

25 years of Social Work Research

Summary of Data Cleaning Procedures

Overview

This document summarizes data cleaning procedures of the data file processed by the `ebscoBWR` function. It should be noted that that `ebscoBWR` function merged the search results from the first- and second-level searches performed on the three databases (psycInfo, Social Science Abstracts, and Social Work Abstract) hosted by the EbscoPlatform. This function merged all the search results, transformed the data into a structured data file, and removed all duplicate article records. No further transformations or cleaning was performed by the `ebscoBWR` function. Provided below are basic descriptive summaries of the results from the full search before further cleaning.

The main file for analysis is `fullIDF.R`, and a separate file is just for psycInfo `psycInfoDF.R`. The same cleaning procedures need to be performed on both data files prior to analysis.

Initialization of workspace

```
# Clear workspace
rm(list=ls())

# Load required packages
library(ggplot2)
library(dplyr)

# Read data files

load("/Users/beperron/Git/SocialWorkResearch/Data/ebscoFULL.R")
load("/Users/beperron/Git/SocialWorkResearch/Data/proQuest.R")
load("/Users/beperron/Git/SocialWorkResearch/Data/ebscoPsycInfo.R")

#load("/Users/SSW/Documents/GitHub/SocialWork100Years/Data/ebscoFULL.R")
#load("/Users/SSW/Documents/GitHub/SocialWork100Years/Data/proQuest.R")
#load("/Users/SSW/Documents/GitHub/SocialWork100Years/Data/ebscoPsycInfo.R")
```

Combines EbscoHost and ProQuest

This section ... wrangles and eliminates duplicate article titles

```
#Creates variable with highest article ID from ebscohost database
ebscoIDmax<-max(ebscoBWR.df$articleID)

#Recodes article ID in ProQuest database in advance of merger with ebscohost
proQuestBWR.df<-mutate(proQuestBWR.df, articleID=articleID+ebscoIDmax)

#Creates data frames that contain only article ID and titles from respective
#databases
```

```

ebscoTitles<-filter(ebscoBWR.df, attributes=="article")
proTitles<-filter(proQuestBWR.df, attributes=="article")

#Combines titles into single data frame
jointTitles<-rbind(ebscoTitles, proTitles)

#Forces article title text to lower case
jointTitles$record<-tolower(jointTitles$record)

#Strips out white space and punctuation to ensure small differences in data
#entry are detected during duplicate elimination
jointTitles$record<-gsub(" ","", jointTitles$record)

jointTitles$record<-gsub("[[:punct:]]", "", jointTitles$record)

jointTitles<- jointTitles[duplicated(jointTitles$record), ]

#Screen out duplicated records by articleID. The articleID must be used
#because the duplicate title contains other article attributes
jointTitles.dup.ID <- jointTitles$articleID

#Combines EbscoHost and ProQuest Wrangles, and strips out article attributes
#from duplicate entries
combined.DF<-rbind(ebscoBWR.df, proQuestBWR.df)
full.df<-combined.DF[!(combined.DF$articleID %in% jointTitles.dup.ID), ]

rm(list=setdiff(ls(), "full.df"))

```

Data cleaning and pre-processing

Narrow results to past 25 years

Inspection of the data indicated notable delays in indexing for the years 2014 and 2015. Therefore, these years were excluded from the current study. Thus, this study includes journals indexed through the end of 2013.

```

year.25 <- as.character(c(1900:1987, 2014, 2015))
years <- c("2014", "2015")
#ebsco.remove <- full.df[full.df$record %in% years, ]
ebsco.remove <- full.df[full.df$record %in% year.25, ]
ebsco.remove <- ebsco.remove[, "articleID"]
full.df <- full.df[!(full.df$articleID %in% ebsco.remove),]

```

Initial descriptive summary

```

number.of.articles <- length(which(full.df$attributes == "article"))
range.of.years <- range(full.df$record[full.df$attributes == "pubYear"])

```

```
# Number of articles
cat(number.of.articles)
```

```
## 32986
```

```
# Range of years
cat(range.of.years)
```

```
## 1988 2013
```

Exclusion of journals not meeting supplemental search terms

The search for journals involved a two-stage process. It began by extracting all records from journals based on the social work journal list of Hodge and L (APA cite). Journals were excluded if they did not contain one of the major search terms from the supplemental list. The majority of these journals were captured in the supplemental search query because the journal used one of the supplemental search terms in a special issue. The mission and scope of journals were also reviewed when it was unclear whether the journal was specific to social work research (e.g., *The Clinical Supervisor*). The supplemental search terms included:

- social work
- social welfare
- social casework (or social case work)
- social service
- human service
- social development
- social environment

```
exclusions.1 <- c(
  "American Journal of Preventive Medicine",
  "Behavior Modification",
  "Brain and Cognition",
  "Canadian Journal of Community Mental Health",
  "Canadian Journal on Aging",
  "Early Child Development and Care",
  "Early Education and Development",
  "Employee Assistance Quarterly",
  "Evaluation and Program Planning",
  "General Hospital Psychiatry",
  "International Review of Social Research",
  "Journal of Applied Behavioral Science",
  "Journal of Applied Rehabilitation Counseling",
  "Journal of Community & Applied Social Psychology",
  "Journal of Health & Human Services Administration",
  "Journal of Nonverbal Behavior",
  "Journal of Psychosomatic Research",
  "Learning Disabilities Research & Practice",
  "Professional Development in Education",
  "PROFILE Issues in Teachers' Professional Development",
  "Psychotherapy",
  "Systems Research and Behavioral Science",
  "Teaching and Teacher Education",
```

```

    "The Scientific Review of Mental Health Practice",
    "The Clinical Supervisor")

ebsco.remove <- full.df[full.df$record %in% exclusions.1, ]
ebsco.remove <- ebsco.remove[, "articleID"]
full.df <- full.df[!(full.df$articleID %in% ebsco.remove),]

rm(ebsco.remove)

```

Exclusions of social work publications

- Journals that did not have at least one full year of indexed records (except journals that subsequently changed titles). This exclusionary criterion was consistent with the working definition of a social work journal.

```

exclusions.2 <- c(
  "Issues in Social Work Education",
  "Maatskaplike Werk/Social Work",
  "Pediatric Social Work",
  "Critical Social Work"
)

ebsco.remove <- full.df[full.df$record %in% exclusions.2, ]
ebsco.remove <- ebsco.remove[, "articleID"]
full.df <- full.df[!(full.df$articleID %in% ebsco.remove),]

rm(ebsco.remove, exclusions.2)

```

Fix journal titles

Additional pre-processing was conducted on journal titles to resolve spelling errors or other minor title differences in the database. For example, *Social Work With Groups* was exported as a separate journal than *Social Work with Groups* (note the capitalization difference in the word *with*.) We also applied a global substitution of *and* with *&*, as this was a common source for differences.

	record	N
1	Administration in Social Work	616
2	Advances in Social Work	239
3	Affilia	186
4	Arete	360
5	Aretê	1
6	Australian Social Work	245
7	British Journal of Social Work	1154
8	Canadian Social Work	52
9	Canadian Social Work Review	498
10	Child & Adolescent Social Work Journal	728
11	Child & Family Social Work	571
12	Children & Schools	239
13	China Journal of Social Work	98
14	Clinical Social Work Journal	823
15	Computers in Human Services	179

16	Contemporary Rural Social Work	27
17	Ethics & Social Welfare	115
18	European Journal of Social Work	459
19	Families in Society	1612
20	Health & Social Work	993
21	International Journal of Social Welfare	388
22	International Social Work	1038
23	Journal of Analytic Social Work	47
24	Journal of Applied Social Sciences	47
25	Journal of Baccalaureate Social Work	60
26	Journal of Comparative Social Welfare	66
27	Journal of Ethnic & Cultural Diversity in Social Work	257
28	Journal of Evidence-Based Social Work	149
29	Journal of Family Social Work	300
30	Journal of Gay & Lesbian Social Services	518
31	Journal of Gerontological Social Work	1219
32	Journal of HIV/AIDS & Social Services	218
33	Journal of Human Behavior in the Social Environment	785
34	Journal of Independent Social Work	74
35	Journal of Multicultural Social Work	21
36	Journal of Practice Teaching in Health & Social Work	40
37	Journal of Progressive Human Services	98
38	Journal of Psychosocial Oncology	627
39	Journal of Religion & Spirituality in Social Work	199
40	Journal of Social Service Research	615
41	Journal of Social Work	255
42	Journal of Social Work & Human Sexuality	45
43	Journal of Social Work Education	995
44	Journal of Social Work Practice	382
45	Journal of Social Work Practice in the Addictions	355
46	Journal of Social Work Research & Evaluation	99
47	Journal of Social Work Values & Ethics	50
48	Journal of Social Work in Disability & Rehabilitation	153
49	Journal of Social Work in End-of-Life & Palliative Care	152
50	Journal of Sociology & Social Welfare	929
51	Journal of Teaching in Social Work	533
52	Journal of Technology in Human Services	292
53	Journal of the Society for Social Work & Research	65
54	Practice	176
55	Prevention in Human Services	91
56	Professional Development	218
57	Psychoanalytic Social Work	154
58	Qualitative Social Work	290
59	Reflections	834
60	Research on Social Work Practice	1122
61	Revista de Asistentă Socială	220
62	Rural Social Work	39
63	School Social Work Journal	155
64	Smith College Studies in Social Work	496
65	Social Casework	164
66	Social Development	691
67	Social Development Issues	279
68	Social Service Review	728
69	Social Work	1628

