

Framing Feminism: An Examination of Narratives in US Newspapers between 1980 and 2009

DH-412: History and the Digital, EPFL

Ryan Chan, An Bui Duc Khanh, Thomas Radinger

6 June 2023

Abstract

The impact of feminism in our society and the widespread effects of media are both areas of great interest and relevance. We explore how feminism has been framed in newspaper media in the United States from 1980 to 2009 by seeking to operationalize framing through two computational methods, topic modelling (with LDA) and collocation. The results seem to largely validate existing work in this area, though we have found some new insights, such as additional potential frames related to the arts and culture and feminism outside just a “Western” context.

Introduction

Feminism as a project challenging patriarchal power relations through both activism and critical thought (D'Ignazio & Klein, 2020) has undeniably influenced contemporary society, reshaping social norms, attitudes, and policies around gender equality. In the United States, legislative achievements such as the Equal Pay Act (1963) and the Supreme Court decision *Roe v. Wade* (1973) emerged as clear manifestations of these changes. The former addressed systemic gender-based wage discrimination, and the latter asserted women's constitutional right to access abortion (Freedman, 2002). More recently, the emergence of the #MeToo movement has spotlighted the pervasive problem of sexual harassment (Jaffe, 2018).

In the context of these changes, the media's role in public discourse becomes crucial, especially through its practice of 'framing'. Framing is a process through which the media shapes the presentation of events or issues by emphasizing certain aspects, making them more noticeable, meaningful, or memorable to their audiences (Entman, 1993). It underscores the influence of the media in defining problems, interpreting causes, making moral judgements, and suggesting potential remedies (Chong & Druckman, 2007).

The framing of feminism by the media significantly impacts the movement's public perception. By controlling how the public perceives feminism, the media has the power to mobilize public support for feminist causes or also to marginalize and trivialize them. This impact, whether reinforcing or challenging prevalent views, further extends to influence the progression of the feminist project and its success in achieving its goals (van Zoonen, 1992).

Generally the body of research examining the ways in which feminism has been framed in the media in the United States remains small, sparse and qualitative. Previous work typically

focused on a few print media, such as *The New York Times*, and looked at the 1960s to the 1990s (e.g. Ashley & Olson, 1998; Huddy, 1997; Terkildsen & Schnell, 1997). Lind & Salo (2002) represents a notable exception as they analyze the framing of feminism in a different medium, namely television, between 1993 and 1996. Mendes (2012) compares the framing of feminism between British and US newspapers between 1968 and 1982, and extends her analysis by looking also at the year 2008.

However, there is a notable contention around the media's framing of feminism. Some scholars argue that the media adopts a negative perspective when covering feminism, ultimately contributing to public backlash against feminism (Freedman, 2002). And most of the framing analyzed in the cited literature represents feminism as a divisive, radical, or threatening force, fostering misunderstanding and resistance towards the goals of gender equality. This contention underscores the importance of examining the media's framing of feminism, thereby making it a compelling subject for our study.

This paper embarks on an exploratory journey into the framing of feminism in US newspapers from 1980 to 2009, thereby extending and complementing these earlier research efforts. Our investigation seeks to contribute a nuanced understanding of the evolution of media narratives about feminism over this time period, and how these have shaped public discourse around the feminist movement. Specifically, we are interested in assessing whether the frames employed in US newspapers match those identified in the previous literature, whether they changed across individual decades, and whether they differ between selected newspapers.

Related work

The framing of feminism in US news media

Exploring previous research on the ways in which feminism has been framed in the news in the United States serves as a foundation for our investigation. Several studies have deployed varied frames to examine the representation of feminism. Terkildsen & Schnell (1997) utilized five frames: sex roles, feminism, political rights, economic rights, and antifeminism, while Ashley & Olson (1998) employed four frames: importance, illegitimacy, deviance, and events.

Despite not conducting a content analysis, Rhode (1997) suggested that the media have applied five frames: neglect/inattention, demonization, personalization and trivialization, polarization, and blurring the focus, indicating a focus on self- rather than social transformation.

Lind & Salo (2002) took a more synthesized approach, building on this previous research from the 1990s. They proposed six guiding frames, including demonization, personalization and trivialization, a goals frame, a victimization frame, an agency frame, and a site of struggle frame.

Since we mostly rely on the frames put forth by Lind & Salo (2002) to inform our own study, we provide some further information on what their frames refer to:

- The first frame is the "demonization" frame, wherein feminism and feminists are depicted as deviant or as opposing conventional societal norms. This frame essentially presents feminists in a negative light, portraying them as the social 'other'.

- The second frame is "personalization and trivialization", which focuses on superficial aspects such as appearance, style, or personal qualities rather than substantive issues or contributions. This can downplay the importance of feminist activists and their work.
- The third frame is the "goals" frame, which reflects attention to civil rights, reproductive rights, workplace rights, and other feminist goals for equality. This frame offers a platform to highlight the objectives of the feminist movement.
- The fourth frame, "victimization", represents feminists as weak and vulnerable, often emphasizing their struggles and hardships rather than their achievements.
- The fifth frame, the "agency" frame, counters the victimization frame by representing feminists as capable individuals who assert their agency to bring about social change.
- Finally, the sixth frame, "site of struggle", focuses on the different settings where feminist activities take place. This includes diverse locations such as home, school, religion, and work, illustrating the broad scope of feminist engagement in different spheres.

Computational methods and framing

As described in the introduction, we are seeking to operationalize framing, aided by an understanding of framing as a theory, and also how that has been already applied computationally in the field.

Traditionally, framing has been done in two methods: either inductively or deductively (Vreese, 2005). The inductive approach means that frames “emerge” from the material during the analysis, and no frames are defined prior to reading the texts. A deductive approach instead takes pre-existing, defined frames and uses those during the investigation. In addition to these two approaches, there are both generic and issue-specific frames, the first being applicable to a wide variety of topics, “transcending thematic limitations,” whereas the latter is only relevant to specific topics (Vreese, 2005). It is from this understanding that we focus on issue-specific frames of feminism, and decide to operationalize framing both from an inductive and deductive approach, with the computational tools of topic modelling and colocation, respectively.

The use of topic modelling, specifically LDA as an inductive approach to framing has been both praised (DiMaggio et al., 2013; Gilardi et al., 2021) and questioned (Field et al., 2018; Nicholls & Culpepper, 2021). We will describe some of these limitations in the Limitations section, but regardless, we believe that LDA on our corpus can still contribute something new to existing research and provide a further testing ground for the suitability of topic modelling for framing.

Where topic modelling has been employed for framing analysis, it has almost exclusively been done so with an inductive approach, in order to take advantage of the exploratory nature of topic modelling (DiMaggio et al., 2013; Gilardi et al., 2021; Wang & Yin, 2022; Ylä-Anttila et al., 2022). We cite Ylä-Anttila et al.’s (2022) approach: (1) frames are operationalized as a connection between concepts; (2) we use theme-specific data, in this case, feminism; and (3) topics are validated in terms of frame analysis. By not directly using topics as frames, we avoid one of the biggest complaints of topic modelling in framing analysis. In addition, it is argued that if we know our texts are already on a specific topic, LDA outputs are best interpreted as “traces of different ways of discussing a topic – that is, frames” (Ylä-Anttila et al., 2022).

In contrast, collocation would be used as the deductive approach to framing. Lind & Salo (2002) use frames that are taken from the existing literature on the framing of feminism as described in the previous section. From these six frames (demonization, personalization & trivialization, goals of movement, victimization, agency, site of struggle), they then sort the collocates into these frames, before conducting further analysis. Though collocation can be used inductively as well (Jaworska & Krishnamurthy, 2012), we find that it is more suited to a deductive approach in comparison to topic modelling.

Data

Data Collection

Our corpus includes 43,124 articles published between 1980 and 2009 inclusive, taken from the Nexis Uni Database by Lexis Nexis under the category of “Major U.S. Newspapers”.

The Nexis Uni Database’s “Advanced Search” capability allows us to search for keywords found within article text and title. We compiled all the articles under this search, as well as metadata information such as the title, publisher, and date. We narrowed the search based on a few factors:

- The keywords “feminism or feminist or post-feminism or third-wave feminism or second-wave feminism or women's movement or women's rights or sex discrimination or gender equality”
- The date of publication, between 1980 and 2009
- The category of “Major U.S. Newspapers”, for a total of 13 different newspapers, all of which are listed below.

The keywords were chosen to reflect the complex history of the word “feminism” itself which changes over time (Freedman, 2002), also in relation to issues addressed by the movement. This was done qualitatively after surveying the literature, such as *The Routledge Companion to Feminism and Postfeminism* (Gamble, 2001).

We performed further analysis on two specific newspapers, *The New York Times* and *Pittsburgh Post-Gazette*. *The New York Times* takes up a sizable portion of our corpus (39%) meaning it has a disproportionate effect on our results, and thus investigating it further could be helpful. *Pittsburgh Post-Gazette* was chosen since it is one of the more conservative-leaning newspapers, compared to the rest of the corpus which tends to be left of centre (Altschiller, n.d.; *Pittsburgh Post-Gazette*, n.d.)

Table 1: Article counts for each newspaper that is present in our corpus.

Paper	Article Count	Paper	Article Count
The New York Times	16,848	Star Tribune (Minneapolis, MN)	1,920
St. Lous Post-Dispatch (Missouri)	4,077	The Philadelphia Inquirer	1,869
St. Petersburg Times (Florida)	3,866	The Tampa Tribune (Florida)	1,461
Pittsburgh Post-Gazette (Pennsylvania)	3,031	PUBLISHER NOT FOUND	1,311
USA TODAY	2,298	Daily News (New York)	1,174
The Atlanta Journal and Constitution	2,334	Philadelphia Daily News count	730
Christian Science Monitor (Boston, MA)	2,086	Journal of Commerce	122

We also include a distribution of the publishers over time, per year. The decade of 1980-1989 shows a skewed distribution of mostly the New York Times and Christian Science Monitor, while later decades of 1990-1999 and 2000-2009 have more variety, though the New York Times is still predominant.

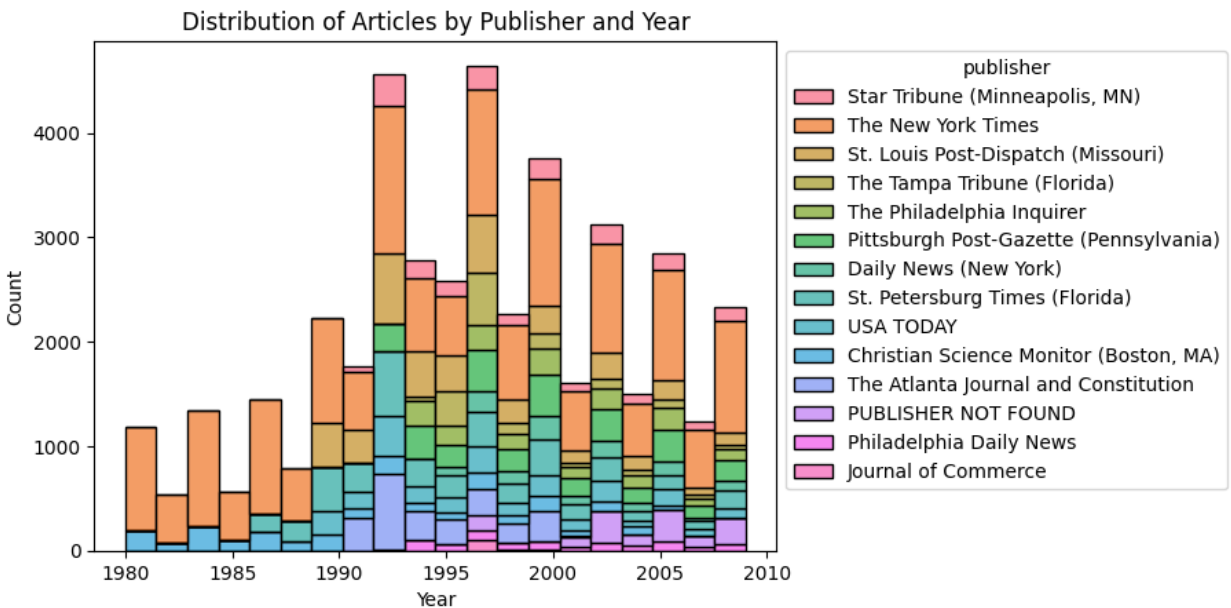


Figure 1: Distribution of articles by publisher and year for our corpus.

Data Pre-Processing

The articles we obtained from Nexis were all Rich Text Files (.rtf), which required further processing before we could continue with our data analysis.

Each .rtf file was converted to a Python string using the striprt library, upon which relevant information was extracted with functions like .strip() and regex matching. The result was turned into a dictionary, and then appended to each other before writing all the articles to a .csv file.

We then did some cleaning on our data. Some examples include:

- 41 articles that had a date earlier than our timeframe, with most being obituaries in the late 2000s. We opted to remove these from our further processing.
- There were 1532 articles where the title and publisher were not found, but we still had the article text. Since our data analysis looks mainly at the text and not the title or publisher, we included these in our processing.
- We also did our analysis by decade, so even though we originally collected 2010 and 2011 information, we removed those in the end.
- There were 4 articles where the year was not found by the regex, so we removed those from the corpus.

With this combined .csv file now complete, we can use the Pandas library to import it as a Pandas DataFrame and employ typical pre-processing techniques with the nltk library. These include removing stopwords, numbers and punctuation; lemmatizing words (only for topic modelling); and finally tokenizing the cleaned text.

Methods

All the code for this project is made publicly available on Github¹.

Topic Modelling

Topic modelling is an unsupervised machine learning technique, typically exploratory, which groups similar words together as a “topic”. We employed the Latent Dirichlet Allocation (LDA) method for our analysis, using Python’s Gensim library to run the model over the cleaned and lemmatized text for each decade in our data (1980-1989, 1990-1999, 2000-2009), as well as articles in New York Times and Pittsburgh Post-Gazette for all three decades.

Our choice of LDA as a model draws upon previous researching outlined in our Related Work section, where it has been employed in similar settings, namely to operationalize framing analysis (DiMaggio et al., 2013; Gilardi et al., 2021; Wang & Yin, 2022; Ylä-Anttila et al., 2022) and to analyze discourse in newspaper corpora (Marjanen et al., 2020; Yang et al., 2011).

¹ <https://github.com/rac99/framing-feminism>

Based on coherence scores and qualitative analysis, we determined that 30 topics provided a good balance between the granularity of topics, its relevance to our research question, and its inner coherence within a generated topic.

The parameters we used for our LDA model were set as: `alpha='auto'`, `eta='auto'`, `passes=10`, `iterations=500`, `random_state=42`, `num_topics=30`. This model was then exported as a visualization with the `pyLDAvis` library. Topic modelling only does the grouping of these words, and so after the topics were formed, we qualitatively determined a name and/or description of each topic. The full list of topics, given names/descriptions, weights of words and the words itself can be found in Appendix A.

Colocation

We employ a colocation technique similar to the one found in Jaworska & Krishnamurthy (2012), but we filter out stopwords (they have kept words like “of” and “is”). We make use of the collocation library from `nlTK`. Their decision to keep these prepositions was to examine the lexico-grammatical patterns, such as “feminism is” and “of feminism”. We are more interested in associated nouns, adjectives and verbs, so we do not include these prepositions.

We filter the top 20 results (as available) based on the likelihood ratio, which “works reasonably well with both large and small text samples and allows direct comparison of the significance of rare and common phenomena” (Dunning, 1993). It is also easily interpretable, as it tells us how much more likely one hypothesis is from another.

The `nlTK` library returns the collocates as a list in the order it was found. This is why the same word can appear twice; for example, [`'feminism'`, `'pacifism'`] and [`'pacifism'`, `'feminism'`] mean that this word shows up both in front of and after our keyword ‘feminism’. A full list of the collocates for each decade, the New York Times (all decades) and the Pittsburgh Post-Gazette (all decades) can be found in Appendix B.

Results

Topic Modelling: Inductively determining frames

First, we performed the LDA, which resulted in a number of topics (see Appendix A). Then, from these topics, we validated which ones were relevant to our study and were coherent, based on the top 30 words. Next, we took these selected topics and placed them into more general groupings and descriptions (what we call a “supergroup”).

Based on past work, our supergroups could already be considered as frames (DiMaggio et al., 2013; Wang & Yin, 2022). Yet, we aim to deepen this by identifying a supergroup that is especially pertinent to ‘feminism’, distinguishing it from the rest of the supergroups. This approach echoes the frame validation process outlined by Ylä-Anttila et al.’s (2022), though it goes beyond this by partitioning the words in the topics into frames, rather than merely using the topics themselves.

As such, we then take the relevant supergroup, and deduce frames from it, similar to how we would with collocates (Lind & Salo, 2002).

Table 2: Supergroups from our LDA models across three decades (1980-1989, 1990-1999, and 2000-2009)

Supergroup	1980-1989 topics	1990-1999 topics	2000-2009 topics
Race and Ethnicity	Topic 0: Race (politics)	Topic 28: Race (education)	
University and College	Topic 8: Academia, intellectualism		
	Topic 26: General school terms (not particular to universities)	Topic 22: University language and names of universities, general	Topic 26: General university words
		Topic 2: College sports	Topic 9: College sports
Law/Judiciary	Topic 17: Judiciary and trials	Topic 5: Judiciary/court (of note: abortion, Ginsburg, discrimination)	Topic 23: Judiciary, (Oconnor, Robert, wade)
	Topic 5: Discrimination in the workplace, lawsuits against sexual discrimination	Topic 15: Discrimination in the workplace, and lawsuits	
		Topic 20: Police events, but neutral, descriptive language	Topic 6: Harassment, rape, police, victims
Goals/Perception of feminism in America	Topic 9: Broadly Women's rights and social issues	Topic 25: Feminism, broadly (woman, sex, power, male, society, culture)	Topic 12: America and feminism - more broadly (includes some notions of race)
Reproductive rights	Topic 22: Abortion (in the courtroom, rally, issues, organizations)	Topic 1: Abortion (more violence than previous decade)	
	Topic 23: Artificial insemination/surrogacy /kind of IVF		

Gender roles	Topic 24: Everyday life and personal relationships	Topic 3: Family words, but also Betty Friedan	Topic 28: Family words
		Topic 23: More familial words but descriptions related to health (birth, doctor, drug)	
		Topic 29: Gender equality in the army	Topic 1: Gender, also terms of equality, science and health, education
Organizational structures		Topic 19: Corporate, advertisements, organizational structure	Topic 27: Corporations, there is one instance of discrimination and lawsuit, but largely lacking apart from general corporate words
National Politics	Topic 11: 1988 Presidential election (dukakis as democratic nominee)	Topic 11: 1992 presidential election	Topic 0: General presidential elections (we see Bush, Gore, Obama, Hillary, Clinton)
	Topic 19: Presidential election (1980)		Topic 5: Rick Santorum, pennsylvania senator, election
		Topic 17: Anita Hill 1991 sexual harassment case	
		Topic 24: Clinton-Lewinsky scandal	
		Topic 14: Shannon Faulkner, first woman allowed to study at The Citadel (military school)	

International Politics	Topic 1: Germany/Germans Topic 2: Cold War	Topic 7: International related words - other nations, and also war	Topic 22: Judaism, but also pakistan, Turkey, taliban, afghanistan, and russian
Arts & culture	Topic 13: America/books/american publishers Topic 15: books (genre of books) Topic 7: Literature, Food and Art	Topic 10: Books/genres	Topic 4: Books and writing
	Topic 20: Film and movies with some reference to gay/homosexual	Topic 27: Film	Topic 2: Films
	Topic 25: Visual art	Topic 12: Visual arts and places of exhibition	Topic 18: Visual arts Topic 17: Art and cooking Topic 3: International art and women artists
	Topic 28: Theatres/plays	Topic 6: Plays/theatre	Topic 16: Plays and theatre Topic 25: Broadway, shows, new york
	Topic 18: Music	Topic 16: Music Topic 18: Pop culture	Topic 24: Music
	Topic 14: Magazines/adult magazines and advertisements		Topic 7: General reading, magazine, editor, published
	Topic 21: Television		Topic 8: TV- we see some new words that are LGBTQ related
	Topic 26: Fashion		

	Topic 16: Philosophy, Psychology and Social theory		
Religious discourse	Topic 27: Roman Catholicism	Topic 4: Religion more generally (Christianity, Judaism, homosexuality)	Topic 13: Religion, but now including Islam, terrorism, Christianity, Catholic
	Topic 6: Jewish/Judaism		

Table 2 introduces many interesting results that could be explored further as an extension, which is described in our Interpretation & Discussion section. In particular, we will explore the supergroup of “arts & culture”, as it isn’t something that has come up frequently in our secondary literature review.

As explained earlier, we want to take our operationalization of framing one step further to compare it to how framing is usually done with LDA, so we take the supergroup of “Goals/perception of feminism in America” to put into frames, given that this inductively generated topic relates most closely to feminism among the others.

