HAL
open science

Information visualisation and library data

Panagiotis Gkioulekas, Panayiota Polydoratou

To cite this version:

Panagiotis Gkioulekas, Panayiota Polydoratou. Information visualisation and library data. ELPUB 2019 23rd edition of the International Conference on Electronic Publishing, Jun 2019, Marseille, France. 10.4000/proceedings.elpub.2019.7 . hal-02142262

HAL Id: hal-02142262

https://hal.science/hal-02142262v1

Submitted on 28 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Information visualisation and library data

A case study of Public Library of Veria, Greece

Panagiotis Gkioulekas and Panayiota Polydoratou

The authors would like to thank Dr. Sarah Anne Murphy, Ohio State University, for her valuable suggestions and information on data visualization, Ms Maria Tolika and Ms Aspasia Tasiopoulou, Public Library of Veria, Greece for their support and for the provision of data.

Introduction

1 Technology and the vast production and availability of online data compel that libraries develop new services that give access to such resources. In addition, libraries hold data about their users, the users' characteristics, their information needs and information use which can be assisting to libraries to better manage their resources and services. However, it is often difficult to interpret such data in a way that is cost effective, valid and at the same time easily readable by both administration and the lay person.

2 Many scientists propose that information visualizations could be regarded as a solution for this problem. According to Card et al. (1999), information visualization is the use of computer-supported, interactive, visual representations of abstract data to amplify cognition. Information visualization allows users to extract data from large databases in an efficient and effective way. Furthermore, the comparison of datasets is more distinct when it is presented through visualization. The usage of computer graphics with a combination of the right label and some parameters such as the color, the distance and the size make the comparison clearer.

3 For libraries, the advantages of visualizing library data include: a) a better understanding of their users, resources and services; b) the potential to improve their services through such process; and c) a better understanding of and a more efficient presentation of the library's data collection entirely.

4 Moreover, library data visualizations could prove a useful tool when it comes to enhancing the decision-making process and the directors/administration could make more successful operational choices.

5 Some of the limitations and the most common problem one can encounter when trying to combine information visualization tools and library data is often the financial, technical or even human resources themselves. Due to lack of specialized IT skills and the necessary knowledge to develop information visualization, the majority of libraries are forced to quit the idea of information visualization or rely on external IT professionals, a process that could increase the libraries' costs. Through this preliminary study we aim to identify potential solutions and/or suggestions to address these limitations.

Library background and data

6 The Public Library of Veria, Greece serves more than 35.000 members residing in a population of approximately 72,000 persons. The library's collection amounts over 70,000 volumes. It has got two branches and two mobile libraries. In 1996, Veria's library was the first that provided access to the World Wide Web in Greece. Its own website has been running since 1997. Furthermore, it has received the Bill and Melinda Gates Foundation "Access to Learning Award 2010" along with a $1,000,000 dollar donation. The Public Library of Veria city, Greece, whose collections' volume, complexity and dynamism are increasing dramatically, requires new tools for analyzing and understanding such data.

Literature review

7 The evolution of technology and the role of libraries to provide access to information, even more publicly and openly nowadays, and serve their users' information needs in an easily accessible and retrievable way is documented in the literature. Already Bates (1989) discussed how the browsing techniques for the online search interfaces could change in order to be more serviceable. However, more and more people show an increasing interest about the role of information visualization on the World Wide Web and the potential benefits that could have for libraries and the services they provide. In a similar effort, Card et al. (1999) offered guidelines for designing information visualization applications. The researchers emphasized ways to strengthen the use of the visualizations, specifically to deploy methods, so that by using vision one can be assisted to interpret information and think in detail and depth. In addition, these guidelines helped us to create the visualization of the current paper using a "reference model," in which three sub processes are included. First, there is the data transformation process, which transforms raw data into data tables that offer structure and ease of manipulation. Second, data tables are mapped onto visual structures that include the application of spatial environment, marks, labels and graphical properties and the last sub process, visual structures are transformed into figures, which involve graphic parameters, such as position, scaling, and clipping. Beagle (1998) claimed that libraries are the principal information centers, and through the descriptive information they produce (metadata) are therefore very important players when it comes to visualizing metadata to assist operational purposes, for example to assist in the formulation of the library's mission. In addition, ALA recently published the articles of H. Chen, (2017) about information visualization and libraries. These articles provide an introduction in the information

