The Legacy of the European Post-Master in Urbanism at TU Delft: A Text Mining Approach

Jane Doe1,✉, and John Q. Doe1

March 17, 2023

Text of abstract

1 Delft University of Technology

✉ Correspondence: [Jane Doe <janedoe@fosg.org>](mailto:janedoe@fosg.org)

Keywords: urbanism; education; post-master; text mining

Highlights: These are the highlights.

# Introduction

The research presented in this paper was prompted by the closure of the European post-Master of Urbanism (EMU) of the Department of Urbanism at the Faculty of Architecture and the Built Environment, TU Delft. The EMU was an advanced master ran jointly by TU Delft, KU Leuven, UPC Barcelona and Università IUAV di Venezia.

Distinctive features of the program:  
- as a postmaster, it only accepted applicants with a prior MSc degree and experience in practice - provides a bridge towards PhD-level research - all students had an exchange semester at one of the other three program universities

In order to describe the legacy of the EMU program, including the distinctive features of its didactic approach, this paper aims to reveal the main topics taught in it and how those topics had evolved through the years of the program. To that end, we employed a text mining approach in which we analysed its output: 96 theses with an average of 1.5627219^{5} words produced over the years for the duration of the program between 2007-2021. The first year of the program was not represented, as PDF files were only available from 2008 onward.

# Methods

Research questions:

* What were the main topics addressed in the EMU program?
* How did those topics evolve throughout the duration of the program between 2007-2021?
* To what extent were the topics and evolution thereof influenced by the assignments given to the students throughout their studies leading up to their thesis?
* To what extent were the topics and evolution thereof influenced by the exchange semester followed by the students?

## Data collection

The theses are available in PDF format with complex layouts typical to an urbanism project in which text and various types of graphics are combined. With the exception of a bilingual English-Spanish thesis, all thesis were written in English.

A total number of 81 out of 96 were analysed.

|  |
| --- |
| Figure 1: Location of EMU theses included in the analysis |

Figure [Figure 1](#fig-map-theses) shows the geographic spread of the theses included in the analysis.

## Data analysis

(Silge and Robinson, 2017a) (Hvitfeldt and Silge, 2021)

### Tokenisation

Word-level tokenisation was applied with the unnest\_tokens() function of the tidytext package. This was preferred over other tokenisers as it produced results that work seamlessly with other tidyverse tools used in the analysis.

### Stop words

For analyses in which word-level tokenisation was chosen, stop words, i.e., common words or words that do not add much meaning, were removed in the data preparation stage. Global, subject-specific and document stop words (Hvitfeldt and Silge, 2021) were differentiated. As all theses were written in English, the pre-made “snowball” lexicon was used to remove global stop words. Subject-specific stop words, we manually constructed a list based on our domain knowledge. Words that are commonly used to structure a theses, such as “preface”, “contents”, and “introduction”, were added as subject-specific stop words as well. Finally, titles and author names were removed as document stop words on document-level.

stop\_words\_custom <- emu\_theses %>%   
 unnest\_tokens(output = word, input = title) %>% ### add words from the title  
 select(word) %>%   
 rbind(., data.frame(word = c("preface", "foreword", "introduction", "conclusion", "thesis",   
 "source", "author"))) %>%   
 # rbind(., data.frame(word = c("city", "urban", "urbanism"))) %>%   
 rbind(., data.frame(word = c("hab", "km")))  
  
emu\_theses\_words <- emu\_theses %>%  
 mutate(author = paste(first\_name, last\_name)) %>%   
 select(author, title, text, -c(graduation\_year, first\_name, last\_name)) %>%   
 unnest\_tokens(output = word, input = text) %>% # remove punctuation, convert to lowercase, seperate all words  
 anti\_join(stop\_words, by = "word") %>% # remove stop words  
 anti\_join(stop\_words\_custom, by = "word")

Stemming is a pre-processing step that needs to be carefully thought through, as it might either degrade topic modeling or produce no meaningful results (Hvitfeldt and Silge, 2021).

emu\_theses\_words %>%   
 mutate(stem = wordStem(words = word,   
 language = "porter")) %>% # Implement the Porter stemming algorithm for English provided by the SnowballC package  
 count(stem, sort = TRUE)

# A tibble: 47,027 × 2  
 stem n  
 <chr> <int>  
 1 fig 4220  
 2 1 3974  
 3 2 3787  
 4 plan 3391  
 5 structur 3290  
 6 natur 3238  
 7 3 3132  
 8 develop 3112  
 9 main 3026  
10 network 2978  
# … with 47,017 more rows

## Topic model

Similar to topic clustering on numeric data, topic modeling is a method of unsupervised classification of topics found in a group of documents (Silge and Robinson, 2017b).

# Results

# Which are the top n most frequently used words in each thesis?  
get\_top\_words\_per\_document(data = emu\_theses\_words,   
 top\_n = 3,   
 title\_col = "title",   
 word\_col = "word")

# A tibble: 251 × 3  
 title word n  
 <chr> <chr> <int>  
 1 'Living on the Edge': Water and the city of Pune fabr… 84  
 2 'Living on the Edge': Water and the city of Pune desh… 73  
 3 'Living on the Edge': Water and the city of Pune adit… 71  
 4 'Territories in between': A comparative permeability study of tw… tib 244  
 5 'Territories in between': A comparative permeability study of tw… ha 155  
 6 'Territories in between': A comparative permeability study of tw… 10 111  
 7 35°6′54″N 33°56′33″E Terra Incognita: Towards common imaginaries… fama… 147  
 8 35°6′54″N 33°56′33″E Terra Incognita: Towards common imaginaries… turk… 107  
 9 35°6′54″N 33°56′33″E Terra Incognita: Towards common imaginaries… 0 106  
10 A Benevolent Future: Towards a strategic spatial framework foste… rela… 216  
# … with 241 more rows

# Which are the top n most frequently used words in all theses`?  
get\_top\_words\_per\_corpus(data = emu\_theses\_words,   
 top\_n = 5,   
 word\_col = "word")

# A tibble: 5 × 2  
 word n  
 <chr> <int>  
1 fig 4220  
2 1 3974  
3 2 3787  
4 3 3132  
5 main 3013

# Which are the top n most frequently used words in all theses`?  
emu\_theses\_top\_20 <- get\_top\_words\_per\_corpus(data = emu\_theses\_words,   
 top\_n = 20,   
 word\_col = "word")

## TO DO: generalise bigrams and trigrams into ngrams function  
  
## Bigrams  
emu\_theses\_bigrams <- emu\_theses |>   
 unnest\_tokens(bigram, text, token = "ngrams", n = 2) |>   
 separate(bigram, into = c("first", "second"), sep = " ", remove = FALSE) |>   
 anti\_join(stop\_words, by = c("first" = "word")) |>   
 anti\_join(stop\_words, by = c("second" = "word")) |>   
 filter(str\_detect(first, "[a-z]") &   
 str\_detect(second, "[a-z]"))  
  
bigram\_counts <- emu\_theses\_bigrams |>   
 count(first, second, sort = TRUE)  
  
bigram\_counts

# A tibble: 242,226 × 3  
 first second n  
 <chr> <chr> <int>  
 1 public space 1114  
 2 public transport 778  
 3 public spaces 746  
 4 urban development 711  
 5 water management 471  
 6 urban fabric 455  
 7 cultural heritage 433  
 8 city centre 416  
 9 spatial planning 380  
10 tu delft 378  
# … with 242,216 more rows

