Remote Instruction for Data Visualization Design—A Report From the Trenches

Jan Aerts, Jannes Peeters, and Jelmer Bot, Hasselt University, 3500, Hasselt, Belgium
Danai Kafetzaki, Hasselt University, 3500, Hasselt, Belgium and also KU Leuven, 3000, Leuven, Belgium
Houda Lamqaddam, KU Leuven, 3000, Leuven, Belgium

In this article, we report on our experiences of running visual design workshops within the context of a master’s level data visualization course, in a remote setting. These workshops aim to teach students to explore visual design space for data by creating and discussing hand-drawn sketches. We describe the technical setup employed, the different parts of the workshop, how the actual sessions were run, and to what extent the remote version can substitute for in-person sessions. In general, the visual designs created by the students as well as the feedback provided by them indicate that the setup described here can be a feasible replacement for in-person visual design workshops.

With the increasing ease and lowering cost of generating data, data-first research has become a significant complement to hypothesis-driven research. Visualization is gaining higher importance in data science and statistics to make sense of these data, and helps an analyst or scientist reason about their data.¹

Data visualizations can consist of generic plots such as scatterplots and barcharts, but more novel visual designs might be necessary to allow reasoning on complex data. Unfortunately, creating innovative, appropriate visual designs is not a trivial task, especially for people who are not familiar with creative visual work, such as students in statistics and data science. For this reason, we want to look at methods that are meant to assist students in the task of design space exploration and ideation for visual design, specifically in an online setting.

Studies have shown the usefulness of, for example, sketching,² brainstorming,³ and collaborative thinking⁴ in the design process. Teaching these skills can be accomplished by running interactive workshops where students are encouraged to be creative and collaborate.

Under some circumstances, however, it might not be possible to run such sessions in person, and they need to be moved to an online setting (e.g., due to social distancing and traveling restrictions as during the COVID-19 pandemic).

In this article, we report on our experiences running online visual design workshops in the context of data visualization, as part of a master’s program in statistics and data science at two universities, using a synchronous, noncollocated setup.⁵ In particular, we will go over the setup and aims of the course as a whole, the role of visual design in it and how it is approached, the technical setup used for teaching visual design, how a session was run, and lessons learned in the process.

VISUAL DESIGN IN STATISTICS AND DATA SCIENCE EDUCATION

The visual design sessions described in this article play an essential role in different data visualization courses and workshops organized by or contributed to by the authors. More specifically, they are an integral part of the data visualization in data sciences courses of the Masters in Statistics and Data Science at Hasselt University and KU Leuven in Belgium, and of the yearly one-week Biological Data Visualization workshop organized together with the European Bioinformatics Institute in Hinxton, U.K. In addition, they are taught as part of the two-week yearly Programming for Evolutionary Biology course (described in Faria et al.’s work⁶) held in Berlin, Germany.

¹ This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/
Digital Object Identifier 10.1109/MCG.2021.3116042
The master’s students typically have a wide range of backgrounds (psychology, sociology, biology, mathematics, economics, etc.). The majority has coding experience in R and SAS, and has followed courses on multivariate statistics. The visualization courses revolve around multistudent projects where the students select a dataset, define open research questions, design visuals and interaction, and implement these. Research questions can change over time and be reformulated based on the insights the students get from the data.

There are many possible approaches for teaching data visualization to students with a statistics or data science background (see, e.g., Lo et al.’s work), ranging from explaining how to use specific data plotting software and libraries (such as Tableau and ggplot2), to more creative ones where new and novel visual designs are explored. We lean toward the latter, as we want to emphasize the creative aspect in data exploration and searching for interesting questions to ask. We specifically aim for the students to explore design space as broadly as possible to develop novel visual designs.