Table 3: Frames across three decades (1980-1989, 1990-1999, 2000-2009)

Frame	1980-1989 top words (topic 9)	1990-1999 top words (topic 25)	2000-2009 top words (topic 12)
Sex, gender and roles	Woman, child, male, female, people, care, equal, role, sex, support, family	Woman, people, sex, male, relationship, female, human, sexual	People, woman, human, leader, family
Career	Job, career, business, organization		
Geography and location	American, national, country, york		American, country, united, america, nation, society, national, south, international
Subjectivity and abstractness		Sense, question, idea, real, feel, mind, true, view, personal, love	
Power		power	power
Society and social issues	Social, change, society, issue, life	society, culture, life	Social, culture, public, community, issue, life, black

Politics and government			War, political, history, government, right
Feminism as a movement	Feminist, movement	Feminist, feminism	Movement, feminist
Other	Percent, time, Study, dr	Time, graphic, Book, Story, image	

Now that we have a table of supergroups (Table 2) and a table of frames (Table 3), our approach should lead to frames that “somewhat correspond to previously identified frames in similar data to ensure validity” (Ylä-Anttila et al., 2022). We have already explained that our supergroups can also be considered frames based on previous research, and so we will look for this “correspondence” that is described by Ylä-Anttila et al. in both our supergroups and our subsequent frames to ensure their validity.

From this, we see that there is considerable overlap between our supergroups, inductively-generated frames, and existing frames in the secondary literature. As a reminder, Lind & Salo (2002) proposed six guiding frames: demonization, personalization and trivialization, a goals frame, a victimization frame, an agency frame, and a site of struggle frame.

Taking for example the frames compiled from Lind & Salo (2002), all the subframes of the “site of struggle” frame (home, education, religion, etc.) that they proposed can be found in our supergroups (law and judiciary, university and college setting, religious discourse, etc.). Some of the frames suggested by Terkildsen & Schnell in 1997 are also similar- they propose “sex roles”, we have a “sex, gender and roles” frame, they propose “political rights”, which we see in the “politics and government” and “society and social issues” frames. We can then validate that our frames are reasonable, as they do not differ wildly from the existing literature.

Colocation: Deductively determining frames

Our results on searching for the top-20 collocates to the word ‘feminism’, along with their likelihood ratio scores can be found in Appendix B.

Given that this is a deductive approach, we purposefully employed frames that were used in the literature, particularly from Lind & Salo (2002). However, there were two groupings that we thought did not fit into these predefined frames, and so we categorized them as “other movements or ideologies” and “theory/classification/history of feminism”.

Table 4: Top-20 Colocations of the word ‘feminism’ sorted into frames across three decades (1980-1989, 1990-1999, 2000-2009)

Frames	1980-1989	1990-1999	2000-2009
Demonization	Radical, militant, byproduct	Radical, stole, militant, fear, illusions	Radical, misguided

Personalization & Trivialization	femininity, brand	brand	Brand, intense
Victimization		Backlash, dead	Harming, dead, triggers
Agency	Modern, rise, oped	Modern, contemporary, rise, professing, alive	Modern, rise, rights
Site of struggle		Campus, american	capital
Other movements/ide ologies	socialism, marxism, ecology, pacifism, anticolonialism, humanism, mormonism	homosexuality	gay, homosexuality, thatcher era, marxism
Theory/Classific ation/History of feminism	wave	Wave, secondwave,	Wave, secondwave, firstwave, movement

To gain a better understanding of the commonalities between these collocates as they show up in each decade, we plot collocates that show up in more than one decade below with a bump plot. Since we found our top-20 collocates based on the likelihood ratio, we were able to see how common each collocate was relative to each other, but that was within a singular decade. Trying to compare that across different training sets may not be an accurate representation of their values, and so we decided to go with a simple ranking of the top-20 collocates. Although we won't be able to provide a specific quantifiable metric, we do believe that this ranking graph gives us some way of comparing between decades without being faulty in its representation of certain values. It is also important to note that if the ordering of the collocates is different, such as ('feminism', 'homosexuality') and ('homosexuality', 'feminism'), they won't count as the same and won't be on the graph.

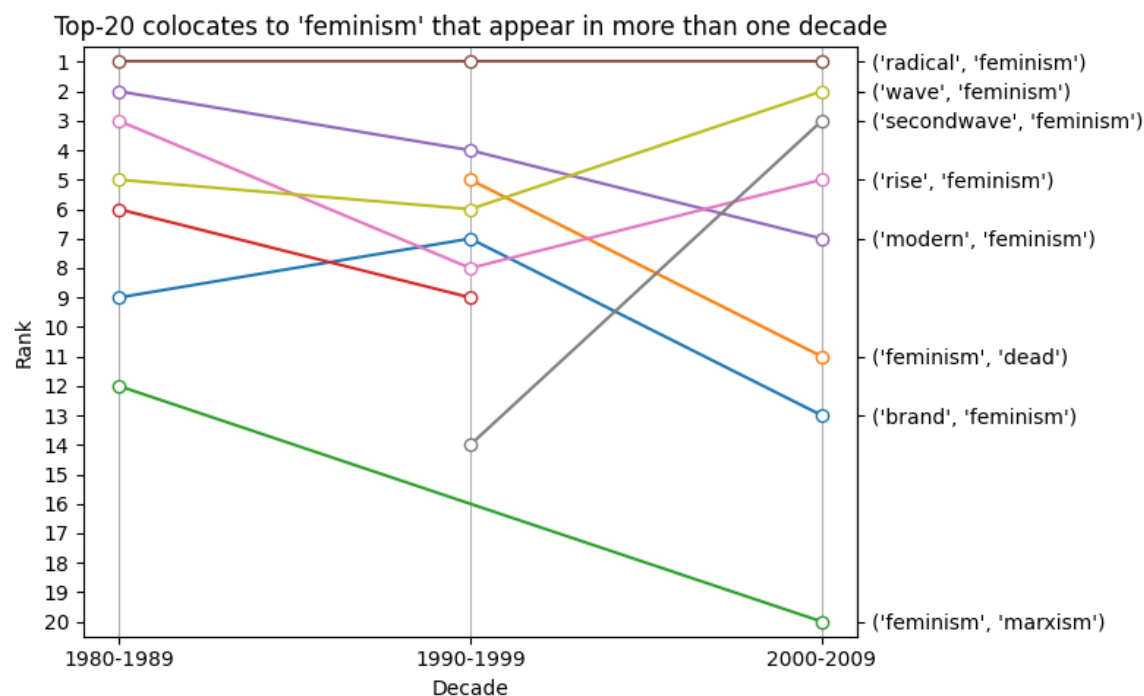


Figure 2: Top-20 collocates to 'feminism' that appear in more than one decade

On the other hand, we also wanted to show the collocates that did not show up more than once in the top-20 collocates per decade, in the same ordering of the pairs. It is possible that they occur in other decades, but their frequency compared to the others was not sufficient to have them appear in the top-20. Here they are listed in order of frequency within each decade.

Table 5: Collocates to the word 'feminism' that only appear in the top-20 in one specific decade

Decade	Collocates
1980-1989	femininity, socialism, ecology, oped, pacifism, anticolonialism, humanism, byproduct, failure, mormonism
1990-1999	stole, backlash, contemporary, fear, illusions, campus, professing, american, alive, word, homosexuality
2000-2009	harming, firstwave, misguided, movement, gay, homosexuality, capital, rights, triggers, thatcher era, mids, intense

Colocation for New York Times and Pittsburgh Post-Gazette

One of the sub questions of interest is how frames might differ across different newspapers. As discussed in the Data Collection section, the two papers we selected were The New York Times and Pittsburgh Post-Gazette. Similar to what was done for the collocates in each decade, we imposed the frames we took from Lind & Salo (2002) upon our collocates,

though there were three other frames that we came up with, two of which were similar to our decadal collocates (Other movements/Ideologies and Theory/Classification/History of feminism), and one new one that we labelled “Geography”. This one only had results for the Pittsburgh Post-Gazette, whereas there were no relevant collocates for this frame in The New York Times.

Table 6: Top-20 Colocations of the word ‘feminism’ sorted into frames across two newspapers, The New York Times and Pittsburgh Post-Gazette

Frames	The New York Times	Pittsburgh Post-Gazette
Demonization	Radical, stole, militant, fear	Radical, stole, misguided
Trivialization and personalization	Femininity, brand	Virginity, brand
Agency	Modern, contemporary, decades, rise	Rise, grown, version, encourages
Victimization	Backlash, triggers	Harming, blow,
Geography		Western, turkey, globalization
Other movements or ideologies	Gay, pacifism, marxism	Islamic, homosexuality, progressive
Theory/Classification/history of feminism	Secondwave, wave, firstwave, movement	Wave

Interpretation & Discussion

Negative framing of feminism: Overlaps with the existing literature

In the secondary literature, the frames proposed by Lind & Salo (2002) are already overwhelmingly negative ones, such as demonization, personalization and trivialization, and victimization. These frames were all clearly represented in our results from the deductive approach with collocates, both in the decadal and newspaper-specific analysis. For instance, the word “radical” was ranked as the top collocate for feminism across all three decades, and is a strongly negative term in its context; we also see words such as “dead”, “failure” and “fear”, which aligns with the negative framing of feminism as a movement.

In our inductive approach, though the supergroups and frames do not definitively invoke negative sentiment, we still see a large amount of overlap with Lind and Salo’s frame of “sites of struggle” represented in our results. All sites of struggle are covered, whether in the judiciary and law (sexual discrimination lawsuits, abortion rulings), universities (college sports with Title

IX as words in those topics), gender roles (descriptions of roles in family, gender quality in the army), and more.

Thus, we show that there is significant overlap with the frames that exist in literature: one, that the negative frames of feminism is seen in our deductive, colocation approach, and two, that the not necessarily negative frames such as the “sites of struggle” are evident in our supergroups and frames from an inductive approach.

Feminism framed in a broader context: Interconnections with other social issues

An interesting result of our approach and the computation of collocates to identify the frames deployed in the news media is the visibility of discussions of feminism in relation to other social issues and movements. In the 1980s, specifically feminism seems to be connected to socialism, ecology, pacifism, anticolonialism and humanism. This may also illustrate the importance of feminism not just as a standalone topic, but as an integral network of related issues and movements. Interestingly, this then shifts to linkages to the gay rights movement in the 1990s and 2000s.

In Figure 3, we can see that this particular individual, the “young ideologue”, supports a few movements, namely both ecology and feminism here. So we see that in this case, its relationship is due to the people who support it, perhaps indicating that people may not be exclusively involved in feminism as a movement, and may participate in other movements and draw upon them.

He stumps not for proletarian revolution, but for ecology, *feminism*, and "individual freedoms." The young ideologue, who admits "I would have killed the ambassador" to spring his comrades in arms, today proclaims: "I don't believe in politics. I believe in people."

Figure 3: Excerpt from article titled: “Brazilian terrorist trades his gun for novelist's pen”. Christian Science Monitor (Boston, MA). Published August 29, 1983. Full article on Github repository.

Figure 4 below shows an example of the discussion between homosexuality and feminism. The article associates feminism (the “feminist message”) with homosexuality (the “contemporary issues and trends”), and this is seen as a positive relationship (“...evolved into a thoroughly modern organization.”) At least in this example, there is a melding of these two movements as almost one and the same, and conveyed as such to the reader.

But the Girl Scouts have also steered clear of the commotion surrounding sexual orientation and religion that has dogged the Boy Scouts of America in recent years. The Boy Scouts' position on homosexuality -- that gay members are not welcome -- has cost that organization financing and members. And troop leaders say the Girl Scouts, traditionally known for the three C's of crafts, cookies and camping, have evolved into a thoroughly modern organization. It has formed an implicitly **feminist** message by plugging into contemporary issues and trends, while maintaining its traditional values.

Figure 4: Excerpt from article titled: "Beyond Crafts and Cookies, Girl Scouts Are Prospering". The New York Times. Published December 25, 2002. Full article on Github repository.

Feminism framed in terms of the metaphorical "wave"

The concept of "waves" has been a popular metaphor used to frame the history of feminism in the United States, in both scholarly work and wider society (Hewitt, 2012; Reger, 2017). Interestingly, our research also observed the use of terms related to this "wave" metaphor, and its prevalence in the collocates seemed to increase over the years. Moreover, it appears that the later decades started referring back to the earlier "waves". For instance, in the 1980-1989 period, the term "wave" was observed, while the 1990-1999 period showed both "wave" and "second wave". Then in 2000-2009, "first wave" made its appearance, in addition to "wave" and "second wave". This common usage of "wave" in conjunction with "feminism" is not necessarily a validation of the wave theory, but rather indicates that this metaphor is embedded in public discourse beyond the confines of academic theory.

Feminism framed beyond "Western" boundaries: Locating feminism in an international context

One of the interpretations from the literature noted that "feminism is firmly located in Western countries, and there are virtually no references to any other places in the world" (Jaworska & Krishnamurthy, 2012). However, our results seem to suggest otherwise.

Our inductive approach led us to a potential "geography and location" frame, which points towards an international perspective, particularly with the word "international" appearing in the 2000-2009 decade. The international politics supergroup is consistent through each decade of analysis, with notable references to non-Western countries such as "Japan", "China", "Turkey", "Israel", "Pakistan", "Afghanistan", among many others.

On the other hand, our deductive approach when analyzing The New York Times and Pittsburgh Post-Gazette was also interesting, as it necessitated the creation of a new frame of "Geography". This frame is constituted by the words "Western", "Turkey" and "globalization," which were in the top-20 collocates. Although our primary focus was on the United States, these findings suggest another potential area for deeper exploration: the global portrayal of feminism. It implies that the depiction of feminism is not confined within national boundaries; it extends beyond, providing an international context

An example of this is in the article snippet found in Figure 5, which describes feminism in the Philippines, as part of a series that the newspaper is running. There is this distancing from "Western feminism", which was reported on in a US newspaper in 1988.

"We're trying to repair centuries of damage," Sister Soledad Perpinan, a **feminist** nun, says today. "We are fighting for our rights, but we are not looking at the West for women's liberation. We were equal long before they even thought about it."

In the Philippines, because a majority of women fall under the poverty line and live in the villages, **feminism** is seen not as a battle of the sexes but as a fight for survival.

Figure 5: Excerpt from article titled: "Feminism as survival // Filipino women try to regain equality". St. Petersburg Times (Florida). Published November 14, 1988. Full article on Github repository.

Deploying similar frames: The New York Times and Pittsburgh Post-Gazette

When comparing the framing of feminism between The New York Times and the Pittsburgh Post-Gazette, the commonalities outweigh the differences. Common frames that emerge from the collocates analysis that emerge in both newspapers reflect the overall frames that we find in our analyses and that were found in the existing literature, and include demonization, trivialization and personalization, and victimization, but also agency.

Differences can be found in relation to other movements or ideologies and the linkage of feminism to other geographies in the Pittsburgh Post-Gazette. Whether these differences stem from differences in political leaning of the two newspapers or are due to other factors, such as local specificities or editorial bias cannot be determined.

Figure 6 and 7 shows the use of the word "radical" in a negative manner in both papers, as associated with the riot grrrl movement, an important part of the history of feminism in the United States.

Figure 6, which comes from the Pittsburgh Post-Gazette, compares the more "sanitized", "new kind" of "diverse" feminism that is positive, compared to the more negative "radical yet short-lived Riot Grrl movement". Figure 7, from The New York Times, describes them as the "teenage, radical, feminist milieu". So we see that their description of the same group is with the same frame.

Ironically, Girl Power is a sanitized version of the radical yet short-lived Riot Grrrl movement, '90s daughter of punk. Riot Grrrl inspired women to take up guitars, to pile-drive their anger into music. It had an industrious fanzine network reflecting a new kind of diverse **feminism**.

Figure 6: Excerpt from article titled: "FEMINISM ROCK; FEMALE SINGERS BRING A MOVEMENT ONCE LED BY ACADEMICS TO A MASS AUDIENCE". Pittsburgh Post-Gazette. Published January 2, 1999. Full article on Github repository.

The story goes back to the teenage radical **feminist** milieu that called itself riot grrrl, an explosion of bands, fanzines and cross-country committees of correspondence, a milieu with an especially strong presence in Olympia.

Figure 7: Excerpt from article titled: "Raising the Stakes In Punk Rock". The New York Times. Published June 18, 2000. Full article on Github repository.

Beyond anticipated frames: the “Arts and Culture”

We also identified other frames that we had not initially anticipated, and that have not been referred to in the previous literature. Notably, “arts and culture” from our inductive analysis. The appearance of “arts and culture” signifies the important influence of feminism in the arts and possibly also the role of feminist theory and intellectuals more generally. Since artistic expression may constitute an arena for challenging traditional gender norms and inequalities, “arts and culture” could also be interpreted as a subframe to the frame of sites of struggle.

Figure 8 illustrates this notion of the role of art and culture in performance art, as a “social lens” through which to view things, and feminism is put up alongside health crises like AIDS and some more mundane topics like government financing. Much more could be explored on the relationship and influence of feminism in the arts and culture sphere.

By looking at the last 20 years, the public can see that performance art didn't operate in a vacuum. "It tells the story of AIDS, *feminism* and Government financing," she said. "It's a social lens on the value of performance art in politics and economics."

Figure 8: Excerpt from article titled: “For an Avant-Garde Center, An End and a Beginning”. The New York Times. Published September 1, 1997. Full article on Github repository.

Limitations & Improvements

Data and Corpora

Based on the way we collected our data, we make a similar remark as Lind & Salo in 2002: “Our corpus could be characterized as a convenience sample; it is not a random sample, and generalizations to other texts should be made with caution, if at all.” Some limits include that a large portion (39%) of the corpus was from one particular newspaper, The New York Times. Moreover, all-but-one of the newspapers in our corpus was left-centre, which can skew our results. In addition, the Nexis Uni Database only provided access to newspapers from 1980 onwards, which meant we could not explore further back in time.

Given that there have been many important events and changes with regards to feminism in earlier decades, such as the 1960s and 1970s, a broader range of dates could provide more insight that we could not go into with our analysis. Nevertheless, compared to the existing literature, we could extend our research up to 2009. Another improvement to expanding the corpus would be to get a better balance of sources from different political leanings.

Methods

An important aspect of framing analysis is also about what is excluded and not just what is included, so ideally looking for the absence of key terms or topics would be of great interest as well. Methodologically, we pre-filtered our corpus based on specific keywords, and so we do not have a “reference corpus” that may be more general to compare our results with. For

instance, how might collocates such as “modern” and “rise” be used with other words in the general newspaper media? This could help us with better determining associated sentiment to the word, and bring more clarity in our choice of frame. Further exploration into these associated words would give greater contextualization of the time period and to our research question.

There are also certain limitations when using topic modelling to operationalize framing. For instance, Nicholls & Culpepper (2021) propose structural topic modelling (STM) rather than LDA, and they found that it was effective, but only with a narrow scope of articles. From their analysis, something that is narrower is able to “extract subtler variations based more closely on tone, emphasis, and repertoire”, whereas a broad dataset that they also used was more on the different subjects being discussed in that newspaper. This can be seen in our analysis: although we were able to find further potential interesting frames such as arts & culture, the results remain relatively broad compared to the collocates, which tend to be more specific and thus provide a stronger justification for the framing employed.

As an extension, there have been other computational methods used to operationalize framing, such as neural networks (Liu et al., 2019) and network analysis (Walter & Ophir, 2019), which both suggest that they are better than topic models to detect frames. This could be explored further, to see how the results may differ with these computational tools.

Conclusion

This paper aimed to understand the framing of feminism in US newspapers from 1980 to 2009 using computational methods. Qualitative insights from the reading of selected articles helped contextualise some of the findings. The results align well with previous literature, and attest to a continuous negative framing beyond the years covered in these previous articles. Frames such as demonization and victimization are present throughout the period we examined. Other frames also show continuity (e.g. sites of struggle for feminism).

Our project, which also included an inductive approach to framing, however also revealed potentially new frames, demonstrating that the discourse around feminism is interrelated with other social issues and ideologies. Over time, feminism was notably linked with socialism, ecology, pacifism, anticolonialism, humanism and the gay rights movements.

This also includes other less-explored aspects in the coverage of feminism with the arts and culture as a site of struggle, or the extension of feminism beyond the US alone. The absence of these narratives from previous research cannot only be explained through our extension of the time period covered since these potential frames were also present prior to 2000. Still, it is important to note that these findings would merit further exploration to validate these narratives as truly new frames or subframes rather than simply the reflection of current events and issues.

This ties in with some methodological insights from our study about the potential and limitations of topic modelling and collocates for the identification of media frames. While it seems possible to discover new frames inductively, framing also includes other dimensions

such as linguistic style, so other methods are clearly needed to fully validate and capture framing, and validate our insights.

