

**School of Engineering  
Department of Financial and Management Engineering**



**Bibliometric Analysis of Journal of American Mathematical Society**

**Charalambidis Stylianos  
2312014147**

**Dissertation Supervisors: Dr. Drivaliaris Dimosthenis  
Dr. Vassilliou Evangelos**

**Dissertation Committee: Dr. Drivaliaris Dimosthenis  
Dr. Constantelou Anastasia  
Dr. Vassilliou Evangelos**

**February 2021**

## **Acknowledgements**

At the beginning of my diploma thesis, I would like to deeply thank the supervising professor Dr. Drivaliaris Dimosthenis for his guidance and encouragement throughout the work. In addition, I would like also to thank the supervisor Dr. Vassiliou Evangelos for his help and his correct instructions to meet the needs of this diploma thesis. I would like to thank them, my classmates, my friends, and my family for their loyalty to me and their patience throughout my years of study.

*Έχω διαβάσει και κατανοήσει τους κανόνες για τη λογοκλοπή και τον τρόπο σωστής αναφοράς των πηγών που περιέχονται στον Οδηγό συγγραφής διπλωματικών εργασιών του ΤΜΟΔ. Δηλώνω ότι, από όσα γνωρίζω, το περιεχόμενο της παρούσας διπλωματικής εργασίας είναι προϊόν δικής μου δουλειάς και υπάρχουν αναφορές σε όλες τις πηγές που χρησιμοποίησα.*

## **Abstract**

The following diploma thesis presents information for the Journal Mathematical Society for the period 1988-2017. We used the Scopus and the Web of Science database to gather the necessary information for our analysis. Then we searched for missing and duplicates values in the information columns. Then we merged the information of the databases with some common columns of information. Also, after the merge, the most important columns were compared between the two sources as well as for missing and duplicate values. Furthermore, bibliometric analysis of the authors of the magazine was done. The proof of Lotka's law for the productivity of the authors. In the end, the analysis of the Volumes was done in terms of the number of articles but also by Volume and Author count, as well as the pages of the articles as well as by Volume and Author count.

## Περίληψη

Η ακόλουθη διπλωματική εργασία παρουσιάζει πληροφορίες για το περιοδικό Mathematical Society για την περίοδο 1988-2017. Χρησιμοποιήσαμε τη βάση δεδομένων Scopus και Web of Science για τη συλλογή των απαραίτητων πληροφοριών για την ανάλυσή μας. Στη συνέχεια, αναζητήσαμε ελλιπείς πληροφορίες και διπλογραφίες στις στήλες πληροφοριών. Στη συνέχεια συγχωνεύσαμε τις πληροφορίες των βάσεων δεδομένων με τα κριτήρια ορισμένων βασικών στηλών πληροφοριών. Επίσης, μετά την ένωση, οι σημαντικές στήλες συγκρίθηκαν μεταξύ των δύο πηγών καθώς και για ελλιπείς πληροφορίες και διπλογραφίες. Πραγματοποιήθηκε βιβλιομετρική ανάλυση των συγγραφέων του περιοδικού. Η απόδειξη του νόμου της Lotka για την παραγωγικότητα των συγγραφέων. Στο τέλος, η ανάλυση των τόμων έγινε από την άποψη του αριθμού των άρθρων αλλά και από τον αριθμό τόμων και συγγραφέων, καθώς και από τις σελίδες των άρθρων καθώς και από τον όγκο και τον αριθμό συγγραφέων.

## Περιεχόμενα

Acknowledgements .....	2
Abstract.....	4
Περίληψη .....	5
1 Introduction.....	8
1.1. Bibliometrics .....	8
1.2. Data collection from the "Journal of the American Mathematical Society" .....	10
2 General informations about our data.....	12
2.1. The export of information to CSV excel .....	12
2.2. Journal of the American Mathematical Society data .....	12
2.3. Finding NA values in Scopus dataset .....	18
2.4. Finding NA values in Web of Science dataset .....	21
2.5. Unique columns in our dataset .....	25
2.6. Identifying Double elements.....	28
2.7. Identification of unique values in specific columns. <b>Error! Bookmark not defined.</b>	
2.8. Number of papers in the dataset .....	30
3 The merge of the files.....	37
3.1. The merge of the information of our dataset.....	37
3.2. The differences in the merged file .....	39
4 Authors characteristics .....	44
4.1. Authors productivity and Lotka's Law .....	44
4.2. Authors characteristics .....	50
5 Journal's Productivity.....	109
6 References.....	115

## List of tables

Table 2-1 Exact Number of Na per columns of Scopus .....	20
Table 2-2Columns with only Na values .....	20
Table 2-3Exact Number of Na per columns of Scopus .....	23
Table 2-4Columns with only Na values .....	24
Table 2-5Columns that only exists in Scopus.....	25
Table 2-6Columns that only exists in Web of Science .....	26
Table 2-7The number of papers in both datasets compared to papers of the Journal .	35
Table 2-8Papers that only exists in Web of Science .....	35
Table 2-9Papers that only exists in Scopus .....	36
Table 3-1The cleaning of the data and the merge file for Scopus .....	38
Table 3-2The cleaning of the data and the merge file for Web of Science .....	39
Table 3-3The differences and the Na of each column in the merged file .....	43
Table 3-4Papers from Scopus in time period 1988-1994 .....	43
Table 4-1Authors productivity .....	45
Table 4-2Lotka's Law .....	48
Table 4-3Authors cooperation .....	50
Table 4-4The number of papers the authors have written and their fractional count	106
Table 5-1The number of papers per Volume .....	110
Table 5-2The number of papers per 5-Volume.....	111
Table 5-3The number of pages per Volume .....	113
Table 5-4The number of pages per 5-Volume .....	114
Table 5-5Mean and median of the number of pages with the volume .....	114

## List of Figures

Figure 1.1Document Search in Scopus .....	10
Figure 1.2 Document Results and Export.....	11
Figure 1.3Document Search in Web of Science .....	11
Figure 1.4 Document Results and Export.....	12
Figure 4.1 Lotka's Law distiribution .....	45
Figure 4.2Top Authors by the papers they have written .....	107
Figure 4.3 Top Authors by their fractional count .....	108
Figure 5.1 Boxplots per 5-Volumes for the Number of papers .....	115

## **1 Introduction**

The data for this analysis was gathered from a journal published by the American Mathematical Society and the full name is Journal of the American Mathematical Society. The information was extracted from this journal that was published from 1988 to 2017. The data was available at the Scopus bibliometric database and was available at the Web of Science. Firstly, we gathered data from the "Journal of the American Mathematical Society" from the database Scopus and then by the other database Web of Science.

With the analysis we did first we see qualitative and then quantitative data of the magazine we analyzed. In fact, for the best analysis and validity of our data, we used two sources of information. The purpose of this study is to analyze the journal in the information it gives us and to have the best possible quality of data so that our analysis is more valid. This process is governed by many stages such as defining our data, finding, and removing blank information as well as duplicate data. Then the union of our information from the two sources with different columns of information. As a result, we have a common file with information from two sources and as much as possible the best quality of information from the databases. Then we looked for differences in columns of information from the two sources and several interesting things were discovered. Finally, parametric statistics were analyzed for some information regarding the productivity and characteristics of the authors as well as for the quality and productivity of the documents.

### **1.1. Bibliometrics**

First, we have scientometrics which deals with the analysis of scientific publications. Nalimov and Mulchenko coined the term as the application of quantitative methods engaged in the analysis of science and is considered as an information process. Infometrics is an interdisciplinary approach to scientific science and efficient information processing. Cybermetrics is a branch of bibliometrics that uses mathematical and statistical techniques to quantify web sites and online resources.

Many research fields use bibliometric methods to investigate the impact in their field, the impact of a set of research, and the impact of a particular document or to identify



impacts in a particular field of research. Bibliography also has a wide range of other applications, such as descriptive linguistics, the development of lexical synonyms, and the evaluation of the reader's use.

The term bibliometry was first used by Paul Otlet in 1934 [1] and was defined as "the measurement of all aspects related to the publication and reading of documents" [2]. When first used in English by Alan Pritchard in a paper published in 1969 entitled "statistical bibliometry or bibliography?" [3]. It defines the term "application of mathematical and statistical methods in books and other media". Price in 1963 published another important work *little science big science*, which represented the first systematic approach to the structure of modern science.

Also, Price formulated the exponential law of growth as an important contribution and thus the law of price was established. The two individuals of this law have to do first with the validity of exponential development, to remain stable for long periods.

The second has to do with its rapid development. The conclusion is that science is growing rapidly and that the productions of scientists are also multiplying. So another important conclusion that the authors of the research will be modern if they develop in a multiplicative way. According to Price (1963), 87.5% of all scientists who have ever lived are still modern something which was believed to be this number in any given this value is known as the synchronous factor (contemporary coefficient). But Price's claims must be treated with caution because the probability that you will analyze must develop according to the exponential law. Where according to this law the growing trend does not stop and consequently science never stops to grow. With the great interest that began to exist, new research publications began to appear. The first periodical in the field of bibliometry was a scientometrics magazine, founded by Tibor Braun (1978), review and book reviews.

Glanzel (2003) used Price's research as the basis of modern technical evaluation. He then divided three sub-areas for modern bibliometrics. First, we have the methodology research. These studies mainly have to do with the methodology used. The transfer of bibliometric research still shows us the evolution of bibliometric indices. Then we have the scientific fields (scientific disciplines) these studies can be carried out from any field. The aim is to study a specific field using bibliometric indices, measurements for finally, we have the science-policy (science policy) a very important area in conducting

research. In this case, we manage the bibliometric studies to assess productivity levels. This research is done by policymakers to decide how to share the available resources.

## 1.2.Data collection from the "Journal of the American Mathematical Society"

With the help of two databases, Scopus bibliometric database and the Web of Science database provided the documents of the "Journal of the Mathematical Society" without charge. Therefore, the Journal's data that is used can simply be found through the database's "Document Search" at Scopus and "Publication Name" at Web of Science.

*Journal of the American Mathematical Society* 1988, American Mathematical Society, viewed 20 May 2020,

Scopus link:

<https://www.scopus.com/search/form.uri?display=basic&zone=header&origin=>

Web of Science link:

<http://login.webofknowledge.com/>

*Journal of the American Mathematical Society* 1988, American Mathematical Society, viewed 20 June 2020.

Moreover, information from other journals and documents, that were published from 1988 until 2017 from the data base of Scopus and from 1995 until 2017 from the data base of Web of Science have used data from the "Journal of the Mathematical Society".

The image shows a screenshot of the Scopus 'Document search' interface. At the top, there are radio buttons for 'Documents' (selected), 'Authors', and 'Affiliations', and a link for 'Advanced' search. A 'Search tips' link is also present. The search bar contains the text 'Journal American Mathematical Society' and 'Source title'. Below the search bar, there is a 'Limit' dropdown menu. Underneath, there is a 'Date range (inclusive)' section with 'Published' selected, and a date range from 1988 to 2017. There is also an option for 'Added to Scopus in the last 7 days'.

Figure 1.1 Document Search in Scopus

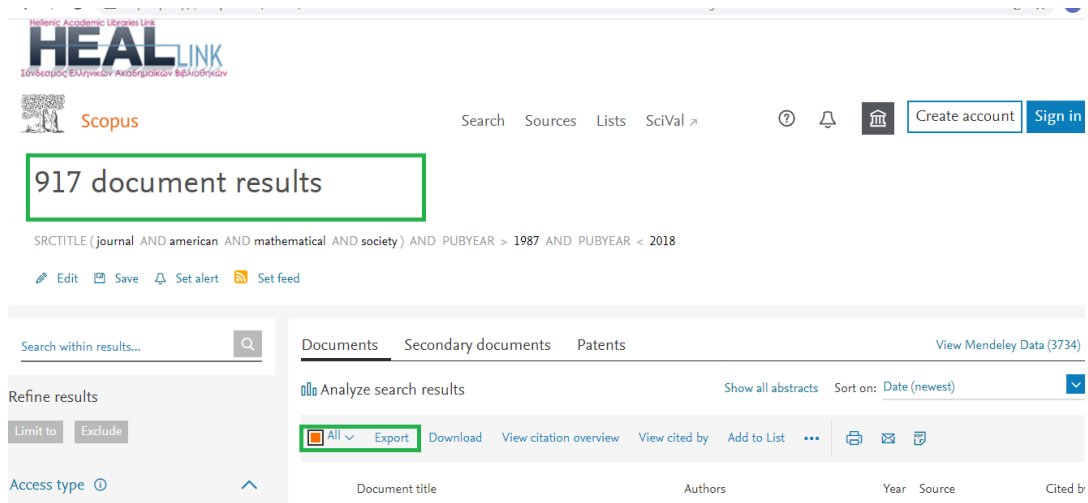


Figure 1.2 Document Results and Export

The results show that there are 917 documents available. Then we exported all the files (Figure 1.2).

