Metadata from ERB

ERB metadata

Abstract

Estonian National Bibliography is a metadata set that aims to collect information on all publications written in any language in Estonia and all texts written in Estonian in whichever country. In this set, only the publications in Estonian language are used. The dataset has been compiled in digital format since 2002. and aggregates the work of multiple institutions and generations in collecting the publication information.

This dataset presents the Estonian National Bibliography dataset in wide instead of long format used in Marc21, with some of the variables that may be useful for text-mining studies. It includes information about

some variables have also been standardized

Information on coding the variables can be found here:

Helpful information on the metadata available can be found here: http://data.digar.ee/#page5

The rules followed in adding information on older books can be found here: http://www.elnet.ee/images/pdf/juhendid/vanaraamat_MARC21.pdf

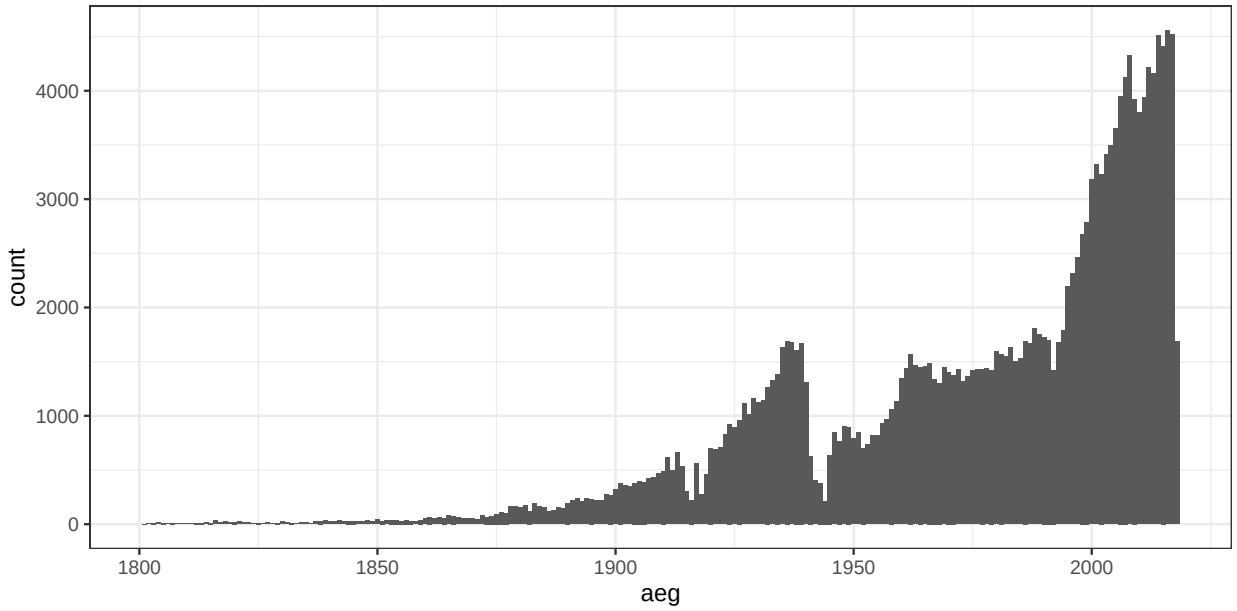
Intro

Converts the dataset into tidy format

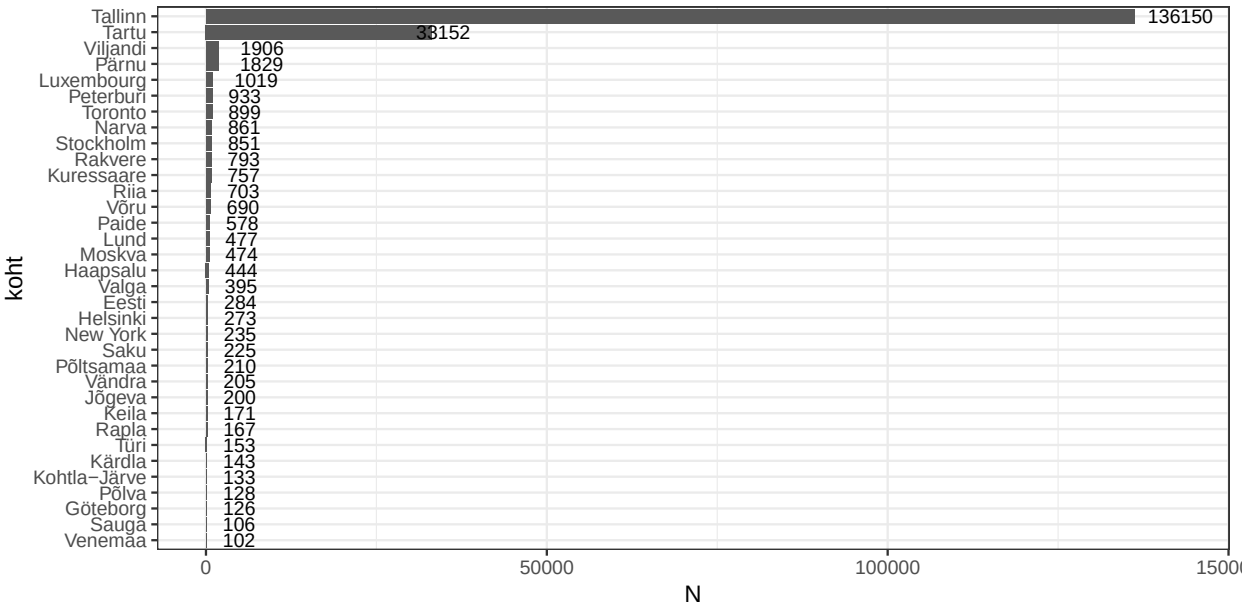
```
## -- Attaching packages ---------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.2
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0

## -- Conflicts ------------------------------------------------------------ tidyverse_conflicts() --
## x dplyr::between()   masks data.table::between()
## x dplyr::filter()    masks stats::filter()
## x dplyr::first()     masks data.table::first()
## x dplyr::lag()       masks stats::lag()
## x dplyr::last()      masks data.table::last()
## x purrr::transpose() masks data.table::transpose()

## here() starts at /media/kalmaar/ADATA SD700/ownCloud/Documents/R/R_corpus_analysis/corp_analysis_repo
```

Summary

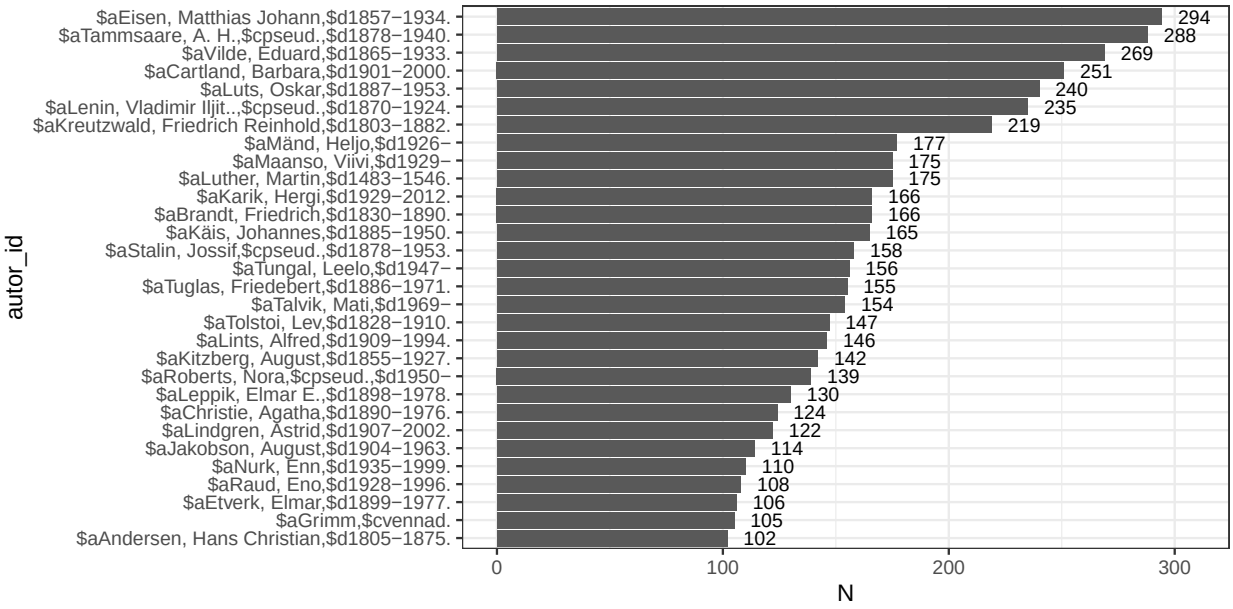
The dataset has altogether information on 193848 printed items. Over time the distribution is the following.



