

# Problem Set 4: Topic Modeling

Your name here:

Due Friday April 15, 2022 by midnight

Please upload answers to all questions (including computational questions and any related graphics and R code *with comments*, in .rmd and knitted pdf forms to Canvas. For any problems that require calculations, please show your work. Finally, when doing simulations or draws from a random distribution in R, please set your seed immediately using `set.seed(2018)`.

## Structural Topic Model

This problem set is meant to provide a brief overview of the `stm` (structural topic model) package.

Structural topic modeling is a way to estimate a topic model that includes document-level meta-data. One can then see how topical prevalence changes according to that meta-data.

Required packages:

```
library(stm)
```

```
## stm v1.3.6 successfully loaded. See ?stm for help.  
## Papers, resources, and other materials at structuraltopicmodel.com
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --  
  
## v ggplot2 3.3.5      v purrr   0.3.4  
## v tibble  3.1.6      v dplyr  1.0.7  
## v tidyr   1.1.4      v stringr 1.4.0  
## v readr   2.0.2      v forcats 0.5.1  
  
## -- Conflicts ----- tidyverse_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()    masks stats::lag()
```

### Question 1

The data you'll be using for this problem set consists of all articles about women published in the New York Times and Washington Post, 1980-2014.

Load the dataset `women-full.csv`. Notice that you have the text of the articles, along with some metadata.

STM has its own unique preprocessing functions and procedure.

The main text you'll be focusing on is the `women$TEXT.NO.NOUN` column, which contains all the text of the articles without proper nouns (previously removed). Use the function `textProcessor` on the text (with `metadata=women`) to preprocess the data, then pull out the `meta`, `vocab`, and `documents` items. You can

then use the `prepDocuments` function on the pulled out `meta,vocab,documents` objects to prepare documents in correct format for analysis with `stm` (make sure to consult the help files on these functions).

```
# read in file
women <- read.csv("women-full.csv")

# pre process data
women_pp <- textProcessor(documents = women$TEXT.NO.NOUN, metadata = women)

## Building corpus...
## Converting to Lower Case...
## Removing punctuation...
## Removing stopwords...
## Removing numbers...
## Stemming...
## Creating Output...

# extract factors
women_prep <- prepDocuments(document = women_pp$documents, vocab = women_pp$vocab, meta = women_pp$meta)

## Removing 19460 of 39403 terms (19460 of 1087166 tokens) due to frequency
## Your corpus now has 4531 documents, 19943 terms and 1067706 tokens.

doc <- women_prep$documents
voc <- women_prep$vocab
meta <- women_prep$meta
```

## Question 2

Estimate a topic model with 15 topics by regressing topical prevalence on `REGION` and `s(YEAR)` covariates. Note that we use `s()` wrapper on year to specify smooth terms for B-spline basis expansion.

Running full model takes a **long** time to finish. For that reason, add an argument `max.em.its` which sets the number of iterations. By keeping it low (15) you'll see a rough estimate of the topics. You can always go back and change this value to estimate the model until its convergence. Save the fitted model as `model`.

Use the following tools to evaluate your model:

- `labelTopics` gives the top words for each topic.
- `findThoughts` gives the top documents for each topic (the documents with the highest proportion of each topic)

```
# fit the model
model <- stm(doc, voc, 15, prevalence = ~ REGION + s(YEAR), data = meta, max.em.its = 15)

## Beginning Spectral Initialization
## Calculating the gram matrix...
## Using only 10000 most frequent terms during initialization...
## Finding anchor words...
## .....
## Recovering initialization...
## .....
## Initialization complete.
## .....
## Completed E-Step (3 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -7.882)
```

```

## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -7.605, relative change = 3.519e-02)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -7.573, relative change = 4.218e-03)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -7.559, relative change = 1.843e-03)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -7.550, relative change = 1.162e-03)
## Topic 1: women, show, one, like, design
## Topic 2: said, women, polic, report, year
## Topic 3: women, team, said, game, world
## Topic 4: year, book, first, one, mrs
## Topic 5: women, said, percent, femal, will
## Topic 6: said, women, one, peopl, woman
## Topic 7: women, work, said, year, men
## Topic 8: women, sexual, sex, rape, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, women, peopl
## Topic 11: women, film, one, said, woman
## Topic 12: said, polit, elect, parti, govern
## Topic 13: women, said, cancer, studi, health
## Topic 14: women, confer, said, deleg, will
## Topic 15: said, women, girl, rape, practic
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -7.544, relative change = 8.200e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -7.539, relative change = 5.961e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -7.536, relative change = 4.491e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -7.533, relative change = 3.481e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -7.531, relative change = 2.759e-04)
## Topic 1: show, design, women, fashion, art
## Topic 2: said, women, polic, report, offici
## Topic 3: women, team, said, game, world

```

```

## Topic 4: book, year, life, first, work
## Topic 5: women, said, femal, percent, will
## Topic 6: said, one, protest, peopl, site
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, rape, sex, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, peopl, day
## Topic 11: women, film, one, like, woman
## Topic 12: polit, said, elect, parti, govern
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, confer, said, will, world
## Topic 15: said, women, girl, rape, case
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -7.530, relative change = 2.214e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -7.528, relative change = 1.804e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -7.527, relative change = 1.496e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -7.526, relative change = 1.265e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Model Terminated Before Convergence Reached