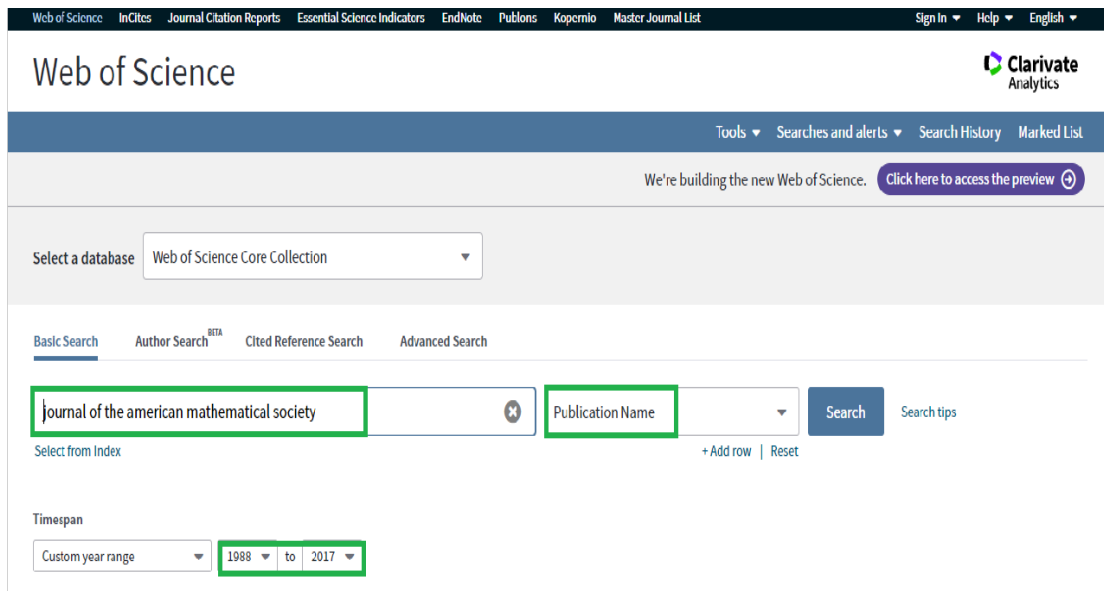


Figure 1.3 Document Search in Web of Science

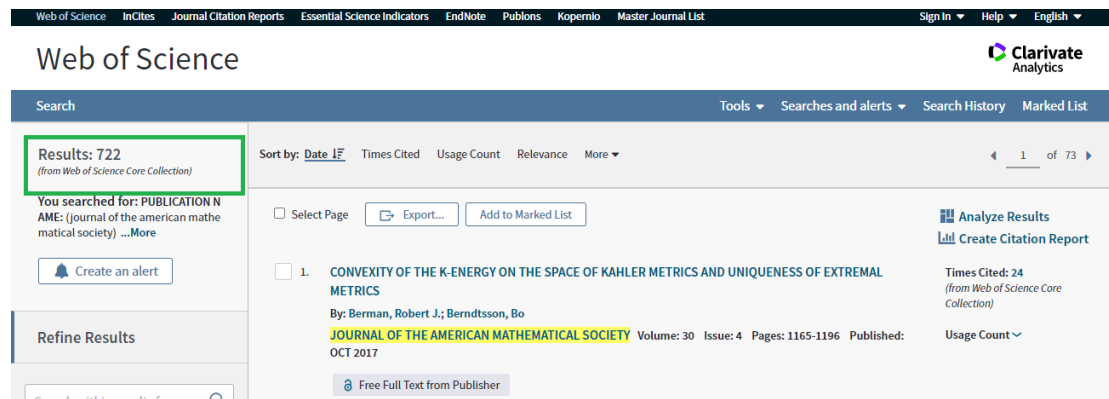


Figure 1.4 Document Results and Export

The results show that there are 722 documents available. Then we exported all the files.

## 2 General information of the data

### 2.1. The export of information to CSV excel

All the information that could be exported was indeed exported. It was then converted to a CSV excel file named JMS for Scopus. Web of Science was a little bit different. It was then converted to two excel files named wos1 and wos2 from Web of Science. In this case, it was necessary to download our documents in two separate files. As there is the possibility to download 500 files each time in the specific source of information. Then two files consisting of 222 and 500 documents were downloaded respectively and then merged into one file named.

### 2.2. Journal of the American Mathematical Society data

The information we have collected in the CSV excel file called jams consists of 45 columns sorted and subcategories containing the selected information, 918 rows corresponding to the 917 documents we found through Document Search plus a row that we exported and contains the document categories that had occurred in the past. The following information is back in Scopus. More specifically through Scopus "What is Scopus Field Code? "Field codes are used in an Advanced search for a term in a specific field. Also, the information we have collected in the Excel file called wos consists of 67 columns that are sorted, subcategorized, and contain the selected

information. 723 rows correspond to the 722 documents we found through Document Search plus a row we exported containing the document categories that were previously displayed. The following information is back on the Web of Science. More specifically through the Web of Science section. Field codes are used in an Advanced search for a term in a specific field. Below are all the columns and information for each of them.

- Authors: (A combined field in the following fields of the author: AUTHLASTNAME and AUTHFIRST, is a variable that can be characterized as a character. An example of this variable is: Wood M.M Brown A., Hertz F.R.)
- Author(s)ID: (Scopus Author Identifier distinguishes between ambiguous names, assigning to each author in Scopus a unique number, grouping all the documents written by this author, is a variable that can be characterized as a character. An example of this variable is: 15064764200)
- Title: (the title of the document, is a variable that can be characterized as a character. An example of this variable is: The distribution of sandpile groups of random graphs)
- Year: (date of issue of the document, is a variable that can be characterized as a numeric. An example of this variable is: 2017,2016)
- Source title: (The title of the journal, book or report in which the document was published, is a variable that can be characterized as a character. An example of this variable is: Journal of the American Mathematical Society)
- Volume: (Recognition for a serial version is a variable that can be characterized as a numeric. An example of this variable is: a set of numbers ranging between 0-30)
- Issue: (Recognition for a serial version, is a variable that can be characterized as a numeric. An example of this variable is: a set of numbers ranging between 1-4)
- Article Number: (A permanent identifier for a document used by some publishers instead of, or in addition, by page numbers. Article numbers can be assigned at the time of publication, so documents can be listed and searched earlier in publication process, is a variable that can be characterized as logical)

- Page Start: (Indicates the first page of a page width in a version, is a variable that can be characterized as a numeric. An example of this variable is: a set of numbers ranging between 0-1300)
- Page End: (Indicates the last page of a page width in one version, is a variable that can be characterized as a numeric. An example of this variable is: a set of numbers ranging between 0-1300)
- Page Count: (A combination field that looks for the fields PAGEFIRST and PAGELAST, is a variable that can be characterized as a logical.)
- Cited by: (is a variable that can be characterized as a numeric. An example of this variable is: a set of numbers ranging between 0-800)
- DOI: (A unique alphanumeric string created to recognize a piece of intellectual property in an online environment, is a variable that can be characterized as a character. An example of this variable is: 10.1090/jams/866)
- Link: (The URL of a website of a reported reference, is a variable that can be characterized as a character. An example of this variable is: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85025842552&doi=10.1090%2fjams%2f866&partnerID=40&md5=ce62cf5e0740fdedcxfb3d2fe07a27>)
- Affiliations: (Organizing a Writer's Collaboration Address, is a variable that can be characterized as a character. An example of this variable is: Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53705, United States; American Institute of Mathematics, 360 Portage Avenue, Palo Alto, CA 94306-2244, United States)
- Authors with affiliations: (AFFIL is a combined field looking for the following author address fields: AFFILCITY, AFFILCOUNTRY and AFFILORG. When searching in the AFFIL field, you can specify if you want all search terms to be in the same relationship, is a variable that can be characterized as a character. An example of this variable is: Wood, M.M., Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53705, United States, American Institute of Mathematics, 360 Portage Avenue, Palo Alto, CA 94306-2244, United States)
- Abstract: (Summary of a document, is a variable that can be characterized as a character)

- Author Keywords: (Keywords assigned to the document by the author, is a variable that can be characterized as a character. An example of this variable is: Measure rigidity; Non-uniform hyperbolicity; Random dynamics; SRB measures; Stiffness of stationary measures)
- Index Keywords: (Controlled Dictionary Terms Assigned to the Document, is a variable that can be characterized as a logical.)
- Molecular Sequence Numbers: (The number attributed to a sequence of amino acids or nucleotides defined or referred to in a document, is a variable that can be characterized as a logical.)
- Chemicals / CAS: (This is the name of a chemical / A numeric identifier assigned to a substance when it enters the CAS registry database, is a variable that can be characterized as a logical.)
- Tradenames: (A name used to identify a commercial product or service, is a variable that can be characterized as a logical.)
- Manufactures: (The name of a manufacturer, such as a device manufacturer, is a variable that can be characterized as a logical.)
- Funding Details: (A combined field looking for funding recognition text as well as the following funding fields: FUND-NO, FUND-ACR, FUND-SPONSOR, is a variable that can be characterized as a character. An example of this variable is: National Science Foundation, NSF: DMS-1147782, DMS-1301690 American Institute of Mathematics, AIM)
- Funding Text 1: (Project support number, is a variable that can be characterized as a character. An example of this variable is: This work was done with the support of an American Institute of Mathematics Five-Year Fellowship, a Packard Fellowship for Science and Engineering, a Sloan Research Fellowship, and National Science Foundation grants DMS-1147782 and DMS-1301690)
- References: (REF is a combined field that looks for the fields REFAUTH, REFTITLE, REFSRCTITLE, REFPUBYEAR, REFPAGE and WEBSITE. When searching in the REF field, you can specify if you want all search terms to be in the same report, is a variable that can be characterized as a character. An example of this variable is: Alfaro, C.A., Valencia, C.E., On the sandpile group of the cone of a graph (2012) Linear Algebra Appl, 436 (5), pp. 1154-1176)

- Correspondence Address: (is a variable that can be characterized as a character. An example of this variable is: Wood, M.M.; Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, United States; email: mmwood@math.wisc.edu)
- Editors: (A combined field looking for the following fields: EDLASTNAME and EDFIRST is a variable that can be characterized as a logical.)
- Sponsors: (Sponsor who provides a grant or funding for the project, is a variable that can be characterized as a logical.)
- Publisher: (Search for books by the publisher called, is a variable that can be characterized as a character. An example of this variable is: American Mathematical Society)
- Conference name: (The name of a conference, is a variable that can be characterized as a logical.)
- Conference date: (The date of a conference, is a variable that can be characterized as a logical.)
- Conference location: (The location of a conference, is a variable that can be characterized as a logical.)
- Conference code: (The code of a conference, is a variable that can be characterized as a logical.)
- ISSN: (A unique identification number assigned to all serial publications, is a variable that can be characterized as a numeric. An example of this variable is: 8940347)
- ISBN: (A unique identification number assigned to all books, is a variable that can be characterized as a logical.)
- CODEN: (A unique six-character code that identifies serial and non-serial posts, is a variable that can be characterized as a logical.)
- PubMed ID: (A unique identifier for all Medline documents, is a variable that can be characterized as a logical.)
- Language of Original Document: (The language in which the original document was written, is a variable that can be characterized as a character. An example of this variable is: English)
- Abbreviated Source Title: (is a variable that can be characterized as a character. An example of this variable is: J. Am. Math. Soc)



- Document Type: (Limits your search to document types, is a variable that can be characterized as a character. An example of this variable is: Article)
- Publication Stage: (is a variable that can be characterized as a character. An example of this variable is: Final)
- Access Type: (Password field code used to filter documents from open access documents, is a variable that can be characterized as a character. An example of this variable is: Open Access)
- Source: (The source for exporting documents, is a variable that can be characterized as a character. An example of this variable is: Scopus)
- EID: (Electronic Identifier is a unique alphanumeric string created to recognize a Scopus record. EIDs are visible through the document export function. They are internal IDs intended for Scopus use only, is a variable that can be characterized as a character. An example of this variable is: 2-s2.0-85025842552).
- ISO:Source abbreviation
- Authors address:Information about the University address
- Digital Object Identifier:Digital Object Identifier
- Unique article identified: Unique Article Identified
- Cited References:Cited References (in wos core collection)
- Subject Categories:Subject Categories
- Usage Count:Usage Count
- Web of Science Categories:Web of Science Categories
- Email Address:The e-mail address of the author(s)
- IDS Number:Identifies an issue of a journal. Used to order copies of articles from a document delivery service
- Reprint Address:The address of the reprint author. It may include reprint author, organization, sub organization, street, city, state or province, zip or postal code
- Bibliographic Database: Which bibliographic database is used
- Authors University: Which is the authors university,Author name ,publication year, source
- Software Review:Contain information which distinguishes each paper

### 2.3. Finding NA values in Scopus dataset

The first thing that we must do is to be cleaning our data by searching for missing values in our data set. Thankfully, with the help of RStudio we will find out if there are any NA values in our data, we find the exact number of NA in each of the information columns we have selected.

Below are two tables detailing all the columns of Scopus and Web of Science as well as the number of Na contained in each column from each source respectively.

Authors	0
Author(s) ID	0
Title	0
Year	0
Source title	0
Volume	0
Issue	0
Art. No.	917
Page start	0
Page end	11
Page count	917
Cited by	2
DOI	148
Link	0
Affiliations	14
Authors with affiliations	3
Abstract	0

Author Keywords	535
Index Keywords	917
Molecular Sequence Numbers	917
Chemicals/CAS	917
Tradenames	917
Manufacturers	917
Funding Details	883
Funding Text 1	888
References	5
Correspondence Address	178
Editors	917
Sponsors	917
Publisher	769
Conference name	917
Conference date	917
Conference location	917
Conference code	917
ISSN	0
ISBN	917
CODEN	917
PubMed ID	917
Language of Original Document	0
Abbreviated Source Title	0
Document Type	0

Publication Stage	0
Access Type	262
Source	0
EID	0

Table 2-1 Exact Number of Na per columns of Scopus

Here we will present a list of columns that consist only of missing values of Scopus:

<b>Scopus</b>
Article number
Page count
Index Keywords
Molecular Sequence Numbers
Chemicals/CAS
Tradenames
Manufacturers
Editors
Sponsors
Conference name
Conference date
Conference location
Conference code
ISBN
CODEN
PubMed ID

Table 2-2 Columns with only Na values

While there are columns that we really shouldn't have missing values and there really are not:

- Source Title
- ISSN

- EID

## 2.4. Finding NA values in Web of Science dataset

In the table below we see all the columns that exist in the file which we downloaded from the Web of Science database. In this table, in addition to all the columns, we also see the exact number Na that exists in each column.