## Trigrams  
emu\_theses\_trigrams <- emu\_theses |>   
 unnest\_tokens(trigram, text, token = "ngrams", n = 3) |>   
 separate(trigram, into = c("first", "second", "third"), sep = " ", remove = FALSE) |>   
 filter(!first %in% stop\_words$word,  
 !second %in% stop\_words$word,  
 !third %in% stop\_words$word) |>   
 filter(str\_detect(first, "[a-z]"),   
 str\_detect(second, "[a-z]"),  
 str\_detect(third, "[a-z]"))  
  
trigram\_counts <- emu\_theses\_trigrams |>   
 count(first, second, sort = TRUE)  
  
trigram\_counts

# A tibble: 129,662 × 3  
 first second n  
 <chr> <chr> <int>  
 1 public space 472  
 2 public transport 440  
 3 socio spatial 316  
 4 tu delft 311  
 5 source http 279  
 6 data source 272  
 7 water management 261  
 8 urban development 248  
 9 public spaces 232  
10 low income 211  
# … with 129,652 more rows

## Bigram tf-idf  
bigram\_tf\_idf <- emu\_theses\_bigrams |>   
 count(title, bigram) |>   
 bind\_tf\_idf(bigram, title, n) |>   
 arrange(desc(tf\_idf))  
  
bigram\_tf\_idf

# A tibble: 306,374 × 6  
 title bigram n tf idf tf\_idf  
 <chr> <chr> <int> <dbl> <dbl> <dbl>  
 1 "Living on the Edges of a Border City: Plan… el ro… 253 0.0322 4.38 0.141   
 2 "The Patchwork metropolis" patch… 124 0.0433 3.00 0.130   
 3 "Urban Resilience in Post-Disaster Reconstr… post … 66 0.0376 3.00 0.113   
 4 "Urban Landscape Infrastructures in Anuradh… sri l… 119 0.0251 4.38 0.110   
 5 "Towards a Critical Urbanism: Evaluation an… criti… 342 0.0219 4.38 0.0959  
 6 "Made in the city of China: Street market a… stree… 113 0.0308 3.00 0.0922  
 7 "Planning for the end of fossil fuel and be… energ… 129 0.0207 4.38 0.0905  
 8 "Contextualizing Edges, the rural case of S… emu20… 57 0.0202 4.38 0.0884  
 9 "Contextualizing Edges, the rural case of S… sanna… 57 0.0202 4.38 0.0884  
10 "Reimagining the New Density in Apartment U… apart… 239 0.0265 3.28 0.0870  
# … with 306,364 more rows

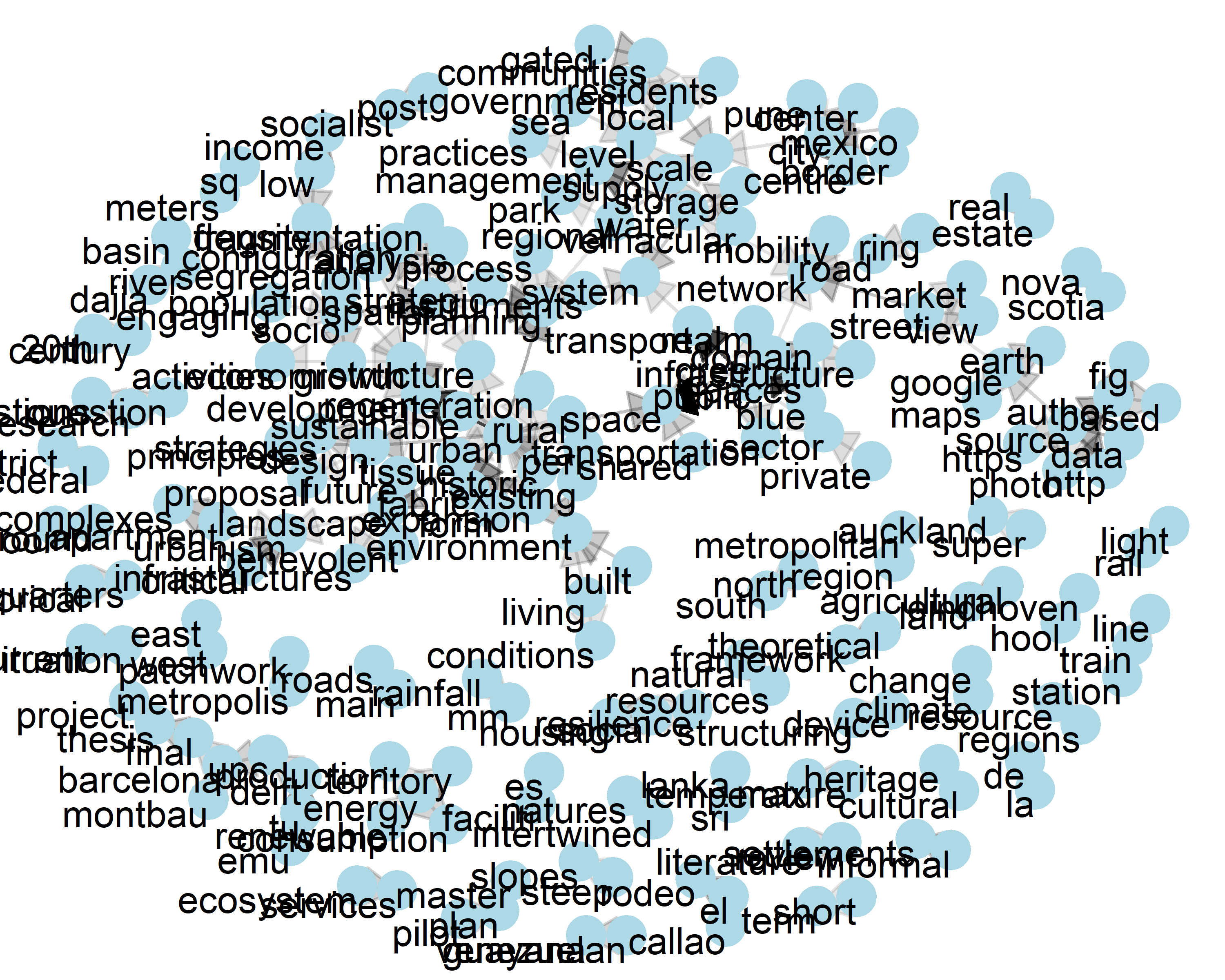
## Trigram tf-idf  
trigram\_tf\_idf <- emu\_theses\_trigrams |>   
 count(title, trigram) |>   
 bind\_tf\_idf(trigram, title, n) |>   
 arrange(desc(tf\_idf))  
  
trigram\_tf\_idf

# A tibble: 182,855 × 6  
 title trigram n tf idf tf\_idf  
 <chr> <chr> <int> <dbl> <dbl> <dbl>  
 1 "Contextualizing Edges, the rural case of … antoni… 57 0.0453 4.38 0.199  
 2 "Contextualizing Edges, the rural case of … sanna … 57 0.0453 4.38 0.199  
 3 "Urban Landscape Infrastructures in Anurad… anurad… 80 0.0389 4.38 0.171  
 4 "Urban Landscape Infrastructures in Anurad… urban … 125 0.0609 2.77 0.169  
 5 "Learning from Vernacular Water-management… vernac… 112 0.0371 4.38 0.163  
 6 "Learning from Vernacular Water-management… water … 112 0.0371 4.38 0.163  
 7 "Bucharest: Between North and South" claudi… 53 0.0355 4.38 0.156  
 8 "Queretaro:\r Exploring sustainable develo… delft … 152 0.0290 4.38 0.127  
 9 "Queretaro:\r Exploring sustainable develo… barcel… 145 0.0276 4.38 0.121  
10 "Queretaro:\r Exploring sustainable develo… upc ba… 145 0.0276 4.38 0.121  
# … with 182,845 more rows