```

```
# list top words for each topic
```

```
labelTopics(model)
```

```

## Topic 1 Top Words:
##   Highest Prob: show, design, fashion, women, art, one, like
##   FREX: coutur, fashion, museum, sculptur, ready--wear, jacket, galleri
##   Lift: ---inch, -ankl, alexandr, armatur, armhol, art-fair, avant
##   Score: coutur, art, artist, fashion, museum, exhibit, cloth
## Topic 2 Top Words:
##   Highest Prob: said, polic, women, kill, report, offici, govern
##   FREX: polic, suicid, kill, attack, investig, suspect, arrest
##   Lift: abducte, charanjit, humanity-soak, male-control, sunil, kalpana, ciudad
##   Score: polic, rape, kill, said, arrest, attack, investig
## Topic 3 Top Words:
##   Highest Prob: women, team, game, said, world, play, olymp
##   FREX: tournament, championship, olymp, soccer, player, game, medal
##   Lift: -america, -foot--inch, -hole, -kilomet, -rank, -round, -trump
##   Score: olymp, championship, tournament, team, player, game, medal
## Topic 4 Top Words:
##   Highest Prob: book, year, life, first, write, novel, work
##   FREX: novel, literari, fiction, book, memoir, novelist, poet
##   Lift: buster, calla, goncourt, identical-twin, italian-american, kilcher, klone

