alfonseca K, Kekatos M. How 2 New Jersey towns illustrated persistent race gap in the COVID pandemic. (2022, May 19). ABCNews. 2022. <https://abcnews.go.com/Health/jersey-towns-illustrated-persistent-race-gap-covid-pandemic/story?id=84315236>.
13. Healthy NJ. 2020. <https://www.nj.gov/health/chs/hnj2020/topics/access-to-health-services.shtml>.
14. Rasmussen AC. Causes and solutions: Mainstream and black press framing of racial and ethnic health disparities. *Howard Journal of Communications*. 2014; 25(3): 257-80.
15. Rose S, Engel D, Cramer N, Cowley W. Automatic keyword extraction from individual documents. *Text mining: applications and theory*. 2010: 1-20.
16. Younis EM. Sentiment analysis and text mining for social media microblogs using open source tools: an empirical study. *International Journal of Computer Applications*. 2015; 112(5).
17. Bravo-Marquez F, Mendoza M, Poblete B. Meta-level sentiment models for big social data analysis. *Knowledge-based systems*. 2014; 69: 86-99.
18. White N. Revealing discrimination - Farmers of color want more from feds. *Burlington County Times (Willingboro, NJ)*. 2021; 1.
19. Calavia-Robertson D. In N.J., the virus too often has cut short the lives of Hispanic men. *Star-Ledger, The (Newark, NJ)*. 2021; 004.
20. Yates R. Data reveals the changing face of opioid addiction Black residents now more likely to die of overdoses than those who are white, a reversal of the epidemic's traditional toll. *Star-Ledger, The (Newark, NJ)*. 2022; 001.
21. Dearing T. Dementia Defensive COVID study could lead to breakthroughs on Alzheimer's The virus attacks the brain. Solving how it does so might reduce the risk of dementia. *Star-Ledger, The (Newark, NJ)*. 2021; 001.
22. Conlin E. Urgency needed to thwart youth mental health crisis, panelists say Urgency needed to thwart youth mental health crisis, Stockton panelists say. *Press of Atlantic City, The (NJ)*. 2022; 6A.
23. Gammage J. Hospitals can now build homes for homeless NJ hospitals can now build housing for those experiencing home-

-
- lessness, connecting need to services. Press of Atlantic City, The (NJ). 2021; 4.
24. Yamagishi D, Nagai H. Development of a Tax-Free Shopping Environment in Japan: An Analysis of Its Representations in a Financial Newspaper. *Tourism Planning & Development*. 2021: 1-20.
25. Brown DK, Yoo J, Johnson TJ. Spreading Ebola panic: Newspaper and social media coverage of the 2014 Ebola health crisis. *Health communication*. 2018.
26. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. *Journal of public health policy*. 2020; 41(4): 410-20.
27. Afrin R, Harun A, Prybutok G, Prybutok V. Framing of COVID-19 in Newspapers: A Perspective from the US-Mexico Border. *Healthcare* 2022; 10(12): 2362.
28. Nunn R, Parsons J, Shambaugh J. Nine facts about state and local policy. Brookings Institution, January. https://www.brookings.edu/wp-content/uploads/2019/01/StateandLocal_Facts_Web_20190128.pdf (accessed Nov 15, 2023). 2019.
29. Dye BA, Duran DG, Murray DM, Creswell JW, Richard P, Farhat T, Breen N, Engelgau MM. The importance of evaluating health disparities research. *American journal of public health*. 2019;109(1): 34-40.
30. USA Today. A devastating trend: Local newspapers are shrinking or disappearing. Congress must act. 2022. <https://www.usatoday.com/story/opinion/todaysdebate/2022/07/19/local-news-organizations-financial-crisis/10085190002/>.
31. Hendrickson C. Local journalism in crisis: Why America must revive its local newsrooms. Brookings. 2019.