Publication Type	0
Authors	0
Book Authors	722
Book Editors	722
Book Group Authors	722
Author Full Names	0
Book Author Full Names	722
Group Authors	722
Article Title	0
Source Title	0
Book Series Title	722
Book Series Subtitle	722
Language	0
Document Type	0
Conference Title	722
Conference Date	722
Conference Location	722
Conference Sponsor	722
Conference Host	722
Author Keywords	375
Keywords Plus	111
Abstract	688
Addresses	40

Reprint Addresses	11
Email Addresses	272
Researcher Ids	630
ORCIDs	583
Funding Orgs	488
Funding Text	489
Cited References	722
Cited Reference Count	0
Times Cited, WoS Core	0
Times Cited, All Databases	0
180 Day Usage Count	0
Since 2013 Usage Count	0
Publisher	0
Publisher City	0
Publisher Address	0
ISSN	0
eISSN	374
ISBN	722
Journal Abbreviation	0
Journal ISO Abbreviation	0
Publication Date	288
Publication Year	0
Volume	0
Issue	0
Part Number	722
Supplement	722
Special Issue	722
Meeting Abstract	722
Start Page	0
End Page	0

Article Number	586
DOI	136
Book DOI	722
Early Access Date	722
Number of Pages	0
WoS Categories	0
Research Areas	0
IDS Number	0
UT (Unique WOS ID)	0
Pubmed Id	721
Open Access Designations	210
Highly Cited Status	722
Hot Paper Status	722
Date of Export	0

*Table 2-3 Exact Number of Na per columns of Scopus*

Here we will present a list of columns that consist only of missing values of Web of Science:

<b>Web of Science</b>
Book Authors
Book Editors
Book Group Authors
Book Author Full Names
Group Authors
Book Series Title
Book Series Subtitle
Conference Title
Conference Date
Conference Location
Conference Sponsor
Conference Host

ISBN
Part Number
Special Issue
Cited References
Supplement
Meeting Abstract
Book DOI
Early Access Date
Host Paper Status
Highly Cited Status

*Table 2-4 Columns with only Na values*

While there are columns that we really shouldn't have missing values and there really are not:

- ISSN
- Source Title
- Article Title

Because there are many columns of information in both sources which are only with empty data, we decided to remove them for a better analysis of our data.



## 2.5.Unique columns in our dataset

I also found that there are columns of information that do not exist in both sources and only exist in one. Below there will be a list of columns of information that exist in Scopus and Web of Science respectively.

<b>Scopus</b>
Link
EID
Source
Access Type
CODEN
Document Type
Abbreviated Source Title
Conference Name
Conference Code
Editors
Sponsors
References
Correspondence Address
Tradenames
Index Keywords
Manufactures
Chemicals/Cas
Molecular Sequence Numbers
Cited by

*Table 2-5Columns that only exists in Scopus*

<b>Web of Science</b>
Hot Paper Status
Date of Export
Highly Cited Status
Open Access Designations
IDS Number
UT (Unique WOS ID)
Research Areas
Early Access Date
Book DOI
Meeting Abstract
Supplement
Special Issue
Journal Abbreviation
Journal ISO Abbreviation
Publication Date
Publication Year
Publisher City
Publisher Address
Times Cited, All Databases
Times Cited, WoS Core

Cited Reference Count
Cited References
ORCIDs
Funding Orgs
Email Address
Researcher Ids
Addresses
Reprint Addresses
Keywords Plus
Conference Host
Conference Title
Conference Sponsor
Book Series Subtitle
Book Series Title
Book Author Full Names
Author Full Names
Book Authors
Book Editors
Book Group Authors

*Table 2-6 Columns that only exists in Web of Science*

## 2.6. Identifying Double elements

Now we will try to find in the following 3 columns of information if there is any information that is repeated. For the information to be correct, the data should not be presented again.

We will do this with the following command: `duplicated ()`: determines which elements of a vector or data frame are duplicates of elements with smaller subscripts, and returns a logical vector indicating which elements (rows) are duplicates.

The information columns we will study are as follows:

- Title
- DOI
- EID

Using the `duplicated` command, we found the following:

We first researched the DOI information column. In which we found 147 double elements. All of which have the Na value. The same thing we see in Web of Science. We found 135 double elements that their values are Na.

As for the other 2 information columns we wanted to study which were EID and Title.

We found out that there are not any double elements because the electronic identification (EID) is unique for every observation a column that only exists in Scopus.

We also searched for double elements in the column named "Title". We found out that there are not any double elements in this column in both of databases.

Next, we searched for double elements in more than one column of information such as:

- Volume, Issue, Page start
- Volume, Issue, Page end

In both cases and in both databases, we have no duplicated values.

In this part, we are going to deal with columns that they have only one value in the entire column and this value is the same in all the rows of these specific columns. We are also going to talk about specific columns that they have only a few values that we are interested in.

So firstly, we search for the content of the column named “ISSN”. We locate that exists only one registration for all the lines of this column and this registration is the numeric integer character “8940347” for Scopus and 0894-0347 for Web of Science. This happened because in every journal newspaper magazine or periodical of all kinds that they had published there is a digit code to identify them. This is the ISSN code and is the same in our research because all the data are published in the same journal, the Journal of the American Mathematical Society.

Next, we search for the registrations in the column named “Publisher”. We figured out that there only 2 values. The first is the Publisher of the journal which is the “American Mathematical Society” and the other value is the Na value. For Web of Science, we found that there is not Na values in this column but only the value “American Mathematical Society”.

Lastly, we searched for the registrations in the column named “Source title” and we discovered that there is only one value in all of the 917 lines of this column and this value is the name of the journal which is “Journal of the American Mathematical Society” as we expected and the same was with Web of Science.

## 2.7. Number of papers in the dataset

Then we studied the number of documents that existed in each source as well as the difference that may have between them.

In this table, we see the number of documents that exist in the two sources but also the number of documents from the magazine.

The differences between the two sources and the magazine are minimal. What is worth saying is that there is only one extra entry in the date 2007 but there we see that the Volume is the same as that of 2008. More clearly, we can see the specific difference in the table where the Web of Science documents are and not there is this entry on the site of the magazine.

Besides, we can see in this table that although the search of the documents was done in both sources with common dates of publication of the files, we see that the Web of Science source has no documents from the dates 1988-1994. Unlike Scopus which contains all the documents from these calendars and based on the documents that are certified in the magazine. Also, this significant difference is shown in the table where they exist in Scopus and not in the other source.

	Year	Volume	Issue	Scopus	Wos	The number of the papers of the American mathematical Society
1	1988	1	1	7	0	7
2	1988	1	2	7	0	7
3	1988	1	3	5	0	5
4	1988	1	4	6	0	6
5	1989	2	1	5	0	5
6	1989	2	2	7	0	7
7	1989	2	3	8	0	8
8	1989	2	4	8	0	8

9	1990	3	1	7	0	7
10	1990	3	2	7	0	7
11	1990	3	3	9	0	9
12	1990	3	4	8	0	8
13	1991	4	1	7	0	7
14	1991	4	2	5	0	5
15	1991	4	3	11	0	11
16	1991	4	4	8	0	8
17	1992	5	1	8	0	8
18	1992	5	2	7	0	7
19	1992	5	3	4	0	4
20	1992	5	4	10	0	10
21	1993	6	1	6	0	6
22	1993	6	2	5	0	5
23	1993	6	3	7	0	7
24	1993	6	4	7	0	7
25	1994	7	1	10	0	10
26	1994	7	2	6	0	6
27	1994	7	3	6	0	6
28	1994	7	4	4	0	4
29	1995	8	1	7	7	7
30	1995	8	2	6	6	6
31	1995	8	3	6	6	6
32	1995	8	4	9	9	9
33	1996	9	1	9	9	9

34	1996	9	2	9	9	9
35	1996	9	3	9	9	9
36	1996	9	4	8	8	8
37	1997	10	1	9	9	9
38	1997	10	2	10	10	10
39	1997	10	3	8	8	8
40	1997	10	4	6	6	6
41	1998	11	1	6	6	6
42	1998	11	2	8	8	8
43	1998	11	3	11	11	11
44	1998	11	4	7	7	7
45	1999	12	1	9	9	9
46	1999	12	2	7	7	7
47	1999	12	3	8	8	8
48	1999	12	4	11	10	11
49	2000	13	1	8	8	8
50	2000	13	2	7	7	7
51	2000	13	3	10	10	10
52	2000	13	4	9	9	9
53	2001	14	1	7	7	7
54	2001	14	2	9	9	9
55	2001	14	3	6	6	6
56	2001	14	4	4	4	4
57	2002	15	1	7	7	7
58	2002	15	2	7	7	7



59	2002	15	3	7	7	7
60	2002	15	4	6	6	6
61	2003	16	1	10	10	10
62	2003	16	2	8	8	8
63	2003	16	3	8	8	8
64	2003	16	4	9	9	9
65	2004	17	1	8	8	8
66	2004	17	2	9	9	9
67	2004	17	3	5	5	5
68	2004	17	4	8	8	8
69	2005	18	1	6	6	6
70	2005	18	2	6	6	6
71	2005	18	3	9	9	9
72	2005	18	4	7	7	7
73	2006	19	1	8	8	8
74	2006	19	2	8	8	8
75	2006	19	3	8	8	8
76	2006	19	4	5	5	5
77	2007	20	1	11	11	11
78	2007	20	2	9	9	9
79	2007	20	3	7	7	7
80	2007	20	4	9	9	9
81	2007	21	1	0	4	0
82	2008	21	1	9	5	9
83	2008	21	2	10	10	10

84	2008	21	3	11	11	11
85	2008	21	4	11	11	11
86	2009	22	1	8	8	8
87	2009	22	2	9	9	9
88	2009	22	3	9	9	9
89	2009	22	4	8	8	8
90	2010	23	1	8	8	8
91	2010	23	2	9	9	9
92	2010	23	3	10	10	10
93	2010	23	4	8	8	8
94	2011	24	1	9	9	9
95	2011	24	2	7	7	7
96	2011	24	3	9	9	9
97	2011	24	4	9	9	9
98	2012	25	1	8	8	8
99	2012	25	2	10	10	10
100	2012	25	3	10	10	10
101	2012	25	4	7	7	7
102	2013	26	1	8	8	8
103	2013	26	2	6	6	6
104	2013	26	3	8	8	8
105	2013	26	4	7	7	7
106	2013	27	1	6	6	6
107	2014	27	2	6	6	6
108	2014	27	3	7	7	7

109	2014	27	4	6	7	7
110	2015	28	1	8	8	8
111	2015	28	2	6	6	6
112	2015	28	3	6	6	6
113	2015	28	4	6	6	6
114	2016	29	1	7	7	7
115	2016	29	2	9	9	9
116	2016	29	3	8	8	8
117	2016	29	4	6	6	6
118	2017	30	1	6	6	6
119	2017	30	2	6	6	6
120	2017	30	3	6	6	6
121	2017	30	4	7	7	7

*Table 2-7The number of papers in both datasets compared to papers of the Journal*

In this table we see the documents which are in the source Web of Science but are different from the magazine.

	Publication Year	Volume	Issue	Wos(anti)
1	2007	21	1	4
2	1999	12	4	1
3	2008	21	1	4

*Table 2-8Papers that only exists in Web of Science*

While In this table we see the documents, which are in the source Web of Science but are different from the magazine.

	Year	Volume	Issue	Scopus(anti)
1	1988	1	1	7
2	1988	1	2	7
3	1988	1	3	5
4	1988	1	4	6
5	1989	2	1	5
6	1989	2	2	7
7	1989	2	3	8
8	1989	2	4	8
9	1990	3	1	7
10	1990	3	2	7
11	1990	3	3	9
12	1990	3	4	8
13	1991	4	1	7
14	1991	4	2	5
15	1991	4	3	11
16	1991	4	4	8
17	1992	5	1	8
18	1992	5	2	7
19	1992	5	3	4
20	1992	5	4	10
21	1993	6	1	6
22	1993	6	2	5
23	1993	6	3	7
24	1993	6	4	7
25	1994	7	1	10
26	1994	7	2	6
27	1994	7	3	6
28	1994	7	4	4
29	2007	27	4	1

*Table 2-9 Papers that only exists in Scopus*

### **3 The merge of the files**

#### **3.1. The merge of the information of our dataset**

Once we have seen what kind of data contains our data and took a first look at what information it gives us, it is time to combine all this information to have the best possible result for study and analysis.

The first association criterion we used was the DOI, which is an information column that contains a unique number for each document. We studied if there are Na prices and double entries for this column. This process was done because during our union there is a high probability that we will not get a reliable result for the study. After all this information that was not useful to us was removed, the union took place. The result was 384 documents having a common Doi. But many documents were still missing to reach the desired result.

Each time the union is done with each of the criteria that we will see below, an archive is created with documents that remain outside the union and this is the file that we join each time with the Na and the duplicate elements so that we can consider our next criterion.