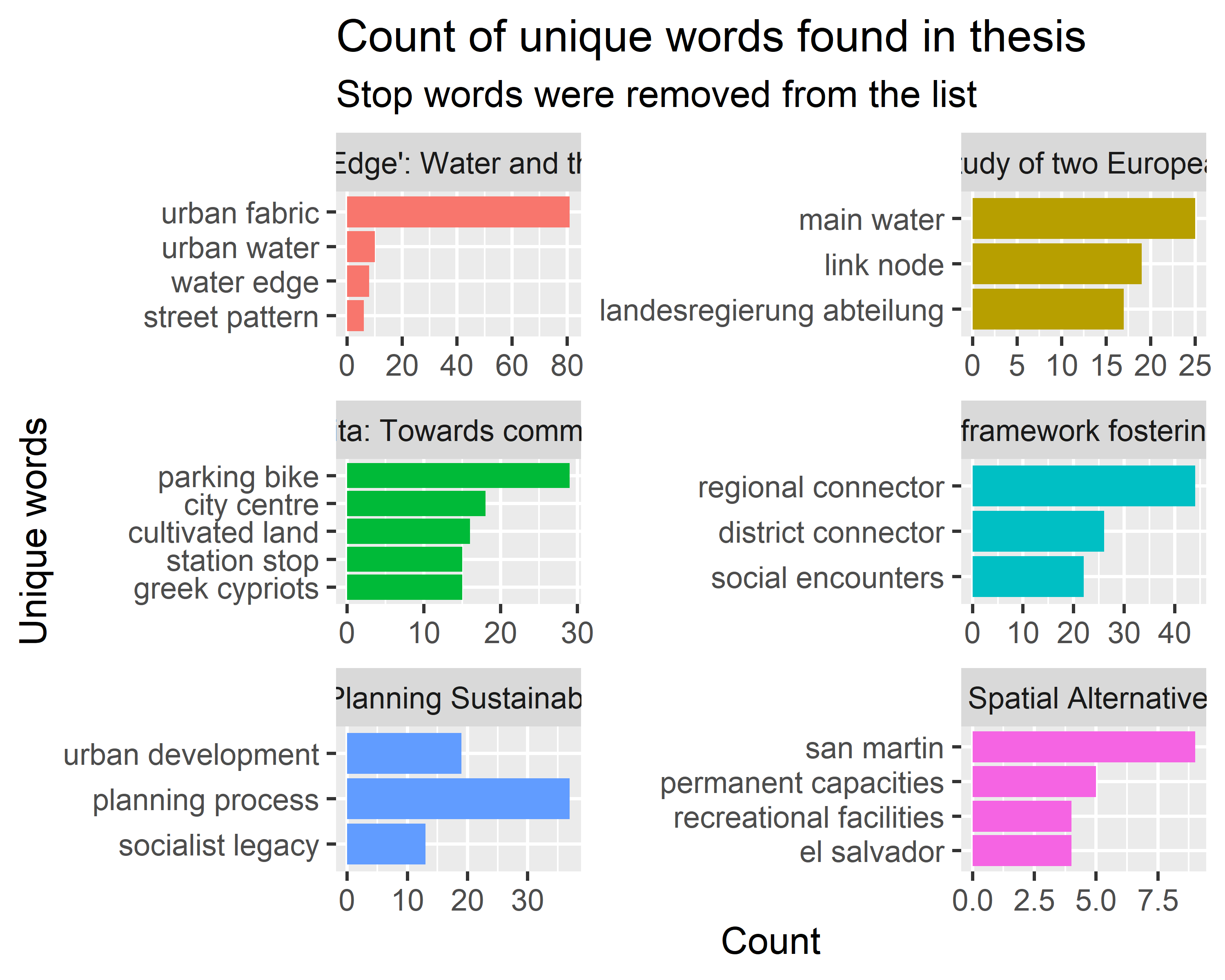
## Show network of bigrams  
bigram\_graph <- bigram\_counts |>   
 filter(n > 100) |>   
 graph\_from\_data\_frame()  
  
bigram\_graph

IGRAPH f05d221 DN-- 205 170 --   
+ attr: name (v/c), n (e/n)  
+ edges from f05d221 (vertex names):  
 [1] public ->space public ->transport public ->spaces   
 [4] urban ->development water ->management urban ->fabric   
 [7] cultural->heritage city ->centre spatial ->planning   
[10] tu ->delft urban ->planning critical->urbanism   
[13] socio ->spatial data ->source source ->author   
[16] urban ->form low ->income public ->domain   
[19] source ->http urban ->landscape urban ->space   
[22] street ->network climate ->change design ->principles   
+ ... omitted several edges

set.seed(2023)  
  
a <- grid::arrow(type = "closed", length = unit(.15, "inches"))  
   
ggraph(bigram\_graph, layout = "fr") +  
 geom\_edge\_link(aes(edge\_alpha = n), show.legend = FALSE, arrow = a) +  
 geom\_node\_point(color = "lightblue", size = 5) +  
 geom\_node\_text(aes(label = name), vjust = 1, hjust = 1) +  
 theme\_void()