Design Fixation
One of the challenges in the creation of such novel visual designs, however, is that many students tend to very quickly translate problem statements into statistical tests, thereby bypassing the exploration phase, and suffer from design fixation, or a “blind adherence to a set of ideas or concepts limiting the output of conceptual design.” Design fixation appears in many different fields of innovation, including industrial design, architecture, software, and interaction design. In teaching data visualization, our experience is that the issue for these students is mainly an “unconscious adherence to the influence of prior designs”: the design space that someone explores to find a solution is by definition a subset of the design space that that person knows.

Research on the creative process is quite extensive, but little research exists that considers data visualization’s specific requirements. Still, some aspects can be found in literature that have an impact on the creative process, which can improve the output of data visualization designs. Some examples are: providing an intrinsic or extrinsic incentive, providing helping examples, or allowing for an incubation period. In addition, the inherent “creative skills” of an individual, the collective or individual dimensions of the process, and the knowledge of a coming evaluation also play a role.

GENERAL COURSE SETUP
The course consists of several parts. First, introductory lectures cover a rationale on the importance of visualization in data science, the history of data visualization, perception, and cognition, examples of visualization of different data types, models for data visualization (most notably Munzner’s work), and the design process. Second, students are guided through a visual design exercise (the topic of the sessions described in this article). Third, they are introduced to web technologies to implement visualizations using HTML, CSS, and JavaScript. Finally, they use the knowledge and skills gained in a group project.

VISUAL DESIGN SESSION SETUP
In the guided visual design sessions, the students get acquainted with the process using a common fictional dataset about COVID-19 infections. The aim of this session is for students to get acquainted with developing novel designs. Inspiration for the visual design session mainly draws from the Double Diamond model (British Design Council; http://www.designcouncil.org.uk), game storming methods, the work by Bill Buxton, and the five-design sheet methodology.

The visual design sessions consist of several parts. First, the students are introduced to the general concepts and the idea of diverge–emerge–converge (or open–explore–close). This is followed by warm-up exercises, and two iterations of diverge–emerge–converge.

Warm-Up Exercises
Warm-up exercises have proven to be very important in our courses to get the students accustomed to the tools they will use in the design session, i.e., pen and paper. Exercises consist of a) “taking a line for a walk” where students draw a single line while a poem is recited by the lecturer; b) drawing eight concepts of increasing complexity in three seconds each, ranging from “table” and “chair,” to “Wednesday” and “existentialism;” and c) drawing two data points in eight different ways.

Diverge
The primary purpose of the divergent phase is to create the raw material that can be combined, modified, or discarded later. Students are given 30 minutes to individually create as broad a range of sketches as possible for a given dataset. Sketches are drawn on A7-sized paper, marked with the student’s initials and a unique identifier, and a clear indication of what the sketch represents. Sketches are meant to convey the idea, not be polished drawings. Therefore, a sketch should only take seconds to create.

Different methods are used in concert to help in ideation. First, we provide an inspiration board with
more than 50 creative data sketches, including work by Duncan Geere (blog.duncangeere.com/tag/sketchbook-pages/) and the Dear Data project (www.deardata.com). Second, students are asked to think about reworking existing visuals (e.g., “how can I create a different version of a line chart”), to glue dissimilar ideas together (e.g., using a node-link diagram to draw geospatial data), or to reverse an idea (e.g., representing data points as links and relationships as nodes in a node-link diagram, as in Longabaugh’s work). Students are specifically asked to suspend critical thinking at this stage, which would lead to the rejection of sketch ideas, as even “bad” sketches might be useful during the later stages, or at least help to map out design space.

As some students are overwhelmed with the range of possibilities for sketching and get stuck as a result of overchoice, we also present them with a way to structure their sketches using two conceptual matrices (see Figure 1). In one matrix, the columns represent the different dimensions in the dataset, whereas the rows represent the number of observations depicted in one visual (i.e., “one,” “a selection,” or “all”). The student can randomly pick a cell in this matrix (e.g., for a dataset on COVID-19, drawing the number of reported infections on a single day, all Mondays, or all days of the last month) and create a sketch for that particular combination. In addition, a second matrix lists the different dimensions both as columns and as rows. Using this matrix as a guide, the student can, for example, focus on drawing infections by city, or tests versus infections. It is made clear that the rows are the base of the sketch, so that, for example, number of tests by virus strain is not the same as virus strain by number of tests.

