UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat d'Informàtica de Barcelona

FIB

Heatmap Perception Study

Degree in Informatics Engineering
Specialization in Computing

Bachelor Degree Final Thesis

Carolina Middel Soria

Director/tutor: Pere-Pau Vázquez Alcocer
Co-director: Elena Molina López
GEP's tutor: Joan Sardà Ferrer

29-04-2022

Abstract

A Heatmap is al representation of data which can essentially be described as a data table denoting sets of values as cells, encoded in colors. Chart Perception is the key to an effective chart design, which has been traced and studied for years to enhance the quality of visualizations. However, heatmaps seem to be forgotten.
In this project, we will study this heatmap's perception, testing common and standardized configurations through simple perceptual tasks.

Resum

Un Heatmap és la representació de les dades que es poden descriure essencialment com a taula de dades que denota conjunts de valors com a cel·les, codificades en colors. La percepció de diagrames és la clau per a un disseny de cartes efectiu, que ha estat rastrejat i estudiat durant anys per millorar la qualitat de les visualitzacions. No obstant això, els heatmaps semblen quedar oblidats.
En aquest projecte, estudiarem la percepció d'aquest heatmap, comprovant configuracions comunes i estandarditzades a través de simples tasques perceptives.

Resumen

Un heatmap es una representación de datos que puede describirse esencialmente como una tabla de datos que denota conjuntos de valores como celdas, codificados en colores. La percepción de los gráficos es la clave para un diseño eficaz de los mismos,
éste se ha rastreado y estudiado durante años para mejorar la calidad de las visualizaciones.
Sin embargo, los heatmaps parecen estar olvidados.
En este proyecto, estudiaremos la percepción de un heatmap, probando configuraciones comunes y estandarizadas a través de tareas perceptivas sencillas.

Acknowledgments

I would like to earnestly acknowledge the director, Pere-Pau Vázquez, and the co-director, Elena Molina, for letting me work for this project. But also for their sincere efforts, their empathy, guidance, feedback and valuable time despite their busy schedules.

I would also like to thank the director colleagues, experts in perception studies, who have helped me with some previous work examples or posted studies to get ideas and form the concrete data and tasks for the actual study.

And finally, appreciate and thank my family and friends for their constant support and their help in gathering participants for the study.

Thanks a lot !

Table of Contents

1. Context and Scope

1.1. Context

This project is the Bachelor Thesis of the Degree in Computer Engineering, with specialization in Computing, by the Facultat d'Informàtica de Barcelona (FIB) of the Universitat Politècnica de Catalunya (UPC).

The project is directed by Pere-Pau Vázquez and co-directed by Elena Molina. The director is currently working at Universitat Politècnica de Catalunya in Barcelona as an associate professor at the Computer Science Department. And is also a member of the Research Center for Visualization, Virtual Reality, and Graphics Interaction, ViRVIG Group [1]. The co-director is currently in her 2nd year of PhD, after obtaining a Master's degree in Innovation and Research in Informatics, both coursed in Universitat Politècnica de Catalunya in Barcelona, Catalonia, Spain.

1.1.1. Introduction

The goal of this Final Bachelor Degree Thesis is to assess heatmap visualizations perception, and obtain reasonable, quantitative and qualitative information for an effective heatmap visualization design and determine possible limits.

The problem and therefore the need to develop this project comes from the almost non-existence studies in Heatmap charts, and their perception in comparison to other visualization techniques. Despite the high importance and utility of heatmaps in various forms of analytics.

The project is divided into two main blocks. The first one, related to data preparation and the different approaches we have followed to get the data for the experiment and its configuration. And second one, the main user study, all its preparation and the results obtained.

Data Preparation

In this first part we discuss the approach we want to give to the data, and we mention the different attempts to get data and validation used for these. Furthermore, we will also discuss the configurations chosen to be shown in the study.

Heatmap Perception User Study

For the User study, we conducted an experiment crowdsourced in Prolific but also shared to family, friends and colleagues to measure the ability of certain tasks on the data samples shown.

1.1.2. Terms and concepts

Every field in the world has its own and special vocabulary. The same holds for chart representation in general, and this project's research. Over the years, just like the discipline itself, the lingo has also developed unique visualization terms that one should know.

So, compiled below are some of the most commonly referred to visualization concepts and key terms that will definitely benefit you to understand the project.

HEATMAP

A heatmap, also known as heat map, matrix chart, or XY heatmap, is a graphical representation of data. A heatmap indicates values across two axis variables as a grid of colored cells. So, basically, it is a table, where each cell in the table can contain a numerical or logical value that determines the cell color based on a color palette [2].

They are well known and used for showing variance across multiple variables. Giving us the perception of how the data represented is clustered or varies over the dimensions, revealing patterns and any correlation existence [3].

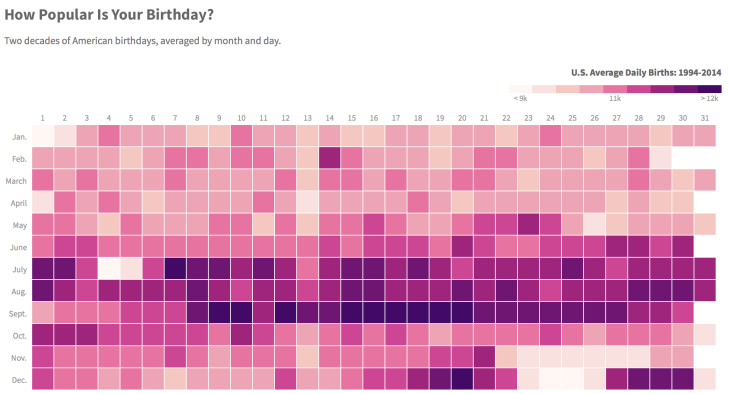


Figure 1. Heatmap representation on birthdays from 1994 to 2014.
From:https://thedailyviz.com/2016/09/17/how-common-is-your-birthday-dailyviz/

CLUSTER

The Cambridge dictionary defines cluster as "A group of similar things that are close together, sometimes surrounding something".

So, when we refer to the term cluster in the project, it would mean a group of data having similar behavior meaning being close and having a similar value. You can see a clear example in the Figure 2, where each cluster is marked with a different color.

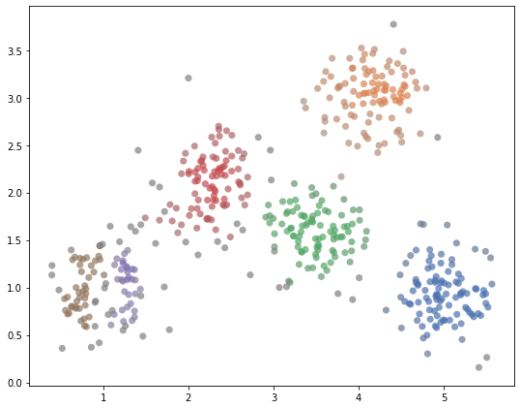


Figure 2. Cluster representation through a scatter plot.
https://towardsdatascience.com/hdbscan-clustering-with-neo4j-57e0cec57560

CROWDSOURCING

The term "crowdsourcing" is actually a contraction of “crowd” and “outsourcing”. It appeared for the first time in 2006 in a *Wired* [4] article, "The Rise of Crowdsourcing" from Jeff Howe [5].

Crowdsourcing is a very broad term with many applications and forms. Therefore, it has more than 40 different definitions in scientific and popular literature. Luckily, the researchers Enrique Estellés-Arolas and Fernando González Ladrón-de-Guevara, from the Technical University of Valencia, developed an integrating definition [6].

So, in this project, crowdsourcing will refer to the act of outsourcing the heatmap perception experiment to a large group of people to perform these tasks virtually.

We will use the crowdsourcing platform to get a part of our research results, not only because we can reach way more public, but for the following advantages:

- Lower costs: offering small rewards for doing the tasks is usually a lot cheaper than formally hiring people to do it.
- Greater speed: as there is a wider pool of people can speed up the task completion process, especially when completing many small tasks in real-time.
- More diversity: we want to evaluate the heatmap perception in general without unconscious or implicit bias established while conducting the experiment. By

crowdsourcing our experiment, we can benefit from others with different backgrounds, values, and life experiences.

1.1.3. Problem to be solved

Reading charts usually presents visible difficulties depending on the configuration of different parameters such as the color palette used, the width and height of different visual marks, pixels' size, and so on.

As mentioned before, the problem encountered is that perception or limits on common visualizations such as heatmaps are still unknown. Hence, the project aims to explore the perceptual limits and best configuration based on a study on Heatmap Perception. Study the visual depictions equally if some parameters are changed such as color palettes, size of marks, etc. evaluated in perception tasks such as cluster detection.

1.1.4. Stakeholders

The project has different parties involved and considered stakeholders, which can be classified in two groups.

First one would be the stakeholders with a direct implication in the project. Which we consider as the director, co-director and researcher, Carolina Middel, who has been responsible for the development of the project. We may also include perception and visualization experts apart from the director and co-director that have affected in some way on the research of the project and the knowledge is reflected in this research area of the project.

Then, the second group, composed of those who have not had any involvement in the project but receive benefits from it. For instance, a clear example of this group could be the research community. Another example would be all the people who use heatmap charts to display data in a company, such as people working on BI or data analytics in any sector.

1.2. Justification

1.2.1. Previous studies

First, we are going to take a brief overview on where and how heatmaps are used in previous studies.

As mentioned, Heatmaps are a pretty common visualization technique. They have multiple applications and has been used in several disciplines, for instance, in bioinformatics to encode gene expression patterns such as the example in Figure 3, in medicine for volume surface visualizations [7], but also in geography to visualize temperature [8] or even for abstract trajectory analysis [9].

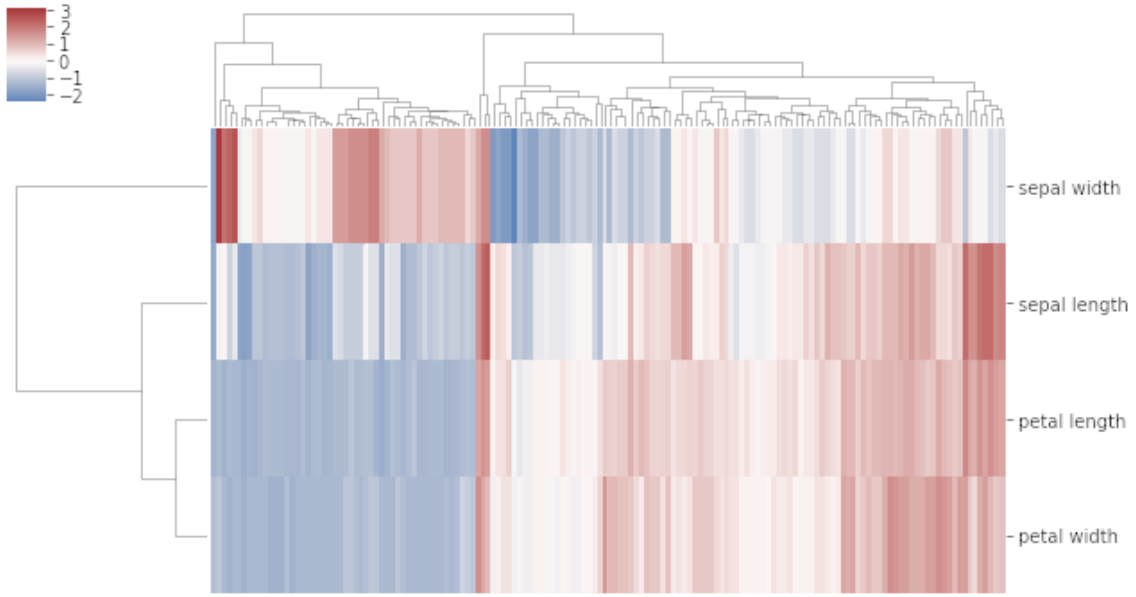


Figure 3. Clustered heatmap, each column a flower specimen, and each row a measurement from that specimen. https://chartio.com/learn/charts/heatmap-complete-guide/

Although all the research gathered, there is no specific study related to visual perception studying heatmap charts. Nonetheless, there are some punctual studies on visualizations' perception which include heatmaps as a secondary chart where we could gather some information from.

For instance, due to the need for group identification in some research areas, usually related to science, cell size can be adjusted (usually enlarged) to facilitate the inspection. But we have also found a workaround to better detect these groups (clusters) heatmaps combined with dendrograms, also known as cluster heatmaps, shown in Figure 3.

1.2.2. Justification

The concept of data visualization could be defined as a graphical representation of data for a more quick and efficient understanding and communication. It is used for many practices and fields nowadays.

However, the existence of heatmaps' research cannot be traced as much as other visualization techniques. Therefore, the knowledge on how to configure the representation or the limits on those types of charts can barely be extracted from some individual previous studies and our own intuition.

Consequently, this is directly affecting heatmap representation and conducting the common heatmap tasks, such as identifying clusters, trends, outliers or certain values. Hence, for all these reasons, this research is needed to expand the knowledge of the visualization community in heatmap charts.

1.3. Scope

1.3.1. Objectives and sub-objectives

As stated in Section 1.1.3, the goal of the project is to study perception on heatmap charts and luckily conclude the best configurations or limits on their configuration. To accomplish this main objective, we have decided to work in different sub-objectives.

THEORETICAL PART

- **Standard parameters**
    - Study heatmap charts and check previous studies standard parameters to use in our heatmap configurations.
- **Python libraries**
    - Explore possible Python libraries, to detect clusters, generate synthetic data or plot charts, for instance: DBScan, HDBscan, K-means, Altair, Seaborn or Numpy…
- **Dataset environment**
    - Explore different datasets. Define a data environment that is familiar to the users so that they can easily understand the survey tasks. And evaluate the data quality.
- **Crowdsourcing platform**
    - Evaluate different crowdsourcing platforms to perform the tasks in.
    - Compare these platforms.
    - Conclude which is the best platform to use.

PRACTICAL PART

- **Data preparation**
    - Gather and prepare the data.

    - Validate the clustering results.
    - Design configuration. Considering cell size, color palette, color gradient, labels, legend…
    - Validate groups with a cluster detection algorithm to assess the veracity.
- **Study preparation**
    - Structure of the study.
    - Explanation and introduction.
    - Database design and info to gather.
    - User prescreening .

MAIN OBJECTIVE

- Conclude if there exists any significant difference on the means between the configurations.
- Present all the conclusions in simple, clear documentation for all the future stakeholders.

1.3.2. Requirements

There are requirements needed to ensure the quality of the project research and ensure the veracity of the results.

- The tasks and visualization should be clear and intuitive for the user.
- To validate the number of clusters in the representation, through cluster detection algorithms and user check. Including the efficiency of the algorithms.
- The use of good programming practices, no complexity, and a readable style.
- To be useful, efficient and reusable for future studies.
- To pass the Anova and Bonferroni test with significant statistics results.

1.4. Methodology and rigor

In a thesis, it is important to discuss the methods used to do the research. This methodology chapter explains how we approached the objectives allowing you, the readers, to evaluate the reliability and validity of the study.

1.4.1. Methodology

The thesis aims to assess visualization design and human perception of heatmaps representations, as an under-researched topic. Therefore, it has been very important to work with a flexible methodology that allows returning to any previous point in case the experimentation doesn't result as we want, as happened several times. However, the methodology used has made those changes and the process itself easier, faster and effective. I am speaking about the Agile methodology. It is characteristic of software development, however, it meets the requirements stated above and can be easily applied to our scope. The idea is to define short development cycles of designing, developing, and testing, to afterwards evaluate them and proceed or discard them to start again.

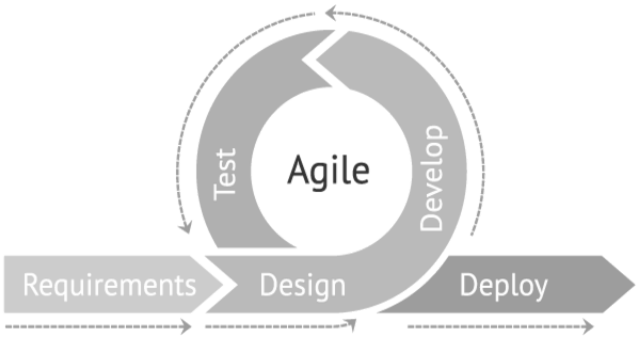


Figure 4. Gantt Chart representation.
From:https://www.price2spy.com/blog/agile-development-methodology/

However, the methodology changed for the last part of the project. The new methodology approach is called Muda.

Muda is about getting rid of waste, as any kind of activity related to the project that does not provide added value. It is characterized as seven original wastes, but the ones applying to the project are:

Figure 5. Muda methodology representation.
From: https://www.ermeo.com/en/blog/7-mudas/

Getting rid of transport and motion, working always at least from the same town, so for this last development phase the student always worked from Barcelona and assisted to the office only once a week.

In addition, we also skipped the waiting. The student reached the director and co-director through a group chat in discord at any time, so they could give quick and more effective feedback. And started working on many things where there was any backlog, like the payment transaction of Prolific for the survey deployment, in the meantime we worked with data collected by friends, family and colleagues, sending the link directly to the website.

Finally, we also reduced overproduction, which has been one of the main causes of the delay caused by the different problems on the data preparation development explained in future sections.

1.4.2. Validation and tools

In order to ease the track of the goals, we will schedule a weekly meeting with the director and co-director to check and evaluate the objectives for the week. And we will work with a checklist with the task to do, so we can manage and keep organized all the tasks needed to do for the week.
Moreover, for the version control, we are using Github [10]. It allows you to track project changes and serves as a backup. Changes will be updated easily, and it is able to track them. We also track project changes during the documentation using Google Docs [11] that also allows us to back up previous versions of the document.