```

## Score: novel, book, fiction, literari, poet, writer, write

## Topic 5 Top Words:

## Highest Prob: women, said, femal, percent, militari, will, compani

## FREX: combat, board, quota, militari, bank, corpor, infantri

## Lift: -combat, cpr, gender-divers, nonexecut, outfitt, r-calif, r-ni

## Score: women, militari, infantri, combat, percent, quota, femal

## Topic 6 Top Words:

## Highest Prob: protest, said, one, site, peopl, young, video

## FREX: orthodox, internet, web, video, rabbi, prayer, site

## Lift: balaclava, grrrl, tehrik-, braveheart, drawbridg, gravesit, guerrilla-styl

## Score: protest, site, orthodox, video, jewish, rabbi, xxxfx

## Topic 7 Top Words:

## Highest Prob: women, work, said, year, percent, men, ese

## FREX: ese, factori, employ, incom, worker, job, market

## Lift: flexim, management-track, nec, nontransfer, rabenmutt, chiho, fumiko

## Score: ese, percent, compani, work, job, women, factori

## Topic 8 Top Words:

## Highest Prob: women, sexual, sex, rape, men, violenc, said

## FREX: harass, sexual, sex, assault, brothel, violenc, behavior

## Lift: offenc, tarun, chaud, much-lov, newt, tiresom, sex-rel

## Score: rape, sexual, harass, violenc, sex, assault, brothel

## Topic 9 Top Words:

## Highest Prob: women, said, right, law, islam, govern, religi

## FREX: islam, religi, veil, constitut, saudi, secular, cleric

## Lift: afghan-styl, anglo-, archdeacon, bien-aim, episcopaci, fez, government-encourag

## Score: islam, law, women, right, religi, ordin, saudi

## Topic 10 Top Words:

## Highest Prob: said, one, famili, peopl, day, like, home

## FREX: villag, room, smile, son, couldnt, recal, sit

## Lift: charpoy, jet-black, mitra, schermerhorn, single-famili, tyson, uja-feder

## Score: villag, husband, fistula, famili, school, girl, said

## Topic 11 Top Words:

## Highest Prob: women, film, one, like, woman, say, play

## FREX: film, theater, movi, charact, actress, documentari, audienc

## Lift: clive, fine-tun, kaffir, nushus, shrew, nushu, cadel

## Score: film, theater, movi, nushu, play, orchestra, femin

## Topic 12 Top Words:

## Highest Prob: polit, elect, parti, minist, presid, govern, said

## FREX: voter, elect, parti, prime, candid, vote, cabinet

## Lift: ernesto, pinbal, influence-peddl, information-servic, kakuei, left--cent, marxist-lenini

## Score: elect, parti, vote, minist, voter, polit, candid

## Topic 13 Top Words:

## Highest Prob: women, said, abort, cancer, health, studi, breast

## FREX: implant, cancer, breast, pill, virus, patient, estrogen

## Lift: acet, adren, ambulatori, analges, anastrozol, antioxid, ashkenazi

## Score: cancer, abort, breast, pill, implant, health, virus

## Topic 14 Top Words:

## Highest Prob: women, said, confer, will, world, organ, right

## FREX: deleg, confer, forum, page, peac, nongovernment, ambassador

## Lift: -glass, barack, brooklyn-born, expansion, foreclosur, guarantor, holden

## Score: deleg, confer, forum, page, palestinian, peac, mrs

## Topic 15 Top Words:

## Highest Prob: said, women, rape, court, case, girl, practic

## FREX: mutil, genit, circumcis, asylum, sentenc, judg, tribun

```
##      Lift: labia, layli, minora, multifait, paraleg, salim, strip-search
##      Score: rape, genit, circumcis, mutil, court, sentenc, prosecutor
```

```
# top documents for each topic
```

```
findThoughts(model, texts = meta$TITLE, n = 2, topics = c(1:15))
```

```
##
```

```
## Topic 1:
```

```
##      KENZO'S CAREFREE STYLES AT AN OFFBEAT SHOWING
```

```
##      A MODERN LOOK, A CLASSIC TOUCH FROM SAINT LAURENT
```

```
## Topic 2:
```

```
##      Assailants Kill 4 Iraqi Women Working for U.S.; Gunmen Follow Van Carrying Laundry Employees
```

```
##      WORLD IN BRIEF
```

```
## Topic 3:
```

```
##      AMERICANS LEAD EAST GERMANS IN TRACK
```

```
##      Russians Chart a New Path
```

```
## Topic 4:
```

```
##      BEST SELLERS: September 6, 1998
```

```
##      BEST SELLERS: September 13, 1998
```

```
## Topic 5:
```

```
##      In Britain, a Big Push for More Women to Serve on Corporate Boards
```

```
##      Poll: Allow women in combat units
```

```
## Topic 6:
```

```
##      Neda's Legacy; A woman's death moves Iranian protesters.
```

```
##      Jewish Feminists Prompt Protests at Wailing Wall
```

```
## Topic 7:
```

```
##      China Scrambles for Stability as Its Workers Age
```

```
##      A high price for a paycheck; Caught between the demands of the corporate workplace and of their
```

```
## Topic 8:
```

```
##      Confronting Rape in India, and Around the World
```

```
##      Sexual Harassment Prosecutions Get Short Shrift in India, Lawyer Says
```

```
## Topic 9:
```

```
##      English Church Advances Bid For Women As Bishops
```

```
##      Egypt Passes Law On Women's Rights; Polygamy Still Allowed for Men
```

```
## Topic 10:
```

```
##      An Old Cinema in Pakistan Has New Life After Quake
```

```
##      Maria Duran's Endless Wait
```

```
## Topic 11:
```

```
##      For France, An All-Purpose Heartthrob
```

```
##      Film: Brazilian 'Vera'
```

```
## Topic 12:
```

```
##      The Widow Of Ex-Leader Wins Race In Panama
```

```
##      Cabinet Defeated in Iceland as Feminists Gain
```

```
## Topic 13:
```

```
##      SECTION: HEALTH; Pg. T18
```

```
##      Dense Breasts May Need Sonograms to Detect Cancer
```

```
## Topic 14:
```

```
##      DISPUTES ON KEY ISSUES STALL KENYA PARLEY
```

```
##      'CHAOTIC' CONDITIONS FEARED AT U.N.'S PARLEY ON WOMEN
```

```
## Topic 15:
```

```
##      Woman Fleeing Tribal Rite Gains Asylum; Genital Mutilation Is Ruled Persecution
```

```
##      Refugee From Mutilation
```

### Question 3

Run a fully-estimated model with larger `max.em.its` (e.g. 200). Using the functions `labelTopics` and `findThoughts`, hand label the 15 topics. Hold these labels as a character vector called `labels`.

```
# run model again but without maximum
model2 <- stm(doc, voc, 15, prevalence = ~ REGION + s(YEAR), data = meta, max.em.its = 200)

## Beginning Spectral Initialization
##   Calculating the gram matrix...
##   Using only 10000 most frequent terms during initialization...
##   Finding anchor words...
##   .....
##   Recovering initialization...
##   .....
## Initialization complete.
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 1 (approx. per word bound = -7.882)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 2 (approx. per word bound = -7.605, relative change = 3.519e-02)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -7.573, relative change = 4.218e-03)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -7.559, relative change = 1.843e-03)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -7.550, relative change = 1.162e-03)
## Topic 1: women, show, one, like, design
## Topic 2: said, women, polic, report, year
## Topic 3: women, team, said, game, world
## Topic 4: year, book, first, one, mrs
## Topic 5: women, said, percent, femal, will
## Topic 6: said, women, one, peopl, woman
## Topic 7: women, work, said, year, men
## Topic 8: women, sexual, sex, rape, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, women, peopl
## Topic 11: women, film, one, said, woman
## Topic 12: said, polit, elect, parti, govern
## Topic 13: women, said, cancer, studi, health
## Topic 14: women, confer, said, deleg, will
## Topic 15: said, women, girl, rape, practic
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -7.544, relative change = 8.200e-04)
```

```

## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -7.539, relative change = 5.961e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -7.536, relative change = 4.491e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -7.533, relative change = 3.481e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -7.531, relative change = 2.759e-04)
## Topic 1: show, design, women, fashion, art
## Topic 2: said, women, polic, report, offici
## Topic 3: women, team, said, game, world
## Topic 4: book, year, life, first, work
## Topic 5: women, said, femal, percent, will
## Topic 6: said, one, protest, peopl, site
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, rape, sex, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, peopl, day
## Topic 11: women, film, one, like, woman
## Topic 12: polit, said, elect, parti, govern
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, confer, said, will, world
## Topic 15: said, women, girl, rape, case
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -7.530, relative change = 2.214e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -7.528, relative change = 1.804e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -7.527, relative change = 1.496e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -7.526, relative change = 1.265e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 15 (approx. per word bound = -7.525, relative change = 1.062e-04)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, report
## Topic 3: women, team, game, said, world

```



```

## Topic 4: book, year, life, first, write
## Topic 5: women, said, femal, percent, militari
## Topic 6: protest, said, one, site, peopl
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, sex, rape, men
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, day
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, parti, minist, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: said, women, rape, court, case
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 16 (approx. per word bound = -7.525, relative change = 8.825e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 17 (approx. per word bound = -7.524, relative change = 7.456e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 18 (approx. per word bound = -7.524, relative change = 6.389e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 19 (approx. per word bound = -7.523, relative change = 5.614e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 20 (approx. per word bound = -7.523, relative change = 5.079e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, year, life, write, first
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, one, site
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, sex, men, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, parti, minist, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: said, women, case, court, rape
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 21 (approx. per word bound = -7.522, relative change = 4.649e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.