So, we used one more join criterion for the number of not united documents. This criterion is the Title of the documents. However, we did not forget to put in this union the documents that had Na and double entries in the Doi column as it is not necessary to have Na and duplicate entries in the title as well. What was observed when joining the title was that initially it could not be done as the source of Web of Science the letters in the titles were in capitals and in Scopus in lower case so could not recognize the titles even if they were the same then we made the letters of both sources in small letters.

Another problem was found, however, because the sources are from all over the world, it was logical and next and the writing language to correspond to the country where it was written, so there were some tones and idioms that in the source of Web of Science did not exist, so it had to be done and the elimination of this problem.

After the previous procedure, we have done the locating and the removal of empty and duplicate elements. Finally, we made the union and the result was 272 documents with a common title.

We did the above procedure not only for an information column but also for a combination of columns. These columns were Page Start, Page End, Volume, Issue. The combinations used were as follows:

Page Start, Page End, Volume, Issue

Page Start, Volume, Issue

Page End, Volume, Issue

In the combination of the columns it was checked separately for each column if there are duplicate entries and blank entries and after we deleted this information we went and made the union of the elements that were not united with the title as well as the Na and duplicate values. The result produced another 57 documents which were merged. The same procedure was done for the other two column combinations Page Start, Volume, Issue, Page End, Volume, Issue, and the result was the union of even 7 and one document respectively.

				<b>Scopus</b>		
	Rows	Na	Dups	Clean of Na and Dups	Merged	No Merged
DOI	917	148	0	769	384	385
Title	533	0	0	533	272	261
PageStart,PageEnd,Volume,Issue	261	3	11	247	57	190
Page Start,Volume,Issue	204	0	43	161	7	154
Page End ,Volume,Issue	197	1	20	176	1	175
Total					721	

Table 3-1The cleaning of the data and the merge file for Scopus

				<b>Wos</b>		
	Rows	Na	Dups	Clean of Na and Dups	Merged	No Merged
DOI	722	136	0	586	384	202
Title	338	0	0	338	272	66
PageStart,PageEnd,Volume,Issue	66	0	1	65	57	8
Page Start,Volume,Issue	9	0	0	9	7	2
Page End,Volume,Issue	2	0	0	2	1	1
Total					721	

Table 3-2The cleaning of the data and the merge file for Web of Science

### 3.2.The differences in the merged file

From the moment the archive was merged as we saw before we decided to investigate if there are differences in the common elements for some columns of information that we deemed to be worth research and study. Initially, we saw in each column that we will analyze below for differences between the columns as well as for blank values in both sources and each one separately.

The first information column we dealt with was the DOI column. In the first approach, it was observed that several of the differences had to do with the difference between uppercase and lowercase numbers within the unique number of each document. So, this change was made from 68 to just 25 different. These differences must do mainly with differences in the digits of these unique numbers. There were also many blank values in the global archive and the largest number of blank data for the DOI in the Web of Science source.

It was then examined to what extent the united file had common titles for its documents. We followed the same technique to make the letters smaller in both sources but here something else was needed. Because we are talking about a global data source various tones were either not placed correctly when recording the titles or were omitted so we

corrected this problem as well. Also, this procedure was followed in advance so that the association with this information column could take place.

What was found was that there was no blank information for this column which is something we were expecting to see. As for the errors observed in this case, a comparison was made with the magazine itself to determine which source is more reliable in the piece of this information. The main error between the 65 different titles between the two sources was the error on the part of the Web of Science was the error of notation of different symbols such as (+, =, &) as well as numerical powers and different types (Mg, SO<sub>3</sub>).

However, there are differences in the titles that have to do with the writing language of the title. We see that on the Web of Science the titles have been translated from French to English as Scopus has kept the French title of the document. In the French language, we see that they are the same titles as presented by the official site of the magazine.

Then the differences with the combinations of the columns Page Start, Page End, Volume, Issue were studied.

But first, we studied each column separately in terms of differences and Na values. The results were as we expected for Volume and Issue, are without differences and empty prices. Besides, no empty values and differences were found even in the combination of Page Start. While in the combination of Page End were found two differences which concern differences in numbers of the pages of the journal as well as two empty information in the source of Web of Science and eight in the source of Scopus. All of the above are confirmed by the combination of all four columns.

Then we studied the columns containing the names of the authors which is called Authors. In the beginning, both sources had this column as one and it contained all the names of the authors. So, for the best analysis of this column, it was decided to break this column into as many columns as needed with the maximum number of authors needed even for a document. An expected result was the Authors 1 column that has no blank value. For differences, we had to do the same procedure for uppercase and lowercase letters as the Web of Science source had it all small and Scopus uses both uppercase and lowercase letters. A first test was performed without changing the letters and concerning the result after the change, the number of errors was found. After all these procedures have been done the errors that arise have to do with differences in



letters in punctuation (such as dots and commas). Also, we found that most documents are written by either one or a small group of authors (two to three authors) compared to a group of four or five authors.

Another column that was studied was the column called Authors with Affiliations. This is also a column of information that was originally one in both sources and we had to separate it for a better analysis. As expected, it is also related to the number of Authors columns and that is why here we have 5 columns for Authors with Affiliations. The differences found here have to do with punctuation, with differences in the author's description beside the Web of Science source also contains the author's name about the other source.

Another column that was worth analyzing is the column that contains the Abstract of the documents. It is worth noting here that the Scopus source contains many empty values 511 compared to the other source which contains only 176 empty values. The differences contained mainly relate to the content of the summary as they either have different content or different in scope.

Then two more columns were studied, these are the language in which the documents are written and the type of document. As for the column of information related to the language of the author we saw as expected that there was no blank information. As far as the mistakes have been made, it has to do with the fact that WoS has translated the documents into English, so about Scopus which has kept a record in the original writing language which is French, we see this remark in 3 documents, although in the titles we have a meeting in Scopus to have more French titles.

Finally, for the Document Type information column we see here that there is no blank value. The differences that exist are for the different characterization of the document from each source for example that one document is characterized by one source as a note and the other as an article.

Below are presented in a table in detail what was mentioned before regarding the differences and the empty variables.

	Differences	Both Na	Na only wos	Na only Scopus
DOI	25	269	204	80
Title	65	0	0	0
Page Start	0	0	0	0
Page End	0	0	8	2
Volume	0	0	0	0
Issue	0	0	0	0
PageStart,Page end, Volume,Issue	2	10	8	2
Page Start, Volume, Issue	0	0	0	0
Page End, Volume, Issue	2	10	8	2
Authors 1	33	0	0	0
Authors 2	15	252	251	250
Authors 3	6	540	540	538
Authors 4	5	690	689	688
Authors 5	1	715	714	715
Authors with Affiliations 1	16	12	1	12
Authors with Affiliations 2	65	335	249	291
Authors with Affiliations 3	27	587	538	547
Authors with Affiliations 4	8	699	687	675
Authors with Affiliations 5	1	717	715	707
Abstract	27	687	176	511
Author Keyword	335	381	361	375

Document Type	13	0	0	0
Language	3	0	0	0

*Table 3-3 The differences and the Na of each column in the merged file*

When the files were merged there were several from the Scopus source which were not merged. This was because the Web of Science source contains data for the journal from 1995 onwards. While Scopus contains data from the first year that the magazine was released. The table below shows the number of records per year for the period 1988-1994.

Year	Volume	Number of papers
1988	1	25
1989	2	28
1990	3	31
1991	4	31
1992	5	29
1993	6	25
1994	7	26

*Table 3-4 Papers from Scopus in time period 1988-1994*

In the table above we find that the Scopus source has more information about the journal, we are studying in terms of the number of subscriptions contained in the two sources. It contains all the documents that were written in those years based on the official site of the magazine itself. Unfortunately, as the above documents do not exist in the Web of Science, we cannot include them in the analysis that will be made below regarding the characteristics of the authors and the productivity of the journal.

## 4 Authors characteristics

This section aims to analyze the most productive authors in the Journal during the period of study. For this purpose, was developed one of the significant laws in bibliometrics studies known as Lotka's law. According to Lotka A.J(1926), in each area of science, there are a lot of authors who publish only one study while a small group of prolific contribute with many publications.

### 4.1. Authors productivity and Lotka's Law

Lotka's law is defined as the total number of authors  $y$  in each subject, each producing  $x$  publications, is inversely proportional to some exponential function  $n$  of  $x$ . Hence, describes the frequency of publications by authors in each subject field. It is expressed as: where

$$y_x = c \times x^{-n}$$

$x = \text{number of publications}$

$y_x = \text{number of authors credited with } x \text{ publications}$

$n = \text{constant, equals 2 for scientific subjects}$

$C = \text{constant}$

Number of papers which have written by authors	Number of authors	frequency(%)
1	999	77.44
2	207	16.05
3	51	3.96
4	18	1.4

5	6	0.47
6	4	0.32
7	1	0.07
8	1	0.07
9	2	0.15
12	1	0.07
Total	1290	100

Table 4-1 Authors productivity

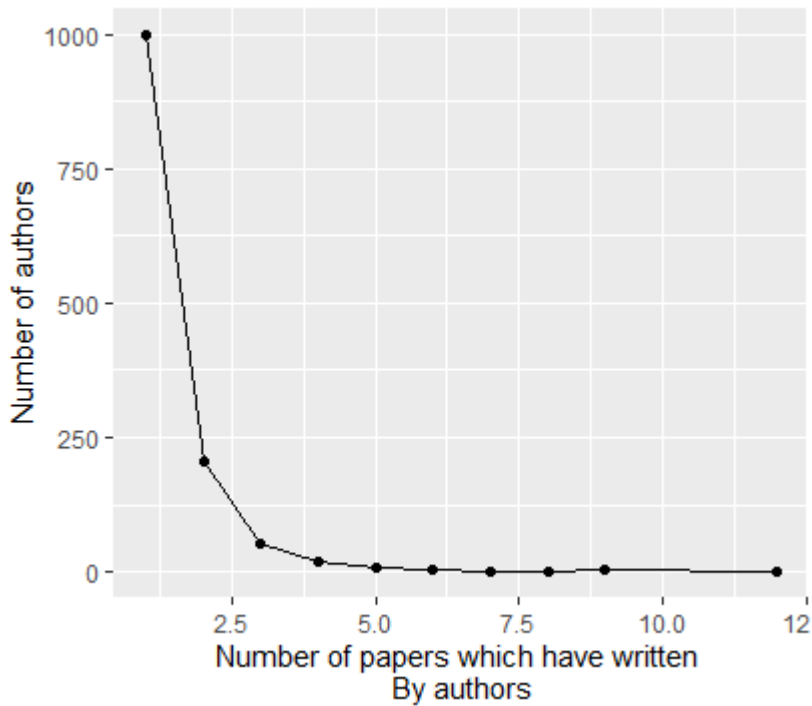


Figure 4.1 Lotka's Law distribution

From the above graph it seems clearly that our data follows Lotka's law distribution.

The first two columns of the table correspond to the data shown in the previous, indicating the number of contributions made by each author. Specifically, column  $x$  corresponds to the number of studies published by the authors in this field. In this case, authors have contributed between one and twelve articles. The next column,  $y_x$ , corresponds to the number of authors publishing a given number of articles. It can be

seen that 999 researchers in this area have contributed a single article, while one author have been the most prolific. As is shown, a total of 1,290 authors ( $\sum y_x$ ) have been involved in producing the articles included in this bibliometric study. The next two columns (X and Y) correspond to the logarithm of the frequency of articles (x) and authors ( $y_x$ ), and the product of these is shown in the next two columns. As the totals for these data are also available at the bottom of the table, it is possible to calculate the frequency of authors with a single article, and those with two, three, etc. (corresponding to column  $y_x / \sum y_x$ ). So far, these data have been directly obtained from the frequency of authors publishing x articles. Once this frequency (the observed frequency) has been obtained, Lotka's law can be applied to obtain the expected frequency of authors publishing x articles. Thus, it is necessary to calculate the n exponent for this particular case. Although Lotka's law proposes a growth in production according to its formula  $y_x = c \times x^{-n}$ , the law has to be tested for the data in question. This will yield the exponent n, which corresponds to the present distribution of author productivity and also indicates whether the data really fit Lotka's law. Consequently, the first step is to calculate the exponent n using the least squares method and according to the following formula:

$$n = \frac{N \sum XY - \sum X \sum Y}{N \sum X^2 - (\sum X)^2}$$

All the data needed for the n formula can be obtained from the table. The only index that requires further work is N, which represents the number of pairs considered. Those authors who have published between one and twelve articles will be considered, so this will represent ten pairs of data (N = 10). There is one specific case where not all the pairs of data are included in the analysis of Lotka's law. This is when  $y_x = 1$  is found at the end of the distribution, corresponding to the highest values of x (number of articles). In such cases, this small group of most prolific authors are excluded from the analysis in order not to overestimate the results. Into the n formula stated above gives:

$$n = \frac{10 \times 3.567297664 - (6.638944282 \times 9.959619767)}{10 \times 5.379791576 - (6.638944282)^2}$$

Thus, the value of n (absolute value) is 3.38, which will then be the specific value of the coefficient in Lotka's formula that will explain author productivity in this case. As the formula of Lotka's law is  $y_x = c \times x^{-n}$ , and n is now known, the only index left to calculate is c. This value is obtained as follows:

$$c = \frac{1}{\sum 1/x^n}$$

For our data, c will take the value 0.851456242, and thus Lotka's formula will be

$$y_x = c \times x^{-n}$$

$$y_x = 0.851456242 \times x^{3.131797597}$$

Here,  $y_x$  represents the expected frequency of authors publishing x documents. To avoid confusion in the nomenclature, let us define the expected frequency of Lotka's law as  $f_e$ , to distinguish it from the observed frequency ( $y_x$ ). By introducing the values taken by the number of articles variable (x) it is now possible to obtain the corresponding expected frequencies. The values taken by  $f_e$  are shown in table, along with its cumulative frequency ( $\sum f_e$ ). As the aim of this analysis is to determine whether these data fit Lotka's law, it is necessary to know the magnitude of the difference between observed and expected frequencies. This difference is shown in the last column of table and is computed by subtracting the cumulative expected frequency from the cumulative observed frequency:  $\sum (y_x / \sum y_x) - \sum f_e$ . This difference, in its absolute value, is shown in the D column.