2. Time planning

2.1. Description of the tasks

This project has taken around 590 hours, distributed in 160 days. It started on September 20, 2021, until April 22, 2022. So, I worked on the project for around 4 hours per day. The oral defense of the project will be April 29.

2.1.1. Task definition

This section describes the necessary tasks to achieve the execution of the project. It will specify the dependencies, duration time, and description for each one of them. You can find the task summary represented in Table 1 and the Gantt chart representation in Figure 6.

Project management

Project management tasks define the goals, resources and needs. Those are the first tasks to complete, as they allow the project to be structured correctly and give consistency to the rest of the tasks.

- **Scope and contextualization.** References the definition of the objectives, the importance of the subject in the field of study as well as the contextualization of the thesis and a brief overview of the previous studies and the justification of the project.
- **Temporal planning.** Describe project tasks, their duration and dependencies. In addition, a Gantt chart is presented to illustrate the planning of the project. It provides different phases, resources, and requirements, necessary for the project's execution.
- **Budget and sustainability.** Define the analysis of the costs and sustainability impact of the project.
- **Meetings.** Referring to the meetings with the tutor/director of the project. In this case, to ensure a correct evolution and fulfillment of the time plan, they are scheduled once a week.

Data preparation

In this first part of the project, all the tasks are related to the data used for the project and their visualization. First, we conducted intensive research on previous studies and heatmap standard parameters to get a general overview of what the study had to be. The tasks in this first part are:

- **Previous studies.** Research heatmaps uses, standard configurations, examples, and define the scope for our User Study.
- **Gather the data.** After defining how we wanted to display the data and what it would represent, we had to actually get the data. First we tried to get real data. We ended up not using it, so we worked on generating them synthetically.
- **Cluster detection.** As one of the tasks of the study is to mark the number of clusters, we must first make sure that the number of clusters indicated is really the number shown through the heatmaps, we did so by validating it with a cluster detection algorithm. In addition, we conducted some user checks on the cluster representation that can be found in the Annex {_}.
- **Plotting data.** Finally, the only thing we were missing was the representation through heatmaps.

Having all this, the only thing left was to carry out the study implementation and deployment, besides the Prolific configuration. The tasks in this second phase are:

- **Implement the web form.** We will use the Flask programming language to implement the web where the experiment will run.
- **Configure Prolific.** Composed by all the small tasks necessary to set up a study in the crowdsourcing platform. These tasks are: configure the study details, study link, study completion, audience prescreening, study cost, bank transfer and publishing the study.
- **Post the experiment.** Publish both, the website on the free host server Python Anywhere and the study on Prolific.
- **Review the answers.** We need to prove the veracity and validity of the users' answers.
- **Statistic analysis of the results.** Review user responses and their validity. For this purpose, we have certain validation questions in the survey that just had to be checked.
- **Draw conclusions.** And to finish this second part, it will be necessary to draw all the conclusions with the results obtained from the study.

The completion of all the previous tasks leads to the two final tasks.

- **Integration of the final document.** Ensure the correctness of all the info gathered in the previous tasks and document them all together.
- **Prepare the oral defense** of the project's presentation. Prepare possible questions, and study the execution details. Besides, rehearsals.

2.1.2. Summary of the tasks

In this section, we will define the duration and dependencies among the tasks in a table format displayed in Table 1.

TASK ID	TASK TITLE	TASK DEPENDENCY	START DATE	DUE DATE	DURATION (h)	% OF TASK COMPLETE
1	Project planning				105	100%
1.1	Scope and contextualization	-	20/09/21	1/11/21	40	100%
1.2	Planning	-	27/09/21	10/04/22	30	100%
1.3	Budget and sustainability	-	4/10/21	1/11/21	25	100%
1.4	Meetings	1.1	20/09/21	1/11/21	10	100%
2	Data preparation				210	100%
2.1	Preliminary research	-	12/9/21	01/04/22	40	100%
2.2	Collect data	-	12/9/21	01/04/22	40	100%
2.3	Synthetic data creation	2.2	1/12/21	01/04/22	70	100%
2.4	Cluster detection	2.3	1/12/21	01/04/22	70	100%
2.5	Chart plotting	2.4	15/02/21	01/04/22	40	100%
2.6	Fine tune	2.5	15/02/21	01/04/22	30	100%
3	Study implementation and analysis				160	100%
3.1	Croudsourcing platform	-	15/01/22	01/04/22	25	100%
3.2	Flask Web implementation	-	01/12/21	05/04/22	90	100%
3.3	SQL Database implementation	3.2	01/01/22	05/04/22	40	100%
3.4	Post the experiment	3.1. 3.2. 3.3	10/04/22	20/04/22	10	100%
3.5	Review the answers	3.4	11/04/22	20/04/22	40	100%
3.6	Statistic Analysis of the results	3.5	12/04/22	22/04/22	70	100%
3.7	Draw conclusions	3.6	14/04/22	22/04/22	50	100%
4	Final tasks				120	55%
4.1	Integration of the final documentation	1. 2 i 3	15/11/21	22/04/22	100	100%
4.2	Prepare the oral defense	4.1	23/04/22	28/04/22	20	10%

Table 1. Summary of the tasks. Prepared by the author.

2.2. Resources

There are some human and material resources needed for the achievement of the project.

2.2.1. Human resources

The main human resources are the researchers responsible for the execution of the project. The student Carolina Middel Soria, with the role of researcher, in charge of the implementation, analysis, and documentation. Also, the project director Pere-Pau Vazquez and co-director Elena Molina, responsible for leading and guiding the researcher. And last, the GEP tutor for supervising and giving advice about the project's management and documentation's correctness.

2.2.2. Material resources

The project requires different hardware and software materials.

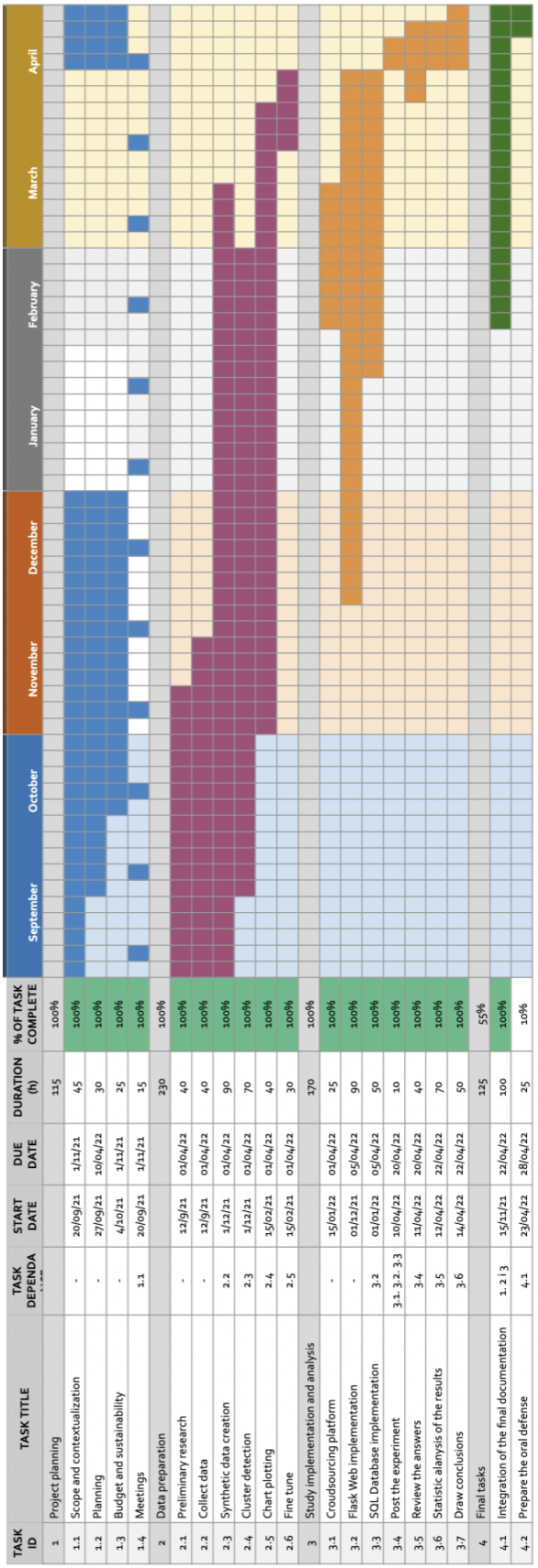
The hardware resource per excellence is the pc or laptop used for the implementation. There will be two different laptops used:

- Apple MacBook Air 13,3": Processor M1 / RAM 8GB / Memory 256GB / 3,2 GHz.
- Xiaomi Mi Air 13,3": Processor Intel Core i5 8th Gen / RAM 8GB / SSD 256GB / 1,60 GHz.

The software resources used are related to all the tools used for the development of the experimentation or the documentation.

- Atenea - Platform to communicate with the GEP's and deliver the weekly work.
- Racó - Platform to check TFG status.
- Github - It will be used as a tool for Version Control due to its accessibility and version recovery.
- Prolific - The crowdsourcing platform to outsource the perception study as a flask web survey.
- PythonAnywhere - Web hosting service based on the Python programming language.
- Grammarly - Check the documents' grammatical mistakes on the spot, with no interruptions to your writing process.
- Google Docs - For the project documentation.

2.3. Gantt chart representation



TASK ID	TASK TITLE	TASK DEPENDA	START DATE	DUE DATE	DURATION (h)	% OF TASK COMPLETE
1	Project planning				115	100%
1.1	Scope and contextualization	-	20/09/21	1/11/21	45	100%
1.2	Planning	-	27/09/21	10/04/22	30	100%
1.3	Budget and sustainability	-	4/10/21	1/11/21	25	100%
1.4	Meetings	1.1	20/09/21	1/11/21	15	100%
2	Data preparation				230	100%
2.1	Preliminary research	-	12/9/21	01/04/22	40	100%
2.2	Collect data	-	12/9/21	01/04/22	40	100%
2.3	Synthetic data creation	2.2	1/12/21	01/04/22	90	100%
2.4	Cluster detection	2.3	1/12/21	01/04/22	70	100%
2.5	Chart plotting	2.4	15/02/21	01/04/22	40	100%
2.6	Fine tune	2.5	15/02/21	01/04/22	30	100%
3	Study implementation and analysis				170	100%
3.1	Croudsourcing platform	-	15/01/22	01/04/22	25	100%
3.2	Flask Web implementation	-	01/12/21	05/04/22	90	100%
3.3	SQL Database implementation	3.2	01/01/22	05/04/22	50	100%
3.4	Post the experiment	3.1 3.2. 3.3	10/04/22	20/04/22	10	100%
3.5	Review the answers	3.4	11/04/22	20/04/22	40	100%
3.6	Statistic alanysis of the results	3.5	12/04/22	22/04/22	70	100%
3.7	Draw conclusions	3.6	14/04/22	22/04/22	50	100%
4	Final tasks				125	55%
4.1	Integration of the final documentation	1. 2	3	15/11/21	22/04/22	100
4.2	Prepare the oral defense	4.1		23/04/22	28/04/22	25

Figure 6. Gantt Chart representation. Prepared by the author.

2.4. Risk and alternatives

Anticipating the risks and obstacles of a project is fundamental. As well as having solutions or alternatives to prevent them and keep our project on track. Thus, in this section, we will describe all the risks predicted that actually happen and the alternatives we worked with to solve them and move forward.

2.4.1. Impossibility of accomplishing the 3D part of the project

The very first approach of the project was to study Heatmap perception on 2D and 3D environments. Due to schedule issues, combining work and the development of the thesis, the student and researcher was not capable of assisting the virtual reality laboratory, so the part of the study in the VR environment was not possible to accomplish.

- **Impact:** Low, since we were prepared for this risk from day one.
- **Alternative:** The 2D heatmap research was extended to evaluate more parameters and configurations.

2.4.2. Deadline of the project

One of the common problems in projects of every kind and field is the deadline and the limited time to work on it. Despite all the tasks that have been planned to assess the project on time, we encountered different backlogs during the development which led to a delay.

- **Impact:** High.
- **Alternative:** We took different workarounds on the data preparation using different libraries or calculations until we got the one working for our study. This led to overworking and losing more than 4 weeks of work and research. We will get into the process' details in further sections.

2.4.3. To not get the results expected on the research

Studying an under-researched topic, you cannot always expect to get the conclusion you hypothetically imagine in the beginning. Our goal is to help other researchers represent data through heatmaps, clarifying the best ways to represent them, but if the results are inconclusive we can also present the limits on the study for future studies.

- **Impact:** Medium.
- **Alternative:** In any case, a non-significant or null result hides information that can generate valuable learning for decision making. What we mean is that

non-significant results still have the potential to show valuable information about the study.

3. Budget and sustainability

After the context is defined and the planning is done, we need to carry the budget and sustainability of the project. We are going to define the different types of costs associated with staff, workspace, and tools and devices used. In addition, to overcome the obstacles that appear and assume the sudden costs, a contingency plan is made for unexpected costs.

3.1. Budget

3.1.1. Staff costs

To estimate the personnel costs, I am going to define the roles and their costs/h per task needed to develop. The workers' salary is estimated with real salaries. The cost for one task will be the sum of all the costs of the personnel. Then, the cost for each worker will be computed by multiplying the cost/h by the number of hours expected for the task. This is shown in the next Table 2. The cost and roles estimation have been extracted from Glassdoor [12].

Role	Cost per hour
Project manager	20€/h
Researcher	17€/h
Developer	15€/h
Analyst	15€/h

Table 2. Personnel costs. Based on the information supplied by the salaries exposed in the IT sector in Glassdoor. Prepared by the author.

Task	Hours	Project Manager	Researcher	Developer	Analyst
Project Planning	115	115	0	0	0
Data preparation	225	0	70	155	0
Study implementation and analysis	170	0	0	75	95
Integration of the final documentation	80	0	45	0	35
Prepare the oral defense	25	0	15	0	10

Table 3. Estimated duration per task for each personnel. Prepared by the author.

Task	Cost (€)
Project Planning	2300€
Data preparation	3515€
Study implementation and analysis	2550€
Integration of the final documentation	1290€
Prepare the oral defense	405€
TOTAL	**10.060€**

Table 4. Estimated cost per task. Prepared by the author.

Role	Hours	Cost (€)
Project manager	115	2300€
Researcher	130	2210€
Developer	230	3450€
Analyst	140	2100€
TOTAL	-	**10.060 €**

Table 5. Estimated working hours and cost per personnel. Prepared by the author.

3.1.2. Generic costs

Amortization

The total cost of the project also takes into account the amortization of the resources. We will consider the software and hardware used for the project.

In the hardware dimension, we have estimated 4 hours of working per day, for 160 days of work. Thus, it has been estimated that 40% of the project will be developed on the Xiaomi Mi Air laptop, and 60% will be done using the Macbook Air laptop. Then, for those resources, we will consider a life span of 3 years.

So, the amortization of those resources is computed by the following equation:

$$\text{Resource} = \textit{Resourse's price } x \frac{1}{life\ span} x \frac{1}{160\ days} x \frac{1}{3\ hours} x\ hours\ used$$

So, the cost of the hardware resources is represented in Table 6.

Hardware	Price	Life span	Hours	Amortization
Xiaomi Mi Air	720€	3 years	236h	118€
MacBook Air M1	1100€	3 years	354h	270,42€
TOTAL	-	-	-	**388,42€**

Table 6. Cost of hardware resources. Prepared by the author.

Thus, we also need to consider Software amortization. There are different Software used in this project, not all of them include fees to use them. The programs used are:

Software	Fee	Cost (€)
Google Docs	Free	0
GitHub	Free	0
Prolific	£175 ≃ 209,47€	209,47
TOTAL	-	**209,47**

Table 7. Cost of Software resources. Prepared by the author.

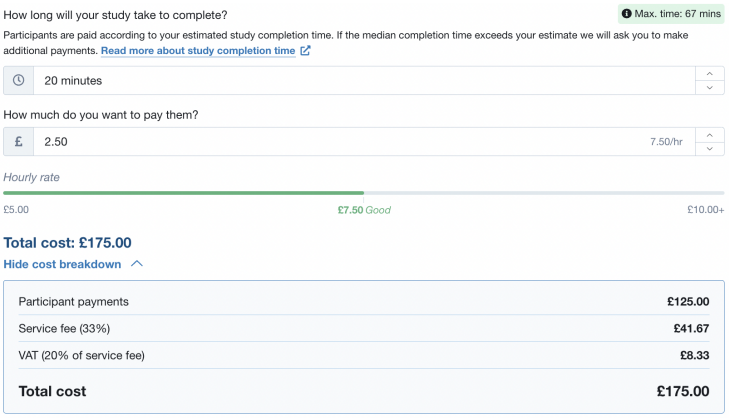


Figure 7. Prolific costs. Captured by the author.

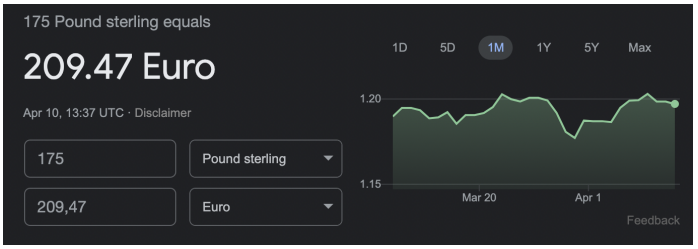


Figure 8. Pounds to Euro conversion. Captured by the author on April 10th.

Internet cost

The internet rate costs 90€ per month. Taking into account that the project lasts 5 months and that the working hours per day are approximately 4h.

Internet cost = 8 $months$ * (90€/$month$) * (4h/24h) = **120 €**

Electric cost

The actual electric cost is 0,2424 €/kWh [13]. We only need to consider the hardware when it is connected. The following Table 8 shows the individual and total costs of the electric consumption.