visualization that could be used by a library. Moreover, extensive literature has also discussed the best practices and efforts related to becoming data librarians (Rice and Southall 2016). The articles of Sarah Murphy (2013, 2015) are also crucial because they analyze tools, such as Tableau Public (http://www.tableau.com/) which was used for this study, that could enhance the decision making process. Furthermore, Murphy stated that information visualization is important for librarians, especially those tasked with assessment, marketing, and other tasks that require considering and understanding data. With the help of information visualization, librarians can leverage large amounts of data that were once impossible to efficiently access and manage. Murphy added that with the help of information visualization, librarians can create visual representations of data which can be very useful and informative. Information visualization will allow them to perform advanced searches of their collections on both complex and more specific topics. Also, in order to be able to decide on be most suitable visualization program is significant to refer to Shneiderman's article (1996). Lastly, public libraries have been also known for facing various challenges such as low financial support and limited resources. O'Malley Voliva (2015) states that these problems can affect the use of information visualization by the public libraries. Therefore, a review of available tools and the selection of a suitable one is important. Also, she states that information visualization is as beneficial for the staff of libraries as it is for their users. Voliva adds that information visualization not only makes data more accessible to everyone who works at the library, but it also helps create transparency across the organization. This is because everyone in the library is able to become a part of decision making processes. Moreover, there is an article of Kyrilidou (2016) that noted the large amount of money that the libraries spend to purchase licensed e-resources. Kyrillidou indicates that by the usage of information visualization, libraries can easily reduce the cost of purchasing e-resources. The example she offers is how information visualization helps libraries identify their users' behaviors about the use of their e-resources. In this study, we created visualizations about some features of the users. In the following section we describe our choice based on specific criteria for a public library.

Methodology

8 The creation of information visualizations requires software that can present all aspects of the data as well as the connections between them. There are several tools that offer various options from which one can choose in order to produce visualizations. However, financial constraints are always a factor that has to be accounted for during the selection of such a tool. For this study, the selection criteria were primarily available open source tools and their utility options.

9 At the beginning, we performed a comparison test of such software. The compared tools were "Many Eyes", "Data Wrapper" and "Tableau Public". There were both advantages and disadvantages for all of them but "Tableau Public" software deemed most suitable for Public Library of Veria data based on the criteria of accessibility (open source), ease of use and support (implementation and use guides) as well as utility options (e.g. visual information display tables that offer filtering and annotation options). The possibility to combine data from different sources to create dynamic graphics and the provision of help with tools and videos (http://onlinehelp.tableau.com/current/pro/desktop/en-us/help.html) were also important. In our decision, pivotal was also the interaction with Dr.

Murphy at Ohio State University who kindly, shared details and suggestions from her experience working with this software.

10 For information about "Tableau Public", background information, awards that it has won and how to become a member and use it one can visit the tool's webpage at http://www.tableau.com.

11 In parallel with the tool's selection there were also meetings and discussions with Veria's Library staff about the data that were available and suitable for this study as well as the appropriate types of visualization. The data were extracted from the library's databases. The Public Library of Veria provided the data to us. Personal data of users, such as nationality and profession, were gathered during their registration at the library and the users consented to their data's usage for such purposes (surveying, analyzing usage patterns, improving library services, etc.). The rest of the data that were used (borrowing, circulation, financing, etc.) were also collected and provided to us by the administration of the Public Library of Veria. Specifically, the data that were made available by the library for this study are related to the nationality of the users, the profession of the users, circulation materials, visitors' access to the library, financial reports (income and expenses) and OPAC's searches (e.g., keyword/phrase searches)., for instance, the way of patron's searching. The data files were in .doc type (MS Office Word) and they were converted in .xls type in order to import and edit them with Tableau Public.

12 The goal was to agree on visualizations types that would be comprehensible and useful to everyone. We are able to create data visualizations which are equally useful and accessible by staff members who feel they were somewhat inexperienced in reading data visualizations.