<b>x</b>	<b>y</b>	<b>X=lgx</b>	<b>Y=lgy</b>	<b>X<sup>2</sup></b>	<b>XY</b>	<b>Y<sub>x</sub>/Σy<sub>x</sub></b>	<b>Σ(y<sub>x</sub>/Σy<sub>x</sub>)</b>	<b>fe</b>	<b>Σfe</b>	<b>D</b>
1	99	0	2.9995	0	0	0.7744	0.7744	0.8514	0.8514	0.0770
	9		6549			18605	18605	56242	56242	37637
2	20	0.3010	2.3159	0.0906	0.6971	0.1604	0.9348	0.0971	0.9485	0.0137
	7	29996	7035	19059	76545	65116	83721	39825	96067	12346
3	51	0.4771	1.7075	0.2276	0.8147	0.0395	0.9744	0.0272	0.9758	0.0014
		21255	7018	44692	18027	34884	18605	84446	80512	61907
4	18	0.6020	1.2552	0.3624	0.7557	0.0139	0.9883	0.0110	0.9869	0.0014
		59991	7251	76233	49356	53488	72093	82361	62873	0922
5	6	0.6989	0.7781	0.4885	0.5439	0.0046	0.9930	0.0055	0.9924	0.0005
		70004	5125	59066	04382	51163	23256	09723	72596	5066
6	4	0.7781	0.6020	0.6055	0.4684	0.0031	0.9961	0.0031	0.9955	0.0005
		5125	59991	19368	93735	00775	24031	12792	85388	38643
7	1	0.8450	0	0.7141	0	0.0007	0.9968	0.0019	0.9975	0.0006
		9804		90697		75194	99225	20818	06207	06982
8	1	0.9030	0	0.8155	0	0.0007	0.9976	0.0012	0.9987	0.0010
		89987		71525		75194	74419	6435	70557	96138
9	2	0.9542	0.3010	0.9105	0.2872	0.0015	0.9992	0.0008	0.9996	0.0004
		42509	29996	78766	55619	50388	24806	74315	44872	20066
12	1	1.0791	0	1.1646	0	0.0007	1	0.0003	1	0
		8125		3217		75194		55128		
	12	6.6389	9.9596	5.3797	3.5672					
	90	44282	19767	91576	97664					

Table 4-2 Lotka's Law



Finally, the Kolmogorov-Smirnov test is applied to verify whether the observed data fit the theoretical distribution according to Lotka's law. The highest value in column (D max) is taken as reference for comparison with the critical value (c.v.) whose general formulation is:

$$c.v. = \frac{1.63}{\left[ \sum y + (\sum y/10)^{1/2} \right]^{1/2}}$$

$$c.v. = \frac{1.63}{\left[ 1290 + (1290/10)^{1/2} \right]^{1/2}}$$

Since the maximum difference (D max) obtained from is 0.077037637, which is smaller than the critical value (0.08070357), the null hypothesis must be accepted. We can therefore conclude that author productivity in this hypothetical research area fits Lotka's law.

Authors cooperation	Number of Papers	Frequency %
1	249	34.67
2	289	40.22
3	149	20.66
4	27	3.6
5	7	0.83

Total	721	100
-------	-----	-----

Table 4-3 Authors cooperation

The table above shows us that the combinations of authors with collaboration of two authors comes first with 290 articles (40.22%) followed by the authorship with one author with 250 articles (34.67%), 149 articles were written by 3 authors (20.66%). Then we have the combination of 4 authors with 26 articles and a percentage (3.6%) Finally, the smallest percentage with just (0.83%) are 6 only articles is from the authorship of 5 authors. Surly, we find that the majority of articles have been made by combinations of authors.

## 4.2. Authors characteristics

At this point, we will see in detail the productivity of the authors as well as the most productive ones based on the number of articles they have written and their total participation throughout the magazine.

We will analyze the characteristics presented by the authors in our journal from the united file we created. The following table lists all the authors who participated in the writing of each article in the magazine. In the Written Papers column is the number of articles each author has written. The contribution column includes the fractional count of each author. The fractional counting method, considering the total number of co-authors in each article which means that  $1/N$  fractional count account for  $N$  authors. Also we have to say that the Contribution number must be smaller than the papers written by the specific author.

Authors	Papers Written	Contributions
scanlon thomas	1	1
abe n	1	0.25
abert m	1	0.5
abouzaid m	1	0.2
abramovich d	2	0.75

abresch u.	1	0.5
abreu m	1	0.5
achlioptas d	1	0.5
acquistapace f	1	0.333333333
adamczak r	1	0.25
adams j	1	0.2
adams s	1	0.5
agler j	1	0.5
agol i	3	1.083333333
alberti g	1	0.333333333
alexander h	1	1
alexandre r	1	0.25
almgren f.j.	1	0.333333333
alon n.	1	0.333333333
ambrosio l	1	0.5
andersen kks	1	0.5
anderson m.t.	1	1
andradas c	1	0.333333333
andrews b	2	1.5
andrews g.e.	2	1
angenent s	2	1.166666667
anick d.j.	1	1
aptekarev a	1	0.333333333

argyros sa	1	0.5
arkhipov s	1	0.333333333
armstrong sn	1	0.333333333
aronson dg	1	0.5
arora s	1	0.333333333
arthur j.	3	3
artstein s	1	0.25
arveson w	1	1
aschbacher m.	2	1.5
aschenbrenner m	1	1
asok a	1	0.5
astala k	1	0.25
atar r	1	0.5
attie o.	1	0.333333333
auroux d	1	0.2
avila a	4	3
babai l	1	0.333333333
babson e	1	0.333333333
bachoc c	1	0.5
backelin e	1	0.5
baik j	1	0.333333333
bakhtin y	1	0.333333333
balakrishnan js	1	0.333333333

balazs m	1	0.333333333
ball km	1	0.25
balogh j	1	0.333333333
bao g	1	0.5
baouendi ms	2	0.833333333
barbasch d	3	1.2
barnetlamb t	1	0.333333333
bartels a	2	0.833333333
barthe f	1	0.25
barvinok a	1	0.5
bass h.	1	0.5
basu s	1	0.333333333
bate d	1	1
bateman m	1	0.5
bayer a	1	0.5
beck j	1	0.5
becker h	1	1
bedford e.	2	1
beilinson a	2	1.333333333
belegradek i	1	0.5
belkale p	2	2
beltran c	1	0.5
bendel cp	2	0.666666667

benjamini i	1	0.5
benkart g	1	0.333333333
benoist y	2	0.833333333
benzvi d	2	0.833333333
berenstein a	1	0.5
berestycki h	1	0.333333333
bergelson v	1	0.5
berkovich vg	1	1
berman rj	1	0.5
bernard p	1	1
berndtsson b	1	0.5
bernstein j	1	0.5
berrick aj	1	0.25
bertram a	2	0.666666667
bertrand d	1	0.5
bestvina m	2	0.833333333
bezrukavnikov r	2	0.833333333
bhargava m	2	1.2
bianchi g	1	0.333333333
bielefeld b.	1	0.333333333
bierstone e.	1	0.5
bigelow sj	1	1
billey s	1	0.5

biran p	1	0.5
birkar c	1	0.25
bismut jm	4	3
bjerklov k	1	0.5
bjorner a.	2	1
blanc j	1	0.5
blekherman g	2	1.333333333
blocha s.	1	1
block j.	2	0.833333333
bombieri e.	1	0.333333333
borcea j	1	0.333333333
borichev a	1	0.5
borisov la	1	0.333333333
boroczky kj	1	0.25
borodin a	1	0.333333333
borwein p	1	0.5
bost j.-b.	1	0.333333333
bott r.	1	0.5
boucksom s	1	0.333333333
bourdon ps	1	0.5
bourgain j	8	5.25
bousch t	1	0.5
bousfield a.k.	1	1

bowditch bh	1	1
bowen l	1	1
boyarchenko m	1	0.5
boyer s	1	0.5
branden p	1	0.333333333
braverman a	1	0.5
braverman m	1	0.5
brendle s	2	1.5
brenier y.	1	1
brenner h	1	0.5
brenti f	1	1
breuer j	1	0.5
breuil c	1	0.25
brezis h	1	0.5
bridgeland t	2	1.333333333
bridson mr	1	1
bringmann k	1	0.333333333
brion m	2	1
brock jf	1	1
broglia f	1	0.333333333
bromberg k	1	1
broto c	2	0.666666667
brown a	1	0.5



brownawell w.d.	1	1
bruillard p	1	0.25
bruin h	1	0.5
brundan j	1	1
bryan j	2	0.7
bryant rl	3	2
brylinski r.	1	0.5
brylinski r.k.	1	1
buch as	2	1.333333333
buechler s	2	0.833333333
bufetov ai	1	1
burdzy k	1	0.5
burger m	1	0.5
burns d.m.	2	1
burq n	2	0.833333333
bushnell cj	2	0.833333333
bux ku	1	0.333333333
buzzard k	1	1
cabre x	1	0.5
caffarelli la	5	2.5
cai jf	1	0.25
cai m	1	0.2
calegari d	4	2.333333333

calegari f	1	1
calta k	1	1
canary r.d.	1	1
cantat s	2	1
caporaso l	2	1.333333333
cappell s.e.	1	0.5
caputo p	1	0.333333333
carbery a	2	0.666666667
cardaliaguet p	1	0.333333333
cascini p	1	0.25
cator e	1	0.333333333
cattani e	1	0.333333333
chan ch	1	0.333333333
chang cy	1	0.333333333
chang mc	1	0.5
chang s.-y.a.	1	0.5
chang s.x.-d.	1	1
chanillo s	1	0.5
charles f	1	0.5
charney r	1	0.5
chazelle b.	1	1
cheeger j.	3	1.333333333
chen gq	1	0.5

chen l	1	0.333333333
chen s	1	1
chen x	5	1.416666667
chenevier g	1	0.5
cherbonnier f	1	0.5
chernousov v	1	0.5
chernov n	1	0.5
chin c	1	1
chinta g	1	0.5
choksi r	1	0.333333333
cholak pa	1	0.333333333
chong ct	1	0.333333333
christ m	4	2.833333333
chung f.r.k.	2	1.5
ciocanfontanine i	1	0.5
clozel l	1	0.5
clutterbuck j	1	0.5
cogdell jw	1	0.5
cohen a	1	0.333333333
cohen fr	1	0.25
cohn h	3	1.333333333
coifman r.r.	1	0.333333333
colding th	1	0.5

colliander j	1	0.2
colliotthelene jl	2	0.833333333
colmez p	1	0.5
coman d	1	0.333333333
concini c.d.	1	0.333333333
conlon d	1	0.333333333
connes a	1	0.5
conrad b	3	0.916666667
constantin p	2	1.5
cooper d	1	0.333333333
cordoba d	1	0.5
cornea o	1	0.5
cortinas g	1	0.333333333
corwin l.	1	1
coutand d	1	0.5
culler m.	1	0.5
cuntz j	2	1
dahlberg b.e.j.	1	0.5
dahmen w	1	0.333333333
dai x.	1	1
damanik d	2	1
darmon h	1	0.5
darvas t	1	0.5

daskalopoulos g	2	0.833333333
davis jf	1	0.25
davis mw	1	0.5
de concini c.	1	0.333333333
de faria e	1	0.5
de jong aj	2	1
de melo w	1	0.5
deift p	1	0.333333333
delecroix v	1	0.5
deligne p	2	1.333333333
delort j.-m.	1	1
demarco l	1	1
dembo a	1	0.25
denef j	2	1
deninger c	2	2
denisov s	1	0.333333333
derksen h	2	0.833333333
deroin b	1	0.5
devore r	2	0.833333333
diaconis p.	1	0.5
diamond f	2	0.583333333
dickenstein a	1	0.333333333
dinh tc	1	0.5

diperna r.j.	1	0.5
dodson b	1	1
dokchitser t	1	0.2
dokchitser v	1	0.2
dolgachev i	1	0.5
dolgopyat d	2	1
donaldson s	3	0.833333333
dong b	1	0.25
donnelly h.	1	0.5
donoho dl	1	0.5
dormy e	1	0.5
dosev d	1	0.333333333
dougherty r.	1	0.5
downey r	1	0.333333333
dritschel ma	1	0.5
dubedat j	2	2
dudas o	1	0.5
duits m	1	0.5
dunfield n	1	0.25
duong xt	1	0.5
dupont c	1	0.5
dvir z	1	1
dwyer w.g.	1	0.5