```

```

## Completing Iteration 22 (approx. per word bound = -7.522, relative change = 4.226e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 23 (approx. per word bound = -7.522, relative change = 3.829e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 24 (approx. per word bound = -7.522, relative change = 3.489e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 25 (approx. per word bound = -7.521, relative change = 3.252e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, first
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, parti, minist, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, women, rape
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 26 (approx. per word bound = -7.521, relative change = 2.983e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 27 (approx. per word bound = -7.521, relative change = 2.765e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 28 (approx. per word bound = -7.521, relative change = 2.550e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 29 (approx. per word bound = -7.521, relative change = 2.351e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 30 (approx. per word bound = -7.520, relative change = 2.185e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, kill, women, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, novel
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, site, one

```

```

## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: case, court, said, women, rape
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 31 (approx. per word bound = -7.520, relative change = 2.011e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 32 (approx. per word bound = -7.520, relative change = 1.806e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 33 (approx. per word bound = -7.520, relative change = 1.683e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 34 (approx. per word bound = -7.520, relative change = 1.582e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 35 (approx. per word bound = -7.520, relative change = 1.480e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, kill, women, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, stori
## Topic 5: women, said, femal, compani, percent
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, girl
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, studi
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, rape, women
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 36 (approx. per word bound = -7.520, relative change = 1.384e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 37 (approx. per word bound = -7.520, relative change = 1.256e-05)
## .....
## Completed E-Step (2 seconds).

```

```

## Completed M-Step.
## Completing Iteration 38 (approx. per word bound = -7.519, relative change = 1.150e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 39 (approx. per word bound = -7.519, relative change = 1.090e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 40 (approx. per word bound = -7.519, relative change = 1.085e-05)
## Topic 1: show, design, women, fashion, art
## Topic 2: said, polic, kill, offici, women
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, stori
## Topic 5: women, said, femal, compani, men
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, girl, peopl
## Topic 11: film, women, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, studi
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, rape, women
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 41 (approx. per word bound = -7.519, relative change = 1.076e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 42 (approx. per word bound = -7.519, relative change = 1.046e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Model Converged

```

```

# print topics
labelTopics(model2)

```

```

## Topic 1 Top Words:
## Highest Prob: show, design, women, fashion, art, like, one
## FREX: coutur, sculptur, ready--wear, fashion, sauna, jacket, museum
## Lift: sportswear, ready--wear, ---inch, --shoulder, --top, -ankl, -piec
## Score: coutur, art, artist, fashion, museum, sculptur, kimono
## Topic 2 Top Words:
## Highest Prob: said, polic, kill, offici, women, report, attack
## FREX: polic, kill, bomb, suicid, attack, milit, gun
## Lift: belarus, burned-, charanjit, jat, maquiladora, n-control, opel
## Score: polic, kill, arrest, investig, abduct, rape, bomb
## Topic 3 Top Words:
## Highest Prob: women, team, game, said, world, play, olymp
## FREX: olymp, championship, tournament, hockey, soccer, medal, player
## Lift: -game, -hole, -rank, -round, abbi, advil, anti-dop