Case study

Visualization Results

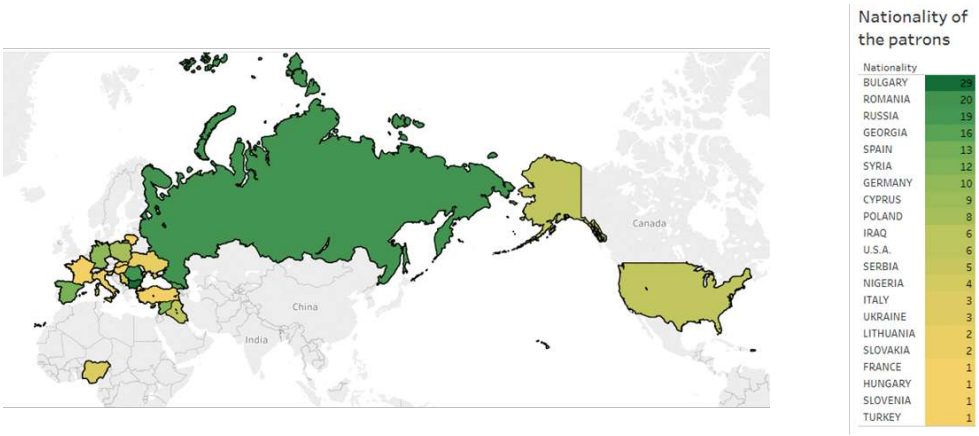
13 In the following sections, we present the visualizations' results; we discuss them and also provide some insights and suggestions how this could be used by the library. Specifically, the visualisations refer to some users' characteristics (nationality and profession), registrations (at branches and the mobile library), borrowing materials (range of time), financial operations of the library (income and expenses) and catalogue searches.

User characteristics (nationality, profession)

14 This visualization shows the nationality of the users. The green colour represents the countries from which most users originate and the yellow one represents the countries where the fewest users come from. The visualization type that is used is "map". Veria is geographically located in the north east of Greece. The Public Library of Veria serves a considerable number of non-native habitants. This information could enhance the decision-making process and for acquisitions and collection development. Looking at the registered users who are not Greeks and have declared their nationality, one can see people originating from Bulgaria, Romania, Russia, Georgia, etc. The library could make decisions about purchasing material from authors of such nationalities or even organize cultural events between such countries for example. Another example is the observation of the increasing number of users from Syria, most of them refugees. The Public Library

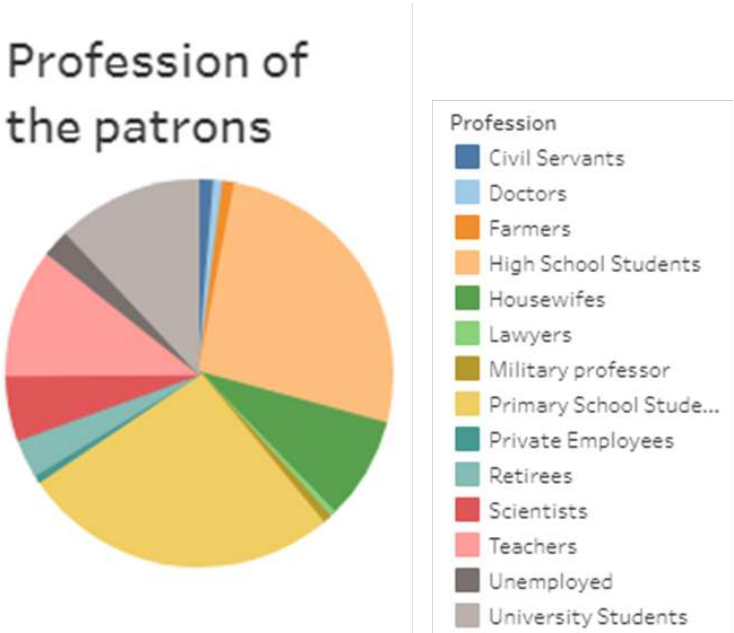
of Veria could provide Greek language learning programs, school support activities as well as job seeking services to these people in order to integrate them into the social environment.

Figure 1: Nationality of the users



15 The second visualization indicates the profession of the users. Each colour of this "pie" chart corresponds to a specific profession. The Public Library of Veria could extract some information from this visualization and subsequently adapt its services to suit more users' professions. For instance, the large number of pupils can lead the library to purchase more children's and teenagers' literature and perhaps engage in educational activities for specific student groups. In addition, the Public Library of Veria can discern the small percentage of farmers who are members of the library. Veria is a city based on the rural economy, so the small number of farmers shows that the library should adapt its services and material in order to stimulate their interest.

