

# Supplementary Material

## Contents

<b>1</b>	<b>Interactive visualizations</b>	<b>1</b>
<b>2</b>	<b>Intercoder reliability</b>	<b>1</b>
<b>3</b>	<b>Number of publications for each journal</b>	<b>1</b>
<b>4</b>	<b>Women proportion among authors of different racial groups</b>	<b>2</b>
<b>5</b>	<b>Citation Analysis</b>	<b>2</b>
5.1	Read data . . . . .	2
5.2	Deal with univariate outliers . . . . .	2
5.3	Build model . . . . .	4
5.4	Diagnostics . . . . .	5
5.5	Run Model . . . . .	8

## 1 Interactive visualizations

Interactive visualizations showing the distribution of race and gender in different countries/regions are available at <https://icaviz.netlify.app/>.

## 2 Intercoder reliability

In the Methods section of the manuscript, we mentioned that

Once coders had completed training, we randomly selected 100 authors among all 11,304 authors for inter-coder reliability and achieved a Krippendorff alpha of 0.91 for race and 0.94 for gender.

Note that when we calculated the gender intercoder reliability here, the sample included 160 subsequently deleted authors whose affiliation data was missing.

## 3 Number of publications for each journal

See Figure 1.

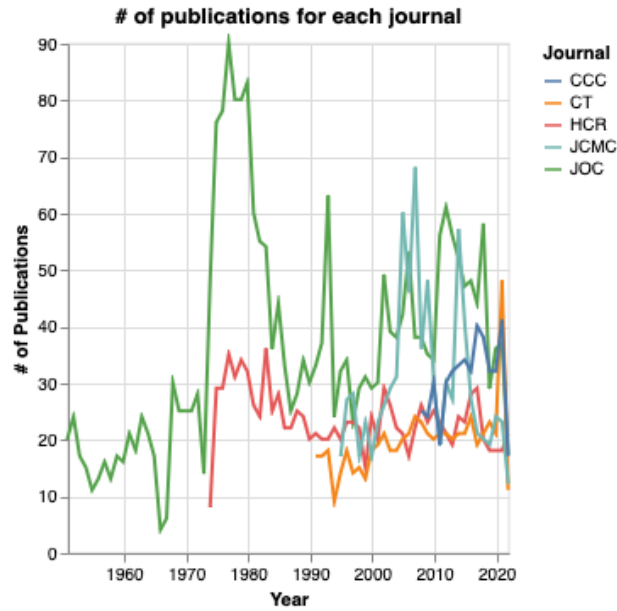


Figure 1: Number of publications for each journal by year

## 4 Women proportion among authors of different racial groups

See Figure 2 and Figure 3.

## 5 Citation Analysis

### 5.1 Read data

The first step is to read data:

```
df <- read.csv("../data/processed/papers_to_study_expanded.csv")
df$Year.Distance.from.2022 = abs(df$year - 2022)

# do the log10 transformation on citation counts
df$gscholar_citation_log10 <- log10(df$gscholar_citation + 0.1)
var_cols <- c(1, 16:17, 20:33, 35:36)
data <- df[, var_cols]
```