Emerge
For the emerge stage, students are asked to discuss the generated sketches in small groups. They perform affinity mapping to identify which sketches have a similar intent, followed by identifying which sketches could act as detailed views for other sketches (e.g., one sketch showing the total number of COVID-19 hospitalizations over a whole year, whereas another presenting a breakdown of the hospitalizations per city), and the SCAMPER framework: Substitute, Combine, Adapt, Modify, Put to another use, Eliminate, Reverse (see Figure 2). Also, they are asked to consider which external data would be interesting to collect and combine with the data they already have.

This collaborative emergent phase is scheduled for approximately 30 minutes.

Converge
Finally, in the convergent stage, students are asked to identify interesting ideas to work out further. To support them, we suggest looking at which are new, useful, and feasible (NUF framework), and performing value mapping in which each student receives several voting tokens that they can put on those ideas that they find most interesting. We emphasize that the technical implementation later in the course should not be a constraint in the first iteration, and that the design section of their final project is graded separately from the implementations. In other words, the importance of the feasibility aspect in the NUF framework is downplayed, to ensure that the students do not limit their visual design based on limited programming experience.

The output of the converge exercise of the first iteration acts as the starting point for a second iteration. Students can optionally perform more iterations.

Differences in In-Person Versus Online Design Teaching
Running the visual design sessions online instead of in-person presented specific hurdles to overcome. As this was the first time the sessions were online, we decided to hold them in real-time instead of

FIGURE 1. Matrix structures to guide students who have difficulty thinking of what to sketch.
asynchronously (i.e., where each student could perform the exercises in their own time). First, connectivity issues might arise. Students are spread over many countries with varying quality of Internet connection. Second, students might be in different time zones, which means that the session could be either early or late in the day. In our case, we had students join from Yangzhou, China, to New York, USA, and Bogota, Colombia. Third, students need to be able to share their sketches. In an in-person session, sharing sketches means just putting the pieces of paper on the table. In an online setting, sketches need to be photographed, maybe uploaded to cloud services, imported into the collaboration tool, and resized if necessary. Finally, it is essential for the teaching team to identify those students that are struggling in generating ideas. In an in-person setting, this is trivial as the team can see which students do not have many sketches. In contrast, in the online version, this can only be assessed after the students have uploaded their sketches.

**TECHNICAL SETUP**

To successfully run a remote interactive visual design session, we set the following requirements of the software to be used. Students need to be able to a) communicate in separate groups, b) maintain “connection” to the rest of the groups and the teaching team, c) make quick sketches and share these with their peers, d) discuss and organize sketches that they created, and e) indicate that they need support. Teaching assistants and lecturer need to be able to a) interact with the groups separately, b) address all students simultaneously, and c) keep an overview of how students are doing during the exercises. In addition, the system should allow for self-organization, in that it provides a clear overview of the situation and allows for dynamic adaptation.

Based on these requirements, we opted for a combination of pen and paper, GatherTown (gather.town) for general communication, and Miro (miro.com) for collaboration. Note that our use of these software platforms should not be seen as an endorsement, and the setup that we used can, to different extents, be duplicated using other software solutions.

**SESSION PREPARATION**

In the following, we describe how the different tools were set up before the start of the session.

**Communication Platform**

A workshop room was established in the communication platform (see Figure 3). This room provides proximity chat (i.e., people can hear each other when they are located close to each other), but also featured a single “spotlight square” (i.e., standing in this position makes everyone in the room hear you, see you, and optionally view your screen). This spotlight square was used for giving general instructions and announcements to the students. In addition, different “private spaces” were set up, consisting of a table and six chairs. Everyone can hear each other in a private space, but cannot be heard by someone outside of the space (i.e., not sitting on a chair at the same table) and vice versa. Each table is numbered; this number corresponds to a working space in the collaboration platform (see the following section).
Collaboration Platform
Because of the large number of students in the course (i.e., more than 90), three parallel collaboration boards were prepared, each managed by one teaching assistant (see Figure 4). Within each board, different sections were created per group of four students. Each section was numbered corresponding to a table in the communication platform. Different areas were created, which would guide the consecutive parts (warm-up, and diverge–emerge–converge) of the visual design exercise.