Hardware	Power (W)	Time (h)	Consumption (kWh)	Cost (€)
Xiaomi Mi Air	46	236	10,856	2,63
MacBook Air M1	3,3	354	1,168	0,28
TOTAL	-	-	-	**2,91€**

Table 8. Estimated electric consumption cost. Prepared by the author.

Travels costs

To transport from my hometown Blanes to Barcelona. I use the T-casual card for public transport, and it costs 45,05 € per 10 trips. Then, I use 2 trips per week, one to go to Barcelona on Sundays and then another one on Fridays to come back home. So, if the project is carried out in 15 weeks I will need 60 travels equivalent to 6 cards with a total cost of **270,30 €**.

Workspace costs

This project will mainly be developed in my student flat, rented in Barcelona. The rental cost of the flat is 1.235/month. However, I share the flat with 4 more people, so my rental fee ends up being 270€ plus dispenses per month. Hence, the cost of the workspace per 8 months is **2.160 €**.

Generic cost of the project

Thus, Table 9 summarizes all the generic costs, stated one by one in the previous sections.

Concept	Cost (€)
Amortization cost	388,42
Internet cost	120
Electric cost	2,91
Travel cost	270,30
Workspace cost	2.160

TOTAL generic cost	**3.151,10**

Table 9. Total generic cost of the project. Prepared by the author.

3.1.3. Contingencies

It is important to consider a cost overrun to cover obstacles and contingencies for any project. In this case, since it is a research project using innovative technologies, the probability of encountering problems during the project is high. Therefore, it has been decided to add a 15% cost overrun. You can see the total contingency of the project represented in Table 10.

Type	Cost	Contingency
Hardware amortization	388,42	58,263
Software	209,47	31,4205
Internet	120	18
Electricity	2,91	0.43
Transportation	270,30	40,54
Workspace	2160	324
TOTAL	**3.151,10**	**472,66**

Table 10. Estimated working hours and cost per personnel. Prepared by the author.

3.1.4. Total cost

Finally, the total budget estimation of the project. To get this content, we summed all the previous costs together. In Table 11, it is shown the total cost of the work.

Type	Cost (€)
Staff costs	10.060
Generic costs	3.151,10
Contingencies cost	472,66
TOTAL	**13.683,76**

Table 11. Total cost. Prepared by the author.

3.1.5. Management control

In big projects, it is very probable that the budget and time estimations will not be 100% fulfilled due to obstacles or other unexpected events. Consequently, we need to define a model for controlling the potential budget deviations.

We present our management control:

1. Deviation of staff cost per task:

$(estimated\ cost\ -\ real\ cost) * real\ hours$

2. Deviation of task development duration:

$(estimated\ hours\ -\ real\ hours) * real\ cost$

3. Total deviation in task development:

$total\ estimated\ cost\ -\ total\ real\ cost$

4. Total deviation of resources (human, software, hardware, or material):

$total\ estimated\ cost\ -\ total\ real\ cost$

5. Deviation of the total incident cost:

$incident\ estimated\ cost\ -\ incident\ real\ cost$

6. Total hour deviation:

Total $estimated\ hours\ -\ total\ real\ hours$

3.2. Sustainability

Nowadays, taking part in conducting a sustainability analysis in real projects is really important and may have a big impact. The principles of sustainability are the foundations of what this concept represents. Therefore, sustainability is made up of three pillars: the economy, society, and the environment. These principles are also informally used for profit, people, and the planet [14]. The author of the project will assess all three pillars through some questions [15], but also with a self-assessment on sustainability.

3.2.1. Self-assessment

Throughout my studies in Computer Engineering, sustainability has been a subject often present, as well as the importance of caring about the environment, and support to long-term ecological balance.

Personally, I have always been aware of the need for an economic analysis to ensure the viability of a project. However, the environmental and social dimensions are also significant, but not typically considered.

Giving the environmental dimension importance is fundamental for all projects, but especially those requiring a huge consumption of resources. However, a research project, such as this one, requires just a few resources, so it is less related to the quality of the environment and its impact on it. Nevertheless, the various conferences and papers on the environment made me realize the need to address sustainable solutions, and take on to consider minimizing the impact of my work on the environment.

Finally, I got to understand that any project must cover a need in society, therefore, this project's purpose is to help analysts and researchers to represent data effectively with heatmaps, and to know which are the limits and boundaries from those visualizations. On a personal level, this project will give me the necessary knowledge to continue my development in the field of data representation.

The process of designing this regenerative world starts by spending time understanding how the systems of life work and encouraging it with our projects.

3.2.2. Environmental impact

Have you estimated the environmental impact of the project?

The realization of this project does not have a major environmental impact, only the manufacturing of the hardware components used in the project, the electricity consumed to develop it, and the transportations realized by the researchers for the face-to-face meetings.

Did you plan to minimize its impact, for example, by reusing resources?

One way to minimize the initial environmental impact is to use only one computer for the development of the project. We cannot reduce electricity consumption using previous studies because no one has studied the heatmaps' performance. But we can also minimize the transportation impact by meeting online when possible.

How will your solution improve the environment with respect to other existing solutions?

The solution will give knowledge to researchers, scientific or analysts on representing data through heatmaps. Due to this, the electricity consumption would be decreased, as they should no longer experiment to get the best representation.

3.2.3. Economic dimension

Reflection on the cost you have estimated for the completion of the project.

This reflection, as well as the management control, are stated in Section 3.1. We have considered human, hardware, software, and material resources. Moreover, the estimated costs also include other costs, such as contingencies and incidental costs, as well as their possible effects on the budget.

How is the economic issue currently solved related to the problem that you want to address?

The economic issue or cost related to the problem exists because there is no previous research on how to use heatmaps, when to use them and which configuration to represent the data. One way to reduce the economic cost in the project would be by doing the experiments through an open-source platform or test the experiments with researchers we know personally through a free form. But it would increase the developing time.

How will your solution improve economic issues with respect to other existing solutions?

In case we succeed in this project, we would be able to show the best configurations and limits on heatmaps representation. This would prevent other people from experimenting with all the configurations and reduce all the costs included in our project.

3.2.4. Social dimension

What do you think you will achieve -in terms of personal growth, from doing this project?

In terms of personal and professional growth, this project will introduce me to the research world, learn about doing experiments, asking for feedback from specialists and observing

results, and extracting conclusions. Moreover, this project is real research and has helped me to gain experience in formal writing documentation and planning correctly and organized. Last, it has also taught me how to assess sustainability aspects in a project.

How is currently solved the problem that you want to address (state of the art)?

The problem is solved by observing standard values and patterns in previous studies and experimenting with these and new configurations of representations and data to get results to get limits on perceptions and conclusions.

How will your solution improve the quality of life with respect to other existing solutions?

The aim of this project is to help people wanting to represent data through heatmaps to know the best configurations and the limits for this type of visualization. Hence, with the research done, people would have more time to focus on other important tasks rather than focusing on how to represent the data.

4. Study

4.1. Related work

Perception has been long studied in order to improve its quality and effectiveness of visualization.

The purpose of this project is to assess heatmap charts perception in order to communicate guidelines for an effective and expressive heatmap design. The goal is to make it easy to understand and unambiguous so that it can accelerate and improve data analysis and decision-making tasks. In addition, we would also define any limitation or any other conclusion resultant from the study that can be reused by heatmap consumers or future researchers.

To get the results and conclusions on heatmap perception, we conducted an experiment on the charts through simple and basic tasks, usually referenced as low-level tasks or elementary perceptual tasks. This practice has been used to study perception on visualization in lots of previous research. For instance, Cleveland and McGill [16], in 1984, published one of the most referenced studies on visualization's perception, where they presented for the first time the concept of "elementary perceptual tasks". They talk about visual encoding, which refers to the properties used to encode data in a visualization, including position, area, volume, length, shading, direction, angle, curvature, and color, shown in Figure 7.

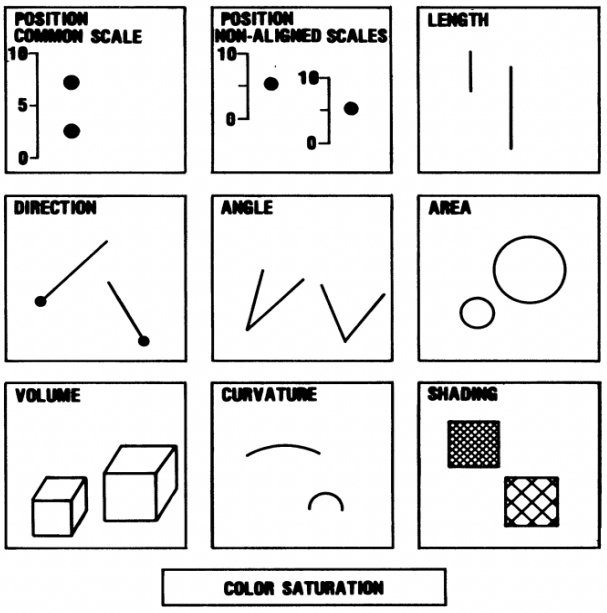


Figure 9. Elementary perceptual tasks to evaluate visualizations. From:
http://faculty.washington.edu/aragon/classes/hcde511/s12/readings/cleveland84.pdf

In addition, some research like [17] suggests that users are an integral part of a visualization process,as Figure 10 shows, meaning the visualization's effectiveness depends on perception, but also cognition, and its tasks and goals. Therefore, we decided to consider users' data in the results' evaluation process, gathering demographic information in the survey, but also a personal evaluation of the experiment, and the time to complete each task performed.

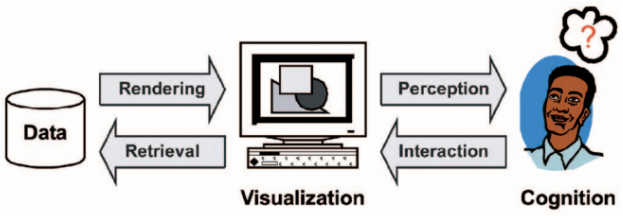


Figure 10. Visualization process.
https://ieeexplore-ieee-org.recursos.biblioteca.upc.edu/stamp/stamp.jsp?tp=&arnumber=1260759

Further analysis and research on previous work is explained on each appropriate section of the following block that will get into the definition of the study.

4.2. Study definition

Once we had an overview of what a perception study must be, we had to define the key parts of the study:

- Data scope. It refers to the definition of the data environment, what the data is going to represent. As we have seen in the previous section, the perception and the results that we obtain depend on many factors, one of them, the human. As we cannot control previous studies or the user's environment in general, it is essential that the data represents a familiar and intuitive environment with an easy and familiar logic for the users. So that this affects the results as little as possible.

- Configuration parameters. Referring to all those parameters that can be modified in a Heatmap, and that therefore, we are interested in estimating to see if they can affect the perception of the data in a positive or negative way. Among these parameters are cell size, heatmap size, cell shape, color palette, color continuity, legend configuration, grid lines, etc.

- Tasks. As discussed in the previous section, we have evaluated perception conducting elementary perceptual tasks assisting many visual encodings.

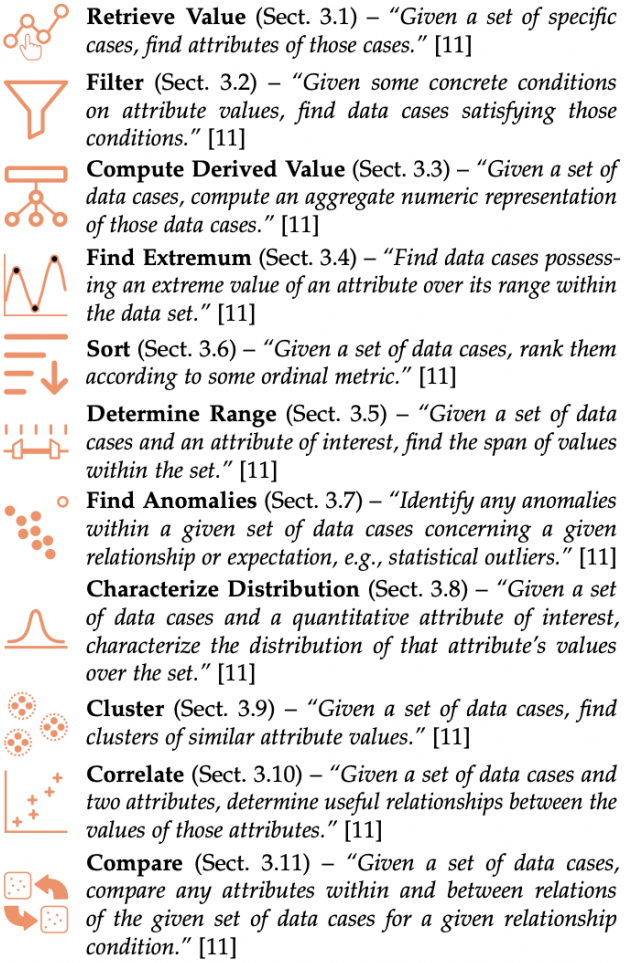


Figure 11. Possible tasks to evaluate perception for different charts.
https://www.researchgate.net/publication/353284356_A_Survey_of_Perception-Based_Visualization_Studies_by_Task

4.2.1. Scope

As discussed above, many factors affect the understanding and therefore the responses of the user study. To facilitate the evaluation of the tasks and omit possible misunderstandings or answers unreliability, we have decided to focus the data environment on a simple topic such as bike sharing in the city of Barcelona, so that users can quickly and easily locate the data, and therefore the tasks they perform in the study.

Thus, the heatmaps will represent the average number of bicycles per station (Y axis) for a day (X axis) from a bike sharing system. As you can see in the next Figure 12.

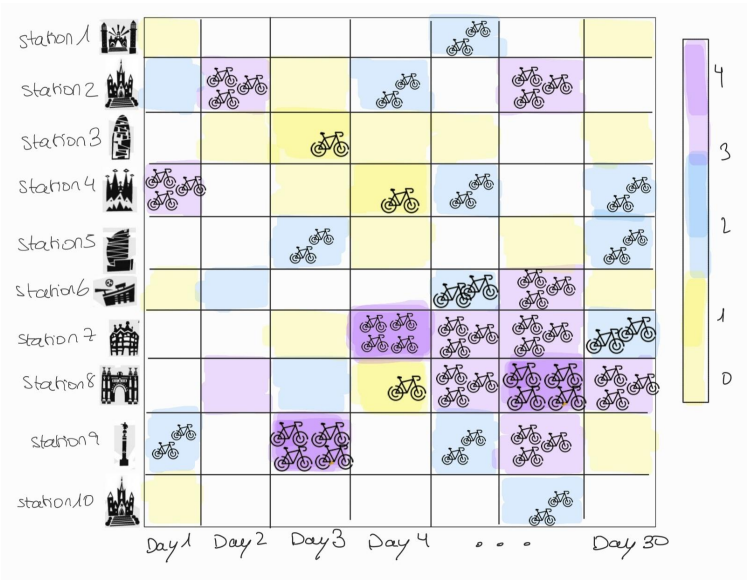


Figure 12. First study representation sketch. Prepared by the author.

4.2.2. Standard parameters on Heatmaps

To begin with the data preparation, we first researched standard used configurations in heatmaps on previous work. This way we could already extract some insides for the design and data needed.

This information is complemented with the information stated in the heatmap technical report conducted by the director and co-director[18] before the study.

The technical report reviews classical and modern visualization references, for instance [19][20][21][22]. In addition, there is also a systematic analysis conducted on research papers from the two major conferences on visualization: EuroVis, the European Visualization conference, organized by EuroGraphics. And, Vis, the visualization conference, organized by IEEE. Resulting in 300 papers with a total of 75 heatmap-like charts to evaluate. Finally, since these papers were showing most of the time heatmaps representations used in science, they also analyzed common charts' usage in non-scientific publications, by performing a search in Google Images for the word "heatmap chart" and concluding over the first 122 results.

4.2.2.1. Color Palettes

There were a great variety of palettes included in the research. However, the most common were: Blues, oranges and greens in sequential values, and viridis and magma for diverging

quantities. In addition, red-green palettes were pretty common, despite the potential well-known color-blind people problems.

Knowing this, we finally decided to work with two different palettes, using the divergent palette Viridis and the sequential Blues palette. We decided to also associate them with data values in two different ways, as a discrete set of colors, where each one is associated with a numeric range, but also as a continuous function between the color and the value.

4.2.2.2. Dimensions

Regarding the dimensions of the graphs, meaning the number of cells in the graphs, they also found a great variety of values. In the papers' charts, the values start from 4x4 until hundreds, they also mentioned that the great majority shows a relatively small number of cells. However, in the common google search heatmaps, the dimensions goes from 12-15 cells to 30-40 per dimension.

Therefore, although the first idea was to represent a number of 120 cells per dimension, we decided to reduce it. Then, while working with time data in one of the axes, the first idea was to work with 24 cells in each dimension and represent the hours in a day. However, we finally decided to represent the average number of bikes per day and stations, having then 30 cells per dimension.

4.2.2.3. Cell size

Although there are previous studies that say there may be cases where these are as small as one pixel [23], in the papers' cases, the cells were clearly visible, occupying more than 10 pixels wide. And the majority of the Google images also show relatively big cells, most of them rectangular.

In our case, to fit the image nicely in the survey we needed it to be around 500px, which led us to finally work with a 525x470px with the margin which ended being 390px of the chart, having a cell size of 13x13.

4.2.2.4. Tasks

The most common tasks performed were: correlation exploration matrices, and search for patterns, shapes, or trends.

Although those are the common tasks, it is difficult to present to the users as a simple task to mark the correlation, patterns or trends in the data. Thus, with the information about the

tasks presented in Figure 11 and the experiments performed in [24], we decided to perform 3 different tasks:

- Mark the number of clusters shown in the heatmap.
- Guess the value contained in a circled cell in the chart.
- Click on a cell in the heatmap containing the indicated value.

