

Data Memes for Personal Visualization

Darius Coelho, Ayush Kumar, and Klaus Mueller

Visual Analytics and Imaging Lab, Computer Science Department, Stony Brook University, NY, USA and SUNY Korea, Songdo, Korea

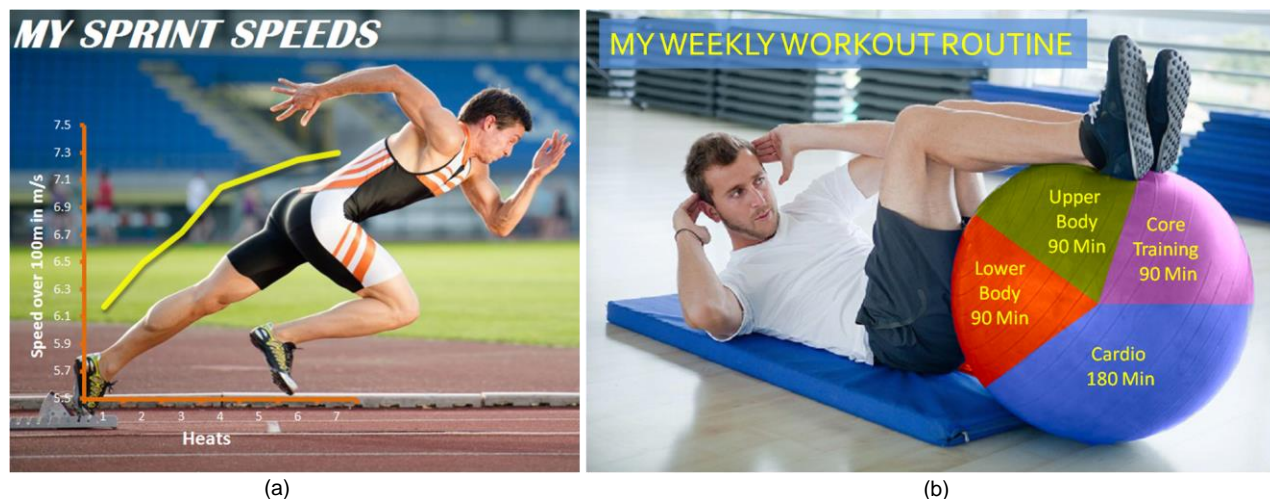


Figure 1. Personal visualizations (a) shows Joe's speed over multiple heats that he ran (b) shows Mike's workout routine for the week

ABSTRACT

We introduce the concept of Data Memes as artistic visuals of data in which users can merge data visualizations with an image such that the structure of the image supports the user's intended meaning (or interpretation) of the data. Since Data Memes can represent very personal views, it is natural to employ them in the visualization of personal data. Hence, in this paper we focus on the use of Data Memes for this purpose and also explain how they can be constructed with relative ease. Data Memes as a general concept are meant to engage users with the visualization and persuade a viewer to look at the data. Communicating personal achievements is often the wish of posters in social networks, and although we have not tested this yet, we believe that Data Memes augmented with pertinent personal data have good potential to achieve higher levels of attention in these circles, and elsewhere.

Keywords: Chart junk, infographics, personal visualization.

1 INTRODUCTION

The information age is upon us and in recent years millions of terabytes of data have been generated every day. A growing chunk of this data is related to our personal lives – information about ourselves, our communities, and issues that have a personal relevance [9]. With the access to such data, people want to make sense of it, and to do so they turn to visualization. While there are a multitude of solutions available to visualize these data, they are rarely used in a personal context.

Recently researchers have been exploring the area of *Personal Visualization* and *Personal Visual Analytics*. Personal visualization caters to the wide variety of requirements of people

who are not necessarily visualization experts or analysts. These people bring personal interpretations to their data and wish to design visualizations that can represent these personal views [20]. In our paper we describe a design framework that allows users to express their opinion about the data by framing them as charts into a self-selected and self-designed visual context. This has good potential for fostering their own engagement with the data as well as that of others with whom these design products are shared.

Fig. 1 shows some examples of the personal visualizations of personal data our system can assist users in creating. Here we note that our system can also be used for other types of data, but the opportunity for users to frame these data into a personalized visual context always remains and is a hallmark of our system [5].

