University of Aegean

Technical School

Department of Financial Management & Engineering

DIPLOMA THESIS

Bibliometric Analysis of the Editorial Board of the Annals of Mathematics for the period 1874-2020.

Edited by:

Kolovos Apostolos

Supervisors:

Dr. Vasileiou Evangelos

Dr. Drivaliaris Demosthenes

2021

ACKNOWLEDGMENT

To those who were by my side during my successes and failures.

ABSTRACT

The purpose of this thesis is the collection, processing, and analysis of the data of the editorial team of the Annals of Mathematics for the whole period of its release so far, 1874-2020. The data collection will be performed using the primary information provided by the issues of the journal that is accessible in electronic form and secondary information concerning demographic features of the editorial team members and will be transformed into an editable format for the analysis. The above data will be evaluated qualitatively, detecting any inconsistencies and omissions within their values, and a set of rules and algorithms will be developed for their qualitative recast. In addition, the result above process will be used to consolidate the information from different sources, developing additional rules and algorithms for either treatment of inconsistencies between different sources or exploiting noncommon information. The data will then be quantified using statistics and econometric methods aimed at investigating and identifying to investigate and identify patterns and trends in the period under study. More specifically, they will be formulated and studied questions about the size of the editorial team, its lifespan, and changes made to it.

Keywords: Bibliometrics, Annals of Mathematics

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

CHAPTER 2

DATA

CHAPTER 3

EDITORIAL BOARD ROLES ACROSS TIME

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

The term *statistical bibliography* seems to have been first used by E. Wyndham Hulme in 1922 when he delivered two lectures as the Sandars Reader in Bibliography at the University of Cambridge. Subsequently, the lectures were published as a book. Although the debt has never been explicitly recognized through citations, Hulme anticipated modern work on the history of science. He used the term to mean the illumination of the processes of science and technology through counting documents. Hulme both summarized the results of Cole and Eales and produced original work on the growth of UK patents (relating these to social processes in the UK) and on the changes displayed in the *International Catalogue of Scientific Literature* (relating changes in subject and country production of literature to international developments).

The term then seems to have been ignored for twenty-two years until Gosnell used it in a paper on the obsolescence of literature without acknowledging its previous use. Presumably, the term was used in the original thesis.

The next use of the term was in 1962 by Raisig in a critical essay on citation studies. Since that time the term does not seem to have been used except by the present writer.

The definition and purpose of statistical bibliography have been variously stated as:

- 'to shed light on the processes of written communication and of the nature and course of development of a discipline (in so far as this is displayed through written communication), through counting and analyzing the various facets of written communication' (a slightly different version appears in Pritchard[6]);

- 'the assembling and interpretation of statistics relating to books and periodicals … to demonstrate historical movements, to determine the national or universal research use of books and journals, and to ascertain in many local situations the general use of books and journals.

In their way these statements of the definition and use of statistical bibliography are adequate and the second describes well the main uses to which the statistics (usually collected through citation studies, abstracts journal counts, or usage studies) have been put.

The present writer has never found the terms *statistical bibliography* at all satisfactory, and, to judge from discussions with many other workers in the field, this feeling is general. A measure of the unsatisfactory nature of the term is the fact that it has been used only four times in forty-six years. The term is clumsy, not very descriptive, and can be confused with statistics itself or bibliographies on statistics. This latter point was made by M. G. Kendall upon receiving a copy of my paper[6] and he suggested that the name of the subject be changed. This provided the final impetus for this paper.

Therefore, it is suggested that a better name for this subject (as previously defined) is BIBLIOMETRICS, i.e., the application of mathematics and statistical methods to books and other media of communication. An intensive search of the literature has failed to reveal any previous use of this term and an approach to the OED again failed to find that the term has been used before.

This term is that, whilst this particular combination is a neologism and therefore to be treated with a certain amount of suspicion, it has very close links to the accepted, and analogous 'biometrics', 'econometrics', and 'scientometrics'. The latter term is a Russian one for the application of quantitative methods to the history of science and overlaps with bibliometrics to a considerable extent. *En passant*, it is greatly to be regretted that the very logical Russian term for studies of all types on the processes of science Scientology has such unfortunate connotations in the West.

In conclusion, it is to be hoped that this term BIBLIOMETRICS will be used explicitly in all studies which seek to quantify the processes of written communication and will quickly gain acceptance in the field of information science. *(Pritchard, A)*

The dataset for this analysis consists of the editorial boards' information of the Annals of Mathematics, published by the Mathematics Department of Princeton University since March 1884, with the homonymous title. In dataset was considered and the data of The Analyst which was the previous name published from 1874 to 1883.

In the examined dataset there are three editions of publishing, where the numbering is repeatable:

1874-1883, as The Analyst where volumes range from 1 to 10.

1884-1899, as The Annals of Mathematics volumes, ranging from 1 to 12.

1899-2020, as The Annals of Mathematics volumes, ranging from 1 to 192.

1.1 DATA COLLECTION PROCESS

Initially, all the requested information must be aggregated into datasets to be analyzed. The source from which information will be collected is JSTOR. In figure 1.1 is displayed the structure of libraries of JSTOR where the volumes are sorted by decade were published. Every decade must be collected all the issues year by year to start structuring the dataset of volumes. The information collected is in volume information, front matter, back matter, and index of each issue.

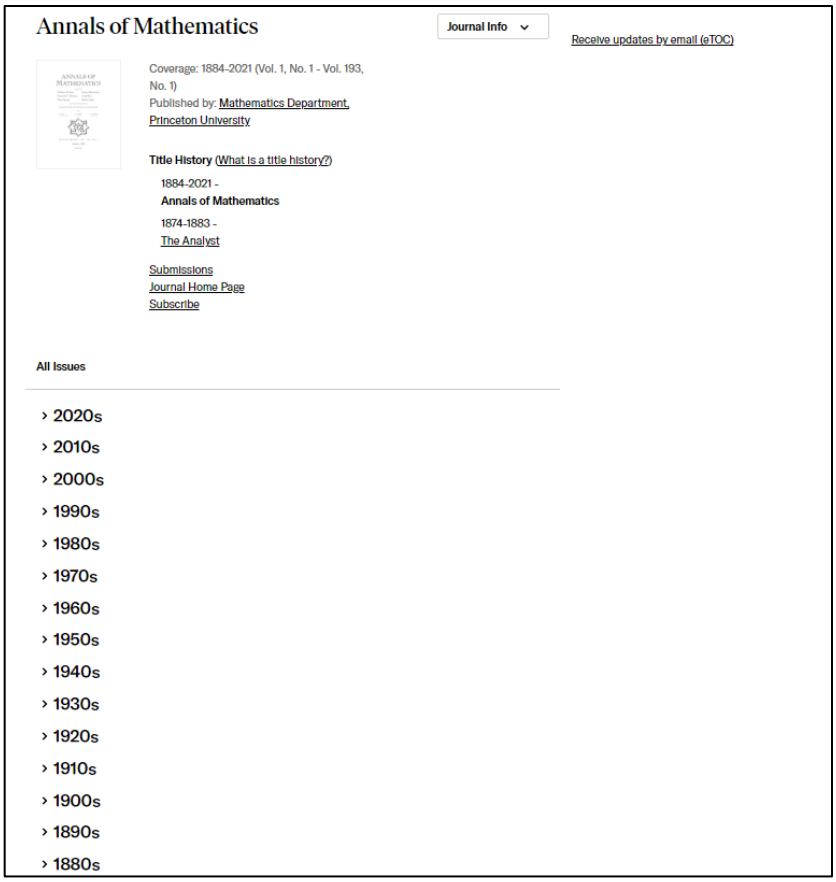


Figure 1.1 JSTOR page of Annals of Mathematics

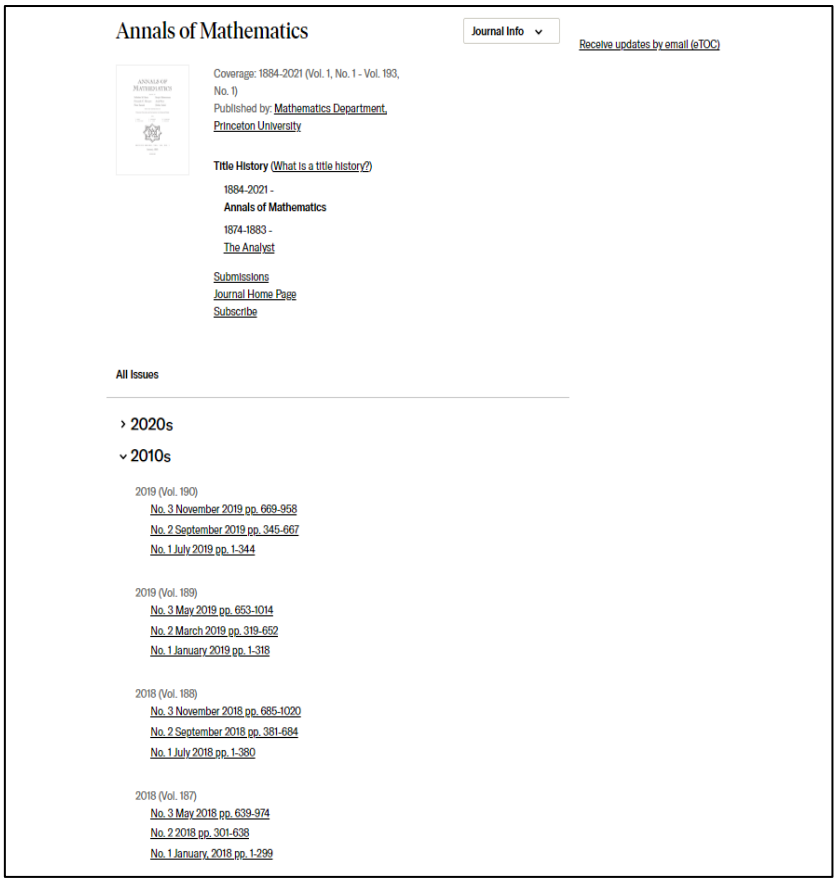


Figure 1.2 Volumes and Issues of Annals of Mathematics

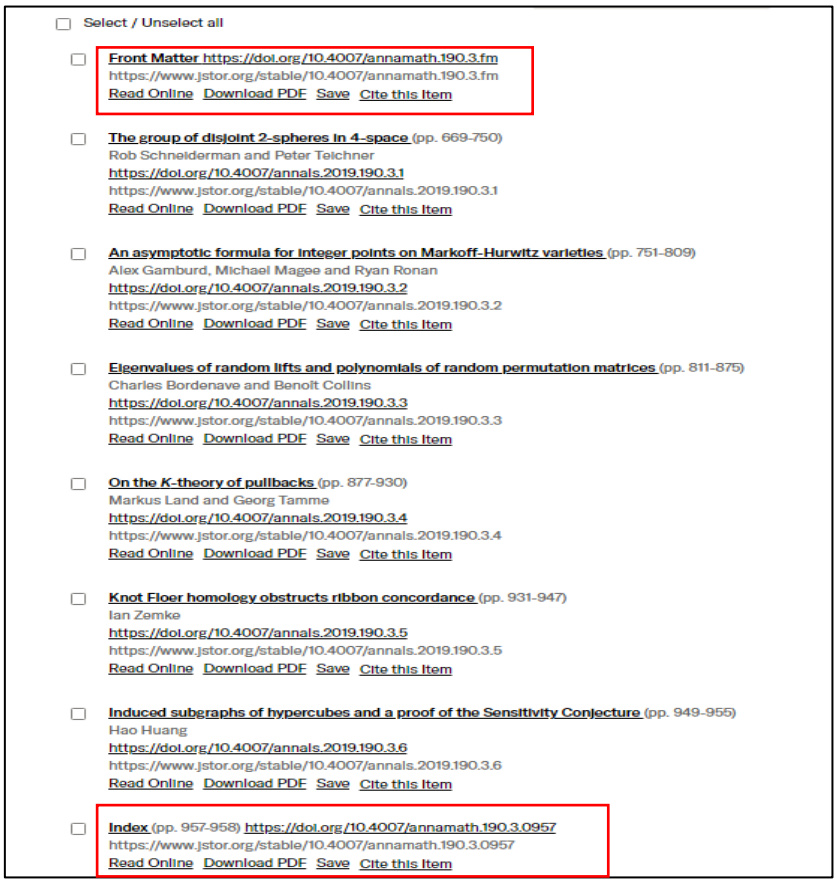


Figure 1.3 Source of information for issue

For each issue, the front matter, index, volume information, and back matter from the print issue can be downloaded in electronic format *(Figure 1.3)*.

The downloaded files on JSTOR, have the following format:

- *Year* (**Vol.** *Number of Volume*)

  No. *Issue Month Year* **pp.** *Including pages*

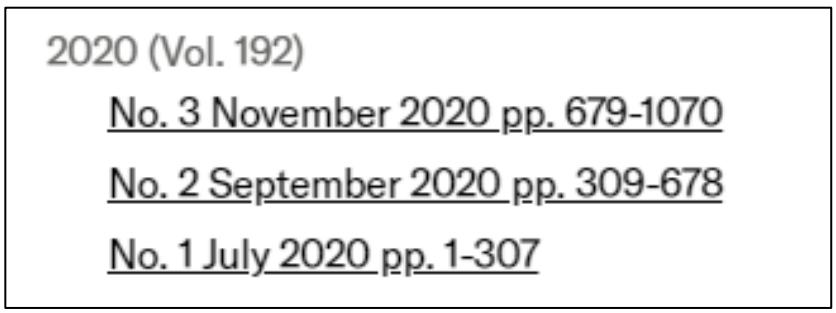


Figure 1.4 JSTOR format per Year, Issue.

The files that were downloaded from JSTOR were named according to by:

- Volume_issue number_category of info (except front matter, which is null)

  (e.g. 192_03 is referring to the 192nd volume of the 3rd issue and the info is from the front matter).

  (e.g. 192_03_INDEX is referring to the 192nd volume of the 3rd issue and the info is from the index).

  (e.g. 142_03_VI is referring to the 142nd volume of the 3rd issue and the info is from volume information).

As a result, are created tables were gathered all the information which can collect from every issue.

1.2 COVERS AND ACQUIRING DATA

Starting with the front matter, on tab 1 are the main editors of the issue. On the center of the page where is tab 2, there are the sections of editors and associate editors for each issue consisting of names of the editorial board in the form of Name Surname. Finally, down below is the number of volumes, issue, month, and year of issue on tab 3.

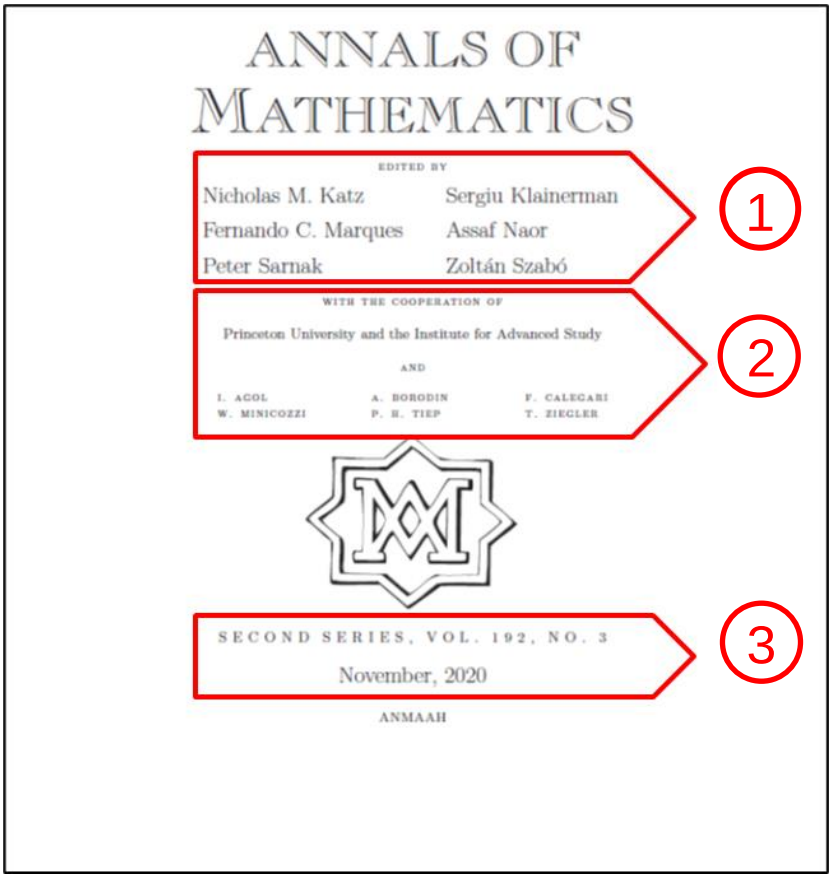


Figure 1.5 Front Matter

Below is a sample of the index that JSTOR has uploaded where can be seen all the topics and editors of the volume. The topics are sorted by all the names of the editors that were participated in the volume. Also, it is included and the back matters which are similar with the index, with the differences that referred to the topics of the issue (Figure 1.7).

ANNALS OF MATHEMATICS
Vol. 192, Nos. 1, 2, 3
(2020)

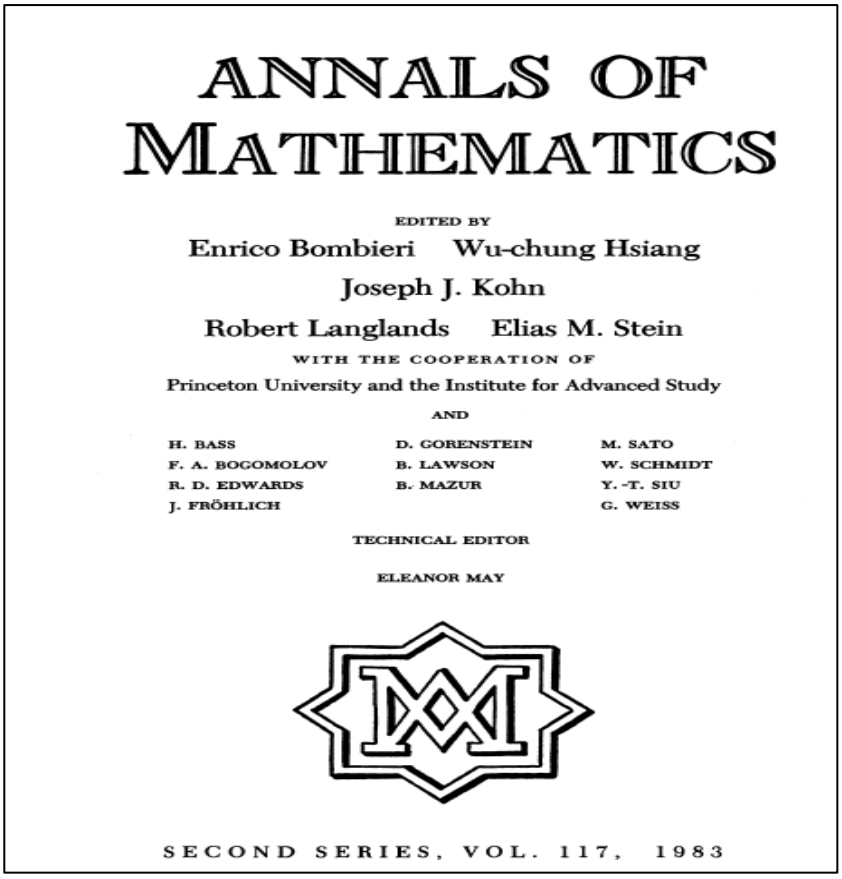
INDEX

Figure 1.6 Index

TABLE OF CONTENTS

Figure 1.7 Back Matter

In addition, must be referred one last category of information. That is the editorial boards, it is similar with front matters as format, but there is not the number of issues. The reason is observed this phenomenon is because those volumes have the same editorial board for all their issues.



ANNALS OF MATHEMATICS

EDITED BY

Enrico Bombieri Wu-chung Hsiang

Joseph J. Kohn

Robert Langlands Elias M. Stein

WITH THE COOPERATION OF

Princeton University and the Institute for Advanced Study

AND

H. BASS
F. A. BOGOMOLOV
R. D. EDWARDS
J. FRÖHLICH
D. GORENSTEIN
B. LAWSON
B. MAZUR
M. SATO
W. SCHMIDT
Y.-T. SIU
G. WEISS

TECHNICAL EDITOR

ELEANOR MAY

SECOND SERIES, VOL. 117, 1983

Figure 1.8 Editorial Board

1.3 DATA LOGGING ON TABLE

Therefore, the following process is to archive information that is collected from source to specific data frames, with the purpose to concentrate details and be trouble-free data to be processed. After the gathering of issues will be created three data frames.

- Information for the editors and the editorial positions that possess in issues.
- Information from the sources of data (Editorial boards, indexes, front matters, back matter).
- The spot of information inside sources.

Those three tables were created to be able to compare the information of each table between them and to be able to come to conclusions about the reliability of the data that were gathered. The raw data are rarely in the correct format, are without errors, are complete, and have all the correct labels and codes that were needed for analysis. For that reason, the process after the collection of data is cleaning.

1.4 DATA CLEANING

Data Cleaning is the process of transforming raw data into consistent data that can be analyzed by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted. It is aimed at improving the content of statistical statements based on the data as well as their reliability. The below processes will be applied to all datasets. The data cleaning process also involves merging data from different sources, extracting subsets, or reshaping the data in some way. All in all, data cleaning is the process of turning raw data, into a data set that can be used in visualization, modeling, and decision-making.

1.4.1 Identify and remove prepending or trailing white spaces

To ensure the integrity of the datasets was removed white spaces were because there is a possibility to affect the correctness of conclusions resulting. Trailing whitespace is any spaces or tabs after the last non-whitespace character.

1.4.2 Identify typos on datasets

Imports of the datasets were checked for typographical errors. Technically, the term includes errors due to mechanical failure or slips of the hand or finger. First, to be able to check for possible spelling mistakes was measured the distance of dissimilarity of the characters between the editors. The first method that was applied to detect typos was the comparison of surnames based on Levenshtein distance and the total number of appearances of the pairs that were arisen. In information theory, linguistics, and computer science, the Levenshtein distance is a string metric for measuring the difference between two sequences. Informally, the Levenshtein distance between two words is the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one word into the other. *(Wikipedia)*

The Levenshtein distance between two strings a, b (of length |a| and |b| respectively) is given by lev (a, b):

$$\operatorname{lev}(a,b) = \begin{cases} |a| & \text{if } |b| = 0, \\ |b| & \text{if } |a| = 0, \\ \operatorname{lev}(\operatorname{tail}(a), \operatorname{tail}(b)) & \text{if } a[0] = b[0] \\ 1 + \min \begin{cases} \operatorname{lev}(\operatorname{tail}(a), b) \\ \operatorname{lev}(a, \operatorname{tail}(b)) \\ \operatorname{lev}(\operatorname{tail}(a), \operatorname{tail}(b)) \end{cases} & \text{otherwise.} \end{cases}$$

Equation 1. Levenshtein

Where the tail of some string x is a string of all but the first character of x, and x[n] is the nth character of string x, starting with character 0. Was created symmetric comparison table where had on the x and y-axis the unique surnames of the authors with the purpose to identify which surnames had the smallest difference of characters between them, as there is a high possibility of the existence of misspelling. The range of differences that were included for examining, reaches two characters.

The second method that was applied to detect typos in editorial members was the comparison of the combination of surnames/names based on Jaro–Winkler similarity and will be checked the timeline of appearances of editorial members in

issues by year. Jaro–Winkler similarity uses a prefix scale p which gives more favorable ratings to strings that match from the beginning for a set prefix length l. Given two strings S_1 and S_2, their Jaro-Winkler similarity sim_w is:

$$sim_w = sim_j + \ell p(1 - sim_j)$$

Equation 2. Jaro-Winkler similarity

- Sim_j is Jaro's similarity for strings S_1 and S_2.
- l is the length of a common prefix at the start of the string up to a maximum of 4 characters.
- p is a constant scaling factor for how much the score is adjusted upwards for having common prefixes.

The Jaro similarity Sim_j of two given strings S_1 and S_2 is

$$sim_j = \begin{cases} 0 & \text{if } m = 0 \\ \frac{1}{3}\left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right) & \text{otherwise} \end{cases}$$

Equation 3. Jaro-Winkler distance

- $|s_i|$ is the length of the string s_i.
- m is the number of matching characters.
- t is the number of transpositions.

Two characters from S_1 and S_2 respectively, are considered *matching* only if they are the same and not farther than $\left\lfloor \frac{max(|s_1|,|s_2|)}{2} \right\rfloor - 1$ characters apart.

Finally, this comparison will appear with two graphs, and where deemed necessary it will proceed in replacement of potential errors.

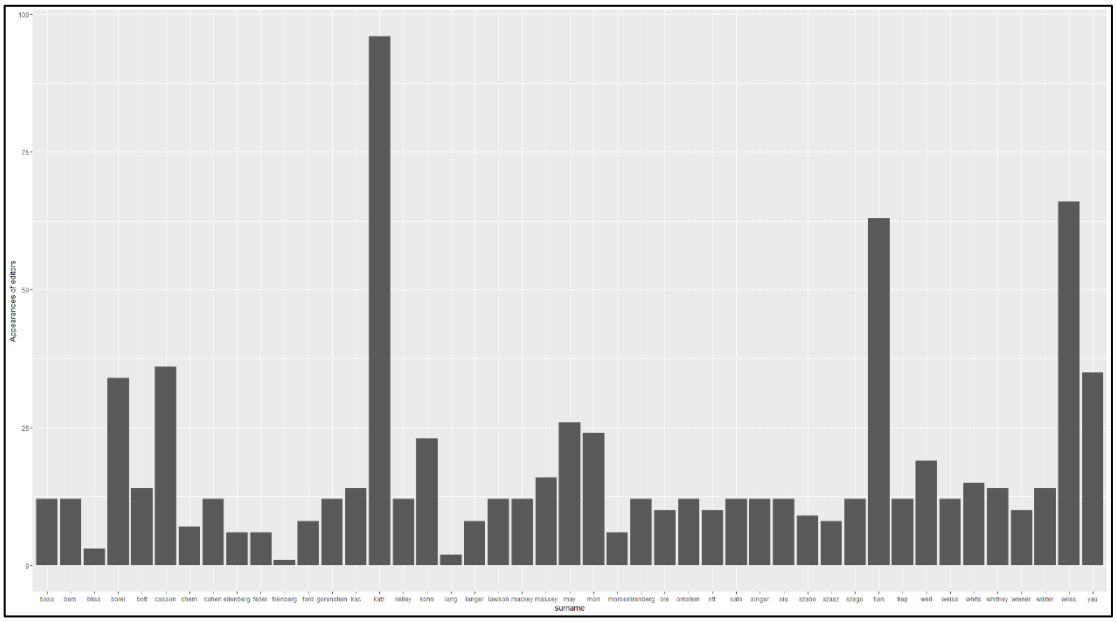


Figure 1.9 Number of appearances in issues of surnames after Levenshtein distance.

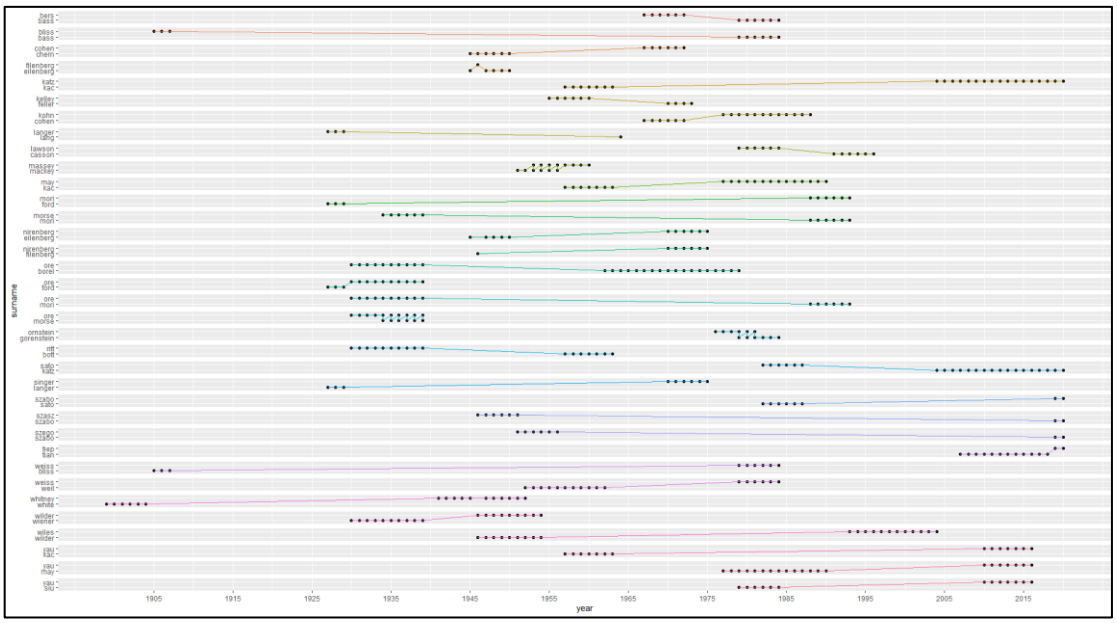


Figure 1.10 All the combinations of Levenshtein distance and comparison their appearance with the year of appearing.

The editor with the name Eilenberg was spotted with the possibility of a typo because he appears only once in volume 47 and has string distance from Filenberg only one character. Also, from the graph can be assumed that is a possible typo because the specific editor appears between the appears of editor Filenberg.

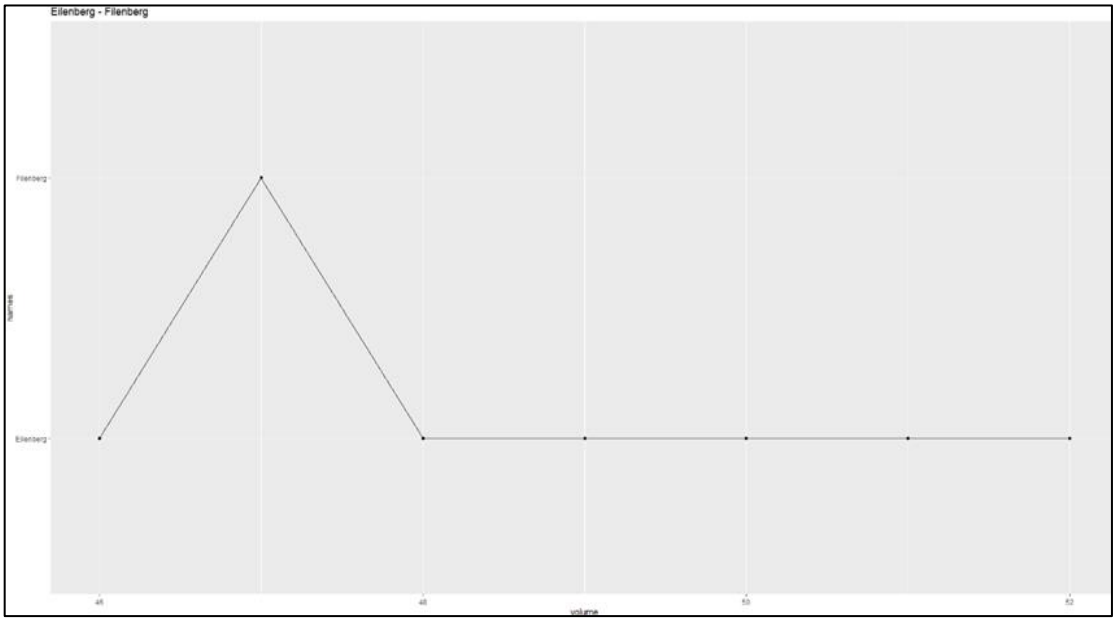


Figure 1.11 Pair of Eilenberg – Filenberg.

Subsequently, will follow the second method of checking that of Jaro–Winkler similarity. In the examined editorial members' combinations of surname/name were settled with $\boldsymbol{l}$=**0,15** and $\boldsymbol{p}$ =**0,1**. The results will be clustered based on the distance of difference and will be displayed the clustered results as a dendrogram. That dendrogram was created with a purpose to cluster editors that were displayed with different ways between issues to merge them, and the other purpose was to spot typos between imports.

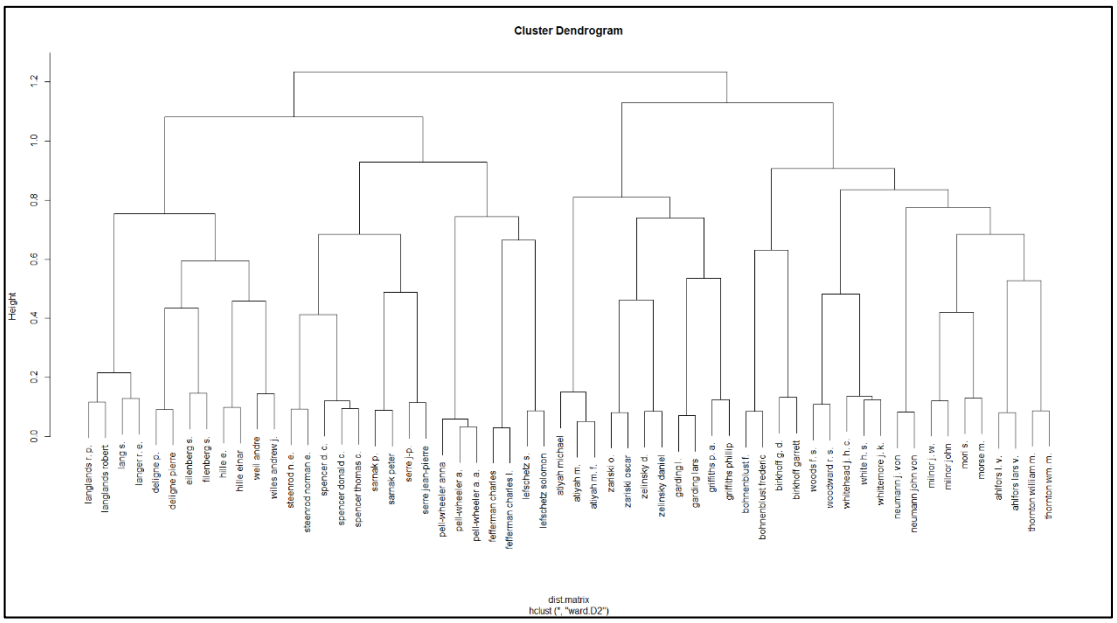


Figure 1.12 All the combination of surname/name of editorial members after Jaro-Winkler distance.

Have means to conclude whether it is the same editor with a different name or a spelling one it will be made a time-series graph with the grouped member names and compare the timelines between appearances. After checks by the groupings of the members, conclusions emerged about the real identities of the editorial participants. Arise four categories:

- Spelling mistakes (confirm conclusion about combination Filenberg-Eilenberg).
- Synonyms (either speaking for same family or other).
- Same identity (the same editor has been written differently).
- Different identity (different editors appears to compare because they have small string distance)

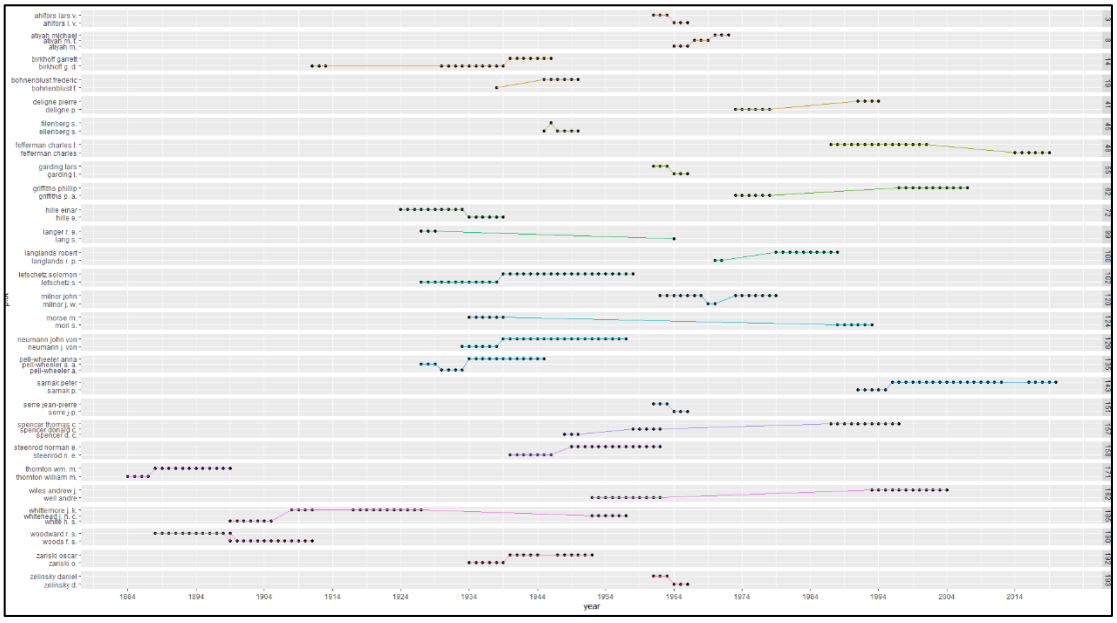


Figure 1.13 Timeline of clustered editors dendrogram.

With the above methods that have been applied can now happen a reliable analysis of the actual size of the editors that appear in the issues. At first, the sight had been recorded, 231 unique editors, due to the reasons mentioned above. After, the cleaning process unique editors ended up being 206. When the same editor appears in two or more different ways, it is kept the import with the highest appears, if there are equal appears on the same editor, the full name is kept.

1.4.3 Identify inconsistencies in datasets

After all the necessary checks have been made for spelling mistakes, data must be checked for inconsistencies. Inconsistences in datasets are generally compounded by the possibility for duplicated imports, the design and structure of our database, and the appearances of multiple tables within a database that deal with the same data but may receive it from different inputs. Below, will be categorized the inconsistencies of raw data that was met in the dataset and will be explained the procedures that will follow so that the data be robust and ready to be analyzed. In addition, for each process, there will be:

- bar plot with the number of issues we have on the editorial board.
- the pie chart which will be displayed each time the percentages of the number of issues which we have an editorial board that we have compared to the percentages of the number of unknown editorial boards of issues

Primarily, "duplication" just means that had repeated data in the dataset. This could be due to things like data entry errors or data collection methods. All the duplicated imports were checked and removed from the dataset. Subsequently, it is needed to check the design and structure of the database as it is a problem for the dataset to be consistent. At first sight, our datasets are displayed as they are:

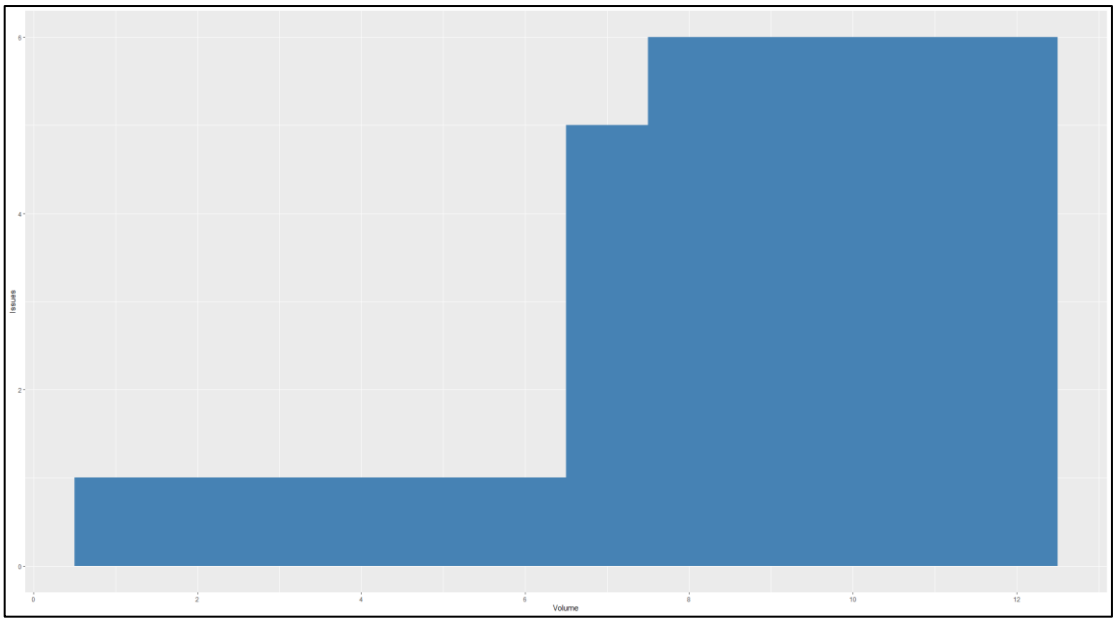


Figure 1.14 The bar plot of the number of editorial boards appears per volume in The Annals of Mathematics 1st period of publishing without processing.

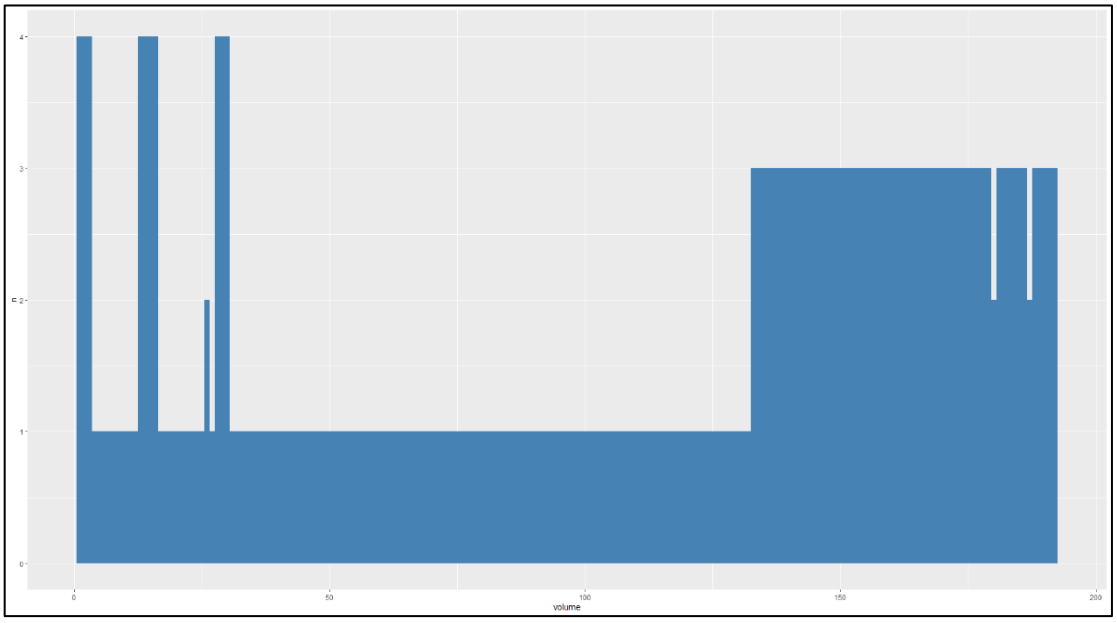


Figure 1.15 The bar plot of the number of editorial boards appears per volume in The Annals of Mathematics 2nd period of publishing without processing.