Finally, while our corpus was limited in its temporal range and bias towards The New York Times and left-leaning newspapers in general, we could still make a comparison in the framing of feminism between newspapers of different political leanings, something that had not been done previously. Interestingly, the framing seems generally consistent across both papers. Differences are small and could potentially be explained due to different local contexts.

In conclusion, our computational analysis has advanced our understanding of the framing of feminism in US newspapers, confirming known trends, but also revealing new potential frames that merit further exploration. Our insights highlight the value in scrutinizing media representations of issues and events which can be framed in certain ways.

References

- Altschiller, D. (n.d.). *Research: WR150: Educated Electorate: Newspapers - which way do they lean?* Retrieved June 6, 2023, from <https://library.bu.edu/blumenthal/bias>
- Ashley, L., & Olson, B. (1998). Constructing Reality: Print Media's Framing of the Women's Movement, 1966 to 1986. *Journalism & Mass Communication Quarterly*, 75(2), 263–277. <https://doi.org/10.1177/107769909807500203>
- Chong, D., & Druckman, J. N. (2007). Framing Theory. *Annual Review of Political Science*, 10(1), 103–126. <https://doi.org/10.1146/annurev.polisci.10.072805.103054>
- D'Ignazio, C., & Klein, L. (2020). *Data Feminism*. The MIT Press.
- DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, 41(6), 570–606. <https://doi.org/10.1016/j.poetic.2013.08.004>
- Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, 19(1), 61–74.
- Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Field, A., Kliger, D., Wintner, S., Pan, J., Jurafsky, D., & Tsvetkov, Y. (2018). Framing and Agenda-setting in Russian News: A Computational Analysis of Intricate Political Strategies. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3570–3580. <https://doi.org/10.18653/v1/D18-1393>
- Freedman, E. (2002). *No Turning Back: The History of Feminism and the Future of Women*. Ballantine Books.
- Gamble, S. (Ed.). (2001). *The Routledge Companion to Feminism and Postfeminism*. Routledge.

- Gilardi, F., Shipan, C. R., & Wüest, B. (2021). Policy Diffusion: The Issue-Definition Stage. *American Journal of Political Science*, 65(1), 21–35. <https://doi.org/10.1111/ajps.12521>
- Hewitt, N. A. (2012). Feminist Frequencies: Regenerating the Wave Metaphor. *Feminist Studies*, 38(3), 658–680.
- Huddy, L. (1997). Feminists and feminism in the news. In *Women, Media and Politics—Paperback—Pippa Norris—Oxford University Press*.
<https://global.oup.com/ushe/product/women-media-and-politics-9780195105674>
- Jaffe, S. (2018). The Collective Power of #MeToo. *Dissent*, 65(2), 80–87.
<https://doi.org/10.1353/dss.2018.0031>
- Jaworska, S., & Krishnamurthy, R. (2012). On the F word: A corpus-based analysis of the media representation of feminism in British and German press discourse, 1990–2009. *Discourse & Society*, 23(4), 401–431. <https://doi.org/10.1177/0957926512441113>
- Lind, R. A., & Salo, C. (2002). The Framing of Feminists and Feminism in News and Public Affairs Programs in U.S. Electronic Media. *Journal of Communication*, 52(1), 211–228.
<https://doi.org/10.1111/j.1460-2466.2002.tb02540.x>
- Liu, S., Guo, L., Mays, K., Betke, M., & Wijaya, D. T. (2019). Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence. *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, 504–514. <https://doi.org/10.18653/v1/K19-1047>
- Marjanen, J., Zosa, E., Hengchen, S., Pivovarova, L., & Tolonen, M. (2020). *Topic modelling discourse dynamics in historical newspapers* (arXiv:2011.10428). arXiv.
<http://arxiv.org/abs/2011.10428>
- Mendes, K. (2012). ‘Feminism rules! Now, where’s my swimsuit?’ Re-evaluating feminist discourse in print media 1968–2008. *Media, Culture & Society*, 34(5), 554–570.
<https://doi.org/10.1177/0163443712442701>

Nicholls, T., & Culpepper, P. D. (2021). Computational Identification of Media Frames: Strengths, Weaknesses, and Opportunities. *Political Communication*, 38(1–2), 159–181.

<https://doi.org/10.1080/10584609.2020.1812777>

Pittsburgh Post-Gazette. (n.d.). [Web page]. Library of Congress, Washington, D.C. 20540 USA.

Retrieved June 5, 2023, from <https://www.loc.gov/item/lcwaN0011431/>

Reger, J. (2017). Finding a Place in History: The Discursive Legacy of the Wave Metaphor and Contemporary Feminism. *Feminist Studies*, 43(1), 193–221.

<https://doi.org/10.15767/feministstudies.43.1.0193>

Rhode, D. L. (1997). Media Images/Feminist Issues. In *Feminism, Media, and the Law*. Oxford University Press.

Terkildsen, N., & Schnell, F. (1997). How Media Frames Move Public Opinion: An Analysis of the Women's Movement. *Political Research Quarterly*, 50(4), 879–900.

<https://doi.org/10.2307/448991>

van Zoonen, E. A. (1992). The Women's Movement and the Media: Constructing a Public Identity. *European Journal of Communication*, 7(4), 453–476.

<https://doi.org/10.1177/0267323192007004002>

Vreese, C. H. (2005). News framing: Theory and typology. *Information Design Journal*, 13(1), 51–62. <https://doi.org/10.1075/idjdd.13.1.06vre>

Walter, D., & Ophir, Y. (2019). News Frame Analysis: An Inductive Mixed-method Computational Approach. *Communication Methods and Measures*, 13(4), 248–266.

<https://doi.org/10.1080/19312458.2019.1639145>

Wang, C., & Yin, H. (2022). A Framing Analysis and Regional Comparison of Newspaper Media Reports of COVID-19 in Shanghai: Based on the LDA Topic Model. *Studies in Media and Communication*, 10(2), 312–319.

Yang, T.-I., Torget, A., & Mihalcea, R. (2011). Topic Modeling on Historical Newspapers.

Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, 96–104. <https://aclanthology.org/W11-1513>

Ylä-Anttila, T., Eranti, V., & Kukkonen, A. (2022). Topic modeling for frame analysis: A study of media debates on climate change in India and USA. *Global Media and Communication*, 18(1), 91–112. <https://doi.org/10.1177/17427665211023984>

Appendix A

LDA Model Topics

A.1: 1980-1989 LDA Model Topics

Topic #	Topic Name	Word weightings (top 30 words)
0	Race	0.029*"black" + 0.022*"mondale" + 0.013*"white" + 0.012*"mayor" + 0.009*"ferraro" + 0.009*"campaign" + 0.009*"city" + 0.008*"hart" + 0.007*"jackson" + 0.007*"glenn" + 0.007*"york" + 0.006*"san" + 0.006*"koch" + 0.006*"south" + 0.006*"governor" + 0.005*"race" + 0.004*"district" + 0.004*"john" + 0.004*"francisco" + 0.004*"democratic" + 0.004*"president" + 0.004*"mondales" + 0.004*"candidate" + 0.004*"washington" + 0.004*"primary" + 0.004*"aide" + 0.003*"news" + 0.003*"percent" + 0.003*"gary" + 0.003*"week"
1	Germany/Germans	0.045*"german" + 0.039*"west" + 0.030*"germany" + 0.029*"berlin" + 0.020*"cheese" + 0.008*"cheryl" + 0.007*"lieberman" + 0.006*"sheedy" + 0.006*"baxter" + 0.006*"mail" + 0.006*"labour" + 0.005*"rockwell" + 0.005*"bristol" + 0.005*"frankfurt" + 0.004*"east" + 0.004*"abernathy" + 0.004*"roseland" + 0.004*"arbuss" + 0.004*"keohane" + 0.004*"hitler" + 0.003*"arbus" + 0.003*"deutsche" + 0.003*"arboretum" + 0.003*"jihan" + 0.003*"dickstein" + 0.003*"pierson" + 0.003*"columnist" + 0.003*"bonn" + 0.003*"moat" + 0.002*"deterrence"
2	Cold War	0.008*"government" + 0.007*"american" + 0.007*"country" + 0.007*"people" + 0.007*"united" + 0.006*"union" + 0.005*"soviet" + 0.004*"official" + 0.004*"police" + 0.004*"city" + 0.004*"nation" + 0.004*"national" + 0.004*"leader" + 0.004*"day" + 0.004*"military" + 0.003*"party" + 0.003*"political" + 0.003*"war" + 0.003*"international" + 0.003*"minister" + 0.003*"time" + 0.003*"labor" + 0.003*"president" + 0.003*"west" + 0.003*"america" + 0.003*"worker" + 0.003*"month" + 0.002*"sport" + 0.002*"foreign" + 0.002*"nuclear"
3	Italy/Italians	0.062*"italian" + 0.026*"italy" + 0.025*"moore" + 0.014*"tyler" + 0.013*"edinburgh" + 0.012*"annie" + 0.012*"levy" + 0.010*"rame" + 0.009*"mary" + 0.008*"aclu" + 0.008*"polly" + 0.007*"alicia" + 0.007*"dario" + 0.007*"italianamerican" + 0.005*"alien" + 0.005*"lorin" + 0.005*"milan" + 0.005*"lurie" + 0.005*"jones" + 0.005*"franca" + 0.004*"immigration" + 0.004*"buffo" + 0.004*"visa" + 0.004*"mistero" + 0.004*"fo" + 0.003*"casa" + 0.003*"tanzi" + 0.003*"parson" + 0.003*"undocumented" + 0.003*"mtm"
4	Urban life	0.020*"street" + 0.011*"city" + 0.010*"sunday" + 0.010*"york" + 0.008*"park" + 0.008*"avenue" + 0.008*"saturday" + 0.008*"center" + 0.006*"west" + 0.006*"east" + 0.006*"tomorrow" + 0.005*"day" + 0.005*"hall" + 0.005*"music" + 0.005*"house" + 0.005*"ticket" + 0.005*"admission" + 0.005*"monday" + 0.004*"weekend" + 0.004*"art" + 0.004*"village" + 0.004*"island" + 0.004*"event" + 0.004*"program" + 0.004*"fashion" + 0.004*"concert" + 0.004*"shop" + 0.004*"friday" + 0.004*"jazz" + 0.004*"tour"
5	Discrimination in the workplace, lawsuits against sexual discrimination	0.022*"discrimination" + 0.017*"federal" + 0.015*"department" + 0.014*"sex" + 0.013*"job" + 0.012*"city" + 0.010*"civil" + 0.009*"company" + 0.009*"right" + 0.009*"commission" + 0.009*"program" + 0.009*"pay" + 0.008*"employee" + 0.007*"education" + 0.007*"insurance" + 0.006*"woman" + 0.006*"law" + 0.006*"government" + 0.006*"suit" + 0.006*"college" + 0.006*"title" + 0.006*"aid" + 0.005*"employment" + 0.005*"action" + 0.005*"filed" + 0.005*"office" + 0.005*"official" + 0.005*"worker" + 0.005*"administration" + 0.005*"sexual"
6	Jewish/Judaism	0.059*"jewish" + 0.032*"jew" + 0.019*"rabbi" + 0.017*"craft" + 0.016*"harris" + 0.015*"israel" + 0.010*"israeli" + 0.009*"station" + 0.009*"judaism" + 0.008*"duncan" + 0.008*"orthodox" + 0.007*"christine" + 0.007*"anchor" + 0.006*"carpenter" + 0.006*"jury" +

		0.005*"trial" + 0.005*"appearance" + 0.005*"metromedia" + 0.005*"synagogue" + 0.005*"kansa" + 0.004*"religious" + 0.004*"hebrew" + 0.004*"roth" + 0.004*"reform" + 0.004*"television" + 0.004*"trilling" + 0.004*"demoted" + 0.004*"community" + 0.004*"coanchor" + 0.003*"yiddish"
7	Literature, Food and Art	0.017*"joyce" + 0.008*"levine" + 0.006*"fanny" + 0.006*"zuckerman" + 0.005*"solomon" + 0.005*"food" + 0.005*"roth" + 0.004*"jong" + 0.004*"cook" + 0.003*"rossetti" + 0.003*"picasso" + 0.003*"luce" + 0.003*"stassinopoulos" + 0.003*"kafka" + 0.003*"updike" + 0.003*"cooking" + 0.003*"erica" + 0.003*"shaw" + 0.002*"fielding" + 0.002*"queen" + 0.002*"philip" + 0.002*"tan" + 0.002*"percy" + 0.002*"geometric" + 0.002*"nutritional" + 0.002*"liaison" + 0.002*"kruger" + 0.002*"vienna" + 0.002*"clare" + 0.002*"hampton"
8	Academia, intellectualism	0.028*"university" + 0.023*"professor" + 0.013*"history" + 0.011*"study" + 0.010*"american" + 0.009*"culture" + 0.009*"feminist" + 0.008*"book" + 0.008*"literature" + 0.007*"student" + 0.007*"scholar" + 0.006*"academic" + 0.005*"faculty" + 0.005*"harvard" + 0.005*"library" + 0.005*"college" + 0.005*"intellectual" + 0.005*"english" + 0.005*"literary" + 0.005*"historian" + 0.004*"social" + 0.004*"critic" + 0.004*"political" + 0.004*"bloom" + 0.004*"department" + 0.004*"feminism" + 0.004*"gilbert" + 0.004*"princeton" + 0.003*"criticism" + 0.003*"movement"
9	Women's rights and social issues	0.094*"woman" + 0.008*"percent" + 0.007*"child" + 0.006*"feminist" + 0.006*"male" + 0.006*"time" + 0.006*"job" + 0.006*"female" + 0.005*"people" + 0.005*"issue" + 0.005*"family" + 0.005*"life" + 0.004*"movement" + 0.004*"american" + 0.003*"social" + 0.003*"study" + 0.003*"care" + 0.003*"dr" + 0.003*"national" + 0.003*"change" + 0.003*"equal" + 0.003*"role" + 0.003*"country" + 0.003*"sex" + 0.003*"society" + 0.003*"career" + 0.003*"york" + 0.002*"business" + 0.002*"organization" + 0.002*"support"
10	? incoherent	0.072*"japanese" + 0.044*"textbook" + 0.044*"japan" + 0.035*"green" + 0.013*"papandreou" + 0.012*"american" + 0.012*"kelly" + 0.008*"tokyo" + 0.006*"school" + 0.006*"falwell" + 0.006*"publisher" + 0.006*"christian" + 0.006*"humanism" + 0.006*"reading" + 0.006*"schmidt" + 0.005*"humanist" + 0.005*"ministry" + 0.005*"fundamentalist" + 0.004*"kohl" + 0.004*"petra" + 0.004*"formula" + 0.004*"connor" + 0.004*"andreas" + 0.004*"religion" + 0.004*"parent" + 0.004*"greece" + 0.004*"acoustic" + 0.003*"newton" + 0.003*"majority" + 0.003*"orlando"
11	1988 presidential election- Dukakis was the democratic nominee	0.049*"dukakis" + 0.030*"jackson" + 0.018*"hampshire" + 0.013*"gore" + 0.013*"gephardt" + 0.012*"scout" + 0.011*"nunn" + 0.009*"southern" + 0.009*"pollard" + 0.008*"pollster" + 0.008*"iowa" + 0.007*"babbitt" + 0.007*"boy" + 0.006*"robb" + 0.006*"michael" + 0.006*"simon" + 0.006*"primary" + 0.006*"hamilton" + 0.005*"jesse" + 0.005*"analyst" + 0.005*"scoutmaster" + 0.004*"harwood" + 0.004*"spacecraft" + 0.003*"south" + 0.003*"massachusetts" + 0.003*"troop" + 0.003*"sam" + 0.003*"sen" + 0.003*"albert" + 0.003*"tuesday"
12	Money, investments, and property in Florida	0.018*"tax" + 0.014*"florida" + 0.013*"tampa" + 0.011*"editor" + 0.010*"bank" + 0.009*"petersburg" + 0.008*"county" + 0.007*"drug" + 0.006*"shah" + 0.006*"child" + 0.006*"money" + 0.006*"time" + 0.005*"november" + 0.005*"letter" + 0.005*"people" + 0.005*"barry" + 0.005*"loaddate" + 0.005*"business" + 0.005*"clearwater" + 0.004*"credit" + 0.004*"estate" + 0.004*"property" + 0.004*"investment" + 0.003*"income" + 0.003*"plath" + 0.003*"firm" + 0.003*"pinellas" + 0.003*"hughes" + 0.003*"support" + 0.003*"jr"
13	America, books, and US publishers	0.028*"book" + 0.010*"story" + 0.010*"author" + 0.008*"american" + 0.008*"life" + 0.007*"york" + 0.007*"writer" + 0.006*"war" + 0.005*"john" + 0.005*"history" + 0.005*"press" + 0.005*"house" + 0.004*"time" + 0.004*"biography" + 0.004*"fiction" + 0.004*"robert" + 0.004*"published" + 0.004*"written" + 0.004*"volume" + 0.004*"simon" + 0.004*"america" + 0.003*"william" + 0.003*"collection" + 0.003*"james" + 0.003*"family" + 0.003*"poet" + 0.003*"account" + 0.003*"knopf" + 0.003*"woman" + 0.003*"tale"
14	Magazines, adult magazines, advertisements	0.057*"magazine" + 0.036*"pornography" + 0.020*"sale" + 0.018*"publisher" + 0.017*"advertising" + 0.015*"company" + 0.014*"editor" + 0.014*"lauper" + 0.014*"steinem" + 0.013*"market" + 0.013*"industry" + 0.011*"publication" + 0.011*"playboy" + 0.010*"material" + 0.010*"business" + 0.009*"product" + 0.008*"sell" + 0.008*"publishing" + 0.007*"advertiser" + 0.007*"sold" + 0.007*"video" + 0.007*"australian" + 0.007*"store" + 0.007*"book" + 0.007*"ordinance" + 0.007*"adult" + 0.006*"circulation" +

		0.006*"sexually" + 0.006*"cover" + 0.006*"marketing"
15	Books, but more on genres of books	0.011*"life" + 0.006*"woman" + 0.006*"writer" + 0.006*"book" + 0.005*"story" + 0.005*"time" + 0.004*"writing" + 0.004*"character" + 0.004*"black" + 0.003*"people" + 0.003*"love" + 0.003*"read" + 0.003*"sense" + 0.003*"reader" + 0.003*"history" + 0.003*"author" + 0.003*"american" + 0.002*"feminist" + 0.002*"write" + 0.002*"death" + 0.002*"written" + 0.002*"subject" + 0.002*"novel" + 0.002*"wrote" + 0.002*"power" + 0.002*"war" + 0.002*"fiction" + 0.002*"idea" + 0.002*"experience" + 0.002*"form"
16	Philosophy, Psychology and Social theory	0.009*"human" + 0.009*"book" + 0.007*"love" + 0.005*"language" + 0.005*"sexual" + 0.005*"god" + 0.005*"freud" + 0.005*"view" + 0.005*"theory" + 0.004*"feminist" + 0.004*"belief" + 0.004*"sex" + 0.004*"idea" + 0.004*"social" + 0.004*"dr" + 0.004*"question" + 0.003*"political" + 0.003*"nature" + 0.003*"body" + 0.003*"science" + 0.003*"behavior" + 0.003*"moral" + 0.003*"word" + 0.003*"modern" + 0.003*"religion" + 0.003*"liberation" + 0.003*"power" + 0.003*"people" + 0.003*"person" + 0.003*"religious"
17	Judiciary, trials	0.060*"court" + 0.031*"law" + 0.023*"judge" + 0.022*"justice" + 0.016*"decision" + 0.016*"supreme" + 0.012*"right" + 0.011*"federal" + 0.010*"ruling" + 0.010*"lawyer" + 0.009*"discrimination" + 0.008*"appeal" + 0.007*"legal" + 0.006*"ruled" + 0.006*"civil" + 0.006*"sex" + 0.005*"constitutional" + 0.005*"opinion" + 0.005*"amendment" + 0.005*"trial" + 0.004*"public" + 0.004*"attorney" + 0.004*"bork" + 0.004*"term" + 0.004*"government" + 0.004*"judicial" + 0.003*"issue" + 0.003*"district" + 0.003*"congress" + 0.003*"jury"
18	Music	0.058*"song" + 0.046*"music" + 0.034*"album" + 0.030*"singer" + 0.021*"rock" + 0.020*"band" + 0.020*"record" + 0.015*"pop" + 0.014*"blue" + 0.013*"singing" + 0.010*"sound" + 0.010*"folk" + 0.009*"voice" + 0.009*"lyric" + 0.008*"space" + 0.007*"concert" + 0.007*"hit" + 0.007*"recording" + 0.006*"sullivan" + 0.006*"sing" + 0.006*"guitar" + 0.005*"musician" + 0.005*"liverpool" + 0.005*"madonna" + 0.005*"single" + 0.005*"performer" + 0.005*"vocal" + 0.005*"released" + 0.005*"radio" + 0.004*"astronaut"
19	Presidential election (1980)	0.020*"reagan" + 0.019*"president" + 0.015*"party" + 0.013*"political" + 0.013*"republican" + 0.011*"democratic" + 0.010*"issue" + 0.010*"campaign" + 0.010*"right" + 0.009*"candidate" + 0.009*"national" + 0.009*"house" + 0.009*"democrat" + 0.008*"senator" + 0.008*"vote" + 0.008*"amendment" + 0.007*"support" + 0.006*"election" + 0.006*"committee" + 0.006*"senate" + 0.006*"presidential" + 0.006*"conservative" + 0.006*"administration" + 0.005*"convention" + 0.005*"leader" + 0.005*"congress" + 0.005*"white" + 0.005*"organization" + 0.004*"representative" + 0.004*"carter"
20	Film and movies, with some reference to gay/homosexual	0.085*"film" + 0.030*"movie" + 0.014*"director" + 0.012*"directed" + 0.010*"homosexual" + 0.008*"gay" + 0.006*"rated" + 0.005*"scene" + 0.005*"star" + 0.005*"hollywood" + 0.005*"picture" + 0.005*"festival" + 0.005*"violence" + 0.004*"borks" + 0.004*"bork" + 0.004*"maker" + 0.004*"screenplay" + 0.004*"documentary" + 0.004*"screen" + 0.004*"studio" + 0.004*"story" + 0.004*"filmmaker" + 0.004*"script" + 0.003*"cinema" + 0.003*"audience" + 0.003*"sexual" + 0.003*"french" + 0.003*"camera" + 0.003*"character" + 0.003*"based"
21	Television	0.070*"television" + 0.036*"network" + 0.031*"series" + 0.026*"news" + 0.019*"cbs" + 0.016*"abc" + 0.014*"program" + 0.013*"comic" + 0.012*"float" + 0.012*"viewer" + 0.012*"nbc" + 0.009*"producer" + 0.009*"sidote" + 0.009*"broadcast" + 0.009*"episode" + 0.009*"station" + 0.007*"cable" + 0.007*"reporter" + 0.007*"commercial" + 0.007*"medium" + 0.006*"executive" + 0.006*"pb" + 0.006*"programming" + 0.006*"channel" + 0.006*"reed" + 0.006*"turner" + 0.006*"president" + 0.005*"rating" + 0.005*"entertainment" + 0.005*"murder"
22	Abortion (in the courtroom, rally, issues, organizations, etc.)	0.100*"abortion" + 0.018*"right" + 0.016*"clinic" + 0.015*"antiabortion" + 0.013*"health" + 0.012*"organization" + 0.012*"life" + 0.011*"smeal" + 0.011*"law" + 0.010*"national" + 0.009*"court" + 0.008*"pregnancy" + 0.008*"fetus" + 0.008*"feminist" + 0.008*"issue" + 0.007*"supreme" + 0.007*"pill" + 0.007*"woman" + 0.007*"decision" + 0.006*"yard" + 0.006*"birth" + 0.006*"doctor" + 0.006*"drug" + 0.006*"roe" + 0.006*"planned" + 0.006*"prochoice" + 0.005*"march" + 0.005*"rally" + 0.005*"medical" + 0.005*"current"