70	Social Work & Christianity	193
71	Social Work & Social Sciences Review	220
72	Social Work & Society	220
73	Social Work Education	844
74	Social Work Forum	27
75	Social Work Research	579
76	Social Work Research & Abstracts	67
77	Social Work With Groups	647
78	Social Work in Health Care	1066
79	Social Work in Mental Health	295
80	Social Work in Public Health	192
81	The British Journal of Social Work	289
82	The Journal of Baccalaureate Social Work	258

Title merging

Some social work journals changed titles over their history. For example, *Families in Society* had three titles that preceded the current title: *Social Casework*, *Journal of Social Casework*, and *The Family*. The database that contributed the largest number of articles to the study database indexed the majority of journal titles under the journal's current name, and provided an additional data field for the historical name. However, historical names from Social Science Abstracts and Social Work Abstracts were not merged with the current name. Thus, the last step of the data cleaning procedures was to conduct these merges.

```
full.df$record <- ifelse(full.df$record ==
  "The British Journal of Pyschiatric Social Work" |
  full.df$record ==
  "The British Journal of Social Work", # Old
  "British Journal of Social Work",    # New title
  full.df$record)

full.df$record <- ifelse(full.df$record ==
  "Computers in Human Services",      # Old title
  "Journal of Technology in Human Services", # New title
  full.df$record)

full.df$record <- ifelse(full.df$record ==
  "The Family" | full.df$record ==      # Old title
  "Journal of Social Casework" | full.df$record == # Old title
  "Social Casework",                    # Old title
  "Families in Society",                # New title
  full.df$record)

full.df$record <- ifelse(full.df$record ==
  "Social Work in Education",          # Old title
  "Children & Schools",                # New title
  full.df$record)

full.df$record <- ifelse(full.df$record ==
  "Journal of Independent Social Work" | full.df$record == # New title
  "Journal of Analytic Social Work",                      # Old title
  "Psychoanalytic Social Work",                          # Old title
  full.df$record)
```

```

# Journals were originally merged after suffixes were eliminated (after the
# colon). Thus, Practice: Social Work in Action (new title) was shortened to
# Practice (old title). This section resolves two of these issues to be
# consistent in the comparison with Hodge.

practice.temp <- grep("^Practice$", full.df$record)
full.df$record[practice.temp] <- "Practice: Social Work in Action"

practice.temp <- grep("^Reflections$", full.df$record)
full.df$record[practice.temp] <- "Reflections: Narratives of Professional Helping"

```

Post-processing summary

Revised database

This section provides a summary of the revised database after the initial processing with the `ebscoBWR` function and the subsequent data cleaning. An additional code is included in this summary to indicate whether the journal title overlaps with the list of Hodge (indicated by *H*).

Joining by: "articleID"