Figure 2: Profession of the users

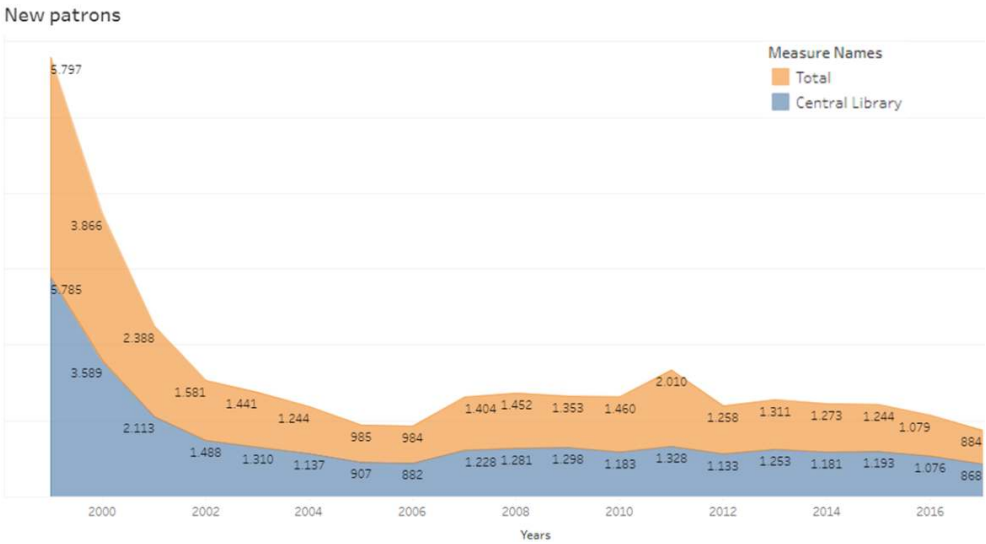


Registrations

16 Through the next visualizations, the librarians of the Public Library of Veria can examine the efficiency of each branch of the Library when it comes to registration of users as well as any increase or decrease of user numbers and what this could imply for the library services and/or user information needs.

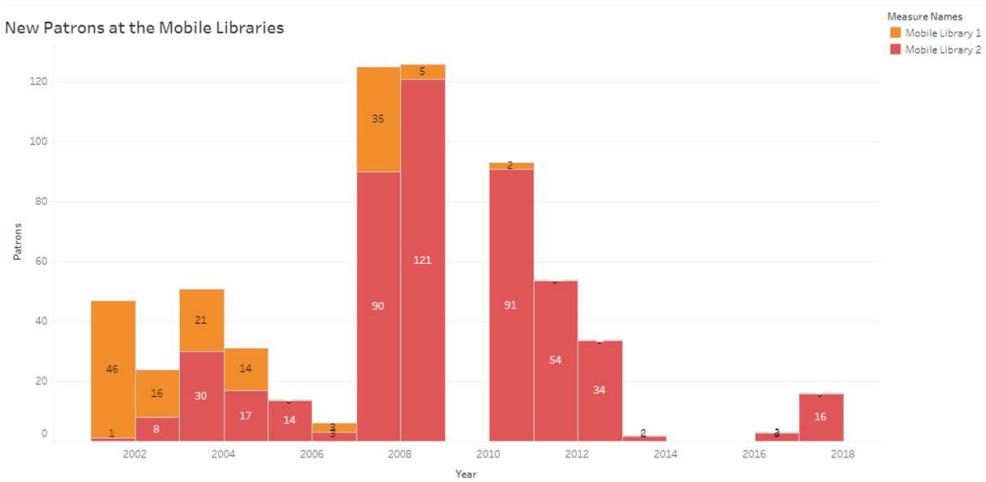
17 The visualization type that was used in figure 3 was "area chart." The purpose of this visualization was to present the correlation between the total registrations and the registrations of the central library. The use of two different colours helps the reader to observe the data and their connection instantly. It is important to note that if we want to add more data to this visualization, we could do it from the Excel immediately.

