

USING FRIDAY SERMONS AS DATA

Ihsan Kahveci

BACKGROUND AND CONTEXT

Friday Sermons (Khutbe) is a one-two page long text that includes certain verses from Quran and Hadith plus a weekly agenda or information for the Muslim society. Those agendas range from daily practices like personal hygiene to popular issues like Black Friday. Hutbe texts must be read out loud by imams during the Friday pray **in all of the 88000 mosques in Turkey and it is enacted by law that imams must stay on the text.**

Presidency of Religious Affairs (Diyanet Isleri Baskanligi a.k.a Diyanet) is a state institution in founded on May 3, 1924 after the Kemalist Secular revolution in Turkey. The aim was to prevent Arabic influence on Turkish Muslim majority by nationalizing Islamic doctrine and practices. It is one of the biggest government institutions with a budget of almost \$2 billion. **One of the main objects of Diyanet is to prepare and distribute Friday Sermons (hutbe) to the imams of all the mosques in Turkey.**

A study undertaken by the DIB in 2014 with a sample size of 21,632 individuals and randomly selected based on census data revealed that **56% percent of the male population reported regularly attending to Friday prayers (DIB, 2014)**. This amount increased to 65% in rural areas. In addition, 15% percent of the population stated that they are attending Friday prayers in an irregular basis.

There are other Islamic institutions across the world that are also responsible for preparation and distribution of the Friday sermons such as Ahmadiyya Muslim Community in United Kingdom and Islamic Society of Boston Cultural Center (ISBCC) in US, both are long-established NGOs and heavily affiliated with the Muslim community in their regions. There are also other state institutions like General Authority of Islamic Affairs & Endowments in United Arab Emirates and Alharamain in Saudi Arabia.

OBJECTIVES:

The broader aim of my project:

I am planning to conduct a comparative content analysis research on the Hutbe texts produced by organizations above and hopefully more. The requirements for a sermon archive to be included in the dataset is whether they are produced on a weekly basis over 10-year period. The second requirement is the availability of English version or long summary.

The comparison will be made around the question below:

- Does the weekly content respond to key political events regarding Muslims in the region?
- Differences in responding to world-wide events like 9/11 or Charlie Hebdo?
- Is there a trend towards and against radicalization?

Hypotheses:

- H1: There will be a distinctive topic that represents nationalism.
- H2: The prevalence of nationalism topic will increase over time.

The goals for this course:

- Using LDA Topic Modelling to come up with distinctive topics.
- Plot graphs to understand change over time.
- Test my hypotheses.
- Basically, writing the first draft of the methodology chapter.

Data Preparation

This is my sample dataset that I scraped from Diyanet's website:

```
In [2]: documents = pd.read_csv('sermons_df', sep=';', index_col=0)
documents
```

```
Out[2]:
```

	date	sermon
0	2019-02-01	\nDATE: 01.02.2019\n[pic]RIFQ (GENTLENE...
1	2018-03-02	\nLOCATION: NATIONWIDE\nDATE: 02.03.2018...
2	2019-12-27	\nDATE: 27.12.2019\n ...
3	2017-08-25	LOCATION\t: NATIONWIDE\n\n\tDATE\t: 25.08.2017...
4	2019-10-25	\nDATE: 25.10.2019\n[pic]\n\n FOR...
...
152	2019-04-12	\nDATE: 12.04.2019\n ...
153	2017-09-01	\nLOCATION: NATIONWIDE\nDATE: 01.09.2017...
154	2017-02-03	\n LOCATION: NATIONWIDE\n DATE: 03....
155	2019-08-23	\nDATE: 23.08.2019\n[pic]\n SOCIAL HARMS OF...
156	2019-11-29	\nDATE: 29.11.2019\n[pic]\n SINCERITY: THE ...

157 rows × 2 columns

```
In [5]: print("\033[1m" + "Original 57th Sermon: " + "\033[0;0m")
print(' ')
print(doc_sample)
```

Original 57th Sermon:

LOCATION : NATIONWIDE
DATE : 13.01.2017

[pic]

TO BE ABLE TO BE A SERVANT ALLAH LOVES

A Blessed Friday to You, Brothers and Sisters!

In the verse I have recited, Allah Almighty shows us the way to be honored with His love and favor: "Say, [O Muhammad], 'If you should love Allah, then follow me, so Allah will love you and forgive you your sins. And Allah is Forgiving and Merciful.'"[i]

Brothers and Sisters!

The Almighty stated the deeds that will move us away from His grace in His Book like He showed us in it the ones that will bring us to His favor. He stated the ones honored with His love as well as those left without His love and mercy. Come, let's hear in today's khutbah who the servants our Lord loves are.

????????? ??????? ?????????????????? Allah loves those pure in the spiritual and material sense. Then come, let's protect our disposition against all evil. Let's make our mind and our heart the center of goodness and nice things, not of evil and ugliness.

????? ??????? ??????? ?????????????????? Allah loves those who repent. Then, honorable brothers and sisters! Come, let's be the servants we are and repent. Let's voice our submission to our Lord and the remorse for our sins. Let's take refuge in His vast mercy. Let's not forget that repentance is like a new start to life.

????? ??????? ??????? ?????????????????? Allah loves those who avoid disobey Him. Then come, let's always be aware of our duties and obligations towards Him. Let's live our lives in accordance with the purpose and reason of our

him. Let's live our lives in accordance with the purpose and reason of our creation.

????? ??????? ??????? ?????????????????? Allah loves those who trust Him. Then come, after doing our duty, ?????????? ??????? ?????????? ?????????? let's say "Allah suffices, what a great trustee He is!" Let's never lose hope in our Lord's grace, help, and blessing.

????????? ??????? ?????????????????? Allah loves the doers of good and those who seek His favor in everything they say and do. Then come, brothers and sisters! Let's act with the awareness that our Lord sees us at every moment.

????????? ??????? ?????????????????? Allah loves the patient. He is with the patient. Then come, let's adopt patience and quiet as a maxim with the awareness that life is a test.

????? ??????? ??????? ?????????????????? Allah loves who are just. Then come, let's elevate justice at all times. Let's not forget that justice is the key to trust and peace. It is indispensable for a humane life.

Dear Brothers and Sisters!

Come, now look at those the Almighty does not love.

????? ??????? ??? ??????? ?????????????????? Allah does not love aggressors, those who cross the line. Then, we must be constructive and unifying, not destructive and divisive.

????????? ??? ??????? ?????????????????? Allah does not love the oppressors. Because oppression is one of the biggest sins. It means disregarding what's right and just. Then, we must always stand against the oppressors and with the oppressed.