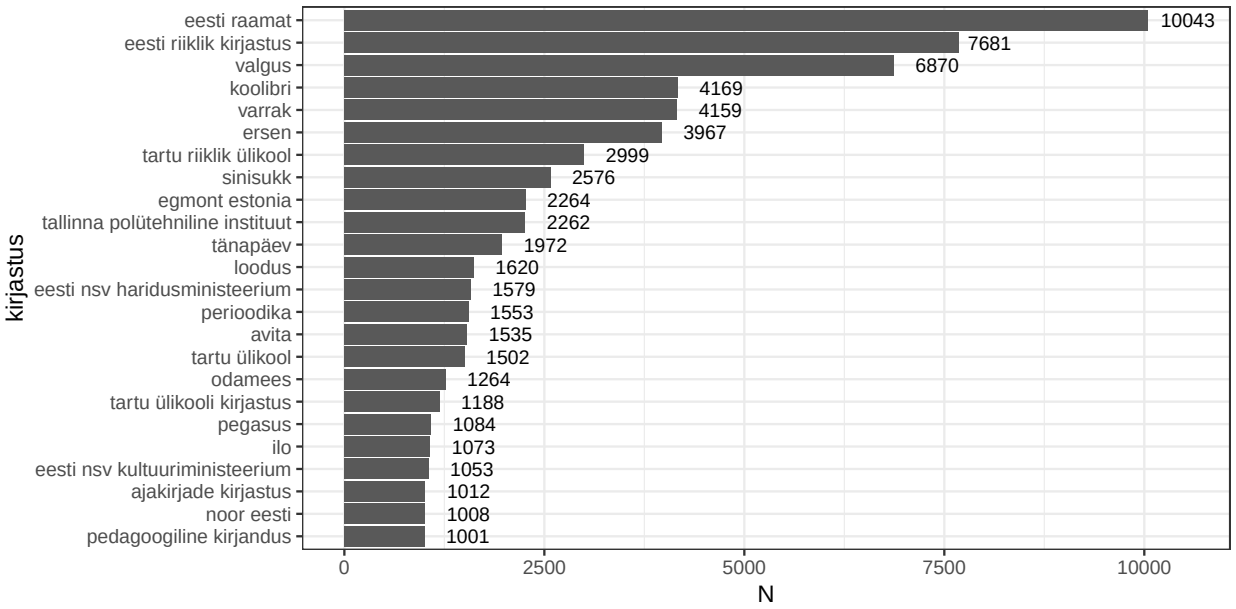
data-1.bb



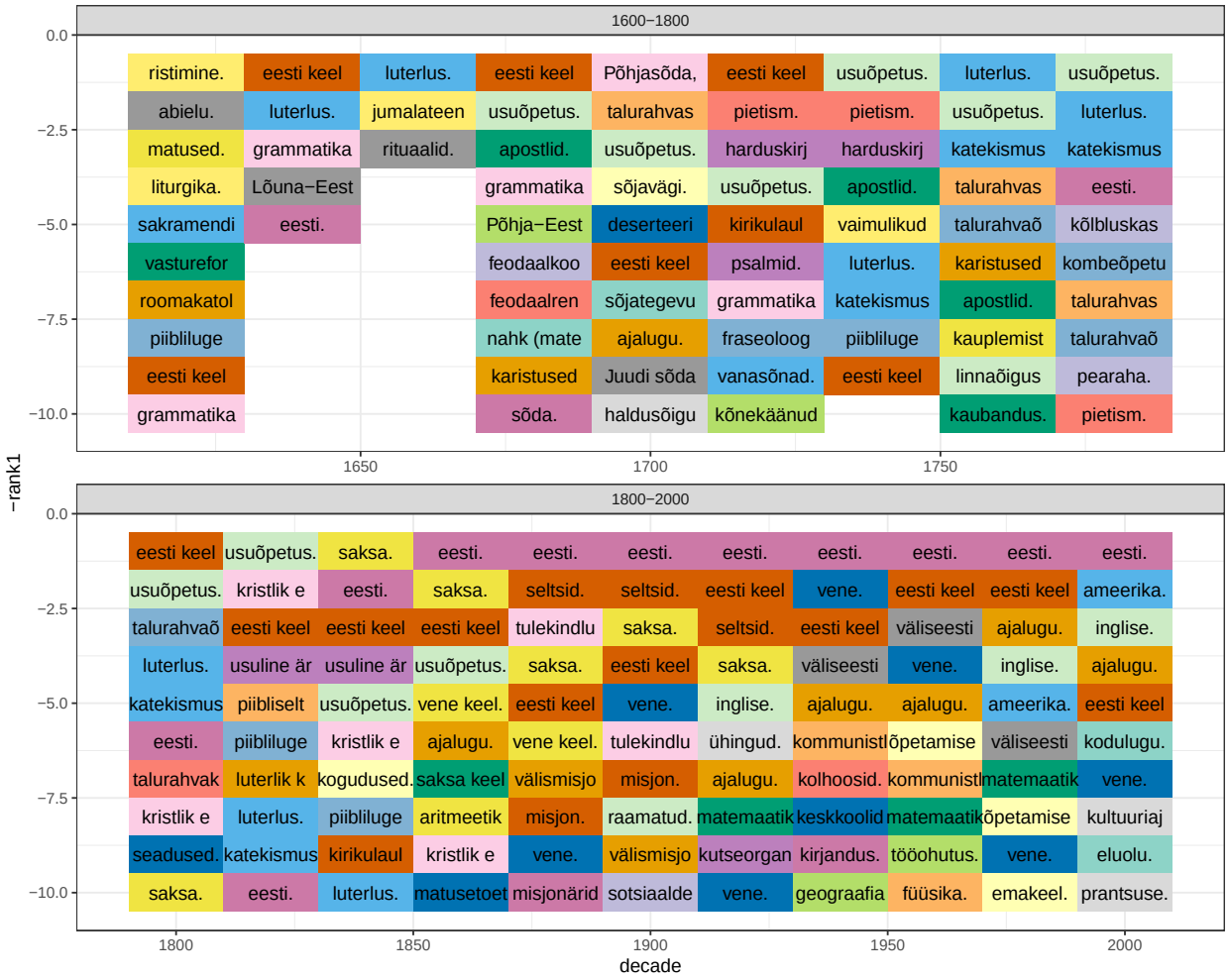
cities-1.bb

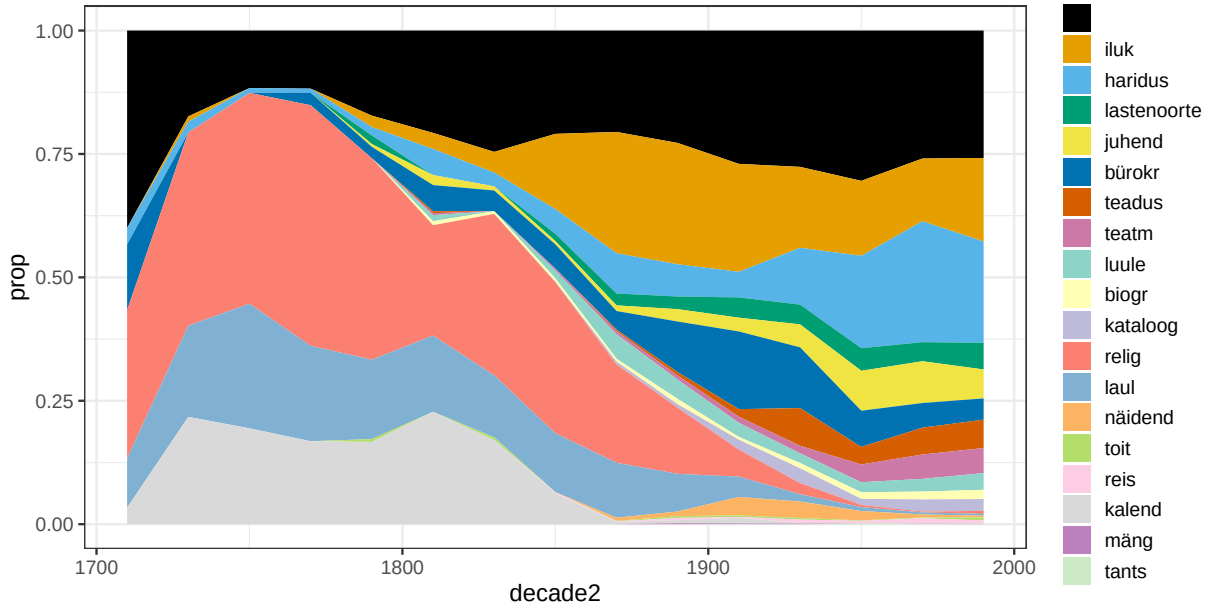


cities-2.bb



publishers-1.bb


1600–1800
1800–2000
−rank1
decade
ristimine. eesti keel luterlus. eesti keel Põhjasõda, eesti keel usuõpetus. luterlus. usuõpetus.
abielu. luterlus. jumalateen usuõpetus. talurahvas pietism. pietism. usuõpetus. luterlus.
matused. grammatika rituaalid. apostlid. usuõpetus. harduskirj harduskirj katekismus katekismus
liturgika. Lõuna−Eest grammatika sõjavägi. usuõpetus. apostlid. talurahvas eesti.
sakramendi eesti. Põhja−Eest deserteeri kirikulaul vaimulikud talurahvaõ kõlbluskas
vasturefor feodaalkoo eesti keel psalmid. luterlus. karistused kombeõpetu
