

# Baby Names Explorer - Final Report

Magdalena Pöhl, Tsafou Kyriakoula and Michael Oppermann

**Abstract**—The Baby Names Explorer is a web - based application created for the purpose of the course Visualization of the Department of Informatics at the University of Vienna, in the spring semester of the academic year 2013 - 2014. This visualization is aimed at highlighting the similarities between the hundred most famous names among the cities of Vienna and England, for the time interval from 2000 to 2012. The presentation of the results is made by using graphs and tables with a view to optimize the visualization tools, thus giving the user a simple and easily understandable application.

**Index Terms**—Naming trends, baby names, naming globalization, visualization tools, slopegraph, linegraph

---

## MOTIVATION

It is increasingly observed nowadays the globalization phenomenon even to the naming patterns, so that children are more and more citizens of the world. We firmly believe a culture is mirrored in naming practices and changes in society are reflected in the names chosen by the parents. According to our observations from everyday life as well as from different surveys or articles, there is a considerable speculation as the parents realise that their offsprings will be part of a more wide - ranging network and not only of their specific home town, which affect the name choices about them.

The rapid evolution of technology providing Internet access almost to all the developed parts of the world either by clicking a mouse, or by touching a small screen, the film and soap series production growth and the advertisements which dominate in every aspect of our social lives are some of the major factors that result and intensify the internationalization in our changing society. Consequently people being influenced by the same sources tend to make decisions about names in a similar way, often following their favorite celebrities' standards.

Thinking about English as the first internationally spoken language, we found the comparison, about the most common baby names, between Austria and United Kingdom, as a reasonable one. Our goal is to study one German and one English speaking country, different in their culture, manners and customs, habits and point out any similarities and disputes, concerning the naming tendency. Vienna and England are selected for the data collection and processing as the most representative examples for their countries and at this phase of the project the interval of the time is focused from 2000 to 2012. Furthermore a distinguish between the sexes is desired, as there is a long-head belief that people tend to be more conservative when choosing a name for a son than for a daughter.

The extent of use of our project idea range between all those social scientists who are interested in exploring the naming analogies or differences between these two countries, in an easy and comprehensive way. Young students with specialization in sociology or specifically making some demographic studies can apparently be assisted and be offered useful findings. Via our tool, provided just in a web interface we also inspire all these young parents who would like to get some ideas in naming their children.

It is worth mentioning at this point that all the necessary information that are imported and used for our diagramming and statistical conclusions were found through the Internet. Specifically the Austrian Open Government Data Portal (<http://data.gv.at/>) was used for all the necessary information on Vienna data collection. About the England's statistics the equivalent British one (<http://data.gov.uk>) was needed. Particularly noted that the data on England were treated appropriately by us, in order to isolate only the ones we were interested in our research and to exclude other cities references, maintaining the extent of our project at reasonable levels for the current purposes.

## 1 RELATED WORK

Before the start of reporting the related work, which inspired our project effort, it is thought as useful to mention that given names have not been widely analyzed so far, because of the privacy security restrictions and therefore the unavailability of appropriate data collections. For these reasons and for the fact that the operators which influence the name selection vary constantly, it is still an open study object, by sociologists.

Despite that, there are two primary portals that someone is able to explore baby names, also with the help of visualization tools via some graphs and heat maps. One can find them under the names of "Name Voyager" (<http://www.babynamewizard.com/voyager>) and "Name Trends" (<http://nametrends.net/>). We shortly present each of them in turn

### 1.1 Name Voyager

The Name Voyager is an applet, custom developed by Martin Wattenberg, inspired by the Laura Wattenberg's book "The Baby Name wizard". With the specific tool the user has the ability to see a graph with the name popularity, according to his selection on names or even on some characters. The whole app is connected with the given names in the United States, in consonance with the statistical data exported by the Social Security Administration (SSA). The time span is being ranged from the early 80s until the year of 2013.

Referring to its running and the appearance of the visualized material, it is required the Java extension on the browsers. The name voyager could be characterized as a dynamic graph due to its feature of zooming in the initial diagram in accordance with the year and name preferences, until one reaches a single name graph. Furthermore the application implements a heat map which shows geographic name trends in the past four decades, featuring a map display view and a timeline one, under the name of "Name Mapper".