We have coined a new term for these types of visualizations – *Data Meme*. Wikipedia defines the notion of Meme "an idea, behavior, or style that spreads from person to person within a culture". A more recent phenomenon is that of Internet Meme. These are images augmented with text in which the creator's choice of image puts forward his/her idea about the message the text conveys. They are often reused with the same core message, and as a result gain viral dissemination. Similarly, our Data Memes are images augmented with related data in which the designer's choice of image puts forward his/her idea about the message the data convey.

Our Data Memes appeal to the masses by embracing mainly basic charts, specifically pie, line, and bar charts. Following the definition put forward by Pousman et al. [15], they fall into the set casual information visualization techniques catering to a wide variety of users. The chosen and further enhanced image that becomes part of the visualization is itself a powerful element of the visualization as it can encode a multitude of information. The image is linked to the topic of the data being displayed, with the choice of image putting forth the designer's point of view about the data. At the same time it is also capable of enhancing memorability and engaging viewers. We foresee that users may use the Data Memes to post their own data on social media or blogs in order to (1) engage viewers to consider the information or (2) put across their own point of view about the data.

Emails: {dcoelho, ayush.kumar, mueller}@sunykorea.ac.kr

In the following, Section 2 reports on related work. Section 3 describes our system by ways of several use examples. Section 4 presents a few examples, and Section 5 ends with conclusions.

2 RELATED WORK

2.1 Visual Embellishments in Charts

Our Data Memes are significantly related to *chart junk* which are visual embellishments of charts. The notion of chart junk has received growing attention in the field of information visualization. Visualization guru Edward Tufte strongly discourages the use of visual embellishments in charts. He opines that visualizations that contain chart junk are distracting to users and make the chart difficult to read [18].

On the other hand, famous designer Nigel Holmes [13] believes that chart junk will pique a viewer's interest in the chart. The study by Bateman et al. [1] concluded that while chart junk and plain graphs were equally comprehended by users, chart junk was more memorable. Stephen Few [7] responded that the study was flawed in some ways, but then ended by stating that properly designed chart junk can be useful provided it does not lead to distractions or misinterpretations. Borkin et al. [3] investigated which aspects of a visualization make it memorable. Other authors [2][11] have made similar basic observations. Harrison et al. [8] studied the first impression that infographics have on people, concluding that colorful infographics with not much text and image areas appeal to most people. Viegas et al. [19] observed that chart junk visualizations created by artists, while memorable, can change the consumer's attitude or change the way people think.

Chart junk so far has been mainly created by artists like Peter Ørntoft [14]. He created simple but beautiful infographics by dressing a model with the chart or designing it physically and then photographing it. Nuage Vert by Evans et al. [6] highlights segments of smoke in an image to show power consumption.

All of these artist-generated chart junk creations are commonly very attractive and inspirational, but it remains difficult for main stream users to create good chart junk with existing digital tools [23]. Also the usability of a visualization is dependent on the aesthetics [4] and the style used in the design

[13]. This inspired us to create a novel design tool we call *MemViz* which allows non-artists to design charts with visual embellishments. MemViz morphs charts such that they fit well in areas of an image. For the work that is subject of this paper we use an enhanced version of our tool, appropriately called *MemeViz* that enables users to create personal visualizations of personal data.

2.2 Personal Visualization

Researchers have been investigating the use of visualizations for personal and casual use. Some of the work explores how both experts and non-experts can make personal data engaging and gain insights into personal data. Leung et al. [10] projected a visualization of a user's social network data and found that it triggered social interactions when viewed by others and acted as a mode of self expression. Xiong et al. [22] also visualize social network data with a flower metaphor in a manner that is easy for non-experts to comprehend. Wood [21] tries to express his personal experiences in the visualization of his cycling data. Rodgers et al. [16] observe that users were able to easily understand their energy consumption data through abstract visualizations with less detail as compared to traditional techniques. Wang et al. [20] studied the impact of design for personal visualization and observed that the abstract techniques were better at engaging the user but at the cost of determining concrete data.