Instructions to Students
It was crucial that students have the necessary material on hand at the start of the session. This material consisted of a pen, paper, and a means of uploading photos. It was made clear that the pen used should allow for clear marks even after uploading sketches onto the collaboration platform.

To make students acquainted with the communication and collaboration platforms, they were invited prior to the design session. Specifically, the communication platform was used for lectures preceding the design session. In addition, students were previously invited to a “Getting to know” board of the collaboration platform where they had to perform some simple exercises, such as adding a sticky note, drawing a circle, placing a comment, and uploading a picture.

RESULTS
At the start of the sessions, students gathered in the communication platform and took a seat at one of the tables. It was important that at least one chair remained empty to allow a teaching assistant or the lecturer to join the group. Teaching assistants and the lecturer could use the spotlight to address the whole student group. In parallel, the students would access the collaboration tool. With this combination, they could work together while still converse with the others at the table.

Warm-Up Exercises
The warm-up exercises acted as a way to get the students familiar with their working material, to put them at ease for a type of workshop that they were not used to, and to test if taking pictures and uploading worked well; all in a fun setting. This part of the session also helped the teaching team to calibrate the time needed for the organizational and technical part of asking questions, uploading images, and organizing sketches on the platform.
Diverge
In the divergent phase, students worked individually and tried to create a wide variety of sketches based on the description of a fictitious COVID-19 dataset, including five example data records listing, among other things, number of performed tests, hospitalizations, etc., per municipality per day. Initially, 20 minutes were set aside for this exercise but this was extended to 30 after student feedback. A clock shown at the bottom of the collaboration platform screen showed the students how much time they had left.

As it was important to gauge how students were doing, teaching assistants and lecturer walked from table to table in the communication platform and asked the students to hold up their sketches in front of the camera. If a student struggled with generating ideas, the teaching team would direct them to the auxiliary material provided at the top of the collaboration board, and provided suggestions. Although students were working individually, they had the opportunity to talk to their peers at the same table as well as a teaching team member when the latter joined the table. Students could also raise their hands at any time to call over a teaching team member.

Intermittently during the divergent phase, students were asked to upload their sketches onto the collaboration platform, in order to a) assess whether they were stuck or not, and b) keep a connection with the students.

Similar to in-person sessions, the breadth of designs generated by the students differed significantly, where some students created a limited number of sketches that resembled standard plots (e.g., barchart, line chart) whereas others created graphics that were more inventive (see Figure 5).

Emerge
During the emerge phase, students switched to collaborative work rather than sketching individually. The teaching assistants and lecturer joined the different tables to provide guidance if the groups had difficulty in organizing their drawings. Students were allowed to include sketches from other groups if these made interesting combinations with their own work.

We found that the initial level of interaction within the different groups varied significantly. The level of conversation at a table could be assessed by the appearance of speech bubbles in the communication platform (see Figure 3). Whereas some groups had lively debates, in other ones, the participants were hesitant to actively contribute to the discussion. The latter was exacerbated by the fact that the students had not met each other in person yet due to the COVID-19 pandemic. Nevertheless, after an initial period, all groups were able to discuss their content.

Figure 6 shows a screenshot of students collaborating in organizing and combining sketches.
Converge
In the converge phase, the students discussed which ideas to develop further. A clear difference was seen between groups who approached this methodically with voting versus groups who selected these ideas under consensus. Some groups connected the different ideas into stories and interaction flows. For most groups, the convergent phase flowed organically from the emergent one, without a clear boundary between them.