Figure 3: Users' registrations at the Central library vs branches



18 Figure 4: The Public Library of Veria has two mobile libraries (or "bookmobiles") which are vehicles designed for use as a library. Their role is to expand the reach of the library by transporting books to potential readers, providing library services to people in underserved locations. There are two of them in order to cover more remote areas. Figure 4 shows a comparison between the mobile libraries concerning the new registrations. In this way, the library can examine which regions have more registrations and then to plan more bookmobile routes at that areas. In addition, focusing on a better comprehension, the "stacked bars" type was used. The advantage of this visualization is the presentation of two data simultaneously, the first one is the total registrations at the mobile libraries and the second one is the registrations at each mobile library. In this way, the process of analysis is accelerated.

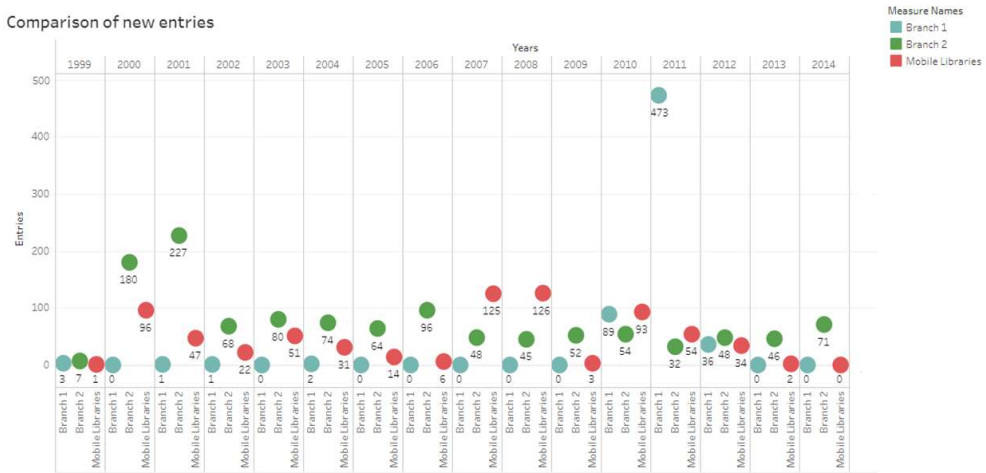
Figure 4: Comparison of the registrations



19 Figure 5: An attempt was made to examine the three regional libraries (the two branches and the mobile library) on new registrations by the way of "circle views" type. The "circle view" is another visualization option for comparative analysis. Figure 5 has quite a bit of information packed into a single visualization, such as the registrations of each branch,

the registrations per year, the comparison between them. The position and the colour of the circles combined with the number of the registrations make the comparison more noticeable. Every branch can obtain access to this visualization for greater transparency in the organization and everyone can take part in the decision making process. The "circle view" is a tool with enables us to include several different fields into a visualization.