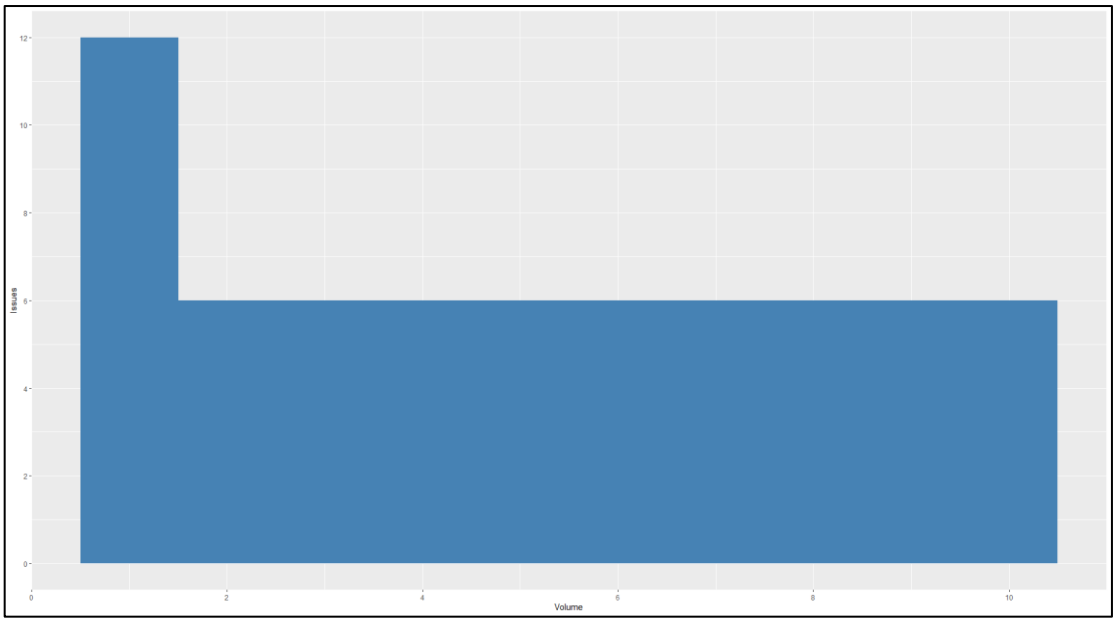


Figure 1.16 The bar plot of the number of editorial boards appears per volume in The Analyst without processing.

To secure a robust dataset to analyze, the dataset with the editorial boards must be compared with the recorded editorial members of each volume and the dataset in which all the issues that published. The comparison is resulting that the percentage of unknown volumes being large 41,44%. While gathered volumes there is from the editorial boards is 58,56%. All the issues are included in unknown issues that their editorial board does not directly know from the source.

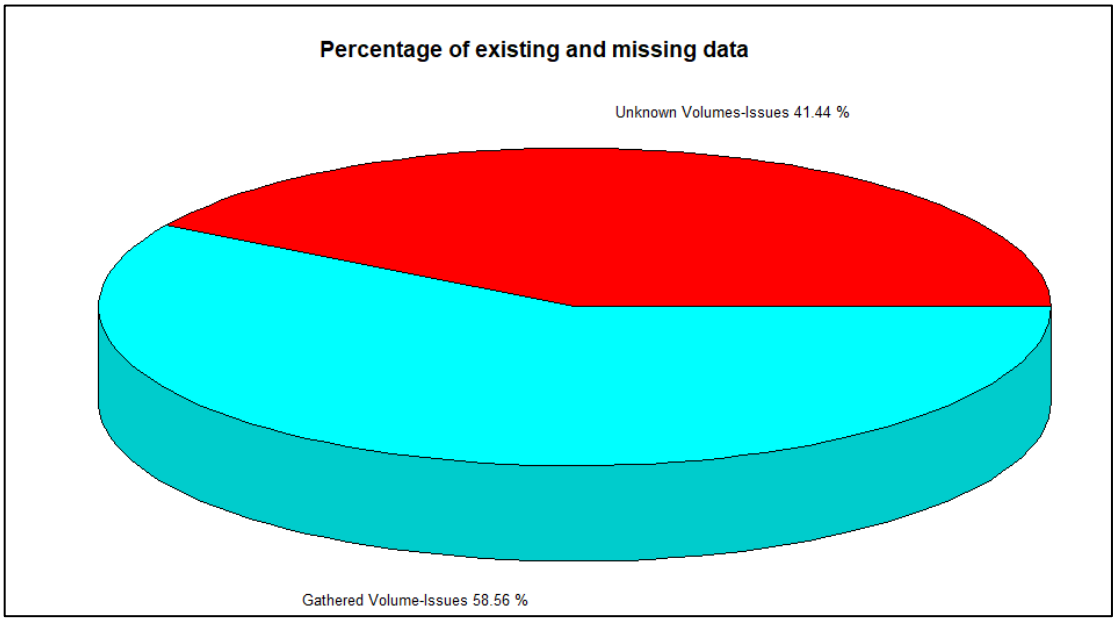


Figure 1.17 Pie chart of the percentage of known and unknown editorial boards

The reason that is illustrated large percentage of unknown issues is:

- the appearance of volume information, because there are periods where information about the participants can be collected from the editorial board of the volume and there is only one issue that was downloaded from the JSTOR, and
- missing issues

For the first reason, must be spotted the cases where this phenomenon occurs so that merge. In graph below is visualized with the red line the number of volumes of information was observed in each issue.

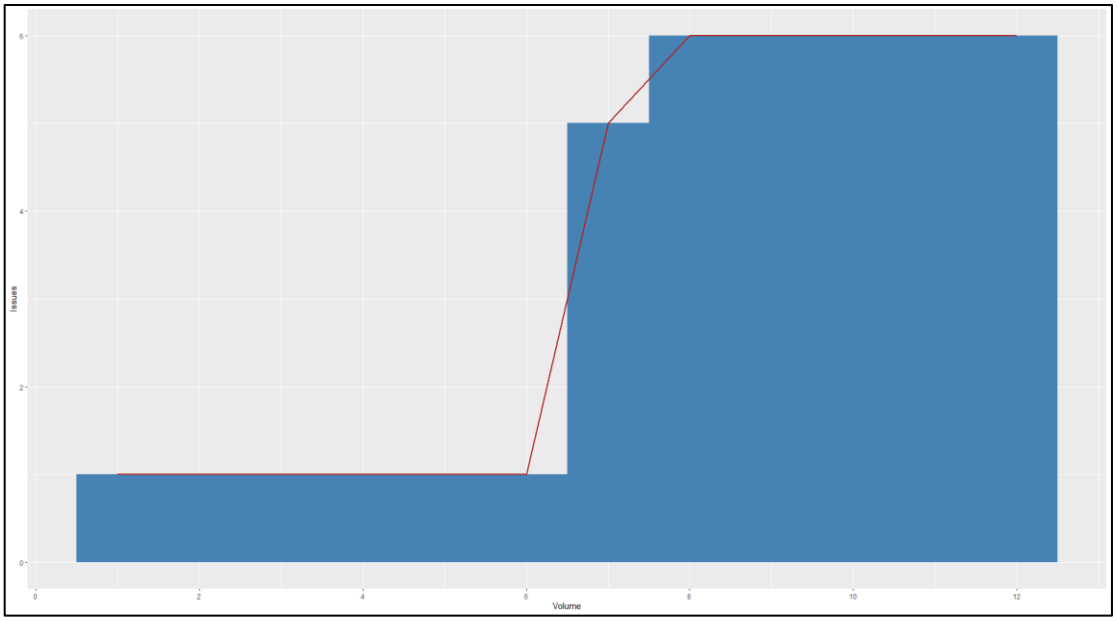


Figure 1.18 Editorial boards appear per volume in The Annals of Mathematics 1st period of publishing as referred to JSTOR with comparison the number of editorial boards there is in our dataset.

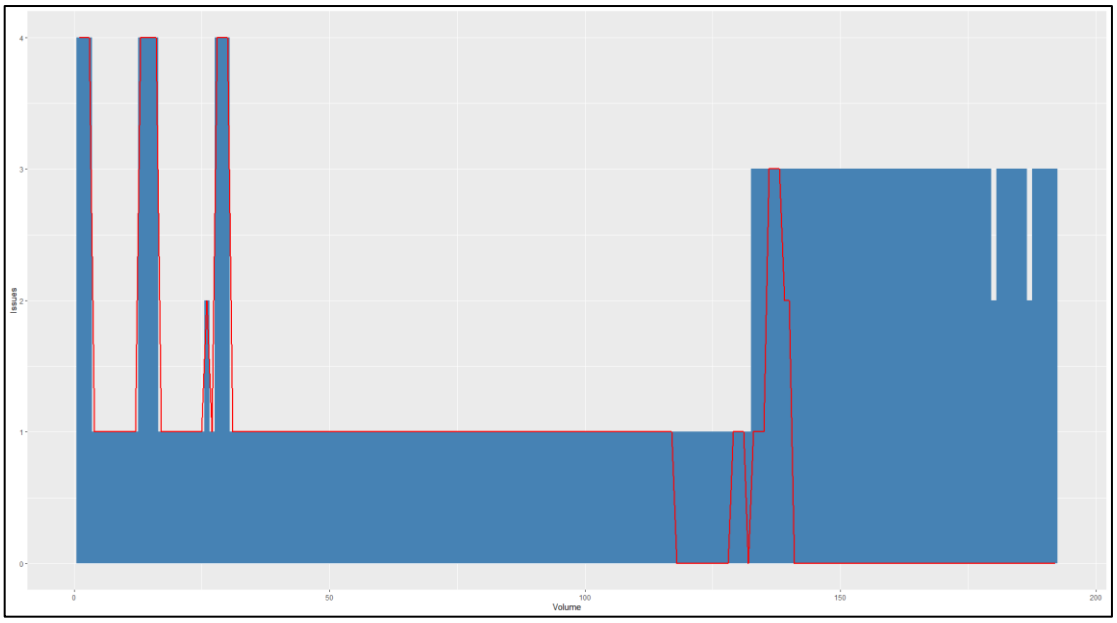


Figure 1.19 Editorial boards appear per volume in The Annals of Mathematics 2nd period of publishing as referred to JSTOR with comparison the number of editorial boards there is in our dataset.

To arrive at a reliable structure of examining datasets will be compared the raw datasets and the dataset that all the volumes were collected from JSTOR. For this reason, must be identified through this process all the issues that were not merged, within these issues that were not merged include those that do not have an editorial board (missing issues). To overcome this obstacle will be combined the new table that emerged, that did not have an editorial board at first sight but was grouped within the volume information and referred to an editorial board. The visualization of the dataset changed dramatically.

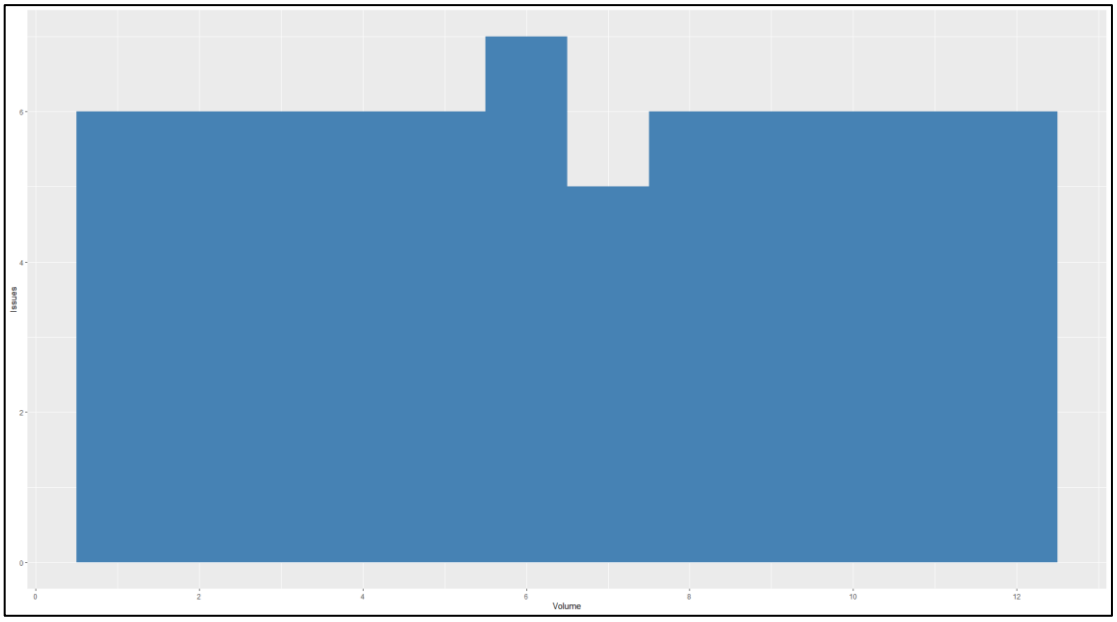


Figure 1.20 The Annals of Mathematics 1st period of publishing after the addition of editorial boards that was condensed in volumes.

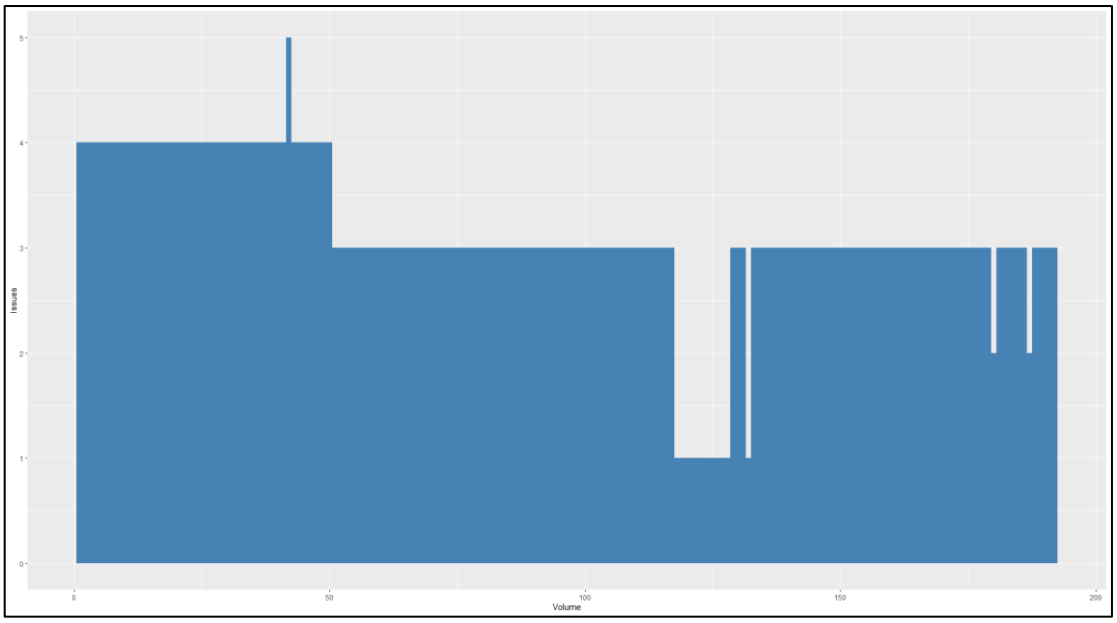


Figure 1.21 The Annals of Mathematics 2nd period of publishing after the addition of editorial boards that was condensed in volumes.

In the following graph, can be conjectured that the missing editorial boards there are when there is no editorial board to volume, but it is published. The missing volumes are displayed throughout the gap between the bar plot and the yellow line, in which the yellow line shows the number of published volumes based on JSTOR. The percentage of known volumes is raised to 96,6%, while the missing editorial boards are defined as 3,4%. In the next sub-paragraphs, the purpose is to eliminate missing values in the dataset so that there will explain this chapter on how to manage these specific values and complete them.

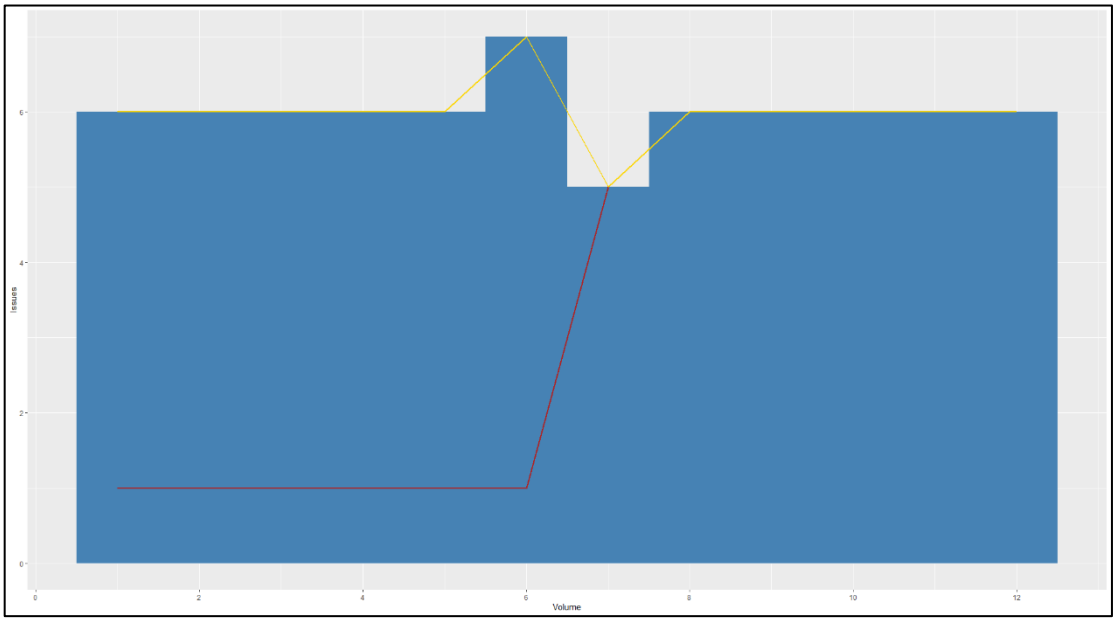


Figure 1.22 The Annals of Mathematics 1st period of publishing after the process and with comparison the number of editorial boards there are on JSTOR.

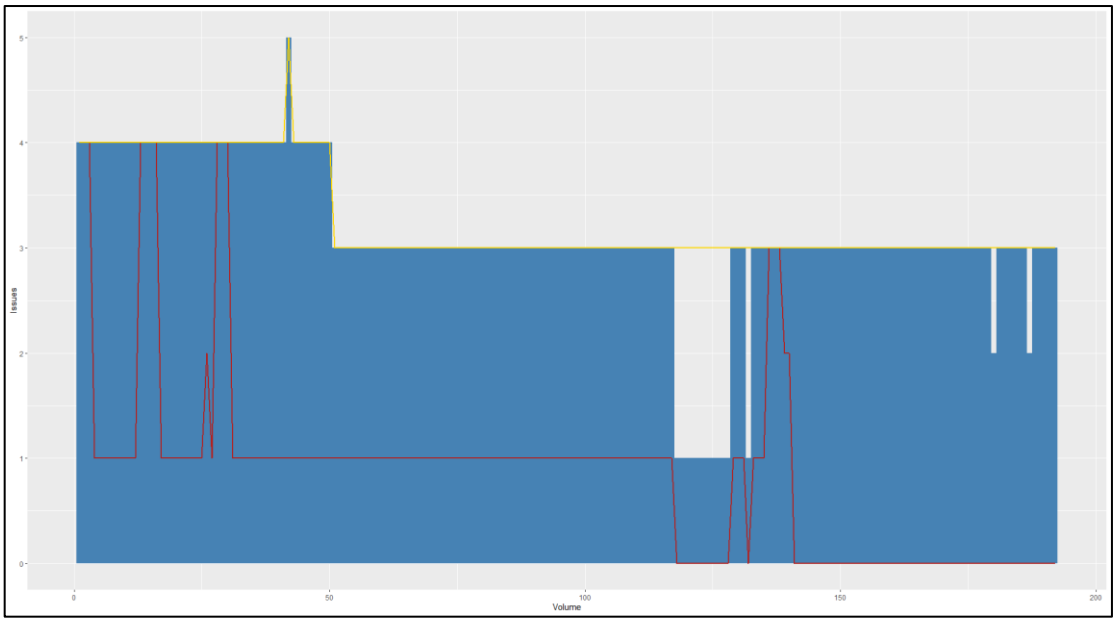


Figure 1.23 The Annals of Mathematics 2nd period of publishing after the process and with comparison the number of editorial boards there are on JSTOR.

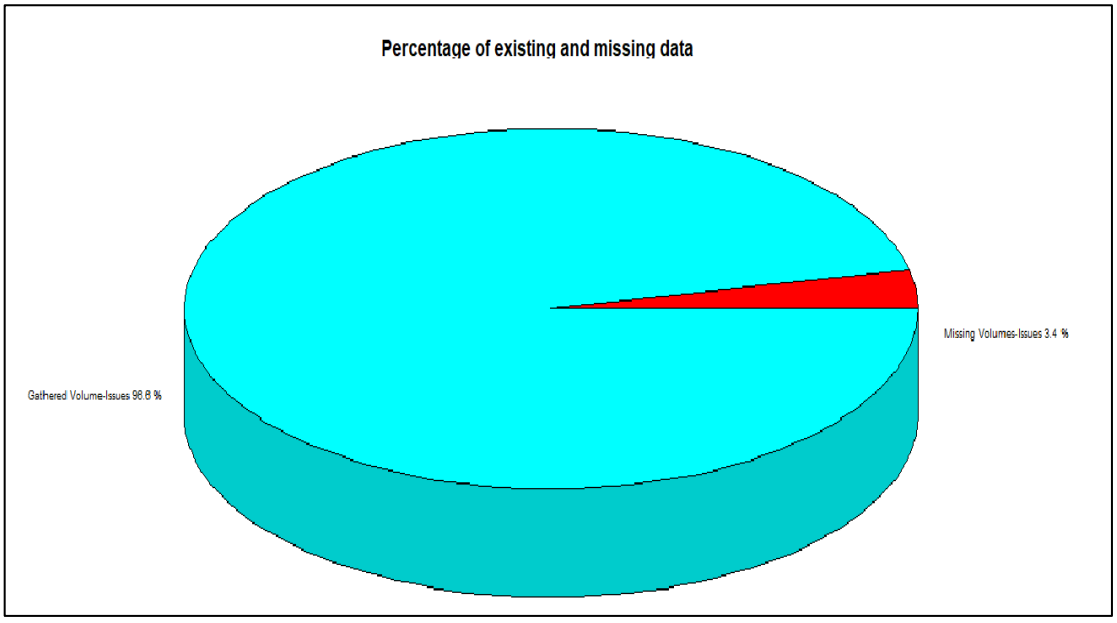


Figure 1.24 Pie chart of the percentage of known and missing editorial boards.

1.4.4 Transform all entries to lowercase and delete any accents

In this phase, data inside datasets will be converted to have common standards. Those standards are referred to above. To be all imports lowercase and in English (without accents).

UTF-8 is a variable-width character encoding used for electronic communication. Defined by the Unicode Standard, the name is derived from **Unicode (or Universal Coded Character Set) Transformation Format** – 8-bit. UTF-8 is by far the most common encoding for the World Wide Web, accounting for over 96% of all web pages, and up to 100% for some languages, as of 2021. *(Wikipedia)*

ASCII abbreviated from **American Standard Code** for Information Interchange, is a character encoding standard for electronic communication. ASCII codes represent text in computers, telecommunications equipment, and other devices. Most modern character-encoding schemes are based on ASCII, although they support many additional characters. *(Wikipedia)*

1.5 MISSING VALUES

Before jumping to the methods of data imputation, one must understand the reason why data goes missing.

• **Missing at Random (MAR):** Missing at random means that the propensity for a data point to be missing is not related to the missing data, but it is related to some of the observed data

• **Missing Completely at Random (MCAR):** The fact that a certain value is missing has nothing to do with its hypothetical value and with the values of other variables.

• **Missing not at Random (MNAR):** Two possible reasons are that the missing value depends on the hypothetical value (e.g., People with high salaries generally do not want to reveal their incomes in surveys) or a missing value is dependent on some other variable's value (e.g., Let's assume that females generally don't want to reveal their ages! Here the missing value in age variable is impacted by gender variable). *(Towards Data Science)*

As we saw in a previous sub-paragraph, we have missing volumes in which we do not have information about the editorial board, but we know which volume they are. For that reason, these issues that are missing will be integrated into the data table where all the editors are mentioned per issue, but the columns concerning the identity of the editors will remain blank.

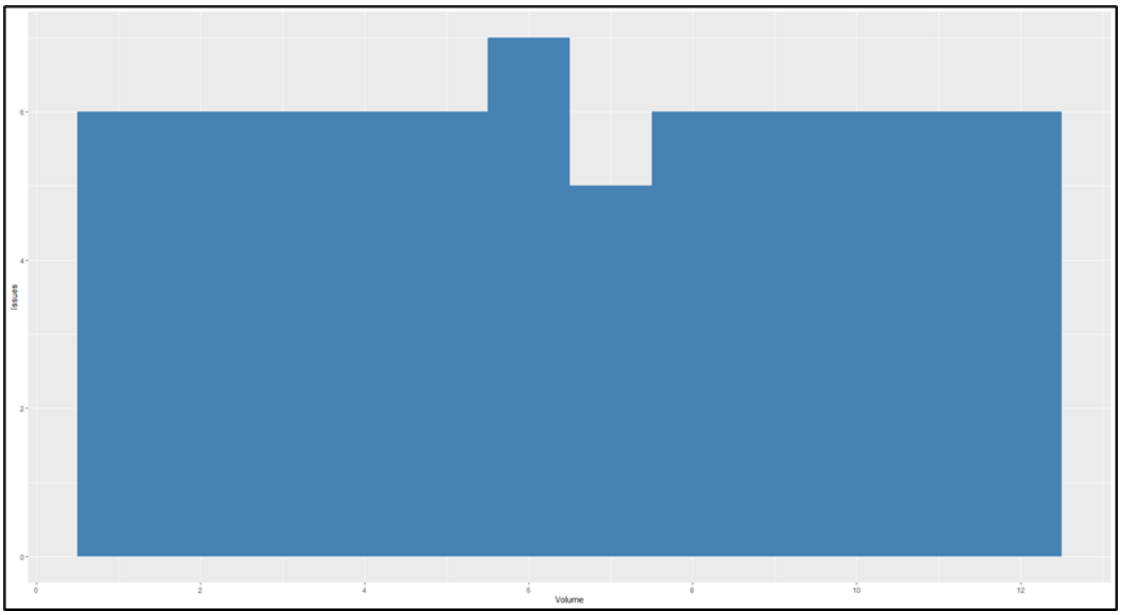


Figure 1.25 Final visualization of The Annals of Mathematics 1st period of publishing after addition of missing and condensed editorial boards.

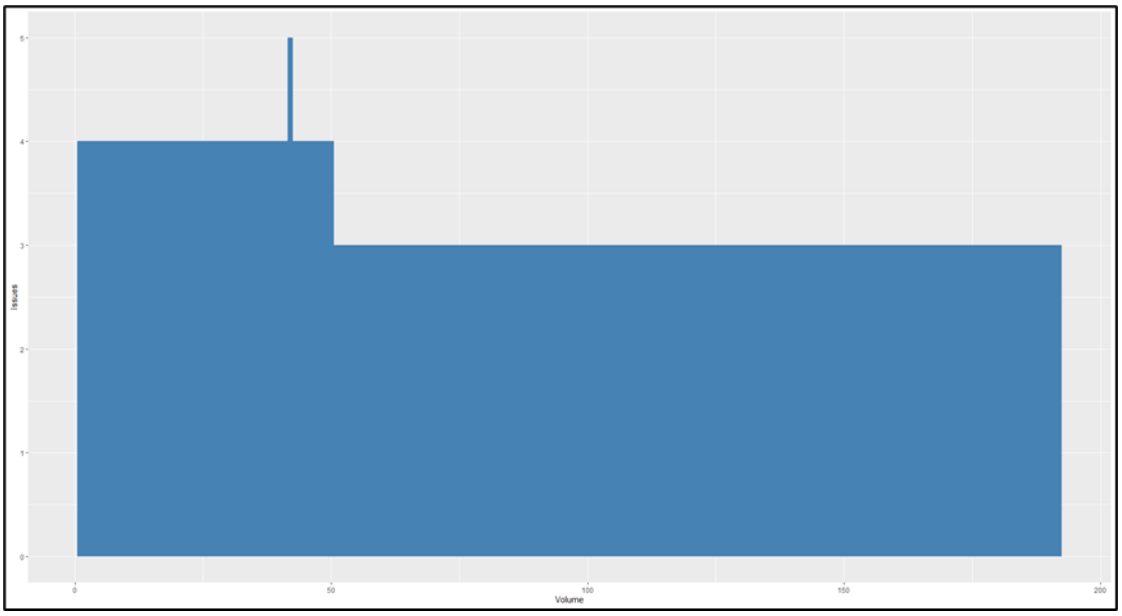


Figure 1.26 Final visualization of The Annals of Mathematics 2nd period of publishing after addition of missing and condensed editorial boards.

First, will be needed to know the number of missing values that have in the dataset and what part it takes in total data. In the sequel, will be separated the categories of missing values and will be managed in different ways. From the graphs with missing values can be concluded that missing values mainly appear in the months, but also in columns that concerning the identity of the editorial members after the last process. Underneath is represented the check that was held with the purpose to be understood the percentage of null imports.

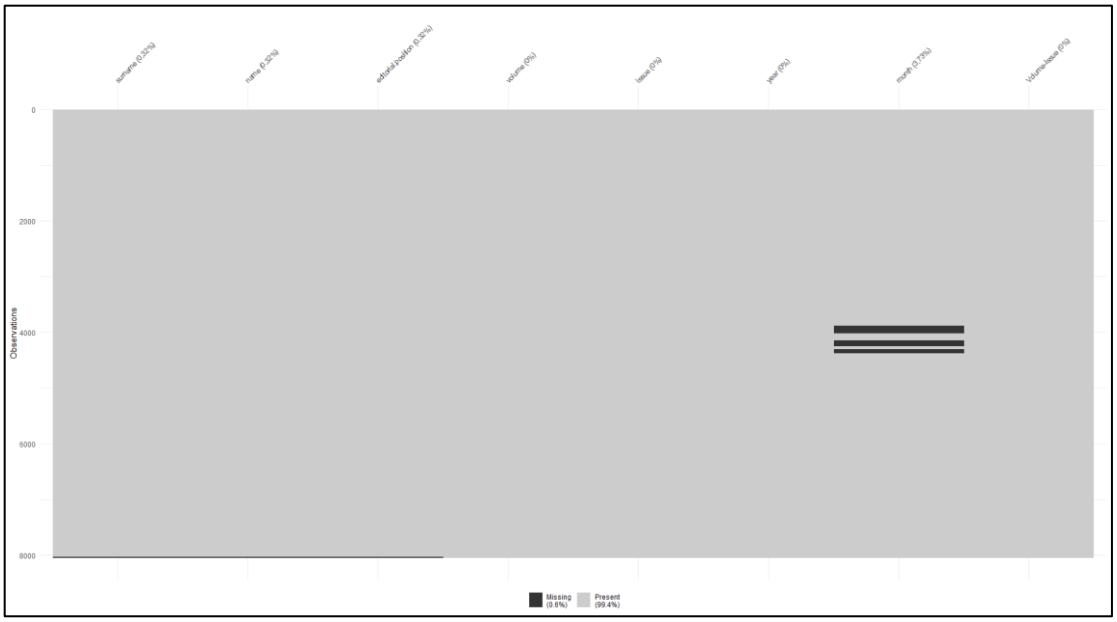


Figure 1.27 The Annals of Mathematics 2nd period of publishing NA values percentage graph.

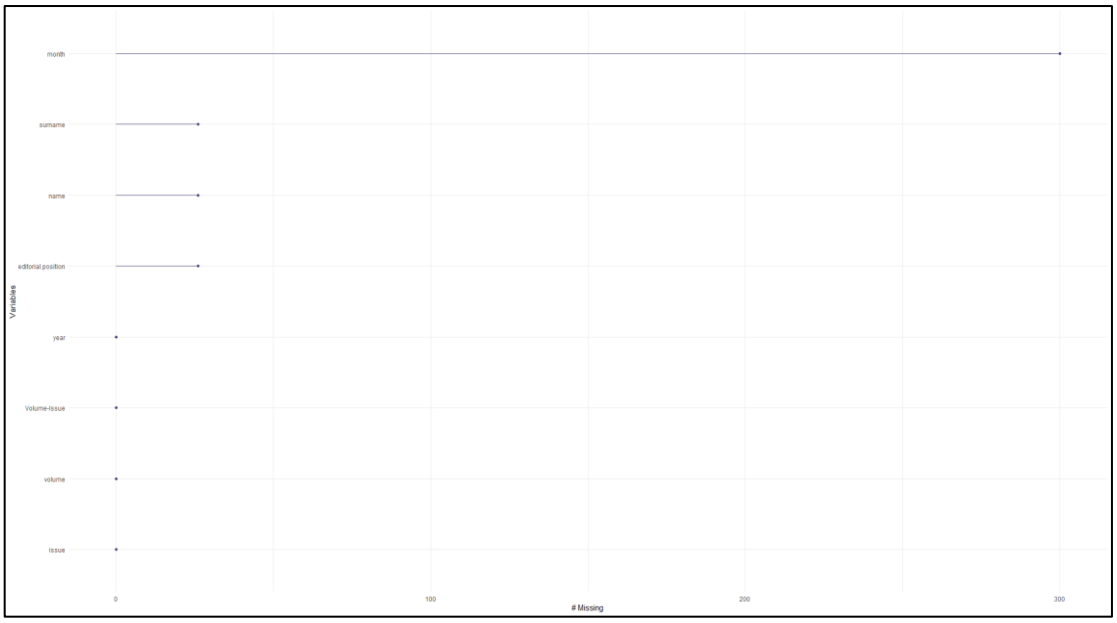


Figure 1.28 The Annals of Mathematics 2nd period of publishing NA values quantitative graph.

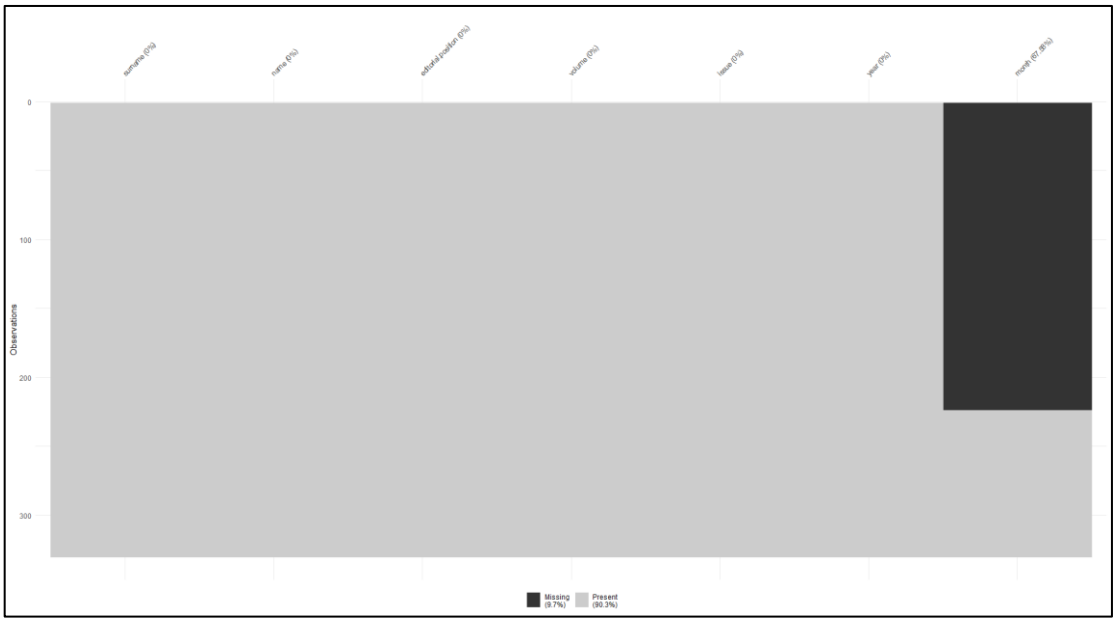


Figure 1.29 The Annals of Mathematics 1st period of publishing NA values percentage graph.

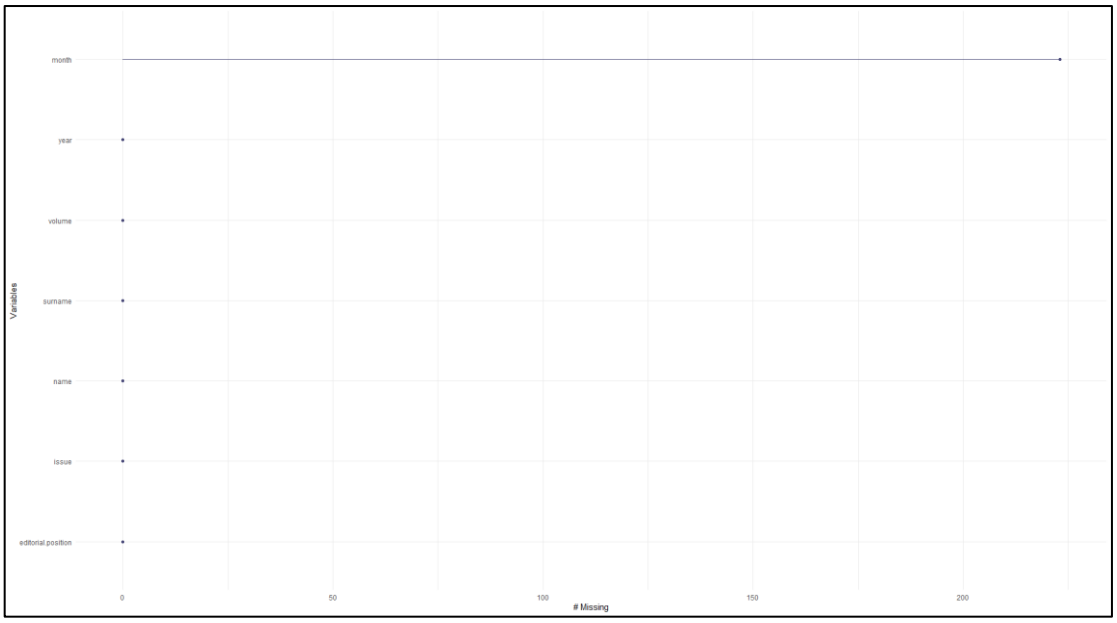


Figure 1.30 The Annals of Mathematics 1st period of publishing NA values quantitative graph.

In both cases, it is observed that they have in common missing values only in months. For the missing months of issues, made a hypothesis. The sequence of months in which the issue is published is proportional to the total number of issues during the year. For example, if in one year 6 issues have appeared then the published step is every 2 months, based on the above assumption made. Only in the unknown months will be a supplement. The publishing sequence is subtracted from the last month when it was known the month. After, the process that was explained presenting data goes to 99,9% for the 2nd period of publication of Annals of Mathematics, but for the 1st period of publication of Annals of Mathematics goes to 100%.

Figure 1.31 The Annals of Mathematics 1st period of publish, NA values percentage graph, after the process for filling missing editorial boards.

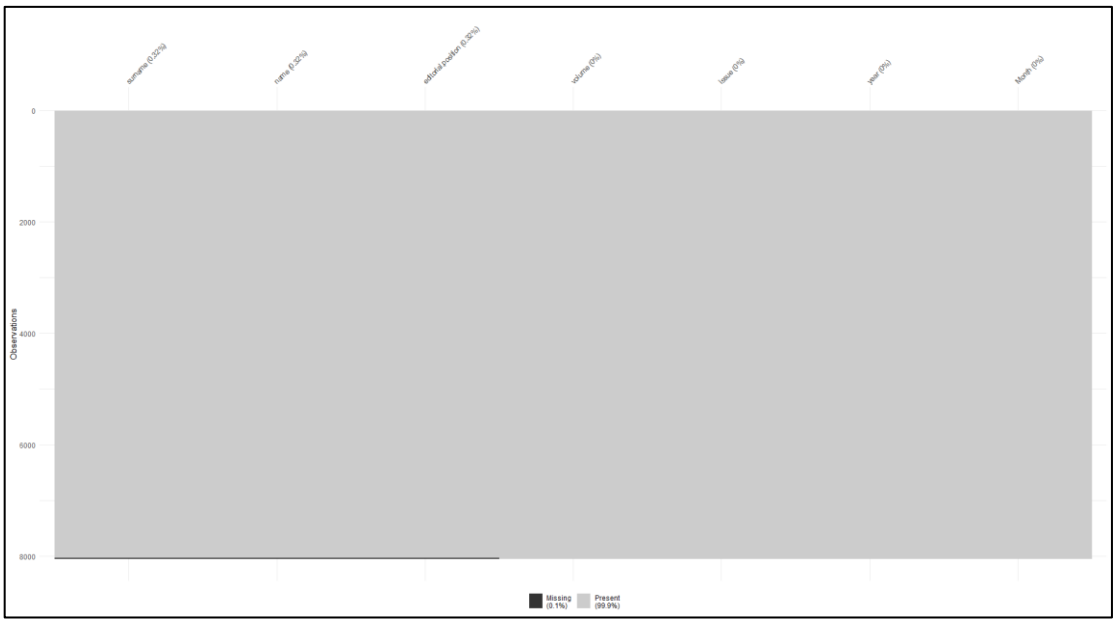


Figure 1.32 The Annals of Mathematics, NA values percentage graph, with missing editorial boards.

Subsequently, columns that concern the identity of the editors in each missing issue should be managed, be robust and consistent with the examined dataset. Of 26 missing editorial boards, the blank values as mentioned above will be filled with editorial board combinations. To achieve this, the method that will be applied is the comparison of the previous one, with the next known table of editorial members. If the combination is the same, they will be filled with the editorial board that was compared. After, the upper process remains 6 editorial boards missing, because the editorial board before and after is different. It is concluded that our dataset is almost robust with a tiny deviation from 100%. The missing values are dropped under 0,1%, while before they reached almost 0,32% on columns which were relevant with the identification of editors, while to all dataset was 0,1%. Must be referred that the missing editorial boards were considered that have the same editorial board with the corresponding known volume. The issues 121-1,121-2 are filled with the imports of the editorial board of 121-3, corresponding cases are volumes 126 and 132.