### 5.2 Deal with univariate outliers

```
hist(data$gscholar_citation_log10)
```

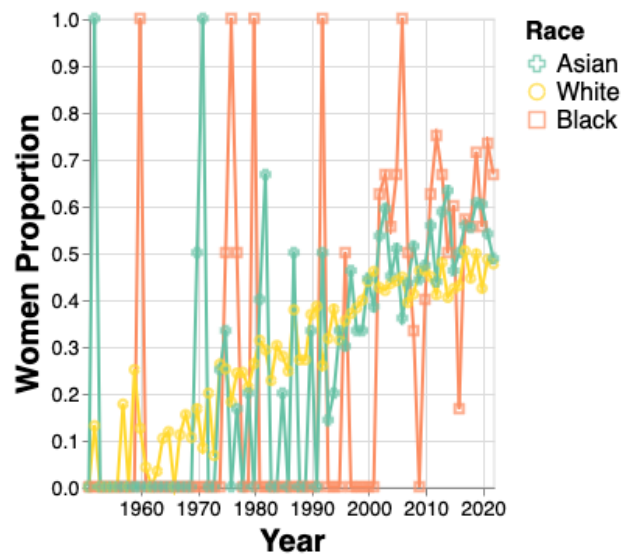


Figure 2: Women proportion

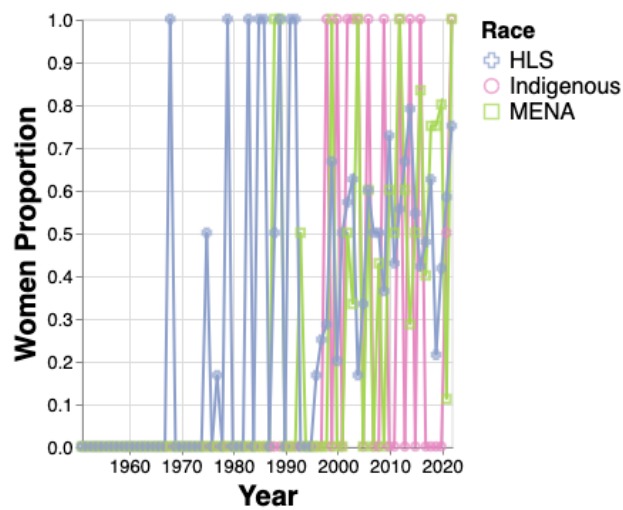
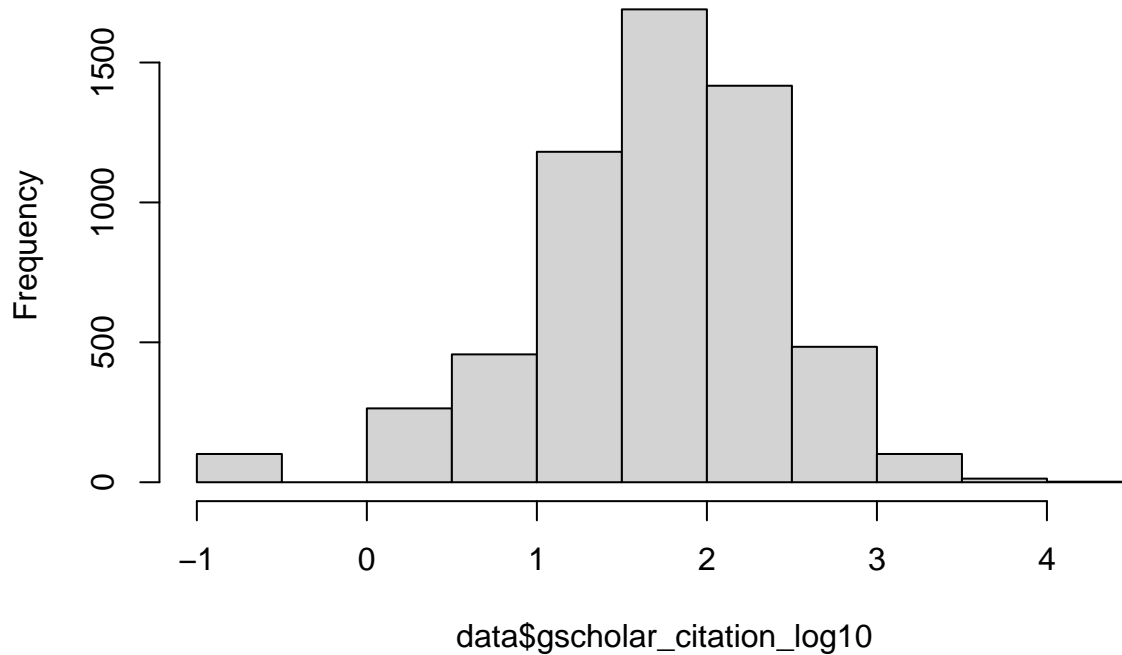


Figure 3: Women proportion

## Histogram of data\$gscholar\_citation\_log10



```
data$gscholar_citation_log10_stdized <- scale(  
  data$gscholar_citation_log10,  
  center = T,  
  scale = T  
)  
data$outlier0 <- ifelse(data$gscholar_citation_log10_stdized > 3.3 |  
  data$gscholar_citation_log10_stdized < -3.3,  
  1, 0  
)  
table(data$outlier0)
```

```
##  
##    0    1  
## 5607  103
```

As can be seen, there are 103 outliers.