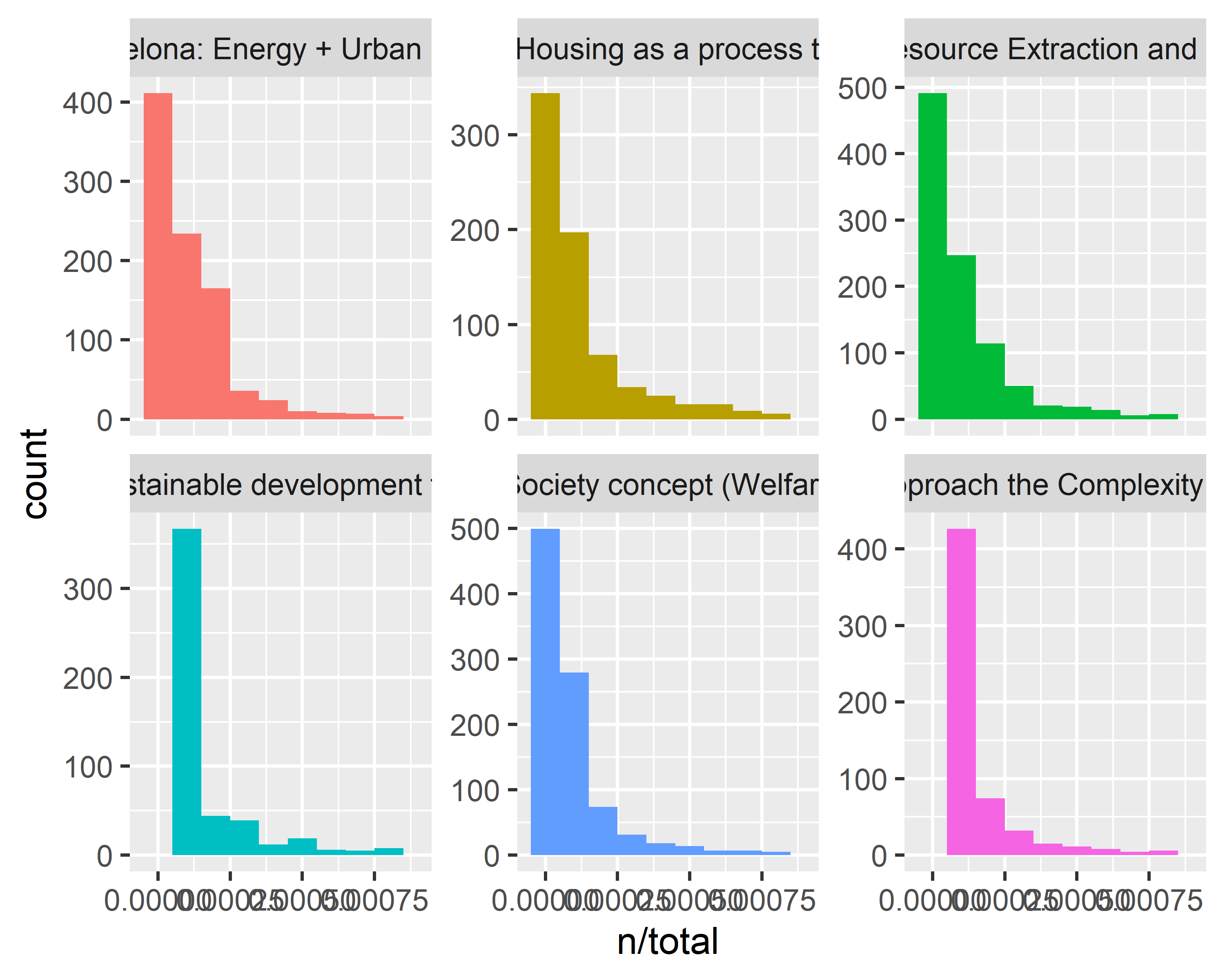
## Higgh frequency bigrams ----  
  
emu\_theses\_bigrams\_top <- emu\_theses\_bigrams %>%  
 group\_by(title) %>%  
 count(bigram, sort = TRUE) %>%  
 slice\_max(n, n = 20) %>% ### top\_n is superseded; better use slice\_max instead  
 # filter(n > 50) ### filter words with a minimum count  
 ungroup() %>%   
 mutate(bigram = reorder(bigram, n))  
  
emu\_theses\_bigrams\_top\_all <- emu\_theses\_bigrams %>%  
 count(bigram, sort = TRUE) %>%  
 slice\_max(n, n = 20) %>% ### top\_n is superseded; better use slice\_max instead  
 # filter(n > 50) ### filter words with a minimum count  
 mutate(bigram = reorder(bigram, n))  
  
emu\_theses\_bigrams\_top %>%  
 filter(title == unique(title)[1:6]) |>   
 ggplot(aes(x = bigram, y = n, fill = title)) +  
 geom\_col() +  
 xlab(NULL) +  
 coord\_flip() +  
 labs(y = "Count",  
 x = "Unique words",  
 title = "Count of unique words found in thesis",  
 subtitle = "Stop words were removed from the list") +  
 facet\_wrap( ~ title, scales = "free", ncol = 2) +  
 theme(legend.position = "none")



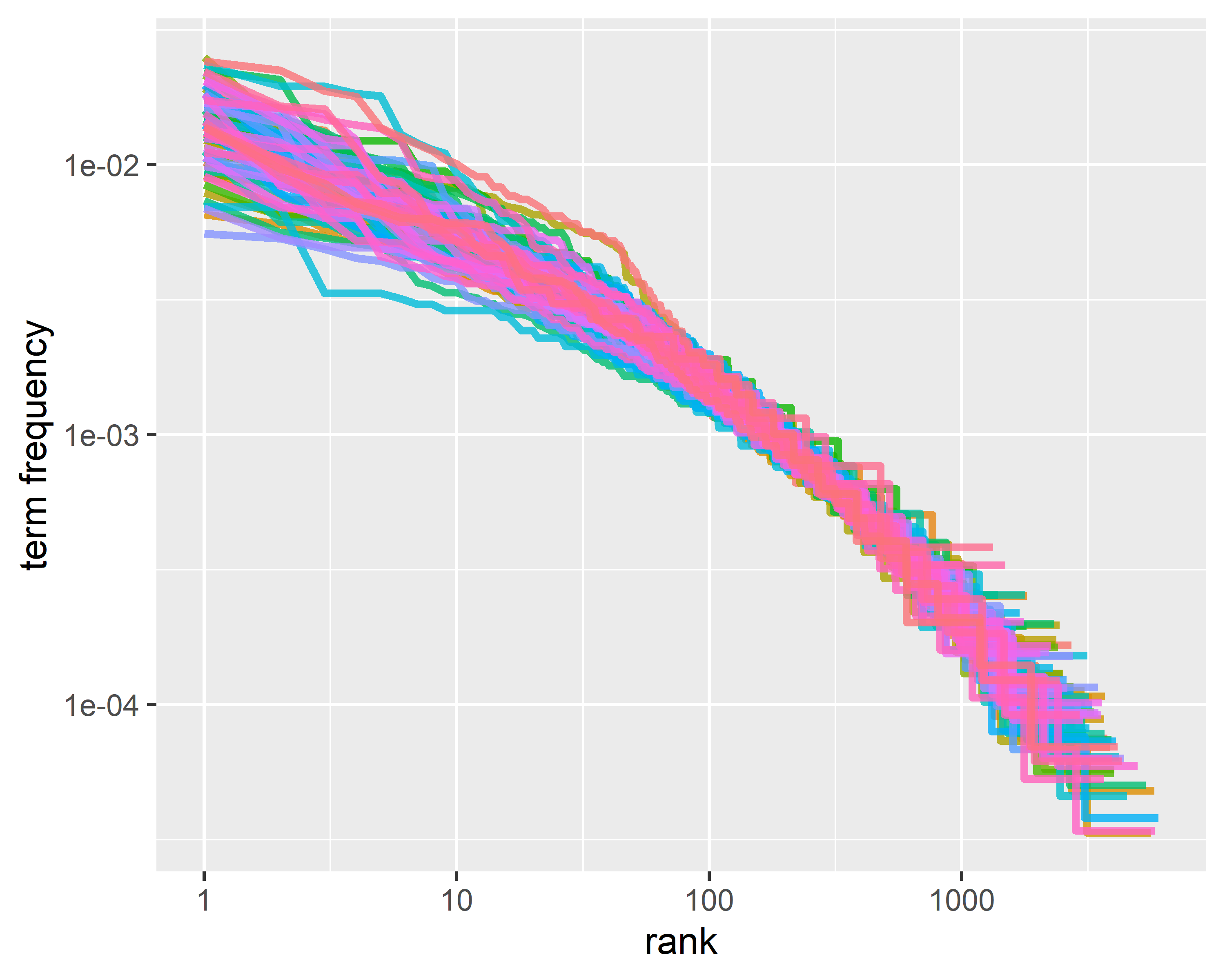
emu\_theses\_words

# A tibble: 860,226 × 3  
 author title word   
 <chr> <chr> <chr>   
 1 Lauren Abrahams Futures in Common? directions  
 2 Lauren Abrahams Futures in Common? shrinking   
 3 Lauren Abrahams Futures in Common? lauren   
 4 Lauren Abrahams Futures in Common? abrahams   
 5 Lauren Abrahams Futures in Common? emu   
 6 Lauren Abrahams Futures in Common? june   
 7 Lauren Abrahams Futures in Common? 2012   
 8 Lauren Abrahams Futures in Common? directions  
 9 Lauren Abrahams Futures in Common? shrinking   
10 Lauren Abrahams Futures in Common? lauren   
# … with 860,216 more rows