23	Artificial insemination, surrogacy - some kind of IVF	0.042*"baby" + 0.019*"nora" + 0.018*"surrogate" + 0.017*"egg" + 0.015*"whitehead" + 0.014*"stern" + 0.012*"child" + 0.011*"mary" + 0.011*"beth" + 0.011*"sperm" + 0.010*"mother" + 0.009*"harrington" + 0.009*"donor" + 0.009*"contract" + 0.006*"beauvoir" + 0.006*"couple" + 0.006*"birth" + 0.006*"surrogacy" + 0.006*"embryo" + 0.005*"elizabeth" + 0.005*"forster" + 0.005*"adoption" + 0.005*"motherhood" + 0.005*"aurora" + 0.005*"bourkewhite" + 0.004*"month" + 0.004*"insemination" + 0.004*"artificial" + 0.004*"goldberg" + 0.004*"audubon"
24	Everyday life and personal relationships	0.054*"child" + 0.041*"family" + 0.039*"mother" + 0.030*"father" + 0.025*"husband" + 0.024*"wife" + 0.017*"daughter" + 0.017*"marriage" + 0.016*"parent" + 0.013*"son" + 0.012*"married" + 0.011*"life" + 0.010*"divorce" + 0.008*"girl" + 0.006*"house" + 0.006*"sister" + 0.006*"time" + 0.005*"couple" + 0.005*"friend" + 0.004*"brother" + 0.004*"boy" + 0.004*"born" + 0.004*"divorced" + 0.004*"relationship" + 0.004*"died" + 0.004*"custody" + 0.004*"yearold" + 0.004*"love" + 0.003*"woman" + 0.003*"lived"
25	Visual art	0.044*"art" + 0.029*"artist" + 0.015*"painting" + 0.013*"museum" + 0.010*"gallery" + 0.008*"exhibition" + 0.008*"image" + 0.006*"color" + 0.006*"photograph" + 0.006*"picture" + 0.005*"drawing" + 0.005*"painter" + 0.005*"sculpture" + 0.005*"form" + 0.005*"figure" + 0.004*"woman" + 0.004*"space" + 0.004*"piece" + 0.004*"contemporary" + 0.003*"view" + 0.003*"york" + 0.003*"painted" + 0.003*"portrait" + 0.003*"photography" + 0.003*"object" + 0.003*"visual" + 0.003*"street" + 0.003*"canvas" + 0.003*"wall" + 0.003*"abstract"
26	General school terms (not particular to universities)	0.040*"school" + 0.029*"student" + 0.023*"college" + 0.017*"university" + 0.013*"dr" + 0.009*"education" + 0.007*"teacher" + 0.007*"class" + 0.006*"people" + 0.006*"program" + 0.005*"graduate" + 0.005*"girl" + 0.004*"campus" + 0.004*"professor" + 0.004*"time" + 0.004*"degree" + 0.004*"medical" + 0.004*"teaching" + 0.003*"boy" + 0.003*"public" + 0.003*"president" + 0.003*"board" + 0.003*"science" + 0.003*"hospital" + 0.003*"faculty" + 0.003*"day" + 0.003*"doctor" + 0.003*"boston" + 0.003*"york" + 0.003*"senior"
27	Roman Catholicism (religion)	0.046*"church" + 0.031*"club" + 0.020*"catholic" + 0.016*"religious" + 0.013*"bishop" + 0.010*"pope" + 0.009*"christian" + 0.007*"woman" + 0.007*"priest" + 0.007*"vatican" + 0.007*"theology" + 0.007*"religion" + 0.006*"john" + 0.006*"god" + 0.006*"roman" + 0.006*"paul" + 0.005*"faith" + 0.005*"american" + 0.005*"council" + 0.005*"rev" + 0.005*"united" + 0.005*"people" + 0.004*"jesus" + 0.004*"meeting" + 0.004*"tradition" + 0.004*"issue" + 0.004*"nun" + 0.004*"teaching" + 0.004*"conference" + 0.004*"cardinal"
28	Theatres/plays	0.035*"play" + 0.028*"theater" + 0.014*"production" + 0.011*"performance" + 0.011*"stage" + 0.010*"director" + 0.009*"broadway" + 0.008*"character" + 0.008*"musical" + 0.007*"company" + 0.007*"actor" + 0.007*"season" + 0.007*"opera" + 0.007*"playwright" + 0.007*"audience" + 0.007*"dance" + 0.006*"role" + 0.006*"comedy" + 0.005*"music" + 0.005*"york" + 0.005*"cast" + 0.004*"drama" + 0.004*"actress" + 0.004*"directed" + 0.004*"played" + 0.004*"festival" + 0.004*"set" + 0.004*"scene" + 0.004*"shakespeare" + 0.004*"john"
29	?	0.011*"time" + 0.010*"woman" + 0.008*"people" + 0.007*"day" + 0.005*"life" + 0.005*"friend" + 0.004*"lot" + 0.004*"girl" + 0.003*"photo" + 0.003*"love" + 0.003*"night" + 0.003*"feel" + 0.003*"graphic" + 0.003*"called" + 0.002*"husband" + 0.002*"talk" + 0.002*"told" + 0.002*"feminist" + 0.002*"mother" + 0.002*"house" + 0.002*"hand" + 0.002*"picture" + 0.002*"eye" + 0.002*"head" + 0.002*"play" + 0.002*"career" + 0.002*"money" + 0.002*"story" + 0.002*"real" + 0.002*"hard"

A.2: 1990-1999 LDA Model Topics

Topic #	Topic Name	Word weightings (top 30 words)
0	Local public affairs and journalism	0.014*"people" + 0.011*"city" + 0.007*"time" + 0.007*"county" + 0.007*"letter" + 0.005*"public" + 0.005*"florida" + 0.004*"money" + 0.004*"community" + 0.003*"column" + 0.003*"day" + 0.003*"tax" + 0.003*"atlanta" + 0.003*"tampa" + 0.003*"local" + 0.003*"article" + 0.003*"service" + 0.003*"petersburg" + 0.002*"editorial" + 0.002*"louis" + 0.002*"read" + 0.002*"pay" + 0.002*"editor" + 0.002*"citizen" + 0.002*"radio" + 0.002*"newspaper" + 0.002*"crime" + 0.002*"school" + 0.002*"person" + 0.002*"country"
1	Abortion (more violence than previous decade)	0.074*"abortion" + 0.038*"clinic" + 0.037*"police" + 0.024*"violence" + 0.012*"doctor" + 0.012*"antiabortion" + 0.012*"officer" + 0.011*"murder" + 0.010*"protest" + 0.010*"sheriff" + 0.009*"protester" + 0.009*"rescue" + 0.008*"crime" + 0.008*"killing" + 0.007*"operation" + 0.007*"death" + 0.007*"planned" + 0.006*"gun" + 0.006*"arrested" + 0.006*"activist" + 0.006*"national" + 0.006*"life" + 0.006*"domestic" + 0.006*"feminist" + 0.006*"killed" + 0.005*"violent" + 0.005*"people" + 0.005*"department" + 0.005*"arrest" + 0.005*"parenthood"
2	College sports	0.026*"team" + 0.026*"sport" + 0.018*"coach" + 0.018*"game" + 0.016*"school" + 0.014*"woman" + 0.013*"player" + 0.012*"football" + 0.011*"title" + 0.011*"basketball" + 0.010*"girl" + 0.009*"men" + 0.009*"athlete" + 0.009*"athletic" + 0.009*"ix" + 0.008*"university" + 0.008*"club" + 0.008*"play" + 0.007*"league" + 0.007*"season" + 0.007*"college" + 0.006*"program" + 0.006*"baseball" + 0.005*"ncaa" + 0.005*"national" + 0.005*"championship" + 0.005*"boy" + 0.005*"soccer" + 0.005*"golf" + 0.005*"director"
3	Family words, but also Betty Friedan	0.017*"died" + 0.016*"woman" + 0.015*"husband" + 0.015*"life" + 0.013*"mother" + 0.012*"daughter" + 0.011*"feminist" + 0.011*"wife" + 0.010*"york" + 0.010*"family" + 0.010*"sister" + 0.009*"photo" + 0.008*"father" + 0.008*"steinem" + 0.008*"wrote" + 0.008*"son" + 0.008*"married" + 0.007*"movement" + 0.007*"friend" + 0.007*"loaddate" + 0.007*"time" + 0.007*"betty" + 0.007*"death" + 0.007*"born" + 0.006*"friedan" + 0.006*"right" + 0.005*"career" + 0.005*"helped" + 0.005*"gloria" + 0.005*"age"
4	Religion more generally- Christianity, Judaism, and homosexuality	0.041*"church" + 0.020*"gay" + 0.019*"catholic" + 0.018*"religious" + 0.017*"god" + 0.015*"christian" + 0.011*"religion" + 0.011*"lesbian" + 0.010*"promise" + 0.010*"bishop" + 0.009*"keeper" + 0.008*"jewish" + 0.008*"conference" + 0.007*"priest" + 0.007*"people" + 0.007*"pope" + 0.007*"faith" + 0.006*"jesus" + 0.006*"prayer" + 0.006*"rev" + 0.005*"hooter" + 0.005*"homosexual" + 0.005*"jew" + 0.005*"spiritual" + 0.005*"christ" + 0.005*"bible" + 0.005*"vatican" + 0.005*"leader" + 0.004*"community" + 0.004*"baptist"
5	Judiciary/court, but of note are the words Ginsburg, abortion, discrimination	0.047*"court" + 0.032*"law" + 0.019*"justice" + 0.019*"right" + 0.016*"supreme" + 0.015*"judge" + 0.013*"decision" + 0.010*"federal" + 0.009*"legal" + 0.009*"abortion" + 0.008*"civil" + 0.007*"ruling" + 0.007*"amendment" + 0.006*"woman" + 0.006*"appeal" + 0.005*"issue" + 0.005*"lawyer" + 0.005*"opinion" + 0.005*"government" + 0.005*"discrimination" + 0.004*"protection" + 0.004*"ginsburg" + 0.004*"constitutional" + 0.004*"congress" + 0.004*"constitution" + 0.004*"public" + 0.004*"speech" + 0.003*"argument" + 0.003*"majority" + 0.003*"action"
6	Plays/theatre	0.035*"play" + 0.030*"theater" + 0.016*"director" + 0.015*"performance" + 0.014*"production" + 0.010*"directed" + 0.010*"stage" + 0.008*"opera" + 0.008*"theatre" + 0.008*"musical" + 0.008*"dance" + 0.008*"actor" + 0.008*"company" + 0.007*"broadway" + 0.007*"audience" + 0.006*"comedy" + 0.006*"cast" + 0.006*"festival" + 0.006*"character" + 0.006*"shakespeare" + 0.006*"scene" + 0.005*"playwright" + 0.005*"drama" + 0.005*"role" + 0.004*"pageant" + 0.004*"season" + 0.004*"music" + 0.004*"set" + 0.004*"art" + 0.003*"dancer"
7	International	0.018*"country" + 0.017*"government" + 0.015*"united" + 0.015*"american" + 0.012*"war" + 0.011*"nation" + 0.007*"international" + 0.007*"people" + 0.006*"japan" +

	related words - other nations, but also still 'war'	0.006*"political" + 0.006*"economic" + 0.006*"leader" + 0.006*"conference" + 0.006*"japanese" + 0.006*"population" + 0.006*"minister" + 0.005*"foreign" + 0.005*"islamic" + 0.005*"china" + 0.005*"muslim" + 0.005*"party" + 0.005*"peace" + 0.005*"union" + 0.005*"right" + 0.004*"military" + 0.004*"human" + 0.004*"europe" + 0.004*"european" + 0.004*"western" + 0.004*"africa"
8	Urban life	0.051*"street" + 0.030*"city" + 0.025*"york" + 0.025*"west" + 0.018*"east" + 0.014*"village" + 0.013*"manhattan" + 0.012*"avenue" + 0.009*"house" + 0.009*"building" + 0.008*"brooklyn" + 0.008*"town" + 0.008*"mexican" + 0.007*"italian" + 0.007*"israeli" + 0.007*"german" + 0.007*"jewish" + 0.006*"neighborhood" + 0.006*"square" + 0.006*"israel" + 0.006*"mexico" + 0.005*"restaurant" + 0.005*"spanish" + 0.005*"tomorrow" + 0.005*"park" + 0.005*"elderly" + 0.005*"wall" + 0.005*"french" + 0.005*"green" + 0.004*"hotel"
9	Public personalities	0.031*"paglia" + 0.024*"cammille" + 0.023*"bus" + 0.020*"voelz" + 0.015*"donahue" + 0.015*"ram" + 0.012*"shaw" + 0.011*"sanderson" + 0.011*"wagner" + 0.011*"davidson" + 0.011*"schleuder" + 0.010*"rubin" + 0.010*"fernandez" + 0.010*"watt" + 0.009*"wimp" + 0.009*"sweeney" + 0.007*"marco" + 0.006*"rabinovicz" + 0.006*"rooney" + 0.006*"cagney" + 0.006*"huffington" + 0.006*"lacey" + 0.006*"seat" + 0.005*"platt" + 0.005*"ragtime" + 0.005*"gless" + 0.005*"lucas" + 0.005*"rosenzweig" + 0.005*"curtain" + 0.005*"guidebook"
10	Books, genres	0.031*"book" + 0.013*"story" + 0.013*"life" + 0.009*"author" + 0.009*"writer" + 0.007*"american" + 0.007*"history" + 0.004*"war" + 0.004*"john" + 0.004*"house" + 0.004*"writing" + 0.004*"fiction" + 0.004*"published" + 0.004*"literary" + 0.004*"biography" + 0.004*"university" + 0.004*"reader" + 0.003*"poet" + 0.003*"written" + 0.003*"death" + 0.003*"time" + 0.003*"press" + 0.003*"century" + 0.003*"tale" + 0.003*"character" + 0.003*"essay" + 0.003*"family" + 0.003*"read" + 0.003*"collection" + 0.003*"james"
11	Presidential election (1992)?	0.015*"republican" + 0.012*"campaign" + 0.012*"party" + 0.012*"political" + 0.011*"candidate" + 0.009*"senate" + 0.009*"president" + 0.009*"vote" + 0.009*"bush" + 0.008*"issue" + 0.008*"house" + 0.008*"democrat" + 0.008*"democratic" + 0.007*"election" + 0.006*"voter" + 0.006*"congress" + 0.006*"conservative" + 0.006*"national" + 0.005*"senator" + 0.005*"committee" + 0.005*"support" + 0.005*"politics" + 0.005*"leader" + 0.005*"right" + 0.004*"race" + 0.004*"sen" + 0.004*"liberal" + 0.004*"office" + 0.004*"governor" + 0.004*"poll"
12	Visual Arts and places of exhibition	0.068*"art" + 0.036*"artist" + 0.026*"museum" + 0.019*"painting" + 0.016*"gallery" + 0.011*"exhibition" + 0.009*"image" + 0.008*"photograph" + 0.007*"sculpture" + 0.007*"contemporary" + 0.006*"american" + 0.005*"figure" + 0.005*"painter" + 0.005*"piece" + 0.005*"collection" + 0.005*"drawing" + 0.005*"form" + 0.004*"color" + 0.004*"picture" + 0.004*"exhibit" + 0.004*"modern" + 0.004*"portrait" + 0.004*"center" + 0.004*"quilt" + 0.004*"space" + 0.004*"history" + 0.004*"curator" + 0.004*"installation" + 0.003*"view" + 0.003*"painted"
13	Everyday life and personal experiences	0.013*"day" + 0.011*"time" + 0.009*"people" + 0.008*"girl" + 0.007*"friend" + 0.006*"lot" + 0.004*"told" + 0.004*"woman" + 0.004*"photo" + 0.004*"house" + 0.004*"called" + 0.004*"hair" + 0.004*"life" + 0.004*"yearold" + 0.003*"guy" + 0.003*"night" + 0.003*"week" + 0.003*"mother" + 0.003*"talk" + 0.003*"started" + 0.003*"boy" + 0.003*"husband" + 0.003*"daughter" + 0.003*"food" + 0.003*"hand" + 0.003*"hour" + 0.003*"eye" + 0.003*"table" + 0.003*"kid" + 0.003*"family"
14	Shannon Faulkner, first woman allowed to study at The Citadel, a military school (1994) following a lawsuit	0.023*"citadel" + 0.021*"faulkner" + 0.020*"school" + 0.018*"military" + 0.016*"college" + 0.015*"vmi" + 0.015*"virginia" + 0.013*"marriage" + 0.013*"cadet" + 0.011*"scout" + 0.011*"samesex" + 0.010*"program" + 0.009*"shannon" + 0.009*"bobbitt" + 0.009*"singlesex" + 0.008*"hawaii" + 0.008*"allmale" + 0.008*"sex" + 0.008*"education" + 0.007*"institute" + 0.006*"lorena" + 0.006*"institution" + 0.005*"couple" + 0.005*"academy" + 0.005*"admission" + 0.005*"lawyer" + 0.005*"corp" + 0.005*"class" + 0.005*"coed" + 0.005*"hester"