Quantitative Aspects
As the workshops had to be changed at short notice from an in-person setting to online, all efforts were directed toward making it work and no specific attention was given to measuring the output quantitatively. Qualitative evaluation was performed using a survey after the fact (see the “Feedback Survey” section). Features that can be useful in quantitatively appraising the used approach are the number of sketches created by each student, the diversity of those sketches, the number and types of clusters created during the emergent phase, etc. The authors did, however, find difficulties in clearly measuring these. First, even after instruction, some students had not marked each sketch with their initials and a unique number, which meant that in some cases sketches would have to be matched on style to assign them to a particular student. Second, students were allowed to incorporate sketches from the other groups during the emerge phase. Third, even though the students discussed how the sketches could be combined and/or clustered in the emerge phase, they were not instructed to add a label or description for post hoc analysis. Fourth, no method was found to describe to what extent a student had explored design space. One student could, for example, create 20 sketches that are all basically different versions of the same thing (e.g., barchart), whereas another student could create five sketches that each are completely novel. Even though strict instructions could have been given regarding how clusters should be formed, how they should be described, etc., the teaching team found it much more important to remain flexible and let the students use the collaboration space as they

FIGURE 5. Example sketches generated by students during the Diverge phase.

FIGURE 6. Students working on collaborative board manipulating and annotating sketches.
would see fit. In that respect, ensuring that quantitative description of the results is possible was subordinate to providing flexibility.

**FEEDBACK SURVEY**

After the visual design sessions, we asked the students to fill out a short optional survey. Twenty-five students provided this feedback. Questions included what worked and what did not, if they had enough or too much time for different parts of the session, how they experienced the technical setup, and if they were comfortable doing the exercises and managed to “lean into” the creative aspect.

Regarding the Warm-Up Exercises

Feedback about the warm-up exercises varied. Whereas many students found that “easing people into the session with the poem etc worked well” [P1] and helped them “dive into creativity and moving on to the point of the exercise” [P19], others did not see the value of taking a line for a walk and quickdraw exercises [P6, P20].

Regarding the Technical Setup

Regarding the choice of communication and collaboration platforms and how these were prepared, the students were positive overall (both about the tools themselves and the templates that were set up), stating these are “amazing tools for learning” [P11] and “allows to organize ourselves in teams” [P9]. However, several students did report issues with the significant machine resources that the tools required and possible connectivity problems.

Regarding the Creative Aspect and Collaboration

Feedback on the creative aspect of visual design indicated that these sessions helped them step out of their “traditional modus operandi” [P14], and that it “was fun to be challenged creatively” [P4]. Initial nervousness was alleviated after the warm-up exercises. While some groups collaborated very well, others were limited by their group dynamics, leading to a loss of motivation.

Overall, feedback was very positive, and students were enthusiastic. As, for example, P14 indicated: “the problem or actually the beauty of this course is that it unteaches you everything you have learned previously in terms of following a straight path or a certain set of rules. It’s almost ‘against the grain’ of the status quo statistician way of thinking and I love it for that.”

**DISCUSSION**

Student feedback, the work generated during the session itself, and the subsequent designs created in the separate group projects showed that this online approach achieved similar results to the in-person sessions that were run in previous years. As expected and in concordance with in-person sessions, some students were able to generate widely diverse sketches whereas the design space explored by others remained limited. In this sense, the remote workshop can be seen as a feasible alternative in case in-person teaching is not possible. Nevertheless, converting the workshop to a remote setting required significant preparation and resources. These included creating the necessary infrastructure and templates in both the communication and collaboration platforms, as well as covering the financial cost for upscaling the communication platform above the free tier.