dyatlov s	1	1
ebenfelt p	1	0.333333333
edelman a.	1	0.333333333
efimov ai	1	0.2
ein l.	3	1.333333333
eisenbud d	2	0.833333333
ekholm t	1	0.5
eliashberg y.	1	1
ellingsrud g	1	0.5
ellis d	1	0.333333333
emerton m	1	1
enriquez b	1	0.5
epstein c.l.	1	0.5
erdelyi t	1	0.5
erdos l	1	0.333333333
eskin a	2	1.5
esnault h	2	0.75
etingof p	1	0.333333333
evans l.c.	1	0.5
eyssidieux p	1	0.333333333
faber c	1	0.2
faltings g	2	2
fan ck	1	1

faou e	1	0.333333333
farb b	1	0.5
farrell f.t.	4	1.833333333
fasel j	1	0.5
favre c	2	0.833333333
fefferman c	6	2.666666667
fefferman r.	1	1
feichtner em	1	0.333333333
feldman m	2	0.833333333
felouzis v	1	0.5
fenley sr	1	1
fiebig p	1	1
filaseta m	1	0.2
finkelberg m	1	0.5
fintushel r	2	1
fisher d	1	0.25
fisher y.	1	0.333333333
flicker yz	1	0.333333333
fomin s	3	1.333333333
ford kb	2	1.2
foreman m.	1	0.5
foulon p.	1	0.333333333
fox j	1	0.333333333



francis j	1	0.333333333
frank rl	2	0.583333333
frankl p.	1	0.5
frantzikinakis n	1	0.5
freed ds	1	0.333333333
freedman m	1	0.333333333
freedman mh	1	0.333333333
frenkel e	3	1.083333333
friedgut e	2	0.833333333
friedl s	1	0.5
friedlander em	2	0.666666667
friedlander j.	1	0.25
friedlander j.b.	1	0.333333333
friedman r	1	1
friedman s.d.	1	1
frohman c	1	0.5
frossard e	2	0.666666667
fuchs e	1	0.5
fukaya k.	1	0.333333333
furman a	2	1.25
gabai d	4	2.833333333
gaboriau d	1	0.5
gaitsgory d	2	0.583333333

galatius s	1	0.333333333
gan wt	1	0.5
ganapathy mk	1	0.333333333
ganchev ha	1	0.2
garban c	1	0.333333333
gardner rj	1	0.333333333
garoufalidis s	1	0.5
gautam s	1	0.5
gee t	2	0.666666667
geiss c	1	0.333333333
geisser t	2	1
gelbart s	1	0.5
gelfand s	1	0.333333333
geraghty d	1	0.333333333
gerard p.	1	0.5
gerardvaret d	1	0.5
germain p	1	0.333333333
getzler e	1	1
gibney a	1	0.333333333
gieseker d	1	0.5
gillet h.	1	0.333333333
gimenez o	1	0.5
ginzburg d	2	0.666666667

ginzburg v	4	2.166666667
giordano t	1	0.25
glasner e	1	0.5
goldstein m	1	0.5
goncharov ab	2	2
goodman j.e.	2	0.833333333
gordon c.m.a.	1	0.5
gorelik m	2	2
gottsche l	1	1
gouezel s	1	1
gouvea f.	1	0.5
gover ar	2	1.5
gowers w.t.	1	0.5
graber t	1	0.333333333
graham r.l.	1	0.5
granville a.	3	1.25
grayson dr	1	1
green b	1	0.5
green m.	1	0.5
gressman pt	1	0.5
grieser d	1	0.5
griffiths pa	1	0.5
grochenig k	1	0.5

grodal j	1	0.5
gromoll d.	1	0.5
gromov m.	2	1.333333333
gross bh	2	1.2
gross m	1	0.5
grossman p	1	0.5
grove k.	2	1
guedj v	1	0.333333333
gues o	1	0.25
guillarmou c	1	1
guillemin v.	1	1
gunnells pe	1	0.5
gunturk cs	1	1
guo y	1	1
guralnick rm	2	0.75
guth l	1	1
habegger n.	1	0.5
haberl c	1	0.5
hacon cd	3	1.25
haesemeyer c	1	0.333333333
hafner j.l.	1	0.5
haglund j	1	0.333333333
haiman m	4	2.833333333

hain r	2	2
hainzl c	1	0.25
halbout g	1	0.5
hales tc	1	0.5
halpernleistner d	1	1
hambleton i.	1	0.5
hamel f	1	0.333333333
hamilton r	1	0.5
hani z	1	0.333333333
hansen ac	1	1
harrington la	3	1.333333333
harris j	2	0.666666667
harris m	2	1.333333333
harvey r.	1	0.5
hass j	1	0.5
hatami h	1	0.5
hausel t	1	0.5
he x	1	0.5
hedenmalm h	1	0.5
hein hj	1	1
helfgott ha	1	0.5
henniart gm	3	1.083333333
hertz fr	1	0.5

herzig f	1	0.25
hesselholt l	2	1
hida h	3	3
hildebrand a.	1	0.25
hirachi k	1	0.5
hiraga k	2	0.666666667
hjorth g	1	1
hochster m.	1	0.5
hoffman c	1	0.333333333
hoffman d.	2	1.333333333
hofmann s	1	0.25
hongler c	1	0.5
hopkins mj	2	0.666666667
hosono s	1	0.333333333
host b	1	0.5
hovey m	1	0.333333333
howe r.	1	1
hrubes p	1	0.333333333
hrushovski e	4	2.833333333
hsieh ml	1	1
huang js	1	0.5
huang xj	1	1
hubbard j.	2	0.666666667

huh j	1	1
huneke c.	1	0.5
ichino a	2	0.666666667
idziak p	1	0.333333333
iii	1	0.333333333
ikeda t	2	0.666666667
imanuvilov oy	1	0.333333333
ioana a	1	1
ionescu ad	3	1.5
israel a	1	0.333333333
iwaniec t	2	0.583333333
iwanieca h.	1	0.333333333
iwao y	1	0.5
jackson s	1	0.5
jacquet h	1	0.333333333
jager t	1	0.5
jaikinzapirain a	1	1
james k	1	1
jech t.	1	0.5
jerison d	4	1.833333333
jiang d	1	0.333333333
johansson k	1	0.333333333
johnson wb	2	0.666666667

jones l.e.	3	1.5
jones p.w.	1	0.333333333
jones vfr	1	0.5
jonsson m	2	0.833333333
joseph a	1	0.333333333
jr.	2	0.666666667
judge cm	1	1
junge m	1	0.5
juteau d	1	0.333333333
kac v.g.	2	0.833333333
kachkovskiy i	1	0.5
kahle m	1	0.333333333
kahn b	1	0.25
kahn j	2	1.333333333
kaledin d	1	0.5
kalinin b	1	0.25
kang ng	1	1
kang sj	1	0.333333333
kantor wm	1	0.25
kaplan a	1	0.333333333
kapovich m	1	0.5
kapovitch v	1	0.5
kapranov mm	2	1.5



karagueuzian db	1	0.5
karu k	1	0.25
kashiwara m	2	0.8333333333
kassabov m	1	0.25
kassaei pl	1	1
katz n.m.	2	2
katz nh	1	0.5
katzarkov l	1	0.2
katzman m	1	0.5
kawakita m	1	1
kawamata y	1	1
kazarian me	1	0.5
kazhdan d.	5	2.25
kechris a.s.	3	1.3333333333
kedlaya ks	3	1.6666666667
keel m	2	0.5333333333
keel s	2	0.8333333333
keeler ds	1	1
kenig c	2	0.5833333333
kenig ce	5	1.9999999999
kenyon r	1	0.3333333333
kenyon rw	1	0.5
khanin k	1	0.3333333333

khovanov m	1	0.5
kiderlen m	1	0.333333333
kiem yh	1	0.5
kim hh	1	0.333333333
kim jl	1	1
kim k.h.	3	1.166666667
kim m	2	1.333333333
king a	1	0.333333333
kirillov jr aa	1	1
kirwan f.	2	2
kiselev a	2	1.5
kisin m	4	4
klainerman k	1	0.5
klainerman s	3	1.5
kleiner b	1	1
klemm a	1	0.25
klingenberg w.	1	0.5
knop f	2	2
knutson a	2	0.833333333
kollar j.	6	4.5
kolster m	1	0.333333333
komlos j	1	0.333333333
kontorovich a	2	0.666666667

konyagin s	1	0.2
kook w	1	0.333333333
korarov l	1	0.5
kostant b.	1	0.5
kostanta b	1	1
kostlan e.	1	0.333333333
kottwitz r.e.	1	1
kovacs sj	2	1
kovalev lv	1	0.333333333
kowalczyk m	1	0.333333333
kozma g	1	0.5
kra i.	1	1
krantz s.g.	1	0.5
kreck m.	1	0.5
kremnizer k	1	0.5
kresch a	1	0.333333333
krieger j	1	0.5
krivelevich m	1	1
kudla ss	1	0.333333333
kuhn nj	1	0.333333333
kulkarni r.	1	0.5
kumar a	1	0.5
kumar s	1	1

kutzko pc	1	0.333333333
kytola k	1	0.5
labourie e.f.	1	0.333333333
lackenby m	1	1
lafforgue l	1	1
lagarias jc	1	0.5
lam t	1	1
lamel b	1	0.5
landau h.j.	1	1
lando sk	1	0.5
landsberg jm	1	1
lannes d	1	1
lapid e	1	0.333333333
laredo vt	1	0.5
larsen mj	2	1
lau e	1	1
lauritzen n	1	0.333333333
lawler g	1	0.333333333
lazarsfeld r.	4	1.833333333
lebeau g	2	0.833333333
lebrun c.	3	1.833333333
leclerc b	1	0.333333333
lee j.m.	1	0.5

lee jr	1	0.333333333
lee sy	1	0.5
lee yp	1	0.5
leibman a	1	0.5
leinert m	1	0.5
lemire n	1	0.333333333
lemou m	1	0.333333333
lempert l	4	4
lempp s	1	0.2
lenagan th	1	0.5
lenstra h.w.	1	0.333333333
lessmann o	1	0.5
letzter g	1	0.333333333
leung nc	1	0.5
levasseur t	1	0.5
levenberg n	1	0.333333333
levi r	1	0.333333333
levine l	1	0.333333333
levine m	1	0.25
lewis jl	1	0.5
li c.	1	0.333333333
li hf	1	0.5
li j	3	1.5

li p	1	0.5
li t	2	2
lian bh	1	0.333333333
lian z	1	0.5
lieb eh	2	0.666666667
liebeck mw	1	0.5
liggett tm	2	0.666666667
lin f	2	0.833333333
lin x.-s.	1	0.5
lind d	1	0.5
lindenstrauss e	1	0.25
linnell p	1	0.5
lions p.l.	1	0.333333333
littelmann p	1	1
litvak ae	1	0.25
liu ccm	2	0.833333333
liu k	1	0.333333333
liu t	1	0.333333333
liu tp	1	0.5
liu y	1	0.5
liverani c	1	0.5
loehr n	1	0.333333333
loeser f	2	1

long d.-g.	1	1
long dd	1	0.333333333
losev i	1	1
lott j	1	0.5
louveau a.	2	0.833333333
lovasz l	2	0.833333333
lu j.-h.	1	0.5
lubetzky e	1	0.5
lubotzky a	4	2.25
luck wl	1	0.333333333
luczak t.	1	0.5
luecke j.	1	0.5
luli gk	1	0.333333333
lunts v	2	1
luszti g.	12	8.833333333
lutwak e	1	0.25
lyubarskii yi	1	0.5
lyubich m.	2	1
m<u+0153>glin c	1	0.5
machedon m	1	0.5
macri e	1	0.5
madsen i	1	0.333333333
magidor m.	1	0.5

maier h.	1	0.25
mairresse j	1	0.5
majda a.	1	0.5
makarov ng	1	0.5
malle g	1	0.5
mallet-paret j.	1	0.5
malliaris m	1	0.5
manolescu c	1	1
marcus b.	1	0.5
margalit d	1	0.333333333
markovic v	2	2
marks as	1	1
markvorsen s	1	0.5
marmi s	2	0.666666667
marques fc	1	0.333333333
marshall de	1	0.5
marshall s	2	2
martel y	2	0.833333333
martin d.a.	2	1
maslen dk	1	0.5
masuda m	1	0.5
masur h	1	0.5
mateu j	1	0.333333333



mather j.n.	1	1
mathieu o.	1	1
matignon m	1	0.5
matousek j	1	0.5
matsuki k	1	0.25
matsuo s	1	0.5
matui h	1	0.25
matusevich lf	1	0.333333333
mauldin rd	1	0.5
maulik d	2	0.75
maurey b	2	0.833333333
mautner c	1	0.333333333
mayboroda s	1	0.25
mazur b	2	0.833333333
mazzeo r	1	0.333333333
mccann rj	1	0.333333333
mccarthy j.e.	1	1
mcclure je	1	0.5
mccord c	1	0.5
mccullough s	1	0.5
mccurley k.s.	1	0.5
mcduff d.	2	1.5
mcintosh a.	1	0.333333333

mckenzie r	1	0.333333333
mckernan j	2	0.75
mclaughlin s	1	0.5
mcmullen ct	3	3
mcnamara pj	1	0.333333333
meeks w.h.	3	1.333333333
mehats f	1	0.333333333
meinrenken e	1	1
meiri c	1	0.5
meise r	1	0.333333333
merkurjev as	2	1.5
merle f	3	2
mess g.	1	0.5
metivier g	1	0.25
meyerhoff r	1	0.333333333
mikhalkin g	1	1
miller e	1	0.333333333
miller js	1	0.2
miller sd	1	0.5
milley p	1	0.333333333
milman p.d.	1	0.5
milne js	1	0.5
milnor j.	1	0.5