### 1.2 Name Trends

Name trends is also a similar app with the previous one, as it visualizes the baby name tendencies again in the United States, at the same period of time, from the aforesaid database source too. The difference between these two online tools resides in that the second one does not contain dynamic graphs. It is possible for the user to get the appropriate information by mouseover the desired points of the graph, but the idea of getting in a separate way the results of less graphs is user friendlier, thus one could say that it is less comprehensive than the first one explorer. However it has to be admitted that despite the interface and the any diagramming shortcomings, this application has still a compact layout and valuable results are made out.

Having described the two major related works that were found on the Internet, it is being mentioned that none of them offers the option of making comparisons between different countries. In this way the

existence of a web application, which will provide comparative data on the user, not only with regard to the passage of time, but also the country distinction. This is the precise idea - problem that we would like to solve with our project, making a hesitant start.

Particularly important for our project implementation was the incorporation of a slopegraphs set, by the developer Edward Tufte. With the use of this visualized approach we found exactly what we wanted to do. The so called slopegraphs are consisted by two facing columns for the data representation and whenever there is a connection between them, a sloping line indicates it. Moreover the different colour of the lines is in correspondence with the distinct slope angle, showing in grey the more popular names in the left - column country, than this one on the right. Respectively with the blue lines, are figured out, all these names, which are more popular in England, in contrast with the Vienna on the left diagrams side

## 2 APPROACH

Firstly it is needed to be commented that the approach used to visualize the desired data and draw conclusions, about the naming similarities between the two countries under investigation, is web based. Actually it is a simple visualization solution, because that alternative was considered as easier to be accessible to the general public, to the simple end user, than a desktop application.

By continuing our analysis of our visualization tool description, the reason that the main graph was implemented as a slopegraph, has to be explained. Although the use of these particular graphs is not very widespread, it is strongly believed that is one of the perfect ways to illustrate the comparison between the names of the two regions. The common names are shown in a remarkably minimalistic way without adding any other information, just using lines thereby indicating the connection. In this way the user can instantly realize the similarities, without unnecessary redundancy of information. Each connecting line get bolder (Figure 1) when it is being mouseovered and the user can click on it and get more information.



Fig. 1. Bold Line During a Mouseover.

The above mentioned information are shown on the right side of the graph container where the selected name with its rank and count are stated. The countries display remains consistent with the standards of the graph for uniformity, so all the details about the city of Vienna are shown on the left, where for England on the right. Node's - each name's information are shown in a simple and compact way without causing fatigue to the user from unnecessary data.

Apart from the provided statistics in a text form, it was deemed essential to add a diagram showing the changes at the name's trending. For this data representation a line chart was selected, so that the data can be displayed a series of data points in a proportionate way. Blue and red, two colours with high contrast were picked out for the two lines, maintaining a strong and clear differentiation between the two countries. To facilitate the reading of the diagram there are thirteen small circles, each one per year, which are analyzed with passing over the mouse pointer, as the rank and count are displayed in a black translucent small box.

Above the node container, there is a control panel, where the user has the possibility to select the desired configuration for the results, presented in this. For the date selection there is a drop-down list with thirteen records in accordance with the time space from 2000 to 2012. Directly below the list there is a set of radio buttons naming "Girls" and "Boys", for the choosing either the female or the male names. Next to the buttons there is a checkbox by means of which all these names that are no common to both of the countries are removed from the graph, in order to help the end users who only want to focus on the similarities.

Our visualization tool is accomplished with one more data container, which is used for comparing the name rankings between Vienna and England, in a more centralized way and using classification. For this purpose a ranking table has been used, where the most social countries for both of the countries are exhibited one after another, starting displaying from the most famous baby name in Vienna. It should also be noted that this table can be accessed by the user just by clicking the second tab above the main container, in which according to the selected tab it is shown either the slopegraph, or the table, as explained. The tabs options was decided, due to its ease of use and because the fact that the average over online user is accustomed to this interface for changing the information that wants to be displayed.

## 3 IMPLEMENTATION

We implemented our interactive visualization - solution in the Javascript dynamic programming language, which is the most commonly used in web browsers, written using notepad and the browser's debug console. The data retrieval from the excel spreadsheets and therefore their use in the graphs creation, was fulfilled by exploiting the D3 library, which enables the developer to bind arbitrary data to a Document Object Model. In addition to jQuery library and the Bootstrap framework assisted us in our implementation. While our project is a web application it is almost unnecessary to mention that we used the HTML and CSS files. Regarding the hosting of our site, the UNET - webserver (ZID) was used, by us.