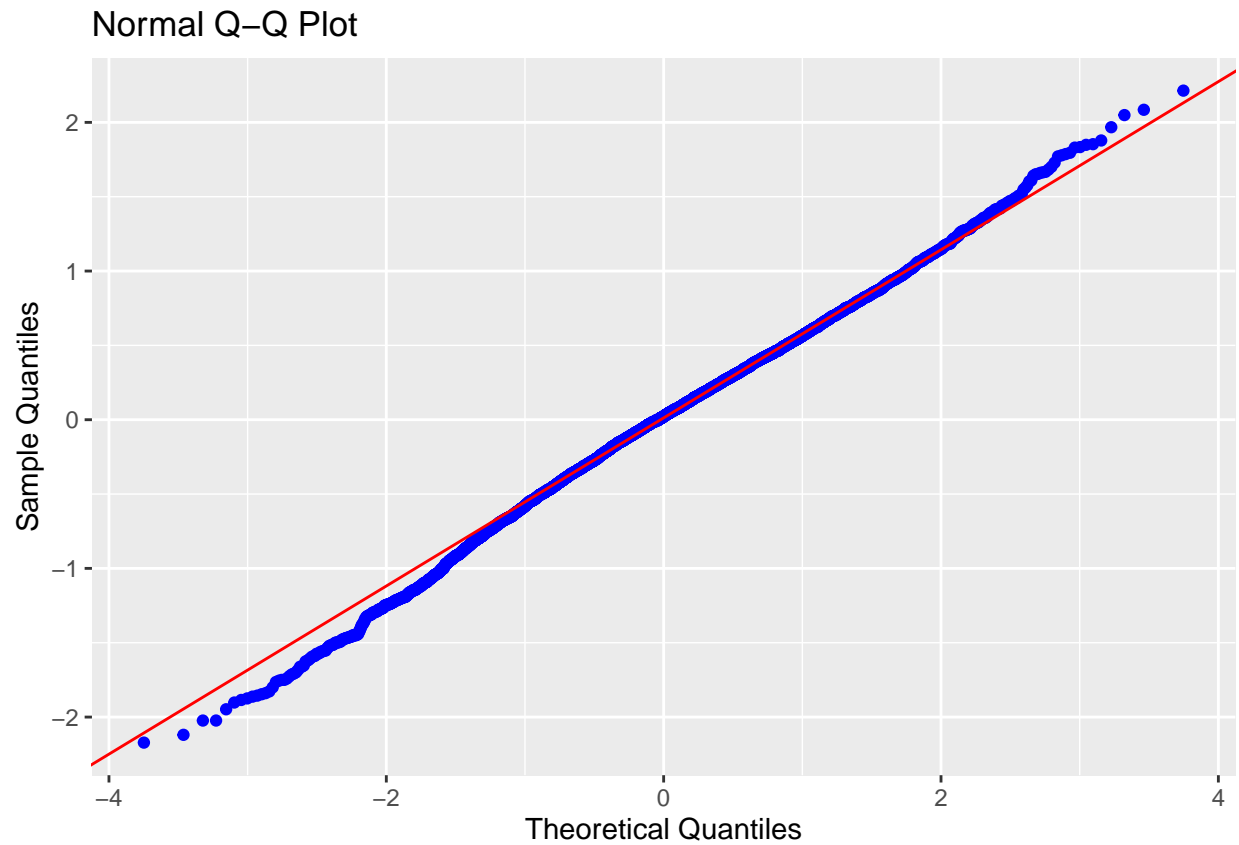
### 5.3 Build model

```
# '10.1111/j.1460-2466.1977.tb02133.x', '10.1111/j.1460-2466.1952.tb00171.x' These two papers have no g  
datanew <- subset(data, outlier0 == 0)  
datanew$outlier0 <- NULL  
datanew$gscholar_citation_log10_stdized <- NULL
```

```
model <- lm(gscholar_citation_log10~., datanew)
```