????????? ??? ??????? ?????????????????? Allah does not love the prodigals. Then, we must avoid squandering what we have earned and lavishly using the blessings Allah has bestowed upon us.

????????? ??? ??????? ?????????????????? Allah does not love those who spoil and make trouble. Then, we must altogether protect our peace, our unity and togetherness, our brotherhood and relationship.

????????? ??? ??????? ?????????????????? Allah does not love the arrogant, those who boast. Then, we must elevate with humility and solemnity.

????? ??????? ??? ??????? ?????????????????? Allah never loves those who betray Himself, His Holy Book, His Prophet, their own values, and their own brothers. Then, we must stay away from all sorts of betrayal, hypocrisy, and deception.

????????? ??? ??????? ?????? ?????????? ??????? Allah does not love anybody who persists in impiety and sin. Then, we must avoid being blind and

oblivious to the countless blessings bestowed upon us by the Almighty and the goodness we have been blessed with.

Brothers and Sisters!

May the Almighty Lord grant us to be among those happy servants of His who have been honored with His love. As I conclude my khutbah, I would like to share with you this piece of information: As they always do, our great nation has shown great favor to the aid campaign launched by the Presidency of Religious Affairs a short while ago with the slogan "Don't Let Humanity Die in Aleppo." Under the campaign, an aid of 133,745,000 TL has been raised so far from abroad and within the country. The raised fund is sent to our oppressed brothers and sisters in Aleppo to tend to their wounds. May the Almighty accept your help. May He always protect our country and our great nation that are the hope of the oppressed, the downtrodden, and the destitute.

[i] Ali 'Imran, 3/31.

Written by: General Directorate of Religious
Services

```
In [6]: print("\033[1m" + "Tokenized and Lemmatized 57th Sermon: " + "\033[0;0m")
print(' ')
print(preprocess(doc_sample))
```

Tokenized and Lemmatized 57th Sermon:

```
['abl', 'servant', 'allah', 'love', 'bless', 'friday', 'brother', 'sister', 'vers', 'r', 'r', 'love', 'favor', 'muhammad', 'love', 'allah', 'follow', 'allah', 'love', 'forgiv', 'er', 'sister', 'almighti', 'state', 'deed', 'away', 'grace', 'book', 'like', 'show', 'onor', 'love', 'leav', 'love', 'merci', 'come', 'hear', 'today', 'khutbah', 'servant', 'spiritu', 'materi', 'sens', 'come', 'protect', 'disposit', 'evil', 'mind', 'heart', 'ugli', 'allah', 'love', 'repent', 'honor', 'brother', 'sister', 'come', 'servant', 'r', 's', 'sin', 'refug', 'vast', 'merci', 'forget', 'repent', 'like', 'start', 'life', 'all', 'awar', 'duti', 'oblig', 'live', 'live', 'accord', 'purpos', 'reason', 'creation', 'allah', 'suffic', 'great', 'truste', 'lose', 'hope', 'lord', 'grace', 'help', 'bless', 'favor', 'come', 'brother', 'sister', 'awar', 'lord', 'see', 'moment', 'allah', 'love', 'patien', 'quiet', 'maxim', 'awar', 'life', 'test', 'allah', 'love', 'come', 'elev', 'ust', 'peac', 'indispens', 'human', 'life', 'dear', 'brother', 'sister', 'come', 'look', 'ggressor', 'cross', 'line', 'construct', 'unifi', 'destruct', 'divis', 'allah', 'love', 'n', 'mean', 'disregard', 'right', 'stand', 'oppressor', 'oppress', 'allah', 'love', 'p', 'sh', 'bless', 'allah', 'bestow', 'allah', 'love', 'spoil', 'troubl', 'altogeth', 'prot', 'ood', 'relationship', 'allah', 'love', 'arrog', 'boast', 'elev', 'humil', 'solemn', 'a', 'prophet', 'valu', 'brother', 'stay', 'away', 'sort', 'betray', 'hypocrisi', 'decept', 'impieti', 'avoid', 'blind', 'oblivi', 'countless', 'bless', 'bestow', 'almighti', 'g', 'hti', 'lord', 'grant', 'happi', 'servant', 'honor', 'love', 'conclud', 'khutbah', 'lik', 'ation', 'show', 'great', 'favor', 'campaign', 'launch', 'presid', 'religi', 'affair', 'paign', 'rais', 'abroad', 'countri', 'rais', 'fund', 'send', 'oppress', 'brother', 'si', 'i', 'accept', 'help', 'protect', 'countri', 'great', 'nation', 'hope', 'oppress', 'dow', 'general', 'director', 'religi', 'servic']
```

Bag of Words

In the end we created a bag of words corpus of each sermon documents. If we go back to our sample sermon which is the 57th sermon, we can read the bag of words matrix for as below:

```
In [13]: bow_doc_57 = bow_corpus[57]

for i in range(len(bow_doc_57)):
    print("Word {} (\\"{}\\") appears {} time.".format(bow_doc_57[i][0],
                                                         dictionary[bow_doc_57[i][0]],
                                                         bow_doc_57[i][1]))
```

```
Word 3 ("affair") appears 1 time.
Word 8 ("avoid") appears 3 time.
Word 9 ("away") appears 2 time.
Word 15 ("bestow") appears 2 time.
Word 32 ("construct") appears 1 time.
Word 37 ("deed") appears 1 time.
Word 46 ("evil") appears 2 time.
Word 55 ("forgiv") appears 2 time.
Word 61 ("grace") appears 2 time.
Word 71 ("imran") appears 1 time.
Word 81 ("mind") appears 1 time.
Word 83 ("nation") appears 2 time.
Word 89 ("oppress") appears 4 time.
Word 92 ("patient") appears 2 time.
Word 107 ("seek") appears 1 time.
Word 115 ("stand") appears 1 time.
Word 121 ("today") appears 1 time.
Word 130 ("accept") appears 1 time.
Word 131 ("accord") appears 1 time.
Word 140 ("awar") appears 3 time.
Word 161 ("decept") appears 1 time.
Word 166 ("duti") appears 2 time.
Word 167 ("earn") appears 1 time.
Word 203 ("look") appears 1 time.
Word 233 ("servant") appears 4 time.
Word 234 ("short") appears 1 time.
Word 240 ("spiritu") appears 1 time.
Word 249 ("thing") appears 1 time.
Word 254 ("trust") appears 2 time.
Word 265 ("abl") appears 1 time.
```