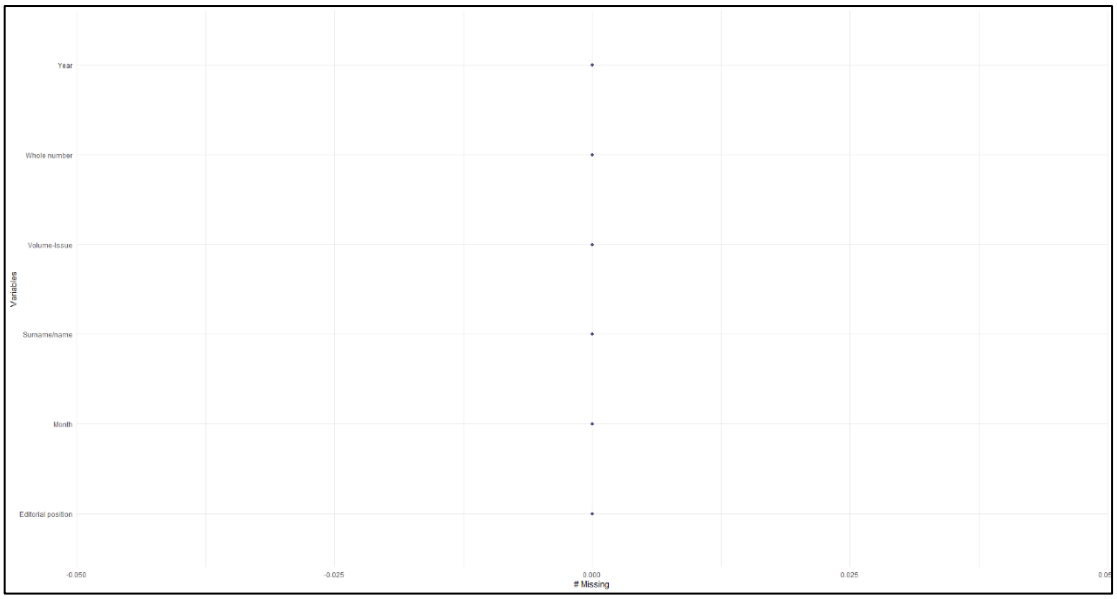


Figure 1.33 The Annals of Mathematics, NA values quantitative graph.

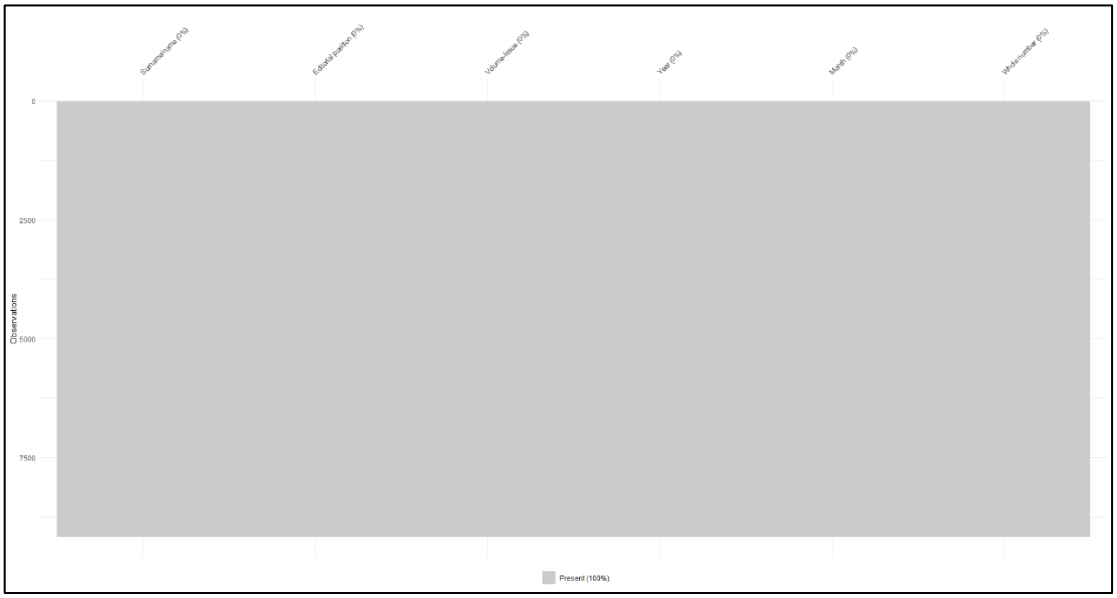


Figure 1.34 The Annals of Mathematics, NA values percentage graph, after the process for filling missing editorial boards.

Finally, yet important, as referred Annals of Mathematics is divided into two publish periods and has also, The Analyst. In the final data frame, the periods must be unified to analyze the data. Because they have repetitive numbering of issues to ensure the uniqueness of each issue, it is created a unique key, that of the whole number. Must be referred that the numbering of the whole number is sorted from the past to present.

1.6 OUTLINE OF THE PAPER

1.6.1 Data

From the data with the processing can have conclusions, either through the calculation of the observations or from their observation. Through data analysis, people can study these phenomena. Data analysis can be regarded as seeking answers to various questions regarding the phenomena. For that reason, chapter 2 will be presented with all the basic datasets and will set procedures and best guide pipelines to have accurate analyzes. Will be set tasks through questions that must be answered in the following chapters.

1.6.2 Editorial board roles across time

The roles that make up our data table will be recorded; with the processes, their summations will be recorded and then, if deemed necessary, the roles will be changed to make our result more efficient. Then, for our results to be understood, the results will be displayed on a timeline. Then, the same procedures will be applied for the combinations that contribute to the issues of our journal. Finally, will be an investigation for the percentage they occupy in each role and role combinations.

1.6.3 Editorial board size across time

Analysis of the size of the editorial boards across the time, as well as the combinations of positions that compose the editorial board, the analyzes will be done with the criteria of time and the number of issues. In addition, changes and changes in the editorial boards will be checked and analyzed.

1.6.4 Number of individual editorial board members per role

Will be reported independently to each editor the sets of impressions they have in each role within the journal. In addition, there will be an investigation into the appearance of an editor in over one role.

1.6.5 Duration of participation per unique editorial board member

Will be analyzed the number of issues that took part every member at first as roles and next to per member the activity in the journal and meanwhile will be represented a timeline to understand the time series of these phenomena

.

CHAPTER 2
DATA

Data are individual facts, statistics, or items of information, often numeric, that is collected through observation. In a more technical sense, data are a set of values of qualitative or quantitative variables about one or more persons or objects, while a datum is a single value of a single variable. *(Wikipedia.com)*

Qualitative variables are non-statistical and are typically unstructured or semi-structured. This data isn't necessarily measured using hard numbers used to develop graphs and charts. Instead, it is categorized based on properties, attributes, labels, and other identifiers. On the other side, quantitative data is statistical and is typically structured in nature – meaning it is more rigid and defined. This data type is measured using numbers and values, making it a more suitable candidate for data analysis. *(g2.com)*

2.1 PRESENTATION OF DATASETS

They are three main datasets, at the first dataset it is recorded every issue, the names of editors, and collected information was gathered on both the editor and the issue. The following columns are from the first dataset, do not need to provide information about the names of the column, dimension of dataset is **[8839 (rows) x 6 (columns)**]:

- Surname/name
- Editorial position
- Volume-Issue
- Year
- Month
- Whole number

A second dataset had recorded from every issue the appearance of information and where it gathered; dimension of dataset is ***[765(rows) x 8(columns)]***. More specifically the following columns are from the second dataset, all columns related to the source of the information have values,

- Volume
- Issue
- Year
- Month
- Front matter
- Index
- Back matter
- Volume information

The third dataset had recorded from every issue the appearance of information and where it was gathered; with the only difference, it is recorded the number of pages that gathered. More specifically the following columns are from the third dataset dimension of the dataset is ***[765(rows) x 13(columns)]***:

- Volume
- Issue
- Year
- Month
- Front matter
- Index
- Back matter
- Editorial board
- Total pages front matter
- Total pages index
- Total pages back matter
- Total pages volume information

2.2 PROCEDURES & GUIDE PIPELINES

Sun Tzu is regarded as one of the greatest military strategists. During the twentieth century, The Art of War grew in popularity and saw practical use in Western society as well. It continues to influence many competitive endeavors in the world, including culture, politics, business, and sports, as well as modern warfare. What does the strategist say about tactics?

"Strategy without tactics is the slowest route to victory.

Tactics without strategy is the noise before defeat."

(Sun Tzu, Art of War)

The data that takes center stage in bibliometric analysis tends to be massive (e.g., hundreds, if not thousands) and objective in nature (e. g., number of citations and publications, occurrences of keywords and topics), though its interpretations often rely on both objective (e.g., performance analysis) and subjective (e.g., thematic analysis) evaluations established through informed techniques and procedures. In other words, bibliometric analysis is useful for deciphering and mapping the cumulative scientific knowledge and evolutionary nuances of well-established fields by making sense of large volumes of unstructured data in rigorous ways. Therefore, bibliometric studies that are well done can build firm foundations for advancing a field in novel and meaningful ways—it enables and empowers scholars to (Journal of Business Research)

- gain a one-stop over-view,
- identify knowledge gaps,
- derive novel ideas for investigation,
- position their intended contributions to the field.

No.	Step	Best practice (questions) guidelines
Step 1	Define the aims and scope of the bibliometric study	• What are the aims and scope of the study? • Is the scope of the study large enough to warrant the use of bibliometric analysis?
Step 2	Choose the techniques for bibliometric analysis	• What bibliometric analysis techniques should be chosen to meet the aims and scope of the study?
Step 3	Collect the data for bibliometric analysis	• Do the search terms exemplify the scope of the study? • Is the coverage of the database adequate for the study? • Is the data free of errors such as duplicates and erroneous entries? • Does the final dataset fulfil the requirements of the bibliometric analysis techniques chosen for the study?
Step 4	Run the bibliometric analysis and report the findings	• Can the bibliometric summary be easily understood by readers? • Does the writing align with the bibliometric summary presented? • Does the writing explain the peculiarities and implications of the bibliometric summary? • Does the writing align with the target outlet for publication?

Figure 2.1

2.2.1 Defined the aims and the scope of the bibliometric study

Foremost, must be defined the aims and the bibliometric study it is aimed to have a display of our data from Annals of Mathematics and to draw qualitative conclusions. It will be set two criteria of analysis, time, and the number of issues so that the multitudes of the publications it was discussed as well as for the periods to which they refer can be understood. After that, the datasets must be evaluated if it is large enough to warrant the use of bibliometric analysis. This paper has 3 main datasets as referred, the first dataset which includes all the editors that served in every issue has dimensions after processing, **8839(rows) x 6(columns)** and that happened after multi-comparisons with the other 2 main datasets. Hence, the first dataset is complete and has almost zero missing values **(<0.1%)**, so from now on the information and our study will be done from it.

2.2.2 Choose the techniques for bibliometric analysis

Incidental, the second phase of the process is to choose the techniques for bibliometric analysis to achieve the tasks that had been set in the previous phase. To achieve an effective and conclusive analysis of the dataset will be performed an examination of our repository based on the questions that will be set in every chapter, which answer tasks that will be set. In more detail, the tasks that are to be raised are mentioned in ***Figure 2.2.***

2.2.3 Collect the data for bibliometric analysis

As referred to in the previous chapter, thorough checks have been made for the final configuration of the data so that to draw as most accurate conclusions as possible. To summarize all the processes that took place to the dataset:

- Check for duplicates and their removal.
- Check for missing volumes and the treatment which follows them.
- Check about different names of similar people and treatment of cases.

All the upper processes are referred to and analyzed extensively in subparagraph ***1.4 Data Cleansing***.

2.2.4 Run the bibliometric analysis and report the findings

In the end, beyond the analyzes, the conclusions that emerge should be understood by the reader, explaining both the methodology and the reason chosen.

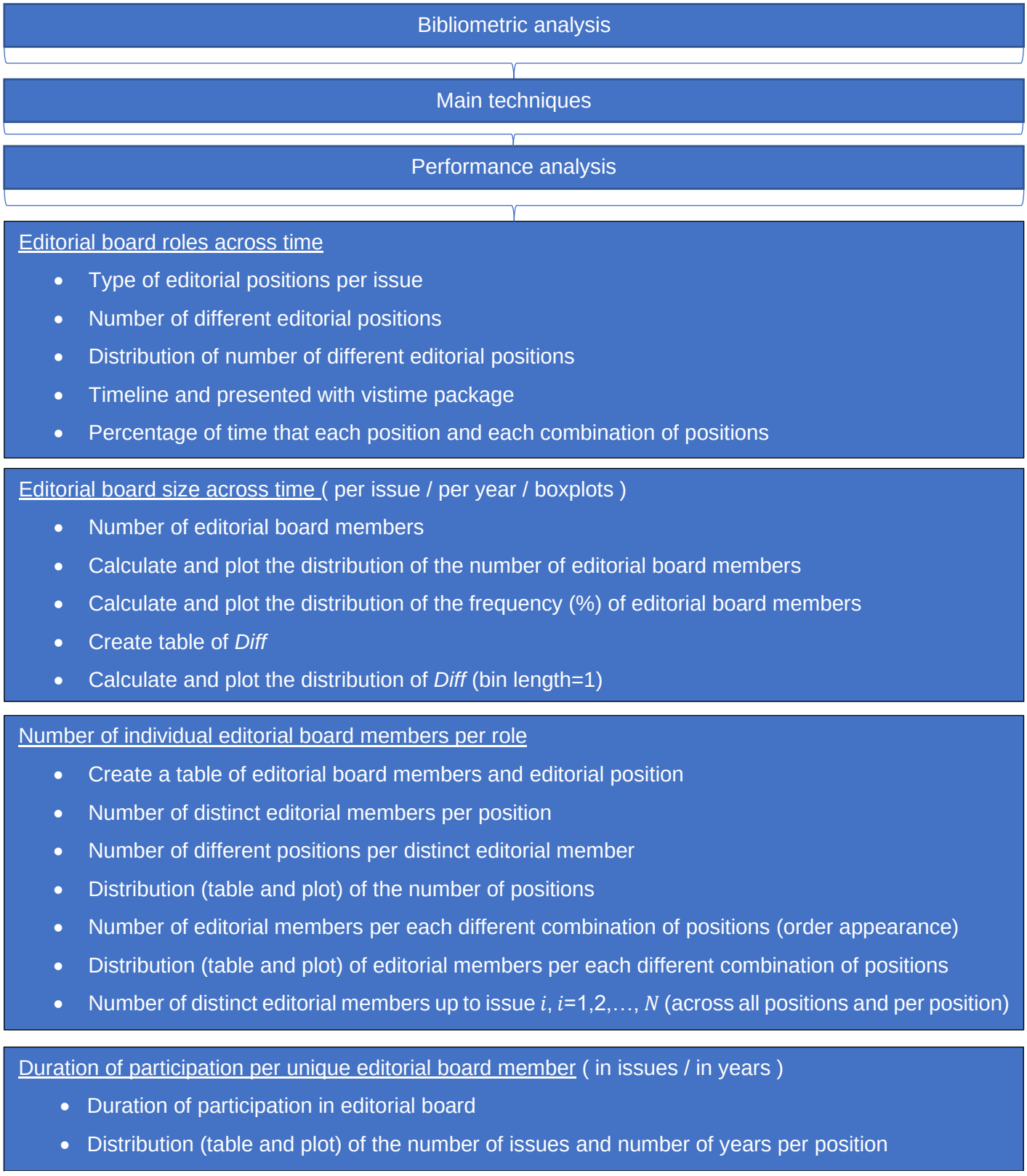

Bibliometric analysis
Main techniques
Performance analysis
Editorial board roles across time
Type of editorial positions per issue
Number of different editorial positions
Distribution of number of different editorial positions
Timeline and presented with vistime package
Percentage of time that each position and each combination of positions
Editorial board size across time (per issue / per year / boxplots)
Number of editorial board members
Calculate and plot the distribution of the number of editorial board members
Calculate and plot the distribution of the frequency (%) of editorial board members
Create table of Diff
Calculate and plot the distribution of Diff (bin length=1)
Number of individual editorial board members per role
Create a table of editorial board members and editorial position
Number of distinct editorial members per position
Number of different positions per distinct editorial member
Distribution (table and plot) of the number of positions
Number of editorial members per each different combination of positions (order appearance)
Distribution (table and plot) of editorial members per each different combination of positions
Number of distinct editorial members up to issue i, i=1,2,…, N (across all positions and per position)
Duration of participation per unique editorial board member (in issues / in years)
Duration of participation in editorial board
Distribution (table and plot) of the number of issues and number of years per position

Figure 2.2

Any task (question) implies two parts: a target, i.e. what information needs to be obtained, and the constraints, i.e. what conditions this information needs to fulfill. The target and constraints can also be viewed as unknown and known (specified) information, respectively; the goal is to find the initially unknown information corresponding to the specified information. The simplest example is to find the value of an attribute corresponding to a certain specified reference. Our task typology has its origin in the ideas expressed by Jacques Bertin in his Semiology of Graphics (Bertin 1967/1983). Like Bertin, it is distinguished tasks according to the level of data analysis ("reading level", in Bertin's terms) but additionally considers the division of data components into referrers and attributes:

- Elementary tasks refer to individual elements of the reference set; for example, "what is the number of total editors in Annals of Mathematics?"
- Synoptic tasks involve the whole reference set or its subsets; for example "during the entire period, what was the average number of editors?".

The tasks are further divided according to the target ("question type", in Bertin's terms), i.e. what is the unknown information that needs to be found. At the elementary level, the target may be one or more characteristics (attribute values) or one or more references (referrer values). A graphical illustration of this group of questions is shown in *Figure 3.1*. Here:

- ***R*** is the reference set of a dataset
- ***C*** is the character set
- ***f*** is the data function, which defines the correspondence between the elements of the reference and characteristic sets
- ***r*** is some specified reference, i.e. an element of the set ***R***.

The corresponding characteristic, which is determined by the data function, is unknown and needs to be found. This task target is indicated in the picture by a question mark. The opposite case is when some attribute value is specified, and the goal is to find the corresponding reference or references. Hence, the references are the target, and the attribute value defines the constraint. This kind of task is illustrated graphically in the right part of *Figure 3.1*. Here, the question mark indicates the unknown reference. In *Figure 3.1*, schematic representations of the task of determining the character that corresponds to a given reference ***r*** according to the data function ***f***, and the task of determining the reference that corresponds to a given characteristic ***C***.

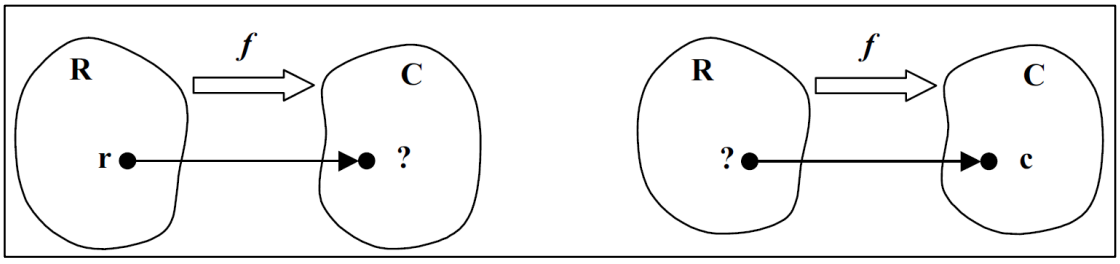


Figure 2.3 Schematic representations of the elementary tasks.

In tasks of the second type, inexact constraints are very often used. This means that a subset or range of attribute values is specified rather than an individual value. Therefore, the graphical representation of such a task could be modified, as shown in Fig. 3.2. In tasks of finding references corresponding to given characteristics, the

characteristics may be specified imprecisely, that is, a subset (designated as ***C'***) rather than an individual element of the character set ***C'*** may play the role of a task constraint. The target comprises all references corresponding to any of the elements of the subset ***C***.

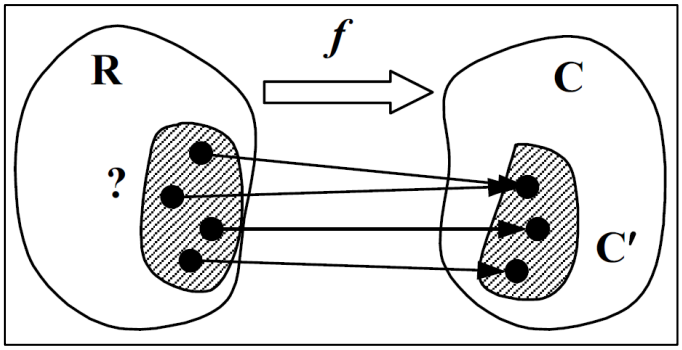


Figure 2.4 Schematic representations of the synoptic tasks.

In addition, there are tasks in which two or more questions are included. This is an effect of the presence of two referential components in the respective datasets. The questions require one to find combinations of values of referrers corresponding to values of the attribute, i.e. the value of each referrer needs to be determined. This is illustrated in Figure 3.3. However, nothing prevents us from formulating simpler questions, in which a value of one referrer is specified, and only the value of the other referrer needs to be found. For a dataset with multiple referential components, values of either all or just some of the referrers may be task targets. This picture illustrates a case with two referrers. On the left, the values of both referrers are unknown and need to be found. On the right, the value of one referrer is specified, i.e. is a part of the task constraint, and only the value of the other referrer is the task target.

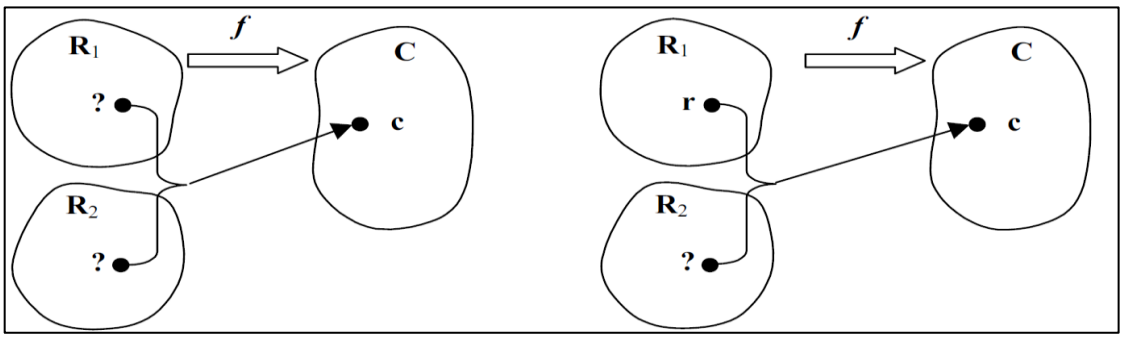


Figure 2.5 Schematic representations of the multiple referential components.

For synoptic tasks, it can be extracted the following methods:

- **Association**: Perception or description of a (sub)set of references as a unified whole based on similarity of their characteristics, i.e. close values of one or more attributes corresponding to these references.
- **Differentiation**: Perception or description of some references or subsets of references as differing from others by their characteristics.
- **Arrangement**: An idea or description of how characteristics are arranged, for an ordering of references, for example, a trend in a characteristic that changes over time.
- **Distribution summary**: A general idea or description of how characteristics are distributed over a reference set: how varied they are, what values occur most frequently, whether there are outliers (a few values greatly differing from the rest), etc.

(Natalia Andrienko, Gennady Andrienko)

CHAPTER 3
EDITORIAL BOARD ROLES ACROSS TIME

Primitively, must mention that there will be an analysis of our dataset regarding the editorial position. The analysis will be done in the first phase per issue and then per year of publication. It will become a matrix and the appearance of each editorial position in the issue will be marked with “X”, and with NA all the positions that do not appear in the issue, after must be converted to 1 all the “X” of the dataset so that they can be measurable. Additional, every process will be plot, so that can be understood conclusions or any interlocutors and visually. Our purposes in this process are:

- To investigates how many and which editorial positions appear, setting the criteria of issue and year.
- Find all the different combinations that appear in the editorial boards
- To define in both roles and combinations timeline
- Percentage of time that each position and each combination of positions were available

In the dataset of Annals of Mathematics, there are nine editorial positions from 1874 to 2020 with 765 issues that be published. The editorial positions that were gathered from volumes are : ***associate editors***, ***edited and published by***, ***edited by***, ***editor***, ***editor-in-chief***, ***editorial board***, ***founded by***, ***technical edited by*** and ***with the cooperation of***.

3.1 EDITORIAL POSITIONS APPEARANCES

In the total in all issues, the number of appearances in issues the editorial positions is:

Editorial position	Group	Number of total appearances	Number of editors	Percentage of number of editors
associate editors	A	60	222	2,54%
edited and published by	B	66	66	0,76%
edited by	C	633	2864	32,77%
editor	D	42	42	0,48%
editor-in-chief	E	18	18	0,21%
editorial board	F	6	12	0,14%
founded by	G	12	12	0,14%
technical edited by	H	170	170	1,94%
with cooperation of	I	557	5335	61,03%
TOTAL		**1564**	**8741**	**100,00%**

Once it is known all of them, the next step is to display where all these editorial positions have appeared in issues. To be able to cross-reference the information about

editorial positions, it will be displayed the appearance of editorial positions with two criteria that were referred upper in the introduction of this chapter.

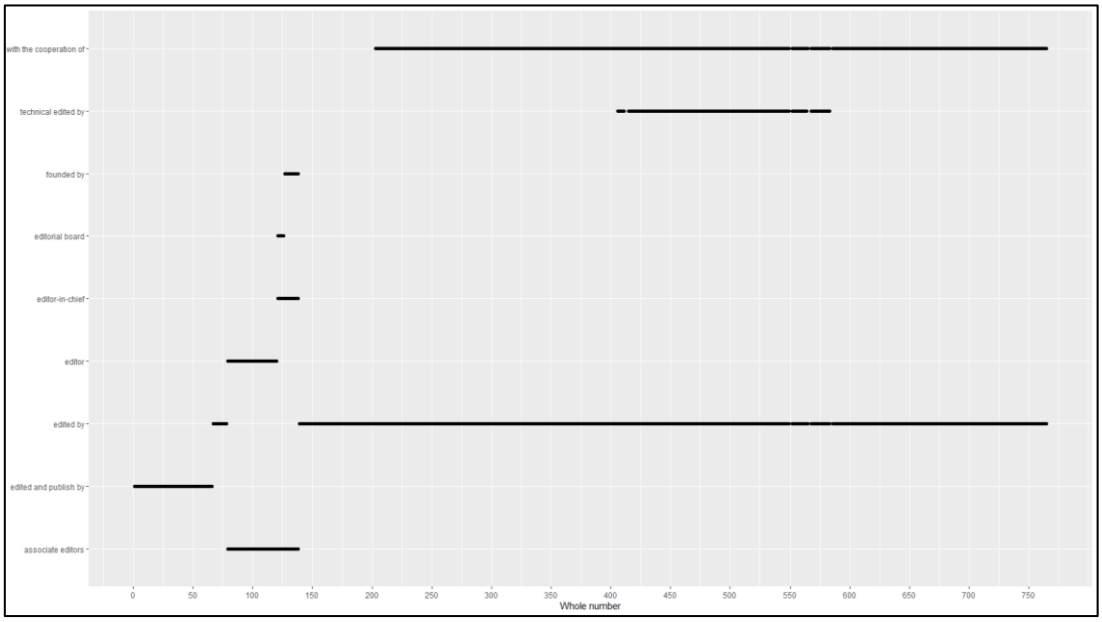


Figure 3.1 Graph editorial positions per issue.

The purpose of the below graph is to understand possible changes in the appearances behavior of editorial positions, to come to possible conclusions or even new questions to be explored. On the x-axis the years are displayed and the y-axis is the editorial positions, also inside the bar plots, the number is the number of total months that position has appeared continuously.

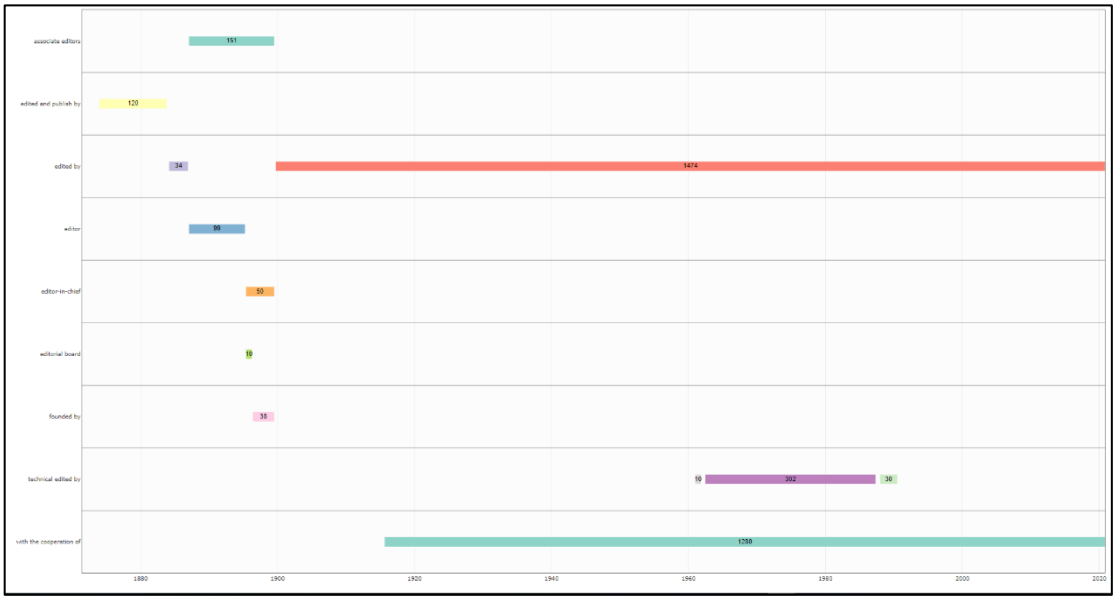


Figure 3.2 Timeline of editorial positions and the months that were available.

From Figure 3.5, at first sight, there are positions of editors that appeared at first publications of the journal and there are not established or appeared for a short interval. Also, the interval that should be discussed is 1886-1899, that is the period when the position “edited by” disappeared, while appeared new editorial positions. The position “edited by” was established, when was reappeared, and all the others which were appeared that period just had stopped after a few years.

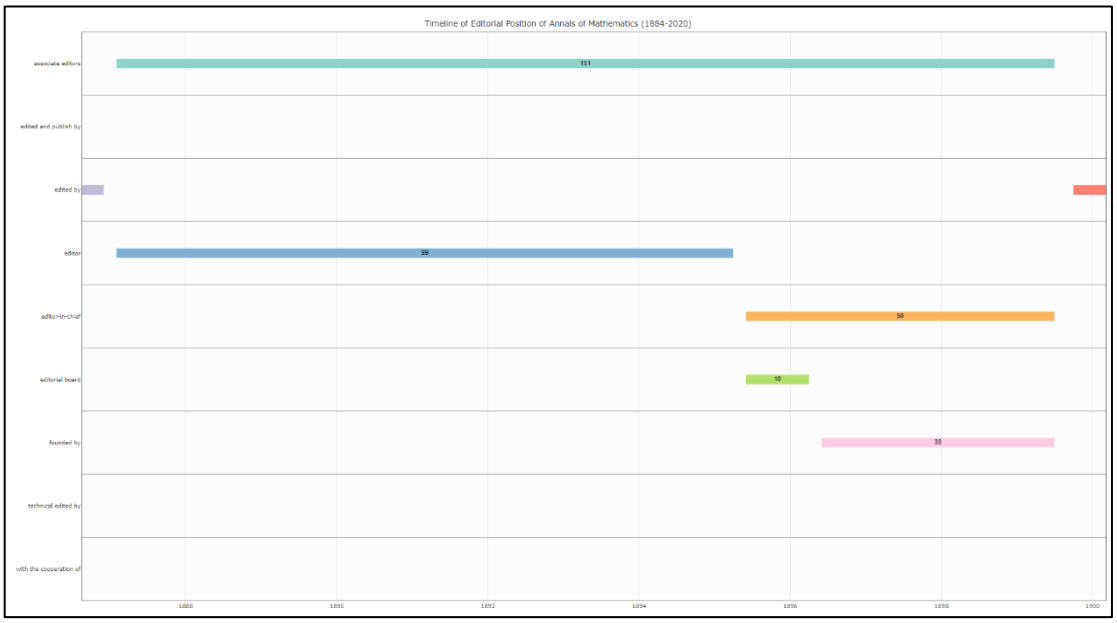


Figure 3.3 Timeline of editorial positions enlarged to the period the position “edited by” was not available the period (1886-1899).

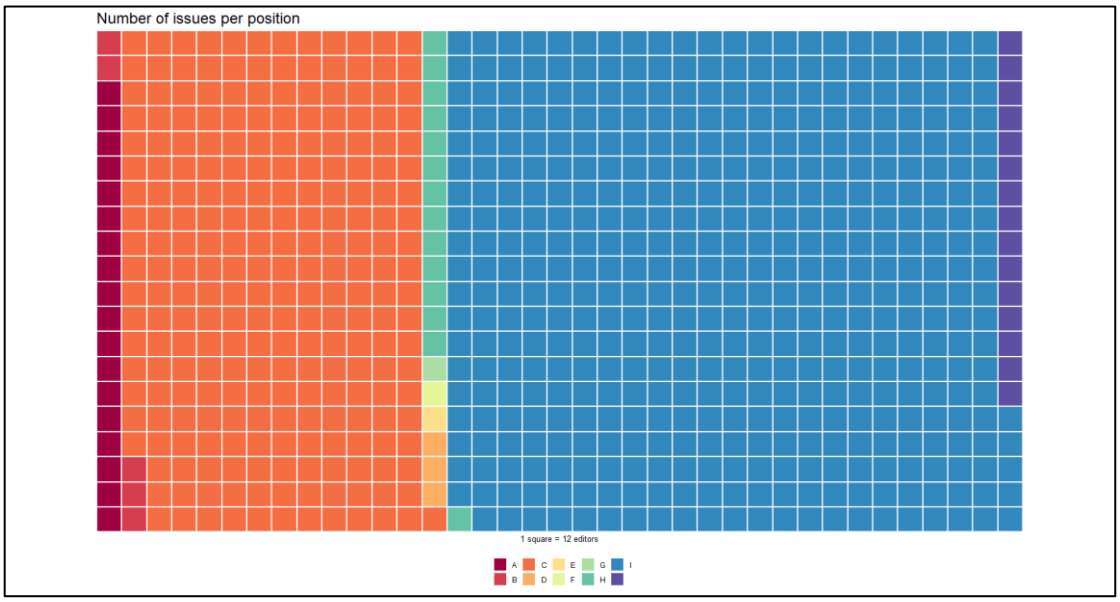


Figure 3.4 Waffle chart of the number of editorial positions per issue. Every square is equal with 12 editors

3.2 COMBINATIONS OF EDITORIAL POSITION

In this subsection will be analyzed the combinations of positions as they appear in the journal. Will be calculated their sizes and will be presented a timeline of the appearance of editorial positions combinations so that be understood how they change in the passage of publications. In total, our dataset has seven combinations of editorial boards that appeared during the publication of all issues. *Table 2* shows the summation of all combinations and their number. Also, every combination will be added in a unique group where this grouping will be used to display combinations in the timelines.

Group	Combinations	Total of appearances
A	edited and publish by	66
B	edited by	76
C	associate editors + editor	42
D	associate editors + editor-in-chief + editorial board	6
E	associate editors + editor-in-chief + founded by	12
F	edited by + with the cooperation of	389
G	edited by + technical edited by + with the cooperation of	174
Total		**765**

Table 1. Sums of total appearances of each combination in the dataset.

The sums of the combinations of editorial positions of the above table will be displayed in the following bar plot. The larger the bar the more impressions it has in our dataset.

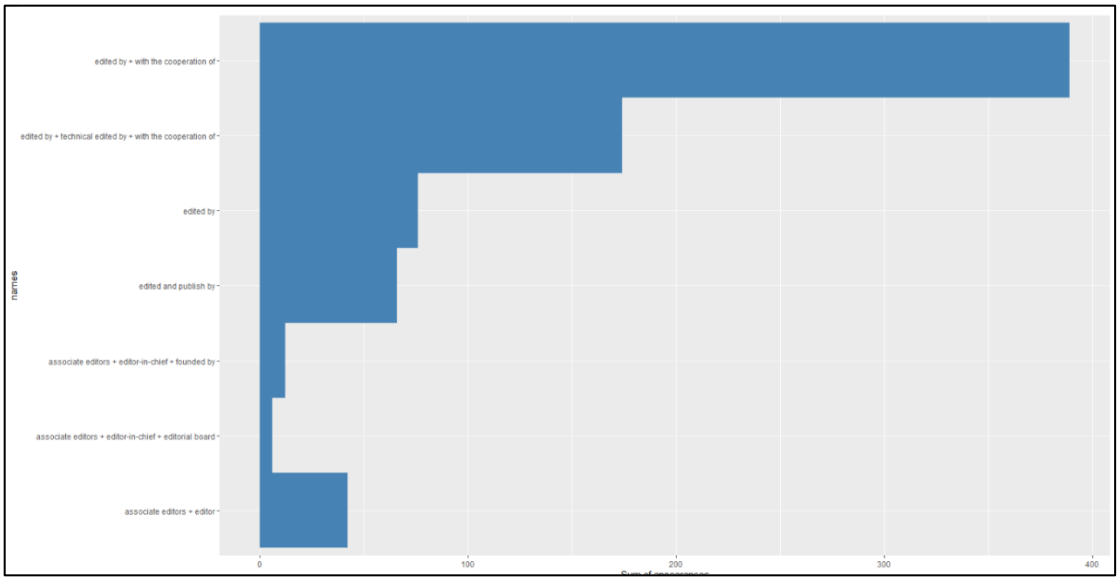


Figure 3.5 Bar plot of total number appearances per editorial position.

From the moment it has recorded the number of appearances of the roles, it would be interesting to record the order in which they appear. Here is a list of editorial

positions in chronological order in Annals of Mathematics in addition is recorded the number of appearances of every continuous combination in the corresponding interval. Finally, this list will be presented in a timeline bar plot.

Start Whole number	Combination Editorial position	start	End Whole number	end	Observations	Group
1	edited and publish by	1874-01-01	66	1883-11-01	66	A
67	edited by	1884-03-01	78	1886-12-01	12	B
79	associate editors + editor	1887-02-01	120	1895-04-01	42	C
121	associate editors + editor-in-chief + editorial board	1895-06-01	126	1896-04-01	6	D
127	associate editors + editor-in-chief + founded by	1896-06-01	138	1899-07-01	12	E
139	edited by	1899-10-01	202	1915-06-01	64	B
203	edited by + with the cooperation of	1915-09-01	405	1960-11-01	203	F
406	edited by + technical edited by + with the cooperation of	1961-01-01	411	1961-11-01	6	G
412	edited by + with the cooperation of	1962-01-01	414	1962-05-01	3	F
415	edited by + technical edited by + with the cooperation of	1962-07-01	564	1987-05-01	148	G
565	edited by + with the cooperation of	1987-07-01	565	1987-07-01	1	F
568	edited by + technical edited by + with the cooperation of	1988-01-01	583	1990-07-01	16	G
586	edited by + with the cooperation of	1991-01-01	765	2020-11-01	180	F

Table 2. Ordered combinations of editorial positions.

Below, is the display of table 3 in the timeline, so that the combination of editorial positions over time of journal publication. The main combination as referred to upper is “edited by + with the cooperation of” and can be understood from the timeline was a desired combination that several attempts were made to establish, until 1991 where it was established until 2020. For a large period proximately 29 years, the editorial position “technical edited by” was a supportive position to “edited by + with the cooperation of” for the periods 1962-1987 *(148 issues)*, 1961-1962 *(6 issues)*, 1988-1991 *(16 issues)*.

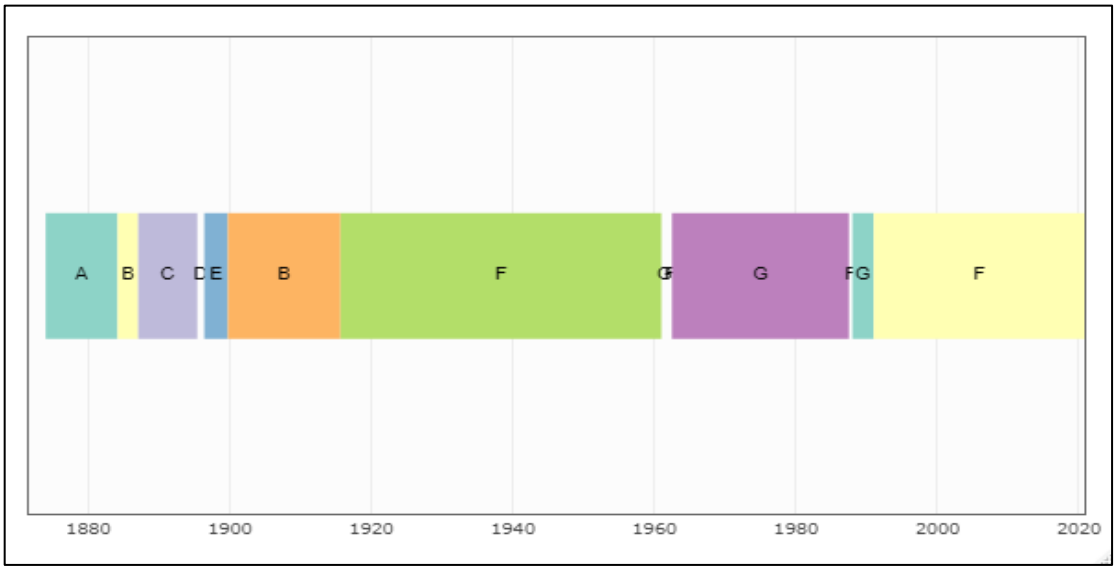


Figure 3.6 Timeline of editorial positions combinations that appear in issues.

3.3 PERCENTAGE OF EDITORIAL POSITIONS

The below table is created from table 1 and table 2. The column with the name “Total appears” is referred to the appearances of the position of editor in volumes, so that table 1. The percentage of appearance is resulting from:

$$\frac{\text{Number appearance of position}}{\text{Total appearances from all positions in volumes}}$$

From the other side, the column with the name “Combination editorial position appears” is referred to the appearances of combinations position of editor in volumes, so that table 2. The percentage of appearance is resulting from:

$$\frac{\text{Number appearance of combination}}{\text{Total appearances from all combinations in volumes}}$$

Editorial position	Total appears	Combination editorial position appears	Percentage of Total appears	Percentage of Combination editorial position appears
associate editors	60	0	7.84	0.00
edited and publish by	66	66	8.63	8.63
edited by	639	76	83.53	9.93
editor	42	0	5.49	0.00
editor-in-chief	18	0	2.35	0.00
editorial board	6	0	0.78	0.00
founded by	12	0	1.57	0.00
technical edited by	174	0	22.75	0.00
with the cooperation of	563	0	73.59	0.00
edited by + with the cooperation of	0	389	0.00	50.85
edited by + technical edited by + with the cooperation of	0	174	0.00	22.75
associate editors + editor-in-chief + founded by	0	12	0.00	1.57
associate editors + editor-in-chief + editorial board	0	6	0.00	0.78
associate editors + editor	0	42	0.00	5.49

Table 3. Percentage of editorial positions that occupy in the dataset and the percentage that occupy the combinations in the dataset.