	title	first	last	n	H
	Administration in Social Work	1988	2013	617	Y
	Advances in Social Work	2000	2012	239	Y
	Affilia	2007	2013	188	Y
	Arete	1988	2013	360	Y
	Aretê	1997	1997	1	N
	Australian Social Work	2006	2013	245	Y
	British Journal of Social Work	1988	2013	1450	Y
	Canadian Social Work	1999	2012	52	Y
	Canadian Social Work Review	1988	2013	500	Y
	Child & Adolescent Social Work Journal	1988	2013	732	Y
	Child & Family Social Work	1997	2013	573	Y
	Children & Schools	2003	2013	242	Y
	China Journal of Social Work	2009	2013	98	Y
	Clinical Social Work Journal	1988	2013	829	Y
	Contemporary Rural Social Work	2010	2013	27	N
	Ethics & Social Welfare	2009	2013	115	N
	European Journal of Social Work	2000	2013	461	Y
	Families in Society	1988	2013	1792	Y
	Health & Social Work	1988	2013	998	Y
	International Journal of Social Welfare	2000	2013	392	Y
	International Social Work	1988	2013	1049	Y
	Journal of Applied Social Sciences	1994	1999	47	Y
	Journal of Baccalaureate Social Work	2005	2013	60	Y
	Journal of Comparative Social Welfare	2006	2012	66	Y
	Journal of Ethnic & Cultural Diversity in Social Work	2000	2013	261	Y
	Journal of Evidence-Based Social Work	2004	2013	154	Y
	Journal of Family Social Work	1995	2013	305	Y
	Journal of Gay & Lesbian Social Services	1994	2013	518	Y
	Journal of Gerontological Social Work	1988	2013	1225	Y

Journal of HIV/AIDS & Social Services	2005	2013	218	Y
Journal of Human Behavior in the Social Environment	1997	2013	793	Y
Journal of Multicultural Social Work	1999	2000	23	N
Journal of Practice Teaching in Health & Social Work	1999	2005	40	N
Journal of Progressive Human Services	2000	2013	101	Y
Journal of Psychosocial Oncology	1988	2013	627	Y
Journal of Religion & Spirituality in Social Work	2004	2013	199	Y
Journal of Social Service Research	1988	2013	621	Y
Journal of Social Work	2001	2013	257	Y
Journal of Social Work & Human Sexuality	1988	1993	45	N
Journal of Social Work Education	1990	2013	999	Y
Journal of Social Work Practice	1999	2013	388	Y
Journal of Social Work Practice in the Addictions	2001	2013	359	Y
Journal of Social Work Research & Evaluation	2000	2005	99	Y
Journal of Social Work Values & Ethics	2004	2012	50	Y
Journal of Social Work in Disability & Rehabilitation	2001	2013	155	Y
Journal of Social Work in End-of-Life & Palliative Care	2005	2013	154	Y
Journal of Sociology & Social Welfare	1988	2013	934	Y
Journal of Teaching in Social Work	1988	2013	542	Y
Journal of Technology in Human Services	1988	2013	471	Y
Journal of the Society for Social Work & Research	2010	2013	66	N
Practice: Social Work in Action	2005	2013	176	Y
Prevention in Human Services	1988	1993	93	N
Professional Development	1998	2012	218	Y
Psychoanalytic Social Work	1988	2013	278	Y
Qualitative Social Work	2002	2013	292	Y
Reflections: Narratives of Professional Helping	1995	2013	838	Y
Research on Social Work Practice	1991	2013	1132	Y
Revista de Asistentă Socială	2010	2013	220	N
Rural Social Work	2000	2004	39	Y
School Social Work Journal	2002	2013	155	Y
Smith College Studies in Social Work	1988	2013	497	Y
Social Development	1992	2013	699	N
Social Development Issues	1992	2013	281	Y
Social Service Review	1988	2013	736	Y
Social Work	1988	2013	1638	Y
Social Work & Christianity	2005	2013	195	Y
Social Work & Social Sciences Review	1989	2013	224	Y
Social Work & Society	2003	2012	220	Y
Social Work Education	1999	2013	846	Y
Social Work Forum	2004	2009	27	Y
Social Work Research	1994	2013	580	Y
Social Work Research & Abstracts	1988	1993	67	N
Social Work With Groups	1988	2013	651	Y
Social Work in Health Care	1988	2013	1076	Y
Social Work in Mental Health	2002	2013	295	Y
Social Work in Public Health	2007	2013	196	Y
The Journal of Baccalaureate Social Work	1995	2013	258	N

Comparison with Hodge

This section provides further comparison of the journal titles with Hodge by reporting on the *intersection*, and *differences*.