Figure 5: Registration of users at the regional libraries

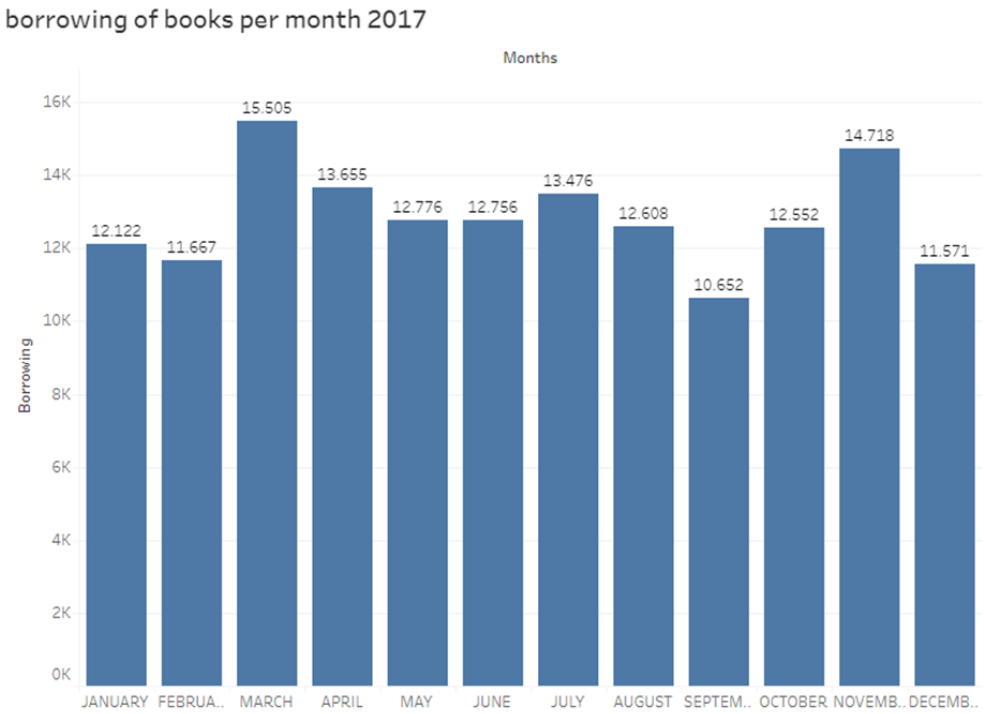


Borrowing

20 Visualizations that depict circulation data are shown below. The term "Borrowing" refers to all the material, such as books, vinyl, e-books, audiobooks, DVDs, journals, which the users have borrowed. Librarians and users have access to these data. For this reason, the visualizations should have simple structure in order to be comprehensible by everyone.

21 Figure 6: In this figure once can see a definite presentation of the number of the borrowing items per month on 2017. The user understands the flow of borrowing immediately. It also is an indicator for library staff about the months that the library material was more in request and also, due to this demand, perhaps library space was more in use. The bars highlight the months with the most borrowing items and the number above each bar indicates the exact number of them.

Figure 6: Borrowing of books 2017 (monthly view)



22 Figure 7: In this figure, the "lines" type was used. It is a simple type that shows the trend of borrowing the last seven years. In this way, the users will not get lost in the complexity of visualization. Figure 7 shows that the number of borrowings had been raising until 2012. Since then, there was a declining path of borrowing activities until the previous year. This data can trigger an interrogation of any specific events or factors that could have affected library use during this period (2013-2017) and seek solutions to any causes/ problems. More detailed information (monthly borrowings) is presented in Figure 8. This visualization presents a comparison between annual borrowings per month. The visualization type is "side by side bars". Each year has a different colour to make the comparison of data more distinct.

Figure 7: Borrowing of materials 2010-2017 (yearly view)

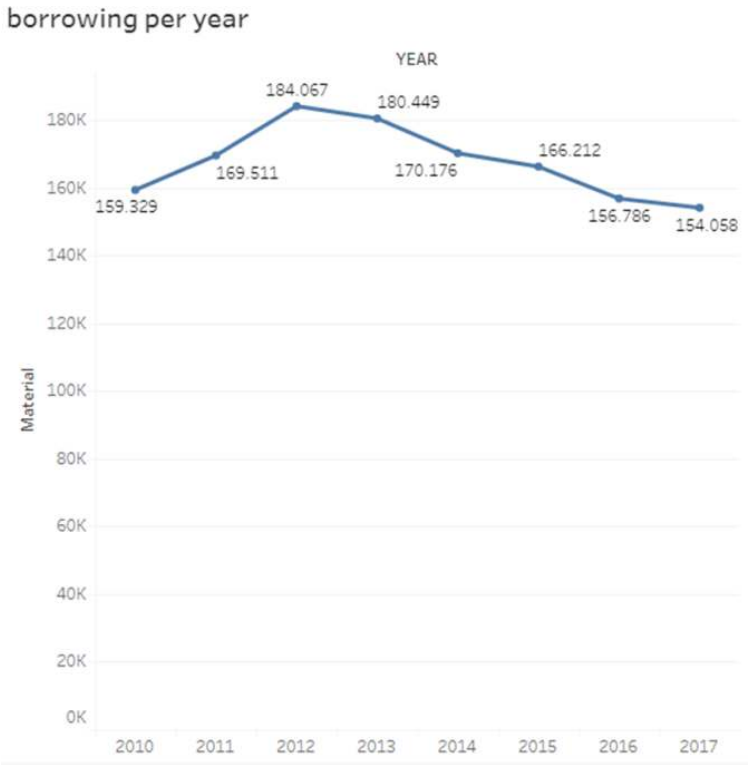
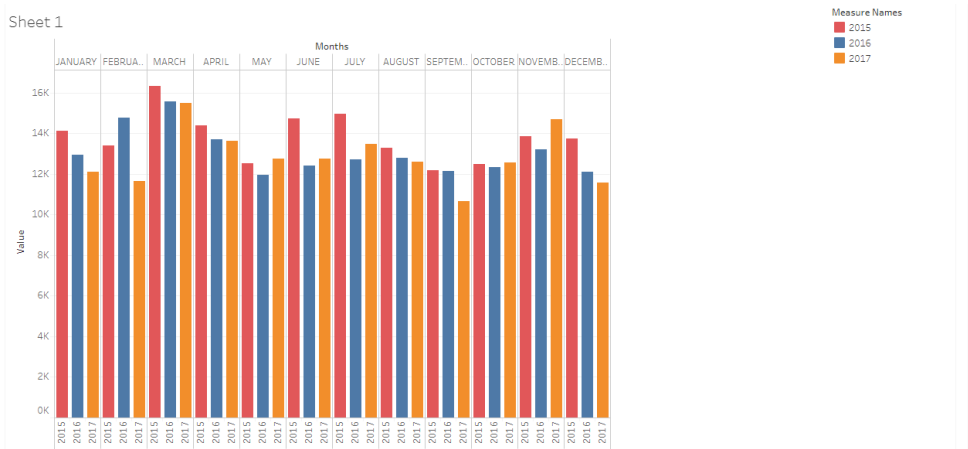