ming p	1	0.333333333
minicozzi wp	2	1.5
minsky y.n.	1	1
mir n	1	0.5
mirzakhani m	1	1
mischaikow k	1	0.5
mitter s	1	0.333333333
mockenhaupt g.	1	0.333333333
mohammadi a	1	0.5
moller jm	1	0.333333333
moller m	1	1
monod n	1	1
moore jt	1	1
morel s	1	1
mori s.	3	2.5
moriwaki a	1	1
morris r	1	0.333333333
morrison d.r.	1	1
morrison i	1	0.333333333
moscovici h	1	0.5
mostow g.d.	1	1
moussa p	2	0.666666667
moy a.	2	1

mozes s	2	0.75
mukhin e	1	0.333333333
muller s.	1	0.5
muller w.	1	1
mulmuley kd	1	1
munoz c	1	0.333333333
munshi r	1	1
muscalu c	1	0.333333333
mustata m	1	1
nachmias a	1	0.5
nadel a.m.	1	1
nadirashvili n	2	0.833333333
nadler d	2	0.833333333
nakajima h	1	1
naor a	2	0.583333333
napier t	1	0.5
narayanan h	1	0.333333333
nazarov f	2	0.666666667
neeman a	1	1
nelson pd	1	0.333333333
neves a	1	0.333333333
nevins t	1	0.5
nevo a.	1	1

ng sh	1	0.25
ni l	1	1
nitsure n.	1	1
nolin p	1	0.5
norine s	1	0.5
noy m	1	0.5
nystrom k	1	0.5
oblomkov a	1	0.5
odell e	1	0.5
oh h	4	1.833333333
okikiolu k	1	1
okounkov a	4	1.233333333
oliver b	2	0.666666667
olla s	1	0.5
olshanski g	1	0.333333333
onninen j	1	0.333333333
ono k	1	0.333333333
oort f	2	1.5
orlik s	1	0.5
orlov do	2	0.7
ormes ns	1	0.333333333
osher s	1	0.25
otto f	1	0.333333333

ozsvath p	1	0.5
pach j	1	0.5
pajor a	1	0.25
pak i	1	0.5
pandharipande r	7	3.116666667
pandzic p	1	0.5
papanikolas ma	1	0.333333333
parapatits l	1	0.5
pardo lm	1	0.5
pardon j	1	1
pardon w.l.	1	0.5
pareschi g	2	1.5
parusinski a	1	0.5
parzanchevski o	1	0.5
patakfalvi z	1	0.5
paul a	1	0.2
pedersen e.k.	1	0.5
peeva i	1	0.5
peller vv	1	1
perelman g.	1	1
peres y	2	0.75
perrinriou b	1	1
perthame b.	1	0.333333333

pete g	1	0.333333333
petersen p.	1	0.5
peterzil y	1	0.333333333
petrie t	1	0.5
petrosyan a	1	0.333333333
petrova g	1	0.5
phillips r.	1	0.5
piatetskishapiro ii	1	0.5
pillay a	3	1.166666667
pilpel h	1	0.333333333
pink r	1	0.5
pipher j	1	0.25
pisier g	1	1
pitale a	1	0.333333333
pitt nje	1	1
pixton a	2	0.833333333
planchon f	1	0.333333333
poletsky ea	1	0.333333333
pollack d	1	0.333333333
pollack r.	3	1.166666667
pollen d.	1	1
poltoratski a	1	1
pomerance c.	2	0.533333333

pommersheim j	3	2
ponce g.	2	0.666666667
poon y.-s.	1	0.5
poonen b	6	3.75
popa m	1	0.5
popa s	3	2
popov vl	1	0.333333333
postnikov a	1	0.333333333
pottharst j	1	0.333333333
pragacz p	1	0.5
prause i	1	0.25
preiss d	1	0.333333333
procesi c.	2	0.666666667
propp j	1	0.333333333
puder d	1	0.5
putnam if	1	0.25
pyber l	1	0.5
qing j	1	0.5
quastel j	1	0.333333333
quillen d	2	1
quinn f.	1	1
quint jf	1	0.5
rabenpedersen u	1	0.333333333



rabinowitz p.h.	1	0.5
radulescu f.	1	1
rains e	1	0.5
rallis s	2	0.666666667
ramachandran m	1	0.5
ramachandran n	1	0.5
ramakrishnan d	1	0.333333333
ramirez ja	1	0.333333333
raphael p	3	1.333333333
rapinchuk as	1	0.333333333
ratiu t.	1	0.5
raugel g.	1	0.5
ravenel dc	1	0.333333333
reeder m	2	1.5
reich h	1	0.5
reichstein z	1	0.333333333
reid aw	1	0.333333333
reid m	1	0.333333333
reiner v	1	0.333333333
reiten i	1	0.5
reshetikhin n	1	0.5
rezk c	1	1
reznikov a	1	1

rhoades rc	1	0.333333333
richthammer t	1	0.333333333
rider b	1	0.333333333
rietsch k	2	2
ritter j	2	1
roberts r	1	0.333333333
robin graham c	1	0.5
rockmore dn	2	1
rodl v	2	1
rodnianski i	1	0.5
rodnianski r	1	0.5
rogawski j	1	0.333333333
rognés j	1	0.333333333
rohde s	1	0.5
rohlfs j.	1	0.5
rolin jp	1	0.333333333
ronyai l	1	0.333333333
ros a	1	1
rosenthal h.	1	1
rotger v	1	0.5
rothschild lp	2	0.833333333
rouquier r	1	0.5
roush fw	3	1.666666667

rowell ec	1	0.25
roy mf	1	0.333333333
rubinstein ya	1	0.5
rucinski a	1	0.5
rudelson m	1	0.5
runde v	1	1
ryabogin d	1	0.333333333
sadun l	1	0.5
safarov y	1	0.5
saha a	1	0.333333333
saito t	1	0.5
saksman e	1	0.25
salehi gelsefidy a	1	0.5
sam sv	1	0.5
samotij w	1	0.333333333
santos f	1	1
sapir m	1	1
sapiro g	1	0.333333333
sarig om	1	1
sarnak p	3	1.333333333
saut j.c.	1	0.5
savin o	1	1
savitt d	1	0.333333333

scanlon t	1	1
scannell kp	1	0.5
schapira p	1	0.5
schechtman g	2	0.666666667
shedler t	1	0.333333333
scheidegger e	1	0.25
scheiderer c	1	0.333333333
schick t	1	0.5
schiffmann o	1	0.333333333
schilling a	1	0.333333333
schlag w	2	1.5
schleicher d	1	0.333333333
schleimer s	1	0.5
schlein b	1	0.333333333
schlumprecht th	1	0.5
schmid w	2	1
schmidt b	1	1
schmidt k	1	0.5
schneider p	1	0.5
schoen c	1	1
schoen r	1	0.5
scholze p	2	1.5
schonbek m.e.	1	1

schramm o	5	2.416666667
schreyer fo	2	0.833333333
schrijver a	1	0.333333333
schroer j	1	0.333333333
schwede s	1	1
schwermer j.	1	0.5
seeger a.	3	2.333333333
segev y	3	1.833333333
seidel p	1	0.5
seip k	1	0.5
seiringer r	2	0.583333333
seitz gm	1	0.333333333
sell g.r.	2	1
semmes s.	2	0.666666667
seppalainen t	1	0.333333333
serfaty s	1	1
serre jp	1	1
seymour p.	1	0.333333333
shah n	1	0.333333333
shah na	2	1.5
shahgholian h	1	0.333333333
shahidi f	1	0.5
shalen p.b.	1	0.5

shalev a	2	1
shalika j	1	0.333333333
shan p	1	0.5
shaneson j.l.	1	0.5
shao qm	1	0.25
shapiro jh	1	0.5
shareshian j	1	0.333333333
sheffield s	2	0.583333333
shelah s.	6	4
shen z	2	0.583333333
shepherd barron ni	1	0.5
shestakov ip	2	1
shiffman b	1	0.5
shin sw	1	0.5
shiple y b	1	0.333333333
shishikura m	1	0.333333333
shkoller s	1	0.5
shub m.	2	0.833333333
shvydkoy r	1	1
sibony n	1	0.5
siebert b	1	0.5
sigal i.m.	1	0.5
simpson c.t.	1	1

simpsonl c.t.	1	1
sinha dp	1	1
sjamaar r	1	0.5
sjostrand j.	1	0.5
skau cf	1	0.25
skora rk	1	1
slaman ta	1	0.333333333
sleator d.d.	1	0.333333333
sly a	1	0.5
smale s.	1	0.5
smania d	1	1
smillie j.	1	0.5
smith gg	2	0.666666667
smith hf	2	0.833333333
smith i	1	0.5
smith j	1	0.333333333
smith jh	1	0.5
smoktunowicz a	1	0.5
snowden a	1	0.5
soare ri	1	0.5
soergel w	2	1.333333333
soffer a.	1	0.5
sogge cd	5	3.166666667

solecki s	1	1
solomon j	1	0.333333333
solovej jp	1	0.25
song j	1	0.5
soskova mi	1	0.2
sottile f	1	1
soudry d	1	0.333333333
souganidis pe	1	0.333333333
soule c.	1	0.333333333
soundararajan k	2	1
spatzier r	1	0.25
speissegger p	1	0.333333333
spencer j.	3	1.5
spivakovsky m	1	1
staffilani g	1	0.2
stafford jt	1	0.5
stanley r.p.	2	2
stanton d	1	0.333333333
starr j	1	0.333333333
steel j.r.	2	1
stefanov f	1	0.5
stefanov p	1	0.333333333
stein m	1	0.333333333



sterbenz j	1	0.5
stern m.a.	1	0.5
stern rj	2	1
stewart c.l.	2	1.5
stolz s.	1	0.5
storm pa	1	0.25
strain rm	1	0.5
strauch m	1	0.5
stromme sa	1	0.5
sturmfels b	3	1.166666667
sudakov b	1	0.333333333
sujatha r	1	0.333333333
suk a	1	1
sun b	2	1.5
sun s	4	1.083333333
sundberg c	1	1
suresh v	3	1
suslin a	2	0.666666667
swartz e	1	1
sweet wj	1	0.333333333
symonds p	2	1.5
szabo e	1	0.5
szabo z	1	0.5

szeftel j	1	0.5
szegedy b	1	1
szekelyhidi g	1	1
szemeredi e	2	0.833333333
szenes a	1	0.5
tadic m	1	0.5
tadmor e.	1	0.333333333
taguchi y	1	0.5
takaoka h	1	0.2
takeda s	1	0.5
takloobighash r	1	0.333333333
talagrand m.	2	2
tamvakis h	1	0.333333333
tannenbaum a	1	0.333333333
tanner j	1	0.5
tao t	9	4.2
tarasov v	1	0.333333333
tardos g	1	0.5
tarjan r.e.	1	0.333333333
tataru d	2	1.5
taubes ch	2	1.5
taylor ba	1	0.333333333
taylor r	4	1.583333333

tcheremchantsev s	1	0.5
teitelbaum j.t.	2	1.5
teleman c	1	0.333333333
terasoma t	1	0.5
thaddeus m	2	1.5
thiele c	1	0.333333333
thom a	1	0.5
thomas h	1	0.5
thomas r.	2	0.833333333
thomas s	1	1
thomassen c.	1	1
thomsen jf	1	0.333333333
thorne ja	1	1
thurston w.p.	2	0.583333333
tian g.	9	4.5
tillmann u	1	0.333333333
toda y	1	1
todorcevic s	1	1
tolsa x	1	0.333333333
tomczakjaegermann n	1	0.25
top j	1	0.5
toro t	1	0.333333333
tosatti v	2	1

totaro b	2	2
trapa p	1	0.2
treil s	2	0.8333333333
treves f	1	0.5
trudinger ns	1	0.5
tschinkel y	1	0.3333333333
tsimerman j	1	1
tsukamoto m	1	0.5
tulyakov d	1	0.3333333333
tuncel s.	1	0.5
uhlenbeck k	1	0.3333333333
uhlmann g	3	1.166666667
umirbaev uu	2	1
vaes s	1	0.5
valeriete m	1	0.3333333333
vallentin f	1	0.5
van den bergh m.	2	1.5
van geemen b	1	0.5
van strien s	3	1.5
van vu v	1	0.5
varchenko a	1	0.3333333333
vargas a	1	0.3333333333
vargas e	2	1

vasserot e	1	0.5
vasseur a	1	0.333333333
vasy a	1	0.333333333
vega l.	3	1
velasco m	1	0.333333333
venkatesh a	1	0.5
verdera j	1	0.333333333
vergne m	2	1
vershynin r	1	0.5
viale m	1	1
vidussi s	1	0.5
viehweg e	1	0.25
vigneras mf	1	0.25
villadsen j	1	1
vilonen k	3	1.083333333
virag b	2	0.833333333
vistoli a	1	0.5
viterbo c	1	1
vogan jr da	1	0.2
vogt d	1	0.333333333
vojta p.	1	1
volberg a	2	1.333333333
vu v	1	0.5

wagner f	1	0.333333333
wagoner j.b.	1	0.333333333
wahl j.	1	1
wainger s	2	0.833333333
walcher j	1	0.333333333
walker a	1	0.5
walker k	1	0.333333333
wallach n.r.	1	1
walther u	1	0.333333333
wan d	3	2.5
wang b	1	0.5
wang l	1	1
wang x	1	0.2
wang xj	1	0.5
wang yg	1	0.25
wang z	1	0.25
warnaar so	1	0.333333333
weber b	1	0.333333333
weibel c	2	0.666666667
weinan e	1	0.333333333
weinberger s.	2	0.833333333
weinkove b	2	1
weinstein a.	1	0.5

