

M4: Final Submission

Nicole Charches
Matr. Nr.: 01506832

Alexander Gelb
Matr. Nr.: 01268620

Benjamin Neckam
Matr. Nr.: 01301917

Axynia Tokareva
Matr. Nr.: 01268620

ABSTRACT

Some blabla here

Index Terms: Visualization—Visualization techniques; GAIA—ESA

1 MOTIVATION

On 19th December 2013 the "European Space Agency", ESA, launched the "Gaia"-satellite to gather more information about our galaxy and create a three dimensional map out of it. Since then data like velocities, positions, errors, parallax and many more were collected of about 2 million stars. All together each star has 50 features in the data set which means the amount of entries is extremely high. But working with such a huge dataset is not easy and most of the time not necessary because scientists only want to focus on a specific area of the galaxy. Therefore they only use subsets of the data. Now the idea was what happens if we take the whole data set and don't look into detail but see the "big picture" and try to find patterns, correlations, outliers or other interesting things.

This could be useful for scientists who are not sure which part of the data they want to explore and for example can focus only on the anomalies or other fascinating sections in the data set.

2 RELATED WORK

What we first did, was searching for similar problems and finding out, what other visualizations looked like in this area. Most of the space visualizations we found online, were in 3d. Examples of them are the ESA Star Mapper¹ or 100000 stars². They both are 3d interactive simulations of the stellar neighbourhood. We also found a 2d-visualization from the ESA, which visualizes the stars from the Gaia dataset³. But we couldn't find anything that has to do with correlations or patterns. All the existing visualizations were about showing the stars themselves with their exact location. So we started prototyping and in the end we even used some of our first basic ideas in the end product of our system. Then we tried to find existing visualization systems. In the VIS course we already learned about Tableau⁴, which allows the user to import Data and create visualization plots based on it. And we also found out about a software named Glue⁵, which helps to create plots and analyze relationships between different data. So we basically tried to create a mixture of both tools and started designing our own system.

Above is a picture, that shows how we imagined the visualization system at the beginning. As we will see later in the end product, our Lo-Fi Prototype looks similar to our visualization system. We kept the sidebar, where the user has basic information about the data and can choose options on how the data will be presented. Also the diagrams and plots appear all on one screen so the user can see how they correlate.



Figure 1: Lo-Fi Prototype of our user interface

3 APPROACH

When we first thought about our visualization design, we tried to think about our user. The system should help to see "the big picture" of a dataset and not only details. So, we thought about plots that could help us to realize that.

What we decided to do, was a Scatterplot Matrix: It really fits into our project, because we can see a whole grid filled with Scatterplots. This way, we can see all Correlations between selected columns of the Dataset. The user has to choose a set of columns, he wants to analyse in the sidebar. The scatterplot has also histograms in the diagonal, so we can see the patterns even better.

Our system can also be used to compute histograms separately. Here the function creates bins for a specific column and shows how many entries are in every bin. While hovering over the bars from our plot, we can also see the exact number of those entries. The user can also set the bin size manually.

Then the system also provides a function to compute a Correlation matrix. For this, the user has also to select a set of columns. For these columns the function computes and displays the correlation coefficient. So it is easy to see how the different columns of a set are connected and if the user needs to analyse it further.

The last part of the system is the PCA. Blabla ber das PCA.

4 IMPLEMENTATION

The whole program was written in JavaScript combined with CSS. Therefore it works on every web browser. For the front-end we decided to use the toolkit "Bootstrap"⁶ which provides different templates so we needn't to spend too much time for the user-interface. To visualize the data the JavaScript library "D3"⁷ was used which offers the ability to let the user interact with the data itself.

¹http://sci.esa.int/star_mapper/

²<http://stars.chromeexperiments.com/>

³<https://gea.esac.esa.int/visualization/index.html>

⁴<https://www.tableau.com/de-de>

⁵<http://www.glueviz.org/en/stable/>

⁶<https://getbootstrap.com/>

⁷<https://d3js.org/>

We encountered many problems which challenged us to implement things properly. One of the problems where to filter the data that all plots can be realized. We figured out that we just can't delete all the entries which contains a NULL or any other non valid values because this would reduce the data set to almost zero entries. Therefore we went through the documentation of the dataset and tried to figure out which features of the stars can be excluded and which not. Unfortunately we can not filter out all unnecessary values at the beginning, we have to do a "pre-filtering" to remove all non integer values and afterwards almost every function has to do their own filtering to adapt the data to their specific task.

5 RESULTS

We started with making a side bar, which shows a few details about the dataset and allows the user to make some specific settings. On top, we can see the Dataset name, rows and columns of our dataset. Firstly, the user can choose which kind of plot he wants to produce in the Options window. "Data Value" is a drop down, in which the user can choose exactly one column, which is what we need for the histogram. If he chooses the Scatterplot, he needs to choose more columns in "Multiple Data". After that, he clicks the "Plot" Button and the chosen diagram will be visible on the display. With the "Remove" Button, he can clear the page.