roomakatol feodaalren sõjategevu grammatika katekismus apostlid. talurahvas
piibliluge nahk (mate ajalugu. fraseoloog piibliluge kauplemist talurahvaõ
eesti keel karistused Juudi sõda vanasõnad. eesti keel linnaõigus pearaha.
grammatika sõda. haldusõigu kõnekäänud kaubandus. pietism.
eesti keel usuõpetus. saksa. eesti. eesti. eesti. eesti. eesti. eesti. eesti. eesti.
usuõpetus. kristlik e eesti. saksa. seltsid. seltsid. eesti keel vene. eesti keel eesti keel ameerika.
talurahvaõ eesti keel eesti keel eesti keel tulekindlu saksa. seltsid. eesti keel väliseesti ajalugu. inglise.
luterlus. usuline är usuline är usuõpetus. saksa. eesti keel saksa. väliseesti vene. inglise. ajalugu.
katekismus piibliselt usuõpetus. vene keel. eesti keel vene. inglise. ajalugu. ajalugu. ameerika. eesti keel
eesti. piibliluge kristlik e ajalugu. vene keel. tulekindlu ühingud. kommunistl õpetamise väliseesti kodulugu.
talurahvak luterlik k kogudused. saksa keel välismisjo misjon. ajalugu. kolhoosid. kommunistl matemaatik vene.
kristlik e luterlus. piibliluge aritmeetik misjon. raamatud. matemaatik keskkoolid matemaatik õpetamise kultuuriaj
seadused. katekismus kirikulaul kristlik e vene. välismisjo kutseorgan kirjandus. tööohutus. vene. eluolu.
saksa. eesti. luterlus. matusetoet misjonärid sotsiaalde vene. geograafia füüsika. emakeel. prantsuse.


prop
decade2
iluk
haridus
lastenoorte
juhend
bürokr
teadus
teatm
luule
biogr
kataloog
relig
laul
näidend
toit
reis
kalend
mäng
tants

genres-1.bb