Word 276 ("brotherhood") appears 1 time.
Word 304 ("forget") appears 2 time.
Word 306 ("happi") appears 1 time.
Word 309 ("help") appears 2 time.
Word 316 ("leav") appears 1 time.
Word 324 ("moment") appears 1 time.
Word 334 ("one") appears 2 time.
Word 350 ("purpos") appears 1 time.
Word 360 ("repent") appears 3 time.
Word 366 ("sens") appears 1 time.
Word 368 ("share") appears 1 time.
Word 369 ("sin") appears 3 time.
Word 374 ("submiss") appears 1 time.
Word 393 ("abroad") appears 1 time.
Word 405 ("countri") appears 2 time.
Word 412 ("destitut") appears 1 time.
Word 417 ("downtrodden") appears 1 time.
Word 424 ("friday") appears 1 time.
Word 431 ("inform") appears 1 time.
Word 435 ("lose") appears 1 time.
Word 450 ("see") appears 1 time.
Word 453 ("show") appears 3 time.
Word 469 ("anybodi") appears 1 time.
Word 475 ("bring") appears 1 time.
Word 489 ("great") appears 4 time.
Word 496 ("hope") appears 2 time.
Word 508 ("oppressor") appears 2 time.
Word 517 ("refug") appears 1 time.
Word 553 ("wound") appears 1 time.
Word 559 ("book") appears 2 time.
Word 560 ("center") appears 1 time.
Word 568 ("creation") appears 1 time.
Word 579 ("grant") appears 1 time.
Word 588 ("khutbah") appears 2 time.
Word 598 ("oblig") appears 1 time.
Word 603 ("rais") appears 2 time.
Word 606 ("reason") appears 1 time.
Word 611 ("send") appears 1 time.

Word 653 ("hear") appears 1 time.
Word 660 ("patienc") appears 1 time.
Word 678 ("stay") appears 1 time.
Word 692 ("arrog") appears 1 time.
Word 716 ("materi") appears 1 time.
Word 739 ("conclud") appears 1 time.
Word 746 ("destruct") appears 1 time.
Word 751 ("disposit") appears 1 time.
Word 752 ("disregard") appears 1 time.
Word 766 ("justic") appears 2 time.
Word 769 ("line") appears 1 time.
Word 782 ("relationship") appears 1 time.
Word 785 ("start") appears 1 time.
Word 794 ("altogeth") appears 1 time.
Word 808 ("favor") appears 4 time.
Word 820 ("piec") appears 1 time.
Word 833 ("uniti") appears 1 time.
Word 861 ("test") appears 1 time.
Word 882 ("remors") appears 1 time.
Word 884 ("voic") appears 1 time.
Word 891 ("adopt") appears 1 time.
Word 903 ("presid") appears 1 time.
Word 908 ("squander") appears 1 time.
Word 959 ("troubl") appears 1 time.
Word 966 ("countless") appears 1 time.
Word 988 ("indispens") appears 1 time.
Word 1004 ("pure") appears 1 time.
Word 1023 ("persist") appears 1 time.
Word 1028 ("spoil") appears 1 time.
Word 1037 ("divis") appears 1 time.
Word 1050 ("togeth") appears 1 time.
Word 1053 ("write") appears 1 time.
Word 1071 ("vast") appears 1 time.
Word 1083 ("nice") appears 1 time.
Word 1091 ("biggest") appears 1 time.
Word 1092 ("boast") appears 1 time.
Word 1093 ("elev") appears 2 time.
Word 1117 ("hypocrisi") appears 1 time.

Deciding number of topics:

After the initial trial I found out the range of k (amount of topics) is between 5 and 20. Then, I run an LDA algorithm for each value from 5 to 20 and compared the equations. For the purpose of better presentation I only added the outputs between 6 and 10 here. If you look closely the topic that represents nationalism and how it changed as k increases, you will rease two important things:

- With $k = 6$, we have a topic that includes a lot of keywords of nationalism but also some major religious keywords.
- With $k = 10$, we have generated more than one topic that represents nationalism and we also have saturated our categories with bunch of additional religious topics.
- With $k = 8$, we reached the optimum point where our topics are very distinctive and the nationalism topic includes very important keywords like "jihad" and "canakkale", latter is one of the most important historical wars in the modern Turkish history

```
In [15]: findK(num_topics=range(6,11,1))
```

Number of topics: 6

Topic: 1

Words: 0.012*"night" + 0.012*"forgiv" + 0.012*"repent" + 0.010*"deed" + 0.010*"truth"
"trade" + 0.008"pbuh" + 0.008*"sin"

Topic: 2

Words: 0.029*"month" + 0.028*"ramadan" + 0.016*"sacrific" + 0.011*"fast" + 0.011*"reac
007*"year" + 0.007*"night" + 0.007*"need"

Topic: 3

Words: 0.033*"mosqu" + 0.025*"children" + 0.024*"famili" + 0.020*"knowledg" + 0.010*"t
008*"week" + 0.008*"build" + 0.007*"companion"

Topic: 4

Words: 0.012*"evil" + 0.010*"deed" + 0.009*"word" + 0.008*"person" + 0.007*"seek" + 0.
0.005*"leav" + 0.005*"respect"

Topic: 5

Words: 0.012*"sunnah" + 0.012*"famili" + 0.011*"religion" + 0.007*"respect" + 0.007*"w
+ 0.006*"person" + 0.006*"societi" + 0.006*"book"

Topic: 6

Words: 0.020*"nation" + 0.013*"oppress" + 0.011*"friday" + 0.011*"homeland" + 0.010*"m
08*"today" + 0.007*"histori" + 0.007*"victori"

Number of topics: 7

Topic: 1

Words: 0.013*"truth" + 0.013*"forgiv" + 0.012*"repent" + 0.011*"night" + 0.011*"servan
"trade" + 0.009"forget" + 0.008*"evil"

Topic: 2

Words: 0.033*"month" + 0.030*"ramadan" + 0.015*"sacrific" + 0.012*"reach" + 0.012*"fas
009*"night" + 0.008*"year" + 0.007*"need"

Topic: 3

Words: 0.033*"mosqu" + 0.023*"famili" + 0.022*"knowledg" + 0.018*"children" + 0.011*"t
008*"build" + 0.008*"week" + 0.007*"civil"

Topic: 4

Words: 0.013*"evil" + 0.012*"deed" + 0.009*"word" + 0.007*"iman" + 0.007*"seek" + 0.00
+ 0.005*"help" + 0.005*"perform"

Topic: 5

Words: 0.016*"sunnah" + 0.014*"famili" + 0.010*"religion" + 0.010*"word" + 0.009*"marr
+ 0.007*"societi" + 0.007*"book" + 0.006*"haram"