emu\_theses\_words\_count <- emu\_theses\_words |>   
 count(title, word, sort = TRUE) |>   
 ungroup()  
  
emu\_theses\_words\_total <- emu\_theses\_words\_count |>   
 group\_by(title) |>   
 summarise(total = sum(n))  
  
emu\_theses\_words\_count <- left\_join(emu\_theses\_words\_count, emu\_theses\_words\_total)  
  
# emu\_theses\_words\_count |>   
# group\_by(title) |>   
# summarise(total = unique(tot  
  
# Preview term frequency distributions in a subset of theses  
ggplot(emu\_theses\_words\_count |>   
 filter(title == unique(title)[1:6]),  
 aes(n/total, fill = title)) +  
 geom\_histogram(show.legend = FALSE, binwidth = 0.0001) +  
 xlim(NA, 0.0009) +  
 facet\_wrap(~ title, ncol = 3, scales = "free\_y")



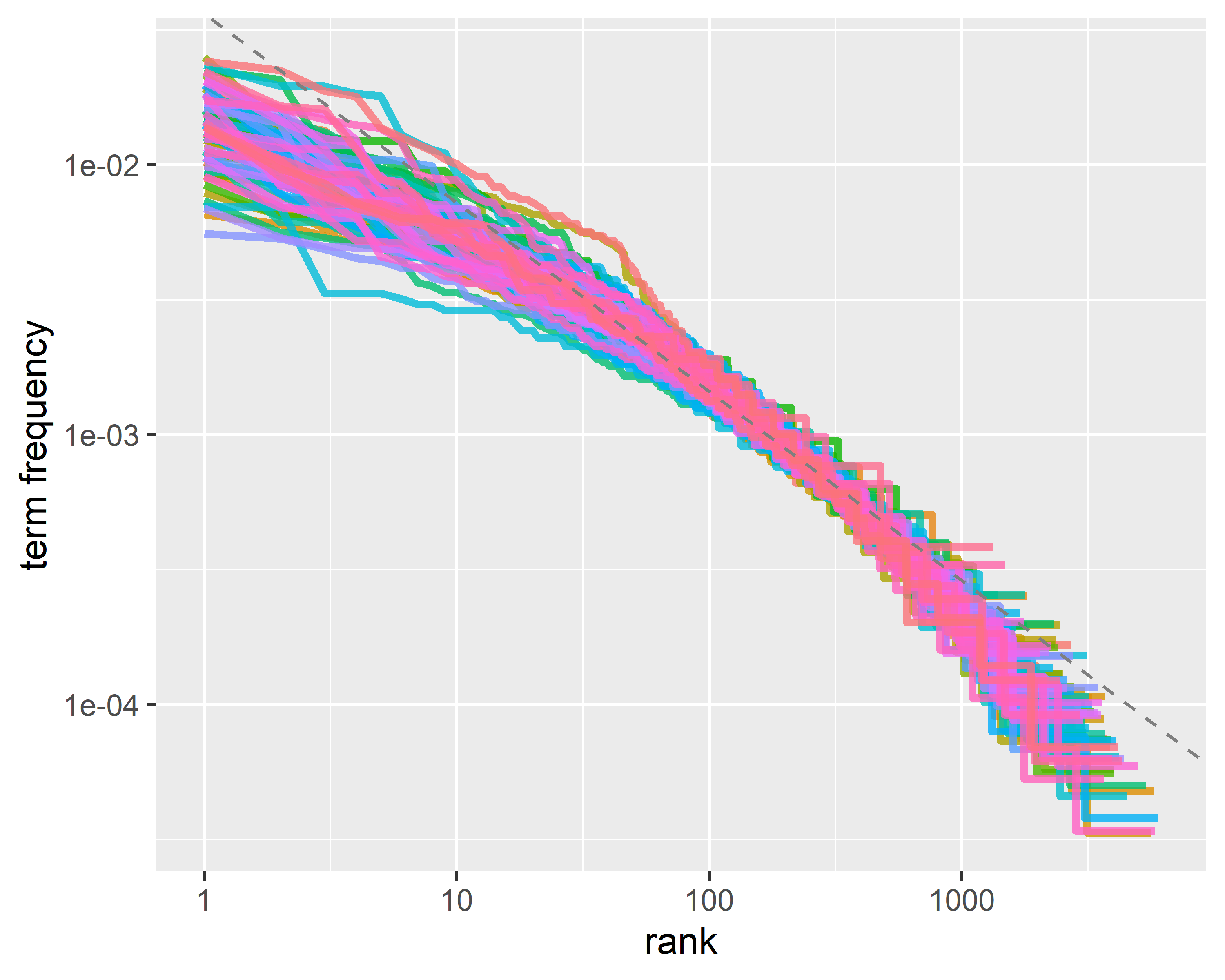
# Examine Zipf's law for the EMU theses  
## Calculate rank and term frequency  
freq\_by\_rank <- emu\_theses\_words\_count |>   
 group\_by(title) |>   
 mutate(rank = row\_number(),  
 `term frequency` = n/total)  
  
## Visualise Zipf's law  
freq\_by\_rank |>   
 ggplot(aes(rank, `term frequency`, color = title)) +  
 geom\_line(size = 1.1, alpha = 0.8, show.legend = FALSE) +  
 scale\_x\_log10() +  
 scale\_y\_log10()



rank\_subset <- freq\_by\_rank |>   
 filter(rank < 500,  
 rank > 10)  
  
lm(log10(`term frequency`) ~ log10(rank), data = rank\_subset)

Call:  
lm(formula = log10(`term frequency`) ~ log10(rank), data = rank\_subset)  
  
Coefficients:  
(Intercept) log10(rank)   
 -1.443 -0.701

freq\_by\_rank |>   
 ggplot(aes(rank, `term frequency`, color = title)) +  
 geom\_line(size = 1.1, alpha = 0.8, show.legend = FALSE) +  
 geom\_abline(intercept = -1.44, slope = -0.7, color = "gray50", linetype = 2) +  
 scale\_x\_log10() +  
 scale\_y\_log10()