4.2.2.5. Limits and bad practices known

After enriching themselves with all this information, they were able to establish some observed limits related to the color palettes. Including the use of a red-green palette as an extended practice in bioinformatics, or inconsistencies in the usage of rainbow palettes.

4.2.3. Crowdsourcing platforms

In the previous section, terms and concepts, in 1.1.2., we gave a brief first explanation of the term crowdsourcing, its history and an overview of its advantages.

Now we will talk about crowdsourcing platforms in perception studies and define why we did finally choose Prolific to conduct the experiment.

We based our decision on previous research, for instance [25], where Jeffrey Heer and Michael Bostock successfully replicated prior experiments on proportional judgments of spatial encodings and alpha contrast adjustment of chart grid lines getting similar design guidelines to prior in-lab work. Moreover, in this same study, the identified benefits of crowdsourcing over laboratory experiments. Such as cost reduction, a faster completion time and access to wider populations.

Our main focus was on Amazon Mechanical Turk (MTurk) and Prolific itself. Then, comparing the different platforms, we decided to deploy our study in Prolific.

Most of the studies reviewed, [26, 27, 28],[[all from different disciplines, agree on the absence of significant difference on Prolific and MTurk's data quality.

Participants, however, seem to be more naïve to common experimental research tasks in Prolific platform [29]. As well as offering a more diverse population in terms of geographical location, ethnicity, etc.

In addition, there is a statement from [30], which reads as follows, "While MTurk is not a focus product of Amazon and has not seen much development in recent years, Prolific is

constantly evolving and – if continuously successful – will probably not only expand its reach further but will also continue to implement functionality requested by its users".

Finally, a recent study, this time conducted for online behavioral research in September 2021 [31], clearly shows a lower data quality on Mturk, even with data quality filters.

4.3. Data preparation

This section is mainly focused on the process of data collection and validation. All the coding and charts produced in this section can be found in the folder "data_generation" among the files submitted.

4.3.1. Methods and alternatives

The data collection process took us many approaches to finally achieve clear and representative data for the study.

The process to follow was to first get the data through the corresponding method, and afterwards, to validate the existence of clusters in the data by a cluster detection algorithm. We needed to do this so that later on, we were able to compare the "real" number of clusters marked by the clustering algorithm against the user's study response.

We are going to talk first about the different cluster detection algorithms we are going to work with.

4.3.1.1. Validation requirements

In this section, we will mention the fundamental requirements that must be fulfilled to determine that the data generated and the charts configuration display is ready to be presented in the study.

Any change or evolution in the creation of the data has been jointly evaluated by the researcher, co-director and director.

The basic requirements are:

- The data may be clearly simulating the representation of a bike-sharing case.
- When data contains any clusters they should be shown in a way that is not too obvious nor too complicated, but that a good balance between the two is found. We should also know for sure when this data contains clusters.
- That all 3 of us agree.

As a last requirement, once the data is ready, the charts will be evaluated with users unrelated to the study.

4.3.1.2. Clustering algorithms

Clustering, also referred to as cluster analysis or detection, is the task of grouping sets of data, in a way that the data points in the same group have more similar characteristics to each other than data in the other groups [32].

The concept of similarity between the data points vary from different methodologies: connectivity models, distribution models, centroid models and density models. And there exist many clustering algorithms for all of these methods, you can find an example of the most popular in Figure 13.

Our work is mainly focused on density models, which isolate different density regions and assign the data points within these regions in the same cluster. The clustering algorithms we used are DBSCAN and HDBSCAN, both are density based. The only difference between them is that HDBSCAN calculates a density function for each cluster, while in DBSCAN all clusters have the same density.

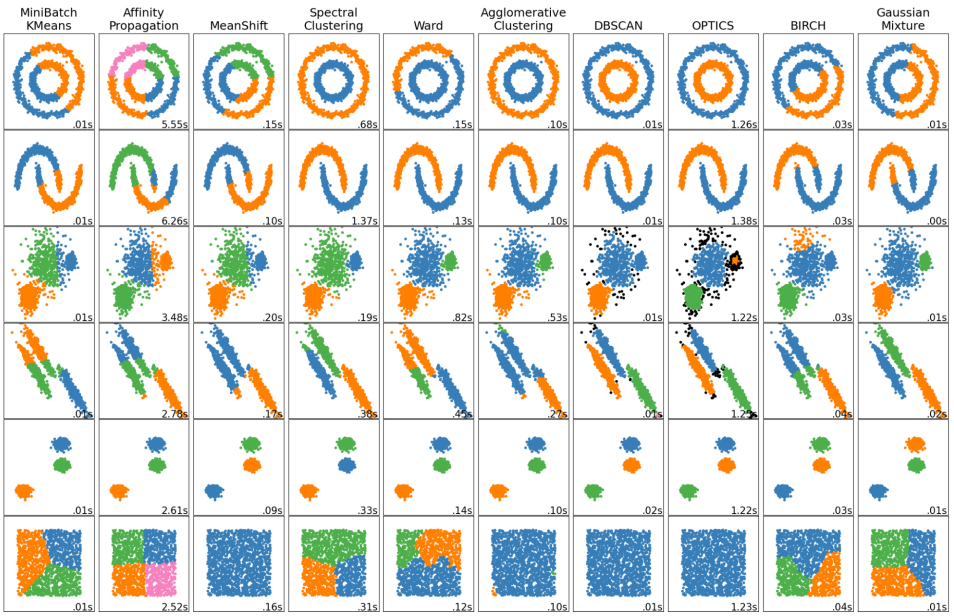


Figure 13. Clustering algorithm comparison.
https://scikit-learn.org/stable/modules/clustering.html

4.3.1.3. Real data

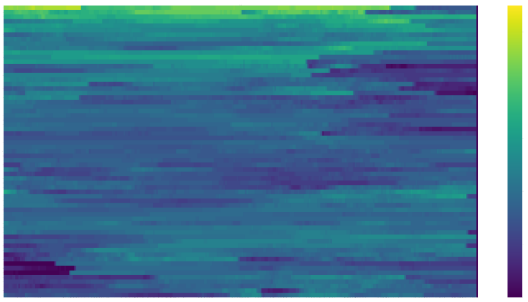
The initial method conducted to obtain the data was to use the actual bike sharing data from Barcelona as it is publicly available data, found on the website of the Ministry of Economic Affairs and Digital Transformation of the Government of Spain, Figure 14.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	station_id	num_bikes_available	num_mechanic_available	num_ebike_available	num_docks_	last_reported	charging_station	status	installed	renting	returning	traffic	last_updated	ttl
2	1	16	14	2	29	1646089166	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
3	2	24	22	2	5	1646089055	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
4	3	8	8	0	18	1646089086	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
5	4	2	2	0	19	1646089118	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
6	5	23	23	0	16	1646089030	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
7	6	10	10	0	28	1646089044	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
8	7	10	10	0	17	1646088994	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
9	8	13	13	0	13	1646088993	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
10	9	2	2	0	25	1646089091	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
11	10	13	13	0	30	1646089088	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
12	11	10	10	0	24	1646088947	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
13	12	6	5	1	19	1646089129	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
14	13	11	11	0	41	1646088944	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
15	14	1	0	1	26	1646089057	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
16	15	11	10	1	13	1646089179	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
17	17	6	6	0	36	1646088961	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
18	18	19	6	13	7	1646088994	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
19	19	29	18	11	1	1646089170	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11
20	20	13	0	13	3	1646089018	TRUE	IN_SERVICE	1	1	1	NA	1646089189	11

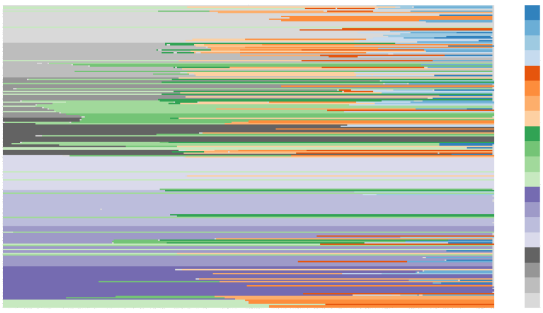
Figure 14. Preview of the Bicing data. Prepared by the author. Data extract from https://datos.gob.es/en/catalogo/l01080193-estado-de-las-estaciones-del-nuevo-bicing-de-la-ciudad-de-barcelona1

Once we had the data, we tried to reorder the key dimensions in order to find patterns (clusters) in the values, Figure 15. Like Tamara Munzner said [23], there are many reordering methods, including matrix permutation, seriation, ordination, biclustering, coclustering and two-mode clustering. However, the variable dimension representing time data, couldn't be reordered.

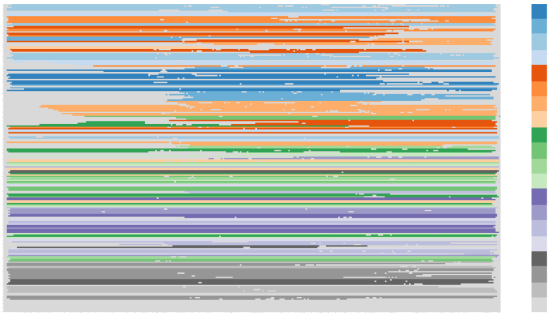
Despite all the experiments conducted, results from the DBSCAN clustering algorithm shown in Figure 15, we couldn't get any profitable results from the clustering analysis, although some patterns were revealed in some charts the results didn't pass the requirements established, so we had to discard this first method to collect the data, and moved to the next one.



Original Data read from the Bicing csv file



DBSCAN, eps = 5, min_samp = 5



HDBSCAN, min_samp =11, min_clus = 200

Figure 15. DBSCAN clustering analysis results on real Bicing data. Prepared by the co-director.

4.3.1.4. Interpolation

Since we could not detect any clusters with the real Bicing data, we came up with the alternative of creating synthetic data. Doing it this way, we also achieved "cleaner" data, which would later help in the heatmap design process and data transformation.

The first approach on generating data synthetically was by interpolation. Interpolation in the mathematics field refers to the estimation of the value of a function of x, f(x), from certain known values of the function, in other words, it is a method of creating new data points based on a set of known data points.

Thus, as you can see in Figure 16, we decided to use 3D interpolation in order to have the attributes of the axes, as well as the value corresponding to each cell that will later encode the corresponding color.

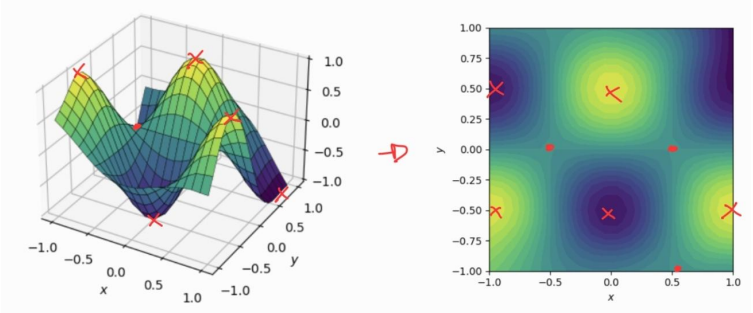
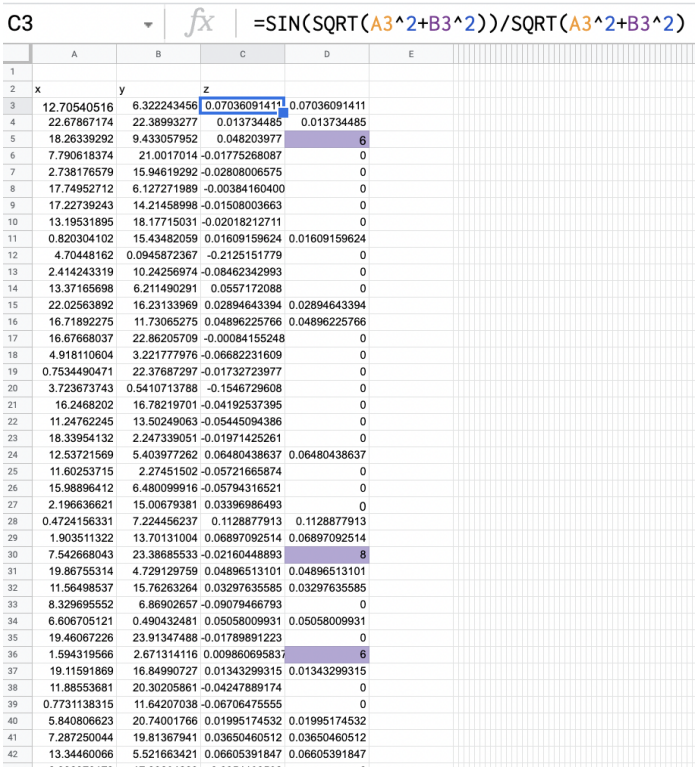


Figure 16. First idea on how to use interpolation to create the data. Images from: https://stackoverflow.com/questions/37872171/how-can-i-perform-two-dimensional-interpolation-using-scipy

To start creating the data, we used the free Google Sheet extension called Calculus Functions. This extension has the function GRIDXYZ, which we used for sampling the

random set from the Mexican hat function $f(x,y)=\frac{\sin\left(\sqrt{(x^2+y^2)}\right)}{\sqrt{(x^2+y^2)}}$. Then we only changed some Z dimension values to get the shape that we want, shown in Figure 17.



C3 | =SIN(SQRT(A3^2+B3^2))/SQRT(A3^2+B3^2)

	A	B	C	D	E
1					
2	x	y	z		
3	12.70540516	6.322243456	0.07036091411	0.07036091411	
4	22.67867174	22.38993277	0.013734485	0.013734485	
5	18.26339292	9.433057952	0.048203977	6	
6	7.790618374	21.0017014	-0.01775268087	0	
7	2.738176579	15.94619292	-0.02808006575	0	
8	17.74952712	6.127271989	-0.00384160400	0	
9	17.22739243	14.21458998	-0.01508003663	0	
10	13.19531895	18.17715031	-0.02018212711	0	
11	0.820304102	15.43482059	0.01609159624	0.01609159624	
12	4.70448162	0.0945872367	-0.2125151779	0	
13	2.414243319	10.24256974	-0.08462342993	0	
14	13.37165698	6.211490291	0.0557172088	0	
15	22.02563892	16.23133969	0.02894643394	0.02894643394	
16	16.71892275	11.73065275	0.04896225766	0.04896225766	
17	16.67668037	22.86205709	-0.00084155248	0	
18	4.918110604	3.221777976	-0.06682231609	0	
19	0.7534490471	22.37687297	-0.01732723977	0	
20	3.723673743	0.5410713788	-0.1546729608	0	
21	16.2468202	16.78219701	-0.04192537395	0	
22	11.24762245	13.50249063	-0.05445094386	0	
23	18.33954132	2.247339051	-0.01971425261	0	
24	12.53721569	5.403977262	0.06480438637	0.06480438637	
25	11.60253715	2.27451502	-0.05721665874	0	
26	15.98896412	6.480099916	-0.05794316521	0	
27	2.196636621	15.00679381	0.03396986493	0	
28	0.4724156331	7.224456237	0.1128877913	0.1128877913	
29	1.903511322	13.70131004	0.06897092514	0.06897092514	
30	7.542668043	23.38685533	-0.02160448893	8	
31	19.86755314	4.729129759	0.04896513101	0.04896513101	
32	11.56498537	15.76263264	0.03297635585	0.03297635585	
33	8.329695552	6.86902657	-0.09079466793	0	
34	6.606705121	0.490432481	0.05058009931	0.05058009931	
35	19.46067226	23.91347488	-0.01789891223	0	
36	1.594319566	2.671314116	0.009860695837	6	
37	19.11591869	16.84990727	0.01343299315	0.01343299315	
38	11.88553681	20.30205861	-0.04247889174	0	
39	0.7731138315	11.64207038	-0.06706475555	0	
40	5.840806623	20.74001766	0.01995174532	0.01995174532	
41	7.287250044	19.81367941	0.03650460512	0.03650460512	
42	13.34460066	5.521663421	0.06605391847	0.06605391847	

Figure 17. Creation of the scattered (x,y,z) data points to later process onto a uniform rectilinear grid based on the natural neighbor algorithm.

Figure 18. Results obtained from the sampling of the interpolation. Prepared by the author.

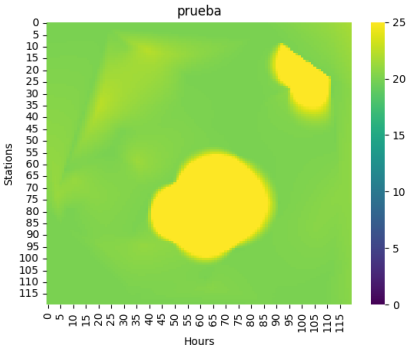


Figure 19. Output obtained from reading the csv file generated with the GRIDXYZ function on Google Sheets. Prepared by the author.

The first attempt on clustering the data using the algorithms was with the original data, read from the csv, shown in Figure 19. The parameters set on the algorithm were adjusted to the needs, in DBSCAN [33], on the other hand for HDBSCAN parameters [34].

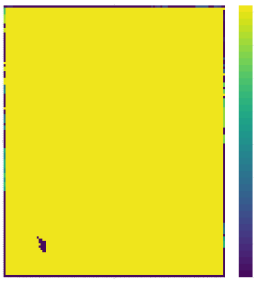
DBSCAN	HDBSCAN
	
Eps = 3 , min_samples = 10	Min_clus_size = 10, cluster_epsilon=3, min_samples=10

Figure 20. Result using the same sample data as the one in Figure 19 after the clustering algorithm DBSCAN and HDBSCAN. Prepared by the author.

After this first try, we decided to scale data, so that all the values are in the same rank, including the dimension values. To do so, we used the method fit_transform(X[, y]) from the class sklearn.preprocessing.MinMaxScaler.