Intersection

This list shows which journal titles are common to both lists.

```
hodge <- readLines("/Users/beperron/Git/SocialWorkResearch/search_specifications/hodge.txt", n = -1)

unique.swHistory <- filter(full.df, attributes == "journal") %>%
  distinct(record)

hodge.swHistory.intersect <- intersect(hodge, unique.swHistory$record)

hodge.swHistory.diff <- setdiff(hodge, unique.swHistory$record)
swHistory.hodge.diff <- setdiff(unique.swHistory$record, hodge)

#hodge.swHistory.union <- union(hodge, unique.swHistory$record)

print(sort(hodge.swHistory.intersect))
```

```
## [1] "Administration in Social Work"
## [2] "Advances in Social Work"
## [3] "Affilia"
## [4] "Arete"
## [5] "Australian Social Work"
## [6] "British Journal of Social Work"
## [7] "Canadian Social Work"
## [8] "Canadian Social Work Review"
## [9] "Child & Adolescent Social Work Journal"
## [10] "Child & Family Social Work"
## [11] "Children & Schools"
## [12] "China Journal of Social Work"
## [13] "Clinical Social Work Journal"
## [14] "European Journal of Social Work"
## [15] "Families in Society"
## [16] "Health & Social Work"
## [17] "International Journal of Social Welfare"
## [18] "International Social Work"
## [19] "Journal of Applied Social Sciences"
## [20] "Journal of Baccalaureate Social Work"
## [21] "Journal of Comparative Social Welfare"
## [22] "Journal of Ethnic & Cultural Diversity in Social Work"
## [23] "Journal of Evidence-Based Social Work"
## [24] "Journal of Family Social Work"
## [25] "Journal of Gay & Lesbian Social Services"
## [26] "Journal of Gerontological Social Work"
## [27] "Journal of HIV/AIDS & Social Services"
## [28] "Journal of Human Behavior in the Social Environment"
## [29] "Journal of Progressive Human Services"
## [30] "Journal of Psychosocial Oncology"
## [31] "Journal of Religion & Spirituality in Social Work"
## [32] "Journal of Social Service Research"
## [33] "Journal of Social Work"
## [34] "Journal of Social Work Education"
## [35] "Journal of Social Work in Disability & Rehabilitation"
```

```
## [36] "Journal of Social Work in End-of-Life & Palliative Care"
## [37] "Journal of Social Work Practice"
## [38] "Journal of Social Work Practice in the Addictions"
## [39] "Journal of Social Work Research & Evaluation"
## [40] "Journal of Social Work Values & Ethics"
## [41] "Journal of Sociology & Social Welfare"
## [42] "Journal of Teaching in Social Work"
## [43] "Journal of Technology in Human Services"
## [44] "Practice: Social Work in Action"
## [45] "Professional Development"
## [46] "Psychoanalytic Social Work"
## [47] "Qualitative Social Work"
## [48] "Reflections: Narratives of Professional Helping"
## [49] "Research on Social Work Practice"
## [50] "Rural Social Work"
## [51] "School Social Work Journal"
## [52] "Smith College Studies in Social Work"
## [53] "Social Development Issues"
## [54] "Social Service Review"
## [55] "Social Work"
## [56] "Social Work & Christianity"
## [57] "Social Work & Social Sciences Review"
## [58] "Social Work & Society"
## [59] "Social Work Education"
## [60] "Social Work Forum"
## [61] "Social Work in Health Care"
## [62] "Social Work in Mental Health"
## [63] "Social Work in Public Health"
## [64] "Social Work Research"
## [65] "Social Work With Groups"
```

Differences

This list shows the articles on Hodge but not the current study.

```
print(sort(hodge.swHistory.diff))
```

```
## [1] "Annual of Social Work"
## [2] "Asia Pacific Journal of Social Work & Development"
## [3] "Black Caucus"
## [4] "Caribbean Journal of Social Work"
## [5] "Critical Social Work"
## [6] "Electronic Journal of Social Work"
## [7] "Indian Journal of Social Work"
## [8] "IUC Journal of Social Work Theory & Practice"
## [9] "Japanese Journal of Social Services"
## [10] "Journal of Changsha Social Work"
## [11] "Journal of Forensic Social Work"
## [12] "Journal of Practice Teaching in Social Work & Health"
## [13] "Journal of Rural Social Work & Social Development"
## [14] "Journal of Social Work in Long-Term Care"
## [15] "Social Work Review"
## [16] "The Hong Kong Journal of Social Work"
```

```
## [17] "The New Social Worker"
## [18] "The Social Worker/Le Travailleur Social"
## [19] "The Spirituality & Social Work Forum"
```

This list shows the articles on the current study but not on Hodge.

```
print(sort(swHistory.hodge.diff))

## [1] "Aretê"
## [2] "Contemporary Rural Social Work"
## [3] "Ethics & Social Welfare"
## [4] "Journal of Multicultural Social Work"
## [5] "Journal of Practice Teaching in Health & Social Work"
## [6] "Journal of Social Work & Human Sexuality"
## [7] "Journal of the Society for Social Work & Research"
## [8] "Prevention in Human Services"
## [9] "Revista de Asistentă Socială"
## [10] "Social Development"
## [11] "Social Work Research & Abstracts"
## [12] "The Journal of Baccalaureate Social Work"
```