## 5.4 Dianostics

```
## Normality  
ols_plot_resid_qq(model)
```

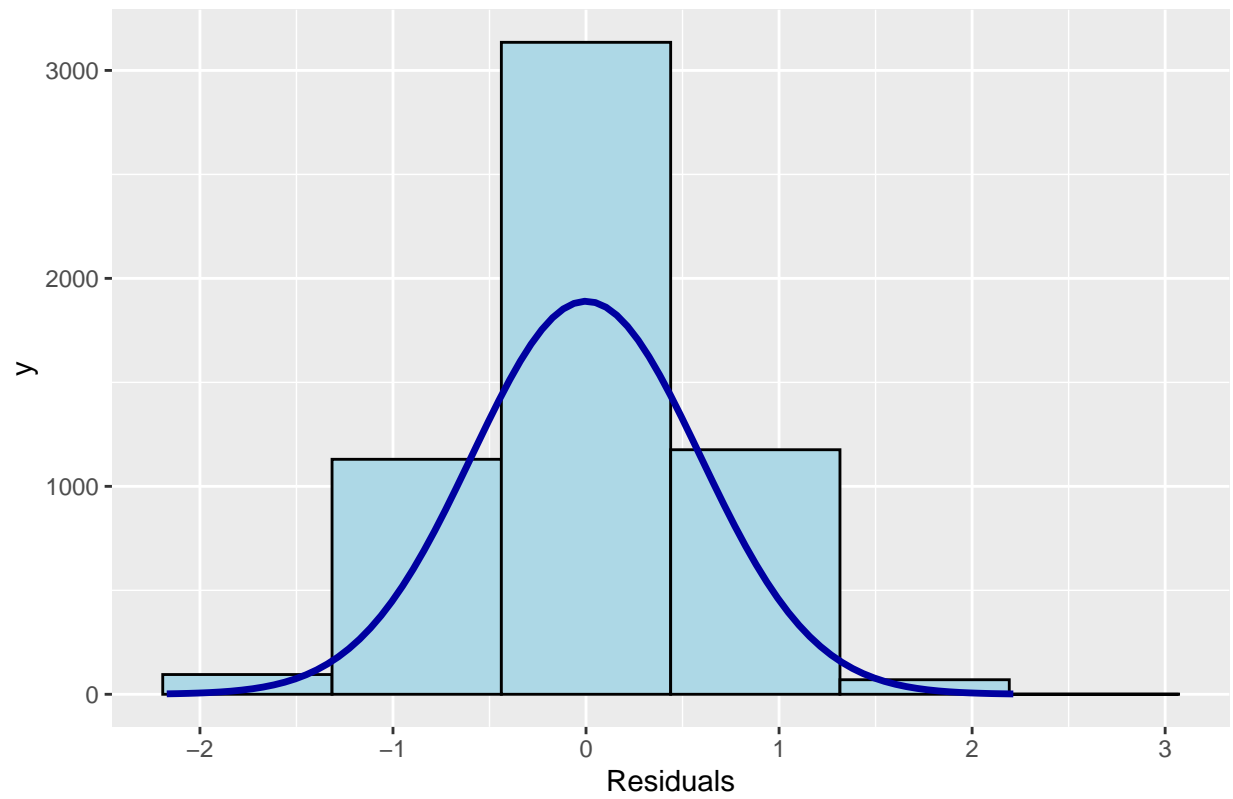


```
#Correlation between observed residuals and expected residuals under normality.  
ols_test_correlation(model)
```

```
## [1] 0.9986378
```