Topic: 6
Words: 0.014*"oppress" + 0.013*"friday" + 0.012*"children" + 0.008*"violenc" + 0.007*"
+ 0.006*"quran" + 0.006*"compass" + 0.006*"today"

Topic: 7
Words: 0.024*"nation" + 0.015*"homeland" + 0.012*"martyr" + 0.010*"victori" + 0.009*"b
0.008*"spiritu" + 0.008*"jihad" + 0.007*"histori"

Number of topics: 8

Topic: 1
Words: 0.015*"forgiv" + 0.014*"repent" + 0.012*"truth" + 0.012*"night" + 0.011*"usuri"
"sin" + 0.008"deed" + 0.008*"earn"

Topic: 2
Words: 0.033*"month" + 0.030*"ramadan" + 0.015*"sacrific" + 0.012*"fast" + 0.011*"reac
009*"share" + 0.008*"year" + 0.006*"zakat"

Topic: 3
Words: 0.048*"mosqu" + 0.030*"knowledg" + 0.023*"children" + 0.013*"teach" + 0.012*"pl
9*"learn" + 0.009*"read" + 0.008*"build"

Topic: 4
Words: 0.013*"evil" + 0.011*"deed" + 0.008*"iman" + 0.008*"word" + 0.007*"seek" + 0.00
0.006*"sincer" + 0.005*"best"

Topic: 5
Words: 0.019*"sunnah" + 0.015*"religion" + 0.013*"word" + 0.010*"book" + 0.009*"truth"
ommand" + 0.006*"enjoin" + 0.006*"harm"

Topic: 6
Words: 0.016*"friday" + 0.015*"oppress" + 0.013*"children" + 0.010*"violenc" + 0.008*"
0.006*"citi" + 0.006*"great" + 0.006*"mosqu"

Topic: 7
Words: 0.025*"nation" + 0.015*"homeland" + 0.013*"martyr" + 0.010*"victori" + 0.010*"c
0.008*"spiritu" + 0.008*"jihad" + 0.008*"histori"

Topic: 8
Words: 0.032*"famili" + 0.011*"trust" + 0.011*"marriag" + 0.007*"respect" + 0.007*"rel
6*"safeti" + 0.006*"spous" + 0.006*"secur"

Number of topics: 9

Topic: 1
Words: 0.022*"truth" + 0.013*"usuri" + 0.013*"trade" + 0.012*"servant" + 0.010*"earn"
lth" + 0.009*"forget" + 0.008*"path"

Topic: 2
Words: 0.027*"sacrific" + 0.013*"gratitud" + 0.012*"share" + 0.011*"reach" + 0.011*"br
0.009*"zakat" + 0.009*"taqwa" + 0.008*"wealth"

Topic: 3

Words: 0.051*"mosqu" + 0.031*"knowledg" + 0.018*"children" + 0.014*"teach" + 0.012*"pl
09*"read" + 0.009*"book" + 0.008*"build"

Topic: 4

Words: 0.016*"evil" + 0.012*"deed" + 0.009*"iman" + 0.009*"seek" + 0.008*"word" + 0.00
0.006*"person" + 0.006*"need"

Topic: 5

Words: 0.021*"sunnah" + 0.014*"religion" + 0.010*"word" + 0.010*"book" + 0.009*"haram"
"children" + 0.007"halal" + 0.007*"societi"

Topic: 6

Words: 0.015*"children" + 0.014*"friday" + 0.013*"oppress" + 0.009*"violenc" + 0.008*"
+ 0.007*"khutbah" + 0.006*"great" + 0.006*"aqsa"

Topic: 7

Words: 0.025*"nation" + 0.016*"homeland" + 0.014*"martyr" + 0.011*"victori" + 0.010*"c
"spiritu" + 0.008"çanakkal" + 0.007*"histori"

Topic: 8

Words: 0.031*"famili" + 0.012*"trust" + 0.012*"marriag" + 0.008*"women" + 0.008*"parad
07*"greet" + 0.007*"secur" + 0.007*"safe"

Topic: 9

Words: 0.029*"month" + 0.022*"ramadan" + 0.016*"forgiv" + 0.015*"night" + 0.013*"rep
8*"year" + 0.007*"opportun" + 0.007*"need"

Number of topics: 10

Topic: 1

Words: 0.019*"truth" + 0.015*"repent" + 0.014*"usuri" + 0.014*"trade" + 0.012*"forgiv"
"sin" + 0.009"earn" + 0.009*"gain"

Topic: 2

Words: 0.022*"sacrific" + 0.020*"ramadan" + 0.015*"month" + 0.012*"fast" + 0.012*"shar
010*"brotherhood" + 0.009*"zakat" + 0.008*"help"

Topic: 3

Words: 0.055*"mosqu" + 0.033*"knowledg" + 0.014*"children" + 0.013*"teach" + 0.013*"pl
010*"learn" + 0.010*"read" + 0.009*"build"

Topic: 4

Words: 0.012*"evil" + 0.009*"seek" + 0.008*"iman" + 0.008*"need" + 0.008*"deed" + 0.00
+ 0.006*"best" + 0.005*"natur"

Topic: 5

Words: 0.018*"word" + 0.011*"truth" + 0.011*"haram" + 0.010*"greet" + 0.009*"person" +
eti" + 0.009*"thing" + 0.008*"religion"

Topic: 6

Words: 0.019*"children" + 0.014*"oppress" + 0.014*"friday" + 0.008*"violenc" + 0.007*"0.006*"mosqu" + 0.006*"compass" + 0.006*"great"

Topic: 7

Words: 0.026*"nation" + 0.015*"homeland" + 0.013*"martyr" + 0.010*"victori" + 0.009*"c008*"futur" + 0.008*"land" + 0.007*"histori"

Topic: 8

Words: 0.029*"famili" + 0.011*"marriag" + 0.010*"trust" + 0.008*"respect" + 0.007*"dee07*"entrust" + 0.006*"secur" + 0.006*"safe"

Topic: 9

Words: 0.025*"month" + 0.016*"ramadan" + 0.014*"night" + 0.011*"forgiv" + 0.011*"deed" *"evil" + 0.007*"year" + 0.007*"servitud"

Topic: 10

Words: 0.039*"sunnah" + 0.023*"religion" + 0.022*"book" + 0.018*"quran" + 0.013*"examp0.010*"enjoin" + 0.010*"practic" + 0.010*"pbuh"

Out[15]: <gensim.models.ldamulticore.LdaMulticore at 0x1a234664d0>

LDA Model with 8 topics:

This is the equation of our model. Later, there will be an interactive visualizaiton that presents everytrhing more user-friendly. Note, I assigned *passes*=20 because I have very few number of documents and I want my model to read each document 20 times before creating topics.


```
In [16]: lda_model = gensim.models.LdaMulticore(bow_corpus, num_topics=8, id2word=dictionary,
                                                passes=20, workers=2, random_state=57)

for idx, topic in lda_model.print_topics(-1):
    print('Topic: {} \nWords: {}'.format(idx+1, topic))
```

Topic: 1

Words: 0.015*"forgiv" + 0.014*"repent" + 0.012*"truth" + 0.012*"night" + 0.011*"usuri"
"sin" + 0.008"deed" + 0.008*"earn"

Topic: 2

Words: 0.033*"month" + 0.030*"ramadan" + 0.015*"sacrific" + 0.012*"fast" + 0.011*"reac
009*"share" + 0.008*"year" + 0.006*"zakat"

Topic: 3

Words: 0.048*"mosqu" + 0.030*"knowledg" + 0.023*"children" + 0.013*"teach" + 0.012*"pl
9*"learn" + 0.009*"read" + 0.008*"build"

Topic: 4

Words: 0.013*"evil" + 0.011*"deed" + 0.008*"iman" + 0.008*"word" + 0.007*"seek" + 0.00
0.006*"sincer" + 0.005*"best"

Topic: 5

Words: 0.019*"sunnah" + 0.015*"religion" + 0.013*"word" + 0.010*"book" + 0.009*"truth"
ommand" + 0.006*"enjoin" + 0.006*"harm"

Topic: 6

Words: 0.016*"friday" + 0.015*"oppress" + 0.013*"children" + 0.010*"violenc" + 0.008*"
0.006*"citi" + 0.006*"great" + 0.006*"mosqu"

Topic: 7

Words: 0.025*"nation" + 0.015*"homeland" + 0.013*"martyr" + 0.010*"victori" + 0.010*"c
0.008*"spiritu" + 0.008*"jihad" + 0.008*"histori"

Topic: 8

Words: 0.032*"famili" + 0.011*"trust" + 0.011*"marriag" + 0.007*"respect" + 0.007*"rel
6*"safeti" + 0.006*"spous" + 0.006*"secur"

The classification of each sermon by dominant topic and the relevant words inside the text can be seen as follow:

```
In [27]: df_dominant_topic.head(30)
```

```
Out[27]:
```

	Document_No	Dominant_Topic	Topic_Perc_Contrib	Keywords	Text
0	0	5.0	0.5447	friday, oppress, children, violenc, masjid, kh...	[rifq, gentl, allah, love, gracious, affair, h...
1	1	0.0	0.5938	forgiv, repent, truth, night, usuri, trade, se...	[ethic, commerc, islam, jumu, mubarak, dear, b...
2	2	1.0	0.6066	month, ramadan, sacrific, fast, reach, night, ...	[life, contempl, honor, muslim, vers, recit, a...
3	3	1.0	0.9962	month, ramadan, sacrific, fast, reach, night, ...	[sacrific, search, close, allah, honor, believ...
4	4	6.0	0.9966	nation, homeland, martyr, victori, caus, bodi,...	[soldier, mehmetcik, prayer, dear, muslim, ble...
5	5	2.0	0.5535	mosqu, knowledg, children, teach, place, civil...	[messag, revel, human, read, jumu, mubarak, ho...
6	6	3.0	0.4692	evil, deed, iman, word, seek, person, need, he...	[funer, etiquett, duti, travel, etern, life, j...
7	7	3.0	0.6251	evil, deed, iman, word, seek, person, need, he...	[worship, spiritu, world, jumu, mubarak, honor...
8	8	7.0	0.6545	famili, trust, marriag, respect, relat, greet,...	[preserv, famili, environ, righteous, peac, fu...
9	9	6.0	0.5493	nation, homeland, martyr, victori, caus, bodi,...	[juli, rebirth, nation, jumu, mubarak, honor, ...
10	10	5.0	0.7733	friday, oppress, children, violenc, masjid, kh...	[qud, ceas, pain, jumu, mubarak, honor, believ...
11	11	0.0	0.9955	forgiv, repent, truth, night, usuri, trade, se...	[merci, forgiv, door, repent, esteem, believ, ...
12	12	1.0	0.6575	month, ramadan, sacrific, fast, reach, night, ...	[ramadan, infaq, honor, believ, almighty, alla...
13	13	3.0	0.7885	evil, deed, iman, word, seek, person, need, he...	[faith, allah, dear, muslim, almighty, allah, ...
14	14	6.0	0.8988	nation, homeland, martyr, victori, caus, bodi,...	[turkish, victori, battl, gallipoli, esprit, c...
15	15	0.0	0.9957	forgiv, repent, truth, night, usuri, trade, se...	[usuri, illicit, earn, honor, brother, sister,...
16	16	6.0	0.6186	nation, homeland, martyr, victori, caus, bodi,...	[bless, worth, world, health, honor, muslim, v...
17	17	4.0	0.9957	sunnah, religion, word, book, truth, forget, l...	[main, sourc, islam, sunnah, jumu, mubarak, ho...
18	18	6.0	0.9295	nation, homeland, martyr, victori, caus, bodi,...	[make, life, livabl, religion, spiritu, honor,...
19	19	5.0	0.6169	friday, oppress, children, violenc, masjid, kh...	[today, unit, terror, bless, friday, brother, ...
20	20	6.0	0.6225	nation, homeland, martyr, victori, caus, bodi,...	[pure, food, pure, generat, honor, muslim, mes...
21	21	4.0	0.6261	sunnah, religion, word, book, truth, forget, l...	[believ, role, model, bless, friday, brother, ...
22	22	5.0	0.6969	friday, oppress, children, violenc, masjid, kh...	[wise, piec, advic, honor, believ, peopl, almi...
23	23	0.0	0.9957	forgiv, repent, truth, night, usuri, trade, se...	[bara, night, occas, forgiv, purif, salvat, ju...
24	24	6.0	0.8359	nation, homeland, martyr, victori, caus, bodi,...	[materi, spiritu, purif, cleanli, honor, musli...
25	25	7.0	0.7580	famili, trust, marriag, respect, relat, greet,...	[protect, home, planet, earth, honor, muslim, ...
26	26	0.0	0.2734	forgiv, repent, truth, night, usuri, trade, se...	[individu, public, right, honor, muslim, proph...
27	27	4.0	0.9961	sunnah, religion, word, book, truth, forget, l...	[haya, pious, shyness, essenc, islam, moral, h...
28	28	7.0	0.9962	famili, trust, marriag, respect, relat, greet,...	[mawlid, nabi, honor, muslim, know, today, fri...
29	29	3.0	0.9644	evil, deed, iman, word, seek, person, need, he...	[adha, honor, muslim, today, adha, except, tim...