From the table can be seen, editorial positions “edited by” and “edited and publish by”, those two positions that are displayed in two columns, as an independent position, and as a combinate position. That values are both columns because there are periods that were the only editorial positions in volume, so are recorded as a combination and as the editorial position of volume. Below there is a pie chart about the percentage of editorial board members, it can be considered that the main structure of the editorial board is” edited by + with the cooperation of “with a percentage of 50,59%. In addition, from the comparison of ***Table 3*** and ***Figure 3.10*** that “technical edited by” is a supporting role. Because from 1915 to 2020 appears the main

combination of" edited by + with the cooperation of "and" edited by + technical edited by+ with the cooperation of "periodically and the total percentage of appearance of that combo is 22,22%.

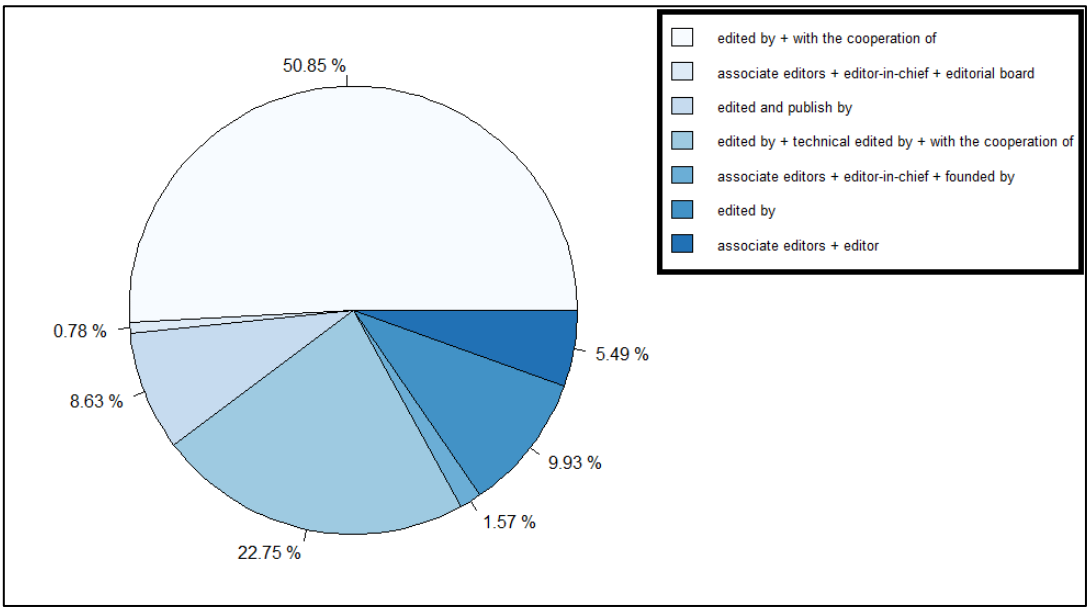


Figure 3.7 Pie chart of combinations appearances in the journal.

For purposes of being able to evaluate and classify the nature of the editorial roles of our dataset, a bar plot will be created with the percentage of roles appearing throughout the publication period of the Annals of Mathematics. The percentages occupy from ***Table 4***, are the percentage that each editorial role occupies in the total of all editions of the magazine

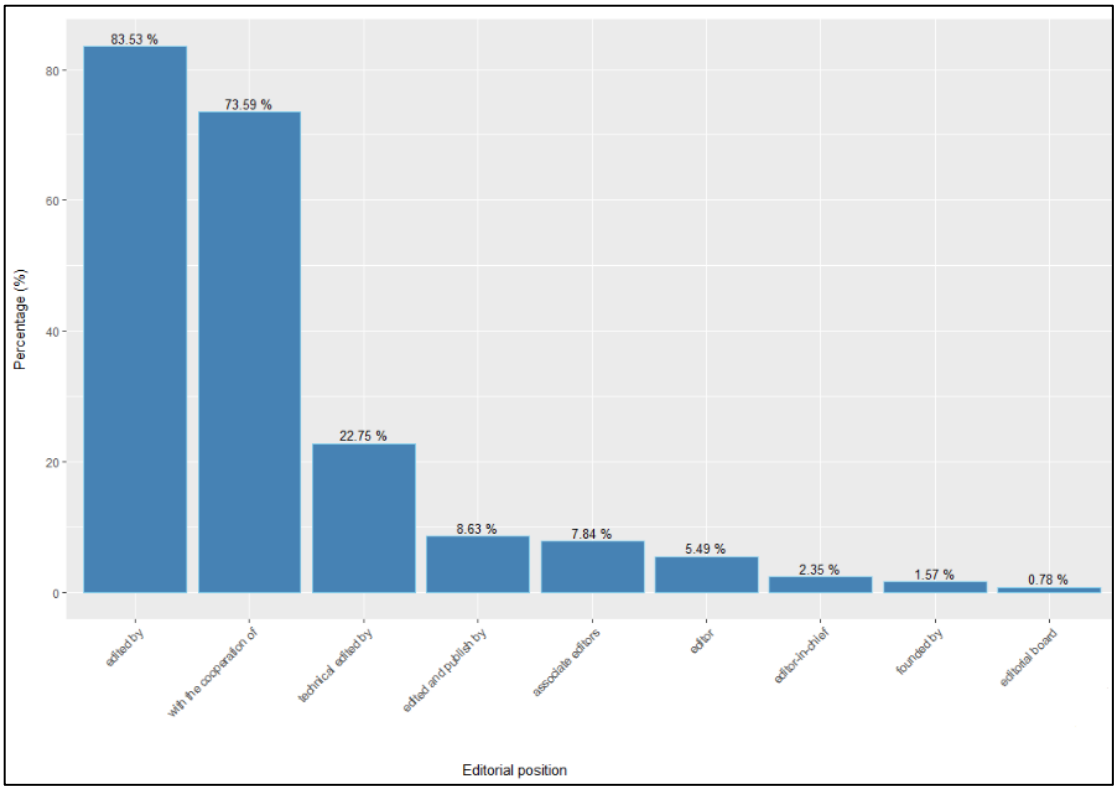


Figure 3.8 Bar plot of percentage of appearance that has every role in journal.

CHAPTER 4
EDITORIAL BOARD SIZE ACROSS TIME

In this chapter, there will be a check for the diversity that exists in the journal as far as the positions that contribute to it and how this is formed over the years while being examined its changes. In addition, it has been observed that some roles reappear at intervals of time in our dataset and others simply cease to exist. Also, there are editorial positions that should be defined as main and supportive as has been done in the previous chapter using selection criteria. In this chapter they need to be clarified:

- Examine the number of editors in editorial boards per issue and the amount that occupy each role per issue, the same process will happen per year, with the difference must be referred to average values
- Examine the number of editors in editorial boards every decade with distribution analysis
- Investigate if there are changes to editorial boards across time and what factors contributed to their appearances
- Categorize editorial positions to mains and supportive using criteria and comparing them
- Categorize changes based on the behavior of editorial board size changes and visualize them on indexes.

4.1 QUANTITATIVE RESEARCH EDITORIAL POSITIONS

Initially, a table will be created which will report in each issue the total number of editors of each role that took part as well as the summation of the editors of all the editorial positions of each issue. After that, it will be known the percentage occupied by each editorial position per issue. A similar process will be followed to other tables for annual analysis but is easy to understand that will be applied to editorial positions the average values per year since I have more than one value for each year. In the following stacked plot are visualizing the summation of editorial positions depending on the issue of publishing. Many questions arise from the diagram, while in the beginning, it starts with a role that the journal contributes to, and the need arises for the creation of new editorial positions.

The first graph represented the total editorial positions and the number of editors of each issue, are following the number of editors of each role per issue. It can be concluded that the roles that are starring are edited by and with the cooperation of, as they are timeless.

Figure 4.2 is presented a stacked plot of the editorial board per issue, which is the same as the first plot of ***Figure 4.1*** with the difference that is displayed in all the roles of the issue that participated. The percentages of editorial position per issue are visualized in the percentage stacked plot in ***Figure 4.3***. The same process will be for annual analysis.

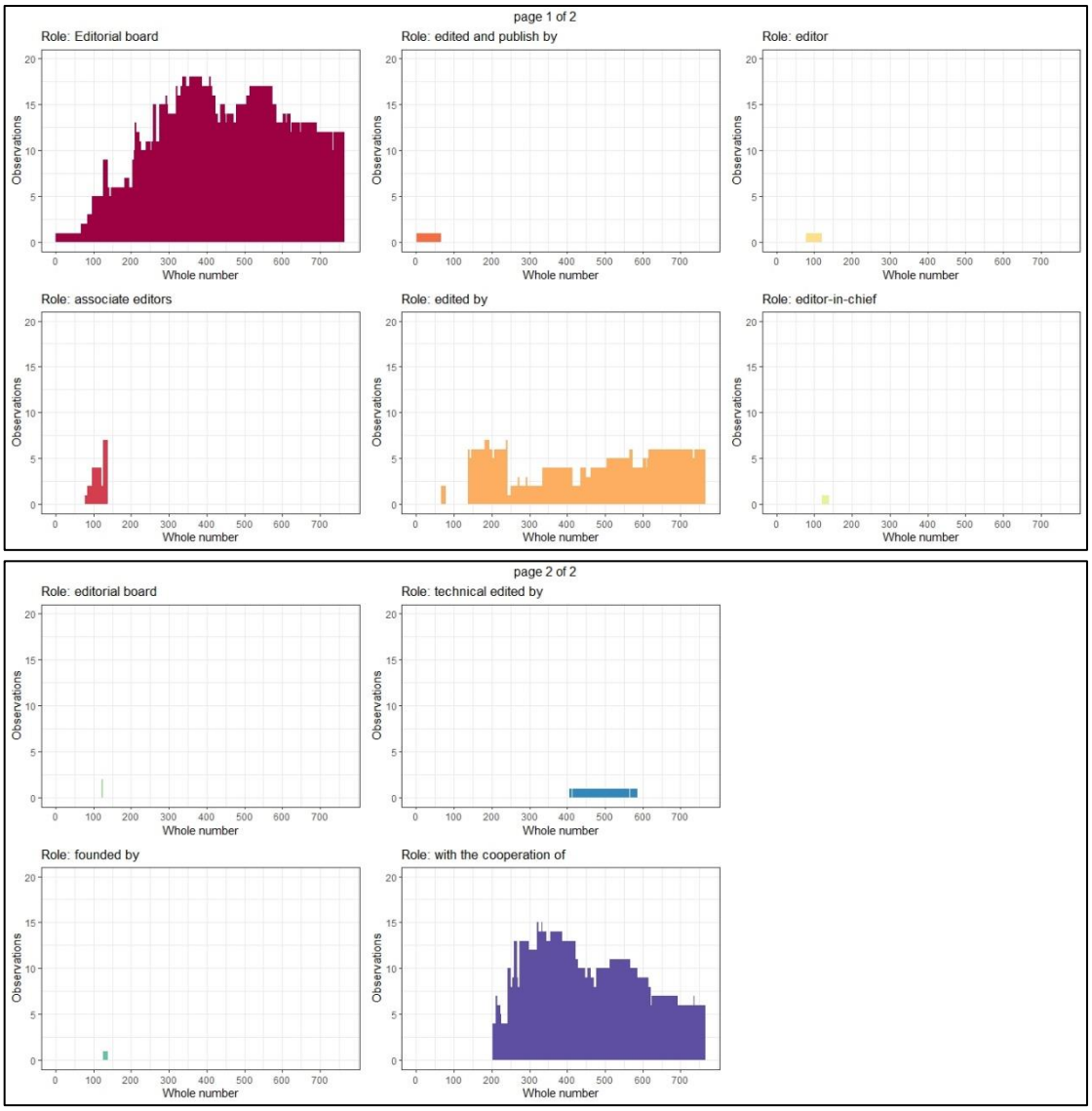


Figure 4.1 The number of editors for each editorial position per issue.

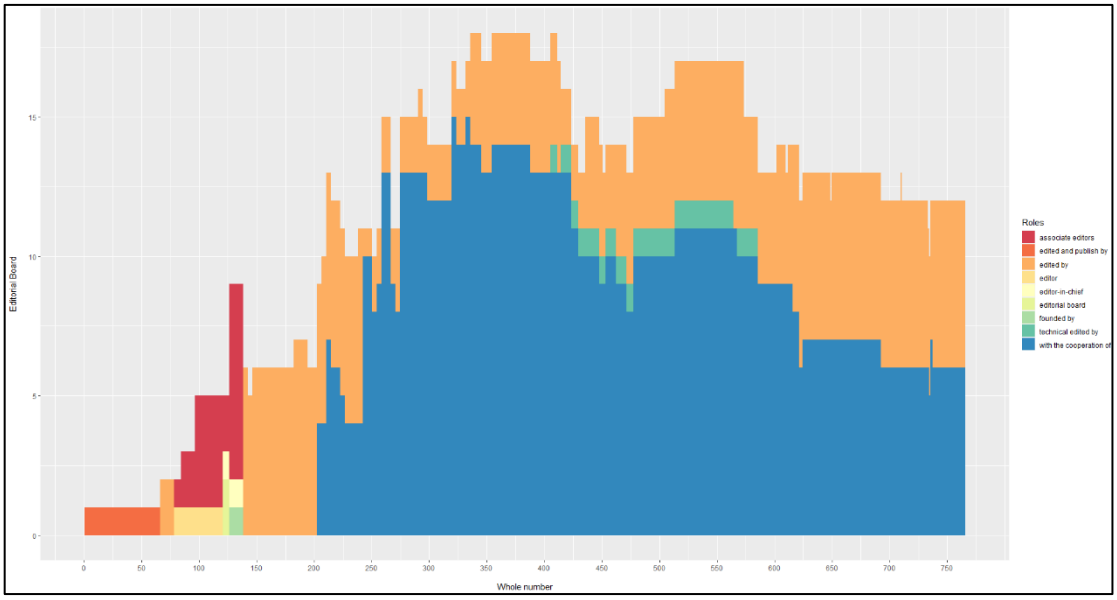


Figure 4.2 The total number of editorial positions per issue.

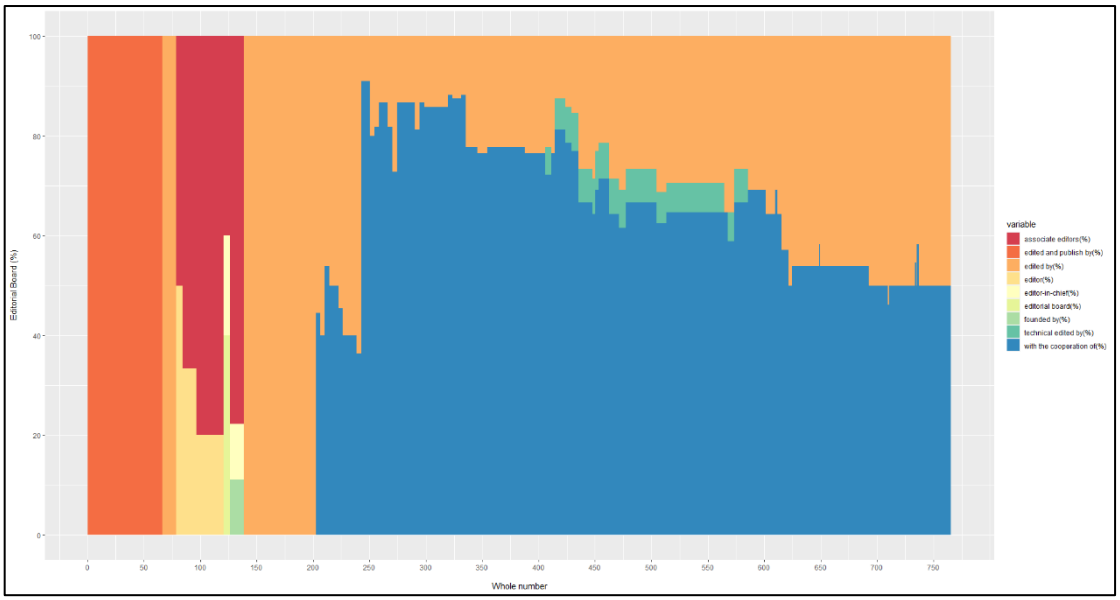


Figure 4.3 The percentage of the total number of editorial positions per issue.

Figure 4.4 The number of editors for each editorial position per year.

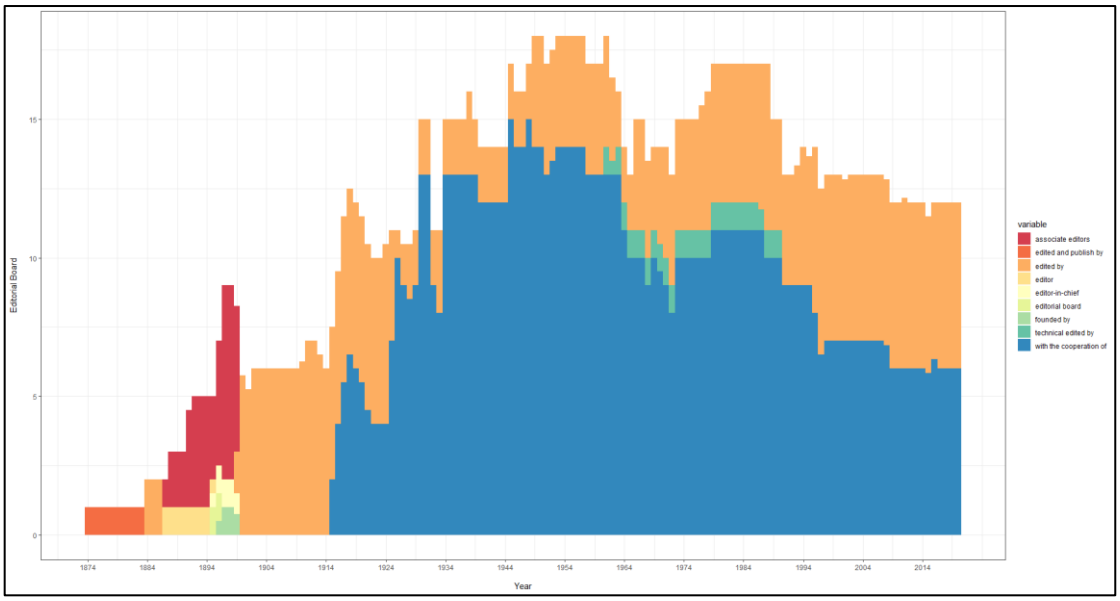


Figure 4.5 The total number of editorial positions per year.

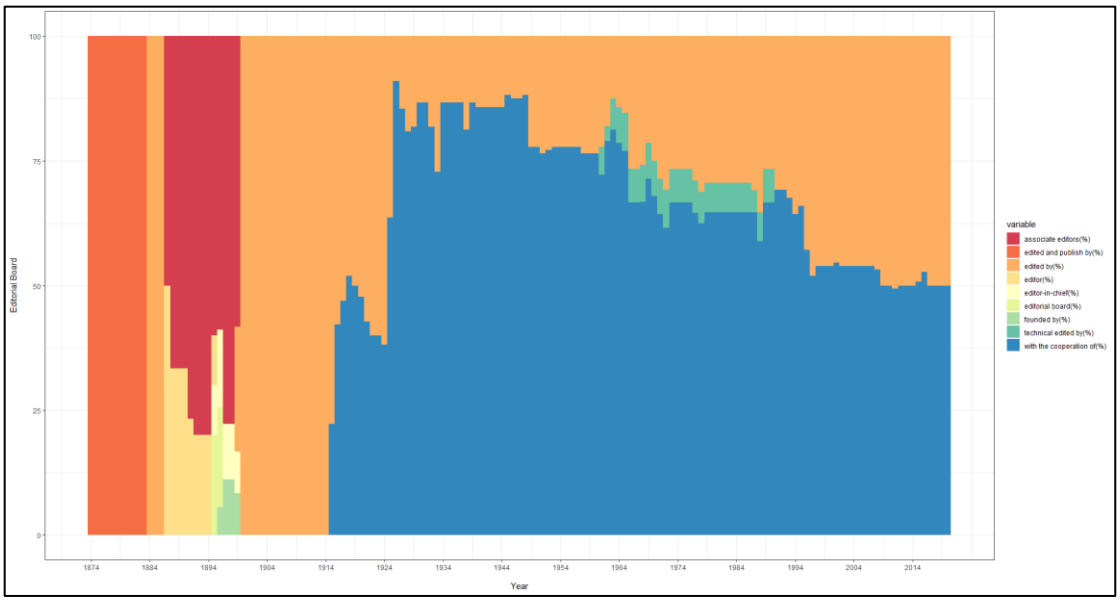


Figure 4.6 The percentage of the total number of editorial positions per year.

4.2 EDITORIAL BOARD SIZE ANALYSIS PER DECADE

Our dataset has ranged from 1884 to 2020, to be able to analyze the size of editorial boards, the annual is not a guide to conclude, for this reason, it will be analyzed per decade. In total, there are 16 decades to our dataset between the upper range of years. For the examination will be used boxplots.

A boxplot is a standardized way of displaying the distribution of data based on a five-number summary ("minimum", first quartile (Q_1), median, third quartile (Q_3), and "maximum"). It can tell you about your outliers and what their values are. It can also tell you if your data is symmetrical, how tightly your data is grouped, and if and how your data is skewed. *(Towardsdatascience.com)*

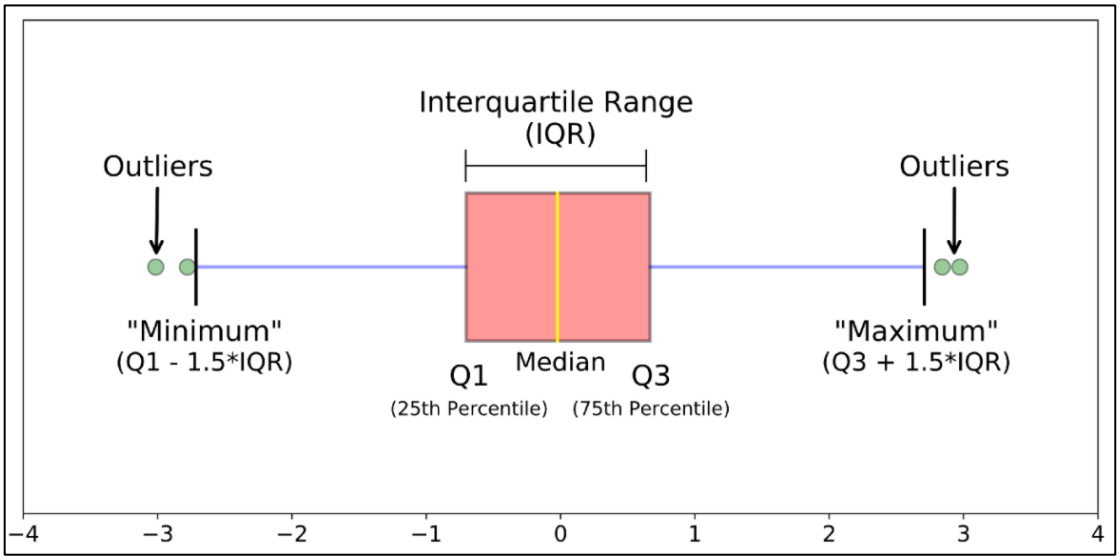


Figure 4.7 Different parts of a boxplot

- **Median (Q_2/50th Percentile)**: the middle value of the dataset.
- **The first quartile (Q_1/25th Percentile)**: the middle number between the smallest number (not the "minimum") and the median of the dataset.
- **Third quartile (Q_3/75th Percentile)**: the middle value between the median and the highest value (not the "maximum") of the dataset.
- **Interquartile range (IQR)**: 25th to the 75th percentile.
- **Whiskers (shown in blue)**
- **Outliers (shown as green circles)**
- **"maximum"**: Q_3 + 1.5*IQR
- **"minimum"**: Q_1 -1.5*IQR

What defines an outlier, "minimum", or "maximum" may not be clear yet. The next section will try to clear that up for you.

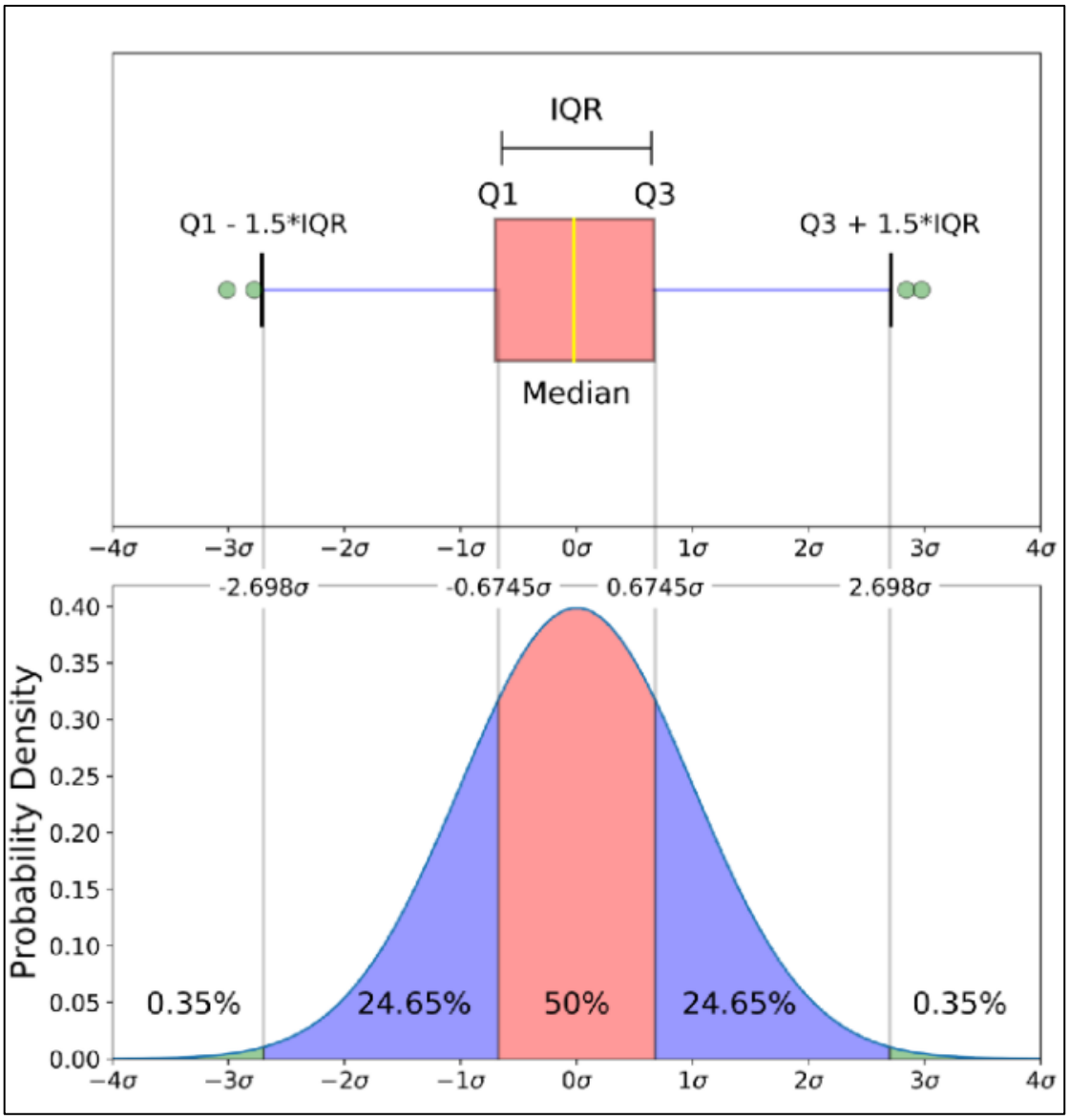
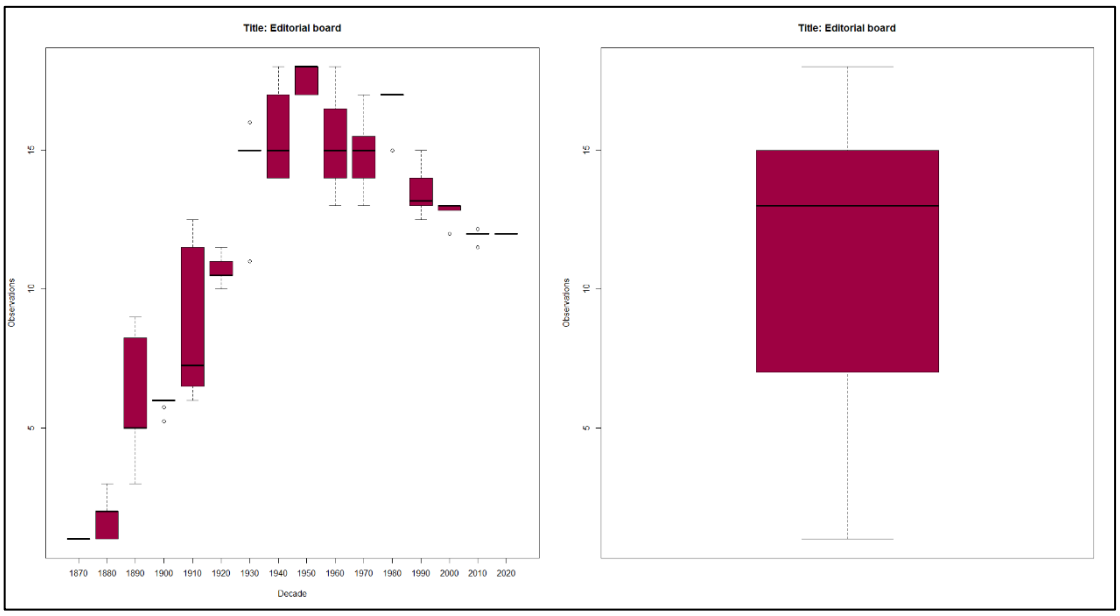


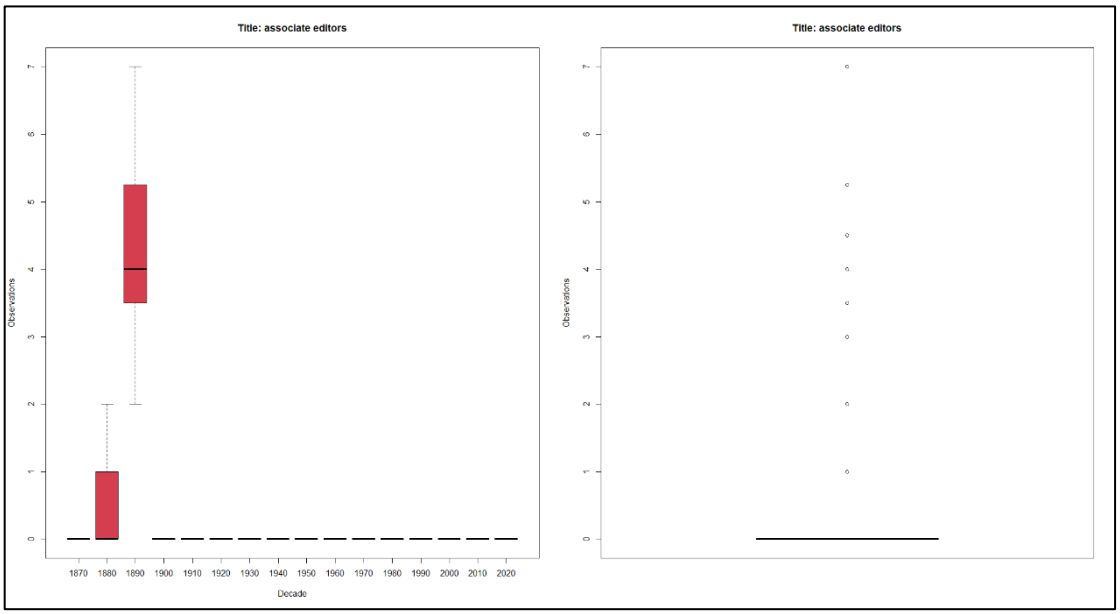
Figure 4.8 Comparison of a boxplot of a nearly normal distribution and a probability density function (pdf) for a normal distribution

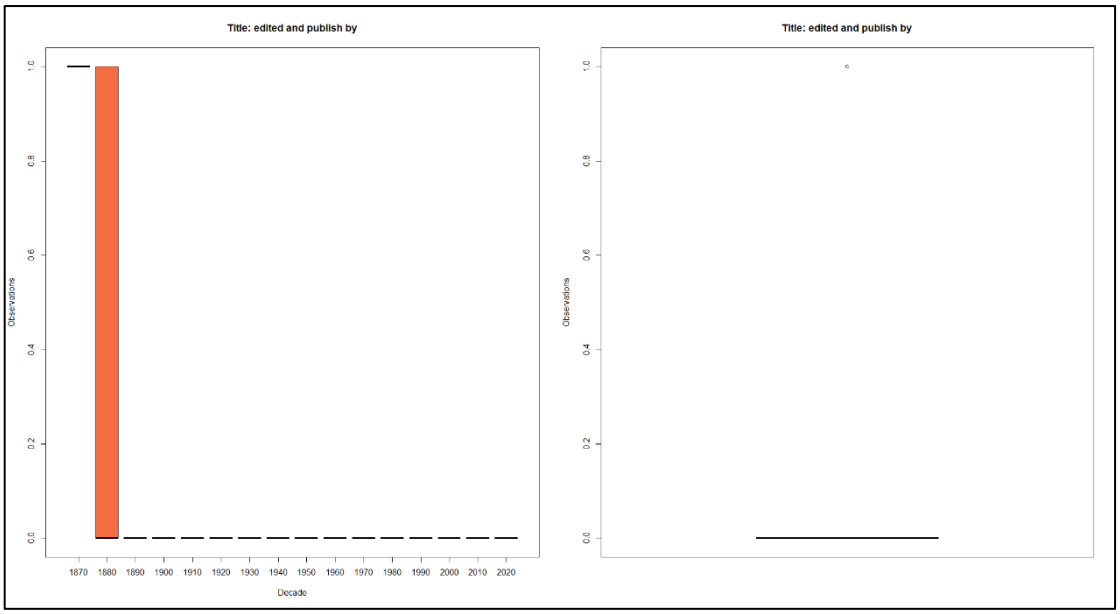
The image above is a comparison of a boxplot of a nearly normal distribution and the probability density function (pdf) for a normal distribution. The reason why it is presented to you this image is that looking at a statistical distribution is more commonplace than looking at a box plot. In other words, it might help you understand a boxplot. This section will cover many things including:

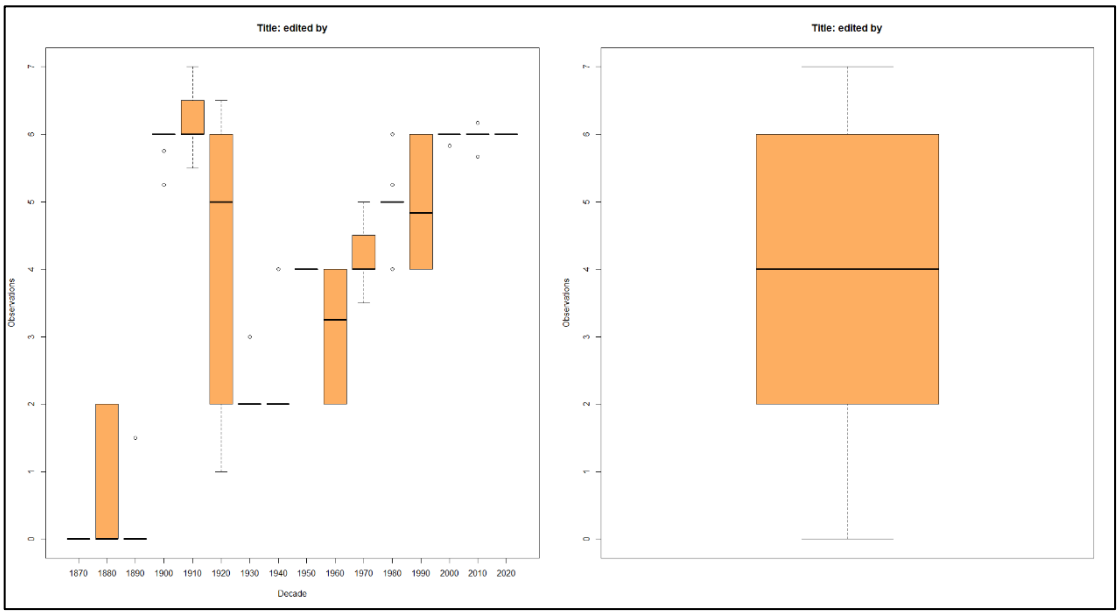
- How outliers are (for a normal distribution) .7% of the data.
- What a “minimum” and a “maximum” are.

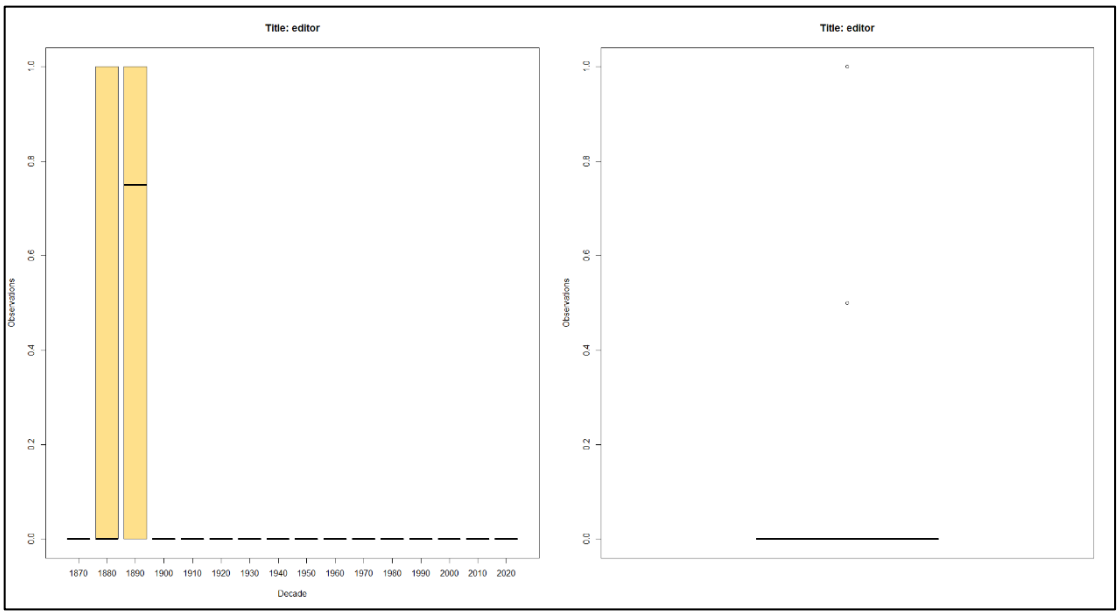


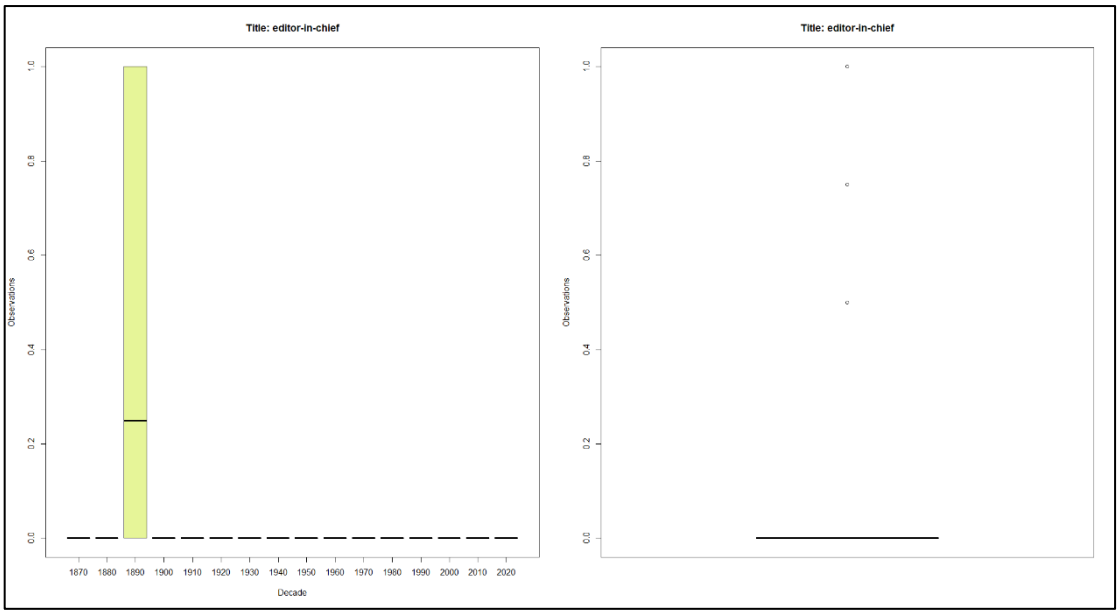
There will be a description of the editorial board from the boxplots on the left, so that can be imprinted a complete description of the number of contributions of the journal. From 1870 to 1920 there is a continuous increase in the number of rates, must be referred that from 1930 to 1940 and 1960 to 1970 was observed stability to median values of each decade that period around to 15 editors, except in the 1950s when the number of editors in editorial board the median price skyrocketed to 18 editors, which also is the maximum number of editorial board positions that is observed. After the 1980s can be concluded from boxplots there was a falling glide to editorial boards per decade with the final number of editorial boards to settle to 12 editors. At the right side, is the boxplot where is displayed for all the values of each role, in this case, the total number of editors. In ***Figure 4.9*** are all the boxplots of each editorial position.