3 DESIGNING THE DATA MEMES

Our experiments have shown that even novice users can employ MemeViz to create eye-catching visualizations that embed basic charts into an image. All can be achieved by ways of a set of dedicated algorithms and interactions which can be executed via simple mouse clicks. The user interface is shown in Fig 2. A user starts off by providing the data file and selecting the type of chart (pie, bar, line) he or she wishes to plot. Next, the tool suggests a set of relevant thematic image backdrops retrieved by a web-scale text-based image search engine (e.g., Google Images, Flickr). To form the search query our tool uses the title of the dataset and the attribute labels since these are intuitive descriptors of the data. Alternatively, the user can also call in a picture from his or her

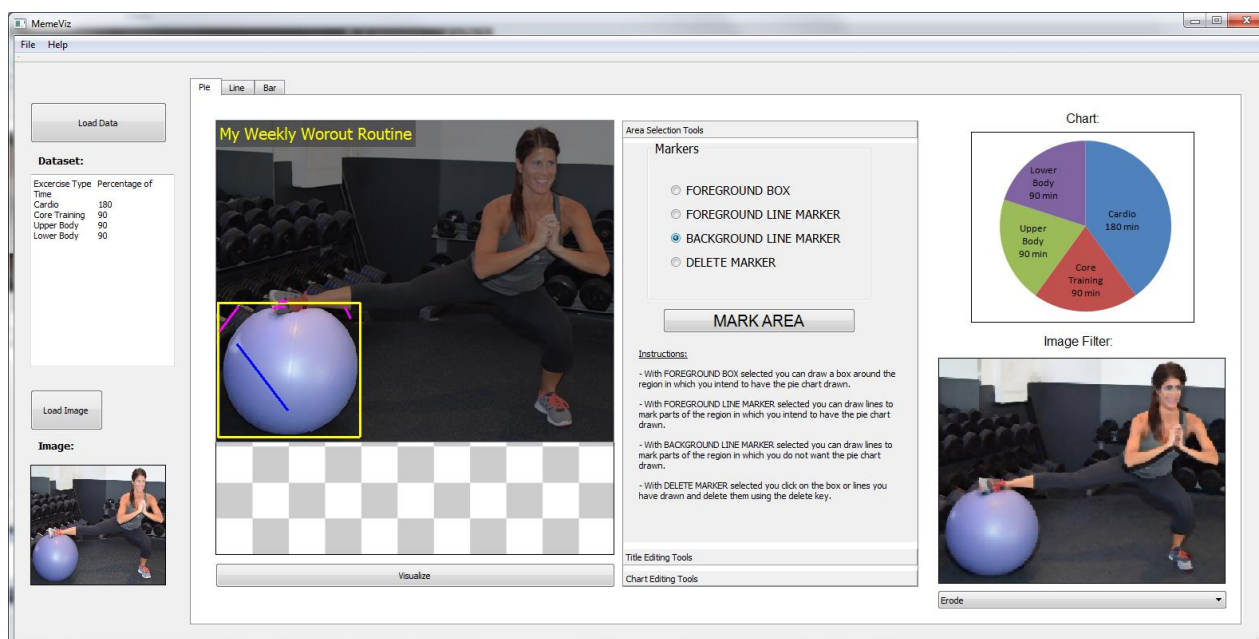


Figure 2. The MemeViz user interface which was used to create all of the Data Memes shown in this paper

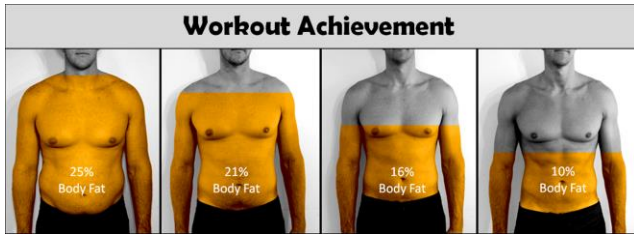


Figure 3. Marcus' body fat levels at different stages of his life.

private photo collection. The latter would likely be a more popular option for the purpose of personal visualization where users wish to post their own data to social media or blog, utilizing a personal image as a carrier.

Once the user has selected an image from the provided set, he or she is given access to an array of processes (described in the following subsections) to fuse the chart with the image. Then, after the chart has been embedded, the user can customize the visualization further by applying a diverse set of visual effects to the backdrop (filters, transparency etc.) or to the chart (color selection, glow, transparency). This customization gives the visualization an illustrative look and feel, and it also reduces image detail that might interfere with the detail of the chart. This finalizes the result.

3.1 Pie Charts

Pie charts and their spatial generalizations use a division of unity paradigm to show numerical proportions. To embed pie charts into images we decided to have users select a region (in most cases an object) in the image after which the tool would divide the region proportionately according to the data.