6 DISCUSSION

6.1 Lessons learned

6.1.1 Benjamin Neckam

There were two things I will take out of this project.

First refers to the implementation and the chosen programming language. Basically I was not very used to JavaScript, to be more precisely I had no experience with it, but I was really impressed by "D3". The opportunities you have to visualize data is breathtaking and allows the programmer to choose over a wide variety of plot types. The only disadvantage that really falls into weight was the performance. If the dataset is bigger than, let's say, 100.000 entries it takes some time to calculate and plot the data and in our case the file contains more than ten times more entries. But all in all it was a very interesting and in my opinion very important for my future to work with JavaScript and "D3" and get deeper into this part of front end programming.

The second part is about the collaboration between the project group of the class and research group of another university. Until now all the projects at the university were from the teachers of the faculty itself but in this case the task was from the "Universitätssternwarte Wien". This means we had to arrange meetings with the project leader and a student which also is a member of the research group and talk about the task in detail and what the aim of this project is. It was fun and interesting to meet all these people but I recognized that even if there are not too many people involved it can easily happen that confusion arises. Like in our case, we started to focus on the wrong details and parts and had to "restart" our project. Therefore I would say communication is the most important part in a project, especially if it is a collaboration of two or more groups, I guess this is the biggest lesson I learned from the project.

7 SEPARATION OF TASKS

8 REFERENCES

- If you receive compilation errors along the lines of "Package ifpdf Error: Name clash, \ifpdf is already defined" then please add a new line "`\let\ifpdf\relax`" right after the "`\documentclass[journal]{vgtc}`" call. Note that your error is due to packages you use that define "`\ifpdf`" which is obsolete (the result is that `\ifpdf` is

defined twice); these packages should be changed to use `ifpdf` package instead.

- The style uses the `hyperref` package, thus turns references into internal links. We thus recommend to make use of the "`\autoref{reference}`" call (instead of "`Figure~\ref{reference}`" or similar) since "`\autoref{reference}`" turns the entire reference into an internal link, not just the number. Examples: ?? and ??.
- The style automatically looks for image files with the correct extension (eps for regular \LaTeX ; pdf, png, and jpg for pdf \LaTeX), in a set of given subfolders (figures/, pictures/, images/). It is thus sufficient to use "`\includegraphics{CypressView}`" (instead of "`\includegraphics{pictures/CypressView.jpg}`").
- For adding hyperlinks and DOIs to the list of references, you can use "`\bibliographystyle{abbrv-doi-hyperref-narrow}`" (instead of "`\bibliographystyle{abbrv}`"). It uses the doi and url fields in a bib \TeX entry and turns the entire reference into a link, giving priority to the doi. The doi can be entered with or without the "`http://dx.doi.org/`" url part. See the examples in the bib \TeX file and the bibliography at the end of this template.

Note 1: occasionally (for some \LaTeX distributions) this hyper-linked bib \TeX style may lead to **compilation errors** ("`pdfendlink ended up in different nesting level ...`") if a reference entry is broken across two pages (due to a bug in `hyperref`). In this case make sure you have the latest version of the `hyperref` package (i. e., update your \LaTeX installation/packages) or, alternatively, revert back to "`\bibliographystyle{abbrv-doi-narrow}`" (at the expense of removing hyperlinks from the bibliography) and try "`\bibliographystyle{abbrv-doi-hyperref-narrow}`" again after some more editing.

Note 2: the "`-narrow`" versions of the bibliography style use the font "PTSansNarrow-TLF" for typesetting the DOIs in a compact way. This font needs to be available on your \LaTeX system. It is part of the "paratype" package, and many distributions (such as MikTeX) have it automatically installed. If you do not have this package yet and want to use a "`-narrow`" bibliography style then use your \LaTeX system's package installer to add it. If this is not possible you can also revert to the respective bibliography styles without the "`-narrow`" in the file name.

DVI-based processes to compile the template apparently cannot handle the different font so, by default, the template file uses the `abbrv-doi` bibliography style but the compiled PDF shows you the effect of the `abbrv-doi-hyperref-narrow` style.

9 BIBLIOGRAPHY INSTRUCTIONS

- Sort all bibliographic entries alphabetically but the last name of the first author. This \LaTeX /bib \TeX template takes care of this sorting automatically.
- Merge multiple references into one; e. g., use [?,?] (not [?][?]). Within each set of multiple references, the references should be sorted in ascending order. This \LaTeX /bib \TeX template takes care of both the merging and the sorting automatically.
- Verify all data obtained from digital libraries, even ACM's DL and IEEE Xplore etc. are sometimes wrong or incomplete.

- Do not trust bibliographic data from other services such as Mendeley.com, Google Scholar, or similar; these are even more likely to be incorrect or incomplete.