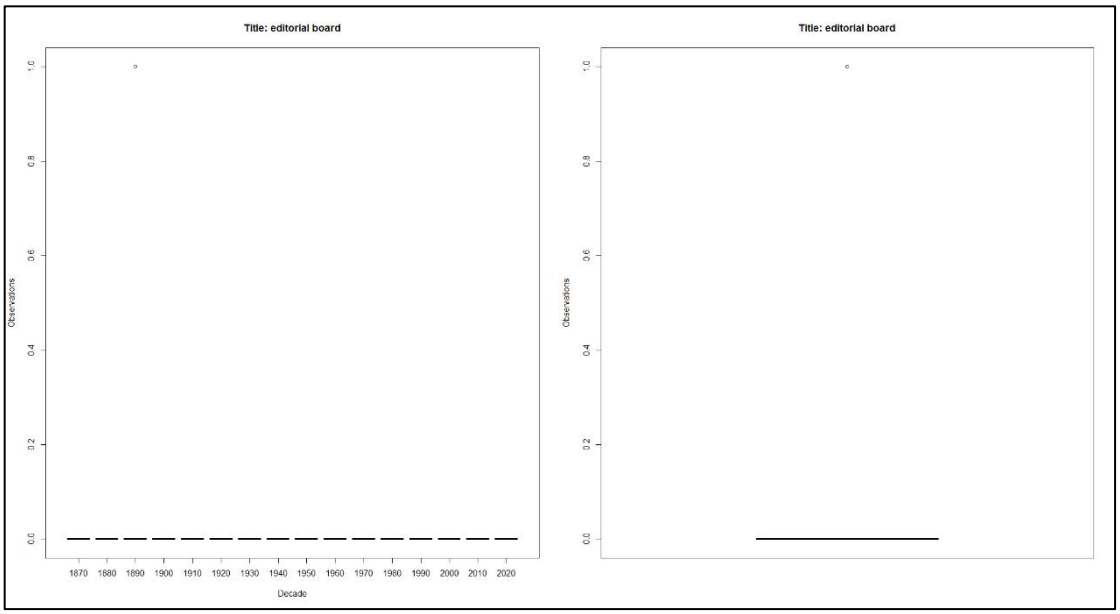


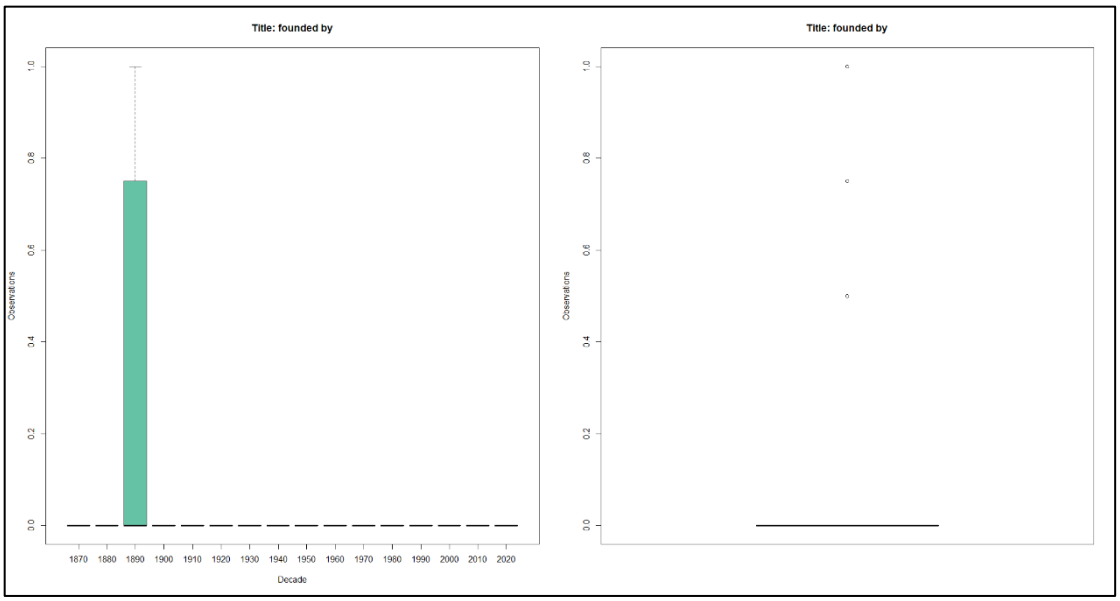

Title: edited and publish by
Observations
Decade
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
0.0 0.2 0.4 0.6 0.8 1.0
Title: edited and publish by
Observations
0.0 0.2 0.4 0.6 0.8 1.0


Title: edited by
Observations
Decade
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
0 1 2 3 4 5 6 7
Title: edited by
Observations
0 1 2 3 4 5 6 7


Title: editor
Observations
Decade
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
0.0 0.2 0.4 0.6 0.8 1.0
Title: editor
Observations
0.0 0.2 0.4 0.6 0.8 1.0


Title: editor-in-chief
Observations
Decade
0.0
0.2
0.4
0.6
0.8
1.0
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Title: editor-in-chief
Observations


Title: editorial board
Observations
Decade
0.0
0.2
0.4
0.6
0.8
1.0
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Title: editorial board
Observations


Title: founded by
Observations
Decade
0.0
0.2
0.4
0.6
0.8
1.0
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020
Title: founded by
Observations

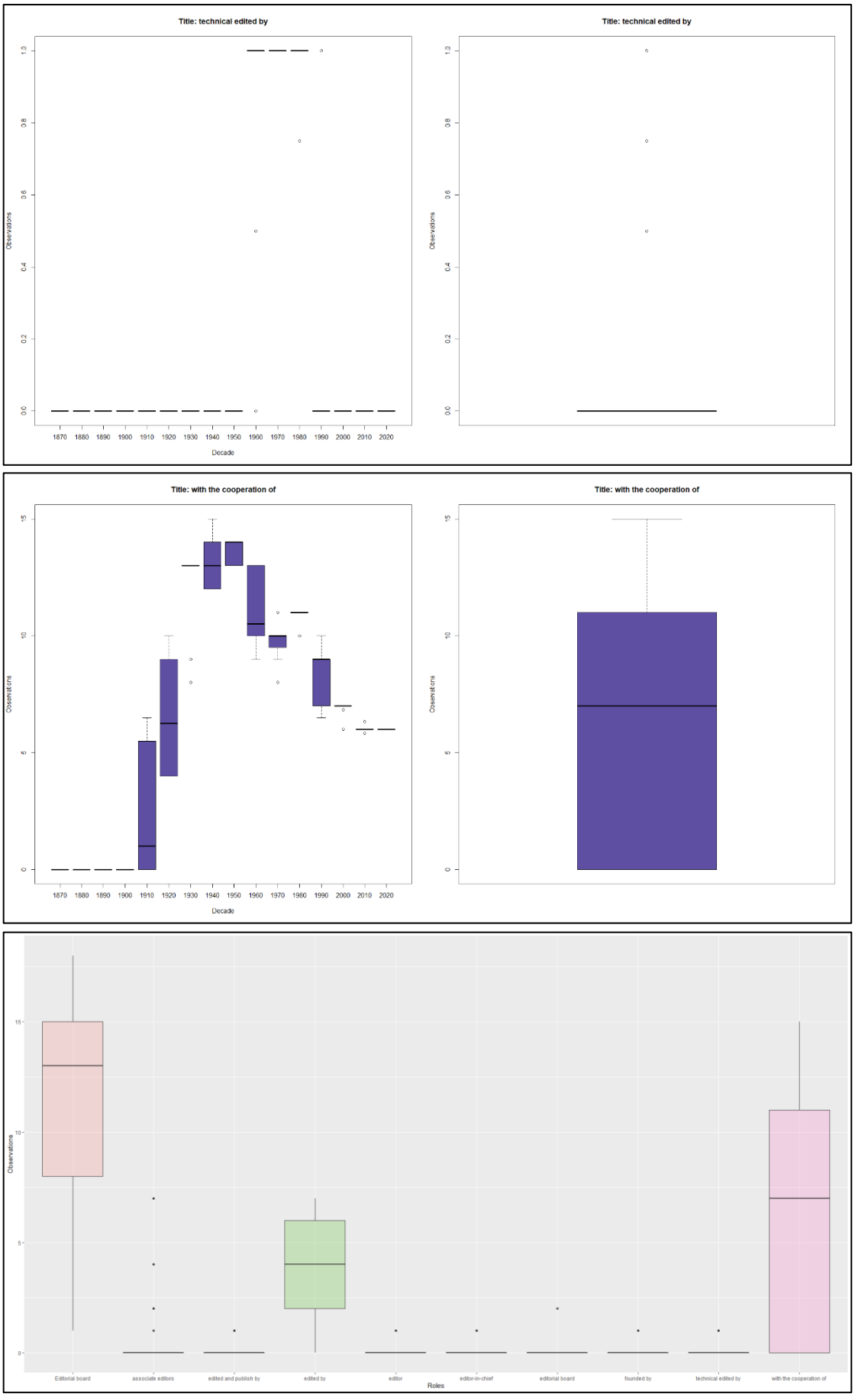
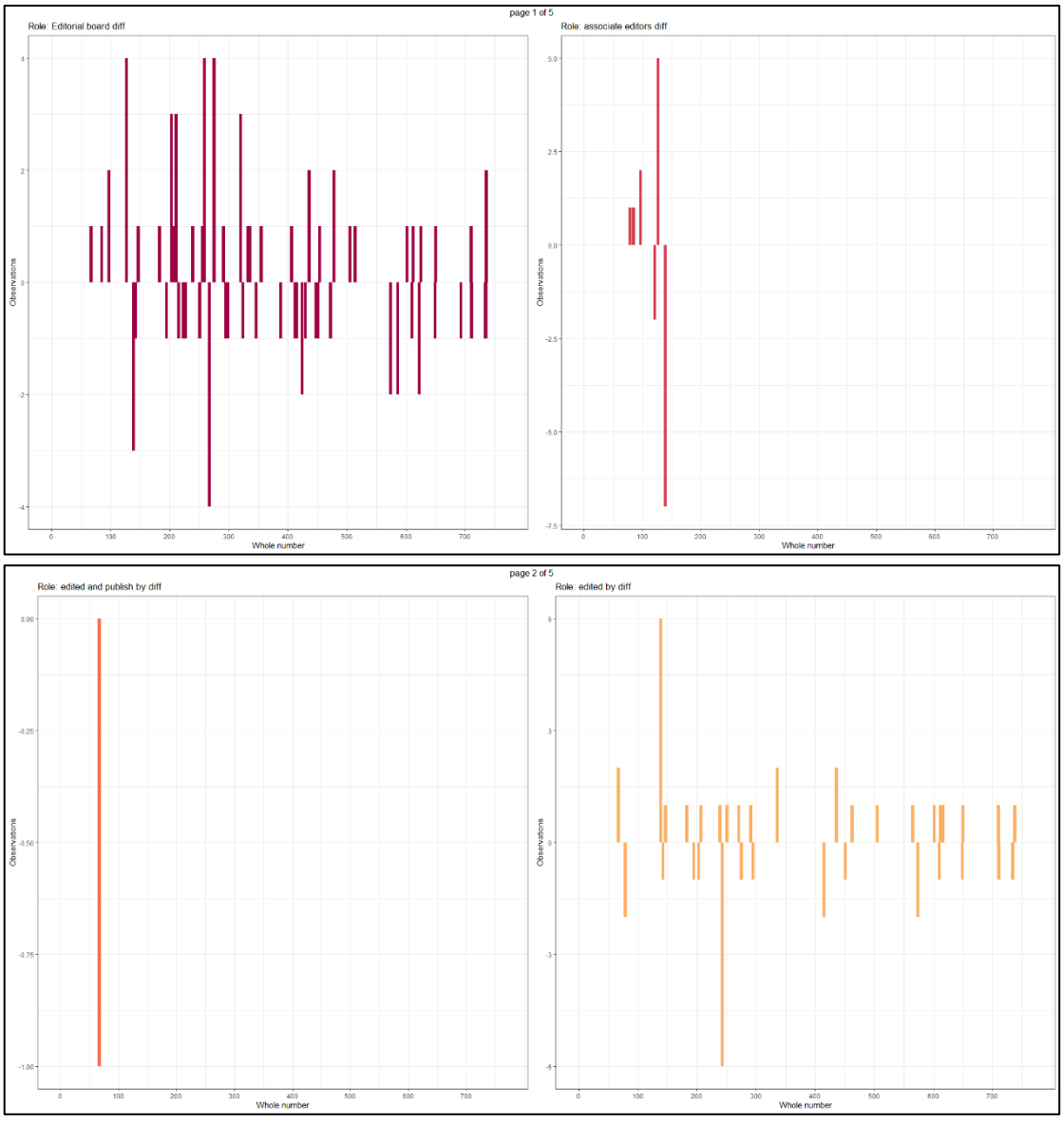


Figure 4.9 Boxplots per editorial positions of Annals of Mathematics.

4.3 CHANGES TO EDITORIAL BOARDS ACROSS TIME

From the previous sub-paragraph's conclusions emerged regarding the changes of the editorial boards that the journal had during its publication from the first issue until today. This paragraph will examine each role about the changes that have occurred into editorial boards across time, to be able to achieve that will be created a table for each role which compare the previous editors' number of issues with the current and will be made line graphs of difference with a purpose to be easier to be understood for the reader. Graphs will follow in which you display for each editorial position the number of the difference from the previous issue and emerges the difference graph per issue. When the bar goes up it means that there is an increase in the issue under consideration, compared to the last known one. Respectively, when a drop-off is observed, a comparison is made with the immediately preceding one and the drop of the bar occurs. Finally, when the bar meets zero there is no difference in the number of editorial positions from the previous issue.

The same graph will be applied and per year must be recapitulated because have been published more than one issue the same year at the annual graph is the average values that concluded in the year. For that reason, the values of differences can be a decimal number at the annual graph, unlike of issue graph which values is only integer.



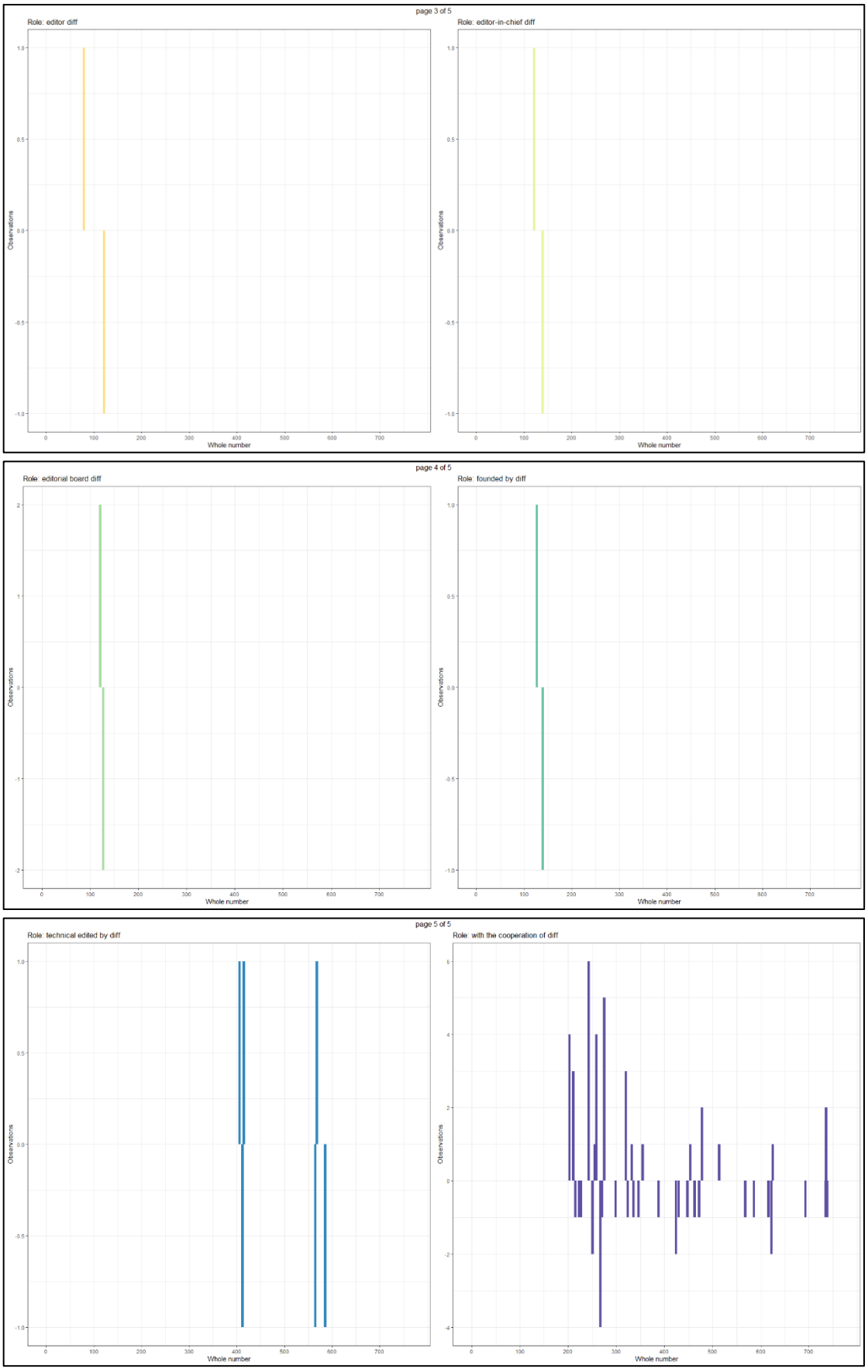


Figure 4.10 Graphs of difference per issue.

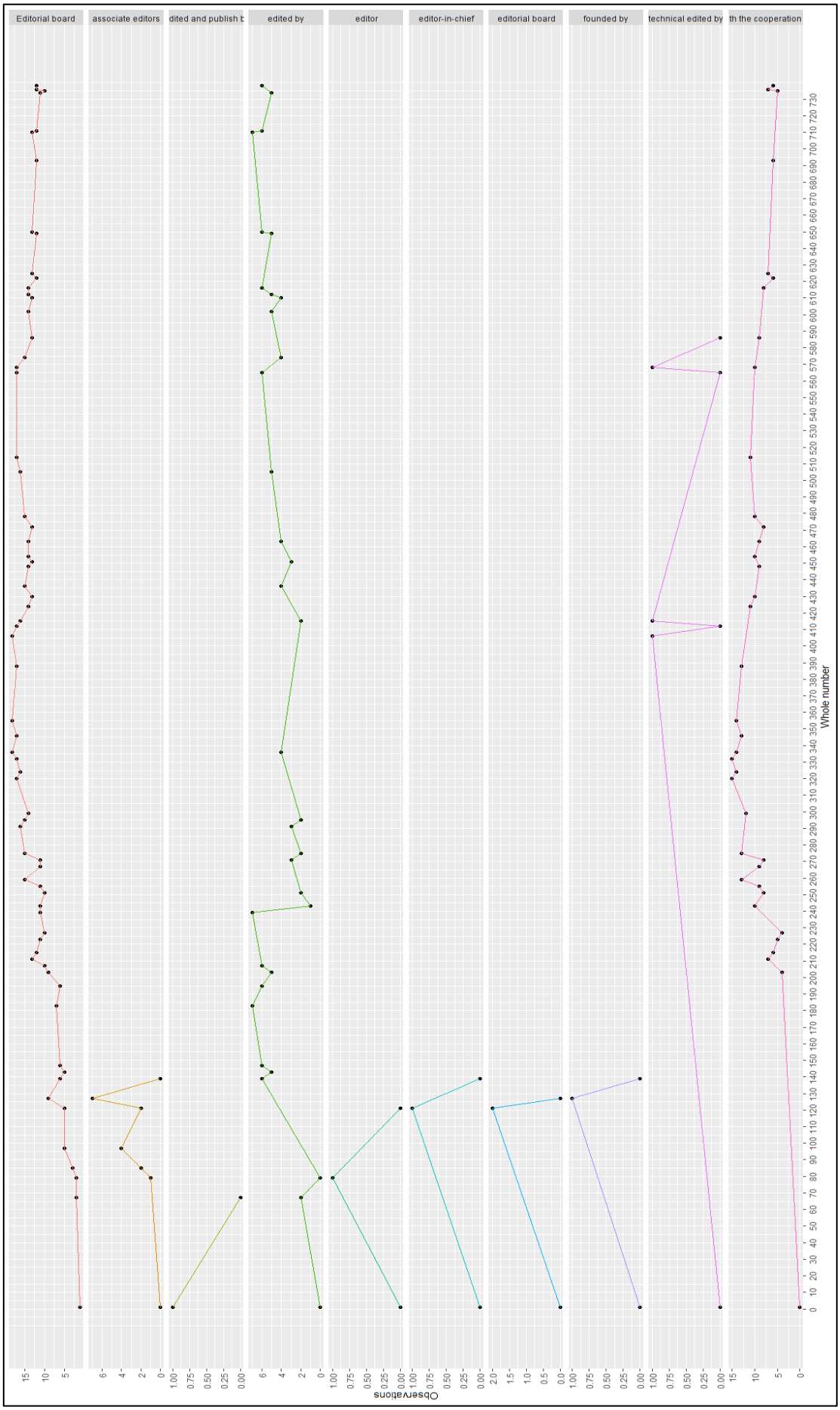


Figure 4.11 The changes to editorial boards across issues and per editorial position.

In ***Figure 4.11*** all the changes of the editorial boards are displayed all the nodes are the number of editors in a specific issue, which is referred to as the x-axis. The graphs are separated equally with the number of editorial positions and the one below the other because the x-axis is common. If a node ends at zero and does not reappear later, it means that it ceases to exist. Also, between nodes the values are equal, for example from issues with whole numbers 1 to 406 the number of editorial positions “technical edited by” is 0.

In the section editorial boards, all the nodes are the changes in the journal, and respectively where there is fluctuation the corresponding change in the editorial position is displayed. In some cases, although there are changes, it is observed that the number of editors in the journal remains stable, that happens because as the number of editors in a particular position increases with the same number decreases in a different position. For that reason, initially, it was a problem for the creation of this graph, and for this, the issues that were changed in all the editorial positions were recorded, and then the editorial board was added which reflects the total change in the journal. About graph, the conclusions that can be drawn from the graph are which editorial positions changed and shaped the editorial board in this way.

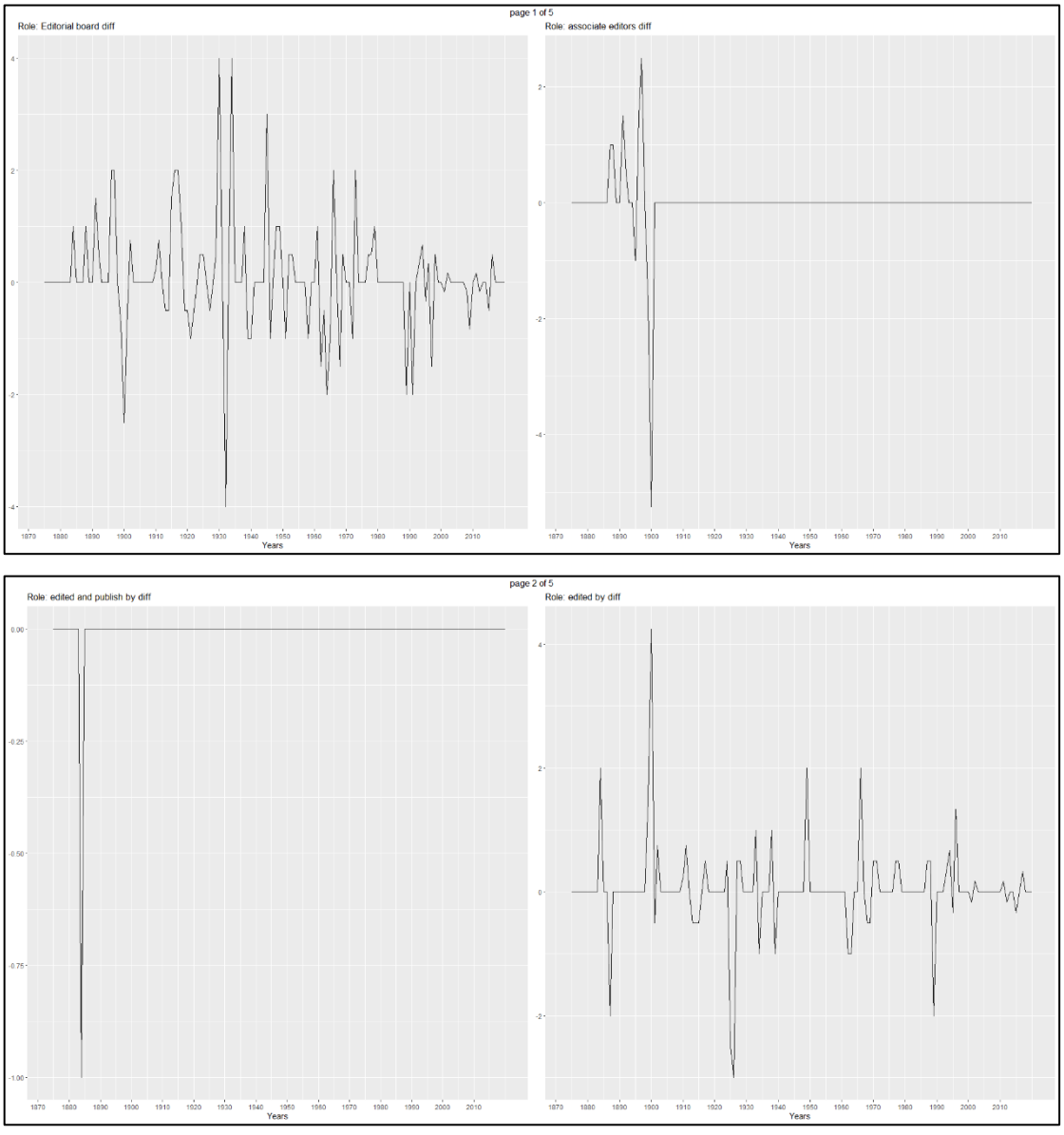


Figure 4.12 Graphs of difference per year.

Below will be presented the results of different editorial boards per issue and per year using the graph of cumulative sum chart in the left and the right barplot with the number of observations, at the same time the table of results will be displayed to understand the behavior of the results. On the tables are included the differences observed between issues, and the number of totals per observation. To be able to structure the cumulative sum chart must be calculated and the sums of observations. Apart from the sums, however, the percentages will also be calculated in one column the observations receive their coverage percentage and in the other the sums of the percentages. In the graph with blue bars is represented the number of differences per issue, while in the beige graph the number of differences per year. In total, the journal has as referred to before 759 issues and is active for 147 years. Must be paid special attention to the realization of the annual chart as each year had more than one issue so for the correct observation of each year, it was considered necessary to record the average number of editors in editorial boards for each year. In this way the integrity of the numbers was not ensured so that they could be analyzed, for this reason, the numbers were rounded and table 5 was created and from there the graph in ***Figure 4.14***. Both graphs are easy to understand editorial boards of issues between them they do not have frequent fluctuations because the percentage of observations are bigger to zero differences.

	Editorial board diff	Observations	Sums Observations	Frequency(%)	Sums Frequency(%)
1	-4	1	1	0.132	0.132
2	-3	1	2	0.132	0.264
3	-2	4	6	0.527	0.791
4	-1	23	29	3.030	3.821
5	0	700	729	92.227	96.047
6	1	20	749	2.635	98.682
7	2	4	753	0.527	99.209
8	3	3	756	0.395	99.605
9	4	3	759	0.395	100.000

Table 4. Grouping results of differences in Editorial boards per issue.

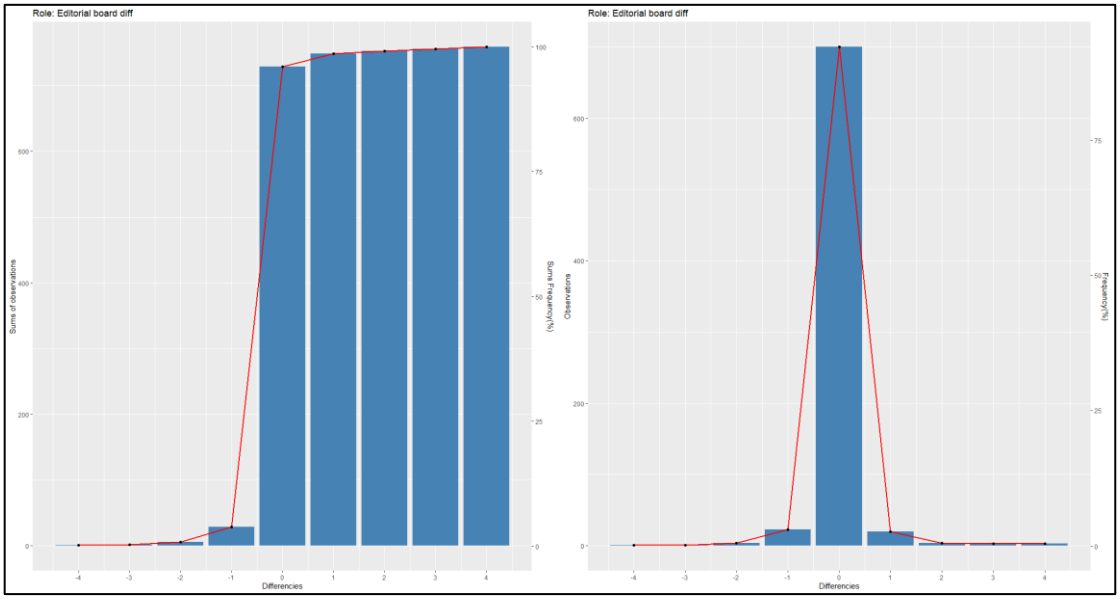


Figure 4.13 Graph of difference of editorial boards per issue.

	Editorial board diff	Observations	Sums Observations	Frequency(%)	Sums Frequency(%)
1	-4	1	1	0.68	0.68
2	-2	7	8	4.76	5.44
3	-1	10	18	6.80	12.24
4	0	107	125	72.79	85.03
5	1	11	136	7.48	92.52
6	2	8	144	5.44	97.96
7	3	1	145	0.68	98.64
8	4	2	147	1.36	100.00

Table 5 Grouping results of differences in Editorial boards per year.

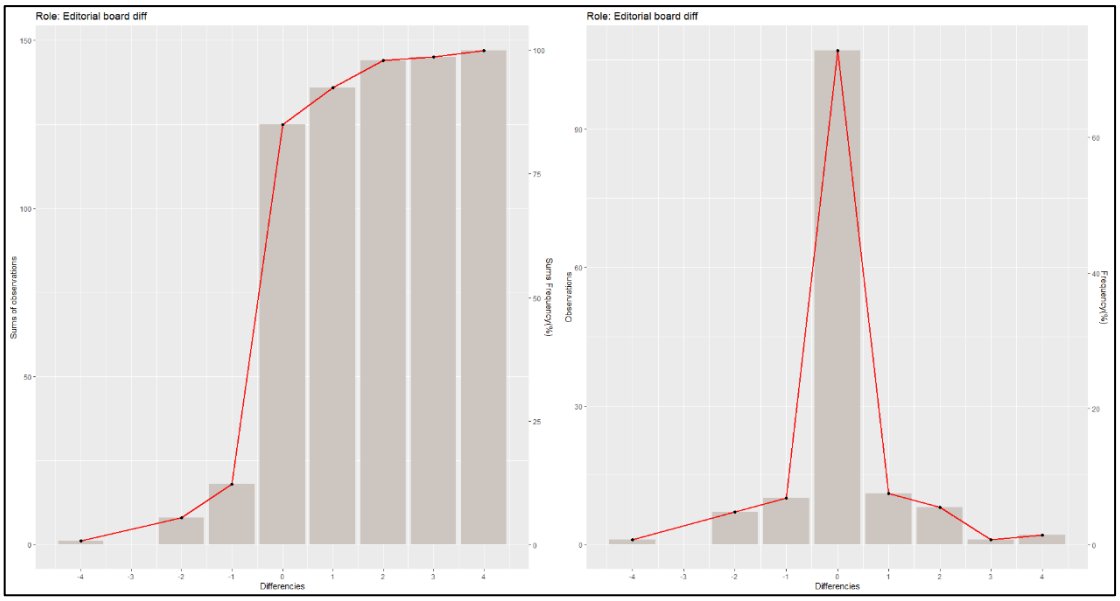


Figure 4.14 Graph of difference of editorial boards per year.

4.4 CATEGORIZE EDITORIAL POSITIONS

Over the years has been noticed from the previous paragraphs that the editorial board is constantly changing from year to year and from issue to issue. This paragraph will summarize the editorial positions in 3 categories (where they will be analyzed below) and will be done analysis every decade. The 3 categories of editorial positions are:

- Main roles
- Supportive roles
- Other roles

By the meaning “supportive role” sometimes refer to a lower-level editor who will use their reasonable endeavors to provide a senior person or group of editors with the necessary support. In our dataset supportive roles are considered associate editors, with the cooperation of and technical edited by. On the other side, main editorial positions can be considered edited by, editor, editor-in-chief and edited and published by. At this phase cannot be categorized the editorial positions founded by an editorial position as the first is a neutral term, while the second is an unspecified term. It can be understood that per issue there are fluctuations in the total number of roles from both sides.

The editorial boards have been analyzed from the 1870s until 2020, and as mentioned, the editorial positions have been converted into 3 categories, were mentioned above. The graph that is presented records the percentage occupied by each category per decade and the number of editorial boards within the decade. It should be noted that the journal has been republished under a different publishing name *“The Analyst”* and as *“Annals of Mathematics”* had 2 republished periods. More specifically:

- 1874-1883, as The Analyst where volumes range from 1 to 10.
- 1884-1899, as The Annals of Mathematics volumes, ranging from 1 to 12.
- 1899-2020, as The Annals of Mathematics volumes, ranging from 1 to 192.

As “The Analyst”, it has been observed that there is an editor to each issue who is active in the editorial position "edited and published by" as can be seen from ***Figure 4.5***. Therefore, since the release period of “The Analyst” was from 1874-1883 and the role has been considered as main occupies 100% in the decade of the 1870s. After “The Analyst”, with the publishes of 1st period of “Annals of Mathematics” new roles began to be established and the first one supportive role has appeared. The 1st period of “Annals of Mathematics” is until 1899, so until then is observed that the supportive roles are more than main roles while at the same time the number of editors in the specific decade is significantly increased compared to the decades which have examined. After the 1900s is started the 2nd period of publishing and is shown by the graph that main roles are included only, and the total number of editors is reduced from the previous decade. From the 1910s to the 1940s supportive roles are reappeared and they occupy a larger percentage from decade to decade. In the 1950s is the peak of editorial positions breaking the barrier of 1000 editors, with 22,66% main roles and 77,34% supportive roles. Finally, from 1960 to 2020 the percentage gap is increasingly converging, approaching the equilibrium of crowds between positions. The decade of 2020 will not be commented on, nor will it be considered as it has only one year of measurement.

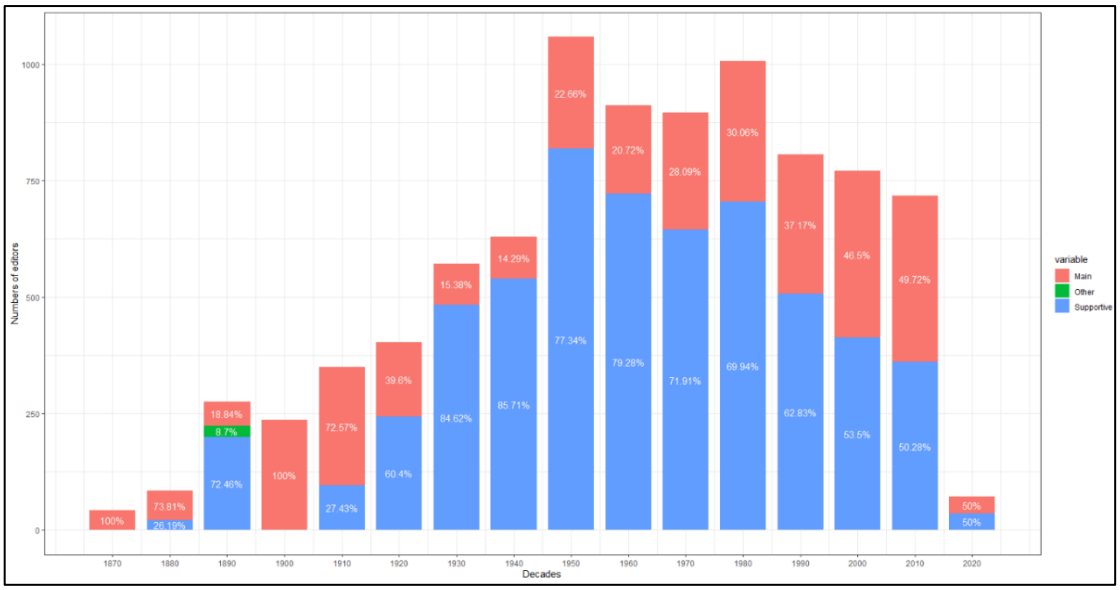


Figure 4.15 Stacked plot per decade with the percentage of main & supportive roles

CHAPTER 5

NUMBER OF INDIVIDUAL EDITORIAL BOARD MEMBERS PER ROLE

As in the previous chapter an analysis was made of the number and activity of editorial boards across time. The number and the structure of the editorial board can change from the editors which are composed. In this chapter, will be checked the editors individually for whom they took part in the editorial board and will look for answers to the following tasks.

- Must be known the number of editors that have been members of the journal.
- In the roles they have served, how many different roles they have performed, and in what order of the roles' appearance.
- Recording of results and examination for all editors if similar observations were observed

5.1 EDITORIAL POSITIONS

To be able to give answers to the tasks assigned above, our first step is to find out the number of editors of the journal. In the dataset, the total number of individual editors is 206. The next step of the process is to be created a table where every editor is recorded when had served the specific editorial position in editorial boards and which was the role that had. As soon as it is known the number of editors, it should be recorded in each role, how many of the editors were active at least once in the specific editorial positions. The total sum of *Table 6* is 235 and the total number of editors is 206, which means some editors were active over two editorial positions.

Editorial position	Total editors
associate editors	8
edited and publish by	1
edited by	57
editor	1
editor-in-chief	2
editorial board	2
founded by	1
technical edited by	4
with the cooperation of	159

Table 6

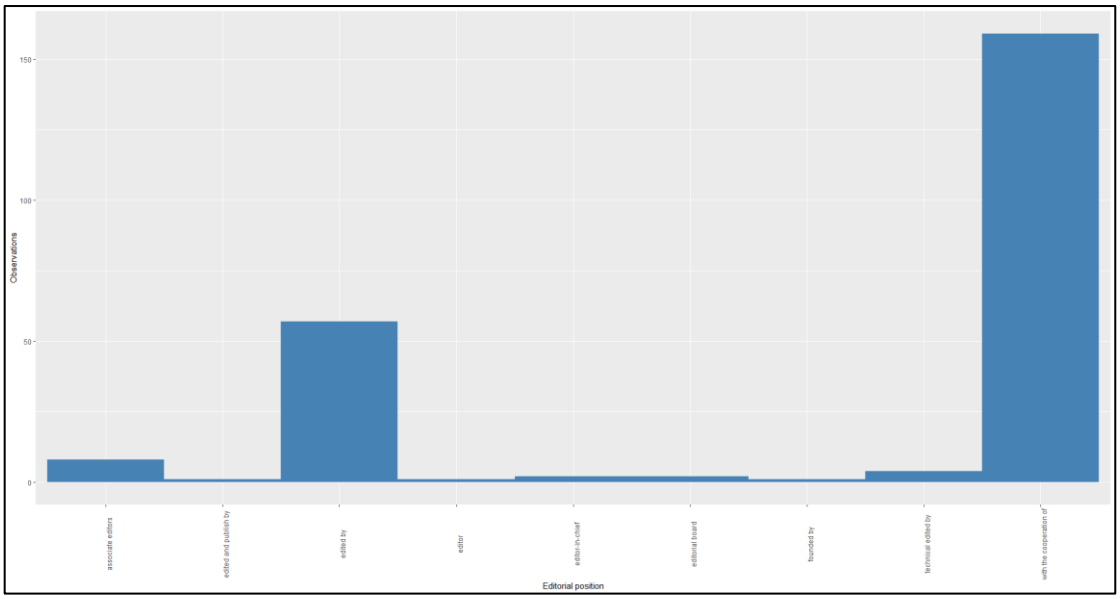


Figure 5.1 Number of editors in editorial positions.

Below is the graph where is displayed the number of editors was served one or more different editorial positions. Most editors stayed to one editorial position with the number of them to raise to 183, while the total of other editors is 23 who change at least one editorial position. Remarkable is that there is an editor who participates in journals with 6 different editorial roles, the name of him is Stone Ormond. After, the barplot is following a pie chart to capture the percentages of the number of different roles.

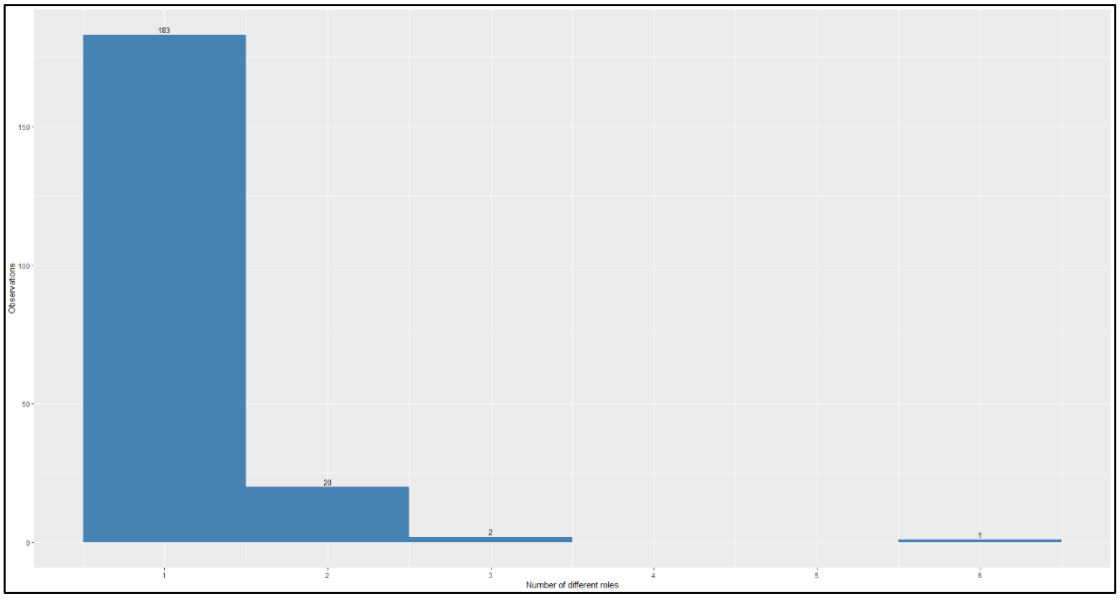


Figure 5.2 Barplot for the sum different roles that editors had served.