DBSCAN	HDBSCAN
	
Eps = 1, min_samples = 5, MinMaxScaler (0, 1)	Min_clus_size = 10, epsilon=3, min_samples= 5, MinMaxScaler (0, 1)

Figure 21. Result using the same sample data as the one in Figure 19 after the clustering algorithm DBSCAN and HDBSCAN. Prepared by the author.

Although the results were better this time, they didn't pass the validation requirements. Meanwhile, we were trying every possible parameter or other clustering algorithm to get better results, but none of them were conclusive. Thus, we tried a different approach on treating data before the clustering, partitioning the data in addition to the scaling.

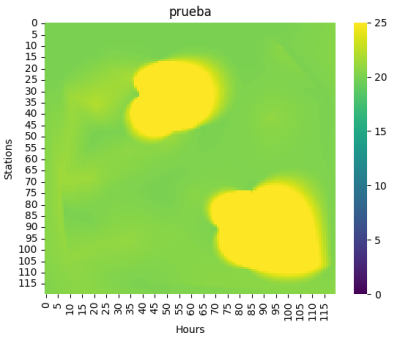


Figure 22. Original read data from csv. Prepared by the author.

The image in Figure 12 shows the original data used for this example. Then, in Figure 23, you can find the different partitions of the data when calling the clustering function. Using HDBSCAN for all the examples with Min_clus_size = 50, epsilon=3, min_samples= 10.

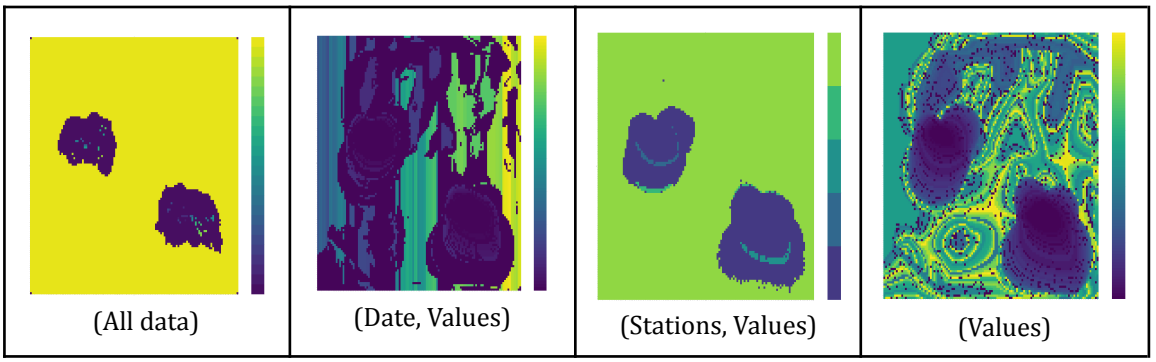


Figure 23. HDBScan output for the different data partitions. Prepared by the author.

The results obtained from the partitioned data (Figure 23) were not detecting the clusters we wanted to mark from the data created by interpolation. Therefore, because of the results obtained and the great amount of time already invested on this method, we decided to discard it and look for another alternative.

4.3.1.5. make_blobs

We conducted this other alternative using the sklearn.datasets module to generate the data, this module is specific for creating many types of datasets. We used the make_blobs samples generator function, which creates isotropic Gaussian blobs for clustering.

The make_blob official user guide is on [35] for further information. The parameters we used for the generation were:

- Input parameters, and meaning:
  - n_samples: (int or array-like), it is the total number of points equally divided among clusters if int, or the amount of points for each cluster if array.
  - n_features: (int), number of features for sample.
  - centers: (int or ndarray of shape (n_centers, n_features))number of centers to generate, or the fixed center locations.
  - cluster_std: (float or float array-like). The standard deviation of the clusters.
- Output / returns:
  - x: (ndarray of shape (n_samples, n_features)) The generated samples.
  - y : (ndarray of shape (n_samples,)) The integer labels for cluster membership of each sample.

You can see in the first chart of Figure 24 the output of the creation with the configuration: `X,_=make_blobs(n_samples=estacions*estacions,centers=clusters,cluster_std=1).` Besides, you can also appreciate how the clustering algorithm worked with this data in the two charts on the right.

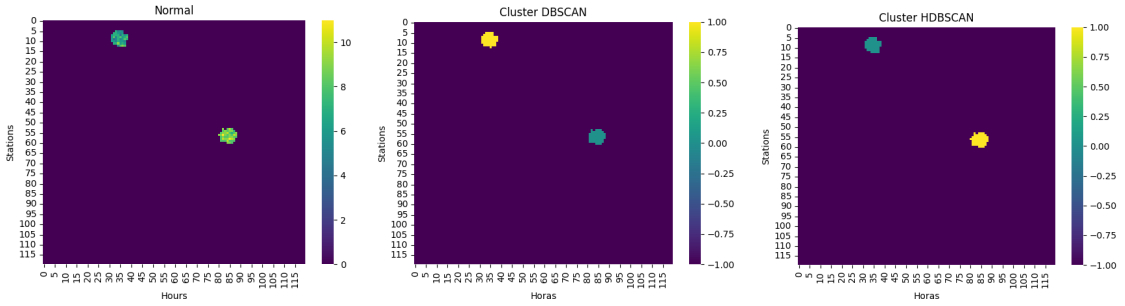


Figure 24. HDBScan output for the different data partitions. Prepared by the author.

This was a big improvement with respect to the other methodologies carried out before. However, we still did not meet the requirement of not clearly displaying the clusters, since we want the task of detecting clusters to be not too difficult, but not too easy either as we mentioned in the section 4.3.1.1. Thus, we proceeded to add some noise to the data.

4.3.2. Adding noise

4.3.2.1. Random

Our first try in generating random data was simply generating random points and joining them with the actual clustered data we had before.

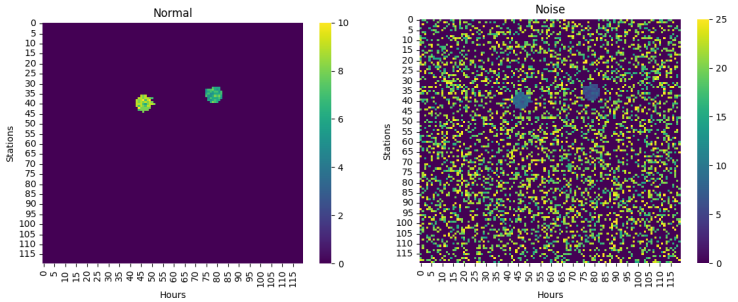


Figure 25. Original data on the left, Original data with random noise on the right. Prepared by the author.

Then, we conducted the clustering algorithms to check if it was still clear for the algorithm to cluster the data.

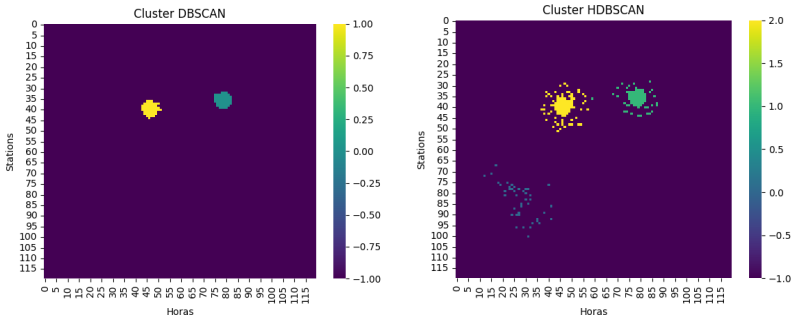


Figure 26. Comparison on the two different clustering algorithms, DBSCAN on the left. Prepared by the author.

As we can see in Figure 26, DBSCAN was more accurate than HDBSCAN on the clustering this time. HDBSCAN detected some noise points as a cluster.

The problem now was that the clusters were not always clearly represented, and although they were detected by the DBSCAN clustering algorithm perfectly for the human eye, it was too difficult to guess. Consequently, we tried to vary the range of noise points and the data so that they were not the same. As you can see in the next Figures.

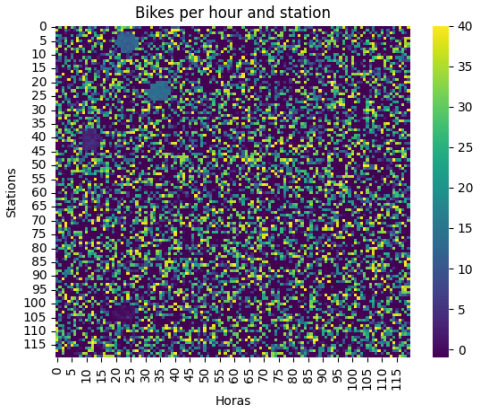


Figure 27. Clusters' values at the extremes from [0-15] and [25-40] and noise at [0-40]

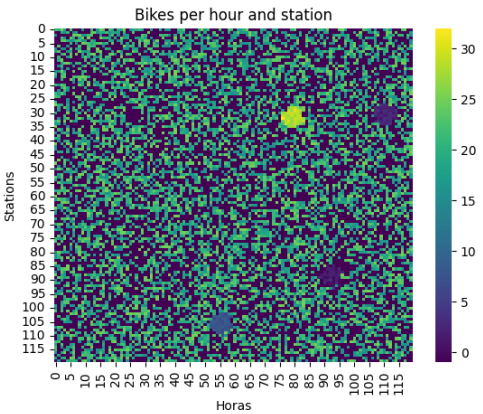


Figure 28. Clusters' values at the extremes from [0-15] and [25-40] and noise at range [15-25]

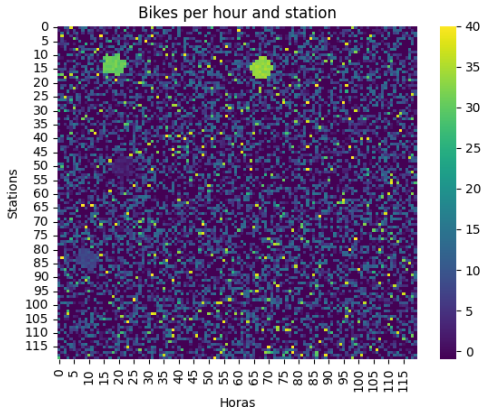


Figure 29. Clusters at the entire range of values[0-40] and noise from [0-15] and [25-40]

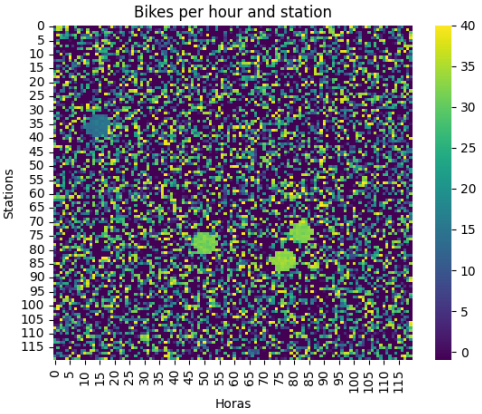


Figure 30. Clusters and noise random for the entire range of values [0-40]

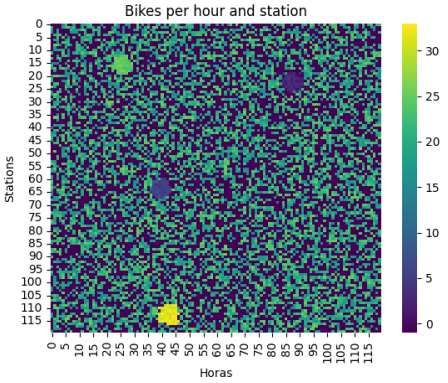


Figure 31. Clusters at the entire range of values[0-40] and noise at range [15-25]

We also worked on the position of the noise data. However, positioning clusters above the noise was making them too clear, and in the opposite case, these were completely hidden, as can be seen in 32.

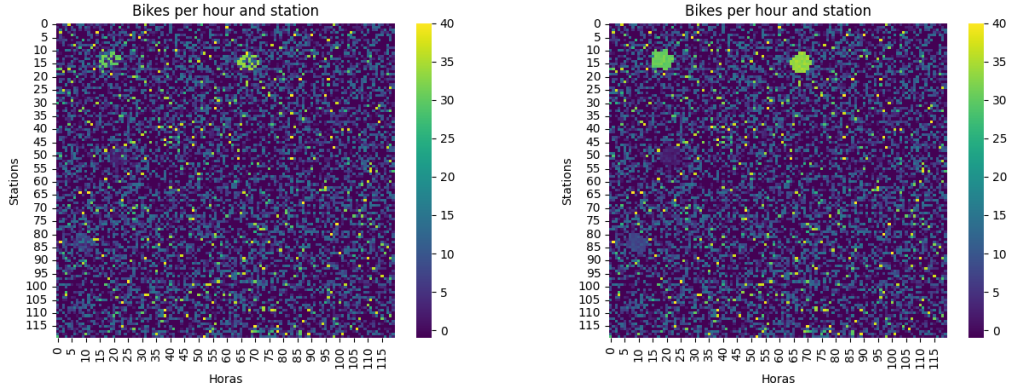


Figure 32. The chart on the left is showing the noise over the clusters, while the one in the right shows the cluster data on top

Unfortunately, none of the above graphs managed to pass the established requirements, as we did not feel that the way they were displayed was correct, there was a very aggressive change between the noise points and the cluesters.

Thus, we tried another workaround, to smooth the data and create the effect wanted, using the Gaussian filter to remove noise and detail from the data, and see if it would be smoother and better displayed. The Gaussian filter can also vary the quantity of blur depending on a parameter called kernel, as it is shown in Figure 33.

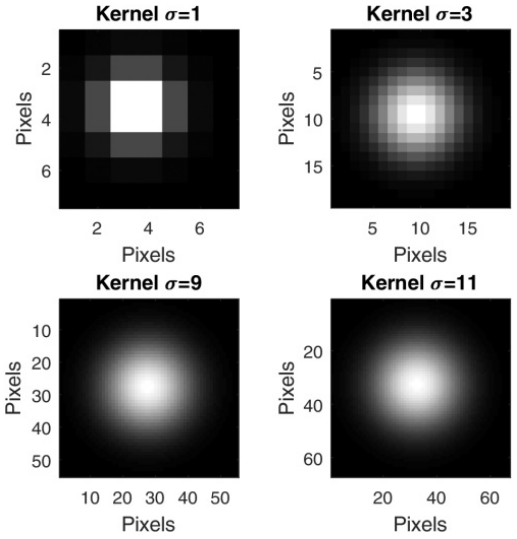
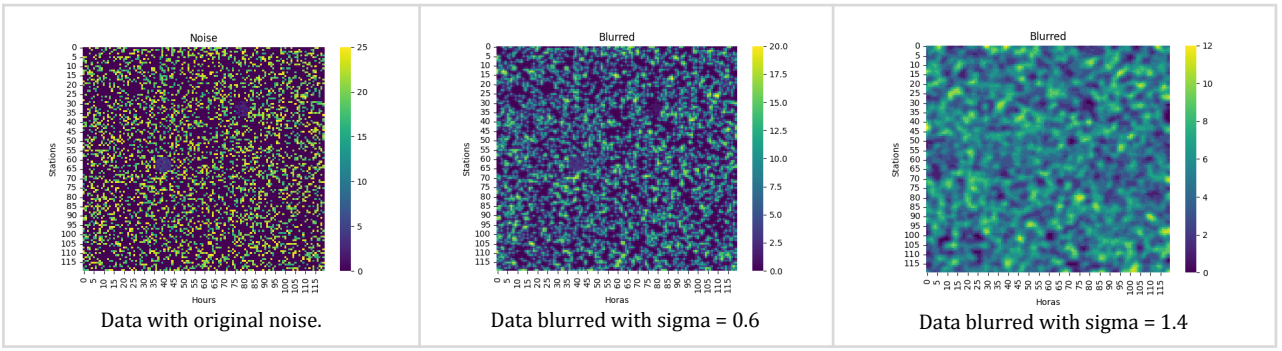


Figure 33. Behavior of the Gaussian filter for different Kernel parameters.
https://www.osapublishing.org/getImage.cfm?img=LmxhcmdlLGFvLTU2LTI4LTgwMTQtZzAwMQ

We used the scipy.ndimage.gaussian_filter to apply the gaussian filter in our data.

```
scipy.ndimage.gaussian_filter(input, sigma, order=0, output=None,
mode='reflect', cval=0.0, truncate=4.0)
```

In order to vary the type of smoothing we played with different values of sigma. In addition, we also worked on smoothing the data horizontally to try to create this effect only in the temporal dimension.



Data with original noise. Data blurred with sigma = 0.6 Data blurred with sigma = 1.4

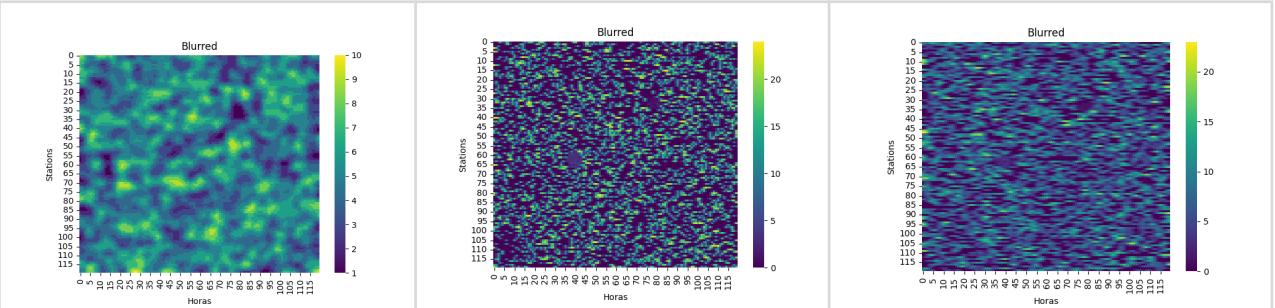


Figure 34. Representation of the data with the different Gaussian filter applied. Prepared by the author.