## 4 RESULTS

From the user's point of view, two specific scenarios are going to be described, illustrating our visualization tools, for a mother and for a young Sociology student, respectively.

### 4.1 Scenario 1

**Task: Parents look for their child's name**

**Persona: Samantha, 27 years old**

#### Background

Samantha is a young woman with degree on economics and is working as an economist at a small company here in Vienna. She loves reading books and travelling, knowing new civilizations, always being excited to their history. Furthermore one could characterize her as cinema addicted, watching quite many films per week and having many favorite actors around the world.

#### Scenario

She has recently acquired a child and with her husband are looking for a good choice about the given name, as both of them want to find a modern and international one, due to their firm belief that their offspring will be citizen of the world. Being interested in the history of England and having on her mind that English is the first internationally spoken language she would like to check out some English names.

Fig. 2. Default View of the Homepage.



So Samantha having been informed about our site's existence, decides to visit it in order to be inspired. By reaching our website she sees at the home page the graph, which shows the most famous names in Vienna and England and the connections - similarities between them. The first view of the graph (Figure 2) is configured to some by default options and these are the most recent year, this one of 2012 and the female pairs demonstration.

Then Samantha immediately wanted to check out only the names which are popular in both countries and so by selecting the option at the checkbox on the right, next to the radio buttons, filtering in this way only the desired results and looking for more international names. This option is shown in the Figure 3 below.

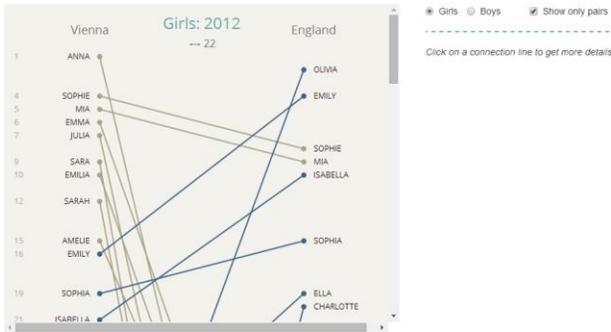


Fig. 3. Pairs Appearance in the Main Graph.

After a short period of time, while seeing the naming similarities, she decided to study the results and get some ideas via the second tab "Table", thinking that a more compact appearance would also be interested. A table with four columns was shown (Figure 4), presenting in a clustered way the rankings of each name from one country to another and at the last column there are the ranking differences, helping to the comparison between them.

Fig. 4. Comparison Table of the Rankings between Vienna & England.

Graph		Table	
<b>Girls: 2000</b> (Compare rankings from Vienna and England)			
Name	Vienna	England	Difference
SARAH	3	34	31
ANNA	4	31	27
SOPHIE	6	7	1
LAURA	9	30	21
VICTORIA	13	52	39
JESSICA	15	5	10
HANNAH	16	9	7
JENNIFER	17	82	65
ISABELLA	21	60	39
NICOLE	24	44	20
ALEXANDRA	28	48	20
NATALIE	33	92	59
REBECCA	37	12	25
MELISSA	46	68	22
EMILY	56	2	54
SOPHIA	70	87	17
EMMA	75	17	58
SAMANTHA	76	58	18
STEPHANIE	84	76	8
ISABEL	87	69	18

## 4.2 Scenario 2

**Task: Sociologist does research on the name globalization**

**Persona: Johannes, 26 years old**

**Background**

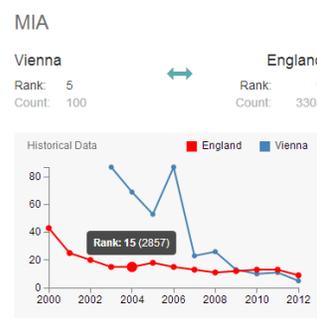
Johannes finished his bachelor studies at the University of Vienna at the department of Sociology and now is attending a master program on Applied Methods and Research Strategies. At this period of time, he is doing a research paper about the globalization and all these particular characteristics that countries may lose or change, with

emphasis on the naming tendencies, due to their connection with the culture of each nation.