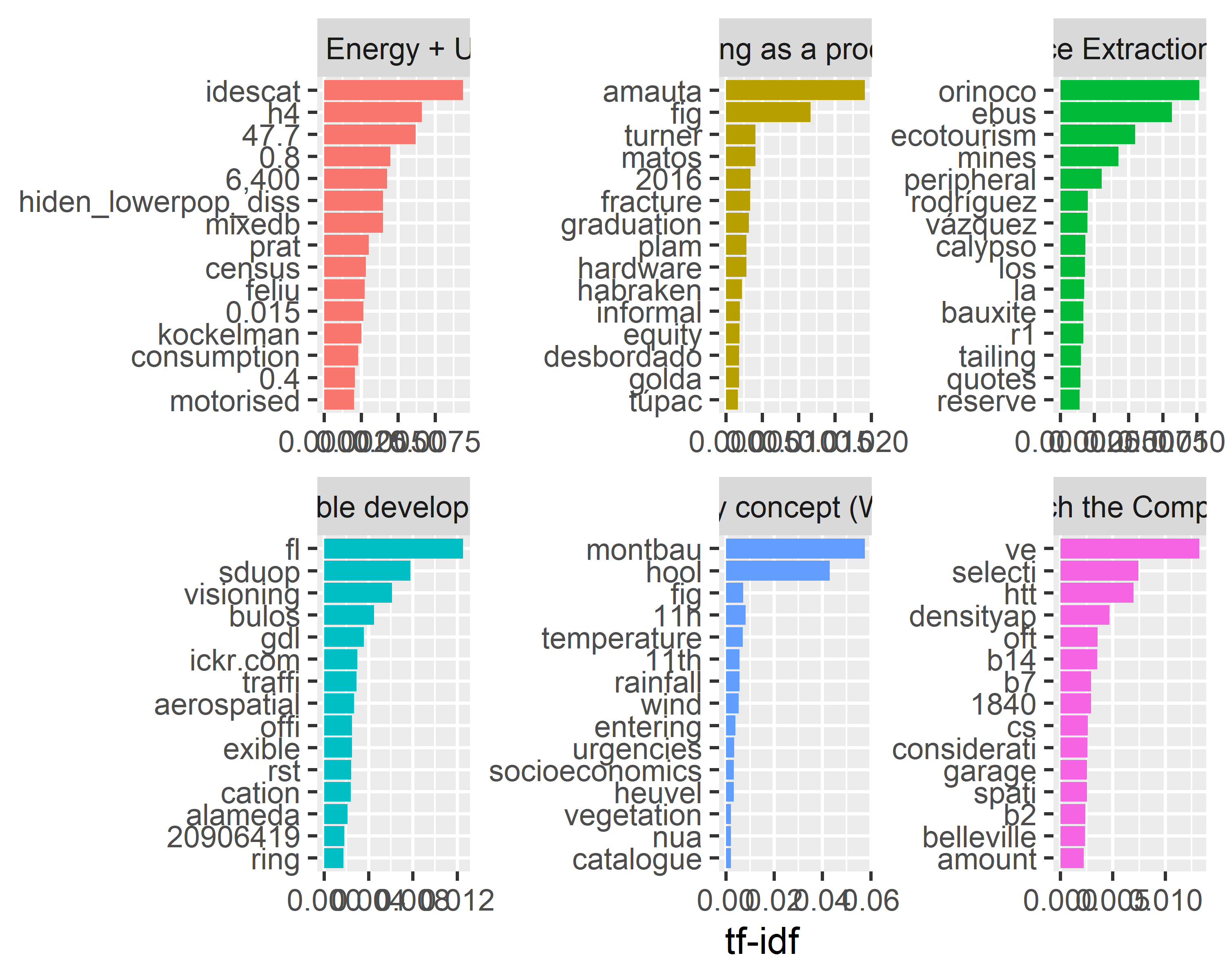
## Calculate tf-idf  
emu\_theses\_words\_count <- emu\_theses\_words\_count |>   
 bind\_tf\_idf(word, title, n)  
emu\_theses\_words\_count

# A tibble: 247,070 × 7  
 title word n total tf idf tf\_idf  
 <chr> <chr> <int> <int> <dbl> <dbl> <dbl>  
 1 "Towards a Critical Urbanism: Evalua… mont… 385 29316 0.0131 4.38 0.0575   
 2 "Barcelona: Energy + Urban Form" 0 326 29821 0.0109 0.119 0.00130  
 3 "Tracing Scopes of Action: Design Pr… 1 326 18818 0.0173 0 0   
 4 "Looking for alternatives in the cit… fig 317 21773 0.0146 0.799 0.0116   
 5 "Tracing Scopes of Action: Design Pr… 2 305 18818 0.0162 0 0   
 6 "Barcelona: Energy + Urban Form" 4 300 29821 0.0101 0 0   
 7 "Towards a Critical Urbanism: Evalua… hool 288 29316 0.00982 4.38 0.0430   
 8 "Overcoming the Paradox of Plenty: R… mini… 278 26297 0.0106 1.49 0.0158   
 9 "Tracing Scopes of Action: Design Pr… 3 276 18818 0.0147 0 0   
10 "Queretaro:\r Exploring sustainable … mobi… 271 14662 0.0185 0.163 0.00300  
# … with 247,060 more rows

emu\_theses\_words\_count |>   
 select(-total) |>   
 arrange(desc(tf\_idf))

# A tibble: 247,070 × 6  
 title word n tf idf tf\_idf  
 <chr> <chr> <int> <dbl> <dbl> <dbl>  
 1 "'Territories in between': A comparative per… tib 244 0.0183 3.69 0.0675  
 2 "Urban Resilience in Post-Disaster Reconstru… duji… 36 0.0138 4.38 0.0604  
 3 "Tehran has a Coast: 35N51E2050" casp… 179 0.0182 3.28 0.0598  
 4 "Halt! [Stay where you are]: There is no lan… nico… 80 0.0160 3.69 0.0589  
 5 "Towards a Critical Urbanism: Evaluation and… mont… 385 0.0131 4.38 0.0575  
 6 "Unfolding urban publicness" enve… 40 0.0131 4.38 0.0575  
 7 "De-fragmenting Athens: Drosscape as a devic… dross 184 0.0214 2.59 0.0555  
 8 "Habitat Embroidery: An Interwoven Landscape… flev… 160 0.0126 4.38 0.0551  
 9 "Industrial Vulnerability" nava… 160 0.0125 4.38 0.0549  
10 "Pilgrimage, power and identity of the place… shri… 230 0.0183 3.00 0.0548  
# … with 247,060 more rows

## Visualise the top n words with the highest tf-idf values for each of a subset of theses  
emu\_theses\_words\_count |>   
 filter(title == unique(title)[1:6]) |>   
 arrange(desc(tf\_idf)) |>   
 mutate(word = factor(word, levels = rev(unique(word)))) |>   
 group\_by(title) |>   
 top\_n(15) |>   
 ungroup() |>   
 ggplot(aes(word, tf\_idf, fill = title)) +  
 geom\_col(show.legend = FALSE) +  
 labs(x = NULL, y = "tf-idf") +  
 facet\_wrap(~ title, ncol = 3, scales = "free") +  
 coord\_flip()



The tf-idf statistic shows high values for the names of thesis locations that tend to be high-frequency thesis-specific words. As such, it shows words that are the most important to one document in a collection of documents (Silge and Robinson, 2017b).

# Discussion

# Conclusion

# Acknowledgements

# References

Hvitfeldt, E., Silge, J., 2021. Supervised Machine Learning for Text Analysis in R, 1st edition. ed. Chapman and Hall/CRC, Boca Raton.