Figure 8: Borrowing of books 2015–2017 (monthly comparison)



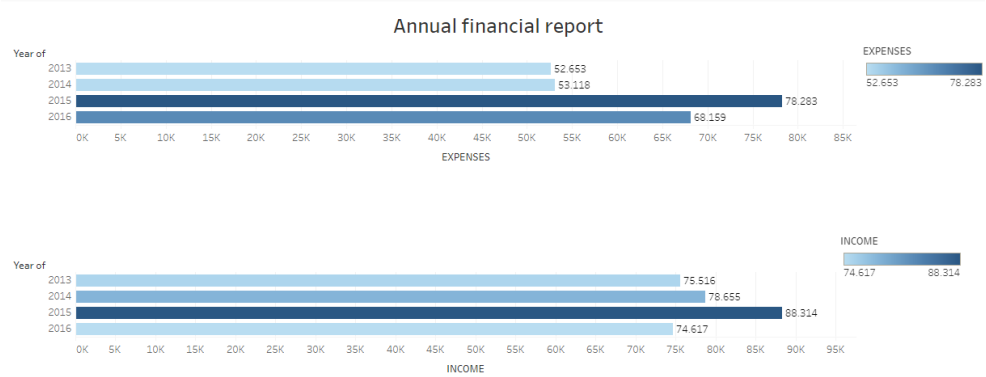
Financial operations

23 Figure 9: The visualization aims to deliver a more complete outlook on the library's finances and more specific the expenses and the incomes, providing numerical data in a format that is much more comprehensible and intuitive than the current setup. An approach like the one in Figure 9 would lead to better understanding of the financial data which will in turn lead to more informed decision and smarter investments. In these days this is very important when funding for libraries is inadequate. The visualization was based on data from the annual reports of the Public Library of Veria relating to its

incomes and expenses for the period of 2013–2016. Due to the fact that finance is a sensitive issue, there could not be any more detailed data.

24 For the visualization of financial data, the "stacked bars" type was used in horizontal format. The expenses and income data are displayed together in order to watch two interrelated items and avoid any error that will cause problems in the Library's financial policy.

Figure 9: Annual financial report



Catalogue search

25 The process of accessing the library's catalogue can be very time consuming. Very often, users are faced with problems due to their lack of familiarity with the library's environment. The main aim of the next visualization is the improvement of the Public Library of Veria's catalogue search. More specifically, the primary objective is to provide a more direct and easier way to find out books based on the author's origin country. Figure 10 was created by the "symbol map" type and shows the number of the writers of each country whose books are in the Public Library of Veria. This visualization could connect with the catalogue of Public Library of Veria so the user could select "Russia", for example, and connect directly to the page of the catalogue named "Russian Literature". For better visualization rendering, it has been necessary to create one that actually works to magnify the main visualization.

Figure 10: Authors by country

Conclusions

26 Technology and the vast production and availability of online data demand that libraries develop new services that will provide easy, efficient and effective access to information that covers the needs of their users. Thus, the procedures of the libraries' modernization should be accelerated and the access to information should be more exciting for the users. To achieve this, libraries need a powerful tool like visualization. It is equally important that librarians should enhance and further develop skills that will allow them to provide such services. Staff development in general but also staff motivation by the libraries to participate in training sessions that are offered by computer software companies should be encouraged. By investing in library staff training on the use of information visualization technologies and tools, libraries can benefit both their staff and the users of library resources. Staff is benefiting by becoming proficient in a new skill set, and users are managing better the information that the library is creating.