```

## Score: olymp, championship, tournament, player, game, medal, team

## Topic 4 Top Words:

## Highest Prob: book, life, year, write, stori, novel, first

## FREX: novel, literari, book, fiction, memoir, literatur, poet

## Lift: barker, buster, calla, carey, chava, co-winn, dani

## Score: novel, book, fiction, literari, poet, writer, memoir

## Topic 5 Top Words:

## Highest Prob: women, said, femal, compani, men, percent, militari

## FREX: combat, board, quota, corpor, infantri, unit, militari

## Lift: -combat, -volunt, auw, baby-blu, bloomberg, board-readi, bradley

## Score: infantri, militari, combat, percent, compani, quota, women

## Topic 6 Top Words:

## Highest Prob: protest, said, peopl, site, one, video, young

## FREX: internet, video, web, orthodox, twitter, site, prayer

## Lift: avi, anat, anti-corrupt, anti-imperialist, arabic-languag, backseat, balaclava

## Score: palestinian, protest, orthodox, video, site, jewish, rabbi

## Topic 7 Top Words:

## Highest Prob: women, work, said, year, percent, ese, famili

## FREX: ese, factori, employ, incom, worker, economi, salari

## Lift: arrack, asst, atsuko, billion-plus, blackjack, british-era, career-mind

## Score: ese, percent, factori, worker, compani, job, econom

## Topic 8 Top Words:

## Highest Prob: women, sexual, men, rape, sex, said, violenc

## FREX: harass, sexual, brothel, sex, violenc, assault, dowri

## Lift: -year-girl, ashoka, chaud, erotica, kingpin, masaomi, much-lov

## Score: rape, sexual, violenc, harass, brothel, assault, prostitut

## Topic 9 Top Words:

## Highest Prob: women, said, right, law, islam, religi, will

## FREX: islam, religi, veil, saudi, constitut, secular, fundamentalist

## Lift: -envelop, afghan-styl, american-islam, anglo-, archdeacon, attaché, augur

## Score: islam, law, religi, ordin, saudi, fundamentalist, divorc

## Topic 10 Top Words:

## Highest Prob: said, one, famili, girl, peopl, like, say

## FREX: villag, son, recal, room, husband, daughter, smile

## Lift: bidi, charpoy, ever-young, first-floor, get-togeth, homeroom, jet-black

## Score: villag, husband, girl, father, famili, mrs, fistula

## Topic 11 Top Words:

## Highest Prob: film, women, one, like, play, woman, say

## FREX: film, theater, movi, actress, audienc, actor, documentari

## Lift: abli, acorn, auteur, beatriz, cinematographi, depression-era, fine-tun

## Score: film, theater, movi, nushu, actor, subtitl, orchestra

## Topic 12 Top Words:

## Highest Prob: polit, elect, minist, parti, presid, govern, said

## FREX: voter, elect, prime, parti, vote, candid, minist

## Lift: -elect, apparatchik, backroom, dae, ernesto, fianna, fusa

## Score: elect, vote, voter, minist, parti, candid, polit

## Topic 13 Top Words:

## Highest Prob: women, said, abort, cancer, studi, health, breast

## FREX: implant, cancer, cervic, virus, breast, chemotherapi, pill

## Lift: --chart, --fact, -billion-ayear, acceleromet, acet, acupressur, adren

## Score: cancer, abort, breast, implant, pill, virus, pregnanc

## Topic 14 Top Words:

## Highest Prob: women, said, confer, will, world, right, organ

## FREX: deleg, confer, forum, peac, nongovernment, ambassador, agreement

```

##      Lift: -paragraph, antifeminist, barack, d-la, foreclosur, foreign-polici, kere
##      Score: deleg, confer, forum, page, peac, palestinian, zionism
## Topic 15 Top Words:
##      Highest Prob: case, said, court, rape, women, law, lawyer
##      FREX: mutil, asylum, circumcis, genit, sentenc, court, prosecutor
##      Lift: catgut, cheerless, clitoridectomi, constablepwashpostcom, efua, evaw, half-fil
##      Score: rape, genit, circumcis, prosecutor, court, mutil, prosecut

# print titles
findThoughts(model2, texts = meta$TITLE, n = 2, topics = c(1:15))

##
## Topic 1:
##      A MODERN LOOK, A CLASSIC TOUCH FROM SAINT LAURENT
##      Designing For Women's Various Lives
## Topic 2:
##      Assailants Kill 4 Iraqi Women Working for U.S.; Gunmen Follow Van Carrying Laundry Employees
##      2 BRITONS KIDNAPPED IN BEIRUT
## Topic 3:
##      AMERICANS LEAD EAST GERMANS IN TRACK
##      RESULTS PLUS
## Topic 4:
##      BEST SELLERS: September 6, 1998
##      BEST SELLERS: September 13, 1998
## Topic 5:
##      Poll: Allow women in combat units
##      In Britain, a Big Push for More Women to Serve on Corporate Boards
## Topic 6:
##      Neda's Legacy; A woman's death moves Iranian protesters.
##      Jewish Feminists Prompt Protests at Wailing Wall
## Topic 7:
##      China Scrambles for Stability as Its Workers Age
##      Women's March Into Work Force Changes Japan's Social Structure; Automation Revolution Lures Wave
## Topic 8:
##      Confronting Rape in India, and Around the World
##      Newsmagazine's Handling of Sex Harassment Complaint Draws Criticism
## Topic 9:
##      Egypt Passes Law On Women's Rights; Polygamy Still Allowed for Men
##      English Church Advances Bid For Women As Bishops
## Topic 10:
##      PIONEER WOMEN OF THE OUTBACK; BOND SPRINGS STATION, Australia
##      An Old Cinema in Pakistan Has New Life After Quake
## Topic 11:
##      Film: Brazilian 'Vera'
##      Women's Status in Vietnam In Documentary Form
## Topic 12:
##      2 New Parties Gain Parliament in Iceland
##      Cabinet Defeated in Iceland as Feminists Gain
## Topic 13:
##      SECTION: HEALTH; Pg. T18
##      Dense Breasts May Need Sonograms to Detect Cancer
## Topic 14:
##      DISPUTES ON KEY ISSUES STALL KENYA PARLEY
##      'CHAOTIC' CONDITIONS FEARED AT U.N.'S PARLEY ON WOMEN
## Topic 15:

```

```
##      Woman Fleeing Tribal Rite Gains Asylum; Genital Mutilation Is Ruled Persecution
##      U.S. Grants Asylum to Woman Fleeing Genital Mutilation Rite

# create the different labels by hand
labels <- c("fashion", "violence", "sports", "literature", "military",
            "protest", "careers", "sexual assault", "islam", "family",
            "film", "politics", "medical", "conference discussion", "FGM")
```

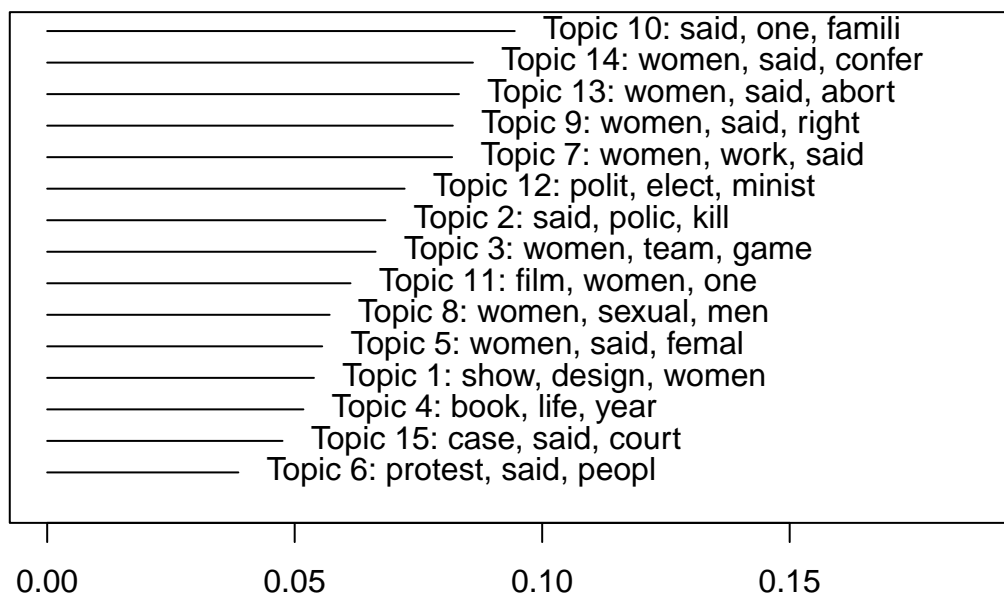
## Question 4

See how the topics are distributed in terms of their prevalence across region. Focus on a single topic of interest for you – What do you notice about the distribution of topic X?