Silge, J., Robinson, D., 2017a. Text Mining with R: A Tidy Approach, 1st edition. ed. O’Reilly Media, Beijing ; Boston.

Silge, J., Robinson, D., 2017b. Text mining with r: A tidy approach, 1st edition. ed. O’Reilly Media, Beijing ; Boston.

### Colophon

This report was generated on 2023-03-17 16:48:03 using the following computational environment and dependencies:

# which R packages and versions?  
if ("devtools" %in% installed.packages()) devtools::session\_info()

─ Session info ───────────────────────────────────────────────────────────────  
 setting value  
 version R version 4.2.0 (2022-04-22 ucrt)  
 os Windows 10 x64 (build 19044)  
 system x86\_64, mingw32  
 ui RTerm  
 language (EN)  
 collate English\_United Kingdom.utf8  
 ctype English\_United Kingdom.utf8  
 tz Europe/Berlin  
 date 2023-03-17  
 pandoc 2.19.2 @ C:/Program Files/RStudio/bin/quarto/bin/tools/ (via rmarkdown)  
  
─ Packages ───────────────────────────────────────────────────────────────────  
 ! package \* version date (UTC) lib source  
 P assertthat 0.2.1 2019-03-21 [?] CRAN (R 4.2.2)  
 P backports 1.4.1 2021-12-13 [?] CRAN (R 4.2.0)  
 P bit 4.0.5 2022-11-15 [?] CRAN (R 4.2.2)  
 P bit64 4.0.5 2020-08-30 [?] CRAN (R 4.2.2)  
 P broom 1.0.2 2022-12-15 [?] CRAN (R 4.2.2)  
 P cachem 1.0.6 2021-08-19 [?] CRAN (R 4.2.0)  
 P callr 3.7.3 2022-11-02 [?] CRAN (R 4.2.2)  
 P cellranger 1.1.0 2016-07-27 [?] CRAN (R 4.2.2)  
 P cli 3.6.0 2023-01-09 [?] CRAN (R 4.2.2)  
 P codetools 0.2-18 2020-11-04 [3] CRAN (R 4.2.0)  
 P colorspace 2.0-3 2022-02-21 [?] CRAN (R 4.2.2)  
 crayon 1.5.2 2022-09-29 [1] CRAN (R 4.2.2)  
 P curl 5.0.0 2023-01-12 [?] CRAN (R 4.2.2)  
 P DBI 1.1.3 2022-06-18 [?] CRAN (R 4.2.2)  
 P dbplyr 2.3.0 2023-01-16 [?] CRAN (R 4.2.0)  
 P desc 1.4.2 2022-09-08 [?] CRAN (R 4.2.0)  
 P devtools 2.4.5 2022-10-11 [?] CRAN (R 4.2.2)  
 P digest 0.6.31 2022-12-11 [?] CRAN (R 4.2.2)  
 P dplyr \* 1.0.10 2022-09-01 [?] CRAN (R 4.2.2)  
 P ellipsis 0.3.2 2021-04-29 [?] CRAN (R 4.2.0)  
 P evaluate 0.20 2023-01-17 [?] CRAN (R 4.2.2)  
 P fansi 1.0.4 2023-01-22 [?] CRAN (R 4.2.2)  
 P farver 2.1.1 2022-07-06 [?] CRAN (R 4.2.2)  
 P fastmap 1.1.0 2021-01-25 [?] CRAN (R 4.2.0)  
 P forcats \* 0.5.2 2022-08-19 [?] CRAN (R 4.2.2)  
 P fs 1.6.1 2023-02-06 [?] CRAN (R 4.2.2)  
 P gargle 1.2.1 2022-09-08 [?] CRAN (R 4.2.2)  
 P generics 0.1.3 2022-07-05 [?] CRAN (R 4.2.2)  
 P ggforce 0.4.1 2022-10-04 [?] CRAN (R 4.2.2)  
 P ggplot2 \* 3.4.0 2022-11-04 [?] CRAN (R 4.2.2)  
 P ggraph \* 2.1.0 2022-10-09 [?] CRAN (R 4.2.2)  
 P ggrepel 0.9.3 2023-02-03 [?] CRAN (R 4.2.2)  
 P glue 1.6.2 2022-02-24 [?] CRAN (R 4.2.0)  
 P googledrive 2.0.0 2021-07-08 [?] CRAN (R 4.2.2)  
 P googlesheets4 1.0.1 2022-08-13 [?] CRAN (R 4.2.2)  
 P graphlayouts 0.8.4 2022-11-24 [?] CRAN (R 4.2.2)  
 P gridExtra 2.3 2017-09-09 [?] CRAN (R 4.2.2)  
 P gtable 0.3.1 2022-09-01 [?] CRAN (R 4.2.2)  
 P haven 2.5.1 2022-08-22 [?] CRAN (R 4.2.2)  
 P here \* 1.0.1 2020-12-13 [?] CRAN (R 4.2.2)  
 P hms 1.1.2 2022-08-19 [?] CRAN (R 4.2.2)  
 htmltools 0.5.4 2022-12-07 [1] CRAN (R 4.2.2)  
 P htmlwidgets 1.6.1 2023-01-07 [?] CRAN (R 4.2.2)  
 P httpuv 1.6.9 2023-02-14 [?] CRAN (R 4.2.2)  
 P httr 1.4.4 2022-08-17 [?] CRAN (R 4.2.1)  
 P igraph \* 1.4.0 2023-02-10 [?] CRAN (R 4.2.0)  
 P janeaustenr 1.0.0 2022-08-26 [?] CRAN (R 4.2.2)  
 P jsonlite 1.8.4 2022-12-06 [?] CRAN (R 4.2.2)  
 P knitr 1.42 2023-01-25 [?] CRAN (R 4.2.2)  
 P labeling 0.4.2 2020-10-20 [?] CRAN (R 4.2.0)  
 P later 1.3.0 2021-08-18 [?] CRAN (R 4.2.0)  
 P lattice 0.20-45 2021-09-22 [3] CRAN (R 4.2.0)  
 P lifecycle 1.0.3 2022-10-07 [?] CRAN (R 4.2.2)  
 P lubridate 1.9.0 2022-11-06 [?] CRAN (R 4.2.2)  
 P magrittr 2.0.3 2022-03-30 [?] CRAN (R 4.2.0)  
 P maps 3.4.1 2022-10-30 [?] CRAN (R 4.2.2)  
 P MASS 7.3-58.1 2022-08-03 [?] CRAN (R 4.2.2)  
 P Matrix 1.5-1 2022-09-13 [?] CRAN (R 4.2.2)  
 P memoise 2.0.1 2021-11-26 [?] CRAN (R 4.2.0)  
 P mime 0.12 2021-09-28 [?] CRAN (R 4.2.0)  
 P miniUI 0.1.1.1 2018-05-18 [?] CRAN (R 4.2.2)  
 P minTEMU \* 0.0.0.9000 2023-02-13 [?] load\_all()  
 P modelr 0.1.10 2022-11-11 [?] CRAN (R 4.2.2)  
 P modeltools 0.2-23 2020-03-05 [?] CRAN (R 4.2.0)  
 P munsell 0.5.0 2018-06-12 [?] CRAN (R 4.2.2)  
 P NLP \* 0.2-1 2020-10-14 [?] CRAN (R 4.2.0)  
 P pillar 1.8.1 2022-08-19 [?] CRAN (R 4.2.1)  
 P pkgbuild 1.4.0 2022-11-27 [?] CRAN (R 4.2.2)  
 P pkgconfig 2.0.3 2019-09-22 [?] CRAN (R 4.2.0)  
 P pkgload 1.3.2 2022-11-16 [?] CRAN (R 4.2.2)  
 P png 0.1-8 2022-11-29 [?] CRAN (R 4.2.2)  
 P polyclip 1.10-4 2022-10-20 [?] CRAN (R 4.2.1)  
 P prettyunits 1.1.1 2020-01-24 [?] CRAN (R 4.2.0)  
 processx 3.8.0 2022-10-26 [1] CRAN (R 4.2.2)  
 P profvis 0.3.7 2020-11-02 [?] CRAN (R 4.2.2)  
 P promises 1.2.0.1 2021-02-11 [?] CRAN (R 4.2.0)  
 ps 1.7.2 2022-10-26 [1] CRAN (R 4.2.2)  
 P purrr \* 1.0.1 2023-01-10 [?] CRAN (R 4.2.2)  
 P R6 2.5.1 2021-08-19 [?] CRAN (R 4.2.0)  
 P Rcpp 1.0.10 2023-01-22 [?] CRAN (R 4.2.2)  
 P readr \* 2.1.3 2022-10-01 [?] CRAN (R 4.2.2)  
 P readxl 1.4.1 2022-08-17 [?] CRAN (R 4.2.2)  
 P remotes 2.4.2 2021-11-30 [?] CRAN (R 4.2.0)  
 P reprex 2.0.2 2022-08-17 [?] CRAN (R 4.2.2)  
 P reticulate 1.28 2023-01-27 [?] CRAN (R 4.2.2)  
 P rlang 1.0.6 2022-09-24 [?] CRAN (R 4.2.2)  
 P rmarkdown 2.20 2023-01-19 [?] CRAN (R 4.2.2)  
 P rprojroot 2.0.3 2022-04-02 [?] CRAN (R 4.2.0)  
 P rstudioapi 0.14 2022-08-22 [?] CRAN (R 4.2.1)  
 P rvest 1.0.3 2022-08-19 [?] CRAN (R 4.2.2)  
 P scales 1.2.1 2022-08-20 [?] CRAN (R 4.2.2)  
 P sessioninfo 1.2.2 2021-12-06 [?] CRAN (R 4.2.0)  
 P shiny 1.7.4 2022-12-15 [?] CRAN (R 4.2.2)  
 P slam 0.1-50 2022-01-08 [?] CRAN (R 4.2.0)  
 P SnowballC \* 0.7.0 2020-04-01 [?] CRAN (R 4.2.0)  
 P stringi 1.7.12 2023-01-11 [?] CRAN (R 4.2.2)  
 P stringr \* 1.5.0 2022-12-02 [?] CRAN (R 4.2.2)  
 P tibble \* 3.1.8 2022-07-22 [?] CRAN (R 4.2.1)  
 P tidygeocoder \* 1.0.5 2021-11-02 [?] CRAN (R 4.2.2)  
 P tidygraph 1.2.3 2023-02-01 [?] CRAN (R 4.2.2)  
 P tidyr \* 1.2.1 2022-09-08 [?] CRAN (R 4.2.2)  
 P tidyselect 1.2.0 2022-10-10 [?] CRAN (R 4.2.2)  
 P tidytext \* 0.4.1 2023-01-07 [?] CRAN (R 4.2.2)  
 P tidyverse \* 1.3.2 2022-07-18 [?] CRAN (R 4.2.2)  
 P timechange 0.2.0 2023-01-11 [?] CRAN (R 4.2.2)  
 P tm \* 0.7-10 2022-12-14 [?] CRAN (R 4.2.2)  
 P tokenizers 0.3.0 2022-12-22 [?] CRAN (R 4.2.2)  
 P topicmodels \* 0.2-13 2022-12-06 [?] CRAN (R 4.2.2)  
 P tweenr 2.0.2 2022-09-06 [?] CRAN (R 4.2.2)  
 P tzdb 0.3.0 2022-03-28 [?] CRAN (R 4.2.2)  
 P urlchecker 1.0.1 2021-11-30 [?] CRAN (R 4.2.2)  
 P usethis 2.1.6 2022-05-25 [?] CRAN (R 4.2.1)  
 P utf8 1.2.3 2023-01-31 [?] CRAN (R 4.2.2)  
 P vctrs 0.5.2 2023-01-23 [?] CRAN (R 4.2.2)  
 P viridis 0.6.2 2021-10-13 [?] CRAN (R 4.2.2)  
 P viridisLite 0.4.1 2022-08-22 [?] CRAN (R 4.2.2)  
 P vroom 1.6.1 2023-01-22 [?] CRAN (R 4.2.2)  
 P withr 2.5.0 2022-03-03 [?] CRAN (R 4.2.0)  
 P xfun 0.37 2023-01-31 [?] CRAN (R 4.2.2)  
 P xml2 1.3.3 2021-11-30 [?] CRAN (R 4.2.0)  
 P xtable 1.8-4 2019-04-21 [?] CRAN (R 4.2.0)  
 P yaml 2.3.7 2023-01-23 [?] CRAN (R 4.2.2)  
  