4.3.2.2. Perlin Noise

Although the random noise results were getting closer and closer to what was expected, one day in a meeting with the director and co-director we decided to try Perlin noise.

The Perlin noise is a type of gradient noise, widely used in computer graphics to increase realism in textures [36]. It is a function that has a pseudo-random appearance and in comparison to the random noise, it offers more natural results providing an ordered sequence of pseudo-random numbers.

In the python noise module we used for the noise implementation, we had to configure some parameters in order to generate it with the shape we wanted [37]:

- **scale**: number that determines at what distance to view the noisemap, it acts like a zoom.
- **octaves**: the number of levels of detail.
- **lacunarity**: amount of detail added or removed at each octave.
- **persistence**: determines how much each octave contributes to the overall shape.

You can see the result obtained using Perlin Noise in Figure 35.

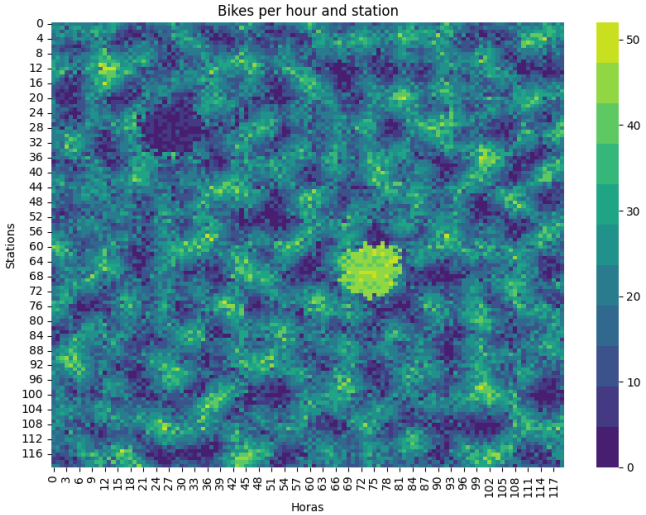


Figure 35. Representation of the data using Perlin Noise. Prepared by the author.

4.3.3. Final data and design configuration

After all the work in developing the data, we finally obtained valid results which met all the requirements. Those were generated with the make_blob function, with a background Perlin noise added.

Therefore, the only thing left was to replicate the charts with all the configurations decided in section 4.2.2. Having:

- Blues as a sequential palette. In both formats continue, and discrete.
- Viridis as a divergent palette. In both formats continue, and discrete.
- 30 cells per dimension.
- Cell size between 10-20px.

In addition, in order to conduct the task of guessing the value in a cell, we needed to somehow mark a cell. As you can see in Figure 36, we used https://colorbrewer2.org/ to fit different colors over our palettes to pick the one that suited the best. Which was red.

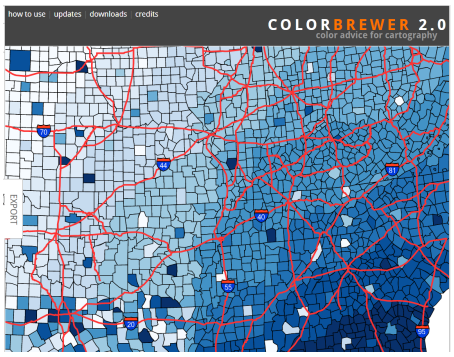


Figure 36. testing the effect and contrast of different colors on our palettes. Prepared by the author.

We also added some transformations to our data, so we could also study if there is any influence on the perception. An example of the final charts are:

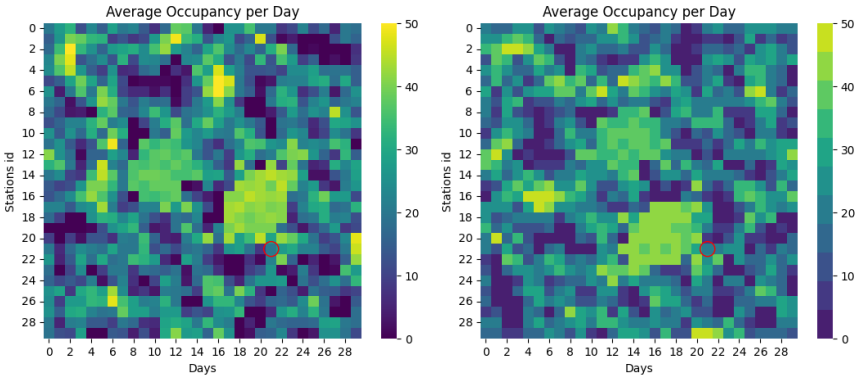


Figure 37. Final charts design, original and transposed. Prepared by the author.

4.4. User Study

This section defines the second part of the study setup, which focuses mainly on the web development of the survey, the database, the host server used and the configurations of the crowdsourcing platform.

You can find screenshots of the final aspect of the survey in the Annex {_}. In addition, you can also access the survey web we used to collect data from our contacts, directly from: http://middeline.pythonanywhere.com

4.4.1. Web development

This section will focus on the development of the website, where the perception study is hosted. We will briefly discuss each of the programming languages and platforms required for the implementation. All the codes and archives can be found in the folder "survey", among the files submitted.

As has already been mentioned, the study has been published in Prolific. This platform offers a number of advantages over its competitors, as explained in section 4.2.3. However, Prolific helps you recruit participants for your studies that are hosted elsewhere, meaning it does not have its own software to develop the survey. Nonetheless, the platform allows the integration of external software to create the surveys such as Qualtrics, SurveyMonkey, Gorilla, Typeform, Google Forms, PsychoPy, Labvanced and more, or any other software using just a link.

These recommended survey integration tools are mostly specialized in psychology, so, using them would have limited us in our experiment development in one way or another.

Thus, we decided to develop a web survey to have more versatility when designing the study. We chose Flask as the web framework for the survey, so it could be developed in Python, a programming language in which the researcher feels comfortable. Besides, we also had the support of the director and his colleague, Jose, who already had some experience in Flask and sample code from the previous bar chart perception study [38].

4.4.1.1. Flask

Flask is a micro web framework for building websites using Python [39] . It is characterized by its simple and scalable core, which also leads to the microframework reference and not just framework. There are lots of frameworks for Python, such as Flask, Tornado, Django...

Flask is based on the Werkzeg WSGI toolkit [40], which implements requests, response objects, and utility functions. And Jinja2 template engine [41], which combines data source with a template and renders a dynamic web page, used in the project to pass Python variables into the HTML templates.

4.4.1.1.1. Folder structure

The structure we followed on our Flask app is:

```
|____persistent_info_users.db
|____heatmap_survey.py
|____static
| |____instructions
| | |____(heatmap charts for the introductory survey in .png)
| |____.DS_Store
| |____css
| | |____(bootstrap css files)
| |____js
| | |____(bootstrap javascript files)
| |____task
| | |____(heatmaps in .png format for the tasks)
| |____data
| | |____ (data csv files to get coords value)
| |____example_cluster.png
| |____example.png
|____proves
| |____(testing files)
|____templates
| |____index.html
| |____evalForm.html
| |____labelEndFailure.html
| |____labelTask.html
| |____demoForm.html
| |____labelInstructions.html
| |____instructions.html
| |____task.html
| |____labelEndSuccess.html
|____respostes.db
```

This follows the diagram of the typical folder structure. With two folders static and templates plus an extra for some testing called “proves”. The static folder contains assets used on the HTML templates, like CSS and JavaScript files, or the Heatmap images. On the other hand, the templates' folder contains only the HTML templates we may want to render.

By default, Flask does not include a database abstraction layer or form validation, instead, Flask works with many extensions to add such functionality to its application. Therefore, we have worked with flask_sqlalchemy, flask_wtf to provide database integration and form validation respectively.

4.4.1.1.2. Programming languages used

Many concepts and programming languages have been used for the development of the web survey. It has been really useful to check these resources [42], [43], [44], [45]. In addition, we used Bootstrap, a popular HTML, CSS, and JavaScript framework for developing responsive, mobile-first websites [46], to design and customize our web easily.

We will now take a brief look at some pieces of code of each programming language:

Python

Python is the programming language that has been used the most in the project, since it has been used both in the development of the web and in the creation and configuration of the data.

```
class taskChart: #Charts used in tasks
    def __init__(self, chart_marked, chart, data, A_numClus, B_valMark, C_valClick):
        self.chart_marked = chart_marked
        self.chart = chart
        self.data = data
        self.A_numClus = A_numClus
        self.B_valMark = B_valMark
        self.C_valClick = C_valClick

charts_task =
[taskChart('0heatmap1_Blues_cont_marker_20', '0heatmap1_Blues_cont', '0data1', 0, 20, random.randint(0, 50)),
 taskChart('0heatmap9_Blues_disc_marker_0', '0heatmap9_Blues_disc', '0data9', 0, 0, random.randint(0, 50)),
 ...
]
```

We created 2 classes in the main code to define the tasks, including their file names, the data file name, the number of clusters existing in the data, the value of the marked cell, and finally the value asked to click on the third task.

We also used python to control the flow of the website. There exist two types of linker functions using Flask.

- redirect function

```
@app.route('/startTask/<int:aux>', methods=['POST','GET'])
def startTask(aux):
    DICT_USERS = shelve.open('persistent_info_users')
    user = request.args.get('u', '')
    # Begin to compute time just before starting the first chart
    DICT_USERS = modifyDictUsers(DICT_USERS, str(user), 2, dt.datetime.now())
    #DICT_USERS[str(user)][2] = dt.datetime.now()
    DICT_USERS.close()
    return redirect('/task/0?u='+str(user))
```

With redirect(url_for()) we are usually sending the control to a method that processes the logic. We usually redirect to another endpoint to initialize variables, store data in the database, or validate answers from a form.

- render_template function

```python
@app.route('/task/<int:question_id>', methods=['POST','GET'])
def task(question_id):
    DICT_USERS = shelve.open('persistent_info_users')
    user = request.args.get('u', '')

    if question_id == NUM_CHARTS_TASK:
        DICT_USERS.close()
        return redirect('/evalForm/0?u='+str(user))
        #Has already finished, so question is number 35

    else:
        if DICT_USERS[str(user)][4]:
            DICT_USERS = modifyDictUsers(DICT_USERS, str(user), 4, False)
            DICT_USERS[str(user)][4] = False

        img_name = DICT_USERS[str(user)][1][question_id].chart #grabs correct Heatmap chart
        img_marked = DICT_USERS[str(user)][1][question_id].chart_marked #Heatmap chart with marker
        C_valClick = DICT_USERS[str(user)][1][question_id].C_valClick #Value to click

        DICT_USERS.close()
        progress = int((float(question_id)/float(NUM_CHARTS_TASK))*100.0)
        return render_template("task.html", img_name=img_name,img_marked=img_marked, question_id=question_id,
progress=progress, C_valClick=C_valClick, user=user)
```

We use render_template when we have already done all the logic, and we must send this data to a template. For example, in the code above, we are checking if the question ID belongs to the last question, meaning the user is done with the tasks and has to perform the evaluation survey. On the other hand, we prepare the data for the next task and then render the task template again with the updated parameters.

HTML

HTML stands for Hyper Text Markup Language, we use it to define the structure of the web pages.

```html
<!doctype html>
<html lang="en">
  <head>
    <title> Heatmaps User Study</title>
    <!-- Required meta tags -->
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no">
  </head>
    <body>
      <div align="center">
          <tr>
            <td colspan="4" height="280">
              <p class="h6 text-center">
                You will now start with the User Study Survey. <br>
                The tasks will be the same as the Introductory Survey ones.
                However, you won't have any option, it will be open values tasks.
               </p>
            </td>
          </tr>
          <div align="center">
            <form method="POST" action="/startTask/0?u={{ user }}">
             <button type="submit" name="Start-button" class="btn btn-primary">Start</button>
            </form>
           </div>
      </div>
    </body>
</html>
```

As you can see in this example. We define a series of elements, which tells the browser how to display the content using the labels, such as the ones in the code above: <head> → where the heading is, <body> → this is the body from the web page, the <p> tag marking it is a paragraph

or even the form tag used to redirect the user to another point.

CSS

CSS is used to describe a web page's appearance. We are mostly using Bootstrap Grid v4.0.0-beta.2 (https://getbootstrap.com). A clear example is the implementation of the CSS code below to develop the functionality to hide the marker on the chart.

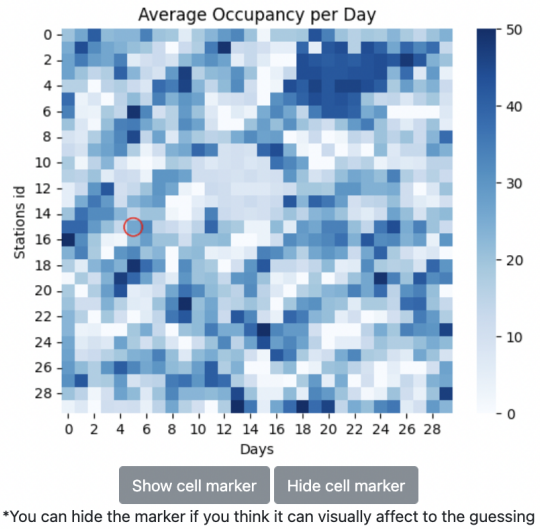


Figure 38. Example of one of the charts shown in the tasks. Where you can appreciate the added functionality to hide the red marker. Prepared by the author.

```
<style>
       .oculto {display:none;}
</style>
```

This is an example of CSS code we used to hide the chart image with the marker when a user clicked a button. This is showing the new image without the marker with the same size, making it seem that the only thing disappearing is the marker, when in fact the whole image is being changed.

Javascript

We also use some basic Javascript functions from Boostrap [46].

```
<script type="text/javascript
        function preventBack() { window.history.forward(); }
        setTimeout("preventBack()", 0);
        window.onunload = function () { null };
</script>
```

This script prevented the user from going back and checking for example the last question's answer, or answer the same question two times.

In addition, we also used Javascript for the development of getting the click coordinates from the chart's image to then read those in the python route and calculate the corresponding value from the graph at those given coordinates. This code can be found in the Annex {_}.

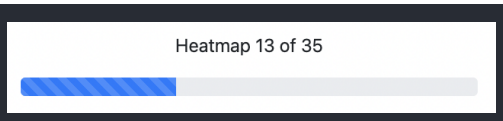
Jinja

We used Jinja to pass variables from the Python code to HTML files.

```
{% if error %}
       <div class="alert alert-danger alert-dismissible fade show" role="alert">
         <strong>Error!</strong> {{ error }}
         <button type="button" class="close" data-dismiss="alert" aria-label="Close">
            <span aria-hidden="true">&times;</span>
         </button>
       </div>
{% endif %}
```

Once the validations are conducted,we send the message to be displayed on the web as an alert. Iif there is any error. Those could be for example: not accepting the terms, leaving blank a task answer or clicking on the wrong coordinates over the image (not pointing a cell).

Another example is the bar showing the process as it progresses.




```
<div class="progress-bar progress-bar-striped"
role="progressbar" style="width: {{progress}}%"
aria-valuenow="{{progress}}" aria-valuemin="0"
aria-valuemax="100"></div>
```

This script prevented the user from going back and checking for example the last question's answer, or answer the same question two times.

4.4.1.1.3. Databases

We used the SQLAlchemy Python library, to get access to a relational database, getting indeed the full power and flexibility of SQL. We configured SQLAlchemy by setting a database URI, in our case, we used sqlite3 instead of the famous ones PostgreSQL, Oracle… SQLite is a C library that provides a disk-based database, which doesn't require any separate server process.

In addition, we also created a "shelf", which is a persistent, dictionary-like object in order to save users info, like the ID, the correct random order assigned for the tasks, or the start time.

```
app.config['SQLALCHEMY_DATABASE_URI'] = 'sqlite:///respostes.db'
app.config['SQLALCHEMY_TRACK_MODIFICATIONS'] = False
db = SQLAlchemy(app)
#In addition
DICT_USERS = shelve.open('persistent_info_users')
```

This code implements the configuration of the SQLAlchemy by setting a database URI to our sqlite path, disables tracking, and creates a database object using the SQLAlchemy class.

Once the databases were created and configured, we created the database table for the answers data, called "AnswersData".

```
class AnswersData(db.Model):
    id                  = db.Column(db.Integer, primary_key=True)
    date                = db.Column(db.String(32))      # '2017-05-10'
    time                = db.Column(db.String(32))      # '08:12:47.292000'
```

```
    user               = db.Column(db.Integer)        # User id
    task               = db.Column(db.String(255))    # Task being performed
    age                = db.Column(db.Integer)        # Age
    gender             = db.Column(db.String(255))    # Gender
    education          = db.Column(db.String(255))    # Education level
    eyesight           = db.Column(db.String(255))    # Eyesight
    display            = db.Column(db.String(32))     # Inches of the display
    understood         = db.Column(db.String(32))     # Understanding
    easiness           = db.Column(db.String(32))     # Easiness of the tasks
    satisfied          = db.Column(db.String(32))     # Satisfaction on the time spent
    comments           = db.Column(db.String(1000))   # Additional comments
    chart              = db.Column(db.String(255))    # Image shown
    chart_marked       = db.Column(db.String(255))    # Image shown
    A_num_clust        = db.Column(db.Integer)        # Task A: Number of Clusters in the Chart
    A_ANS_num_clus     = db.Column(db.Integer)        # Task A: number of Clusters ANSWERED
    B_value_marker     = db.Column(db.Integer)        # B: Value of the marked cell
    B_ANS_value_marker = db.Column(db.Integer)        # B: Value ANSWERED
    C_value_to_click   = db.Column(db.Integer)        # C: Value to pinpoint "click"
    C_ANS_value_clicked = db.Column(db.Integer)       # C: Value to pinpoint "click"
    error_A            = db.Column(db.Integer)        # Error A (Real Clusters - Quantity answered)
    error_B            = db.Column(db.Integer)        # Error B (Real Cell Value - Value answered)
    error_C            = db.Column(db.Integer)        # Error C (Real Value asked - Cell's Value clicked)
    time_spent         = db.Column(db.String(32))     # Time spent to answer
```

Creation of the AnswersData model, which inherits from the db.Model class. We used db.Column class to define the columns for our table, passing as first argument the column type.