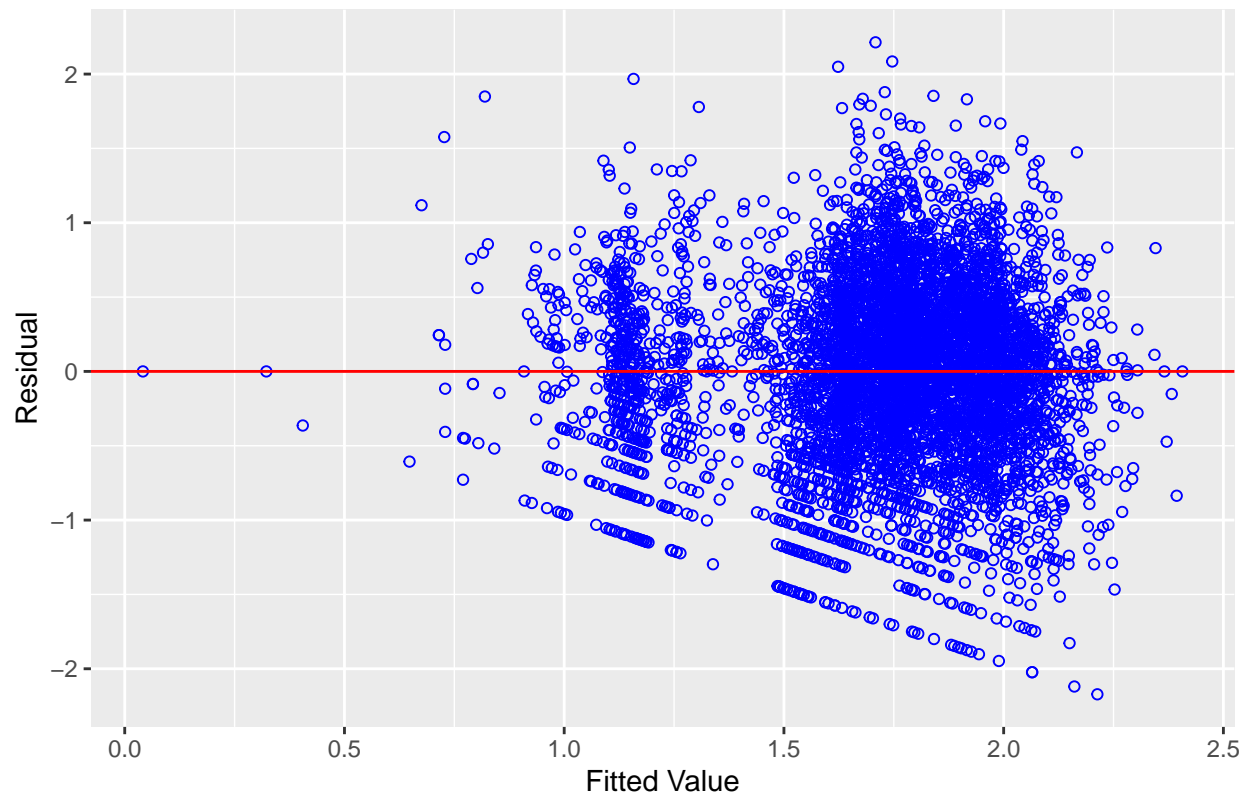
```
ols_plot_resid_hist(model)
```

Residual Histogram



```
## linearity & homoscedasticity  
ols_plot_resid_fit(model)
```

## Residual vs Fitted Values



```
## collinearity diagnostics
# ols_coll_diag(model)
vif(model)
```

	GVIF	Df	$GVIF^{1/(2*Df)}$
## journal	1.803106	4	1.076472
## cross_country	6.619440	1	2.572827
## cross_type	1.180897	1	1.086691
## num_race	12.682847	1	3.561298
## num_country	4.963491	1	2.227889
## cross_gender_and_race	4.109228	1	2.027123
## cross_gender_and_country	4.438181	1	2.106699
## cross_country_and_race	5.693560	1	2.386118
## cross_gender_race_and_country	6.218367	1	2.493665
## numberOfAuthors	2.314327	1	1.521291
## first_author_gender	11.030926	2	1.822439
## first_author_race	88.563799	5	1.565762
## first_author_country	27.159605	64	1.026130
## first_author_afftype	1.203357	1	1.096976
## with_us_authors	8.687016	1	2.947374
## cross_race_details	941.384763	6	1.769351
## cross_gender_details	24.331605	3	1.702270
## Year.Distance.from.2022	1.849162	1	1.359839