Use `plot.STM` on your model to summarize the corpus. Then `estimateEffect` to find the coefficients on the docs across regions — “Asia”, “EECA”, “MENA”, “Africa”, “West”, and “LA” — and visualize the result using `plot.estimateEffect`. Briefly discuss the results in your own words.

```
# summarize the corpus - find most common topics
plot.STM(model2)
```

### Top Topics



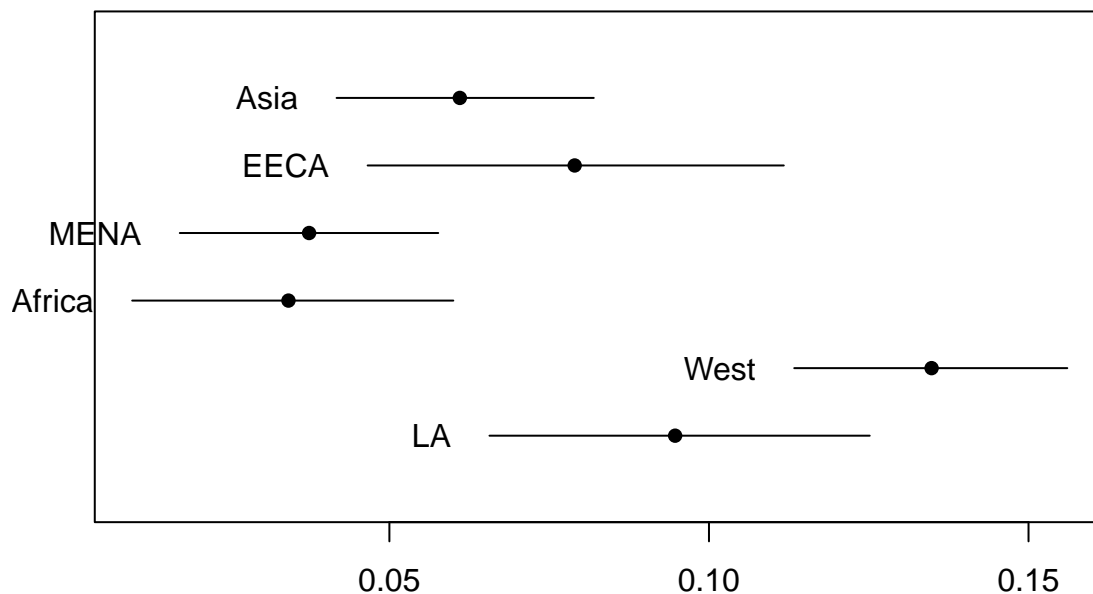
### Expected Topic Proportions

```
est <- estimateEffect(1:15 ~ REGION + s(YEAR), meta = meta, model2)

# name regions
regions <- c("Asia", "EECA", "MENA", "Africa", "West", "LA")

# create plot
plot.estimateEffect(x = est, covariate = "REGION", topics = 3, method = "pointestimate", xlab = "Proportion")
```

## Prevalence of Mentions of Women in Sport by Region



Proportion of Overall Mentions

The above plot shows how mentions of women playing sports (topic 3) are distributed by region. As you can see, the West has a larger share of mentions than other regions, followed by Latin America. The Middle East and Africa have the fewest mentions.

### [Bonus]

Try different specification(s) of the prevalence model. Fully estimate the model until its convergence and explore the result. Briefly discuss the results: is there any new finding?

*# rerun the code but with different specifications - I now include the publication*

*# fit the model*

```
model3 <- stm(doc, voc, 15, prevalence = ~ REGION + PUBLICATION + s(YEAR), data = meta, max.em.its = 200)
```

```
## Beginning Spectral Initialization
```

```
## Calculating the gram matrix...
```

```
## Using only 10000 most frequent terms during initialization...
```

```
## Finding anchor words...
```

```
## .....
```

```
## Recovering initialization...
```

```
## .....
```

```
## Initialization complete.
```

```
## .....
```

```
## Completed E-Step (2 seconds).
```

```
## Completed M-Step.
```

```
## Completing Iteration 1 (approx. per word bound = -7.882)
```

```
## .....
```

```
## Completed E-Step (2 seconds).
```

```
## Completed M-Step.
```

```
## Completing Iteration 2 (approx. per word bound = -7.605, relative change = 3.519e-02)
```



```

## .....
## Completed E-Step (3 seconds).
## Completed M-Step.
## Completing Iteration 3 (approx. per word bound = -7.573, relative change = 4.218e-03)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 4 (approx. per word bound = -7.559, relative change = 1.842e-03)
## .....
## Completed E-Step (3 seconds).
## Completed M-Step.
## Completing Iteration 5 (approx. per word bound = -7.550, relative change = 1.163e-03)
## Topic 1: women, show, one, like, design
## Topic 2: said, women, polic, report, year
## Topic 3: women, team, said, game, world
## Topic 4: year, book, first, one, mrs
## Topic 5: women, said, percent, femal, will
## Topic 6: said, women, one, peopl, woman
## Topic 7: women, work, said, year, men
## Topic 8: women, sexual, sex, rape, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, women, peopl
## Topic 11: women, film, one, said, woman
## Topic 12: said, polit, elect, parti, govern
## Topic 13: women, said, cancer, studi, health
## Topic 14: women, confer, said, deleg, will
## Topic 15: said, women, girl, rape, practic
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 6 (approx. per word bound = -7.544, relative change = 8.207e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 7 (approx. per word bound = -7.539, relative change = 5.962e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 8 (approx. per word bound = -7.536, relative change = 4.492e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 9 (approx. per word bound = -7.533, relative change = 3.482e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 10 (approx. per word bound = -7.531, relative change = 2.760e-04)
## Topic 1: show, design, women, fashion, art
## Topic 2: said, women, polic, report, offici
## Topic 3: women, team, said, game, world
## Topic 4: book, year, life, first, work
## Topic 5: women, said, femal, percent, will
## Topic 6: said, one, protest, peopl, site
## Topic 7: women, work, said, year, percent