weinstein j	1	0.5
weinstein mi	1	0.5
weiss a	2	1
weiss b	1	0.5
wentworth r	1	0.333333333
wenzl h	1	1
werner w	3	1.833333333
wewers s	1	1
weyman j	3	1.166666667
white b	3	3
wick bd	1	0.5
wigderson a	1	0.333333333
wilf h.s.	1	0.5
wilkerson c.w.	1	0.5
wilkie aj	2	1.333333333
williams m	1	0.25
williamson g	2	0.666666667
wilson db	2	0.75
winter d	1	0.5
wittenberg o	1	0.5
wlodarczyk j	2	1.25
wolf m	2	1
wong yl	1	0.25

woo j	1	0.5
wood mm	1	1
woods k	1	0.5
woodward c	1	0.333333333
wooley t.d.	1	1
wright j	3	1.166666667
wu j	1	0.25
wu s	1	1
xi n	1	1
xiao l	1	0.333333333
xu c	1	0.5
xu cj	1	0.25
xu j	1	0.5
xu q	1	0.5
yamaki k	1	1
yamamoto m	1	0.333333333
yampolsky m	2	1
yan l	1	0.5
yang d	1	0.25
yang p.c.p.	1	0.5
yang t	2	0.75
yang y	1	0.333333333
yau ht	1	0.333333333



yau st	3	1.333333333
yehudayoff a	1	0.333333333
yoccoz jc	2	0.666666667
yoshida t	1	0.5
young ls	1	0.5
young mp	1	1
young nj	1	0.5
youssin	1	1
yu g	1	0.2
yu jk	2	1.5
yu sh	1	1
zakeri s	1	0.5
zaslow e	1	0.5
zeilberger d.	1	0.5
zeitouni o	1	0.25
zelati v.c.	1	0.5
zelditch s	1	0.5
zelevinsky a	3	1.333333333
zelikson s	1	0.333333333
zeriahi a	1	0.333333333
zhang g	1	0.25
zhang h	1	0.5
zhang p	1	0.333333333

zhang w	1	1
zhang x	1	0.5
zhanga s	1	1
zhou j	1	0.333333333
zhu cb	1	0.5
zhu x	1	0.5
zhu y	1	1
ziegler g.m.	1	0.5
ziegler t	1	1
zikatanov l	1	0.5
zilber b	1	0.5
zimmer r.j.	1	1
zimmere r.j.	1	1
zimmers r.j.	1	1
zinger a	1	1
zlatos a	1	1
zumbrun k	1	0.25
zvavitch a	1	0.333333333
zvonkine d	1	0.333333333
zworski m.	2	1

*Table 4-4 The number of papers the authors have written and their fractional count*

The graph illustrated the top 10 productive authors of Journal. Author Lutzig.G holds the first position with 12 in journal while Tao t has the second position with 9.

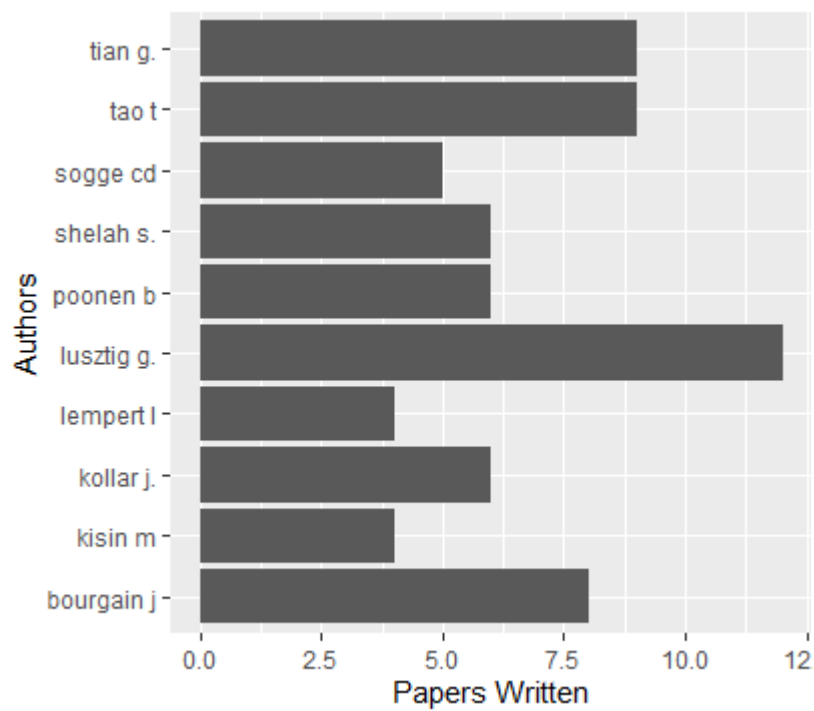


Figure 4.2 Top Authors by the papers they have written

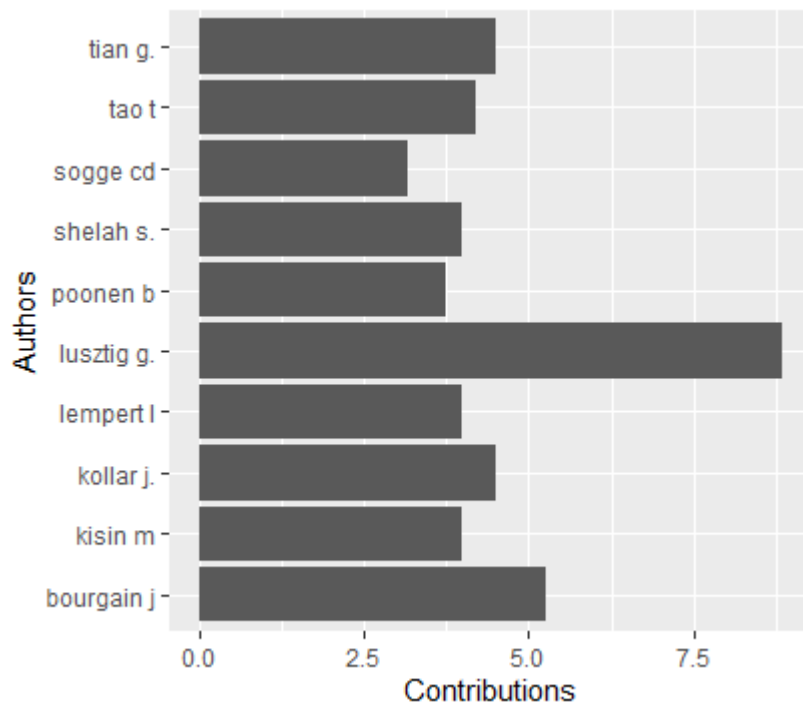


Figure 4.3 Top Authors by their fractional count

The graph illustrated the top 10 authors according to their fractional counts. Author Lutzig.G has the first position with 8.83 in journal while Bourgain J has the second position with 5.25.

## 5 Journal's Productivity

The journal of the American Mathematical Society initiated its publications in 1988. In the time 1988-2017, the journal has published 30 volumes and 120 issues. A volume of a journal is defined as the number of years the publications has been circulated whereas issues as how many times that journal has been published during the year. In this chapter all calculations regarding the publication's characteristics of the journal of the American Mathematical Society, are included. More specifically, the publications of published documents by volume and 5-volume, the number of their pages also are analyzed. We also analyzed the contribute between authors and the number of the papers and the pages they write. In the end the length of the papers is examined.

The following table shows the number of documents that have been written and exist in each Volume as well as the percentage that each Volume has in relation to the total number of documents that exist in the journal.

Volume	Year	Number of papers	%
1	1988	25	2.73
2	1989	28	3,05
3	1990	31	3.38
4	1991	31	3.38
5	1992	29	3.16
6	1993	25	2.73
7	1994	26	2.83
8	1995	28	3.05
9	1996	35	3.82

10	1997	33	3.59
11	1998	32	3.49
12	1999	34	3.72
13	2000	34	3.72
14	2001	26	2.83
15	2002	27	2.94
16	2003	35	3.82
17	2004	30	3.27
18	2005	28	3.05
19	2006	29	3.16
20	2007	36	3.91
21	2008	41	4.47
22	2009	34	3.72
23	2010	35	3.82
24	2011	34	3.72
25	2012	35	3.82
26	2013	29	3.16
27	2014	26	2.83
28	2015	26	2.83
29	2016	30	3.27
30	2017	25	2.73
Total		917	100

*Table 5-1 The number of papers per Volume*

This table shows the number of Volumes per 5 years and respectively the percentage of recent registrations per 5 years. Here we can see that the most productive title is 2008-2012 with a percentage of 19.52% and the total number of documents 179.

Volume	Year	Number of papers	%
1-5	1988-1992	144	15.7
6-10	1993-1997	147	16.03
11-15	1998-2002	153	16.69
16-20	2003-2007	158	17.23
21-25	2008-2012	179	19.52
26-30	2013-2017	136	14.83
Total		917	100

*Table 5-2 The number of papers per 5-Volume*

Below we see in detail the number of pages per Volume from the first year of publication of the magazine and until the year we study the total number of 31820 pages

has been written, with the year with the most pages is 2011 with Volume 24 with 1209 pages.

Year	Volume	Number of pages	%
1988	1	938	2.95
1989	2	793	2.49
1990	3	929	2.92
1991	4	814	2.55
1992	5	958	3.01
1993	6	972	3.05
1994	7	986	3.11
1995	8	1003	3.15
1996	9	1192	3.73
1997	10	985	3.09
1998	11	1023	3.22
1999	12	1202	3.78
2000	13	994	3.12
2001	14	995	3.13
2002	15	996	3.14
2003	16	1014	3.18
2004	17	980	3.08
2005	18	991	3.11
2006	19	1003	3.15
2007	20	1165	3.68



2008	21	1198	3.78
2009	22	1197	3.77
2010	23	1176	3.69
2011	24	1209	3.79
2012	25	1198	3.78
2013	26	1193	3.74
2014	27	1144	3.59
2015	28	1196	3.76
2016	29	1194	3.75
2017	30	1182	3.71
	Total	31820	100

*Table 5-3 The number of pages per Volume*

In the table we calculate the percentage of research articles by 5-volume. The research articles into 5-volumes periods are divided. As before, the five years with the most written is in 2008-2012 with a total number of pages of 5978 and a percentage of 18.81%.

Volume	Year	Number of pages	%
1-5	1988-1992	4432	13.92
6-10	1993-1997	5138	16.13
11-15	1998-2002	5210	16.39
16-20	2003-2007	5153	16.2

21-25	2008-2012	5978	18.81
26-30	2013-2017	5909	18.55
Total		31820	100

*Table 5-4 The number of pages per 5-Volume*

Volume	Year	Mean	Median
1-5	1988-1992	31.02	26.9
6-10	1993-1997	35.30	30.2
11-15	1998-2002	34.34	30.9
16-20	2003-2007	32.79	28.6
21-25	2008-2012	33.56	30.3
26-30	2013-2017	43.99	40.6

*Table 5-5 Mean and median of the number of pages with the volume*

In the table we calculate the percentage of research articles by 5-volume. The research articles into 5-volumes periods are divided. Here we can see that the biggest mean and median have the time period between 2013-2017.

## Boxplot volume per Number of Pages

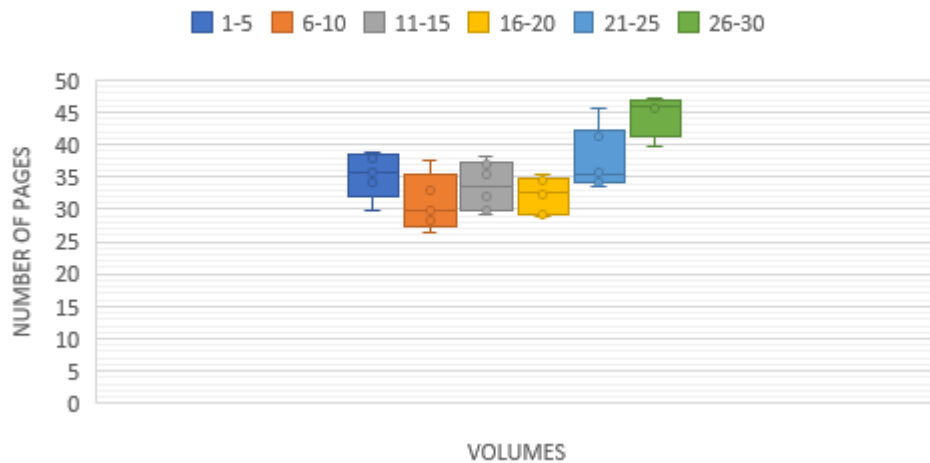


Figure 5.1 Boxplots per 5-Volumes for the Number of papers

Figure(5.1) illustrates the five number summary of each quartile. As it seems 6-10 and 21-25 quartiles have the biggest interquartile range which means that varied the set of observed values. Also, it is noticed that  $\text{median} < \text{mean}$  in all quartiles, so the data are right-skewed.

## 6 References

- [1] Otlet, P. (1934), *Traité De Documentation: Le Livre Sur Le Livre, Théorie Et Pratique*, Editions Mundaneum: Mons, Belgium.
- [2] Rousseau, Ronald (2014), "Library Science: Forgotten Founder of Bibliometrics.", *Nature*, 510 (7504): 218.
- [3] Pritchard, A. (1969) 'Statistical bibliography or bibliometrics?', *Journal of Documentation*, 24: 348–9.