15	Discrimination in the workplace, and also lawsuits	0.024*"discrimination" + 0.018*"job" + 0.016*"employee" + 0.013*"suit" + 0.012*"lawsuit" + 0.011*"company" + 0.011*"filed" + 0.010*"federal" + 0.010*"lawyer" + 0.010*"complaint" + 0.010*"court" + 0.010*"department" + 0.009*"attorney" + 0.009*"sex" + 0.008*"worker" + 0.008*"office" + 0.008*"district" + 0.008*"pay" + 0.008*"law" + 0.007*"employment" + 0.007*"commission" + 0.007*"employer" + 0.007*"settlement" + 0.006*"damage" + 0.006*"loaddate" + 0.006*"claim" + 0.006*"firm" + 0.006*"city" + 0.005*"manager" + 0.005*"policy"
16	Music	0.055*"music" + 0.036*"song" + 0.021*"band" + 0.019*"album" + 0.017*"singer" + 0.017*"rock" + 0.013*"concert" + 0.012*"record" + 0.011*"country" + 0.011*"pop" + 0.009*"jazz" + 0.008*"lyric" + 0.008*"sound" + 0.008*"artist" + 0.007*"musician" + 0.007*"musical" + 0.007*"voice" + 0.007*"blue" + 0.007*"rap" + 0.007*"radio" + 0.007*"orchestra" + 0.007*"love" + 0.006*"guitar" + 0.006*"hit" + 0.006*"recording" + 0.005*"singing" + 0.005*"composer" + 0.005*"tonight" + 0.005*"tune" + 0.005*"folk"
17	Sexual harassment, specifically Anita Hill case (1991)	0.050*"sexual" + 0.034*"harassment" + 0.016*"thomas" + 0.014*"rape" + 0.013*"hill" + 0.012*"sex" + 0.011*"victim" + 0.010*"charge" + 0.007*"behavior" + 0.007*"trial" + 0.007*"abuse" + 0.007*"sexually" + 0.007*"hearing" + 0.006*"judge" + 0.005*"simpson" + 0.005*"lawyer" + 0.005*"anita" + 0.005*"accused" + 0.005*"clarence" + 0.005*"told" + 0.005*"court" + 0.005*"allegation" + 0.004*"law" + 0.004*"story" + 0.004*"jury" + 0.004*"evidence" + 0.004*"crime" + 0.004*"issue" + 0.004*"report" + 0.004*"assault"
18	Pop Culture	0.008*"photo" + 0.004*"time" + 0.004*"graphic" + 0.004*"star" + 0.004*"john" + 0.004*"love" + 0.004*"night" + 0.003*"loaddate" + 0.003*"king" + 0.003*"black" + 0.003*"set" + 0.002*"color" + 0.002*"george" + 0.002*"michael" + 0.002*"madonna" + 0.002*"york" + 0.002*"hit" + 0.002*"david" + 0.002*"lee" + 0.002*"fan" + 0.002*"day" + 0.002*"summer" + 0.002*"prince" + 0.002*"smith" + 0.002*"feminist" + 0.002*"bob" + 0.002*"heart" + 0.002*"white" + 0.002*"guy" + 0.001*"favorite"
19	Corporate, advertisements, organizational structure	0.021*"company" + 0.017*"magazine" + 0.016*"business" + 0.011*"percent" + 0.010*"store" + 0.008*"sale" + 0.008*"issue" + 0.008*"market" + 0.007*"industry" + 0.007*"executive" + 0.007*"publix" + 0.006*"news" + 0.006*"editor" + 0.005*"people" + 0.005*"week" + 0.005*"product" + 0.005*"president" + 0.005*"month" + 0.005*"york" + 0.005*"money" + 0.005*"cover" + 0.005*"advertising" + 0.004*"price" + 0.004*"internet" + 0.004*"network" + 0.004*"time" + 0.004*"corporate" + 0.004*"service" + 0.004*"stock" + 0.004*"ad"
20	Police events, but neutral, descriptive language	0.009*"official" + 0.007*"city" + 0.006*"police" + 0.006*"week" + 0.005*"month" + 0.004*"charge" + 0.004*"plan" + 0.004*"day" + 0.004*"report" + 0.004*"officer" + 0.004*"monday" + 0.003*"federal" + 0.003*"tuesday" + 0.003*"york" + 0.003*"news" + 0.003*"death" + 0.003*"prison" + 0.003*"reported" + 0.003*"island" + 0.003*"wednesday" + 0.003*"car" + 0.003*"air" + 0.003*"chief" + 0.003*"los" + 0.003*"drug" + 0.003*"killed" + 0.003*"north" + 0.003*"angeles" + 0.003*"judge" + 0.003*"trial"
21	Local events	0.019*"sunday" + 0.019*"saturday" + 0.017*"center" + 0.014*"atlanta" + 0.013*"friday" + 0.011*"road" + 0.010*"tampa" + 0.010*"march" + 0.010*"park" + 0.009*"hall" + 0.009*"south" + 0.008*"ticket" + 0.008*"louis" + 0.008*"oct" + 0.008*"event" + 0.008*"thursday" + 0.007*"wednesday" + 0.007*"program" + 0.007*"nov" + 0.007*"festival" + 0.006*"feb" + 0.006*"north" + 0.006*"georgia" + 0.006*"tuesday" + 0.006*"monday" + 0.006*"april" + 0.005*"university" + 0.005*"child" + 0.005*"art" + 0.005*"avenue"
22	University language and names, but very general	0.045*"school" + 0.042*"student" + 0.041*"university" + 0.030*"college" + 0.019*"board" + 0.013*"director" + 0.011*"professor" + 0.011*"president" + 0.011*"organization" + 0.010*"campus" + 0.009*"dr" + 0.008*"national" + 0.008*"education" + 0.008*"naacp" + 0.008*"graduate" + 0.008*"chavis" + 0.007*"faculty" + 0.006*"association" + 0.006*"program" + 0.006*"pittsburgh" + 0.006*"degree" + 0.006*"executive" + 0.006*"harvard" + 0.005*"department" + 0.005*"meeting" + 0.005*"teacher" + 0.005*"columbia" + 0.004*"grant" + 0.004*"money" + 0.004*"senior"
23	Familial words, but also (birth, doctor, drug)	0.041*"child" + 0.025*"family" + 0.020*"woman" + 0.018*"mother" + 0.013*"parent" + 0.011*"care" + 0.009*"life" + 0.009*"father" + 0.009*"health" + 0.008*"baby" + 0.007*"time" + 0.006*"dr" + 0.006*"marriage" + 0.006*"percent" + 0.005*"medical" + 0.005*"age" + 0.005*"husband" + 0.005*"birth" + 0.005*"doctor" + 0.005*"study" +

		0.005*"day" + 0.005*"daughter" + 0.004*"divorce" + 0.004*"people" + 0.004*"kid" + 0.004*"drug" + 0.004*"wife" + 0.004*"support" + 0.004*"girl" + 0.004*"pregnant"
24	Clintons, most focus on the husband/wife, but also mention of affair	0.075*"clinton" + 0.037*"president" + 0.014*"hillary" + 0.013*"house" + 0.013*"white" + 0.012*"jones" + 0.008*"lady" + 0.008*"public" + 0.008*"political" + 0.007*"husband" + 0.007*"wife" + 0.007*"washington" + 0.006*"feminist" + 0.006*"gore" + 0.006*"woman" + 0.006*"american" + 0.005*"news" + 0.005*"press" + 0.005*"time" + 0.005*"office" + 0.004*"people" + 0.004*"loaddate" + 0.004*"arkansas" + 0.004*"rodham" + 0.004*"quayle" + 0.004*"lewinsky" + 0.004*"scandal" + 0.004*"affair" + 0.003*"presidential" + 0.003*"lawyer"
25	Feminism - woman, sex, power, male, society, culture	0.013*"woman" + 0.008*"life" + 0.008*"time" + 0.007*"feminist" + 0.004*"people" + 0.004*"sex" + 0.004*"love" + 0.004*"loaddate" + 0.003*"sense" + 0.003*"power" + 0.003*"male" + 0.003*"feminism" + 0.003*"idea" + 0.003*"real" + 0.003*"feel" + 0.002*"story" + 0.002*"image" + 0.002*"relationship" + 0.002*"graphic" + 0.002*"female" + 0.002*"book" + 0.002*"question" + 0.002*"mind" + 0.002*"human" + 0.002*"society" + 0.002*"culture" + 0.002*"true" + 0.002*"sexual" + 0.002*"view" + 0.002*"personal"
26	Fashion	0.045*"fashion" + 0.026*"dress" + 0.024*"wear" + 0.018*"designer" + 0.018*"model" + 0.016*"clothes" + 0.016*"hughes" + 0.015*"wearing" + 0.014*"hair" + 0.013*"suit" + 0.011*"skirt" + 0.011*"klein" + 0.011*"style" + 0.010*"shirt" + 0.010*"plath" + 0.010*"bra" + 0.008*"shoe" + 0.008*"heel" + 0.008*"jacket" + 0.008*"pant" + 0.007*"body" + 0.007*"hat" + 0.007*"clothing" + 0.007*"short" + 0.006*"breast" + 0.006*"swimsuit" + 0.006*"woman" + 0.006*"black" + 0.006*"collection" + 0.006*"beauty"
27	Film	0.049*"film" + 0.038*"movie" + 0.013*"star" + 0.011*"character" + 0.010*"television" + 0.009*"hollywood" + 0.008*"series" + 0.008*"story" + 0.007*"director" + 0.007*"comedy" + 0.007*"girl" + 0.006*"producer" + 0.006*"channel" + 0.006*"actor" + 0.006*"actress" + 0.005*"picture" + 0.005*"directed" + 0.005*"minute" + 0.005*"scene" + 0.005*"role" + 0.005*"comic" + 0.005*"violence" + 0.004*"starring" + 0.004*"viewer" + 0.004*"screen" + 0.004*"video" + 0.004*"fox" + 0.004*"network" + 0.004*"played" + 0.004*"barbie"
28	Race	0.026*"black" + 0.017*"school" + 0.012*"white" + 0.010*"people" + 0.010*"girl" + 0.010*"university" + 0.010*"american" + 0.009*"study" + 0.009*"professor" + 0.008*"issue" + 0.007*"student" + 0.007*"class" + 0.007*"education" + 0.007*"program" + 0.007*"boy" + 0.006*"social" + 0.006*"society" + 0.006*"history" + 0.006*"political" + 0.006*"teacher" + 0.005*"culture" + 0.005*"movement" + 0.005*"public" + 0.005*"america" + 0.005*"action" + 0.005*"race" + 0.004*"college" + 0.004*"feminist" + 0.004*"minority" + 0.004*"woman"
29	Gender equality, but in the army?	0.334*"woman" + 0.028*"female" + 0.020*"male" + 0.017*"percent" + 0.010*"gender" + 0.010*"feminist" + 0.008*"force" + 0.007*"military" + 0.007*"job" + 0.007*"equal" + 0.006*"equality" + 0.005*"army" + 0.004*"loaddate" + 0.004*"career" + 0.004*"men" + 0.004*"change" + 0.004*"week" + 0.004*"study" + 0.004*"war" + 0.004*"training" + 0.003*"opportunity" + 0.003*"report" + 0.003*"officer" + 0.003*"organization" + 0.003*"issue" + 0.003*"role" + 0.003*"service" + 0.003*"survey" + 0.003*"combat" + 0.003*"gap"

A.3: 2000-2009 LDA Model Topics

Topic #	Topic Name	Word weightings (top 30 words)
0	General presidential elections (we see Bush, Gore, Obama, Hillary, Clinton)	0.017*"president" + 0.014*"clinton" + 0.013*"bush" + 0.011*"republican" + 0.011*"party" + 0.011*"campaign" + 0.010*"political" + 0.009*"candidate" + 0.009*"vote" + 0.008*"white" + 0.008*"democrat" + 0.007*"democratic" + 0.007*"conservative" + 0.006*"election" + 0.006*"hillary" + 0.006*"house" + 0.006*"obama" + 0.006*"national" + 0.006*"voter" + 0.006*"senator" + 0.005*"issue" + 0.005*"presidential" + 0.005*"senate" + 0.004*"people" + 0.004*"liberal" + 0.004*"governor" + 0.004*"support" + 0.004*"gore" + 0.004*"politics" + 0.004*"race"
1	Gender, also terms of equality, science and health, education	0.133*"woman" + 0.018*"female" + 0.012*"male" + 0.012*"girl" + 0.011*"percent" + 0.010*"gender" + 0.007*"sex" + 0.007*"boy" + 0.007*"study" + 0.005*"job" + 0.004*"issue" + 0.004*"university" + 0.004*"time" + 0.004*"role" + 0.004*"center" + 0.003*"school" + 0.003*"change" + 0.003*"dr" + 0.003*"child" + 0.003*"equality" + 0.003*"difference" + 0.003*"program" + 0.003*"career" + 0.003*"education" + 0.003*"health" + 0.003*"age" + 0.003*"equal" + 0.003*"people" + 0.002*"sexual" + 0.002*"science"
2	Films	0.031*"film" + 0.022*"movie" + 0.010*"director" + 0.009*"scene" + 0.006*"story" + 0.005*"star" + 0.005*"minute" + 0.004*"violence" + 0.004*"directed" + 0.004*"min" + 0.004*"hollywood" + 0.003*"picture" + 0.003*"time" + 0.003*"documentary" + 0.003*"rated" + 0.003*"including" + 0.003*"drug" + 0.003*"character" + 0.003*"filmmaker" + 0.003*"french" + 0.003*"starring" + 0.003*"sex" + 0.003*"screen" + 0.003*"staff" + 0.003*"sexual" + 0.003*"life" + 0.003*"love" + 0.003*"plath" + 0.002*"english" + 0.002*"john"
3	International art and women artists	0.036*"mexico" + 0.032*"vagina" + 0.024*"mexican" + 0.023*"japan" + 0.023*"kahlo" + 0.020*"monologue" + 0.019*"frida" + 0.017*"ensler" + 0.016*"japanese" + 0.014*"okeeffe" + 0.013*"rivera" + 0.008*"eve" + 0.008*"kahlos" + 0.007*"diego" + 0.007*"enslers" + 0.006*"okeeffes" + 0.006*"handmaid" + 0.006*"tokyo" + 0.005*"garcia" + 0.005*"vday" + 0.005*"ringgold" + 0.005*"ballard" + 0.005*"mpls" + 0.005*"fischer" + 0.004*"philadelphia" + 0.004*"potter" + 0.004*"hayek" + 0.004*"body" + 0.004*"stieglitz" + 0.004*"av"
4	Books and writing	0.027*"book" + 0.017*"writer" + 0.016*"life" + 0.014*"poet" + 0.014*"writing" + 0.013*"literary" + 0.013*"author" + 0.012*"poem" + 0.012*"king" + 0.010*"biography" + 0.010*"death" + 0.010*"fiction" + 0.010*"poetry" + 0.009*"story" + 0.009*"history" + 0.008*"bloom" + 0.008*"novel" + 0.007*"wrote" + 0.007*"prize" + 0.007*"writes" + 0.007*"memoir" + 0.006*"letter" + 0.006*"essay" + 0.006*"written" + 0.006*"prose" + 0.006*"novelist" + 0.006*"literature" + 0.005*"hughes" + 0.005*"write" + 0.005*"published"
5	Rick Santorum, Pennsylvania senator, election	0.059*"city" + 0.045*"minnesota" + 0.031*"minneapolis" + 0.021*"council" + 0.016*"mayor" + 0.016*"district" + 0.015*"paul" + 0.013*"board" + 0.009*"community" + 0.008*"green" + 0.008*"santorum" + 0.007*"endorsement" + 0.007*"tax" + 0.007*"public" + 0.007*"neighborhood" + 0.007*"resident" + 0.007*"county" + 0.006*"bloomberg" + 0.006*"coleman" + 0.006*"budget" + 0.006*"twin" + 0.006*"casey" + 0.005*"carney" + 0.005*"building" + 0.005*"local" + 0.005*"av" + 0.005*"background" + 0.005*"housing" + 0.005*"ordinance" + 0.005*"seat"
6	Harassment, rape, police, victims	0.013*"police" + 0.011*"sexual" + 0.008*"officer" + 0.008*"department" + 0.007*"city" + 0.006*"sex" + 0.005*"county" + 0.005*"victim" + 0.005*"abortion" + 0.005*"official" + 0.005*"attorney" + 0.005*"report" + 0.005*"charge" + 0.005*"harassment" + 0.005*"woman" + 0.005*"drug" + 0.004*"told" + 0.004*"yesterday" + 0.004*"crime" + 0.004*"health" + 0.004*"office" + 0.004*"loaddate" + 0.004*"trial" + 0.004*"law" + 0.004*"medical" + 0.004*"week" +

		0.004*"federal" + 0.004*"rape" + 0.004*"doctor" + 0.003*"letter"
7	General reading, magazine, editor, published	0.018*"book" + 0.007*"time" + 0.006*"photo" + 0.006*"york" + 0.005*"house" + 0.004*"life" + 0.004*"story" + 0.004*"author" + 0.003*"writer" + 0.003*"graphic" + 0.003*"fashion" + 0.003*"loaddate" + 0.003*"read" + 0.003*"people" + 0.003*"magazine" + 0.003*"day" + 0.003*"white" + 0.003*"called" + 0.003*"reader" + 0.002*"left" + 0.002*"girl" + 0.002*"dress" + 0.002*"body" + 0.002*"wrote" + 0.002*"city" + 0.002*"editor" + 0.002*"feminist" + 0.002*"writing" + 0.002*"friend" + 0.002*"published"
8	TV- we see some new words that are LGBTQ related	0.012*"television" + 0.009*"series" + 0.009*"photo" + 0.009*"girl" + 0.009*"gay" + 0.007*"star" + 0.007*"fox" + 0.007*"sex" + 0.007*"guy" + 0.007*"network" + 0.006*"night" + 0.006*"season" + 0.006*"episode" + 0.006*"week" + 0.006*"news" + 0.005*"producer" + 0.005*"magazine" + 0.005*"abc" + 0.004*"medium" + 0.004*"lesbian" + 0.004*"viewer" + 0.004*"audience" + 0.004*"video" + 0.004*"people" + 0.004*"tonight" + 0.004*"cbs" + 0.004*"graphic" + 0.004*"host" + 0.004*"loaddate" + 0.003*"executive"
9	College sports	0.043*"school" + 0.021*"sport" + 0.021*"team" + 0.015*"girl" + 0.013*"game" + 0.013*"title" + 0.012*"coach" + 0.010*"club" + 0.010*"player" + 0.010*"ix" + 0.010*"national" + 0.009*"boy" + 0.008*"college" + 0.008*"play" + 0.008*"education" + 0.007*"golf" + 0.007*"basketball" + 0.006*"program" + 0.006*"athlete" + 0.006*"football" + 0.006*"league" + 0.006*"athletic" + 0.006*"student" + 0.005*"johnson" + 0.005*"season" + 0.004*"men" + 0.004*"tournament" + 0.004*"director" + 0.004*"opportunity" + 0.004*"soccer"
10	Days of the week	0.018*"saturday" + 0.016*"sunday" + 0.016*"atlanta" + 0.013*"street" + 0.013*"park" + 0.011*"oct" + 0.010*"friday" + 0.010*"center" + 0.010*"festival" + 0.009*"child" + 0.008*"event" + 0.008*"nov" + 0.007*"road" + 0.007*"art" + 0.007*"music" + 0.006*"dec" + 0.006*"ticket" + 0.005*"day" + 0.005*"noon" + 0.005*"city" + 0.005*"admission" + 0.005*"tour" + 0.005*"thursday" + 0.005*"avenue" + 0.005*"student" + 0.005*"childrens" + 0.005*"program" + 0.005*"sept" + 0.005*"holiday" + 0.004*"ave"
11	Florida and tampa again	0.014*"tampa" + 0.011*"tax" + 0.010*"people" + 0.010*"car" + 0.010*"business" + 0.010*"percent" + 0.009*"price" + 0.008*"sale" + 0.007*"internet" + 0.007*"restaurant" + 0.007*"market" + 0.006*"food" + 0.006*"money" + 0.006*"store" + 0.006*"online" + 0.006*"florida" + 0.005*"owner" + 0.005*"sell" + 0.005*"property" + 0.005*"plan" + 0.005*"driver" + 0.005*"customer" + 0.005*"cost" + 0.005*"product" + 0.004*"technology" + 0.004*"tribune" + 0.004*"site" + 0.004*"bay" + 0.004*"consumer" + 0.003*"government"
12	America and feminism - more broadly (includes some notions of race)	0.013*"american" + 0.009*"country" + 0.008*"war" + 0.007*"people" + 0.006*"united" + 0.006*"political" + 0.006*"woman" + 0.006*"america" + 0.006*"history" + 0.006*"government" + 0.005*"social" + 0.005*"right" + 0.005*"nation" + 0.005*"society" + 0.004*"movement" + 0.004*"culture" + 0.004*"issue" + 0.004*"black" + 0.004*"public" + 0.003*"human" + 0.003*"feminist" + 0.003*"national" + 0.003*"life" + 0.003*"loaddate" + 0.003*"south" + 0.003*"international" + 0.003*"power" + 0.003*"leader" + 0.003*"family" + 0.003*"community"
13	Religion, but now including Islam, terrorism, Christianity, Catholic	0.019*"church" + 0.014*"religious" + 0.014*"god" + 0.012*"christian" + 0.012*"muslim" + 0.011*"catholic" + 0.007*"religion" + 0.007*"war" + 0.007*"people" + 0.007*"faith" + 0.007*"iraq" + 0.006*"islamic" + 0.006*"military" + 0.006*"islam" + 0.006*"priest" + 0.005*"jesus" + 0.005*"pope" + 0.005*"terrorist" + 0.004*"john" + 0.004*"attack" + 0.004*"robertson" + 0.004*"rev" + 0.004*"belief" + 0.004*"iran" + 0.004*"leader" + 0.004*"falwell" + 0.004*"bible" + 0.004*"paul" + 0.003*"conservative" + 0.003*"bishop"
14	Names and locations	0.012*"erickson" + 0.010*"ybor" + 0.008*"shank" + 0.008*"kimmel" + 0.007*"kahn" + 0.007*"nyt" + 0.006*"penny" + 0.006*"keenan" + 0.006*"sanchez" + 0.005*"brazilian" + 0.005*"lamott" + 0.004*"elderly" + 0.004*"mcmanus" + 0.004*"coin" + 0.004*"chrysler" + 0.004*"richler" + 0.004*"tatiana" + 0.004*"maathai" + 0.004*"gerber" + 0.004*"apprenticeship" + 0.004*"damato" + 0.003*"meghan" +