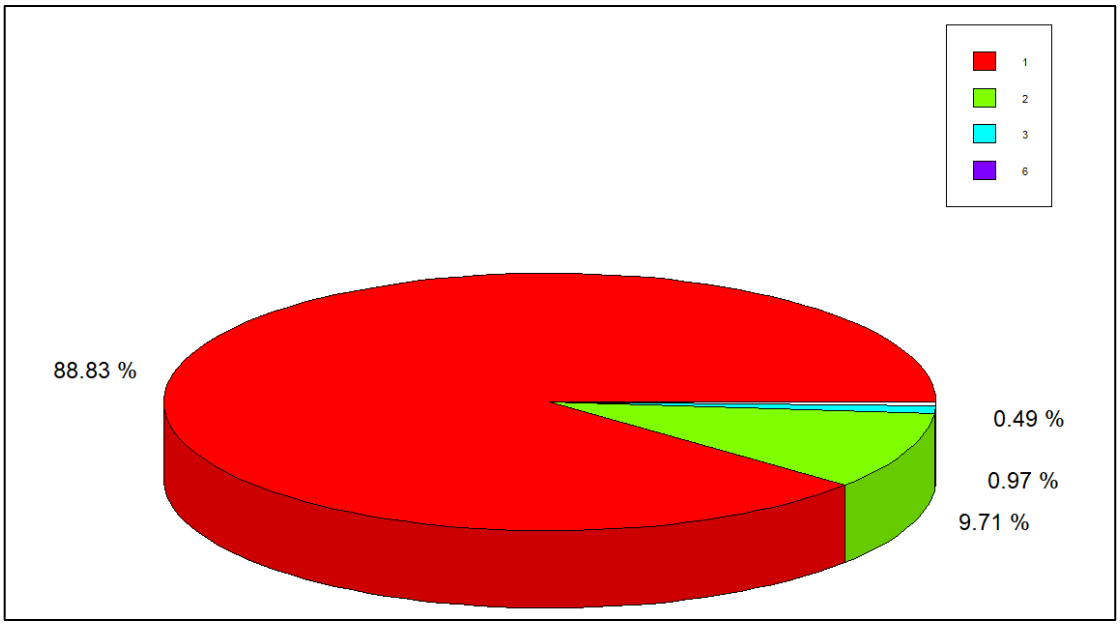


Figure 5.3 Pie chart percentages of the number of different roles.

5.2 COMBINATIONS OF EDITORIAL POSITIONS

Since it was observed that there are editors who were active in more than one role, it is interesting to analyze from which role they started and to whom they went. To locate when an editor changes from one role to another and to be recorded what is the order of the roles' appearance it was created a table with all issues based on the editors' Surname/name and it is recorded the issue and the editorial position that had in issue. Subsequently, it was sorted by Surname/name and number of issues and had added a counter which counts the number of consecutive appearances of the editor in an editorial position. The counter returns to 1 when the name of the editor under review is changed or when his role on the editorial board is changed at the same editor under review. In this paragraph, for now, will be examined the existence of each combination and not the total number of them, to be achieved that the observations with counter are equal to 1 will be chosen. In total, in the dataset, it is observed 242 observations of combinations, while there are 206 editors (per observation), 36 observations of them had served over 1 role. That means that there are combinations in which editors start from one editorial position and return to that editorial position or they are active to another role.

	with the cooperation of	edited by	associate editors	technical edited by	edited and publish by	editor-in-chief	editor	editorial board	founded by, associate editors
with the cooperation of	139	13	0	0	0	0	0	0	0
edited by	13	35	1	0	0	0	0	0	1
associate editors	0	1	4	0	0	0	0	1	0
technical edited by	0	0	0	4	0	0	0	0	0
editorial board	0	0	2	0	0	0	0	0	0
edited and publish by	0	0	0	0	1	0	0	0	0
editor	0	1	0	0	0	0	0	0	0
founded by, associate editors	0	0	0	0	0	1	0	0	0
editor-in-chief	0	0	0	0	0	0	1	1	0

Table 7. The number of editors that served editorial positions from one to another.

combo	Number of editorial members	Percentage(%)
with the cooperation of	139	63.470
edited by	35	15.982
edited by --> with the cooperation of	13	5.936
with the cooperation of --> edited by	13	5.936
associate editors	4	1.826
technical edited by	4	1.826
associate editors --> editorial board	2	0.913
associate editors --> edited by	1	0.457
edited and publish by	1	0.457
edited by --> associate editors	1	0.457
edited by --> editor	1	0.457
editor-in-chief --> founded by, associate editors	1	0.457
editor --> editor-in-chief	1	0.457
editorial board --> associate editors	1	0.457
editorial board --> editor-in-chief	1	0.457
founded by, associate editors --> edited by	1	0.457

Table 8. Combinations of editorial positions table and the percentage of coverage.

If an editor in issue had at the same time two roles in the table will be recorded with two roles but will be separated the roles with a comma (“,”), while when is recorded transition from one role to another, they are recorded with an arrow (“→”). From the results, the most often combinations are from “edited by” to “with the cooperation of” and reversal. The last process of this chapter is the creation of a table where will be displayed the unique number of distinct editorial members up to the issue across all positions and issues. In the end, our results confirm ***Table 6***.

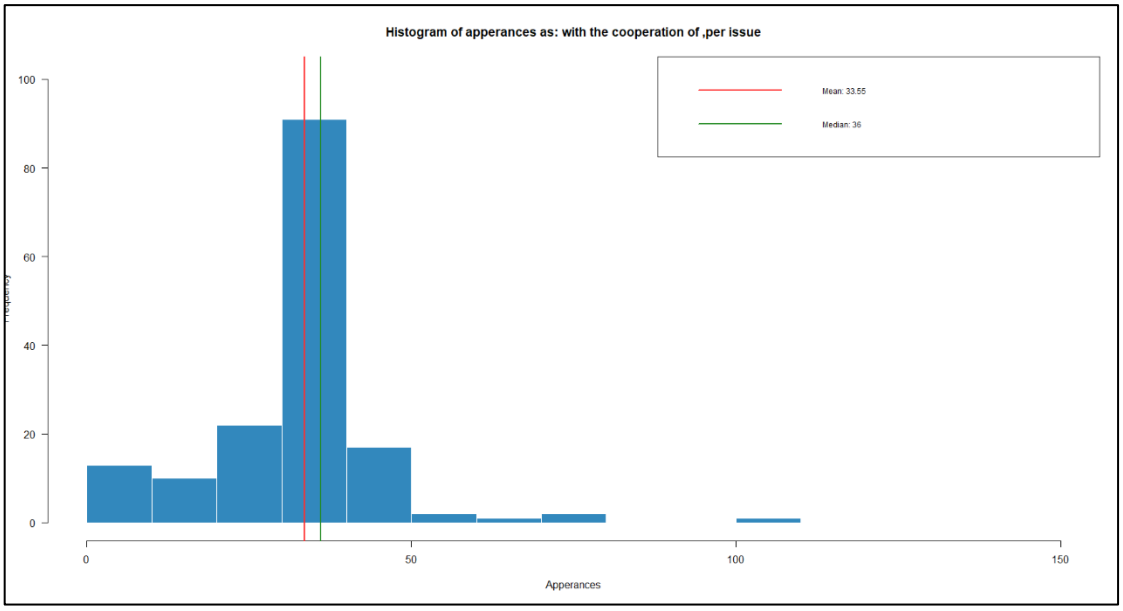
CHAPTER 6
DURATION OF PARTICIPATION PER UNIQUE EDITORIAL BOARD MEMBER

So far, they have been examined with the two criteria set, that of the year and that of the issues first in the editorial boards and secondly in the editors to come to conclusions through their analysis. The last part that will be analyzed in this paper is the duration of participation in the editorial board of each editor that was active in the journal. The duration will be measured in issues and years, the results will be represented in one table across all positions and one table per position and at the same time will create the distribution of the number of issues and the number of years per position. Below it will be analyzed:

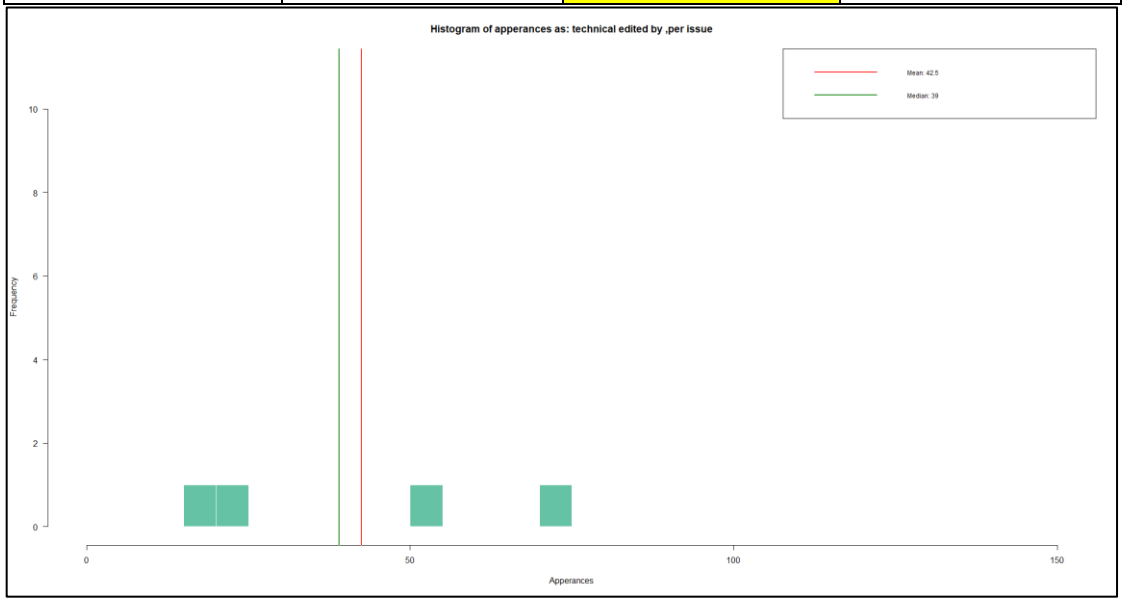
- The duration of participation in the editorial board for each editorial position measured in issues and years and distribution of the number of issues and number of years per position
- For every editorial board member, the time that served per issue, and presentation of their timeline in the journal.
- To the editors who changed editorial positions recording the duration of their participation in each position, and presentation of their timeline.

6.1 DURATION OF PARTICIPATION PER ISSUE AND PER POSITION

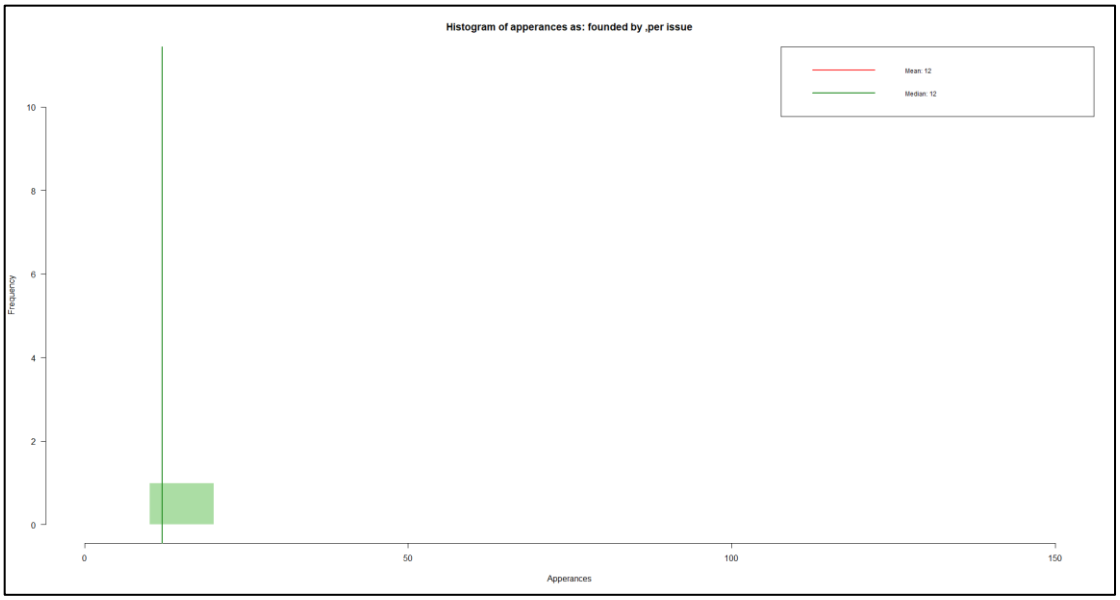
At first, a table is created in which the total of the issues that have served in each role are displayed for all editors, the observations are gathered in bins and visualized to below graphs per editorial position. The bins are separated by 10, starting from 0 to 150 observations. In the editorial position "with the cooperation of", in total have been served until 10 issues,13 editors. Also, in graphs, are visualized the means and medians of positions.



Bins	Observations	Bins	Observations
(0,10]	13	(80,90]	0
(10,20]	10	(90,100]	0
(20,30]	22	(100,110]	11
(30,40]	91	(110,120]	0
(40,50]	17	(120,130]	0
(50,60]	2	(130,140]	0
(60,70]	1	Mean	33,35
(70,80]	2	Median	36



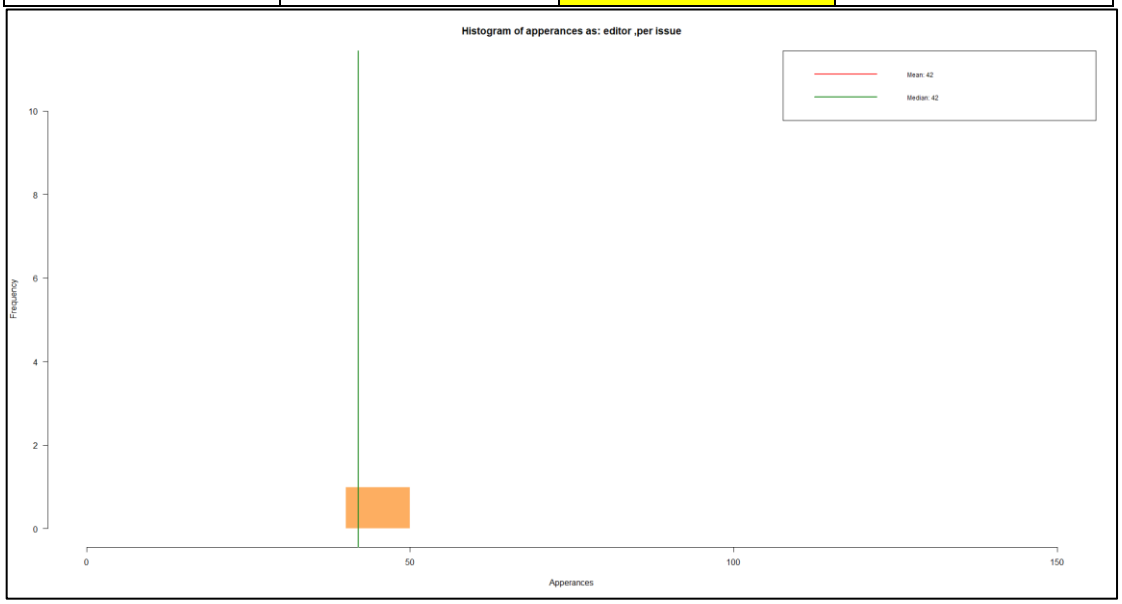
Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	1	(90,100]	0
(20,30]	1	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	0	(120,130]	0
(50,60]	1	(130,140]	0
(60,70]	0	Mean	42,5
(70,80]	1	Median	39



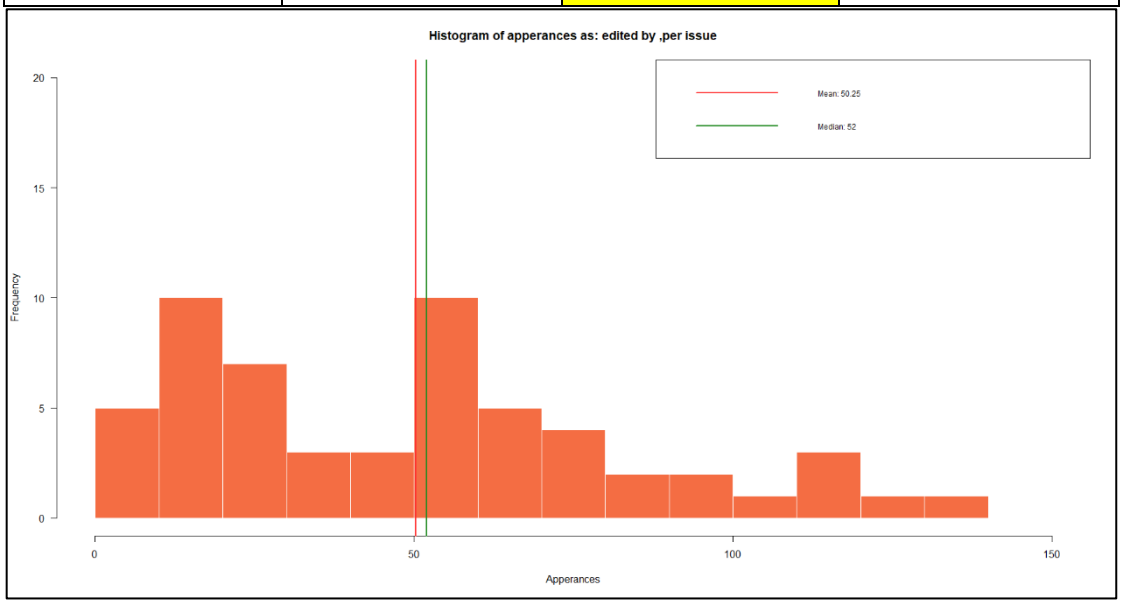
Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	1	(90,100]	0
(20,30]	0	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	0	(120,130]	0
(50,60]	0	(130,140]	0
(60,70]	0	Mean	12
(70,80]	0	Median	12

Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	1	(90,100]	0
(20,30]	0	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	0	(120,130]	0
(50,60]	0	(130,140]	0
(60,70]	0	Mean	6
(70,80]	0	Median	6

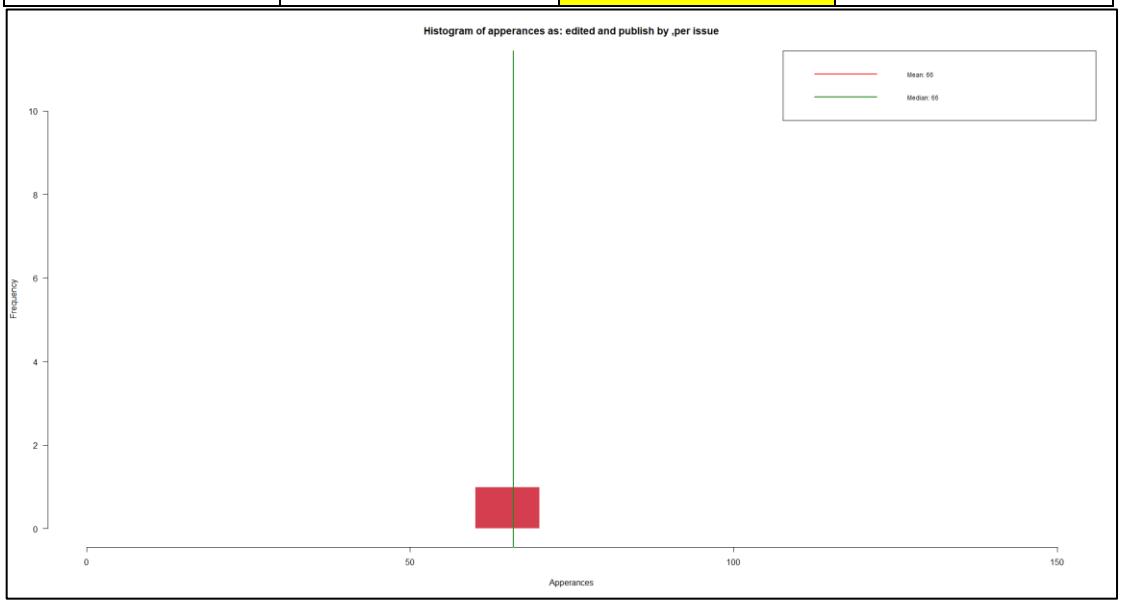
Bins	Observations	Bins	Observations
(0,10]	1	(80,90]	0
(10,20]	1	(90,100]	0
(20,30]	0	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	0	(120,130]	0
(50,60]	0	(130,140]	0
(60,70]	0	Mean	9
(70,80]	0	Median	9



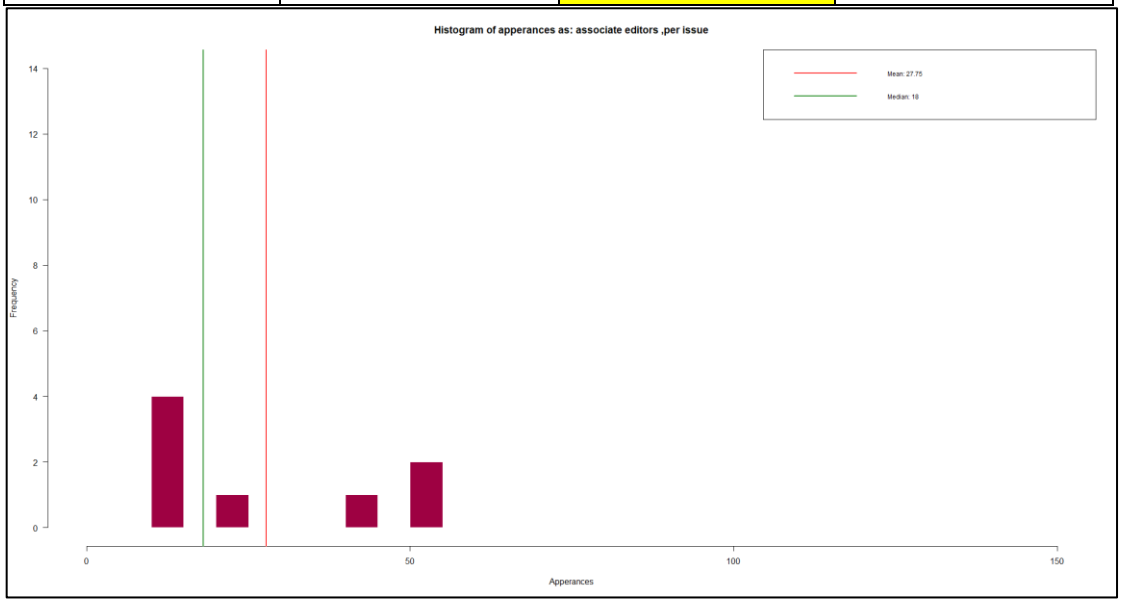
Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	0	(90,100]	0
(20,30]	0	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	1	(120,130]	0
(50,60]	0	(130,140]	0
(60,70]	0	Mean	42
(70,80]	0	Median	42



Bins	Observations	Bins	Observations
(0,10]	5	(80,90]	2
(10,20]	10	(90,100]	2
(20,30]	7	(100,110]	1
(30,40]	3	(110,120]	3
(40,50]	3	(120,130]	1
(50,60]	10	(130,140]	1
(60,70]	5	Mean	50,25
(70,80]	4	Median	52

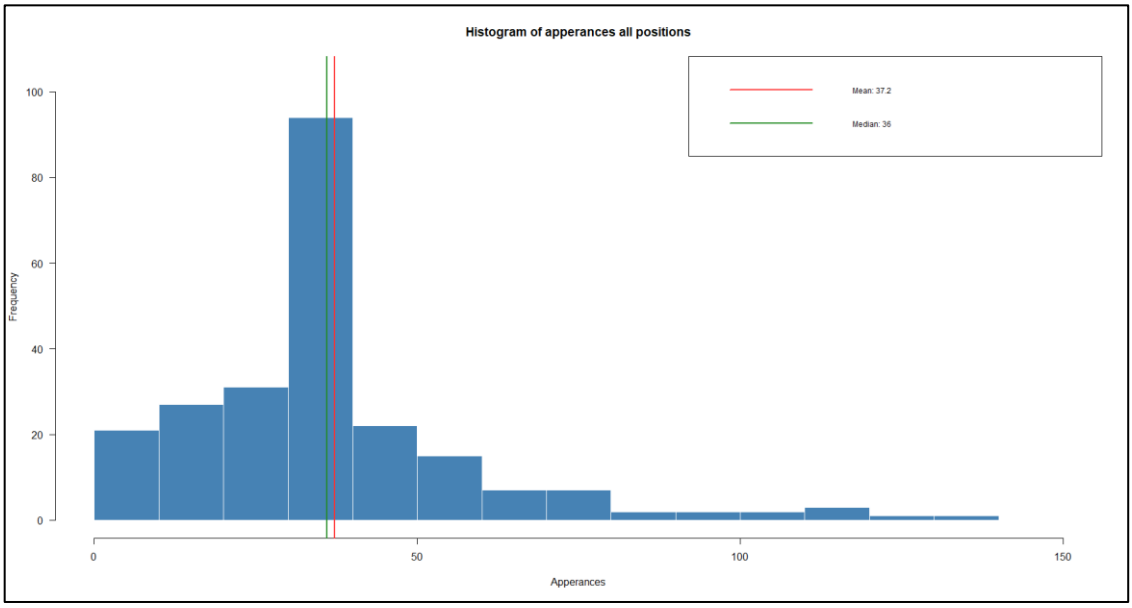


Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	0	(90,100]	0
(20,30]	0	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	0	(120,130]	0
(50,60]	0	(130,140]	0
(60,70]	1	Mean	66
(70,80]	0	Median	66



Bins	Observations	Bins	Observations
(0,10]	0	(80,90]	0
(10,20]	4	(90,100]	0
(20,30]	1	(100,110]	0
(30,40]	0	(110,120]	0
(40,50]	1	(120,130]	0
(50,60]	2	(130,140]	0
(60,70]	0	Mean	27,75
(70,80]	0	Median	18

After checking for each role separately, below is represented the number of editors per number of appearances in issues of the journal across all editorial positions. The purpose of this process is to understand if they remain for a long time editors who participated in the journal. The average duration of participation is 37,2 issues.



Bins	Observations	Bins	Observations
(0,10]	21	(80,90]	2
(10,20]	27	(90,100]	2
(20,30]	31	(100,110]	2
(30,40]	94	(110,120]	3
(40,50]	22	(120,130]	1
(50,60]	15	(130,140]	1
(60,70]	7	Mean	37,2
(70,80]	7	Median	36

6.2 DURATION OF PARTICIPATION PER YEAR AND PER POSITION

In this paragraph, there will be an analysis of the time that each editorial member contributed to each editorial position and in total the unit of measurement will now be time. During the analysis, complications were created regarding the recording of the roles are changed and the editorial position “founded by” is consolidated with “associate editors”, however “associate editors” as a unique editorial position continues to exist. At first, the reason for this conversion was for reasons of usability of the following methods as the "founded by" role was used only by one editor and was used at the same time with another role of the journal. The new roles and their colors are represented below:

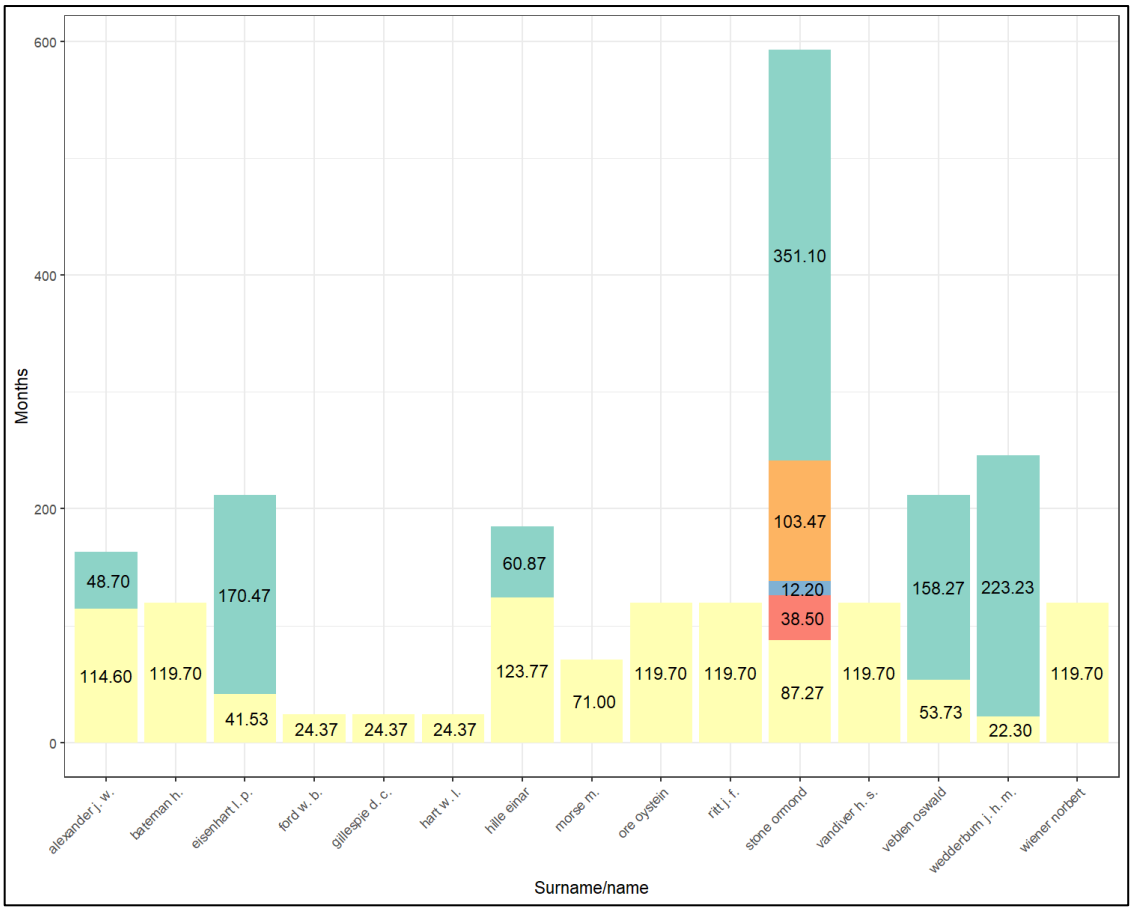
Editorial position	color
edited by	
with the cooperation of	
technical edited by	
founded by,associate editors	
editor-in-chief	
editor	
associate editors	
editorial board	
edited and publish by	

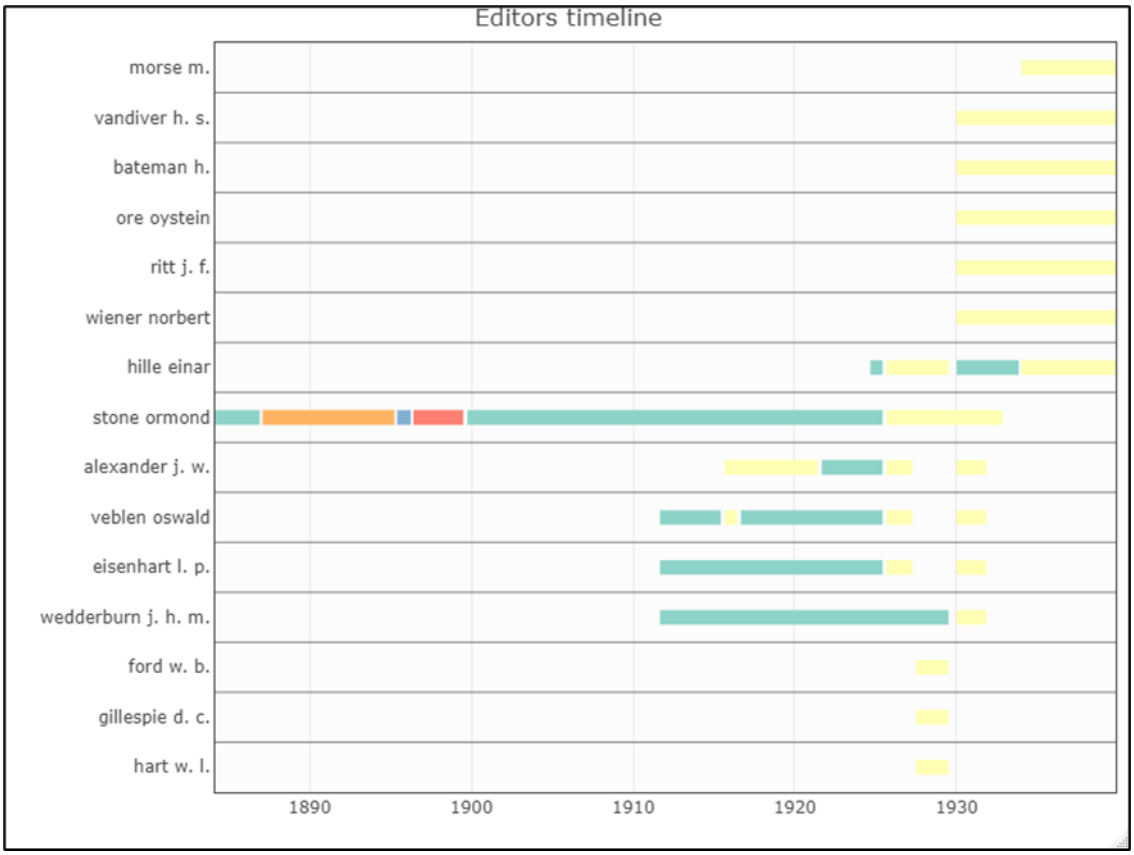
Table 9

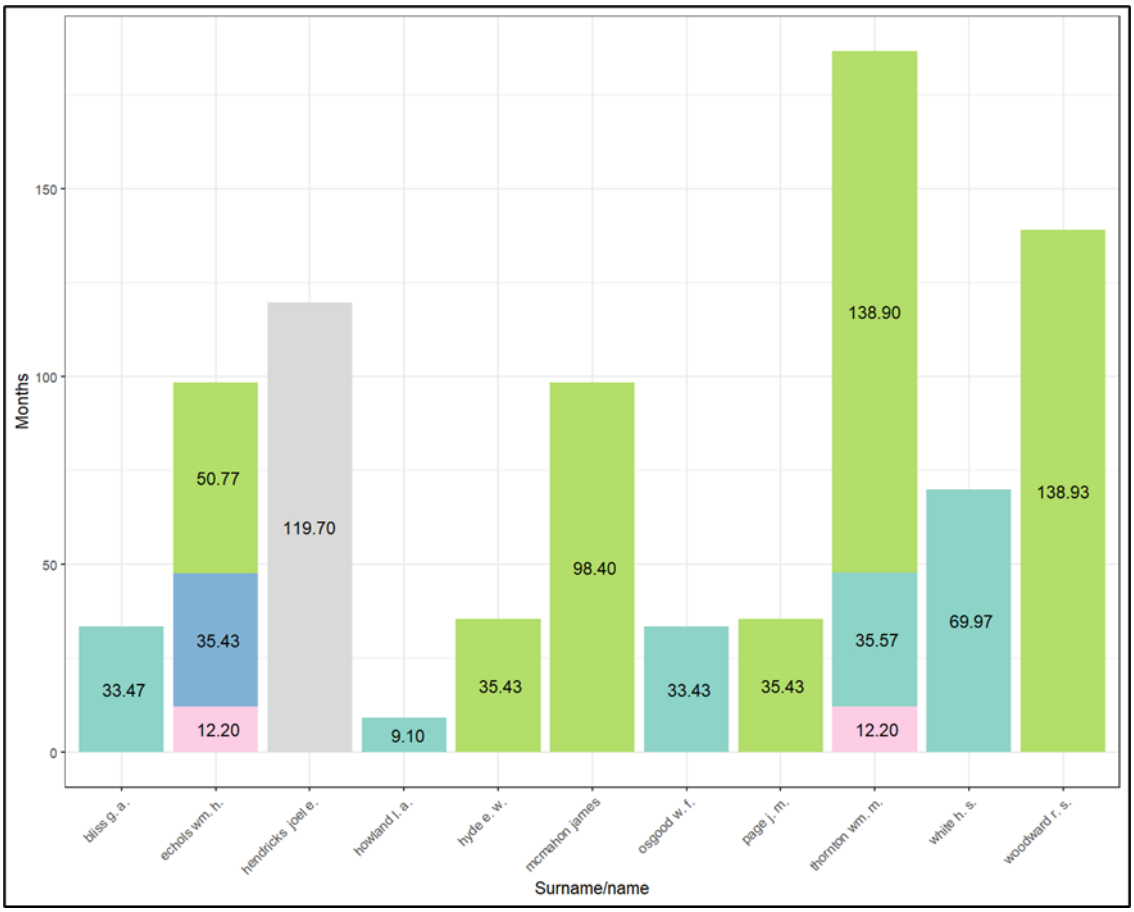
For the analysis to take place, a table was created in which is recorded per individual editorial member the number of participants and the total number of days he participated in each position. Then all the roles are separated for each member who participated in the role for at least one day, if the number of days is 0 then is deleted. For the table to be presented, they will be created stacked plots for each editorial member, the color of the barplot will be the editorial position and the number inside to barplot will be the total number of days in this role. However, for the procedure that will follow to be accurate, there will be a timeline for each member in which will be presented his activity, in the journal. The graphs of this paragraph are separated per 15 editorial members with the purposed to be more distinct and easier to be examined. In total are 206 editors and 15 editors per graph, so are separated into 14 groups of plots.

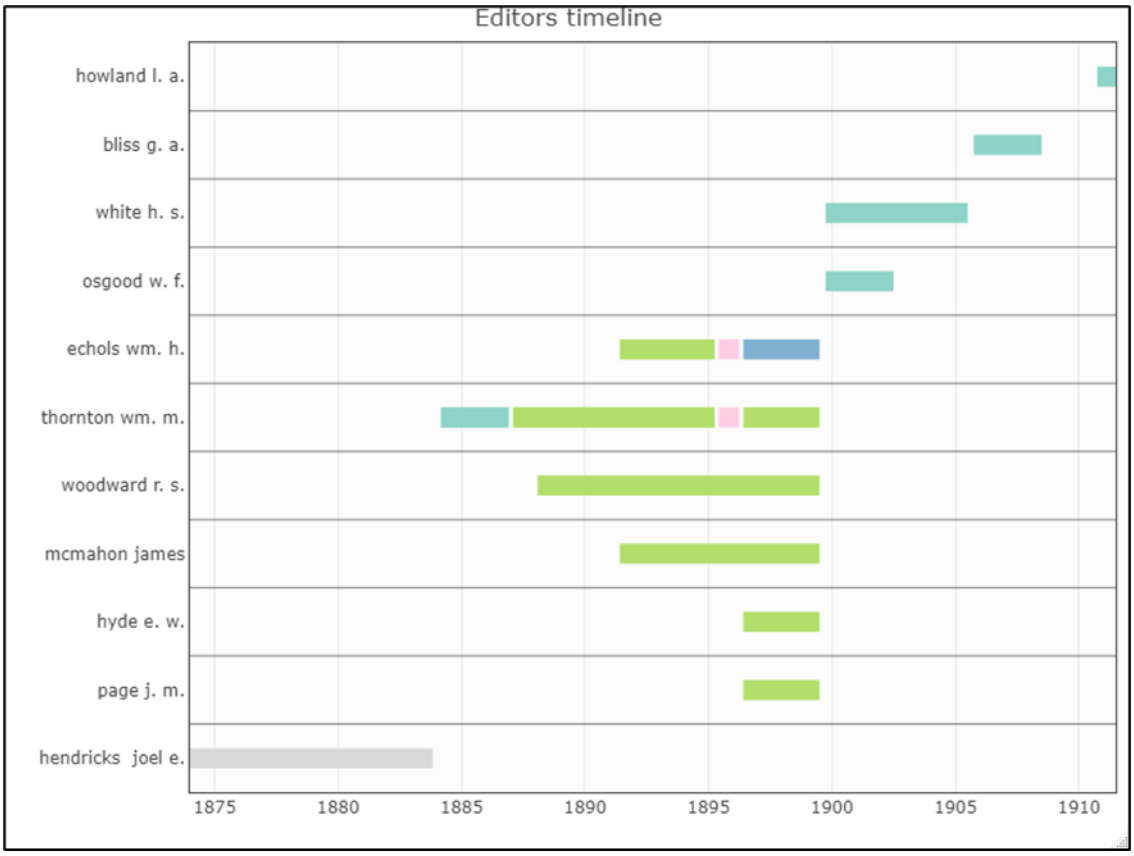
Finally, the editorial members which were participated over one editorial position will be analyzed with a timeline to examine the order of their activity in the journal and will be answered questions such as:

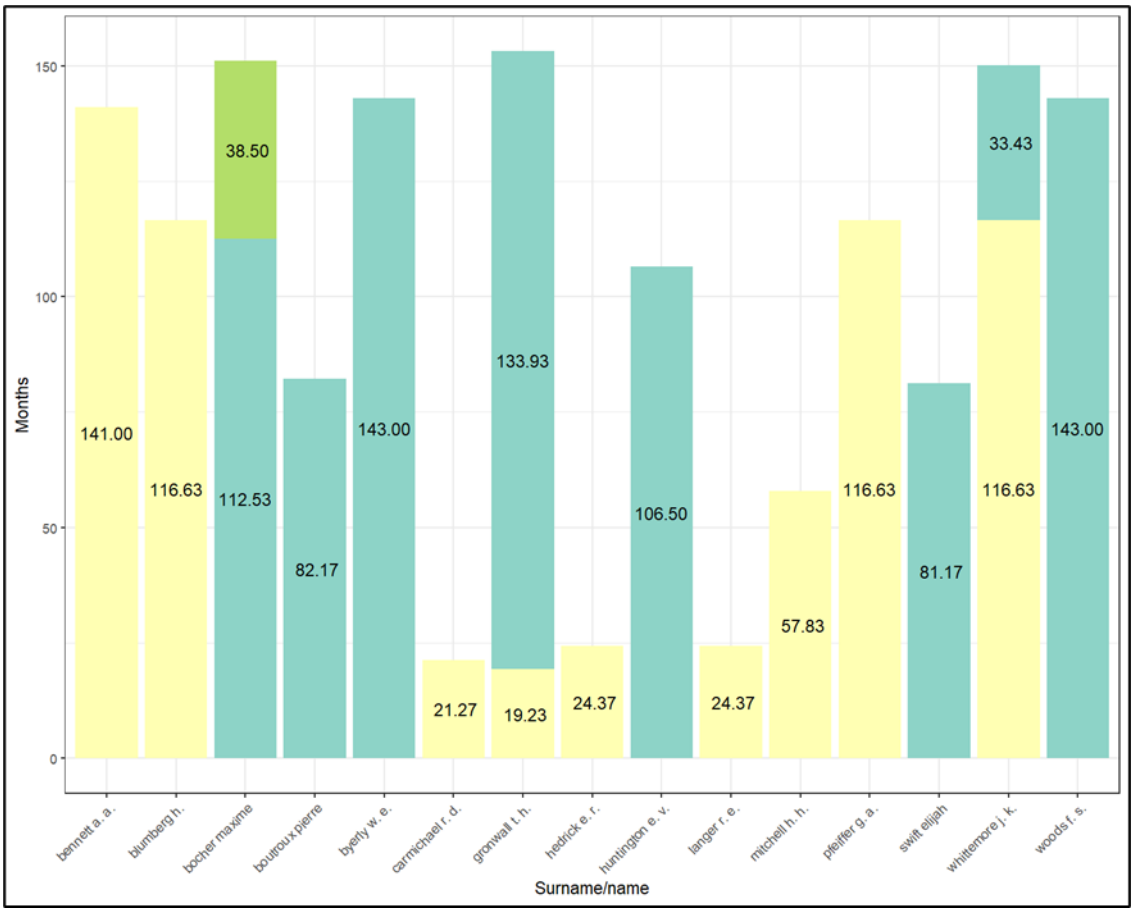
- The number of editors who participate in more than one role, the order of activity, and the time that participates in each role.
- The capacity of the role that is active and the role that is transferred.