Finally, we started working with the table, the basic functions used are:

db.drop_all() -> to delete all existing database tables
db.create_all() -> to create the tables that are associated
db.session.add(*something) -> to add something to the database
db.session.commit() -> to commit the transaction and apply the change to database

4.4.1.2. Python Anywhere

Flask comes with a simple built-in server to run the Flask application, which can be then reached on http://127.0.0.1:5000/. This option was good enough for testing while developing. However, it is not scalable for the production website.

After checking all the options to deploy the app on the official Flask documentation web [39], and doing a search on the internet about the most used, best rated, and also easy to use, we decided to use PythonAnywhere, which gives a fully configured Python environment. Offering:

- Coding Python applications in their web-based editor.
- Save and run the code, validate its output.
- Host the code.

PythonAnywhere has different pricing plans. Including a free hosting plan which we were available to use. This one already offers:

Plan	Beginner: $0/month
Application domain	**your-username.pythonanywhere.com**
Web workers per web app	1
Python, Bash, MySQL consoles	Up to 2
SSH access to your account via ssh.pythonanywhere.com	NO
Access to external internet sites from your code e.g. urllib or wget	Specific sites via HTTP(S) only
Share your consoles with other people	YES
Scheduled tasks Like cron, but simpler	1 daily task
Always-on tasks Tasks that are always running	NO
CPU allowance more info	100 seconds
Bandwidth more info	Low
Private file storage	512MB

Table 12. Summary of the free hosting plan of PythonAnywhere. Prepared by the author with https://www.pythonanywhere.com/pricing/

After all these configurations, we could already find our survey at: middeline.pythonanywhere.com.

4.4.1.3. Prolific

Finally, the last step to get the study up and running was to set up Prolific. You can directly go to Figures 41 to 45 to see all the choices made, or go step by step on the details of each section.

- Study details (Figure 41): where we defined the title of the study, a description of the tasks and the purposes of the study, and a first restriction of the devices and additional software requirements.
- Study link (Figure 42): we had to define the link to the web survey currently hosted on PythonAnywhere, so they could redirect the participants when they sign up.

  In addition, we also needed to mark how we were going to record the participant's Prolific ID, so we can distinguish who has submitted which individual survey response, and afterwards determine through the quality of the answers if we approved or reject the user's submission.

  We decided to redirect the users by URL parameters, as it was the recommended option. To do it like this, we decided to host a new website in PythonAnywhere with

a new domain, to gather the user ID from the URL parameter as you can see in Figure 39, and another one which assigned random user number to send it to family, friends and acquaintances as shown in Figure 40.

```
@app.route('/')
def homepage():
    db.create_all()  # make our sqlalchemy tables
    DICT_USERS = shelve.open('persistent_info_users')
    # We get the user id from the URL parameters
    user = request.args.get('PROLIFIC_PID')
    return redirect('/index/0?u='+str(user));
```

Figure 39. Code to get Prolific ID as the user using URL parameters. Prepared by the author for the Prolific study.

```
@app.route('/')
def homepage():
    db.create_all()  # make our sqlalchemy tables
    DICT_USERS = shelve.open('persistent_info_users')
    # We create the user id and randomize the images for both tasks
    user = random.randint(0,100000000)
    # Check that the user does not exist
    while DICT_USERS.__contains__(str(user)):
        user = random.randint(0,100000000)
    return redirect('/index/0?u='+str(user));
```

Figure 40. Code to generate a random user number for each participant. Prepared by the author for the web survey we sent to family, friends and acquaintances.

- Study completion (Figure 43): when the users have finished the experiment, they must check in one way or another the completeness, so Prolific can afterwards pay if the submission is accepted. In this case, as we have already separated the two studies, the one that collects prolific users will also redirect them back to Prolific with the completion code.
- Audience (Figure 44): in this section, we were able to choose the number of participants we wanted to recruit for the study. Then we could also prescreen the participants on their demographic information including geographic, language, work, education, health and more. Besides, we also had the option to balance the participant on gender.
- Study cost (Figure 45): where we provided an estimated completion time for the entire survey. To get this estimated value as close to reality as possible, we sent it to 2 possible users and asked them to answer the survey and also time themselves while they were doing it. Once we had those times we didn't choose the average , but we took the greater of the two, which was 19 minutes, rounded up to 20. Moreover, we had to establish the payment we wanted to offer. Prolific has a minimum hourly rate amount of at least £5.00 / $6.50 or €6 per hour.

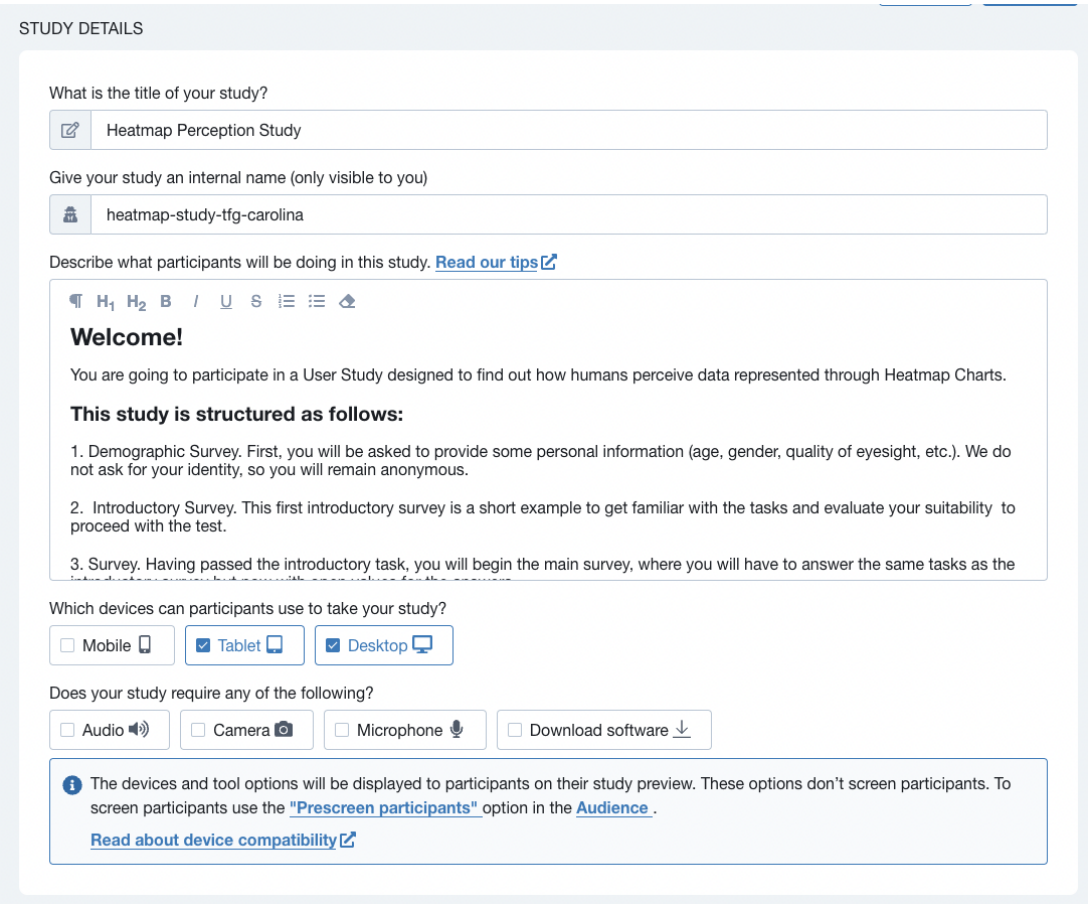


Figure 41. Study details defined on Prolific. Prepared by the author.

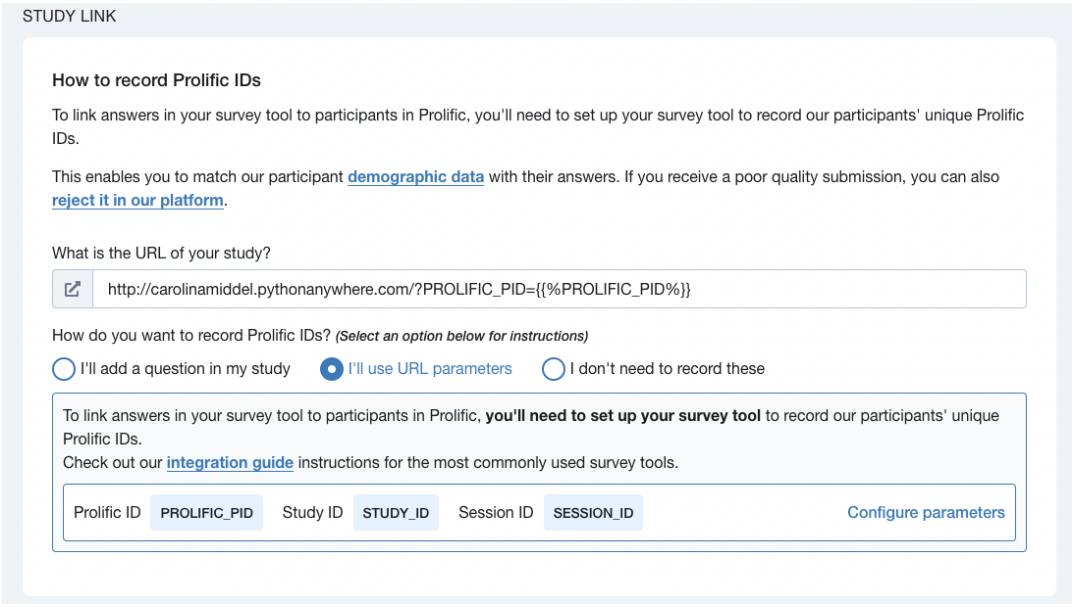


Figure 42. Study link definition on Prolific, where we selected to get the users Prolific IDs by URL parameters. Prepared by the author.

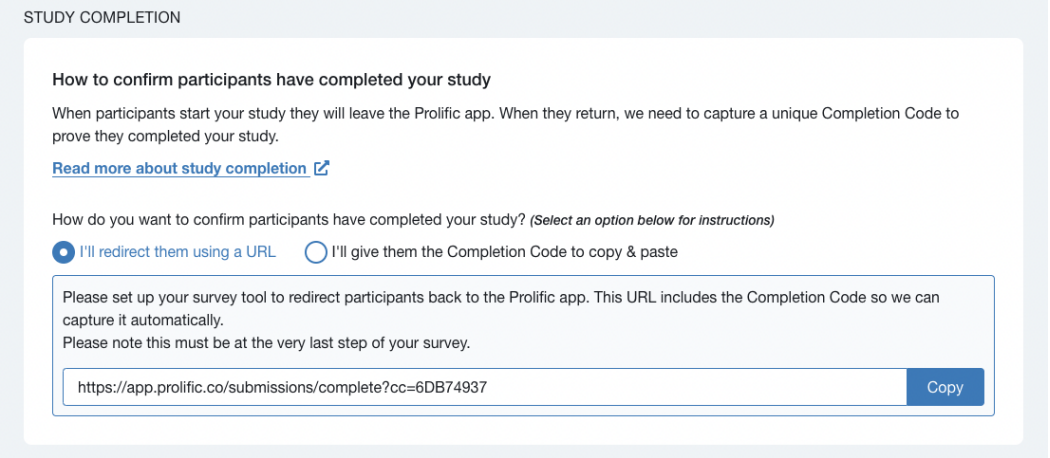


Figure 43. Study completion definition on Prolific. Prepared by the author.

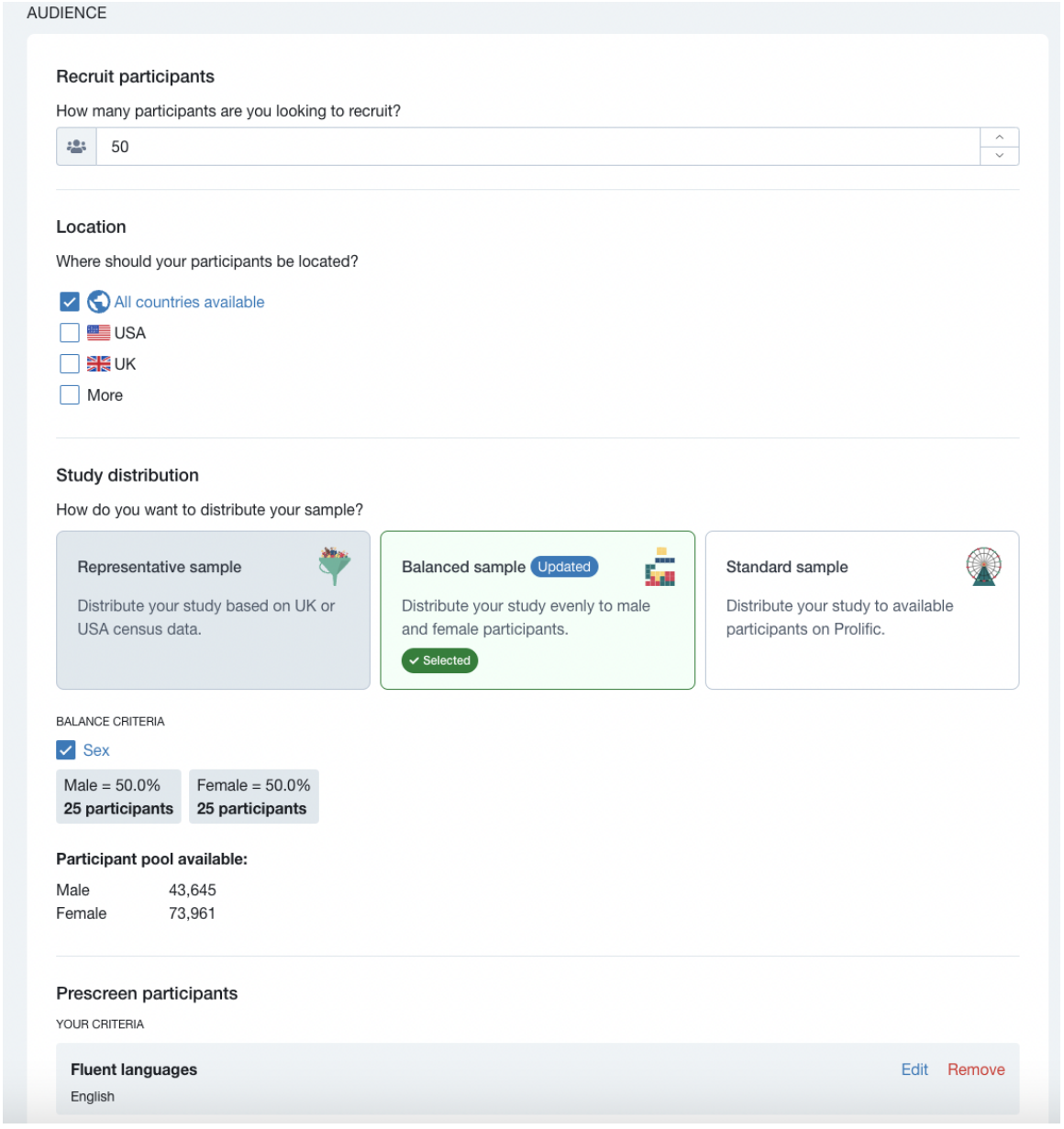


Figure 44. Study Audience definition on Prolific. We decided to first select 50 participants to our study, then we could always increase the number of participants. Also selected users from all around

the world, but equally distributed between men and women. They must know English to participate. Prepared by the author.

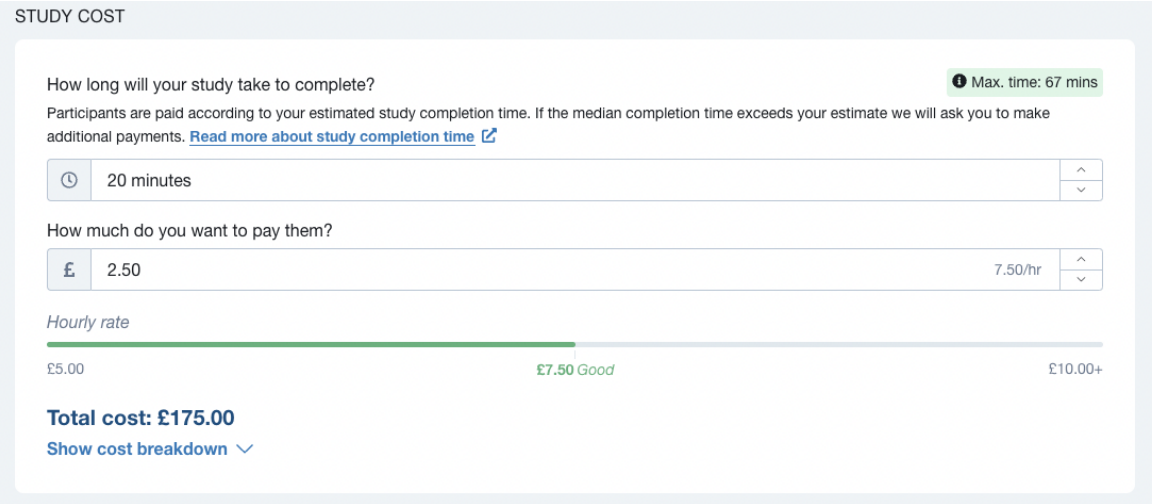


Figure 45. Study costs definition on Prolific. With the study experiment we conducted we got an estimated duration of 20 minutes, then we marked we wanted to pay them £2.5 for the completion of the study which has been conclusive on an hourly rate of £7.5. Prepared by the author.