With this tool, after selecting the backdrop image, the user interactively selects a region into which to plot the chart by marking a bounding box around the region. The tool segments the region using the Grabcut algorithm [17]. The segmentation can be refined by marking highlights on the background and the foreground. The next step involves laying out the chart into the segmented region. Here, users have various options to divide and layout the chart. For pie-style charts they can divide the region radially in two ways: (1) angle-based, i.e. dividing 360° radially proportional to the data, and (2) area-based, i.e. dividing the area of the region radially proportional to the data. In the latter case we use a pixel count to determine the area. For fill-style charts we allow the user to specify an axis across the segmented region and then fill the region along this axis. Again, we divide the region in two ways: (1) based on the length of the axis, and (2) based on the area of each division within the segment for which we use a pixel count as well. Once the region is divided we color the divisions of the region in CIE LAB color space. This retains the texture of the region and only changes the color.

3.2 Line Charts

The line chart is one of the most common visualizations to plot a series of data points. To merge line charts with images, we match the line to structures in the image that have a similar trend to the line. In this case instead of looking for images with closed regions as in the fill chart task, we are now looking for images with features and edges that mimic the trend of the line chart as close as possible.

As the first step the user is asked to place the line chart over a desired region in the backdrop image. We experimented with a number of methods to perform line matching so as to place the

chart automatically in the image, but a solution remains elusive. Instead we now rely on the user to find the features visually by looking at the chart and the image side by side. Although this is not ideal we observed that users were capable of performing this task well. We have provided an interactive interface that allows users to move and scale line charts and so fit them into appropriate positions in the image. However, we specifically prevented the user from rotating the chart as it may lead to line chart trends looking more dramatic.

3.3 Bar Charts

Bar charts are mainly used to show comparison of values. They are also a very popular visualization paradigm. In our application, the bar charts can be merged with images by either overlaying the chart over the image or by treating objects in the image as bars. In this case we make use of techniques used to embed pie charts and overlay line charts.

Once the backdrop has been selected, the user is asked to decide if he/she wants to place the bar chart over the image or fill in objects. In the former case the system applies the same technique that was also used for the line charts where the user can place the chart using the mouse. On the other hand, if the user wishes to fill objects then we follow the segmentation technique used for the pie charts, but in this case for every bar the user must mark a region in the image.

4 EXAMPLES

We shall now give a few examples on what types of personal visualizations might be possible using our tool. All images were obtained using Google image search using the respective activity terms in the search string.

Fig. 1a shows a Data Meme featuring Jim, a short distance runner, who visualizes his progress in sprinting speed via a line plot embedded in a picture that shows him just after setting off. The picture was chosen since the posture at that time aligns well with the line plot. This creates a good synergy between image content and the data which is a prerequisite for a good Data Meme.

Fig. 1b shows Mike during an exercise with his workout routine allocation mapped as a pie chart onto a medicine ball. On the other hand, Fig. 3 shows Marcus at different stages of his life with the corresponding level of body fat mapped like a bar chart. The sequence of pictures was acquired during a long term "Get Lean in 180 Days" program that Marcus underwent. He uses the Data Meme to proudly showcase his amazing progress.

David, who plays an online game League of Legends, created the Data Meme of Fig. 4a with his game stats and an image of a character he plays in the game. In Fig 4b, c, d we see the Data Meme's of Maria, Betty and Anna, created from their fitness data. In Fig. 4e, we see Patrick's Data Meme visualizing the Netflix data of his most watched TV shows. Finally, Fig 4f shows Ryan's Data Meme that visualizes his dataset about gym accidents.

5 CONCLUSION AND FUTURE WORK

We believe that the examples we have presented do a good job in explaining the potential of Data Memes as a means to contextualize the data with an image and so provide for a more engaging viewing experience. In the case of a personal visualization this image context will often involve the user as an actor, or a well-chosen substitute or avatar. Most of the Data Memes we have created here were meant to increase viewer engagement for data usually posted just as plots or numbers on social media or blogs. In some cases, however, they can also be used to put forth a personal point of view about the data.



Figure 4. Personal visualizations created by our designers: (a) shows David's League of Legends stats (b) shows the amount of time Maria ran everyday for a week and (c) shows Betty's bike tension levels during her spinning routine (c) shows Anna's workout routine (e) shows the top 3 most watched television shows Patrick watches on Netflix and (f) show a stat about gym accidents created by Ryan.

Future work will run a user study where subjects would create Data Memes with real and even personal datasets using pictures of themselves. We will also test how engaging these visualizations are and if the user's point of view is put across in the Data Meme.

6 ACKNOWLEDGEMENT

This research was partially supported by NSF grants IIS 1117132 and 1527200 and the MSIP, Korea, under the "ICT Consilience Creative Program" (IITP-2015-R0346-15-1007).