## 5.5 Run Model

```
summary(model)
```

```
##
## Call:
## lm(formula = gscholar_citation_log10 ~ ., data = datanew)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.17181 -0.36909  0.02204  0.39369  2.21343
##
## Coefficients:
##                                     Estimate Std. Error t value
## (Intercept)                        1.7575010  0.4326112   4.063
## journalCommunication, Culture and Critique -0.6191977  0.0385673 -16.055
## journalHuman Communication Research      0.1922003  0.0315584   6.090
## journalJournal of Communication         0.0748179  0.0290478   2.576
## journalJournal of Computer-Mediated Communication 0.2535682  0.0328364   7.722
## cross_countryTrue                     0.1661404  0.0669349   2.482
## cross_typeTrue                        0.0593013  0.0413427   1.434
## num_race                             -0.1527943  0.0682835  -2.238
## num_country                          0.0645221  0.0419850   1.537
## cross_gender_and_raceTrue             -0.1094405  0.0543681  -2.013
## cross_gender_and_countryTrue          -0.2550186  0.0692284  -3.684
## cross_country_and_raceTrue            -0.0986971  0.0956964  -1.031
## cross_gender_race_and_countryTrue      0.0209796  0.1212366   0.173
## numberOfAuthors                      0.0019381  0.0085842   0.226
## first_author_genderM                   0.0483471  0.0300441   1.609
## first_author_genderN                  -0.0823867  0.3489748  -0.236
## first_author_raceBlack                 -0.1022738  0.1313243  -0.779
## first_author_raceHispanic              0.0432094  0.0810781   0.533
## first_author_raceIndigenous            -0.2074777  0.3017309  -0.688
## first_author_raceMiddle Eastern        -0.1638593  0.1279014  -1.281
## first_author_raceWhite                  0.0088196  0.0465751   0.189
## first_author_countryAR                 -0.1670665  0.5992099  -0.279
## first_author_countryAT                  0.0541806  0.4419681   0.123
## first_author_countryAU                  0.0344297  0.4299884   0.080
## first_author_countryBE                  0.0782253  0.4385574   0.178
## first_author_countryBR                 -0.0897198  0.4746324  -0.189
## first_author_countryBS                 -0.3078873  0.7317539  -0.421
## first_author_countryBY                 -0.1763622  0.7317590  -0.241
## first_author_countryCA                  0.1519215  0.4268059   0.356
## first_author_countryCH                  0.2272420  0.4404915   0.516
## first_author_countryCL                  0.2191174  0.4574807   0.479
## first_author_countryCN                 -0.0089431  0.4311088  -0.021
## first_author_countryCO                  0.1698293  0.5186720   0.327
## first_author_countryCR                  0.7622836  0.7371177   1.034
## first_author_countryCZ                  0.9509944  0.4895678   1.943
## first_author_countryDE                  0.2570244  0.4270822   0.602
## first_author_countryDK                  0.2083847  0.4440428   0.469
## first_author_countryEE                 -0.0331776  0.7322282  -0.045
## first_author_countryEG                  0.3530255  0.5491112   0.643
```