```

```

## Topic 8: women, sexual, rape, sex, men
## Topic 9: women, said, right, law, govern
## Topic 10: said, one, famili, peopl, day
## Topic 11: women, film, one, like, woman
## Topic 12: polit, said, elect, parti, govern
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, confer, said, will, world
## Topic 15: said, women, girl, rape, case
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 11 (approx. per word bound = -7.530, relative change = 2.210e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 12 (approx. per word bound = -7.528, relative change = 1.806e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 13 (approx. per word bound = -7.527, relative change = 1.497e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 14 (approx. per word bound = -7.526, relative change = 1.266e-04)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 15 (approx. per word bound = -7.525, relative change = 1.059e-04)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, report
## Topic 3: women, team, game, said, world
## Topic 4: book, year, life, first, write
## Topic 5: women, said, femal, percent, militari
## Topic 6: protest, said, one, site, peopl
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, sex, rape, men
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, day
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, parti, minist, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: said, women, rape, court, case
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 16 (approx. per word bound = -7.525, relative change = 8.840e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 17 (approx. per word bound = -7.524, relative change = 7.478e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.

```

```

## Completing Iteration 18 (approx. per word bound = -7.524, relative change = 6.406e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 19 (approx. per word bound = -7.523, relative change = 5.618e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 20 (approx. per word bound = -7.523, relative change = 5.066e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, year, life, write, first
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, one, site
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, sex, men, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, parti, minist, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: said, women, case, court, rape
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 21 (approx. per word bound = -7.522, relative change = 4.664e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 22 (approx. per word bound = -7.522, relative change = 4.236e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 23 (approx. per word bound = -7.522, relative change = 3.813e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 24 (approx. per word bound = -7.522, relative change = 3.480e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 25 (approx. per word bound = -7.521, relative change = 3.249e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, women, kill, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, first
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like

```

```

## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, women, rape
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 26 (approx. per word bound = -7.521, relative change = 2.999e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 27 (approx. per word bound = -7.521, relative change = 2.756e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 28 (approx. per word bound = -7.521, relative change = 2.534e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 29 (approx. per word bound = -7.521, relative change = 2.341e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 30 (approx. per word bound = -7.520, relative change = 2.177e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, kill, women, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, novel
## Topic 5: women, said, femal, percent, compani
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, like
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, health
## Topic 14: women, said, confer, will, world
## Topic 15: case, court, said, women, rape
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 31 (approx. per word bound = -7.520, relative change = 2.009e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 32 (approx. per word bound = -7.520, relative change = 1.827e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 33 (approx. per word bound = -7.520, relative change = 1.673e-05)
## .....
## Completed E-Step (2 seconds).