### Scenario

So Johannes being informed about our site, where one can easily compare the most popular names in Vienna and in England, faces our homepage and instantly clicks on the connection of his favorite female name Mia. Then on the right side of the main graph, as shown in the Figure 5, a line graph is appeared, with more details indicating the popularity progress of the selected name, in the passage of the decade. Mousehovering each one of the graph circles can also check out the specific ranking information per year and in this way the understanding of the graph becomes easier

Fig. 5. Historical Data Linegraph.



Soon after he is selecting the second option from the two radio buttons on the right side, because he is also interested in learning the naming similarities between the male names.

Fig. 6. Showing Male Names in the Main Graph.



## 4.3 System Performances

Our system's performance could be characterized as quite fast, as the system's response is immediate in real time, without any delay to refresh the data in graphs, in line with the individual user's choice. It has to be noticed that the visualization works normally either on windows or mac devices, regardless of the software version. But crucial for the smooth and normal operation of our application is a good enough internet connection, since it is a web based one and the user has to be online.

## 4.4 Comments on the tool's design and functionality

First of all we should say that during the beta period of our project solution, the feedback, mainly given by our colleagues and friends, was really helpful to identify any errors and omissions. We could say that we received equally positive and negative comments about our project implementation.

First of all we should say that during the beta period of our application, the feedback, mainly given by our colleagues and friends, was really helpful to identify any errors and omissions. We could say that we received equally positive and negative comments about our project implementation.

Another important set of comments is stated on the “Historical Data” graph and in particular about these particular moments to be refreshed and the data to be cleared. What applies to our project so far, is that having selected a name and all the appropriate data are shown on the right, if one would like to change the configuration of the line graph the Historical Data one is going to be there and will not be refreshed, until another name would be clicked on. The idea behind this implementation is simple and aims to help the user to memorizing to compare more easily the results. However a significant amount of users thinks that this aspect of our visualization solution looks like it is a wrong one.

## 5 DISCUSSION

An element of our application that we sometimes consider in a sense, weakness and would like to have solved it, is the “compression failure” of the chart after selecting the option of showing only the pairs. The used term “sometimes” is going to be explained right after.

As it can be realised from the Figure 7 gaps between the common names are created, allowing a dual commentary about it, as from the one hand it could be considered like a positive element and from the other hand it could be regarded as a weakness. The positive aspect is justified by the provided facility to the user to read the chart more clearly, while it is not so dense now.

A strong feature about our project implementation is the fixed size of the diagram, in combination with the scrollbar on the right. We believe that this is a strength for the visualization, by maintaining the uniformity of appearance of the website and not the diversity from one browser to another or from one screen to another.

Another consideration raised by the study of the design and development of our application is that the specific implementation approach, could not be applied to a multitude of names significantly of a greater range. This could not happen because the line graph would be extremely long and this it would be definitely tiring for the user to check it out entirely and finally he would not be able to draw any conclusions.

Completing our project using the Javascript language for the implementation of our visualization tools, what we would like to point, is that it gives the user many possibilities creating remarkable sites, especially in alliance with the D3 library, of course when talking about visualization solutions. Furthermore we would like to give an advice, which is also a personal lesson for everyone, is that whether desktop application are programmed or a web based one, difficulties will always exist but with patience and perseverance can be overcome, while we keep in mind it is not always a time to reinvent the wheel, hence the search is always a good ally.

## REFERENCES

- [1] Tamara Munzner. Visualization Analysis and Design. University of British Columbia: Department of Computer Science, Thursday 24th April, 2014.
- [2] HANSEN, Charles D.; JOHNSON, Chris R. (ed.). *The visualization handbook*. Academic Press, 2005.
- [3] Given Names as Indicator for Measuring Social Change. [http://www.polsoz.fu-berlin.de/en/soziologie/arbeitsbereiche/makrosoziologie/projekte/proj\\_vornamen.html](http://www.polsoz.fu-berlin.de/en/soziologie/arbeitsbereiche/makrosoziologie/projekte/proj_vornamen.html)
- [4] First Names and Globalization. <http://depot.knaw.nl/8747/>
- [5] Summary Slopegraphs. <http://charliepark.org/slopegraphs/>
- [6] JavaScript Library d3.js, <http://d3js.org>
- [7] Bootstrap framework, <http://getbootstrap.com/>
- [8] jQuery library, <http://jquery.com/>