## first_author_countryES	0.2036157	0.4436131	0.459
## first_author_countryFI	-0.1162013	0.4437614	-0.262
## first_author_countryFR	-0.2781594	0.4440562	-0.626
## first_author_countryGB	0.1902383	0.4256283	0.447
## first_author_countryGH	0.1365045	0.5986721	0.228
## first_author_countryGR	0.0537018	0.5459413	0.098
## first_author_countryHK	-0.5647224	0.5181047	-1.090
## first_author_countryHR	0.2661562	0.7328335	0.363
## first_author_countryHU	-0.0685841	0.5192323	-0.132
## first_author_countryIE	-0.1709405	0.5460988	-0.313
## first_author_countryIL	0.0217967	0.4265232	0.051
## first_author_countryIN	-0.0969677	0.4737428	-0.205
## first_author_countryIR	0.6667213	0.7336795	0.909
## first_author_countryIT	0.0239192	0.4436468	0.054
## first_author_countryJP	0.0630023	0.4377805	0.144
## first_author_countryKE	0.7413407	0.7367288	1.006
## first_author_countryKH	0.3940539	0.7317559	0.539
## first_author_countryKR	0.2380853	0.4348270	0.548
## first_author_countryLB	0.3573528	0.4728221	0.756
## first_author_countryLK	-0.0849982	0.7325615	-0.116
## first_author_countryMA	-0.2157315	0.7327270	-0.294
## first_author_countryMO	0.0278520	0.7322781	0.038
## first_author_countryMX	0.1267372	0.5036210	0.252
## first_author_countryMZ	-0.6006328	0.7327921	-0.820
## first_author_countryNG	0.2897573	0.7317651	0.396
## first_author_countryNL	0.1637743	0.4264875	0.384
## first_author_countryNO	0.1862891	0.4545124	0.410
## first_author_countryNZ	0.2226745	0.4505984	0.494
## first_author_countryPH	0.3074057	0.4284483	0.717
## first_author_countryPL	0.1515184	0.4884319	0.310
## first_author_countryPT	-0.2684248	0.6030294	-0.445
## first_author_countryQA	0.3412421	0.4905214	0.696
## first_author_countryRU	-0.3264408	0.5474785	-0.596
## first_author_countrySE	0.1308407	0.4381343	0.299
## first_author_countrySG	0.0233332	0.4335970	0.054
## first_author_countrySI	-0.1105005	0.4882286	-0.226
## first_author_countrySK	0.1043069	0.7357632	0.142
## first_author_countryTH	0.3389994	0.6016546	0.563
## first_author_countryTJ	-0.7856905	0.4729496	-1.661
## first_author_countryTR	0.1684039	0.5198392	0.324
## first_author_countryTW	-0.1226529	0.4536421	-0.270
## first_author_countryUG	-1.3624534	0.7375944	-1.847
## first_author_countryUK	0.0028515	0.4472601	0.006
## first_author_countryUS	0.1661144	0.4268645	0.389
## first_author_countryVN	1.0079261	0.7327678	1.376
## first_author_countryZA	-0.3171598	0.5463729	-0.580
## first_author_afftypeNon Education	-0.3818704	0.0389782	-9.797
## with_us_authorsTrue	0.0362180	0.0576519	0.628
## cross_race_detailsBlack only	-0.0436183	0.1622478	-0.269
## cross_race_detailscross race	0.2000852	0.0906408	2.207
## cross_race_detailsHispanic only	-0.1002020	0.1122239	-0.893
## cross_race_detailsIndigenous only	-0.2969760	0.5198808	-0.571
## cross_race_detailsMiddle Eastern only	0.2420354	0.1711496	1.414
## cross_race_detailsWhite only	0.0250952	0.0582313	0.431