Visualizing LDA model:

This is Mako's favorite visualization of topic models. It is a lot easier than it seems.

```
In [21]: vis
```

Out[21]:

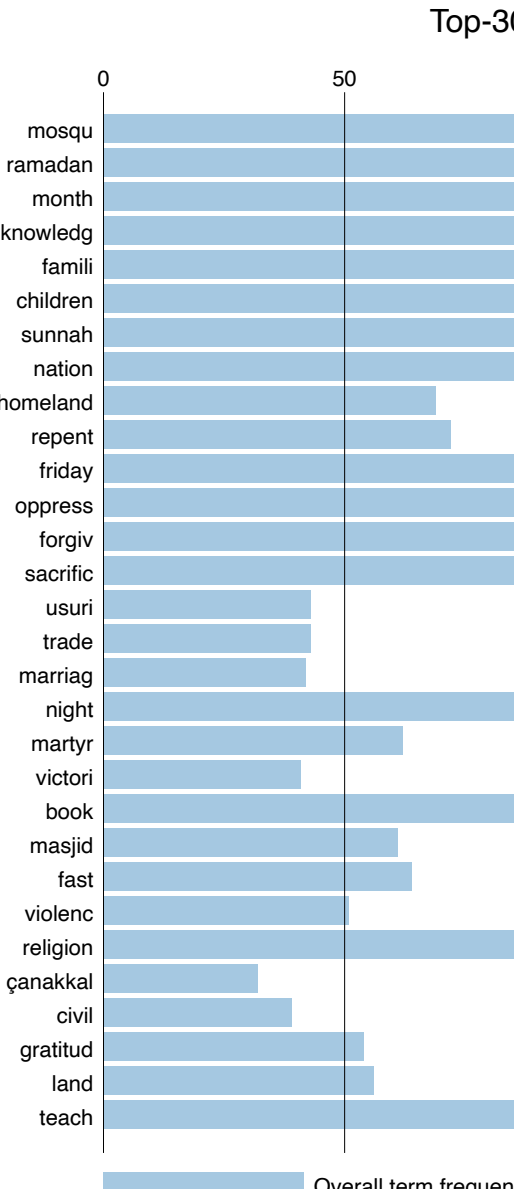
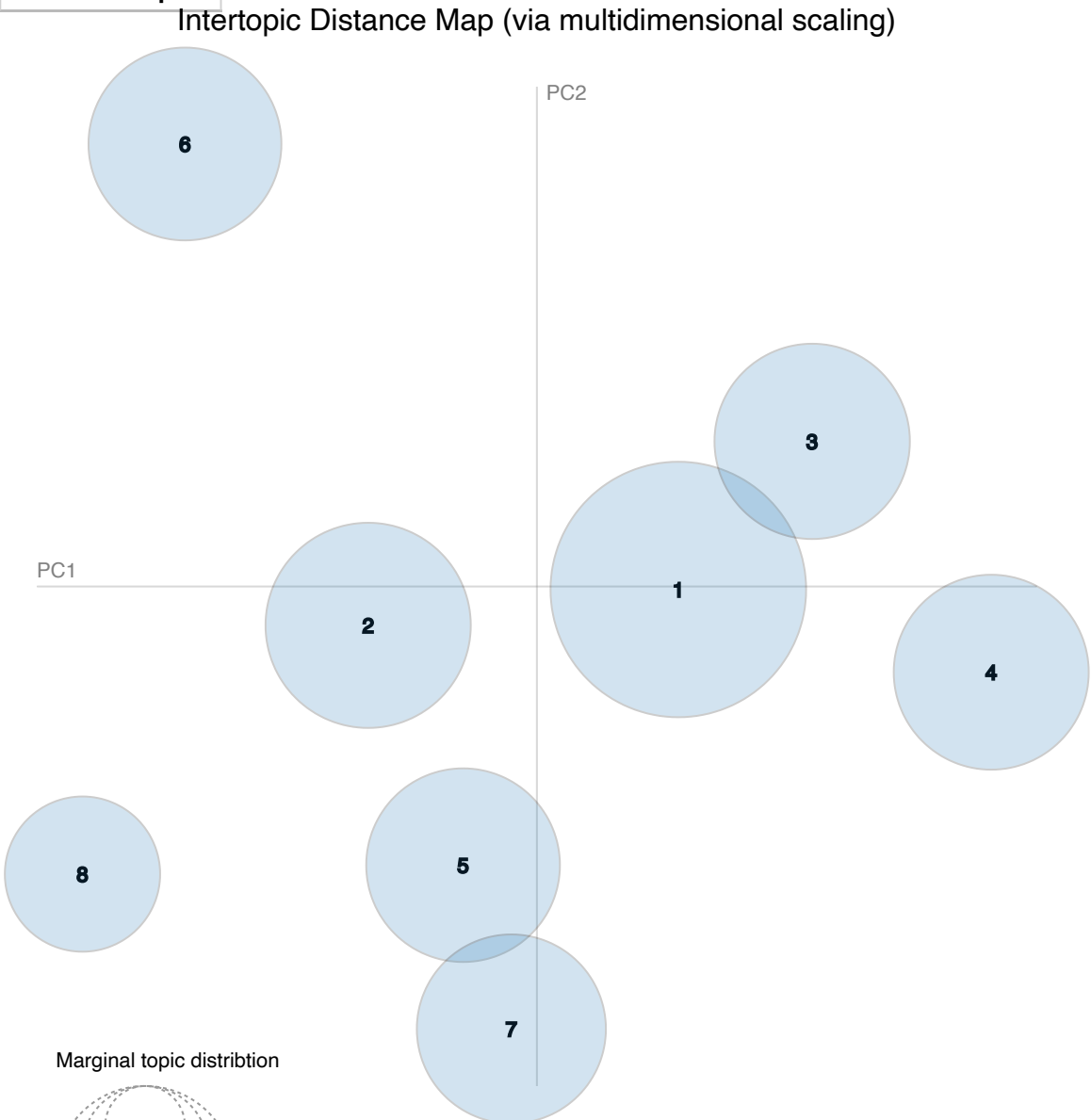
Selected Topic: **0**