27 In this paper, we considered the effect of the information visualization analytics on the operations and activities in Public Library of Veria. Also, our study adopted information visualization techniques using the Tableau Public software to analyze the data. Information visualization has been proven as an effective way to analyze unstructured data, identify outliers and discover new insights. The libraries and the information science community show an increasing interest in the usage of the information visualization to better analyze and understand the library data (Murphy 2013). Every library can create visualizations that will help make it more functional. At the beginning, a library can use the information visualization on a smaller scale to capture their metrics, such as done for this paper. After, the library can start moving to more composite metrics for instance, the attendance of the programs, public computing use and free Wi-Fi consumption. The use of information visualizations can also be extended to digital

libraries. With the number and types of documents in digital library increasing, tools for organizing and presenting the content have to be used. One of these tools could be the visualization that focuses on topic-based organization and structuring. Visualizations have been developed and evaluated for accessing a digital library containing thousands of documents and terabytes of data.

28 Our goal is that this current study provided some preliminary examples of how information visualization tools and library data can be combined. Finally, since our findings and the visualizations are based on data from one public library further research in different, expanded settings and contexts are suggested.

BIBLIOGRAPHY

References

Card, S., J. Mackinlay, and B. Shneiderman. 1999. *Readings in Information Visualization: Using Vision to Think*. San Francisco: Morgan Kaufmann.

Bates, M. 1989. "The design of browsing and berrypicking techniques for the online search interfaces." *Online Review* 13 (5): 407–424

Beagle, D. 1998. "Visualization of Metadata." *Information Technology and Libraries* 18 (4): 192–199.

Chen, H. 2017. "Chapter 1. An Overview of Information Visualization." *ALA TechSource* 53 (3). https://journals.ala.org/index.php/ltr/article/view/6288

Rice, R., and J. Southall. 2016. *The Data Librarian's Handbook*. Facet Publishing.

Murphy, S. 2013. "Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making." *Journal of Web Librarianship* 7 (4): 465–476

Murphy, S. 2015. "How Data Visualization Supports Academic Library Assessment: Three Examples from the Ohio State University Libraries Using Tableau." *College and Research Libraries* 76 (9): 482–486.

Shneiderman, B. 1996. "The eyes have it: a task by Data Type Taxonomy for Information Visualizations." In: *Proceedings 1996 IEEE Symposium on Visual Languages*, 336–43. https://doi.org/10.1109/VL.1996.545307

O'Maley Voliva, C. 2015. "Data Visualization for Public Libraries." *Public Libraries Online* 5 (20 April). http://publiclibrariesonline.org/2015/04/data-visualization-for-public-libraries/

Martha Kyrillidou. 2016. "Business Intelligence and Data Visualization with Tableau in Research Libraries." *Research Library Issues* 288:1–4.

ABSTRACT

This paper comprises an attempt to create information visualizations for the Public Library of Veria city, Greece based on library data. Specifically, it is a formative study about the use of a visualization tool for analysing library data. The goal of this study is to make suggestions for the presentation of the public library's data in order to be more accessible and understandable by the users and the librarians. The data were provided by the Library and the data types that were used were selected after a thorough consultation with the library staff. Consideration was given to the availability of data. The study employed information visualization techniques to create the visual displays of the data. Visualizations were created using the Tableau Public software in an effort to provide a quantitative, analytical, and evidence-based view of how libraries could manage their data. The study also aimed to gather any additional potential uses of visualizations that can be exploited by the libraries in the future. The current research was conducted in a single public library, thus further research in different, expanded settings and contexts is suggested.

INDEX

Keywords: visualization, information visualization, public library, library data, Public Library of Veria, Tableau Public software

AUTHORS

PANAGIOTIS GKIOULEKAS

Department of Library Science and Information Systems, ATEI of Thessaloniki, Greece
panagiotisgkioul@gmail.com
(corresponding author)

PANAYIOTA POLYDORATOU

Vienna University of Technology, IFS, Vienna, Austria
ATEI of Thessaloniki, Greece