		0.003**"julia" + 0.003**"stringer" + 0.003**"cardoso" + 0.003**"bathe" + 0.003**"gold" + 0.003**"vanessa" + 0.003**"mouthpiece" + 0.003**"kermani"
15	Obituaries and Family events	0.013**"service" + 0.009**"church" + 0.009**"age" + 0.009**"family" + 0.008**"friend" + 0.007**"sister" + 0.007**"memorial" + 0.007**"son" + 0.007**"died" + 0.006**"brother" + 0.006**"death" + 0.006**"daughter" + 0.006**"survived" + 0.005**"county" + 0.005**"funeral" + 0.005**"child" + 0.005**"husband" + 0.005**"wife" + 0.005**"grandchild" + 0.005**"john" + 0.005**"mary" + 0.004**"center" + 0.004**"chapel" + 0.004**"school" + 0.004**"hour" + 0.003**"paul" + 0.003**"monday" + 0.003**"visitation" + 0.003**"parent" + 0.003**"florida"
16	Plays and theatre	0.013**"play" + 0.010**"character" + 0.005**"theater" + 0.005**"love" + 0.005**"story" + 0.004**"production" + 0.004**"actor" + 0.004**"role" + 0.004**"photo" + 0.003**"life" + 0.003**"time" + 0.003**"set" + 0.003**"stage" + 0.003**"comedy" + 0.003**"played" + 0.003**"performance" + 0.003**"cast" + 0.003**"john" + 0.003**"loaddate" + 0.003**"director" + 0.003**"actress" + 0.003**"audience" + 0.003**"drama" + 0.003**"graphic" + 0.003**"woman" + 0.002**"star" + 0.002**"directed" + 0.002**"musical" + 0.002**"girl" + 0.002**"scene"
17	Art and Cooking	0.017**"art" + 0.012**"artist" + 0.008**"food" + 0.006**"piece" + 0.006**"pittsburgh" + 0.005**"quilt" + 0.005**"cooking" + 0.005**"photo" + 0.004**"cook" + 0.004**"space" + 0.004**"cup" + 0.004**"kitchen" + 0.004**"feminist" + 0.004**"chicago" + 0.004**"warhol" + 0.003**"add" + 0.003**"video" + 0.003**"performance" + 0.003**"water" + 0.003**"carnegie" + 0.003**"project" + 0.003**"green" + 0.003**"oil" + 0.003**"egg" + 0.003**"chicken" + 0.002**"time" + 0.002**"minute" + 0.002**"medium" + 0.002**"dinner" + 0.002**"recipe"
18	Visual art	0.029**"art" + 0.016**"artist" + 0.015**"museum" + 0.012**"street" + 0.011**"painting" + 0.009**"gallery" + 0.008**"exhibition" + 0.005**"photograph" + 0.005**"image" + 0.005**"drawing" + 0.005**"johnson" + 0.005**"west" + 0.004**"american" + 0.004**"sculpture" + 0.004**"avenue" + 0.004**"york" + 0.004**"smith" + 0.004**"contemporary" + 0.003**"collection" + 0.003**"picture" + 0.003**"modern" + 0.003**"video" + 0.003**"color" + 0.003**"hour" + 0.003**"object" + 0.003**"portrait" + 0.003**"painter" + 0.003**"form" + 0.003**"photography" + 0.003**"design"
19	Missouri and St. Louis	0.071**"louis" + 0.022**"missouri" + 0.016**"boulevard" + 0.014**"charles" + 0.010**"info" + 0.009**"lindbergh" + 0.008**"postdispatch" + 0.008**"stepford" + 0.007**"kirkwood" + 0.007**"webster" + 0.007**"jw" + 0.007**"university" + 0.007**"copy" + 0.007**"walker" + 0.006**"barnes" + 0.006**"washington" + 0.006**"cha" + 0.006**"bowie" + 0.005**"sofia" + 0.005**"illinois" + 0.005**"sign" + 0.005**"schieffelin" + 0.005**"cost" + 0.005**"clayton" + 0.005**"duke" + 0.005**"levin" + 0.004**"ofallon" + 0.004**"bryn" + 0.004**"manchester" + 0.004**"north"
20	?	0.026**"mailer" + 0.024**"ford" + 0.020**"baker" + 0.019**"welch" + 0.017**"borough" + 0.014**"sweden" + 0.013**"swedish" + 0.011**"tritt" + 0.009**"silva" + 0.008**"cristina" + 0.007**"norman" + 0.007**"woodward" + 0.007**"hastings" + 0.006**"greco" + 0.006**"cannon" + 0.006**"sterling" + 0.006**"gilmore" + 0.005**"egan" + 0.005**"arthur" + 0.005**"auerbach" + 0.005**"intel" + 0.005**"moran" + 0.005**"fitch" + 0.004**"maude" + 0.004**"abercrombie" + 0.004**"dotson" + 0.004**"supremacist" + 0.004**"stockholm" + 0.004**"obrien" + 0.004**"norway"
21	Cultural Events	0.025**"art" + 0.021**"center" + 0.012**"march" + 0.012**"library" + 0.011**"hour" + 0.011**"road" + 0.011**"april" + 0.011**"feb" + 0.011**"street" + 0.010**"friday" + 0.009**"theater" + 0.009**"june" + 0.009**"avenue" + 0.009**"saturday" + 0.008**"sunday" + 0.007**"gallery" + 0.007**"tuesday" + 0.007**"thursday" + 0.007**"museum" + 0.007**"jan" + 0.007**"music" + 0.006**"hudson" + 0.006**"public" + 0.006**"monday" + 0.006**"exhibition" + 0.006**"main" + 0.005**"wednesday" + 0.005**"noon" + 0.005**"college" + 0.005**"south"
22	Judaism, but also pakistan, Turkey, taliban, afghanistan, and russian	0.021**"jewish" + 0.011**"rabbi" + 0.008**"beach" + 0.008**"jew" + 0.007**"orthodox" + 0.006**"seder" + 0.006**"protest" + 0.006**"taliban" + 0.006**"afghanistan" + 0.006**"turkey" + 0.005**"people" + 0.005**"afghan" + 0.005**"day" + 0.005**"israel" + 0.005**"passover" + 0.004**"russian" + 0.004**"police" + 0.004**"city" + 0.004**"camp" + 0.004**"tee" + 0.004**"protester" + 0.004**"haggadah" + 0.004**"week" +

		0.004*"pakistan" + 0.004*"synagogue" + 0.003*"hebrew" + 0.003*"country" + 0.003*"community" + 0.003*"woman" + 0.003*"family"
23	Judiciary, (Oconnor, Robert, wade)	0.050*"court" + 0.035*"law" + 0.026*"justice" + 0.022*"judge" + 0.017*"right" + 0.016*"supreme" + 0.016*"abortion" + 0.012*"decision" + 0.012*"discrimination" + 0.011*"robert" + 0.011*"federal" + 0.009*"legal" + 0.007*"lawyer" + 0.007*"ruling" + 0.007*"appeal" + 0.007*"congress" + 0.006*"roe" + 0.006*"amendment" + 0.006*"civil" + 0.005*"opinion" + 0.005*"sex" + 0.005*"issue" + 0.005*"oconnor" + 0.004*"wade" + 0.004*"government" + 0.004*"wrote" + 0.004*"constitutional" + 0.004*"majority" + 0.004*"constitution" + 0.003*"protection"
24	Music	0.031*"music" + 0.023*"song" + 0.020*"band" + 0.014*"rock" + 0.013*"album" + 0.011*"singer" + 0.008*"pop" + 0.008*"record" + 0.007*"sound" + 0.007*"love" + 0.007*"hit" + 0.007*"artist" + 0.006*"country" + 0.005*"hiphop" + 0.005*"photo" + 0.005*"fan" + 0.005*"voice" + 0.004*"lyric" + 0.004*"musician" + 0.004*"guitar" + 0.004*"radio" + 0.004*"star" + 0.004*"recording" + 0.004*"blue" + 0.004*"concert" + 0.004*"tune" + 0.004*"time" + 0.004*"singing" + 0.004*"folk" + 0.003*"solo"
25	Broadway, shows, New York	0.029*"street" + 0.015*"theater" + 0.015*"west" + 0.011*"tomorrow" + 0.009*"tonight" + 0.008*"sunday" + 0.008*"east" + 0.007*"music" + 0.007*"dance" + 0.006*"village" + 0.006*"avenue" + 0.005*"thursday" + 0.005*"broadway" + 0.005*"minute" + 0.005*"center" + 0.005*"play" + 0.005*"night" + 0.004*"york" + 0.004*"manhattan" + 0.004*"opera" + 0.004*"performance" + 0.004*"brantley" + 0.004*"cover" + 0.004*"ticket" + 0.004*"jazz" + 0.004*"wednesday" + 0.004*"tuesday" + 0.004*"hall" + 0.003*"brooklyn" + 0.003*"musical"
26	General university words	0.055*"student" + 0.055*"university" + 0.042*"college" + 0.042*"school" + 0.027*"professor" + 0.015*"class" + 0.015*"dr" + 0.015*"campus" + 0.012*"teacher" + 0.012*"study" + 0.010*"harvard" + 0.010*"summer" + 0.009*"science" + 0.009*"graduate" + 0.009*"education" + 0.009*"faculty" + 0.008*"academic" + 0.007*"degree" + 0.007*"program" + 0.006*"teaching" + 0.006*"yale" + 0.006*"black" + 0.005*"president" + 0.005*"department" + 0.005*"english" + 0.005*"scholar" + 0.005*"wrote" + 0.005*"writing" + 0.005*"taught" + 0.005*"teach"
27	Corporations, there is one instance of discrimination and lawsuit, but largely lacking apart from general corporate words	0.029*"company" + 0.016*"employee" + 0.013*"job" + 0.013*"business" + 0.011*"pay" + 0.011*"discrimination" + 0.010*"executive" + 0.009*"walmart" + 0.009*"worker" + 0.009*"firm" + 0.009*"lawsuit" + 0.008*"manager" + 0.007*"suit" + 0.006*"lawyer" + 0.006*"money" + 0.006*"corporate" + 0.005*"employer" + 0.005*"office" + 0.005*"bank" + 0.005*"percent" + 0.005*"filed" + 0.005*"commission" + 0.005*"employment" + 0.005*"chief" + 0.005*"financial" + 0.005*"federal" + 0.005*"board" + 0.004*"paid" + 0.004*"management" + 0.004*"agency"
28	Family words	0.020*"woman" + 0.012*"time" + 0.011*"child" + 0.011*"life" + 0.009*"mother" + 0.009*"people" + 0.009*"family" + 0.007*"day" + 0.006*"husband" + 0.005*"feminist" + 0.005*"marriage" + 0.005*"father" + 0.005*"wife" + 0.004*"friend" + 0.004*"parent" + 0.004*"daughter" + 0.004*"feel" + 0.003*"loaddate" + 0.003*"love" + 0.003*"married" + 0.003*"told" + 0.003*"kid" + 0.003*"lot" + 0.003*"job" + 0.003*"couple" + 0.002*"baby" + 0.002*"talk" + 0.002*"person" + 0.002*"question" + 0.002*"relationship"
29	Female Characters and Authors	0.048*"miller" + 0.023*"wharton" + 0.022*"wright" + 0.018*"nancy" + 0.018*"memory" + 0.016*"clancy" + 0.014*"ally" + 0.012*"library" + 0.012*"amelia" + 0.010*"freud" + 0.009*"edith" + 0.008*"kelley" + 0.008*"boyle" + 0.008*"drew" + 0.007*"charlotte" + 0.006*"earhart" + 0.006*"mack" + 0.006*"mcbeal" + 0.006*"frank" + 0.006*"wexler" + 0.005*"stefani" + 0.005*"eyre" + 0.005*"lee" + 0.005*"jane" + 0.005*"camilla" + 0.005*"bronte" + 0.005*"abduction" + 0.004*"dsouza" + 0.004*"librarian" + 0.004*"davy"

A.4: The New York Times LDA Model Topics

Topic #	Topic Name	Word weightings (top 30 words)
0	Visual art and locations of display	0.048*"art" + 0.030*"museum" + 0.022*"artist" + 0.014*"painting" + 0.012*"exhibition" + 0.011*"gallery" + 0.010*"street" + 0.006*"american" + 0.006*"sculpture" + 0.006*"contemporary" + 0.005*"york" + 0.005*"collection" + 0.005*"jan" + 0.005*"drawing" + 0.005*"smith" + 0.005*"modern" + 0.004*"avenue" + 0.004*"johnson" + 0.004*"installation" + 0.004*"curator" + 0.004*"painter" + 0.004*"west" + 0.003*"cotton" + 0.003*"thcentury" + 0.003*"object" + 0.003*"design" + 0.003*"century" + 0.003*"photograph" + 0.003*"metropolitan" + 0.003*"center"
1	Female representation in the media	0.023*"woman" + 0.008*"girl" + 0.008*"sex" + 0.006*"love" + 0.005*"character" + 0.005*"female" + 0.005*"sexual" + 0.004*"feminist" + 0.004*"male" + 0.003*"loadate" + 0.003*"beauty" + 0.003*"lady" + 0.003*"graphic" + 0.002*"body" + 0.002*"heroine" + 0.002*"time" + 0.002*"story" + 0.002*"real" + 0.002*"hair" + 0.002*"photo" + 0.002*"fantasy" + 0.002*"scene" + 0.002*"power" + 0.002*"eye" + 0.002*"bad" + 0.002*"boy" + 0.002*"lover" + 0.002*"take" + 0.002*"moment" + 0.001*"jane"
2	Films	0.062*"film" + 0.031*"movie" + 0.011*"directed" + 0.009*"director" + 0.008*"minute" + 0.008*"story" + 0.007*"star" + 0.006*"hollywood" + 0.005*"picture" + 0.005*"starring" + 0.005*"character" + 0.005*"comedy" + 0.005*"documentary" + 0.004*"scene" + 0.004*"screen" + 0.004*"series" + 0.004*"time" + 0.004*"played" + 0.004*"feature" + 0.003*"actor" + 0.003*"filmmaker" + 0.003*"actress" + 0.003*"audience" + 0.003*"love" + 0.003*"producer" + 0.003*"released" + 0.003*"john" + 0.003*"television" + 0.003*"rated" + 0.003*"life"
3	Presidential elections	0.020*"president" + 0.014*"clinton" + 0.012*"republican" + 0.012*"campaign" + 0.012*"political" + 0.010*"party" + 0.009*"candidate" + 0.009*"reagan" + 0.008*"senator" + 0.008*"democratic" + 0.008*"house" + 0.008*"democrat" + 0.008*"vote" + 0.007*"bush" + 0.006*"issue" + 0.006*"conservative" + 0.006*"senate" + 0.006*"national" + 0.006*"white" + 0.005*"election" + 0.005*"governor" + 0.005*"presidential" + 0.005*"washington" + 0.005*"committee" + 0.004*"politics" + 0.004*"voter" + 0.004*"right" + 0.004*"abortion" + 0.004*"public" + 0.004*"support"
4	Cities and Cultures	0.068*"los" + 0.065*"angeles" + 0.024*"australian" + 0.022*"ross" + 0.017*"chicago" + 0.017*"australia" + 0.011*"sydney" + 0.011*"casanova" + 0.009*"armstrong" + 0.009*"tucker" + 0.009*"thai" + 0.007*"krasner" + 0.007*"nussbaum" + 0.007*"california" + 0.007*"flynn" + 0.007*"hyman" + 0.007*"judy" + 0.006*"san" + 0.006*"hanson" + 0.006*"fletcher" + 0.005*"silber" + 0.005*"varda" + 0.005*"thailand" + 0.004*"ana" + 0.004*"mendieta" + 0.004*"bangkok" + 0.004*"maam" + 0.004*"pollock" + 0.004*"rollins" + 0.004*"nairobi"
5	Culture, America, history, politics, intellectualism	0.010*"american" + 0.007*"culture" + 0.007*"social" + 0.007*"history" + 0.006*"political" + 0.005*"black" + 0.005*"society" + 0.005*"idea" + 0.005*"people" + 0.005*"view" + 0.004*"movement" + 0.004*"feminist" + 0.004*"question" + 0.004*"human" + 0.004*"cultural" + 0.004*"power" + 0.003*"study" + 0.003*"century" + 0.003*"public" + 0.003*"form" + 0.003*"issue" + 0.003*"sense" + 0.003*"theory" + 0.003*"intellectual" + 0.003*"sexual" + 0.003*"male" + 0.003*"university" + 0.003*"politics" + 0.003*"subject" + 0.003*"term"
6	Court, judiciary, discrimination and harassment	0.034*"court" + 0.021*"law" + 0.014*"discrimination" + 0.014*"justice" + 0.014*"judge" + 0.013*"federal" + 0.010*"right" + 0.010*"decision" + 0.009*"lawyer" + 0.009*"supreme" + 0.009*"sex" + 0.008*"civil" + 0.008*"sexual" + 0.006*"legal" + 0.006*"harassment" + 0.006*"employee" + 0.005*"appeal" + 0.005*"ruling" + 0.005*"suit" + 0.005*"department" + 0.004*"action" + 0.004*"lawsuit" + 0.004*"commission" + 0.004*"district" + 0.004*"job" + 0.004*"employer" + 0.004*"government" + 0.004*"policy" + 0.004*"filed" + 0.003*"complaint"