 [1] C:/Users/awilczynski/AppData/Local/R/cache/R/renv/library/minTEMU-78ecd818/R-4.2/x86\_64-w64-mingw32  
 [2] C:/Users/awilczynski/AppData/Local/Temp/RtmpiU9t3N/renv-system-library  
 [3] C:/Users/awilczynski/AppData/Local/Programs/R/R-4.2.0/library  
  
 P ── Loaded and on-disk path mismatch.  
  
─ Python configuration ───────────────────────────────────────────────────────  
 python: C:/Users/awilczynski/Desktop/DCC-projects/minTEMU/inst/python/pyvenv/Scripts/python.exe  
 libpython: C:/Users/awilczynski/AppData/Local/r-reticulate/r-reticulate/pyenv/pyenv-win/versions/3.9.7/python39.dll  
 pythonhome: C:/Users/awilczynski/Desktop/DCC-projects/minTEMU/inst/python/pyvenv  
 version: 3.9.7 (tags/v3.9.7:1016ef3, Aug 30 2021, 20:19:38) [MSC v.1929 64 bit (AMD64)]  
 Architecture: 64bit  
 numpy: C:/Users/awilczynski/Desktop/DCC-projects/minTEMU/inst/python/pyvenv/Lib/site-packages/numpy  
 numpy\_version: 1.24.2  
   
 NOTE: Python version was forced by use\_python function  
  
──────────────────────────────────────────────────────────────────────────────

The current Git commit details are:

# what commit is this file at?   
if ("git2r" %in% installed.packages() & git2r::in\_repository(path = ".")) git2r::repository(here::here())

Local: 26-clean-aw C:/Users/awilczynski/Desktop/DCC-projects/minTEMU  
Remote: 26-clean-aw @ origin (git@github.com:UD3-Lab/minTEMU.git)  
Head: [51caab3] 2023-03-17: explicitly stated package name in clean.R