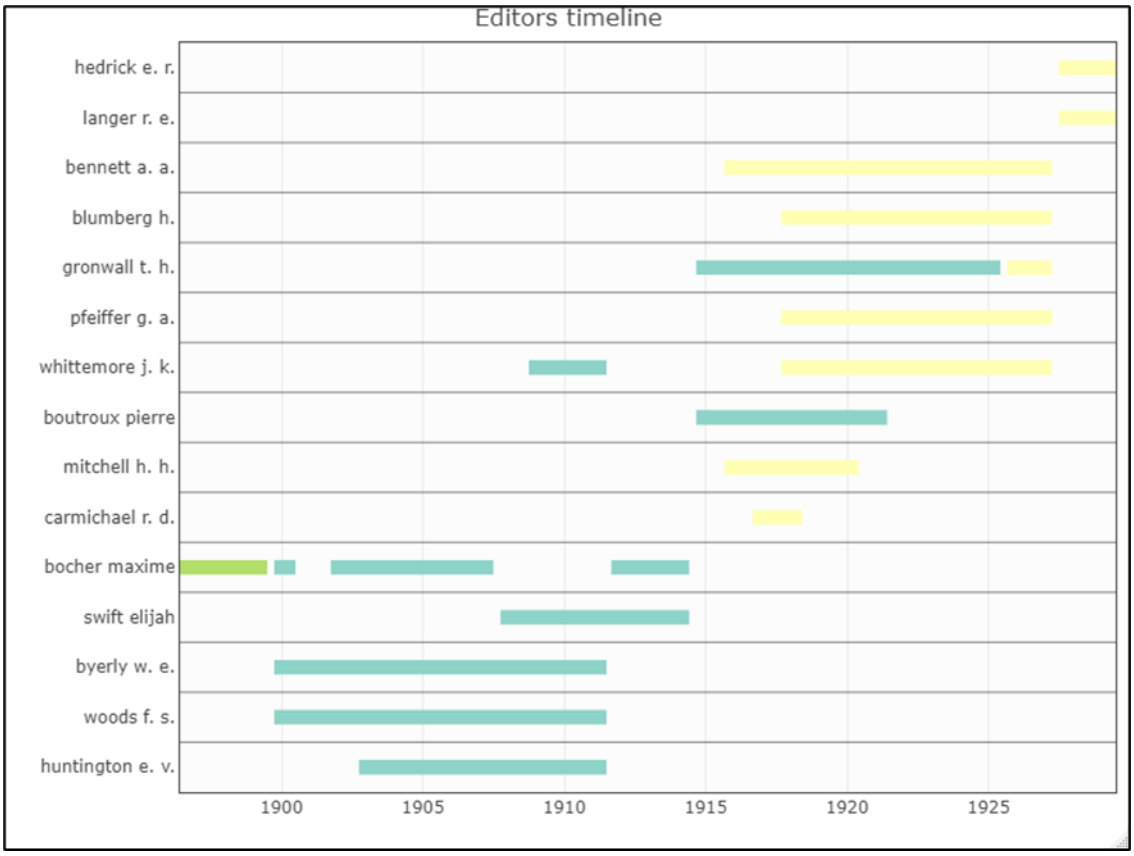

Months
600
400
200
0
48.70
114.60
119.70
170.47
41.53
24.37
24.37
24.37
60.87
123.77
71.00
119.70
119.70
351.10
103.47
12.20
38.50
87.27
119.70
158.27
53.73
223.23
22.30
119.70
alexander j. w.
bateman h.
eisenhart l. p.
ford w. b.
gillespie d. c.
hart w. l.
hille einar
morse m.
ore oystein
ritt j. f.
stone ormond
vandiver h. s.
veblen oswald
wedderburn j. h. m.
wiener norbert
Surname/name

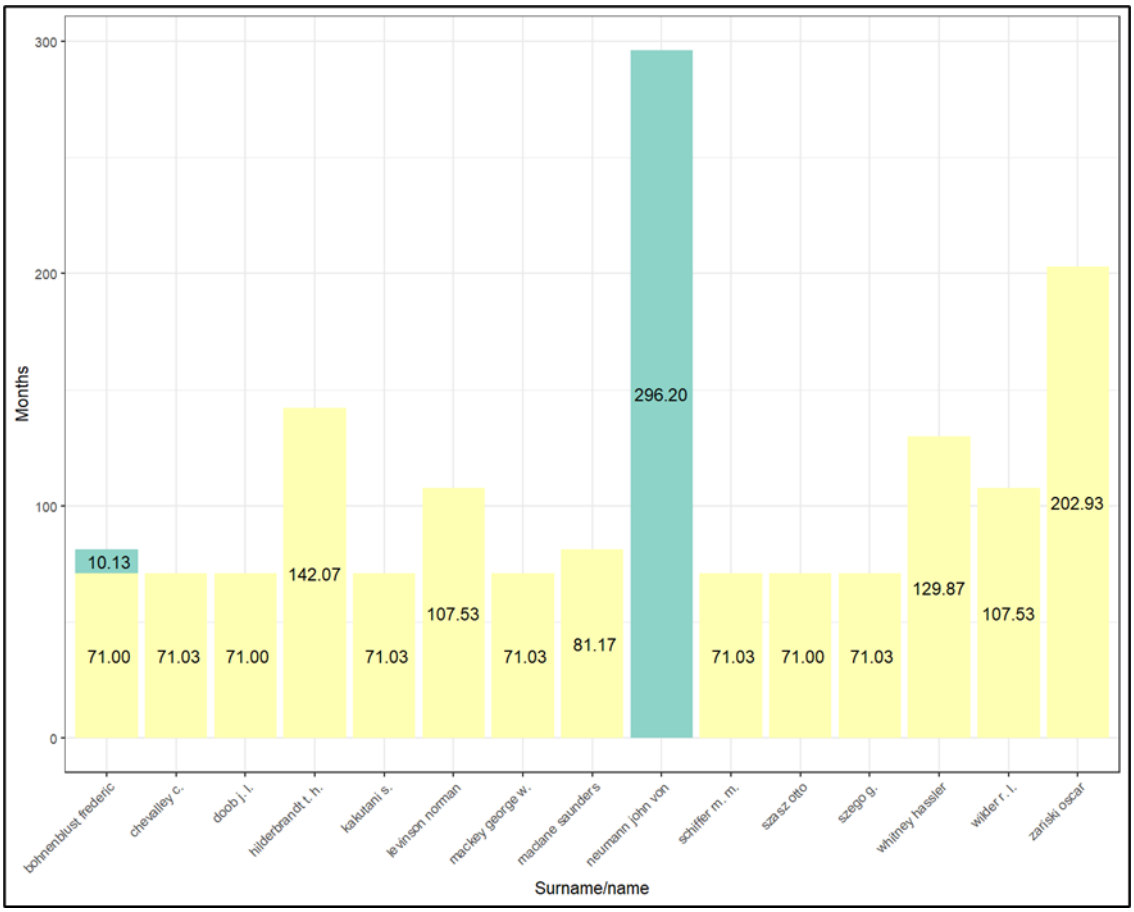

Editors timeline
morse m.
vandiver h. s.
bateman h.
ore oystein
ritt j. f.
wiener norbert
hille einar
stone ormond
alexander j. w.
veblen oswald
eisenhart l. p.
wedderburn j. h. m.
ford w. b.
gillespie d. c.
hart w. l.
1890
1900
1910
1920
1930

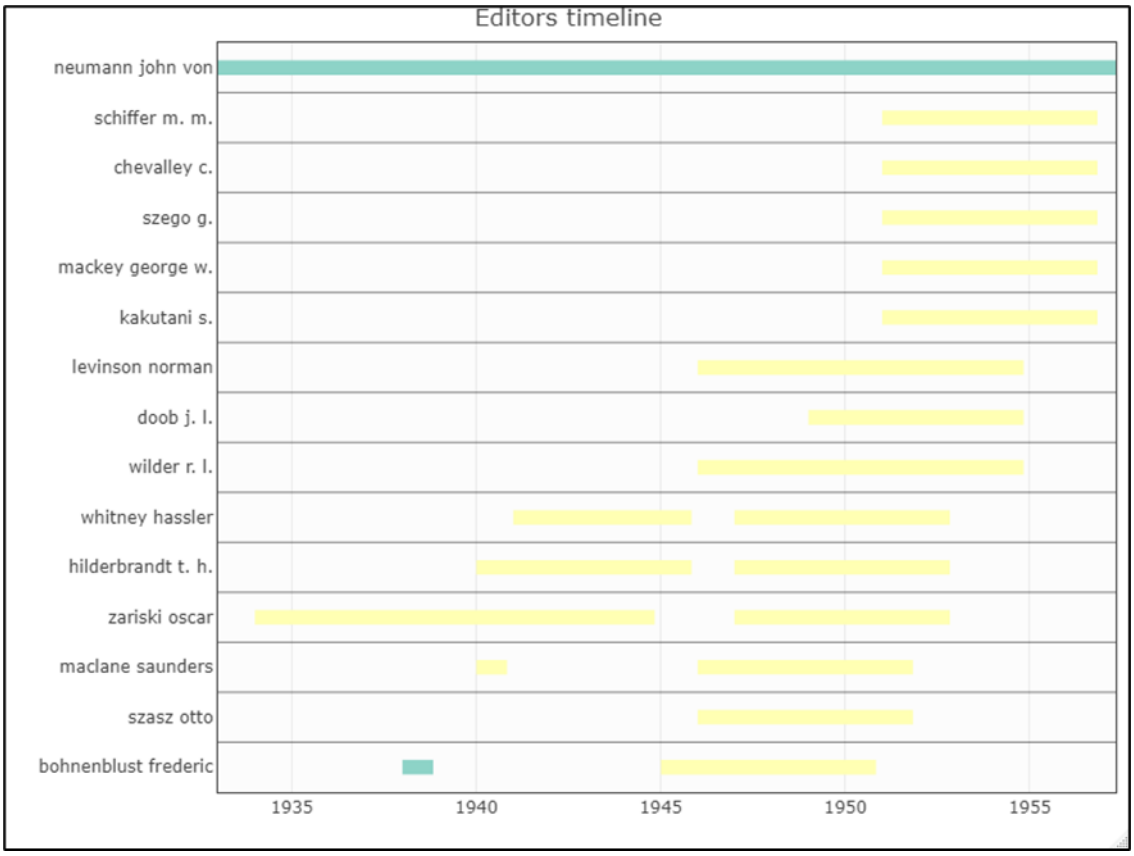

Months
150
100
50
0
33.47
50.77
35.43
12.20
119.70
9.10
35.43
98.40
33.43
35.43
138.90
35.57
12.20
69.97
138.93
bliss g. a.
echols wm. h.
hendricks joel e.
howland l. a.
hyde e. w.
mcmahon james
osgood w. f.
page j. m.
thornton wm. m.
white h. s.
woodward r. s.
Surname/name

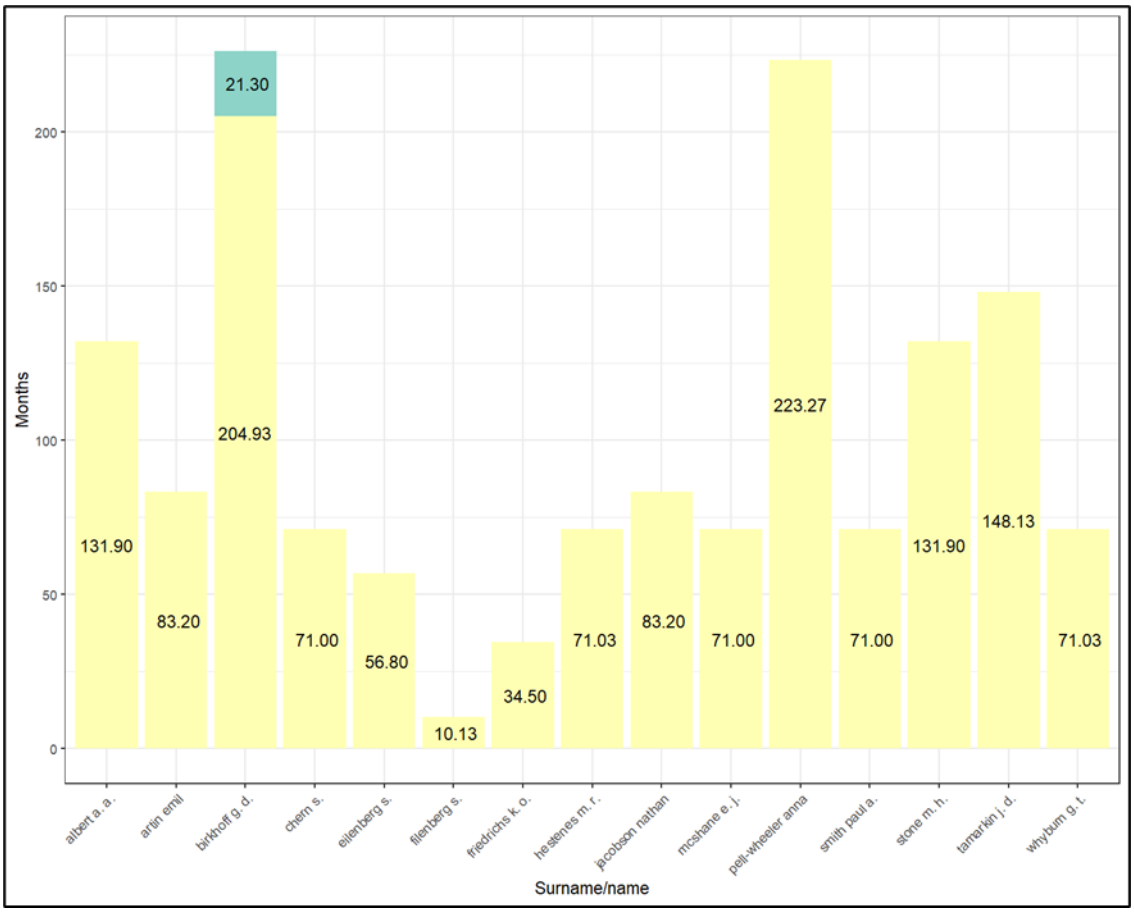

Editors timeline
howland l. a.
bliss g. a.
white h. s.
osgood w. f.
echols wm. h.
thornton wm. m.
woodward r. s.
mcmahon james
hyde e. w.
page j. m.
hendricks joel e.
1875
1880
1885
1890
1895
1900
1905
1910

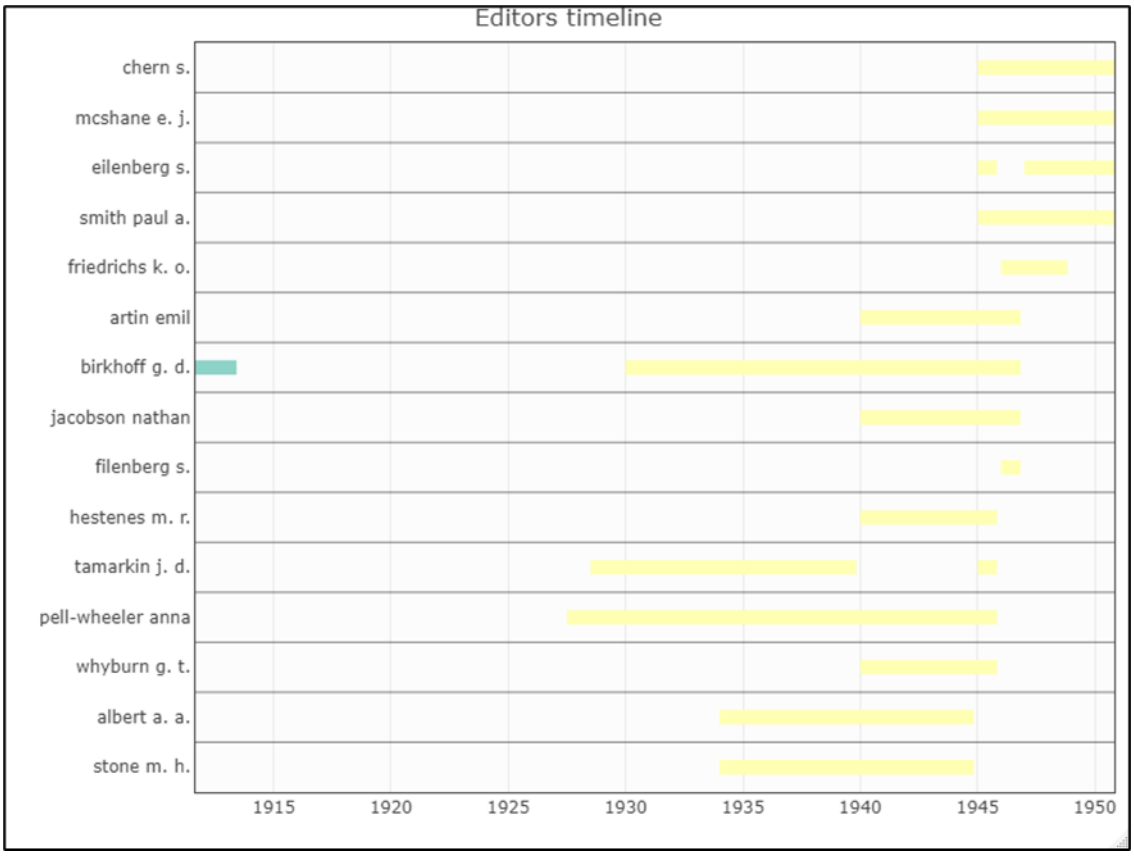

Months
150
100
50
0
141.00
116.63
38.50
112.53
82.17
143.00
21.27
133.93
19.23
24.37
106.50
24.37
57.83
116.63
81.17
33.43
116.63
143.00
bennett a. a.
blumberg h.
bocher maxime
boutroux pierre
byerly w. e.
carmichael r. d.
gronwall t. h.
hedrick e. r.
huntington e. v.
langer r. e.
mitchell h. h.
pfeiffer g. a.
swift elijah
whittemore j. k.
woods f. s.
Surname/name

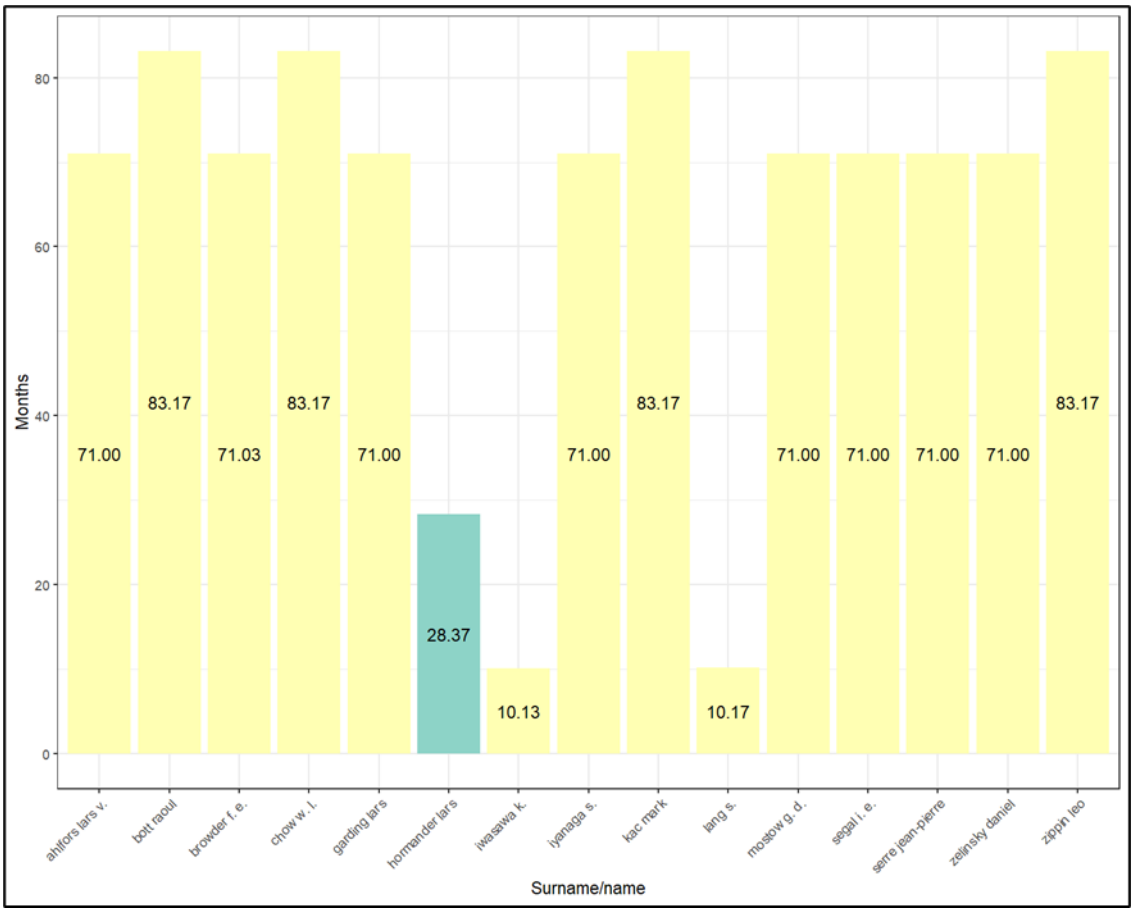

Editors timeline
hedrick e. r.
langer r. e.
bennett a. a.
blumberg h.
gronwall t. h.
pfeiffer g. a.
whittemore j. k.
boutroux pierre
mitchell h. h.
carmichael r. d.
bocher maxime
swift elijah
byerly w. e.
woods f. s.
huntington e. v.
1900
1905
1910
1915
1920
1925

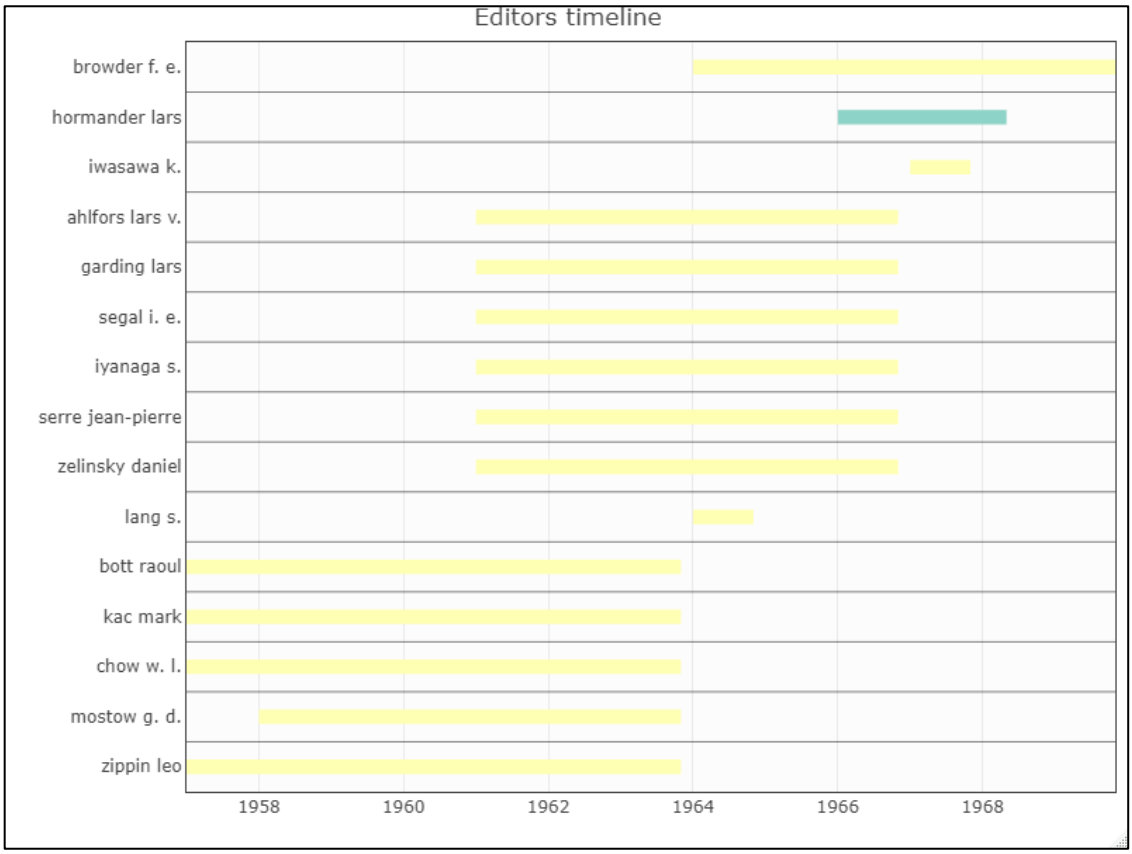

Months
300
200
100
0
10.13
71.00
71.03
71.00
142.07
71.03
107.53
71.03
81.17
296.20
71.03
71.00
71.03
129.87
107.53
202.93
bohnenblust frederic
chevalley c.
doob j. l.
hilderbrandt t. h.
kakutani s.
levinson norman
mackey george w.
maclane saunders
neumann john von
schiffer m. m.
szasz otto
szego g.
whitney hassler
wilder r. l.
zariski oscar
Surname/name

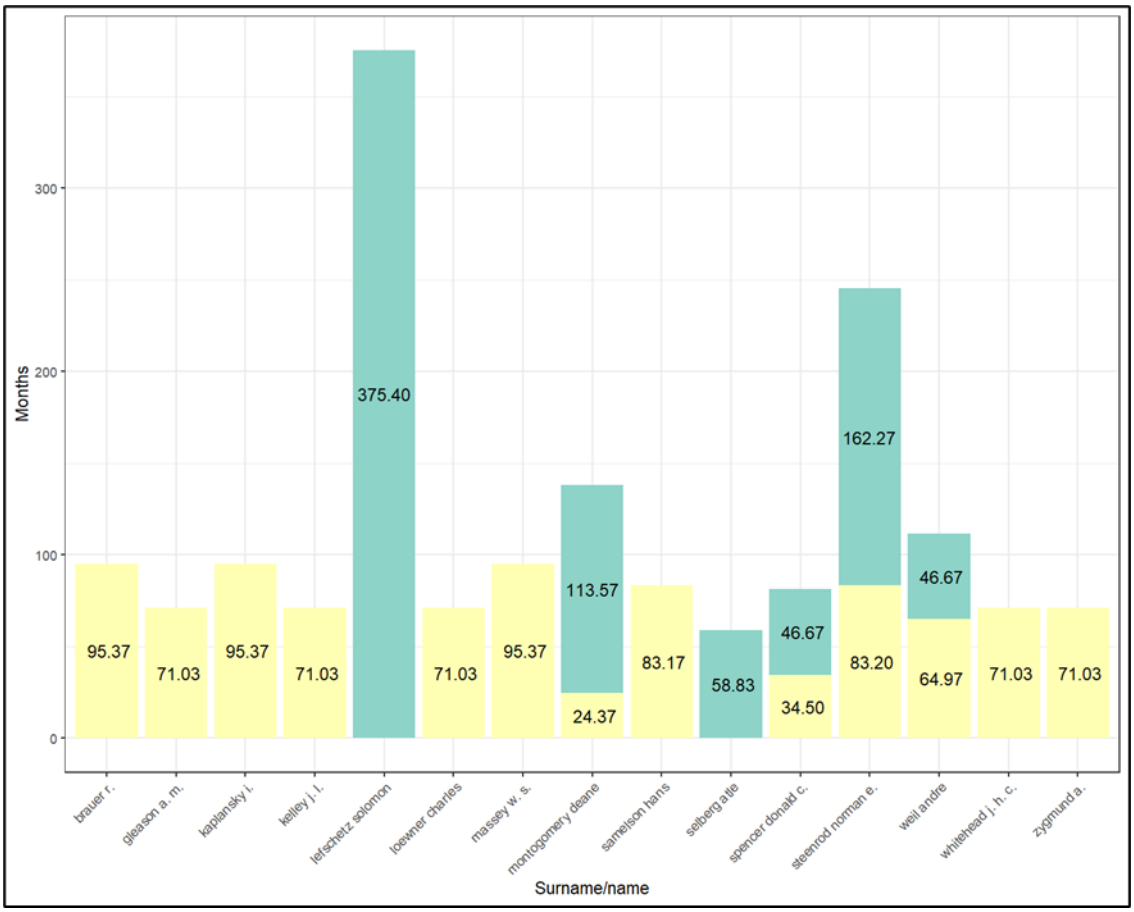

Editors timeline
neumann john von
schiffer m. m.
chevalley c.
szego g.
mackey george w.
kakutani s.
levinson norman
doob j. l.
wilder r. l.
whitney hassler
hilderbrandt t. h.
zariski oscar
maclane saunders
szasz otto
bohnenblust frederic
1935
1940
1945
1950
1955

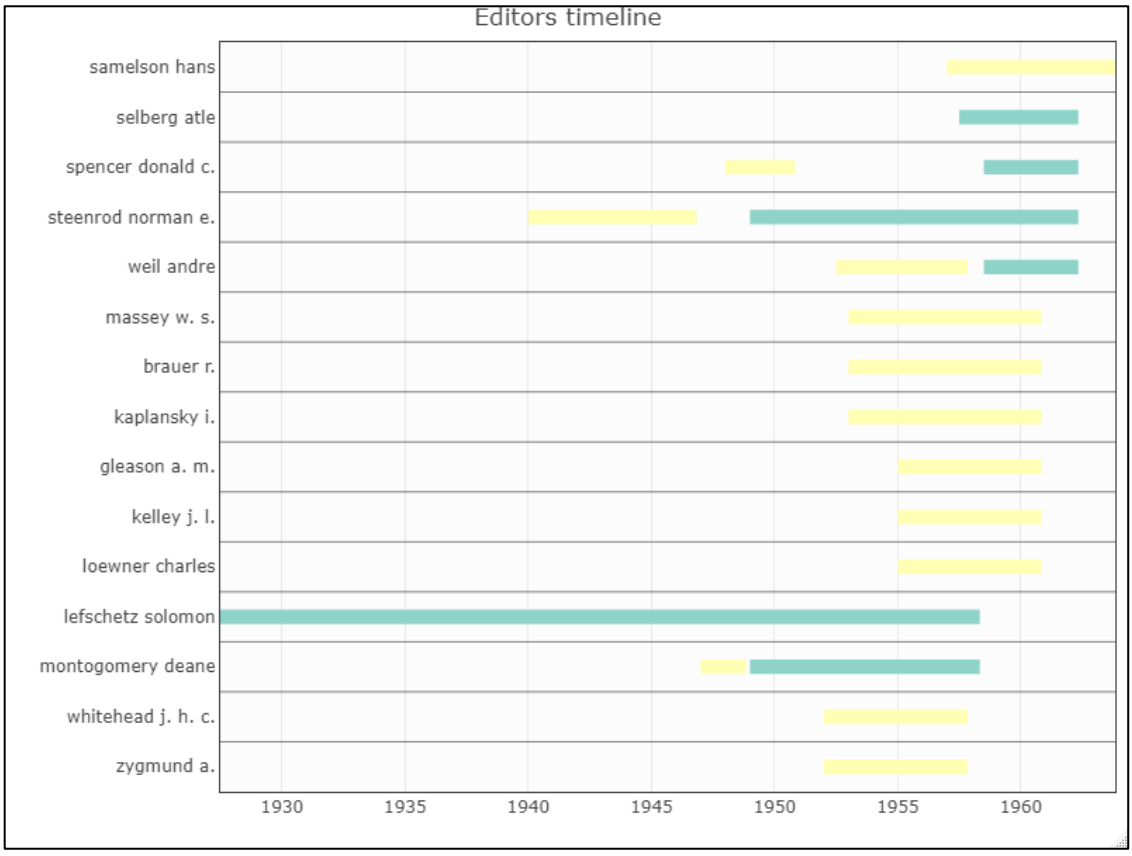

Months
Surname/name
0
50
100
150
200
131.90
83.20
204.93
21.30
71.00
56.80
10.13
34.50
71.03
83.20
71.00
223.27
71.00
131.90
148.13
71.03
albert a. a.
artin emil
birkhoff g. d.
chern s.
eilenberg s.
filenberg s.
friedrichs k. o.
hestenes m. r.
jacobson nathan
mcshane e. j.
pell-wheeler anna
smith paul a.
stone m. h.
tamarkin j. d.
whyburn g. t.

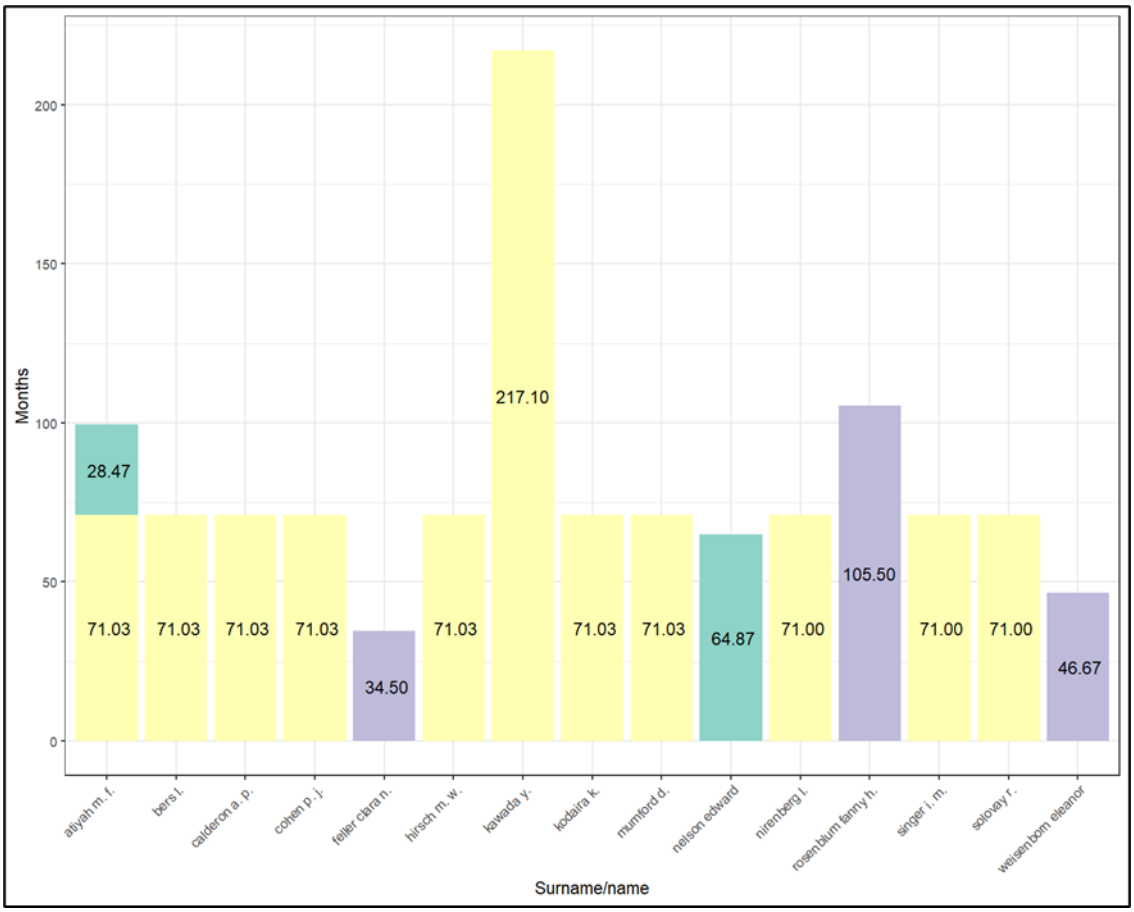

Editors timeline
chern s.
mcshane e. j.
eilenberg s.
smith paul a.
friedrichs k. o.
artin emil
birkhoff g. d.
jacobson nathan
filenberg s.
hestenes m. r.
tamarkin j. d.
pell-wheeler anna
whyburn g. t.
albert a. a.
stone m. h.
1915
1920
1925
1930
1935
1940
1945
1950

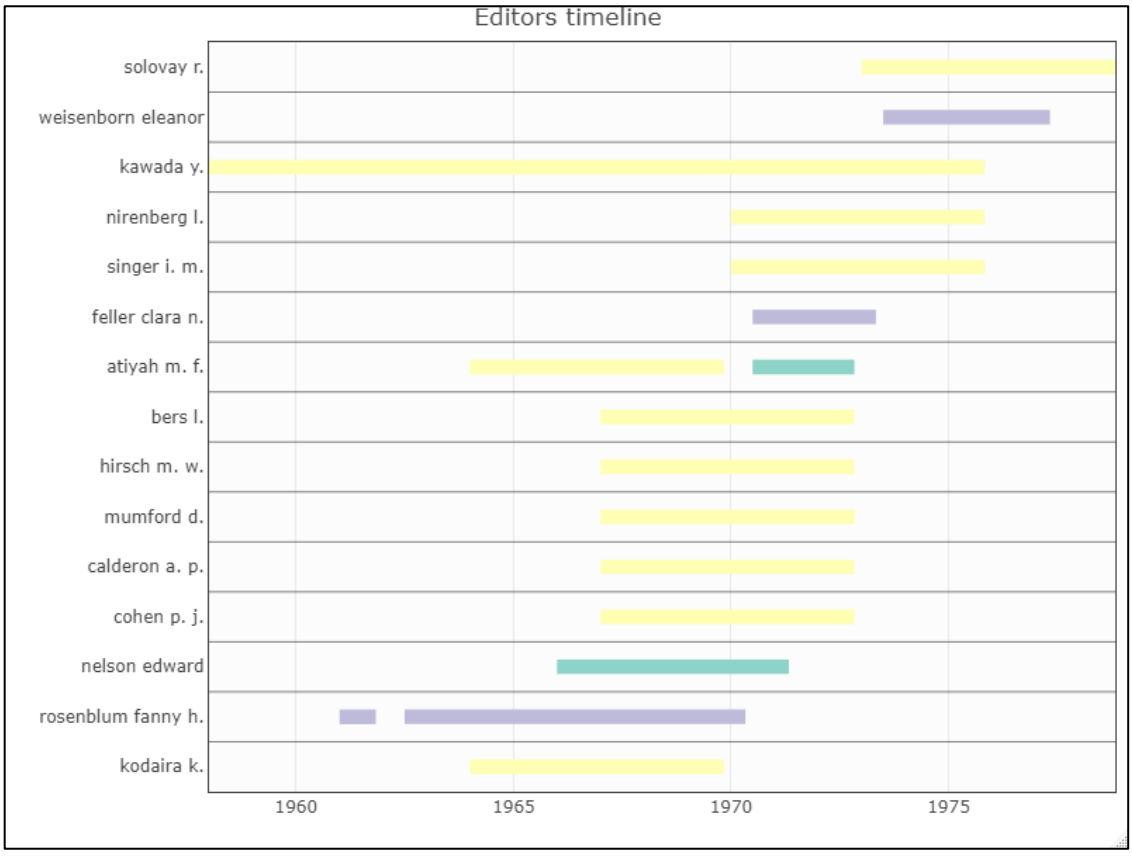

Months
Surname/name
0
20
40
60
80
71.00
83.17
71.03
83.17
71.00
28.37
10.13
71.00
83.17
10.17
71.00
71.00
71.00
71.00
83.17
ahlfors lars v.
bott raoul
browder f. e.
chow w. l.
garding lars
hormander lars
iwasawa k.
iyanaga s.
kac mark
lang s.
mostow g. d.
segal i. e.
serre jean-pierre
zelinsky daniel
zippin leo

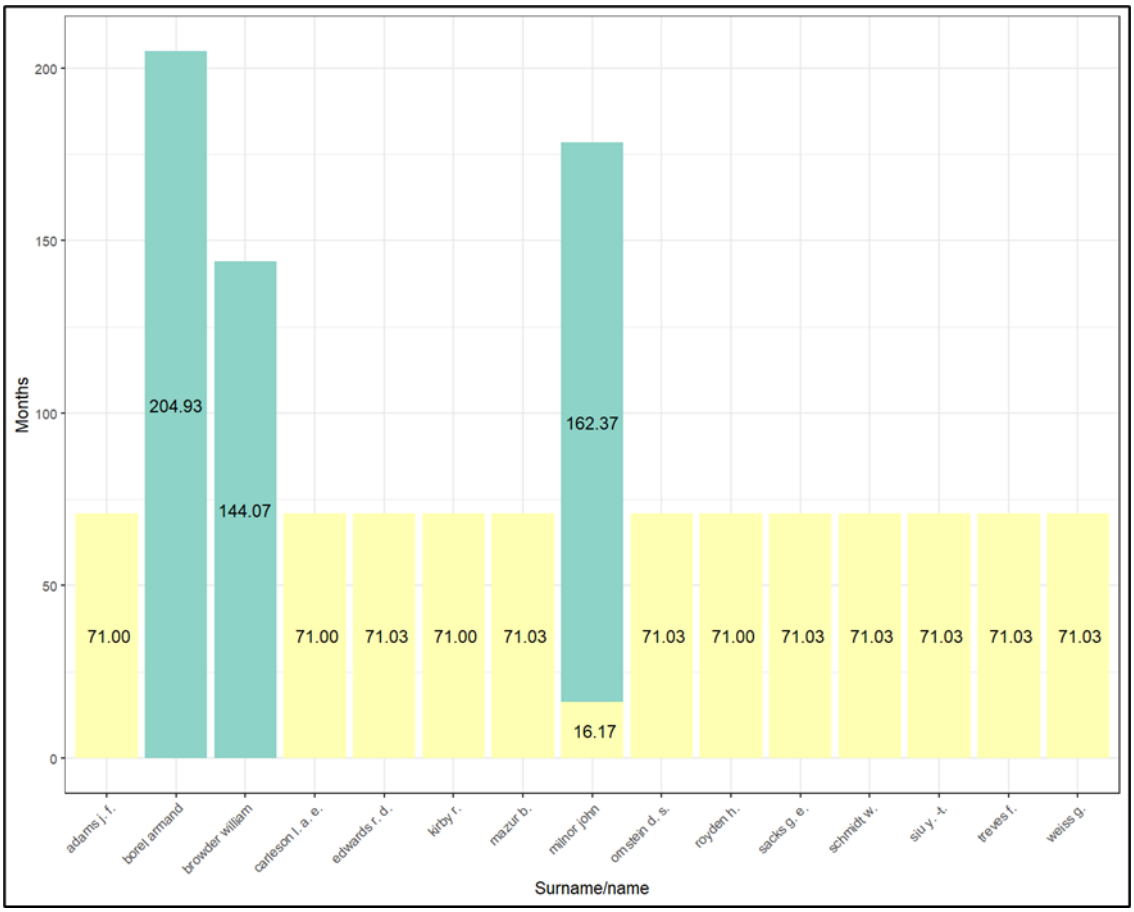

Editors timeline
browder f. e.
hormander lars
iwasawa k.
ahlfors lars v.
garding lars
segal i. e.
iyanaga s.
serre jean-pierre
zelinsky daniel
lang s.
bott raoul
kac mark
chow w. l.
mostow g. d.
zippin leo
1958
1960
1962
1964
1966
1968