7	Judaism, but also psychoanalysis	0.058**jewish" + 0.028**jew" + 0.022**rabbi" + 0.017**freud" + 0.016**israel" + 0.012**orthodox" + 0.008**judaism" + 0.007**religious" + 0.007**synagogue" + 0.007**community" + 0.007**harris" + 0.007**hebrew" + 0.005**witch" + 0.005**god" + 0.005**psychoanalysis" + 0.005**love" + 0.004**holocaust" + 0.004**ritual" + 0.004**psychoanalyst" + 0.004**prayer" + 0.004**psychoanalytic" + 0.004**roth" + 0.004**daly" + 0.003**torah" + 0.003**yiddish" + 0.003**tradition" + 0.003**religion" + 0.003**service" + 0.003**jerusalem" + 0.003**passover"
8	Television, and sports on television	0.017**television" + 0.013**sport" + 0.012**team" + 0.011**game" + 0.009**girl" + 0.008**network" + 0.007**player" + 0.006**news" + 0.006**boy" + 0.006**time" + 0.006**series" + 0.006**club" + 0.005**coach" + 0.005**program" + 0.005**play" + 0.005**title" + 0.005**season" + 0.004**athlete" + 0.004**week" + 0.004**woman" + 0.004**football" + 0.004**national" + 0.004**executive" + 0.004**abc" + 0.004**basketball" + 0.004**radio" + 0.004**medium" + 0.003**cbs" + 0.003**viewer" + 0.003**athletic"
9	Women's Health	0.034**dr" + 0.014**doctor" + 0.014**health" + 0.013**medical" + 0.012**abortion" + 0.010**clinic" + 0.009**patient" + 0.008**child" + 0.008**woman" + 0.008**baby" + 0.008**hospital" + 0.008**drug" + 0.007**birth" + 0.007**sex" + 0.006**pregnancy" + 0.006**disease" + 0.006**week" + 0.005**cancer" + 0.005**control" + 0.005**study" + 0.005**aid" + 0.005**body" + 0.005**pregnant" + 0.004**pill" + 0.004**medicine" + 0.004**egg" + 0.004**treatment" + 0.004**surgery" + 0.004**risk" + 0.004**sexual"
10	Retail and Theatre	0.030**walmart" + 0.026**street" + 0.019**west" + 0.018**theater" + 0.008**brantley" + 0.007**broadway" + 0.007**ticket" + 0.006**walmarts" + 0.006**east" + 0.006**schmidt" + 0.005**chavez" + 0.005**saturday" + 0.005**sunday" + 0.005**mason" + 0.005**friday" + 0.005**musical" + 0.004**hopkins" + 0.004**village" + 0.004**clinton" + 0.004**mercer" + 0.004**preview" + 0.004**thursday" + 0.004**revival" + 0.004**wednesday" + 0.004**tuesday" + 0.003**ben" + 0.003**waterhouse" + 0.003**wayland" + 0.003**frier" + 0.003**oneill"
11	Music	0.037**music" + 0.037**song" + 0.017**rock" + 0.016**album" + 0.016**band" + 0.016**singer" + 0.012**pop" + 0.010**record" + 0.010**voice" + 0.008**sound" + 0.008**lyric" + 0.007**singing" + 0.006**musical" + 0.006**hit" + 0.006**country" + 0.006**audience" + 0.006**love" + 0.005**folk" + 0.005**musician" + 0.005**recording" + 0.005**star" + 0.005**blue" + 0.005**time" + 0.005**artist" + 0.004**concert" + 0.004**performer" + 0.004**rap" + 0.004**photo" + 0.004**guitar" + 0.004**sing"
12	Books and writing	0.059**book" + 0.021**writer" + 0.014**writing" + 0.013**author" + 0.011**wrote" + 0.011**published" + 0.010**reader" + 0.009**literary" + 0.009**read" + 0.009**editor" + 0.007**feminist" + 0.007**literature" + 0.007**letter" + 0.007**write" + 0.007**reading" + 0.006**review" + 0.006**york" + 0.006**fiction" + 0.006**press" + 0.005**novel" + 0.005**written" + 0.005**time" + 0.005**publisher" + 0.005**story" + 0.005**critic" + 0.005**life" + 0.005**woman" + 0.005**english" + 0.004**poet" + 0.004**writes"
13	Cultural events	0.020**street" + 0.017**art" + 0.017**sunday" + 0.016**center" + 0.016**saturday" + 0.011**friday" + 0.011**avenue" + 0.009**child" + 0.009**park" + 0.009**museum" + 0.007**thursday" + 0.007**library" + 0.007**theater" + 0.007**tuesday" + 0.006**noon" + 0.006**road" + 0.006**hour" + 0.006**monday" + 0.006**june" + 0.005**music" + 0.005**march" + 0.005**admission" + 0.005**exhibition" + 0.005**april" + 0.005**wednesday" + 0.005**york" + 0.004**gallery" + 0.004**hall" + 0.004**feb" + 0.004**public"
14	Equality in the workplace, but also as a social issue	0.137**woman" + 0.012**percent" + 0.009**feminist" + 0.009**female" + 0.008**job" + 0.007**family" + 0.007**issue" + 0.007**organization" + 0.006**child" + 0.006**male" + 0.006**abortion" + 0.006**national" + 0.005**right" + 0.005**time" + 0.004**movement" + 0.004**director" + 0.004**equal" + 0.004**support" + 0.004**care" + 0.003**life" + 0.003**pay" + 0.003**york" + 0.003**service" + 0.003**social" + 0.003**role" + 0.003**program" + 0.003**change" + 0.003**sex" + 0.003**gender" + 0.003**worker"

15	Plays/Theatre	0.030**play" + 0.022**theater" + 0.013**production" + 0.010**director" + 0.010**character" + 0.009**stage" + 0.008**performance" + 0.008**actor" + 0.007**broadway" + 0.007**role" + 0.007**audience" + 0.006**comedy" + 0.006**playwright" + 0.005**musical" + 0.005**company" + 0.005**york" + 0.004**cast" + 0.004**actress" + 0.004**set" + 0.004**drama" + 0.004**shakespeare" + 0.004**directed" + 0.004**opera" + 0.004**season" + 0.004**scene" + 0.004**david" + 0.004**john" + 0.003**photo" + 0.003**played" + 0.003**london"
16	Broadway, Shows, New York	0.029**street" + 0.014**west" + 0.013**theater" + 0.011**tomorrow" + 0.010**tonight" + 0.009**dance" + 0.009**music" + 0.008**sunday" + 0.008**east" + 0.006**village" + 0.006**avenue" + 0.005**night" + 0.005**thursday" + 0.005**minute" + 0.005**york" + 0.005**cover" + 0.004**center" + 0.004**manhattan" + 0.004**jazz" + 0.004**ticket" + 0.004**opera" + 0.004**performance" + 0.004**play" + 0.004**broadway" + 0.004**hall" + 0.004**program" + 0.004**band" + 0.004**tuesday" + 0.004**concert" + 0.003**wednesday"
17	University	0.061**school" + 0.046**university" + 0.043**student" + 0.037**college" + 0.024**professor" + 0.018**education" + 0.011**teacher" + 0.011**program" + 0.011**study" + 0.010**class" + 0.010**girl" + 0.010**department" + 0.009**dr" + 0.009**campus" + 0.009**harvard" + 0.009**graduate" + 0.008**faculty" + 0.008**boy" + 0.007**academic" + 0.006**columbia" + 0.006**science" + 0.005**president" + 0.005**degree" + 0.005**institution" + 0.005**teaching" + 0.005**york" + 0.005**board" + 0.005**institute" + 0.004**yale" + 0.004**teach"
18	Government and international relations	0.014**government" + 0.013**country" + 0.012**united" + 0.008**nation" + 0.008**american" + 0.007**people" + 0.007**union" + 0.007**official" + 0.006**war" + 0.006**party" + 0.006**political" + 0.006**leader" + 0.006**city" + 0.005**international" + 0.005**minister" + 0.004**national" + 0.004**soviet" + 0.004**worker" + 0.004**president" + 0.004**foreign" + 0.004**protest" + 0.003**police" + 0.003**economic" + 0.003**europe" + 0.003**security" + 0.003**force" + 0.003**communist" + 0.003**japan" + 0.003**report" + 0.003**month"
19	Assorted words related to companies-stores, advertising, product	0.020**company" + 0.014**business" + 0.011**magazine" + 0.008**time" + 0.008**fashion" + 0.008**york" + 0.006**people" + 0.006**executive" + 0.006**store" + 0.006**sale" + 0.006**market" + 0.005**industry" + 0.005**money" + 0.004**designer" + 0.004**price" + 0.004**month" + 0.004**photo" + 0.004**chief" + 0.004**firm" + 0.004**advertising" + 0.004**bank" + 0.003**office" + 0.003**manager" + 0.003**dress" + 0.003**sell" + 0.003**percent" + 0.003**product" + 0.003**week" + 0.003**editor" + 0.003**day"
20	Street Art and Photography	0.011**street" + 0.010**artist" + 0.009**art" + 0.008**image" + 0.008**photograph" + 0.007**painting" + 0.006**gallery" + 0.005**picture" + 0.005**color" + 0.005**piece" + 0.004**hour" + 0.004**west" + 0.004**johnson" + 0.004**video" + 0.004**form" + 0.004**wall" + 0.003**drawing" + 0.003**admission" + 0.003**figure" + 0.003**student" + 0.003**white" + 0.003**photo" + 0.003**body" + 0.003**space" + 0.002**material" + 0.002**portrait" + 0.002**photographer" + 0.002**painted" + 0.002**time" + 0.002**sunday"
21	New York	0.027**city" + 0.015**york" + 0.013**house" + 0.009**food" + 0.008**building" + 0.008**town" + 0.007**restaurant" + 0.007**island" + 0.006**day" + 0.006**club" + 0.006**park" + 0.006**street" + 0.005**local" + 0.005**hotel" + 0.005**village" + 0.005**dinner" + 0.005**car" + 0.005**mile" + 0.004**neighborhood" + 0.004**water" + 0.004**bar" + 0.004**kitchen" + 0.004**east" + 0.004**garden" + 0.004**shop" + 0.004**space" + 0.004**apartment" + 0.004**table" + 0.003**square" + 0.003**foot"
22	Personalities	0.037**mailer" + 0.023**florio" + 0.017**lerner" + 0.015**buckley" + 0.010**norman" + 0.010**segal" + 0.009**meyer" + 0.008**sabbath" + 0.008**gilmore" + 0.007**summer" + 0.007**madeline" + 0.007**swinton" + 0.007**lurie" + 0.007**jeanne" + 0.006**potter" + 0.006**moreau" + 0.006**waldman" + 0.006**straw" + 0.005**minyan" + 0.005**danish" + 0.005**kureishi" + 0.004**horatio" + 0.004**gerda" + 0.004**delilah" + 0.004**dearborn" + 0.003**peckinpah" + 0.003**jaffe" + 0.003**tilda" + 0.003**dowd" + 0.003**salvi"

23	Family	0.013**time" + 0.011**life" + 0.010**people" + 0.009**child" + 0.009**mother" + 0.007**day" + 0.007**family" + 0.007**woman" + 0.006**father" + 0.006**friend" + 0.006**husband" + 0.004**wife" + 0.004**feel" + 0.004**daughter" + 0.004**told" + 0.003**marriage" + 0.003**parent" + 0.003**left" + 0.003**son" + 0.003**married" + 0.003**lot" + 0.003**black" + 0.002**york" + 0.002**called" + 0.002**talk" + 0.002**love" + 0.002**white" + 0.002**couple" + 0.002**age" + 0.002**hard"
24	Army	0.040**military" + 0.023**army" + 0.021**officer" + 0.013**force" + 0.012**soldier" + 0.011**commander" + 0.009**navy" + 0.009**combat" + 0.009**air" + 0.009**war" + 0.007**captain" + 0.007**faludi" + 0.006**academy" + 0.006**lieutenant" + 0.006**endowment" + 0.006**mcmillan" + 0.005**cadet" + 0.005**citadel" + 0.005**corp" + 0.005**faulkner" + 0.005**marine" + 0.005**helmsley" + 0.005**civilian" + 0.004**greek" + 0.004**command" + 0.004**russo" + 0.004**gen" + 0.004**sexual" + 0.004**shannon" + 0.004**sergeant"
25	Books, but more of a biography/memoir tilt	0.014**life" + 0.011**story" + 0.009**book" + 0.007**american" + 0.006**history" + 0.006**war" + 0.005**author" + 0.005**john" + 0.005**house" + 0.004**time" + 0.004**biography" + 0.003**family" + 0.003**love" + 0.003**simon" + 0.003**death" + 0.003**america" + 0.003**woman" + 0.003**robert" + 0.003**account" + 0.003**knopf" + 0.003**tale" + 0.003**writer" + 0.003**memoir" + 0.003**james" + 0.003**age" + 0.002**fiction" + 0.002**university" + 0.002**schuster" + 0.002**collection" + 0.002**random"
26	New Jersey and Intellectuals	Word: 0.044**jersey" + 0.024**beauvoir" + 0.019**morris" + 0.015**wright" + 0.015**sartre" + 0.012**newark" + 0.011**simone" + 0.010**rutgers" + 0.010**heilbrun" + 0.008**princeton" + 0.007**brunswick" + 0.007**chanel" + 0.006**bailey" + 0.006**nj" + 0.005**bergen" + 0.005**trenton" + 0.005**coco" + 0.005**kean" + 0.005**knitting" + 0.005**paterson" + 0.005**lingerie" + 0.004**montclair" + 0.004**mcgrath" + 0.004**bra" + 0.004**englewood" + 0.004**weinstein" + 0.004**orange" + 0.004**hildegard" + 0.003**hamilton" + 0.003**frank"
27	? more names	0.021**cole" + 0.021**liberty" + 0.018**powell" + 0.017**bryant" + 0.014**wimbledon" + 0.012**billie" + 0.010**meter" + 0.010**henson" + 0.009**connors" + 0.009**dominican" + 0.008**tennis" + 0.007**aclu" + 0.005**lieberman" + 0.005**evert" + 0.005**bilingual" + 0.005**mcenroe" + 0.004**backhand" + 0.004**donohue" + 0.003**trujillo" + 0.003**bryants" + 0.003**blau" + 0.003**willed" + 0.003**pettine" + 0.003**cheek" + 0.002**egger" + 0.002**saperstein" + 0.002**aclus" + 0.002**shanker" + 0.002**cooney" + 0.002**mismatch"
28	Religion- islam, christianity, roman catholicism	0.036**church" + 0.019**catholic" + 0.018**religious" + 0.014**christian" + 0.011**god" + 0.010**bishop" + 0.009**muslim" + 0.009**religion" + 0.008**pope" + 0.008**priest" + 0.007**vatican" + 0.006**faith" + 0.006**conservative" + 0.006**roman" + 0.005**jesus" + 0.005**leader" + 0.005**conference" + 0.005**john" + 0.005**rev" + 0.005**islamic" + 0.005**american" + 0.004**islam" + 0.004**theology" + 0.004**paul" + 0.004**people" + 0.004**cardinal" + 0.004**nun" + 0.004**united" + 0.004**bible" + 0.004**christ"
29	Police, crimes, news reports	0.032**police" + 0.026**rape" + 0.026**crime" + 0.022**victim" + 0.019**violence" + 0.019**trial" + 0.014**murder" + 0.014**prison" + 0.014**criminal" + 0.014**charge" + 0.013**jury" + 0.013**jones" + 0.012**death" + 0.011**lawyer" + 0.011**prosecutor" + 0.009**officer" + 0.008**killing" + 0.008**assault" + 0.007**accused" + 0.007**defense" + 0.007**attorney" + 0.007**investigation" + 0.007**killed" + 0.007**news" + 0.006**gun" + 0.006**jail" + 0.006**convicted" + 0.006**newspaper" + 0.006**witness" + 0.006**arrest"

A.5: Pittsburgh Post-Gazette LDA Model Topics

Topic #	Topic Name	Word weightings (top 30 words)
0	Shannon Faulkner, first woman allowed to study at The Citadel, a military school (1994) following a lawsuit	0.014*"school" + 0.011*"singlesex" + 0.010*"shaw" + 0.008*"citadel" + 0.006*"faulkner" + 0.005*"cadet" + 0.004*"military" + 0.004*"play" + 0.004*"evita" + 0.004*"college" + 0.004*"allen" + 0.003*"peron" + 0.003*"theater" + 0.003*"aire" + 0.003*"gun" + 0.003*"vmi" + 0.003*"buenos" + 0.003*"class" + 0.003*"woman" + 0.003*"public" + 0.003*"district" + 0.003*"shannon" + 0.003*"yeats" + 0.003*"education" + 0.003*"ibsen" + 0.002*"loaddate" + 0.002*"pfa" + 0.002*"musical" + 0.002*"photo" + 0.002*"beat"
1	Family, but more mixed than purely family words	0.008*"time" + 0.008*"woman" + 0.006*"life" + 0.006*"mother" + 0.006*"people" + 0.005*"child" + 0.005*"movie" + 0.004*"day" + 0.004*"love" + 0.004*"story" + 0.004*"family" + 0.003*"film" + 0.003*"father" + 0.003*"photo" + 0.003*"friend" + 0.003*"daughter" + 0.003*"lot" + 0.003*"loaddate" + 0.003*"character" + 0.003*"feminist" + 0.003*"star" + 0.003*"kid" + 0.003*"critic" + 0.002*"feel" + 0.002*"job" + 0.002*"guy" + 0.002*"language" + 0.002*"son" + 0.002*"gay" + 0.002*"play"
2	Pittsburgh	0.011*"pittsburgh" + 0.009*"family" + 0.006*"business" + 0.005*"director" + 0.005*"company" + 0.005*"city" + 0.005*"money" + 0.005*"school" + 0.004*"photo" + 0.004*"people" + 0.004*"program" + 0.003*"country" + 0.003*"club" + 0.003*"center" + 0.003*"child" + 0.003*"time" + 0.003*"service" + 0.003*"community" + 0.003*"president" + 0.003*"national" + 0.003*"woman" + 0.003*"theater" + 0.003*"father" + 0.002*"heinz" + 0.002*"public" + 0.002*"executive" + 0.002*"local" + 0.002*"percent" + 0.002*"job" + 0.002*"university"
3	Police, charges, arrests	0.028*"police" + 0.013*"county" + 0.012*"yesterday" + 0.009*"officer" + 0.006*"city" + 0.005*"authority" + 0.005*"death" + 0.005*"north" + 0.004*"washington" + 0.004*"car" + 0.004*"department" + 0.004*"killed" + 0.004*"street" + 0.004*"told" + 0.004*"jail" + 0.004*"people" + 0.004*"charge" + 0.004*"charged" + 0.004*"judge" + 0.003*"arrested" + 0.003*"bronte" + 0.003*"gun" + 0.003*"march" + 0.003*"school" + 0.003*"robertson" + 0.003*"prison" + 0.003*"chief" + 0.003*"burger" + 0.003*"student" + 0.003*"killing"
4	Specifically Protestant Christianity, denomination specific (presbyterian, methodist), conferences, etc.	0.066*"church" + 0.015*"god" + 0.014*"woman" + 0.013*"conference" + 0.012*"christian" + 0.011*"rev" + 0.010*"jesus" + 0.009*"pastor" + 0.009*"faith" + 0.009*"presbyterian" + 0.008*"religious" + 0.007*"christ" + 0.007*"ministry" + 0.007*"theological" + 0.006*"religion" + 0.006*"theology" + 0.006*"seminary" + 0.006*"jewish" + 0.005*"prayer" + 0.005*"feminist" + 0.005*"united" + 0.005*"methodist" + 0.005*"people" + 0.005*"reimagining" + 0.005*"pittsburgh" + 0.005*"congregation" + 0.005*"community" + 0.005*"denomination" + 0.005*"bible" + 0.004*"leader"
5	Universities and colleges, with a variety of associated fields (law, medical, health, education)	0.020*"student" + 0.020*"school" + 0.017*"university" + 0.012*"program" + 0.011*"college" + 0.011*"woman" + 0.009*"study" + 0.008*"dr" + 0.007*"professor" + 0.006*"campus" + 0.006*"health" + 0.006*"education" + 0.005*"issue" + 0.005*"child" + 0.005*"care" + 0.005*"class" + 0.005*"people" + 0.005*"gender" + 0.004*"association" + 0.004*"day" + 0.004*"pittsburgh" + 0.004*"faculty" + 0.004*"pitt" + 0.004*"science" + 0.004*"institute" + 0.003*"law" + 0.003*"medical" + 0.003*"teacher" + 0.003*"parent" + 0.003*"society"
6	International art and women painters	0.008*"city" + 0.005*"mexico" + 0.004*"people" + 0.003*"mexican" + 0.003*"day" + 0.003*"rivera" + 0.003*"stewart" + 0.003*"art" + 0.002*"french" + 0.002*"museum" + 0.002*"time" + 0.002*"drink" + 0.002*"american" + 0.002*"photo" + 0.002*"water" + 0.002*"child" + 0.002*"kahlo" + 0.002*"muslim" + 0.002*"frida" + 0.002*"tourist" + 0.002*"plaza" + 0.002*"public" + 0.002*"hour" + 0.002*"oil" + 0.002*"school" + 0.002*"husband" + 0.002*"film" + 0.002*"street" + 0.002*"artist" + 0.002*"life"