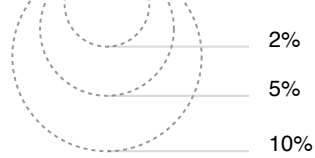
Previous Topic

Next Topic

Clear Topic

Slide to adjust relevance metric
 $\lambda = 1$





Overall term frequency withi

1. $\text{saliency}(\text{term } w) = \text{frequency}(w) * \text{sum}$

2. $\text{relevance}(\text{term } w \mid \text{topic } t) = \lambda * p(w \mid t)$

```
In [22]: #pyLDAvis.show(vis)
```

Based on my observations of topics and closer-reading of relevant texts, I have named the 8 topics as follows:

- belief: Abstract concepts about iman (belief), the existence of evil and meaning of being a muslim.
- praying: Concrete discussions about proper praying, the code of conduct for Friday prayers.
- holy months: includes ramadan, sacrifice, fasting and suggestions for praying in holy months.
- salvation: includes forgiveness and mercifulness of god, how to properly ask for salvation.
- family: suggestions for marriage, solving family problems and gender roles.
- nationalism: Annotations of nation and history, sacredness of homeland and martyr and specific contexts about jihad and historic wars.
- lifestyle: sunnah stands for the life practices of the prophet as an exemplar for a proper religious lifestyle. Also includes words like rules and haram (things are not-permitted by god).
- education: explain the importance of summer schools takes place in mosques to provide religious education to children.

Final Output:

AS I stated earlier, I am interested in seeing the change of the prevalence of certain topics over time instead of the dominant topic for each text. As a result, I have created the following tables for each document. Each row represents one sermon including the date of the sermon and the probability of that sermon to belong to the particular topics in each column. I will use these tables to graph timeseries for each topic

In [25]: topics_df

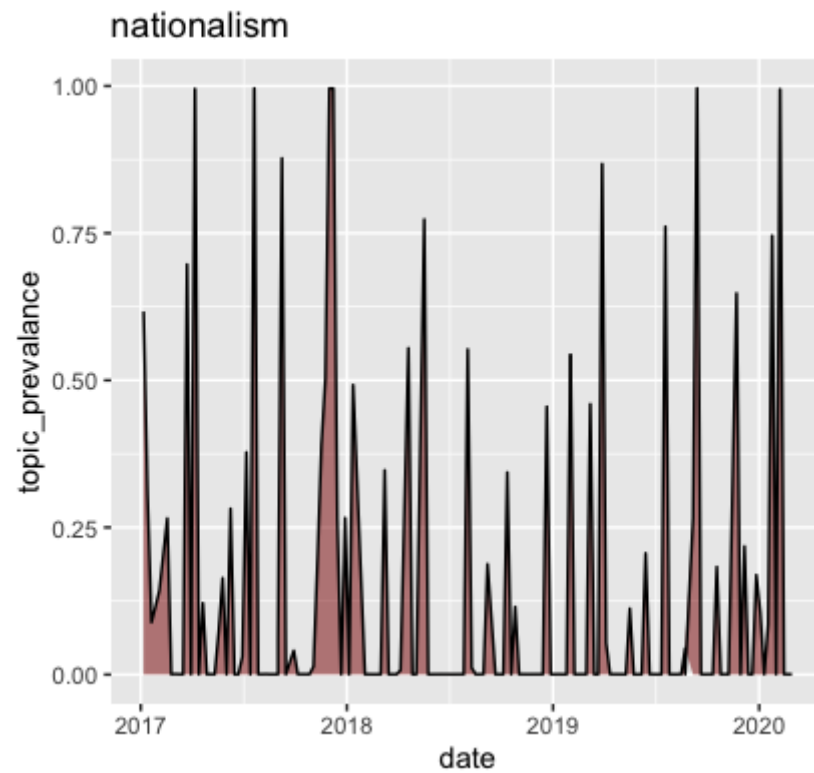
Out[25]:

	date	belief	praying	holy months	salvation	family	nationalism	lifestyle	education
0	2019-02-01	0.000703	0.000703	0.000703	0.451073	0.000704	0.544707	0.000703	0.000704
1	2018-03-02	0.593834	0.000559	0.000559	0.149976	0.175987	0.000559	0.077966	0.000559
2	2019-12-27	0.050748	0.606617	0.000501	0.000501	0.000501	0.169245	0.000501	0.171387
3	2017-08-25	0.000547	0.996174	0.000546	0.000547	0.000547	0.000547	0.000546	0.000546
4	2019-10-25	0.000489	0.000489	0.000489	0.000489	0.000489	0.000489	0.996577	0.000489
...
152	2019-04-12	0.000670	0.000669	0.000669	0.995313	0.000670	0.000669	0.000670	0.000669
153	2017-09-01	0.595874	0.000646	0.000645	0.221821	0.118348	0.061376	0.000645	0.000646
154	2017-02-03	0.000574	0.000574	0.000574	0.000574	0.169881	0.142938	0.000574	0.684310
155	2019-08-23	0.931653	0.000485	0.000485	0.065435	0.000485	0.000485	0.000485	0.000485
156	2019-11-29	0.092015	0.000715	0.000716	0.580406	0.000716	0.000716	0.000716	0.324001

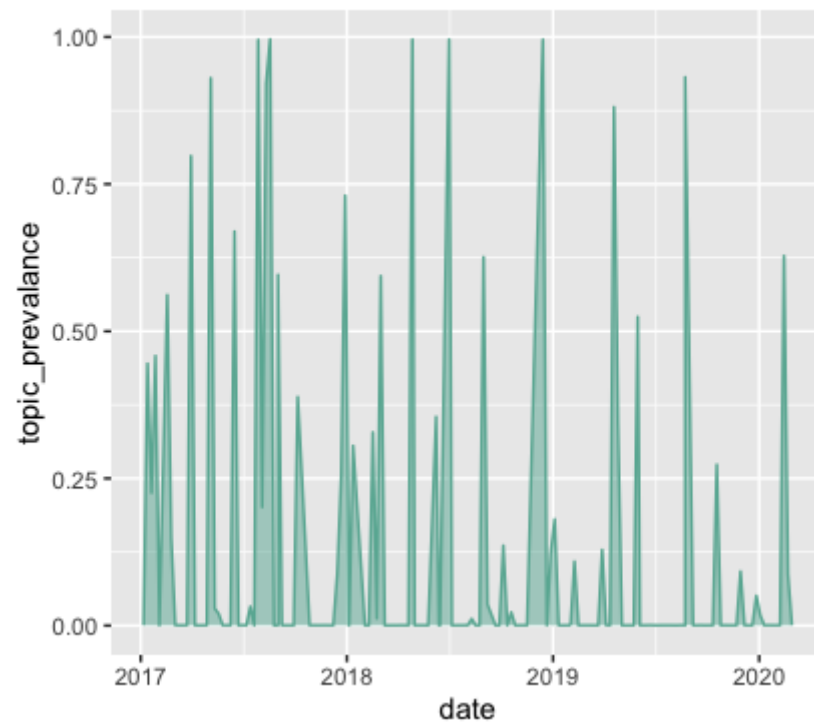
157 rows × 9 columns

Prevalance of topics over time:

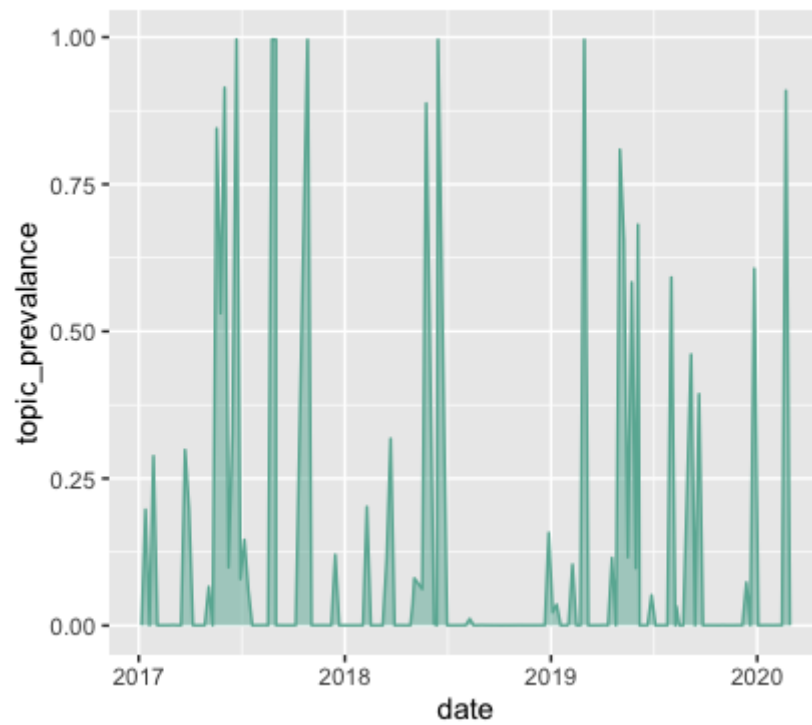
The graphs are area graph. It is probably not the best way to visualize this data since there are many zero points over the years.



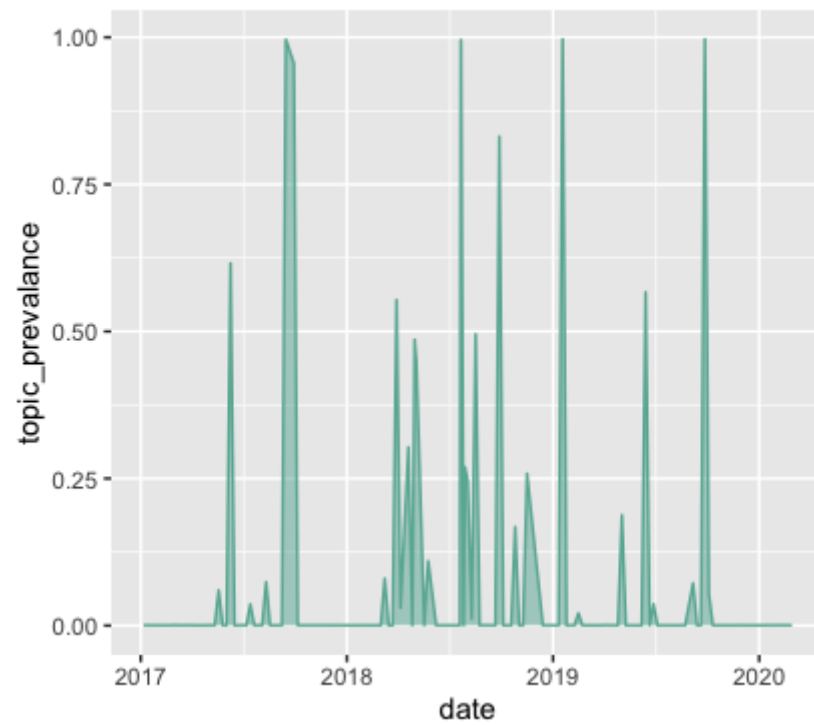
belief



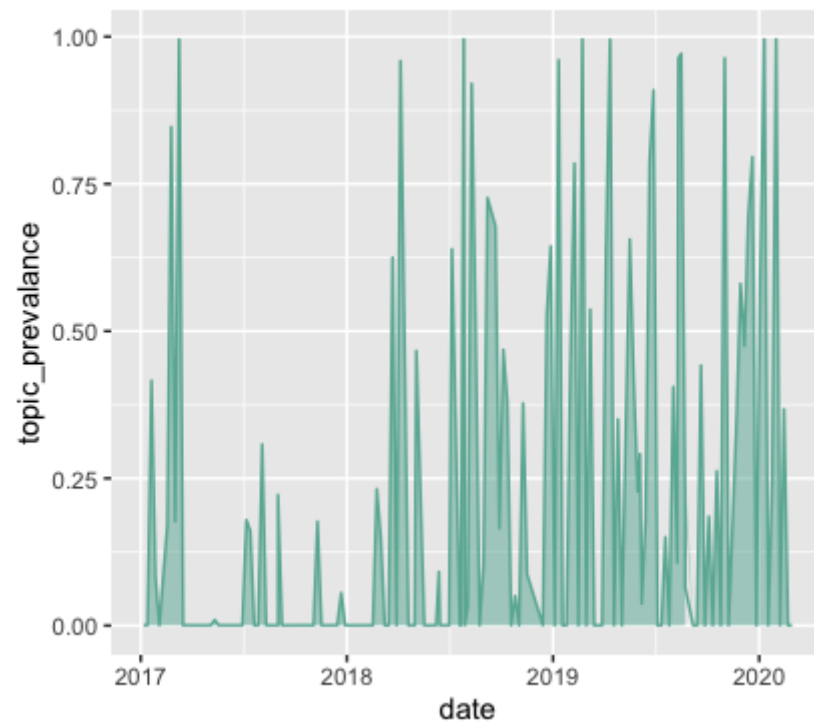
praying



ramadan



salvation



Future work:

- Compare the Turkish and English LDA models. Important for increasing data size.
- Use the time series plots to do close readings.