There are several things that we have learned by running these sessions. First, communication was very important. This was true between students and teaching team, amongst students, and within the teaching team itself. The communication platform used (i.e., gather.town) made this easier because these interactions were more “natural” and people could sit together or walk up to someone to initiate conversation. This was especially important to support students who struggled with the exercise. Second, preparation was key. This included making clear to the students what was expected, setting up the workshop room in the communication platform, and creating a template in the collaboration platform. These gave the students an overview of what needed to be done, and in what order, so that they had a structure to think and work in. Third, it was very important to remain flexible. Not only did this relate to the time allowed for students to complete the exercises, but also to how the students used the templates: these could be used as guides, but the students could also repurpose them and use them as they see fit. Finally, and in concordance with in-person sessions, students needed to be reassured during the exercises. As they were not used to exercise sessions that revolved around creativity, some were uncomfortable initially and were hesitant to make mistakes (although—by definition—no mistakes could be made in most of the exercises). This tension was exacerbated by the fact that either the Internet connection was intermittently lost for some of the students, or their computer was working at the top of their resources. As the students would employ the same approach in their group projects at a later stage, however, this issue was partly mitigated when they started working on their projects and could be guided through the process by their peers.
Apart from the visual design aspect of the workshop, these sessions indicated to be important for their social setting as well. Most of the students had very little social contact with their peers during the months preceding the workshops. As some of them have made casual comments that these workshops were a welcome change in the day-to-day succession of online lectures, we believe that these sessions can aid in alleviating the isolation that the students suffered from.

In the coming years, this course will be made available both in in-person and in remote versions. Even though the setup used was not perfect, we will continue to use the approach described in this article for future online installments, with some adaptations. First, in case the students do not know each other before the course, we will include activities at the start specifically aimed at tackling that issue. Second, a method needs to be included to notify the teaching team and other students if someone has connectivity issues. These notifications cannot be sent through the tools being used (i.e., gather and miro) and should not rely on additional software being installed, but should still be accessible and take privacy into account (e.g., without sharing telephone numbers). Finally, stricter instructions will be given on making sketches identifiable with initials and number (divergent phase) and describing clusters (emergent phase) in order to measure the outcome from the workshops quantitatively. The in-person version of the course will be adapted slightly given our experience with the online setting. Specifically, the sketching phase and discussions will remain in-person with the help of physical pieces of paper, but each student or group of students will be asked to upload these after the exercise onto a shared collaboration board so that the material can be accessed, analyzed, and reworked at a later time as well.

In conclusion, with this article, we report on our experience in running online visual design workshops in the context of data visualization for statistics and data science. It is important to note that this is a post-factum experience report rather than a designed experiment. We described the technical setup and how the separate sessions were run, and show to what extent these can replace in-person visual design sessions. In doing so, we hope to provide general guidelines and ideas for such sessions in the future.

Additional instructions, higher resolution images, and template miro boards are available at https://bit.ly/online-vis-design-supplemental.

ACKNOWLEDGMENTS

The authors would like to thank the students who participated in the courses and provided helpful feedback, and the Data Science Institute and UHasselt M.Sc. Statistics and Data Science for providing the funding for running gather.town. This work was supported in part by Hasselt University BOF under Grant ADMIRE BOF21GP17, in part by the Flemish Government programme “Onderzoeksprogramma Artificiële Intelligente (AI),” and in part by the Research Foundation of the University of Leuven and the Research Foundation Flanders.

REFERENCES


**JAN AERTS** is with Hasselt University, Hasselt, Belgium. He is the corresponding author of this article. Contact him at jan.aerts@uhasselt.be.

**JANNES PEETERS** is with Hasselt University, Hasselt, Belgium. Contact him at jannes.peeters@uhasselt.be.

**JELMER BOT** is with Hasselt University, Hasselt, Belgium. Contact him at jelmer.bot@uhasselt.be.

**DANAI KAFETZAKI** is with Hasselt University, Hasselt, Belgium, and also with KU Leuven, Leuven, Belgium. Contact her at danai.kafetzaki@uhasselt.be.

**HOUDA LAMQADDAM** is with KU Leuven, Leuven, Belgium. Contact her at houda.lamqaddam@kuleuven.be.