7	Army	0.022*"army" + 0.021*"military" + 0.011*"war" + 0.009*"combat" + 0.009*"soldier" + 0.009*"service" + 0.006*"force" + 0.006*"draft" + 0.005*"kennedy" + 0.005*"selective" + 0.004*"unit" + 0.003*"time" + 0.003*"civilian" + 0.003*"serve" + 0.003*"iraq" + 0.003*"troop" + 0.003*"president" + 0.003*"register" + 0.003*"marine" + 0.003*"vietnam" + 0.003*"training" + 0.003*"march" + 0.003*"soviet" + 0.003*"corp" + 0.003*"navy" + 0.003*"defense" + 0.003*"pirate" + 0.003*"afghanistan" + 0.003*"air" + 0.003*"life"
8	Borat (the movie)?	0.004*"public" + 0.004*"jones" + 0.004*"wright" + 0.004*"borat" + 0.004*"bank" + 0.004*"helen" + 0.004*"woman" + 0.003*"black" + 0.003*"life" + 0.003*"talk" + 0.002*"office" + 0.002*"japanese" + 0.002*"japan" + 0.002*"hull" + 0.002*"white" + 0.002*"cohen" + 0.002*"boy" + 0.002*"translation" + 0.002*"director" + 0.002*"loan" + 0.002*"bennett" + 0.002*"titus" + 0.002*"radio" + 0.002*"lender" + 0.002*"week" + 0.002*"imperial" + 0.002*"loaddate" + 0.002*"moll" + 0.002*"girl" + 0.002*"palace"
9	Wuerl was the bishop of Pittsburgh from 1988–2006	0.026*"wuerl" + 0.023*"bishop" + 0.008*"rhodes" + 0.008*"dennis" + 0.006*"scholarship" + 0.005*"carnegie" + 0.005*"connor" + 0.004*"wuerls" + 0.004*"mellon" + 0.004*"university" + 0.004*"school" + 0.004*"pittsburgh" + 0.004*"committee" + 0.004*"rotc" + 0.003*"student" + 0.003*"dalton" + 0.003*"craig" + 0.003*"scholar" + 0.003*"helicopter" + 0.003*"meeting" + 0.003*"physic" + 0.003*"day" + 0.003*"week" + 0.003*"air" + 0.003*"tango" + 0.002*"hospital" + 0.002*"majetich" + 0.002*"mancano" + 0.002*"mass" + 0.002*"grammys"
10	Food and cooking	0.014*"food" + 0.013*"barbie" + 0.011*"cup" + 0.009*"book" + 0.008*"doll" + 0.007*"add" + 0.007*"cook" + 0.007*"oil" + 0.007*"minute" + 0.006*"teaspoon" + 0.006*"egg" + 0.006*"recipe" + 0.005*"tomato" + 0.005*"tablespoon" + 0.005*"day" + 0.005*"fresh" + 0.005*"salt" + 0.005*"onion" + 0.005*"chopped" + 0.005*"people" + 0.005*"zucchini" + 0.004*"smith" + 0.004*"taste" + 0.004*"pepper" + 0.004*"olive" + 0.004*"chicken" + 0.004*"dish" + 0.004*"cooking" + 0.003*"meditation" + 0.003*"garlic"
11	Songs, artists, hit albums	0.013*"song" + 0.009*"artist" + 0.007*"album" + 0.005*"photo" + 0.004*"add" + 0.004*"exhibition" + 0.004*"art" + 0.003*"time" + 0.003*"musical" + 0.003*"color" + 0.002*"video" + 0.002*"lyric" + 0.002*"pop" + 0.002*"music" + 0.002*"piece" + 0.002*"water" + 0.002*"quilt" + 0.002*"red" + 0.002*"center" + 0.002*"love" + 0.002*"feminist" + 0.002*"pittsburgh" + 0.002*"hit" + 0.002*"annual" + 0.002*"aap" + 0.002*"vocal" + 0.002*"image" + 0.002*"spirit" + 0.002*"graphic" + 0.002*"loaddate"
12	Presidential elections, but a lot more names of participants, and more republican names (bush, palin, mccain, cheney)	0.017*"gore" + 0.009*"specter" + 0.008*"bush" + 0.008*"palin" + 0.006*"mccain" + 0.005*"bradley" + 0.005*"republican" + 0.005*"palmer" + 0.005*"people" + 0.005*"gorman" + 0.004*"cheney" + 0.004*"political" + 0.004*"campaign" + 0.004*"american" + 0.004*"palins" + 0.004*"presidential" + 0.004*"time" + 0.003*"candidate" + 0.003*"video" + 0.003*"hampshire" + 0.003*"john" + 0.003*"issue" + 0.003*"vice" + 0.003*"day" + 0.003*"gina" + 0.003*"week" + 0.002*"party" + 0.002*"life" + 0.002*"child" + 0.002*"goldman"
13	Clinton and hillary, but a lot of description on their roles and genders	0.049*"woman" + 0.008*"clinton" + 0.008*"child" + 0.007*"time" + 0.006*"president" + 0.006*"feminist" + 0.005*"husband" + 0.004*"life" + 0.004*"family" + 0.004*"people" + 0.004*"mother" + 0.003*"loaddate" + 0.003*"political" + 0.003*"hillary" + 0.003*"issue" + 0.003*"candidate" + 0.003*"female" + 0.003*"care" + 0.003*"week" + 0.003*"day" + 0.003*"country" + 0.003*"wife" + 0.003*"campaign" + 0.003*"percent" + 0.003*"gender" + 0.003*"american" + 0.003*"house" + 0.002*"national" + 0.002*"school" + 0.002*"note"
14	Benjamin Chavis-NAACP CEO in 1993	0.015*"rowing" + 0.012*"naacp" + 0.011*"chavis" + 0.011*"board" + 0.009*"river" + 0.007*"team" + 0.007*"director" + 0.007*"people" + 0.006*"sport" + 0.006*"girl" + 0.006*"woman" + 0.005*"school" + 0.005*"organization" + 0.005*"boathouse" + 0.004*"row" + 0.004*"college" + 0.004*"pittsburgh" + 0.004*"boat" + 0.004*"city" + 0.004*"executive" + 0.003*"ames" + 0.003*"black" + 0.003*"water" + 0.003*"farrakhan" + 0.003*"meeting" + 0.003*"time" + 0.003*"robert" + 0.003*"level" + 0.003*"morris" + 0.003*"program"

15	Sex scandals in the catholic church	0.021*"boy" + 0.020*"girl" + 0.015*"sexual" + 0.014*"catholic" + 0.012*"vatican" + 0.011*"pope" + 0.010*"john" + 0.010*"bishop" + 0.009*"priest" + 0.009*"school" + 0.008*"paul" + 0.008*"sex" + 0.007*"harassment" + 0.007*"human" + 0.005*"rome" + 0.005*"witch" + 0.004*"woman" + 0.004*"teacher" + 0.004*"difference" + 0.003*"people" + 0.003*"feminism" + 0.003*"cardinal" + 0.003*"nun" + 0.003*"bosco" + 0.003*"report" + 0.003*"latin" + 0.003*"diocese" + 0.003*"behavior" + 0.003*"language" + 0.002*"abuse"
16	College sports	0.029*"team" + 0.028*"school" + 0.025*"sport" + 0.018*"title" + 0.016*"ix" + 0.016*"woman" + 0.015*"coach" + 0.011*"athletic" + 0.011*"athlete" + 0.010*"player" + 0.010*"college" + 0.010*"basketball" + 0.010*"game" + 0.009*"program" + 0.009*"university" + 0.008*"girl" + 0.008*"men" + 0.008*"golf" + 0.007*"female" + 0.007*"football" + 0.007*"play" + 0.007*"opportunity" + 0.005*"season" + 0.005*"ncaa" + 0.005*"championship" + 0.005*"national" + 0.005*"education" + 0.005*"tournament" + 0.005*"field" + 0.005*"federal"
17	Books and writing	0.035*"book" + 0.010*"life" + 0.010*"writer" + 0.008*"author" + 0.007*"story" + 0.006*"writing" + 0.006*"read" + 0.006*"feminist" + 0.006*"people" + 0.005*"woman" + 0.005*"american" + 0.005*"reader" + 0.005*"magazine" + 0.005*"father" + 0.005*"history" + 0.004*"time" + 0.004*"reading" + 0.004*"published" + 0.004*"culture" + 0.004*"editor" + 0.004*"loaddate" + 0.003*"friedan" + 0.003*"study" + 0.003*"writes" + 0.003*"family" + 0.003*"list" + 0.003*"boy" + 0.003*"fiction" + 0.003*"poet" + 0.003*"century"
18	Plays, but more generally	0.005*"play" + 0.005*"music" + 0.005*"time" + 0.004*"star" + 0.004*"night" + 0.004*"photo" + 0.004*"life" + 0.003*"film" + 0.003*"loaddate" + 0.003*"love" + 0.003*"band" + 0.003*"movie" + 0.003*"people" + 0.002*"actor" + 0.002*"day" + 0.002*"character" + 0.002*"story" + 0.002*"american" + 0.002*"set" + 0.002*"city" + 0.002*"black" + 0.002*"john" + 0.002*"stage" + 0.002*"song" + 0.002*"hit" + 0.002*"audience" + 0.002*"graphic" + 0.002*"dance" + 0.002*"feminist" + 0.002*"director"
19	National politics	0.013*"abortion" + 0.010*"right" + 0.010*"president" + 0.007*"republican" + 0.007*"government" + 0.006*"court" + 0.006*"law" + 0.006*"political" + 0.006*"issue" + 0.006*"national" + 0.006*"conservative" + 0.006*"people" + 0.005*"clinton" + 0.005*"american" + 0.005*"bush" + 0.005*"senate" + 0.004*"woman" + 0.004*"clinic" + 0.004*"congress" + 0.004*"robert" + 0.004*"party" + 0.004*"democratic" + 0.004*"house" + 0.004*"supreme" + 0.004*"democrat" + 0.004*"campaign" + 0.004*"white" + 0.004*"yesterday" + 0.004*"leader" + 0.003*"loaddate"
20	Islam/muslim people, and some reference to schools	0.022*"woman" + 0.011*"muslim" + 0.007*"team" + 0.007*"girl" + 0.005*"award" + 0.005*"quinn" + 0.005*"islamic" + 0.005*"hockey" + 0.004*"islam" + 0.004*"head" + 0.004*"school" + 0.003*"college" + 0.003*"time" + 0.003*"american" + 0.003*"prize" + 0.003*"irish" + 0.003*"female" + 0.003*"boy" + 0.003*"veil" + 0.003*"country" + 0.003*"player" + 0.003*"play" + 0.003*"people" + 0.002*"send" + 0.002*"day" + 0.002*"afghanistan" + 0.002*"loaddate" + 0.002*"lady" + 0.002*"feminist" + 0.002*"ice"
21	Exhibitions in Pittsburgh	0.022*"art" + 0.015*"pittsburgh" + 0.013*"artist" + 0.009*"museum" + 0.007*"exhibition" + 0.007*"gallery" + 0.007*"carnegie" + 0.006*"center" + 0.004*"university" + 0.004*"nov" + 0.004*"saturday" + 0.004*"photo" + 0.004*"painting" + 0.004*"event" + 0.004*"theater" + 0.004*"festival" + 0.003*"contemporary" + 0.003*"sat" + 0.003*"sunday" + 0.003*"series" + 0.003*"history" + 0.003*"noon" + 0.003*"international" + 0.003*"friday" + 0.003*"oct" + 0.003*"woman" + 0.003*"american" + 0.003*"park" + 0.003*"ave" + 0.003*"tour"
22	Court, judiciary	0.025*"court" + 0.014*"discrimination" + 0.014*"law" + 0.010*"judge" + 0.009*"sexual" + 0.009*"sex" + 0.008*"federal" + 0.008*"justice" + 0.008*"employee" + 0.007*"lawsuit" + 0.007*"lawyer" + 0.007*"woman" + 0.006*"attorney" + 0.006*"harassment" + 0.006*"suit" + 0.006*"complaint" + 0.006*"filed" + 0.006*"job" + 0.005*"district" + 0.005*"supreme" + 0.005*"loaddate" + 0.005*"decision" + 0.005*"company" + 0.005*"claim" + 0.004*"yesterday" + 0.004*"county" + 0.004*"school" + 0.004*"legal" + 0.004*"pay" + 0.004*"office"

23	Family and career and gender roles	0.065*"woman" + 0.021*"percent" + 0.009*"child" + 0.007*"female" + 0.006*"male" + 0.006*"job" + 0.006*"study" + 0.006*"time" + 0.005*"family" + 0.005*"marriage" + 0.005*"feminist" + 0.004*"force" + 0.004*"life" + 0.004*"career" + 0.003*"survey" + 0.003*"gender" + 0.003*"people" + 0.003*"school" + 0.003*"loaddate" + 0.003*"american" + 0.003*"age" + 0.003*"social" + 0.003*"mother" + 0.003*"decade" + 0.002*"day" + 0.002*"parent" + 0.002*"married" + 0.002*"role" + 0.002*"divorce" + 0.002*"hour"
24	? names	0.017*"hooter" + 0.011*"yard" + 0.011*"station" + 0.007*"molly" + 0.006*"channel" + 0.005*"tom" + 0.005*"restaurant" + 0.005*"hritz" + 0.005*"now" + 0.005*"plath" + 0.005*"cain" + 0.004*"nbc" + 0.004*"phillips" + 0.004*"column" + 0.004*"matthew" + 0.004*"miniseries" + 0.004*"hughes" + 0.004*"chain" + 0.004*"cbs" + 0.004*"guaracha" + 0.004*"saffron" + 0.003*"burrow" + 0.003*"reinhold" + 0.003*"minnie" + 0.003*"waitress" + 0.003*"nwangaza" + 0.003*"ireland" + 0.002*"store" + 0.002*"mee" + 0.002*"wqeds"
25	Politics, with some more names (santorum, casey, limbaugh, clinton)	0.017*"santorum" + 0.010*"guy" + 0.008*"conservative" + 0.007*"political" + 0.006*"liberal" + 0.006*"republican" + 0.006*"limbaugh" + 0.006*"book" + 0.005*"people" + 0.004*"casey" + 0.004*"rick" + 0.004*"family" + 0.004*"power" + 0.004*"male" + 0.004*"word" + 0.004*"term" + 0.004*"democrat" + 0.003*"santorums" + 0.003*"sen" + 0.003*"democratic" + 0.003*"woman" + 0.003*"senator" + 0.003*"correctness" + 0.003*"take" + 0.003*"child" + 0.002*"passage" + 0.002*"culture" + 0.002*"russia" + 0.002*"clinton" + 0.002*"america"
26	?	0.022*"promise" + 0.019*"keeper" + 0.007*"mailer" + 0.005*"cloning" + 0.005*"black" + 0.005*"bernard" + 0.005*"organization" + 0.003*"harmony" + 0.003*"piaa" + 0.003*"oates" + 0.003*"conference" + 0.003*"ticket" + 0.003*"people" + 0.003*"stone" + 0.003*"brazil" + 0.002*"football" + 0.002*"owen" + 0.002*"game" + 0.002*"band" + 0.002*"wife" + 0.002*"event" + 0.002*"square" + 0.002*"embryo" + 0.002*"ambridge" + 0.002*"durbin" + 0.002*"school" + 0.002*"house" + 0.002*"nascar" + 0.002*"sunday" + 0.002*"mall"
27	Women's fashion	0.035*"woman" + 0.010*"girl" + 0.005*"people" + 0.005*"feminist" + 0.004*"day" + 0.004*"time" + 0.004*"photo" + 0.003*"loaddate" + 0.003*"hair" + 0.003*"sex" + 0.003*"fashion" + 0.003*"body" + 0.003*"magazine" + 0.003*"wear" + 0.003*"life" + 0.003*"female" + 0.003*"breast" + 0.003*"black" + 0.002*"feel" + 0.002*"mother" + 0.002*"shoe" + 0.002*"love" + 0.002*"lot" + 0.002*"dress" + 0.002*"model" + 0.002*"wearing" + 0.002*"male" + 0.002*"miller" + 0.002*"child" + 0.002*"sexual"
28	1995 women's conference hosted in Beijing	0.034*"woman" + 0.021*"conference" + 0.017*"beijing" + 0.008*"document" + 0.007*"delegate" + 0.006*"company" + 0.005*"united" + 0.004*"nation" + 0.004*"alcott" + 0.004*"theater" + 0.004*"session" + 0.003*"week" + 0.003*"female" + 0.003*"china" + 0.003*"committee" + 0.003*"violence" + 0.003*"hickton" + 0.003*"solanas" + 0.003*"role" + 0.003*"local" + 0.003*"play" + 0.003*"video" + 0.003*"right" + 0.003*"pittsburgh" + 0.003*"mccarthy" + 0.003*"allen" + 0.003*"wynn" + 0.003*"nongovernmental" + 0.002*"warhol" + 0.002*"song"
29	Mix of marriage, religion, LGBTQ	0.012*"marriage" + 0.010*"mccartney" + 0.009*"catholic" + 0.008*"county" + 0.006*"religious" + 0.006*"christian" + 0.006*"people" + 0.005*"medical" + 0.005*"truth" + 0.005*"promise" + 0.005*"plan" + 0.004*"samesex" + 0.004*"public" + 0.004*"god" + 0.004*"gay" + 0.004*"church" + 0.004*"hawaii" + 0.004*"keeper" + 0.003*"examiner" + 0.003*"law" + 0.003*"allegheeny" + 0.003*"reason" + 0.003*"pittsburgh" + 0.003*"property" + 0.003*"tax" + 0.003*"parliament" + 0.003*"larsen" + 0.003*"view" + 0.003*"moral" + 0.003*"value"

Appendix B

nltk collocates

B.1: nltk collocates 1980-1989

	collocates	Score (likelihood ratio)
1	('radical', 'feminism')	211.40823
2	('modern', 'feminism')	110.028134
3	('rise', 'feminism')	108.688635
4	('feminism', 'femininity')	100.200489
5	('wave', 'feminism')	85.1692345
6	('militant', 'feminism')	82.4704691
7	('socialism', 'feminism')	69.9124613
8	('marxism', 'feminism')	68.75471
9	('brand', 'feminism')	68.4767256
10	('ecology', 'feminism')	63.7234548
11	('feminism', 'oped')	59.2395947
12	('feminism', 'marxism')	59.1796944
13	('feminism', 'pacifism')	56.4287293
14	('pacifism', 'feminism')	45.0592912
15	('anticolonialism', 'feminism')	44.7649793
16	('feminism', 'ecology')	41.9500632
17	('humanism', 'feminism')	41.8263786
18	('byproduct', 'feminism')	39.6912326
19	('failure', 'feminism')	39.2939801
20	('mormonism', 'feminism')	36.4506623

B.2: nltk collocates 1990-1999

	collocates	Score (likelihood ratio)
1	('radical', 'feminism')	935.02396
2	('stole', 'feminism')	637.722486
3	('backlash', 'feminism')	357.650098
4	('modern', 'feminism')	318.921769
5	('feminism', 'dead')	269.429142
6	('wave', 'feminism')	249.324319
7	('brand', 'feminism')	188.318708
8	('rise', 'feminism')	179.937727
9	('militant', 'feminism')	173.582348
10	('contemporary', 'feminism')	166.670697
11	('fear', 'feminism')	165.024319
12	('feminism', 'illusions')	151.503027
13	('feminism', 'campus')	124.158204
14	('secondwave', 'feminism')	93.2755253
15	('professing', 'feminism')	86.333342
16	('american', 'feminism')	85.6798807
17	('feminism', 'alive')	83.2579082
18	('word', 'feminism')	82.8930485
19	('feminism', 'radical')	82.5034467
20	('feminism', 'homosexuality')	80.953403

B.3: nltk collocates 2000-2009

	collocates	Score (likelihood ratio)
1	('radical', 'feminism')	498.051951
2	('wave', 'feminism')	318.521582
3	('secondwave', 'feminism')	309.527104
4	('feminism', 'harming')	253.147895
5	('rise', 'feminism')	229.805854
6	('firstwave', 'feminism')	180.134377
7	('modern', 'feminism')	177.122105
8	('misguided', 'feminism')	175.212792
9	('movement', 'feminism')	129.996455
10	('feminism', 'gay')	115.263287
11	('feminism', 'dead')	112.377037
12	('homosexuality', 'feminism')	94.4840697
13	('brand', 'feminism')	92.4307736
14	('feminism', 'capital')	90.0529219
15	('rights', 'feminism')	80.8432556
16	('triggers', 'feminism')	79.4621933
17	('thatcherera', 'feminism')	76.4037384
18	('mids', 'feminism')	74.8148877
19	('feminism', 'intense')	68.2828988
20	('feminism', 'marxism')	65.0868899

B.4 nltk collocates (The New York Times)

	collocates	Score (likelihood ratio)
1	('radical', 'feminism')	522.5801031636960
2	('secondwave', 'feminism')	326.38058145707000
3	('wave', 'feminism')	295.5118969517540
4	('backlash', 'feminism')	283.6745997696130
5	('stole', 'feminism')	259.36216274828400
6	('rise', 'feminism')	243.05562743286300
7	('militant', 'feminism')	200.3701328948530
8	('modern', 'feminism')	198.2609119617300
9	('firstwave', 'feminism')	184.37753600860700
10	('brand', 'feminism')	139.82660576858700
11	('feminism', 'gay')	105.31872109322300
12	('pacifism', 'feminism')	99.59203589431210
13	('feminism', 'marxism')	94.91767694753340
14	('marxism', 'feminism')	94.91767694753340
15	('contemporary', 'feminism')	93.89258694928070
16	('movement', 'feminism')	83.60263202668890
17	('feminism', 'femininity')	82.56121277852190
18	('fear', 'feminism')	81.45637382831910
19	('decades', 'feminism')	79.135580970788
20	('triggers', 'feminism')	78.10760708593010

B.5 nltk collocates (Pittsburgh Post-Gazette)

	collocates	Score (likelihood ratio)
1	('radical', 'feminism')	120.42644891060600
2	('rise', 'feminism')	51.81929022135450
3	('virginity', 'feminism')	43.015190823955
4	('wave', 'feminism')	41.81512287878060
5	('western', 'feminism')	41.47330314338170
6	('islamic', 'feminism')	35.775121240789400
7	('feminism', 'harming')	32.719515406411300
8	('homosexuality', 'feminism')	30.292739405694400
9	('feminism', 'dead')	29.874107126150600
10	('word', 'feminism')	29.768928086567800
11	('feminism', 'turkey')	29.71547178799040
12	('feminism', 'globalization')	28.743164346628200
13	('blow', 'feminism')	28.201702017016700
14	('brand', 'feminism')	27.437353500046700
15	('progressive', 'feminism')	26.015057334827400
16	('stole', 'feminism')	24.698420200335800
17	('version', 'feminism')	24.688604301326400
18	('misguided', 'feminism')	24.293668739763200
19	('feminism', 'encourages')	22.31704643001630
20	('feminism', 'grown')	22.302328286549600