REFERENCES

- [1] Bateman, S., Mandryk, R., Gutwin, C., Genest, A., McDine, D., Brooks, C. "Useful junk? the effects of visual embellishment on comprehension and memorability of charts." *CHI 2010*, pp. 2573-82.
- [2] Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P. W., Reppa, I., Floridi, L., Chen, M. "An empirical study on using visual embellishments in visualization." *IEEE TVCG*, 18(12), 2012, pp. 2759-2768.
- [3] Borkin, M., Vo, A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. "What makes a visualization memorable?." *IEEE TVCG*, 19(12), 2013, pp. 2306-2315.
- [4] Cawthon, N., & Moore, A. "The effect of aesthetic on the usability of data visualization." *IV'07 Int'l Conf. IEEE*, 2007, pp. 637-648.
- [5] D. Coelho, S. Ha, S. Cheng, S. Mahmood, J. Kim, K. Mueller, "MemViz: A Tool for Creating Memorable Visualizations," *IEEE VIS Poster Abstracts*, Paris, France, November 2014
- [6] Evans, H., Hansen, H., & Hagedorn, J. "Artful media: Nuage vert." *MultiMedia*, IEEE, 16(3), 2009, pp. 13-15.
- [7] S. Few. "The chartjunk debate. a close examination of recent findings", *Perceptual Edge Visual Business Intelligence Newsletter*, April, May, and June 2011
- [8] Harrison, L., Reinecke, K., & Chang, R. "Infographic Aesthetics: Designing for the First Impression." *CHI*, 2015, pp. 1187-1190.
- [9] Huang, D., Tory, M., Aseniero, B., Bartram, L., Bateman, S., Carpendale, S., Tang, A., Woodbury, R. "Personal visualization and personal visual analytics." *IEEE TVCG*, 21(3), 2015, pp. 420-433.
- [10] Leung, M., Tomitsch, M., & Vande Moere, A. "Designing a personal visualization projection of online social identity." *CHI'11 Extended Abstracts*, 2011, pp. 1843-1848.
- [11] Li, H., & Moacdieh, N.. "Is "chart junk" useful? An extended examination of visual embellishment." *Proc. Human Factors and Ergonomics Society Annual Meeting*, Vol. 58, No. 1, (2014).
- [12] Moore, A. V., Tomitsch, M., Wimmer, C., Christoph, B., & Grechenig, T. "Evaluating the effect of style in information visualization." *IEEE TVCG*, 18(12), 2012, pp. 2739-2748
- [13] Nigel Holmes, <http://nigelholmes.com/>
- [14] Peter Ørntoft, <https://www.behance.net/peterorntoft>
- [15] Pousman, Z., Stasko, J. T., & Mateas, M. "Casual information visualization: Depictions of data in everyday life." *IEEE TVCG*, 13(6) (2007): 1145-1152.
- [16] J. Rodgers and L. Bartram, "Exploring Ambient and Artistic Visualization for Residential Energy Use Feedback," *IEEE TVCG*, 17(12), 2011, pp. 2489-2497.
- [17] Rother, C., Kolmogorov, V., & Blake, A. "Grabcut: Interactive foreground extraction using iterated graph cuts." *ACM Transactions on Graphics* 23(3), 2004, pp. 309-314.
- [18] E. Tufte. *The Visual Display of Quantitative Information*. Cheshire (Conn.), 2001.
- [19] Viégas, F., Wattenberg, M. "Artistic data visualization: Beyond visual analytics." *Online Communities and Social Computing*, 2007, pp. 182-191.
- [20] Wang, S., Tanahashi, Y., Leaf, N., & Ma, K. L. "Design and Effects of Personal Visualizations." *CG&A*, IEEE 35(4), 2015, pp. 82-93.
- [21] Wood, J. "Visualizing Personal Progress in Participatory Sports Cycling Events." *CG&A*, IEEE, 35(4), 2015, pp.73-81.
- [22] R. Xiong and J. Donath, "PeopleGarden: Creating Data Portraits for Users," *Proc. 12th Ann. ACM Symp. User Interface Software and Technology*, 1999, pp. 37-44.
- [23] Su, Y. S. "It's easy to produce chartjunk using Microsoft® Excel 2007 but hard to make good graphs." *Computational Statistics & Data Analysis* 52(10), 2008, pp. 4594-4601.