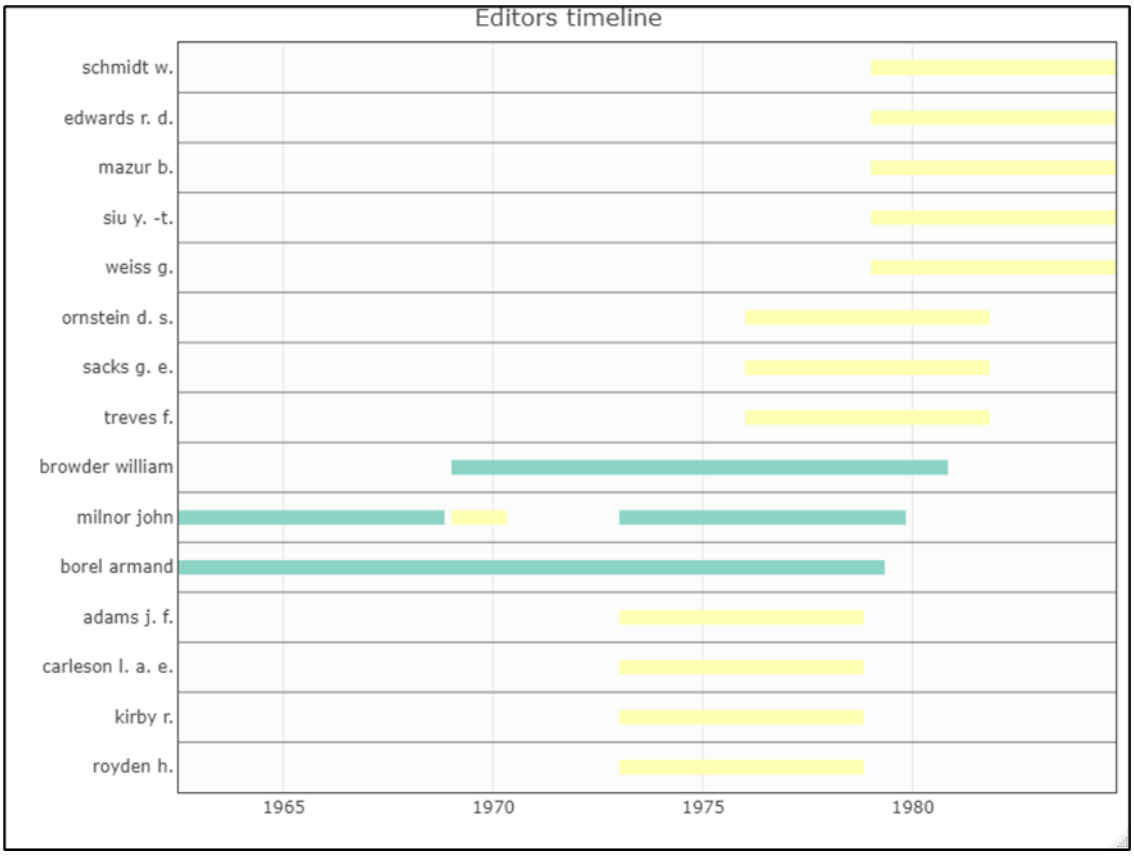

Months
300
200
100
0
95.37
71.03
95.37
71.03
375.40
71.03
95.37
113.57
24.37
83.17
58.83
46.67
34.50
162.27
83.20
46.67
64.97
71.03
71.03
brauer r.
gleason a. m.
kaplansky i.
kelley j. l.
lefschetz solomon
loewner charles
massey w. s.
montogomery deane
samelson hans
selberg atle
spencer donald c.
steenrod norman e.
weil andre
whitehead j. h. c.
zygmund a.
Surname/name

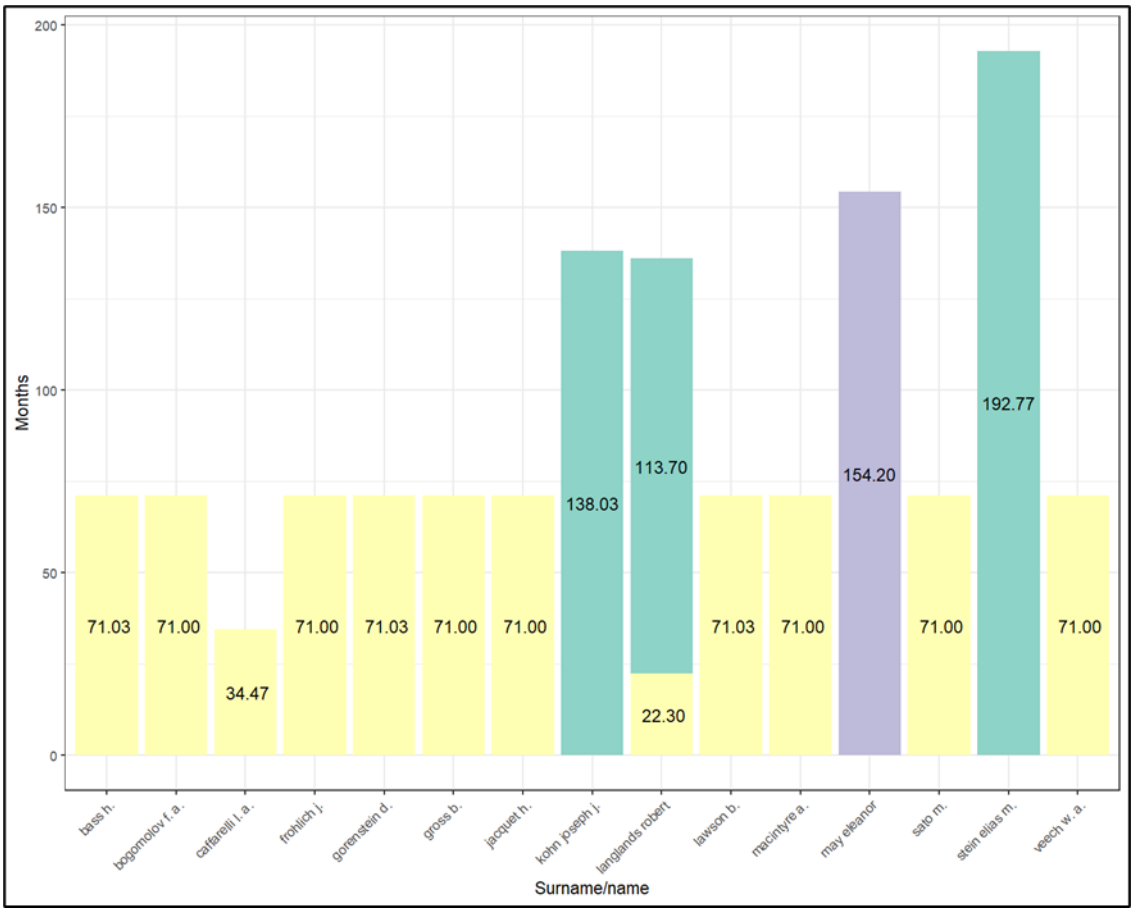

Editors timeline
samelson hans
selberg atle
spencer donald c.
steenrod norman e.
weil andre
massey w. s.
brauer r.
kaplansky i.
gleason a. m.
kelley j. l.
loewner charles
lefschetz solomon
montogomery deane
whitehead j. h. c.
zygmund a.
1930
1935
1940
1945
1950
1955
1960

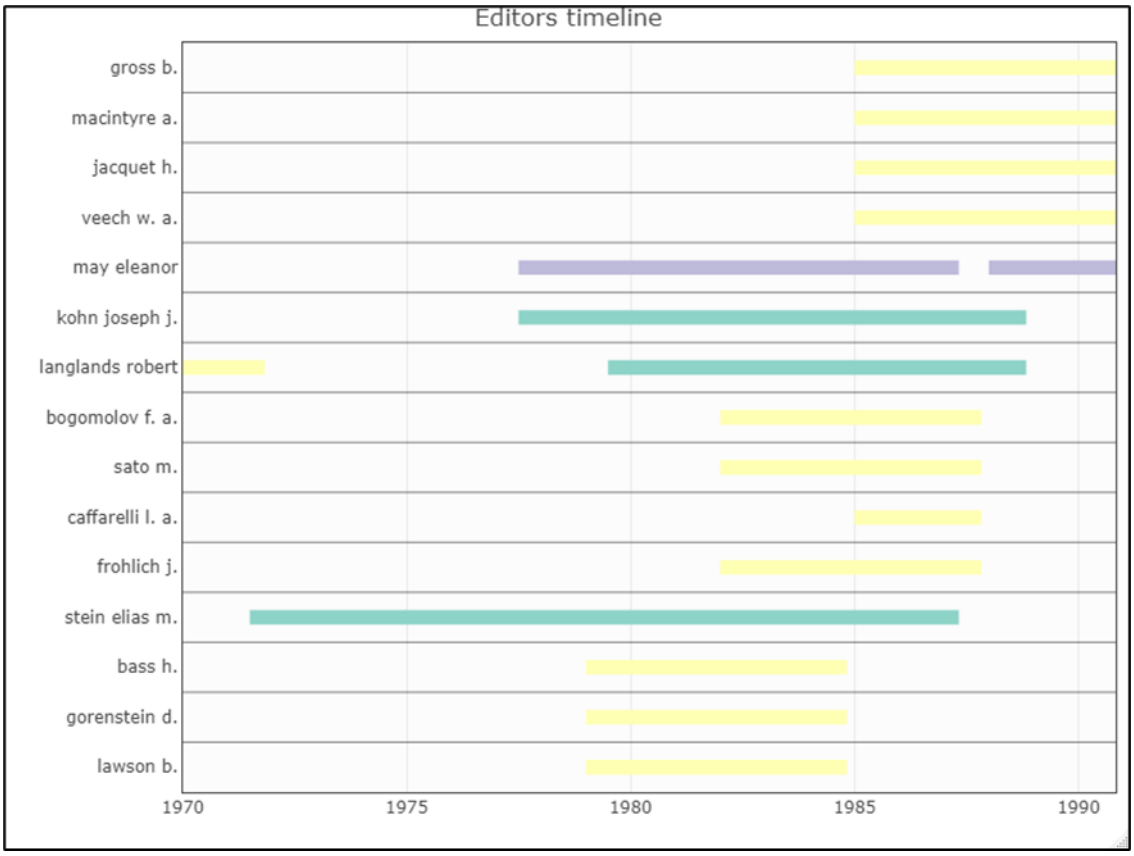

Months
200
150
100
50
0
28.47
71.03
71.03
71.03
71.03
34.50
71.03
217.10
71.03
71.03
64.87
71.00
105.50
71.00
71.00
46.67
atiyah m. f.
bers l.
calderon a. p.
cohen p. j.
feller clara n.
hirsch m. w.
kawada y.
kodaira k.
mumford d.
nelson edward
nirenberg l.
rosenblum fanny h.
singer i. m.
solovay r.
weisenborn eleanor
Surname/name

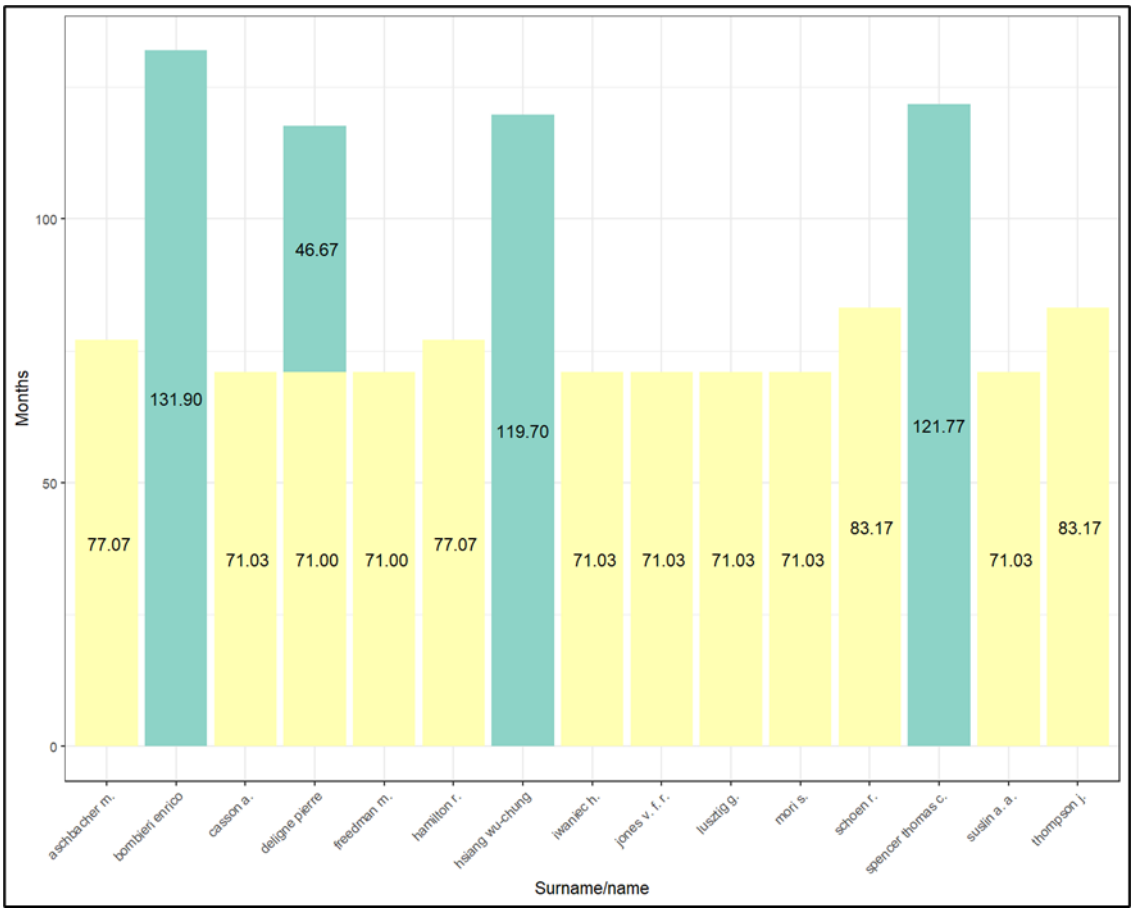

Editors timeline
solovay r.
weisenborn eleanor
kawada y.
nirenberg l.
singer i. m.
feller clara n.
atiyah m. f.
bers l.
hirsch m. w.
mumford d.
calderon a. p.
cohen p. j.
nelson edward
rosenblum fanny h.
kodaira k.
1960
1965
1970
1975

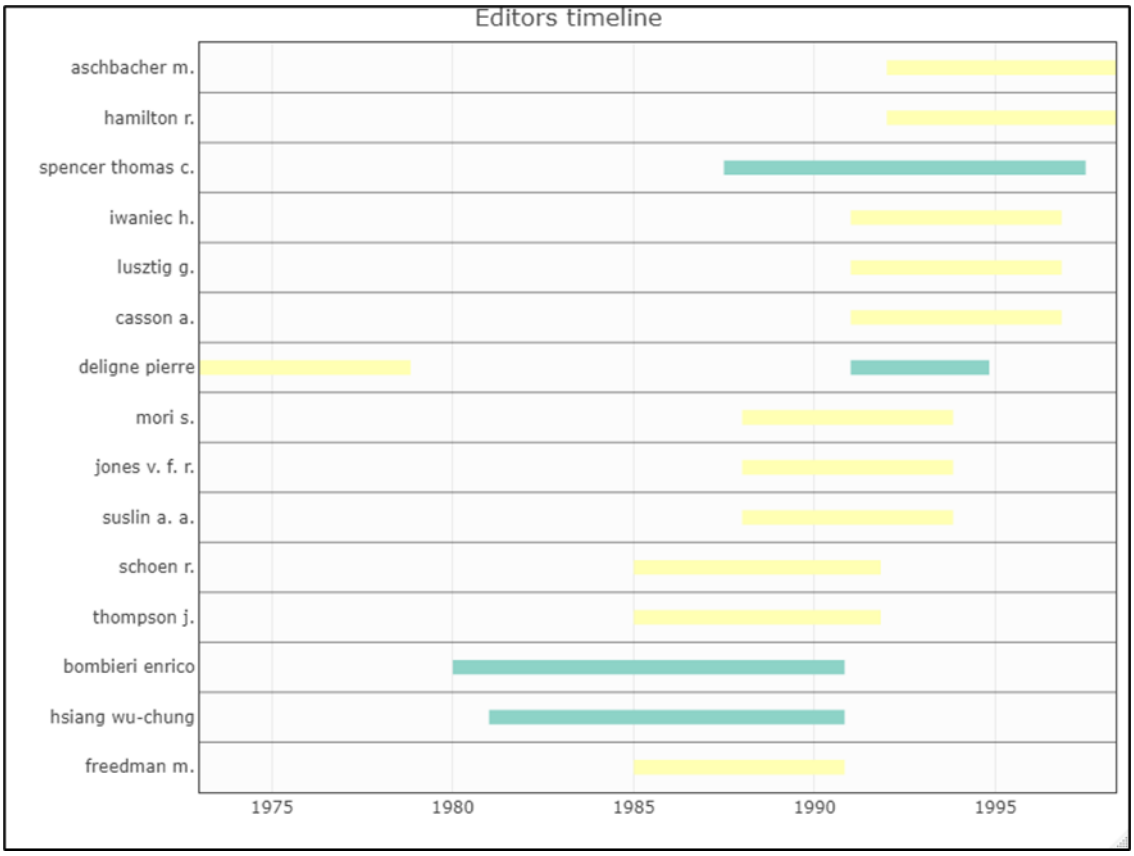

Months
200
150
100
50
0
204.93
144.07
162.37
16.17
71.00
71.00
71.03
71.00
71.03
71.03
71.00
71.03
71.03
71.03
71.03
71.03
71.03
adams j. f.
borel armand
browder william
carleson l. a. e.
edwards r. d.
kirby r.
mazur b.
milnor john
ornstein d. s.
royden h.
sacks g. e.
schmidt w.
siu y. -t.
treves f.
weiss g.
Surname/name

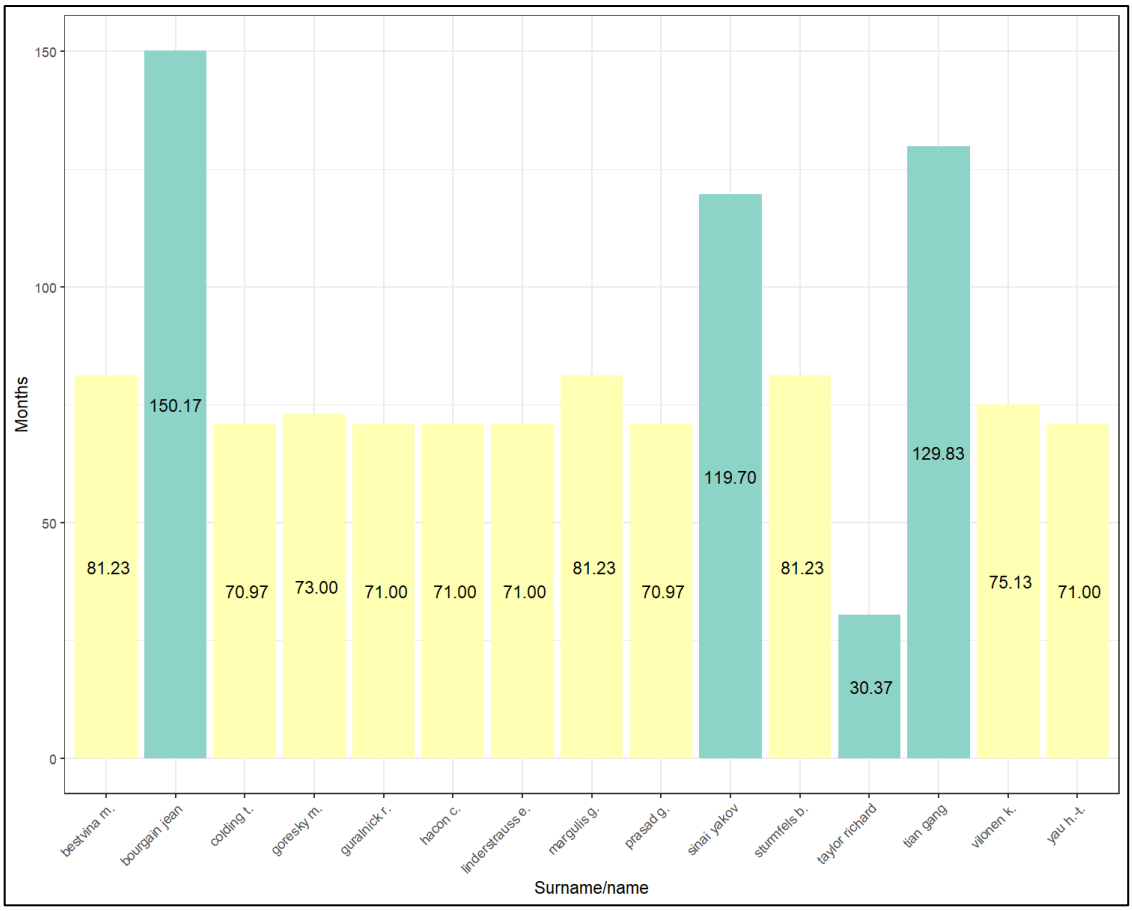

Editors timeline
schmidt w.
edwards r. d.
mazur b.
siu y. -t.
weiss g.
ornstein d. s.
sacks g. e.
treves f.
browder william
milnor john
borel armand
adams j. f.
carleson l. a. e.
kirby r.
royden h.
1965
1970
1975
1980


Months
200
150
100
50
0
71.03
71.00
34.47
71.00
71.03
71.00
71.00
138.03
113.70
22.30
71.03
71.00
154.20
71.00
192.77
71.00
bass h.
bogomolov f. a.
caffarelli l. a.
frohlich j.
gorenstein d.
gross b.
jacquet h.
kohn joseph j.
langlands robert
lawson b.
macintyre a.
may eleanor
sato m.
stein elias m.
veech w. a.
Surname/name

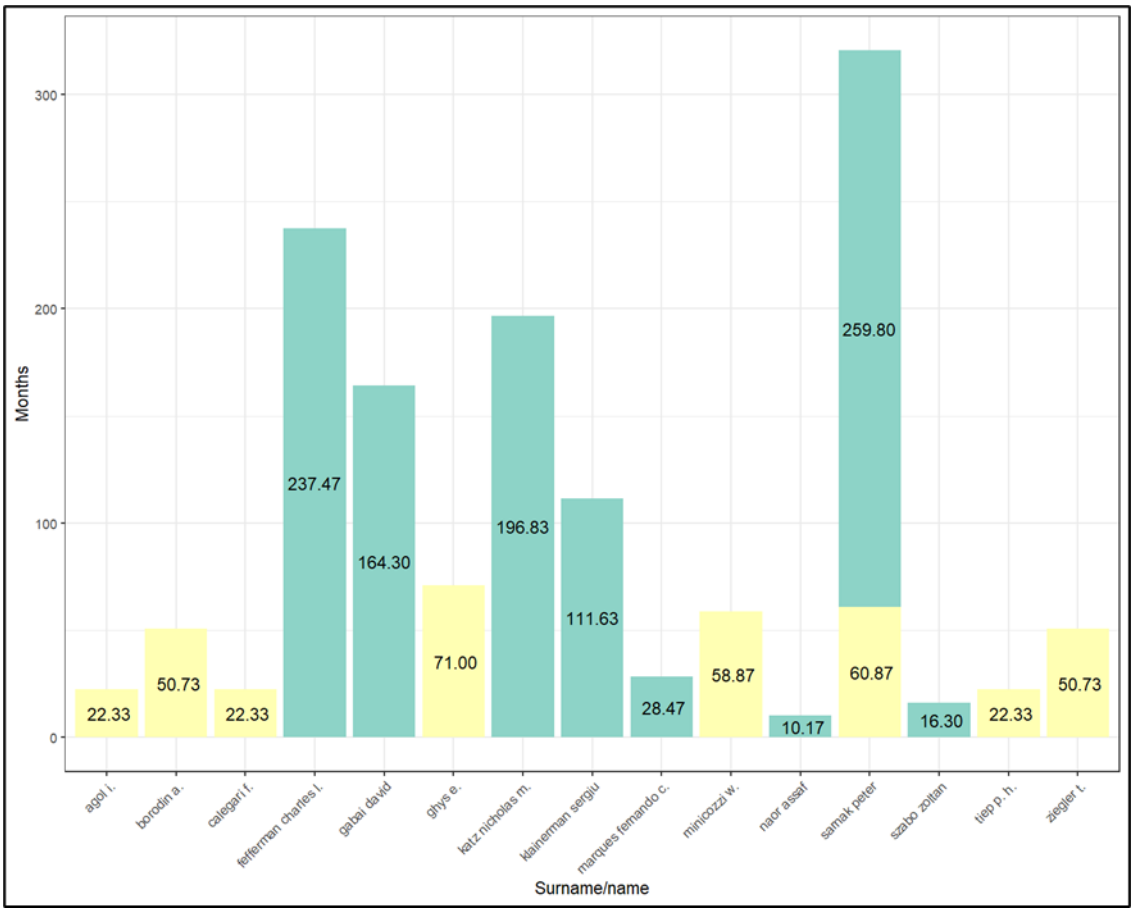

Editors timeline
gross b.
macintyre a.
jacquet h.
veech w. a.
may eleanor
kohn joseph j.
langlands robert
bogomolov f. a.
sato m.
caffarelli l. a.
frohlich j.
stein elias m.
bass h.
gorenstein d.
lawson b.
1970
1975
1980
1985
1990

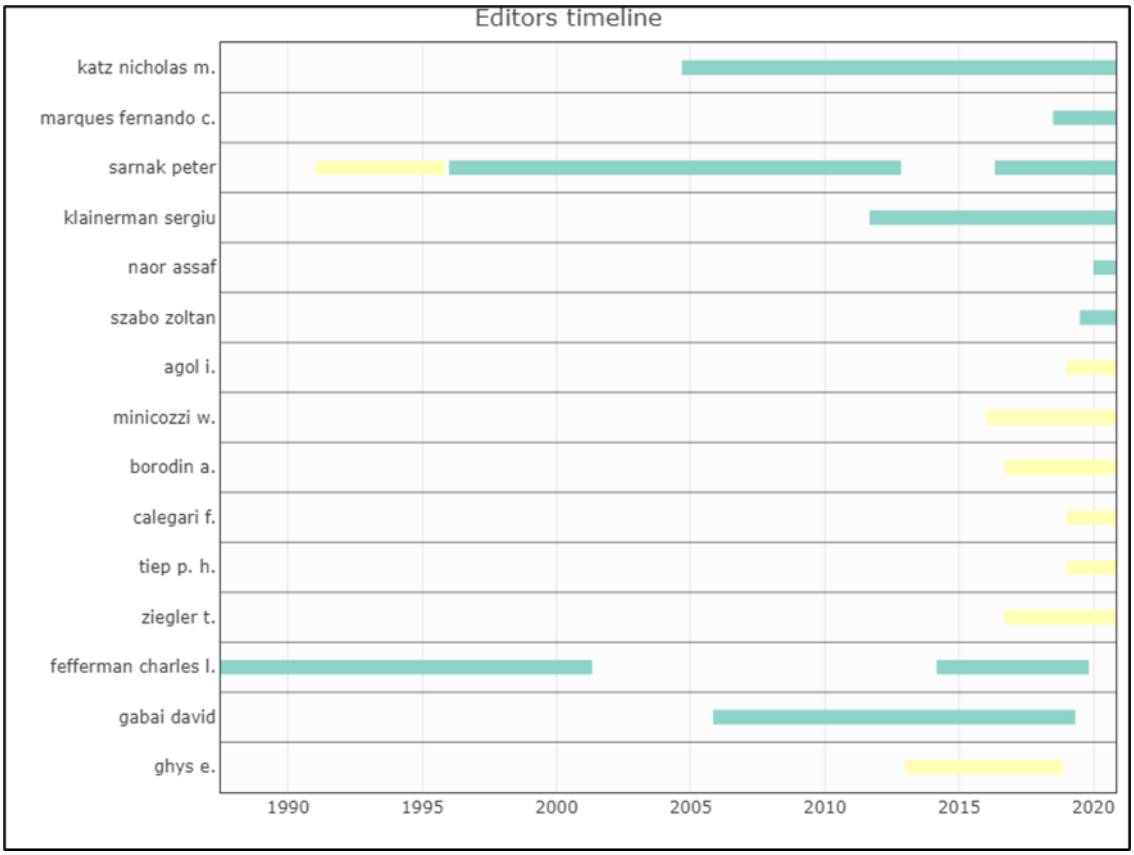

Months
100
50
0
77.07
131.90
71.03
46.67
71.00
71.00
77.07
119.70
71.03
71.03
71.03
71.03
83.17
121.77
71.03
83.17
aschbacher m.
bombieri enrico
casson a.
deligne pierre
freedman m.
hamilton r.
hsiang wu-chung
iwaniec h.
jones v. f. r.
lusztig g.
mori s.
schoen r.
spencer thomas c.
suslin a. a.
thompson j.
Surname/name

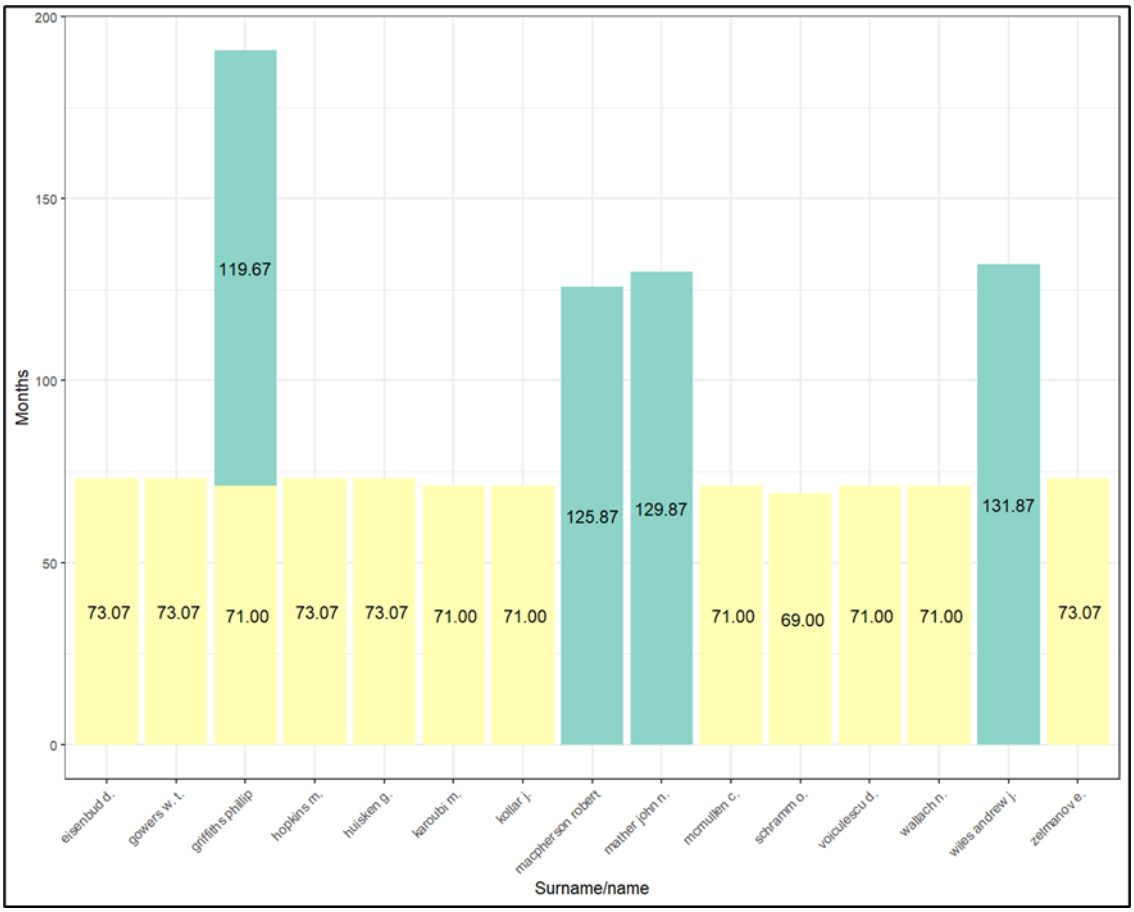

Editors timeline
aschbacher m.
hamilton r.
spencer thomas c.
iwaniec h.
lusztig g.
casson a.
deligne pierre
mori s.
jones v. f. r.
suslin a. a.
schoen r.
thompson j.
bombieri enrico
hsiang wu-chung
freedman m.
1975
1980
1985
1990
1995

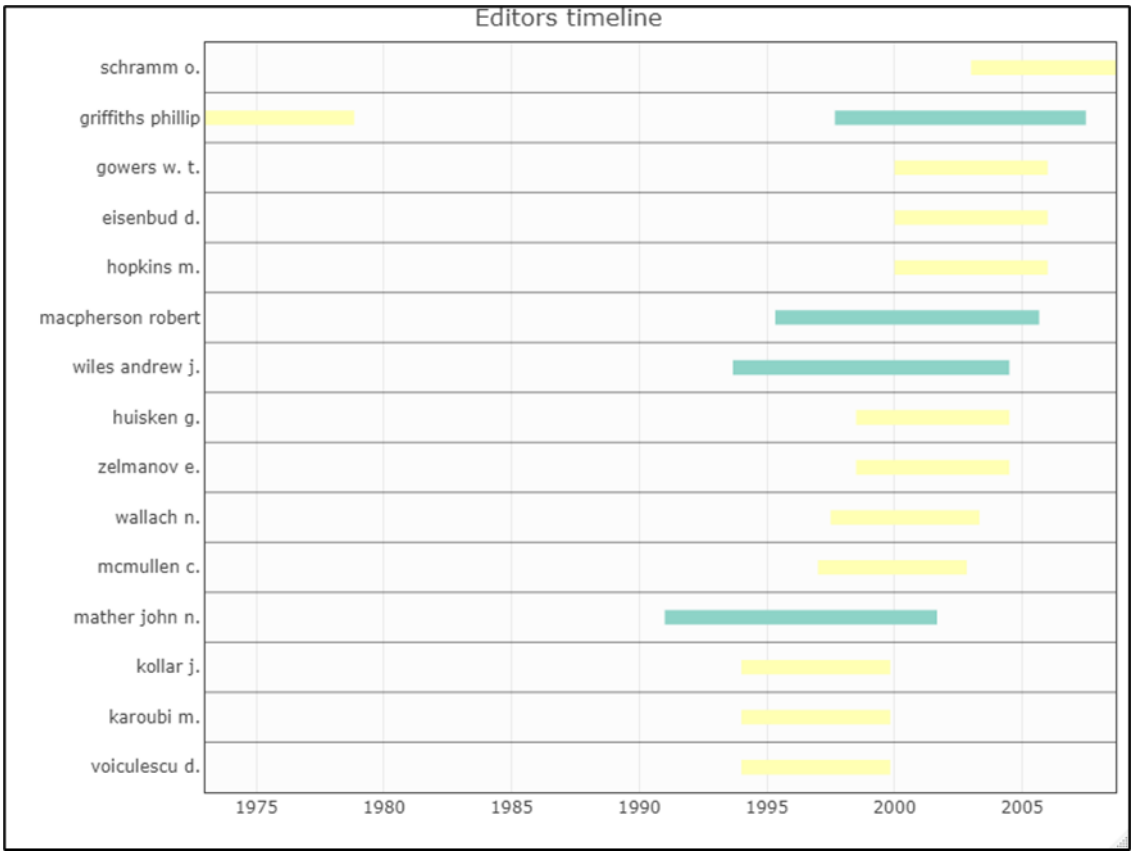

Months
150
100
50
0
81.23
150.17
70.97
73.00
71.00
71.00
71.00
81.23
70.97
119.70
81.23
30.37
129.83
75.13
71.00
bestvina m.
bourgain jean
colding t.
goresky m.
guralnick r.
hacon c.
linderstrauss e.
margulis g.
prasad g.
sinai yakov
sturmfels b.
taylor richard
tian gang
vilonen k.
yau h.-t.
Surname/name


Editors timeline
guralnick r.
hacon c.
tian gang
linderstrauss e.
yau h.-t.
goresky m.
taylor richard
bourgain jean
bestvina m.
margulis g.
sturmfels b.
sinai yakov
prasad g.
colding t.
vilonen k.
2002
2004
2006
2008
2010
2012
2014
2016
2018


Months
300
200
100
0
22.33
50.73
22.33
237.47
164.30
71.00
196.83
111.63
28.47
58.87
10.17
259.80
60.87
16.30
22.33
50.73
agol i.
borodin a.
calegari f.
fefferman charles l.
gabai david
ghys e.
katz nicholas m.
klainerman sergiu
marques fernando c.
minicozzi w.
naor assaf
sarnak peter
szabo zoltan
tiep p. h.
ziegler t.
Surname/name


Editors timeline
katz nicholas m.
marques fernando c.
sarnak peter
klainerman sergiu
naor assaf
szabo zoltan
agol i.
minicozzi w.
borodin a.
calegari f.
tiep p. h.
ziegler t.
fefferman charles l.
gabai david
ghys e.
1990
1995
2000
2005
2010
2015
2020


Months
200
150
100
50
0
119.67
125.87
129.87
131.87
73.07
73.07
71.00
73.07
73.07
71.00
71.00
71.00
69.00
71.00
71.00
73.07
eisenbud d.
gowers w. t.
griffiths phillip
hopkins m.
huisken g.
karoubi m.
kollar j.
macpherson robert
mather john n.
mcmullen c.
schramm o.
voiculescu d.
wallach n.
wiles andrew j.
zelmanov e.
Surname/name


Editors timeline
schramm o.
griffiths phillip
gowers w. t.
eisenbud d.
hopkins m.
macpherson robert
wiles andrew j.
huisken g.
zelmanov e.
wallach n.
mcmullen c.
mather john n.
kollar j.
karoubi m.
voiculescu d.
1975
1980
1985
1990
1995
2000
2005

To be able to resolve the questions in which they were asked, the editorial members who were active in more than two roles were separated to carry out the analysis of the sample. The editorial members and their timeline are represented below:

Figure 6.1 Timeline of editorial members with more than one editorial position.

From ***Table 8*** can be observed all the appearances and the percentage occupied by the editorial positions, as well as the combinations of the editorial positions in the journal. In total, as it said in ***chapter 5 sub-paragraph 2***, they are 36 observations of them that are visualized to the upper timeline. Those observations have corresponded to 23 editors, therefore are more than one combination in editors. The purpose of this analysis is to conclude the order of activities of the specific editorial members. For that reason the observed combinations will be categorized as referred to in ***chapter 4 sub-paragraph 4***, is following table for a reminder.

Main role	Supportive role	Other
Edited by	Associate editors	Founded by
Editor	With the cooperation of	Editorial position
Editor-in-chief	Technical edited by	
Edited and published by		

The recording that will follow concerns the category of editorial position that editors started in the journal and the category of the role that they had when they finished. There is no specific finding as to the direction of the editors in the editorial roles as the results are shared almost equally.

Start from	Finished to	Total Observations
Main	Supportive	10
Supportive	Main	11
Main	Main	1
Supportive	Supportive	1

To be able to understand the sizes of each role, a table will be printed in the form of a waffle table in which all the years that they occupy will be depicted. This graph results from the months that each editor has been active in any position and the numbers you display regarding the total of these values.

associate editors	edited and publish by	edited by	editor	editor-in-chief	editorial board	founded by,associate editors	technical edited by	with the cooperation of
536	120	6621	103	47.6	24.4	38.5	341	11967

Table 10. The aggregate duration of participation in each role by all editorial members is calculated in months.

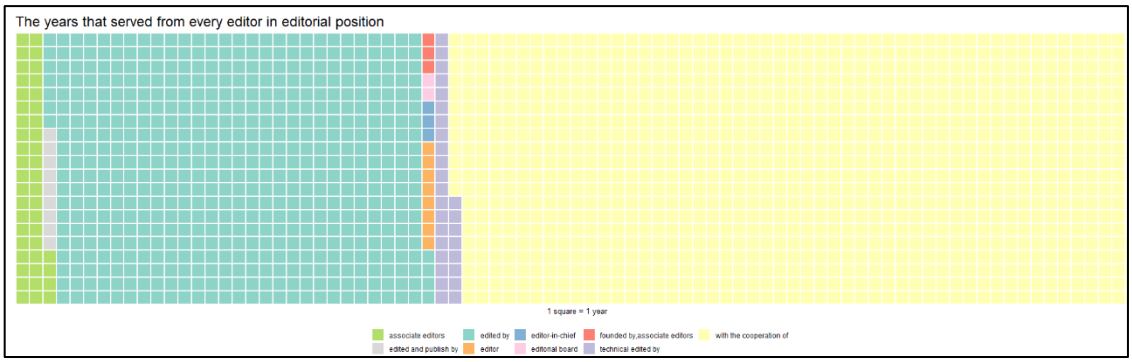


Figure 6.2 Waffle chart from table 10.

It is easy to understand that the roles with the largest activity crowds in the journal during the entire period of their publication were the editorial positions "with the cooperation of" and “edited by”.

BIBLIOGRAPHY

- Ronald Rousseau, Leo Egghe, Raf Guns. (2018). Becoming Metric-Wise: A Bibliometric Guide for Researchers. Chandos Publishing is an imprint of Elsevier, 50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States, The Boulevard, Langford Lane, Kidlington, OX5 1GB, United Kingdom
- Pritchard, A. (1969). Statistical bibliography or bibliometrics? Journal of Documentation, 25(4), 348-349.
- Wikipedia Contributors. (2021). Levenshtein distance. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Levenshtein distance. [18 October 2021]
- Wikipedia Contributors. (2021). UTF-8. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/UTF-8. [18 October 2021]
- Wikipedia Contributors. (2021). ASCII. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/ASCII. [18 October 2021]
- Alvira Swalin. (31 January 2018), 'How to handle Missing Data', [online], towands data science. Available at: https://towardsdatascience.com/how-to-handle-missing-data-8646b18db0d4 [18 October 2021]
- Wikipedia Contributors. (2021). Data. [online] Wikipedia. Available at: https://en.wikipedia.org/wiki/Data. [18 October 2021]
- Devin Pickell. (14 May 2021), 'Qualitative vs Quantitative Data – What's the Difference?', [online], g2. Available at: https://www.g2.com/articles/qualitative-vs-quantitative-data [18 October 2021]
- Naveen Donthu, Satish Kumar, Debmalya Mukherjee, Nitesh Pandey, Weng Marc Lim. (2021). 'How to conduct a bibliometric analysis: An overview and guidelines', Journal of Business Research, vol. 133, pp. 285-296.
- Natalia Andrienko, Gennady Andrienko. (2006), Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach, Springer-Verlag, Berlin Heidelberg
- Michael Galarnyk. (12 September 2018), 'Understanding Boxplots', [online], towands data science. Available at: https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51 [18 October 2021]