## cross_gender_detailsF only	-0.0963608	0.0339207	-2.841
## cross_gender_detailsM only	-0.1588912	0.0298945	-5.315
## cross_gender_detailsN only	0.0839789	0.4176286	0.201
## Year.Distance.from.2022	-0.0054423	0.0006475	-8.405
##	Pr(> t )		
## (Intercept)	4.92e-05	***	
## journalCommunication, Culture and Critique	< 2e-16	***	
## journalHuman Communication Research	1.20e-09	***	
## journalJournal of Communication	0.010030	*	
## journalJournal of Computer-Mediated Communication	1.35e-14	***	
## cross_countryTrue	0.013090	*	
## cross_typeTrue	0.151520		
## num_race	0.025284	*	
## num_country	0.124402		
## cross_gender_and_raceTrue	0.044168	*	
## cross_gender_and_countryTrue	0.000232	***	
## cross_country_and_raceTrue	0.302419		
## cross_gender_race_and_countryTrue	0.862621		
## numberOfAuthors	0.821383		
## first_author_genderM	0.107630		
## first_author_genderN	0.813378		
## first_author_raceBlack	0.436138		
## first_author_raceHispanic	0.594100		
## first_author_raceIndigenous	0.491718		
## first_author_raceMiddle Eastern	0.200199		
## first_author_raceWhite	0.849815		
## first_author_countryAR	0.780400		
## first_author_countryAT	0.902437		
## first_author_countryAU	0.936183		
## first_author_countryBE	0.858439		
## first_author_countryBR	0.850076		
## first_author_countryBS	0.673952		
## first_author_countryBY	0.809555		
## first_author_countryCA	0.721892		
## first_author_countryCH	0.605957		
## first_author_countryCL	0.631982		
## first_author_countryCN	0.983450		
## first_author_countryCO	0.743354		
## first_author_countryCR	0.301116		
## first_author_countryCZ	0.052125		
## first_author_countryDE	0.547322		
## first_author_countryDK	0.638881		
## first_author_countryEE	0.963861		
## first_author_countryEG	0.520314		
## first_author_countryES	0.646257		
## first_author_countryFI	0.793443		
## first_author_countryFR	0.531075		
## first_author_countryGB	0.654922		
## first_author_countryGH	0.819645		
## first_author_countryGR	0.921646		
## first_author_countryHK	0.275771		
## first_author_countryHR	0.716479		
## first_author_countryHU	0.894920		
## first_author_countryIE	0.754276		

```

## first_author_countryIL 0.959245
## first_author_countryIN 0.837826
## first_author_countryIR 0.363529
## first_author_countryIT 0.957005
## first_author_countryJP 0.885574
## first_author_countryKE 0.314335
## first_author_countryKH 0.590250
## first_author_countryKR 0.584030
## first_author_countryLB 0.449809
## first_author_countryLK 0.907634
## first_author_countryMA 0.768446
## first_author_countryMO 0.969661
## first_author_countryMX 0.801319
## first_author_countryMZ 0.412451
## first_author_countryNG 0.692142
## first_author_countryNL 0.700988
## first_author_countryNO 0.681920
## first_author_countryNZ 0.621202
## first_author_countryPH 0.473105
## first_author_countryPL 0.756410
## first_author_countryPT 0.656245
## first_author_countryQA 0.486663
## first_author_countryRU 0.551025
## first_author_countrySE 0.765232
## first_author_countrySG 0.957086
## first_author_countrySI 0.820954
## first_author_countrySK 0.887269
## first_author_countryTH 0.573155
## first_author_countryTJ 0.096719 .
## first_author_countryTR 0.745985
## first_author_countryTW 0.786883
## first_author_countryUG 0.064778 .
## first_author_countryUK 0.994913
## first_author_countryUS 0.697180
## first_author_countryVN 0.169031
## first_author_countryZA 0.561613
## first_author_afftypeNon Education < 2e-16 ***
## with_us_authorsTrue 0.529887
## cross_race_detailsBlack only 0.788065
## cross_race_detailscross race 0.027324 *
## cross_race_detailsHispanic only 0.371963
## cross_race_detailsIndigenous only 0.567861
## cross_race_detailsMiddle Eastern only 0.157367
## cross_race_detailsWhite only 0.666517
## cross_gender_detailsF only 0.004517 **
## cross_gender_detailsM only 1.11e-07 ***
## cross_gender_detailsN only 0.840639
## Year.Distance.from.2022 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5969 on 5510 degrees of freedom
## Multiple R-squared:  0.1726, Adjusted R-squared:  0.1582
## F-statistic: 11.97 on 96 and 5510 DF, p-value: < 2.2e-16

```