```

```

## Completed M-Step.
## Completing Iteration 34 (approx. per word bound = -7.520, relative change = 1.570e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 35 (approx. per word bound = -7.520, relative change = 1.491e-05)
## Topic 1: show, design, fashion, women, art
## Topic 2: said, polic, kill, women, offici
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, stori
## Topic 5: women, said, femal, compani, percent
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, peopl, girl
## Topic 11: women, film, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, studi
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, rape, women
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 36 (approx. per word bound = -7.520, relative change = 1.388e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 37 (approx. per word bound = -7.520, relative change = 1.270e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 38 (approx. per word bound = -7.519, relative change = 1.125e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 39 (approx. per word bound = -7.519, relative change = 1.123e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 40 (approx. per word bound = -7.519, relative change = 1.086e-05)
## Topic 1: show, design, women, fashion, art
## Topic 2: said, polic, kill, offici, women
## Topic 3: women, team, game, said, world
## Topic 4: book, life, year, write, stori
## Topic 5: women, said, femal, compani, men
## Topic 6: protest, said, peopl, site, one
## Topic 7: women, work, said, year, percent
## Topic 8: women, sexual, men, sex, rape
## Topic 9: women, said, right, law, islam
## Topic 10: said, one, famili, girl, peopl
## Topic 11: film, women, one, like, woman
## Topic 12: polit, elect, minist, parti, presid
## Topic 13: women, said, abort, cancer, studi

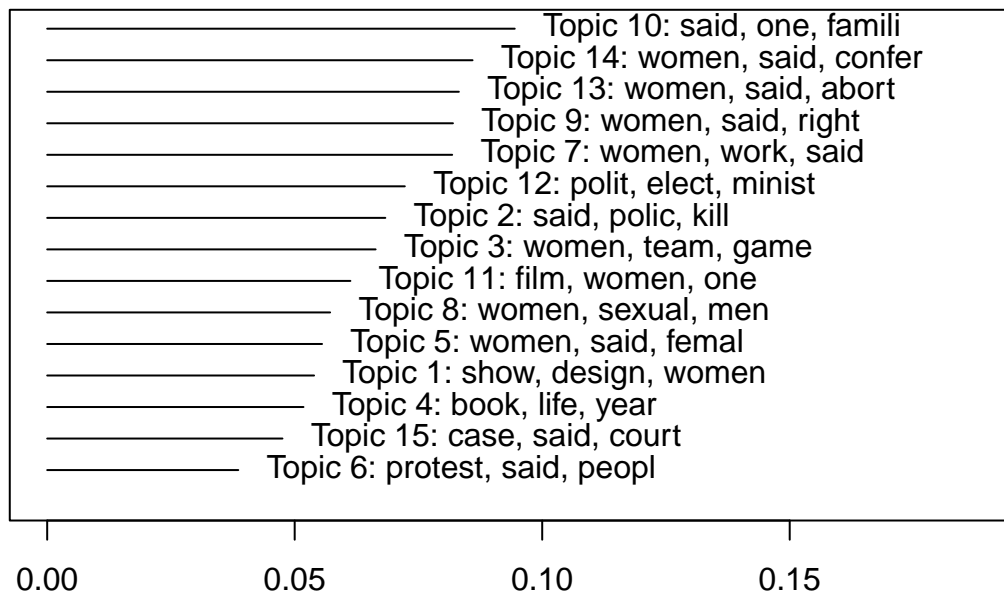
```

```
## Topic 14: women, said, confer, will, world
## Topic 15: case, said, court, rape, women
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Completing Iteration 41 (approx. per word bound = -7.519, relative change = 1.074e-05)
## .....
## Completed E-Step (2 seconds).
## Completed M-Step.
## Completing Iteration 42 (approx. per word bound = -7.519, relative change = 1.039e-05)
## .....
## Completed E-Step (1 seconds).
## Completed M-Step.
## Model Converged
```

```
# list top words for each topic
# labelTopics(model)
#
# # top documents for each topic
# findThoughts(model3, texts = meta$TITLE, n = 2, topics = c(1:15))

# summarize the corpus - find most common topics
plot.STM(model3)
```

## Top Topics



## Expected Topic Proportions

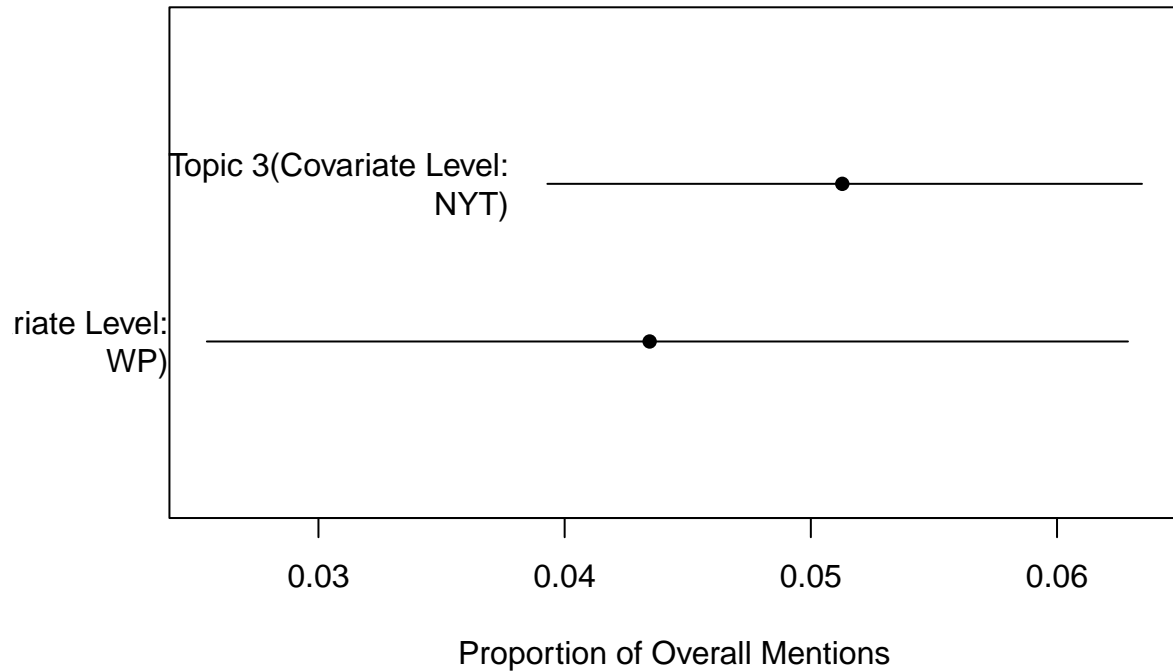
```
est <- estimateEffect(1:15 ~ REGION + PUBLICATION, meta = meta, model3)

# name regions
pubs <- c("NYT", "WashPo")
```

```
# create plot
```

```
plot.estimateEffect(x = est, covariate = "PUBLICATION", topics = 3, method = "pointestimate", xlab = "P
```

## Prevalence of Mentions of Women in Sport by Publication



Above I run the same code but instead for mentions by publication rather than region. I find that the NYT mentions women in sports more often than the Washington Post.