You can use the margins for comments while editing the submission, but please remove the margin-par comments for submission.

- Articles in journal—items to include:
 - author names
 - title
 - journal name
 - year
 - volume
 - number
 - month of publication as variable name (i. e., {jan} for January, etc.; month ranges using {jan #{/}# feb} or {jan #-}# feb})
- use journal names in proper style: correct: “IEEE Transactions on Visualization and Computer Graphics”, incorrect: “Visualization and Computer Graphics, IEEE Transactions on”
- Papers in proceedings—items to include:
 - author names
 - title
 - abbreviated proceedings name: e. g., “Proc.\ CONF_ACRONYNM” without the year; example: “Proc.\ CHI”, “Proc.\ 3DUI”, “Proc.\ Eurographics”, “Proc.\ EuroVis”
 - year
 - publisher
 - town with country of publisher (the town can be abbreviated for well-known towns such as New York or Berlin)
- article/paper title convention: refrain from using curly brackets, except for acronyms/proper names/words following dashes/question marks etc.; example:
 - paper “Marching Cubes: A High Resolution 3D Surface Construction Algorithm”
 - should be entered as “{M}arching {C}ubes: A High Resolution {3D} Surface Construction Algorithm” or “{M}arching {C}ubes: A high resolution {3D} surface construction algorithm”
 - will be typeset as “Marching Cubes: A high resolution 3D surface construction algorithm”
- for all entries
 - DOI can be entered in the DOI field as plain DOI number or as DOI url; alternative: a url in the URL field
 - provide full page ranges AA--BB
- when citing references, do not use the reference as a sentence object; e. g., wrong: “In [?] the authors describe ...”, correct: “Lorensen and Cline [?] describe ...”

10 EXAMPLE SECTION

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua [?, ?]. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est.

11 EXPOSITION

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi. Lorem ipsum dolor sit amet, consetetur adipiscin elit, sed diam nonumy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat [?].

$$\sum_{j=1}^z j = \frac{z(z+1)}{2} \quad (1)$$

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

11.1 Lorem ipsum

Lorem ipsum dolor sit amet (see ??), consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum.

11.2 Mezcal Head

Lorem ipsum dolor sit amet (see ??), consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Table 1: VIS/VisWeek accepted/presented papers: 1990–2016.

year	Vis/SciVis	SciVis conf	InfoVis	VAST	VAST conf	TVCG @ VIS	CG&A @ VIS	VIS/VisWeek incl. TVCG/CG&A	VIS/VisWeek w/o TVCG/CG&A
2016	30		37	33	15	23	10	148	115
2015	33	9	38	33	14	17	15	159	127
2014	34		45	33	21	20		153	133
2013	31		38	32		20		121	101
2012	42		44	30		23		139	116
2011	49		44	26		20		139	119
2010	48		35	26				109	109
2009	54		37	26				117	117
2008	50		28	21				99	99
2007	56		27	24				107	107
2006	63		24	26				113	113
2005	88		31					119	119
2004	70		27					97	97
2003	74		29					103	103
2002	78		23					101	101
2001	74		22					96	96
2000	73		20					93	93
1999	69		19					88	88
1998	72		18					90	90
1997	72		16					88	88
1996	65		12					77	77
1995	56		18					74	74
1994	53							53	53
1993	55							55	55
1992	53							53	53
1991	50							50	50
1990	53							53	53
sum	1545	9	632	310	50	123	25	2694	2546

11.2.1 Duis Autem

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

11.2.2 Ejector Seat Reservation

Duis autem [?]⁸ vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat,⁹ vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.

⁸The algorithm behind Marching Cubes [?] had already been described by Wyvill et al. [?] a year earlier.

⁹Footnotes appear at the bottom of the column.

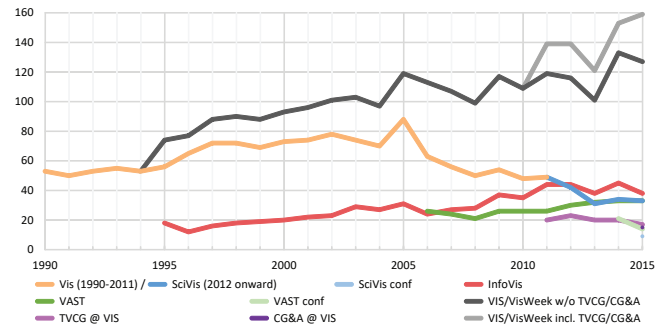


Figure 2: A visualization of the 1990–2015 data from [?]. The image is from [?] and is in the public domain.

Confirmed Ejector Seat Reservation Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat [?]. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

Rejected Ejector Seat Reservation Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie

12 CONCLUSION

Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet. Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum.

ACKNOWLEDGMENTS

The authors wish to thank A, B, and C. This work was supported in part by a grant from XYZ.