Descriptive summary of records

```
number.of.articles <- length(which(full.df$attributes == "article"))
range.of.years <- range(full.df$record[full.df$attributes == "pubYear"])
```

```
# Number of articles
cat(number.of.articles)
```

```
## 32559
```

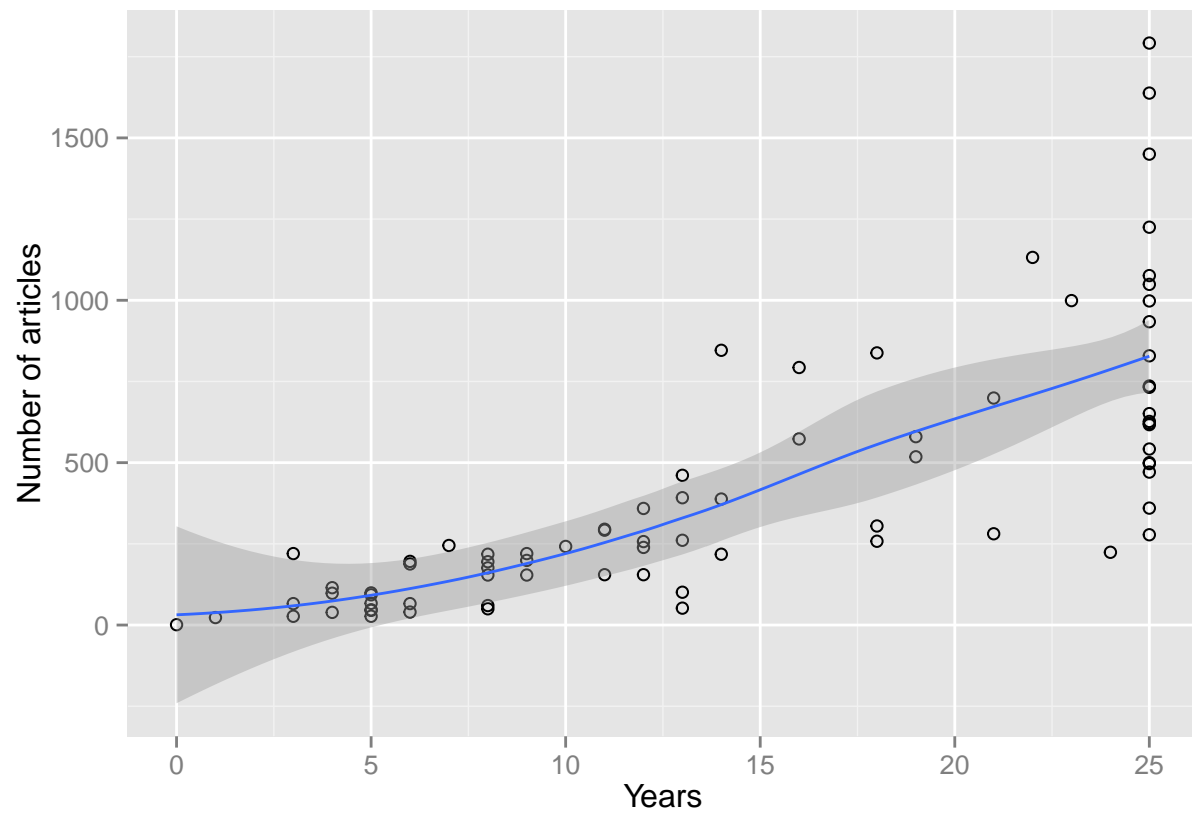
```
# Range of years
cat(range.of.years)
```

```
## 1988 2013
```

```
cor.data <- select(final.list, first, last, n) %>%
  mutate(no.years = as.numeric(last) - as.numeric(first))

ggplot(cor.data, aes(x = no.years, y = n)) +
  geom_point(shape = 1) +
  geom_smooth() +
  xlab("Years") +
  ylab("Number of articles")
```

```
## geom_smooth: method="auto" and size of largest group is <1000, so using loess. Use 'method = x' to c
```



```
cor(cor.data$no.years, cor.data$n)
```

```
## [1] 0.7587
```