4.4.2. Results

In this section and the following ones, we are going to refer to the tasks as:

- Task A, estimation of the number of clusters. Which its error is going to be, the difference of the real number of clusters and the number indicated, referred as error A.
- Task B, estimation of the circled cell's value. Its error is the difference between the real value of the corresponding cell and the value indicated by the user, called error B.
- Task C, the task of clicking on a cell in the heatmap containing the indicated value. Its error is the difference between the value indicated and the value corresponding to the clicked cell, as error C.

Moreover, we are also going to refer to the different heatmaps configuration as:

- Viridis: associated with the heatmaps represented with the divergent palette Viridis.
- Blues: associated with the heatmaps represented with the sequential palette Blues.
- Disc: the heatmap values are configured as a discrete set of colors.
- Cont: heatmap values as a continuous function between the numeric value and color.

4.4.2.1. Hypothesis

The purpose of this project is to assess heatmap charts' perception in order to communicate guidelines for an effective and expressive heatmap design. With the goal of making it easier to understand and unambiguous.

We are going to analyze these concepts evaluating the effect of the different heatmap configurations on the tasks errors. While working on the study and the design configurations, some hypothesis were formulated:

- H1: Discrete palette helps on the tasks of detecting the number clusters.
- H2: Diverging palettes communicates the cell's values better.
- H3: Guessing the number of clusters increases difficulty when there are more clusters.

4.4.2.2. Answers Validation and Data cleansing

The study has been provided to the users in two different ways. We first sent the study to our family and friends, where we reached 53 different users. Then, we also posted the study in the crowdsourcing platform, Prolific, where we reached 67 users.

Each of these studies consists in:

1. Demographic Survey.
2. Introductory Survey.
3. Main Survey.
4. Personal Evaluation Survey.

To validate a user's submission we have determined that the user has had to finish the full study, including the personal evaluation. And the average error of the tasks answers must be:

- Less than 2 for task A.
- Less than 20 for task B and C.

We conducted this filtering through the sql query shown in Figure [46]. After this cleaning, we were left with 38 and 47 users respectively. With the characteristics shown in Figure [47].

```sql
select  USER
from answers_data
where user in
        (select user
        from answers_data
        where task = 'END')
group by user
having
        avg(abs(error_A)) < 2 and
        avg(abs(error_B)) < 20 and
        avg(abs(error_C)) < 20
```

	user
1	55dc405950a1f7000590d658
2	5c2a4cf4f4c9ee00010230e9
3	5ea1dd85692b7c0c73166152
4	5f218ad15490300a85fa19b5
5	5f37e777f9794005a85d7e8e
6	607d8f67c90c228dadbeb3ac
7	60e0ba60d52d208f64ad7cab
8	60e0e420bf302d544c690574
9	6119a55edf346b8f1ec6241e
10	6129ad268b36125cd6cca6b9
11	6153064a78cb05a613f1888a
12	61669b71c26a000b62054637
13	616b139e65144bdd0e61f7ec
14	616f74e919e910895fd75631
15	617018432d3fad3003de4824
16	61728b3e4e680d6c607f514c
17	NOPID290
18	NOPID3385
19	NOPID6295
20	NOPID722

Figure 46. SQL Query to get data meeting the established requirements.

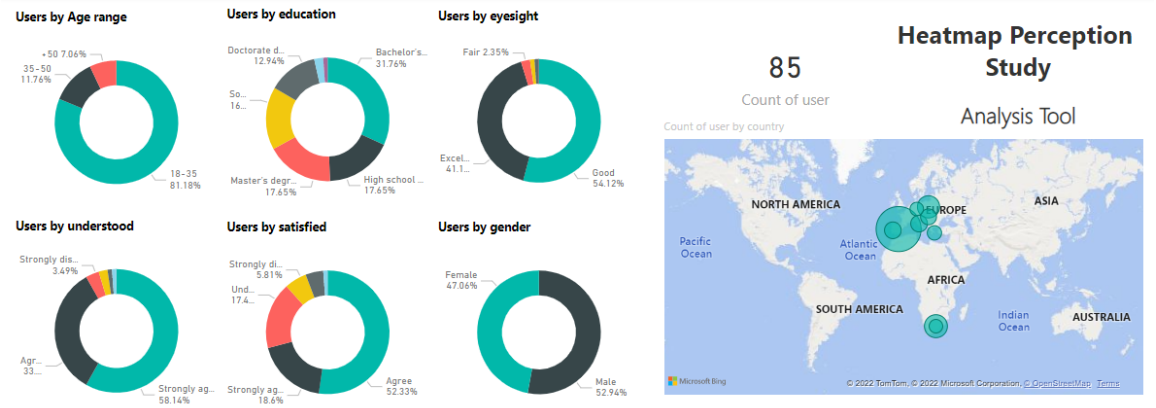


Figure 47. SQL Query to get data meeting the established requirements.

We also had to clean some more data, with typos. Like entering 33 clusters. This could have been just deleted, as a possible case of random answer produced by a lack of interest or a misunderstandment of the task. However, we conducted a study on this user's answers, in Figure [48], and we assumed it was a typo and cleaned it on the data manually changed to 3 clusters.

user	task	chart	chart_marked	error_A
4009750	task	2heatmap1_Blues_disc	2heatmap1_Blues_disc_marker_34	33

AVG error A	AVG error B	AVG error C
0.05555555556	2.648648649	2.783783784
0.1049382716	2.157779401	2.252739226

AVG error A	AVG all	
0.05555555556	0.2258064516	Average
0.1049382716	0.3944062333	Average of the magnitudes of deviations of data from a dataset's mean

Figure 48. The first values show the average error conducted by the user in all 3 tasks. Then, at the bottom you can find a comparison of the user's mean and the overall mean Error conducted for Task A.

4.4.2.3. Statistical results4.4.2.3. Statistical results

First of all, we conducted a study on the empirical probability distribution of the estimation error for the different tasks and heatmap configurations figuring on the study. The results presented in Figure 49, 50, 51, present a normal distribution of the estimation error.

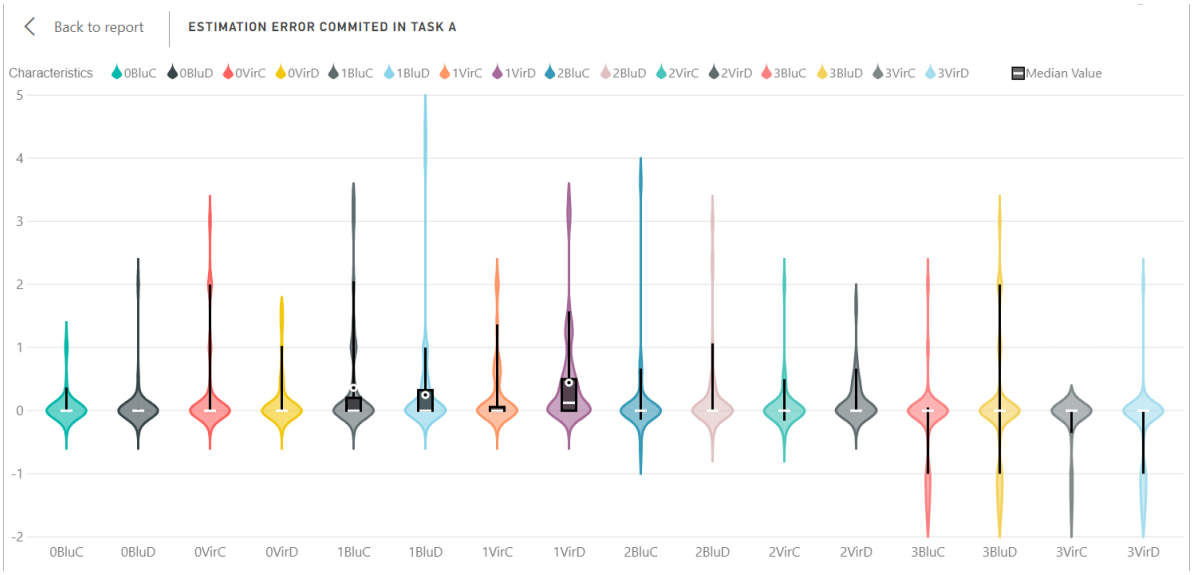


Figure 49. Estimation error committed in task A for all the answers given in the study for each heatmap configuration.

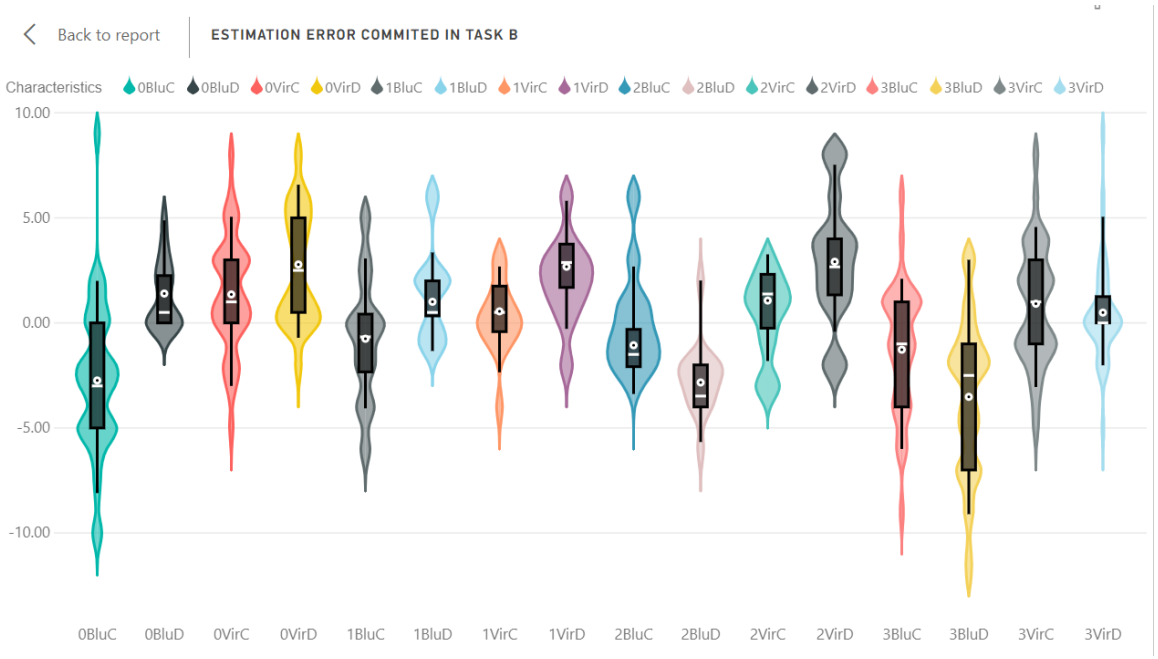


Figure 50. Estimation error committed in task B for all the answers given in the study for each heatmap configuration.

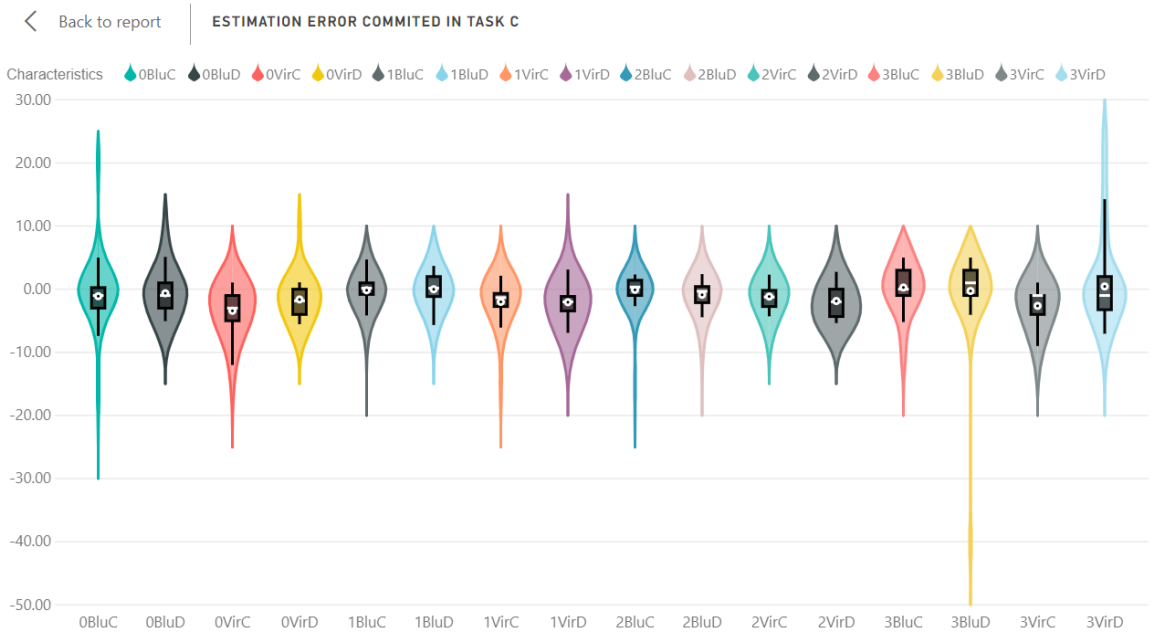


Figure 51. Estimation error committed in task B for all the answers given in the study for each heatmap configuration.

We evaluated the statistics results of the study, through the analysis of the accuracy of the data by the absolute error of the estimated values for the different tasks conducted. We used One-Way analysis of variance, known as ANOVA to study the existence of any significant difference between the means of the subject studied. We used a significance level of $\alpha = 0.05$, besides, we also studied post-hoc Bonferroni's pairwise test with the same significance level ($\alpha = 0.05$).

- H1: Continuous palettes help on the tasks of detecting the number clusters.

For this first hypothesis we checked the ANOVA tests result. The accuracy of the answers was compared using the statistical analysis stated before. Which revealed the existence of significant difference between the means of their absolute errors with a p-value of 0.013113413 ($p < 0.05$). With means (CONT = 0.240128068, DISC =0.34791889). However, the post-hoc analysis with the new Bonferroni corrected alpha level: 0.025, p-value = 0.102893 determined there was not significant improvement in between the two palettes.

As a consequence, we can conclude that there is no significant difference between the estimation of the number of clusters when the palette was discrete.

- H2: Diverging palettes communicates the cell's values better.

The results of the ANOVA tests for the tasks B and C, the ones where the user has to estimate the cell's value. The test has resulted with a p-value greater than 0.05, which led us to reject the null hypothesis, meaning there is a non-significant difference of the means of the errors between the different palettes.

We only compared palettes that are continuous since the user can mark the corresponding value within the range exactly, in the discrete palette you would have to look through the color ranges of the discrete palette.

- H3: Guessing the number of clusters increases difficulty when there are more clusters.

The results of the ANOVA tests showed non significant difference for the means in the absolute error of the charts displaying a different number of clusters. The test has resulted with a p-value of 0.1360122 as shown in Figure 52, which is greater than 0.05. This led us to reject the null hypothesis.

ALL CHARTS
Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0 CLUS	438	115	0.262557	0.6517298
1 CLUS	1126	395	0.350799	0.8572749
2 CLUS	1124	298	0.265125	1.368649
3 CLUS	345	118	0.342029	0.4873273

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	5.410858	3	1.803619	1.849491	0.1360122	2.607842
Within Groups	2953.874	3029	0.975198			
Total	2959.285	3032				

Figure 52. Anova analysis of the clusters

4.4.2.4. Further analysis

There are many different approaches on analyzing the data of a perceptual study. In addition to the ones studied on the hypothesis, we have studied without any significant difference the time spent by the users conducting the task A related to the palette.

Furthermore, we also studied through the plot in Figure 53 the average time spent by the number of cluster in the charts, seeing a great difference between them.

Figure 53. Average time spent by number of clusters in the data.

Another interesting insight we got evaluating the time spent was that the 2 outliers shown in Figure 54, with greater time spent and greater error in task A, were both showing 3 clusters.

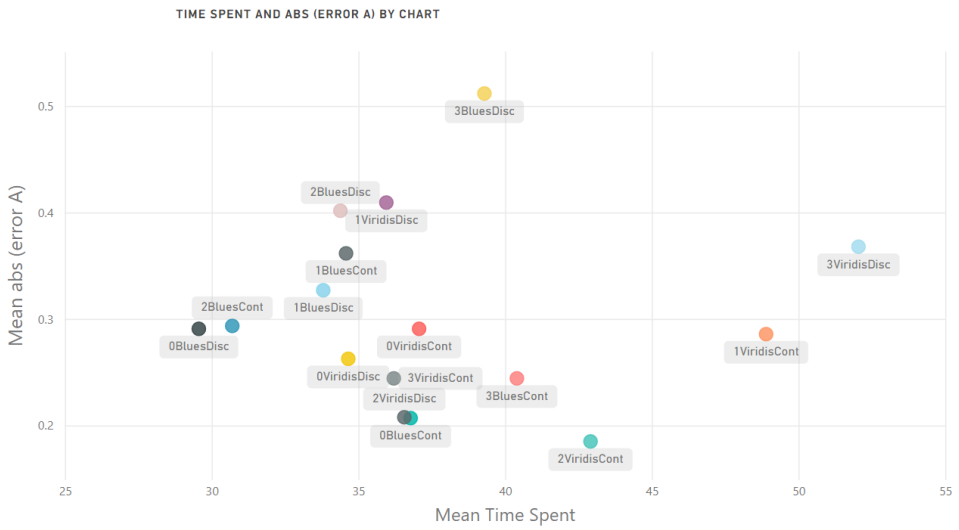


Figure 54. Time spent, Error A distribution by type of chart.

Same happens when we check task B, for the same group of data.

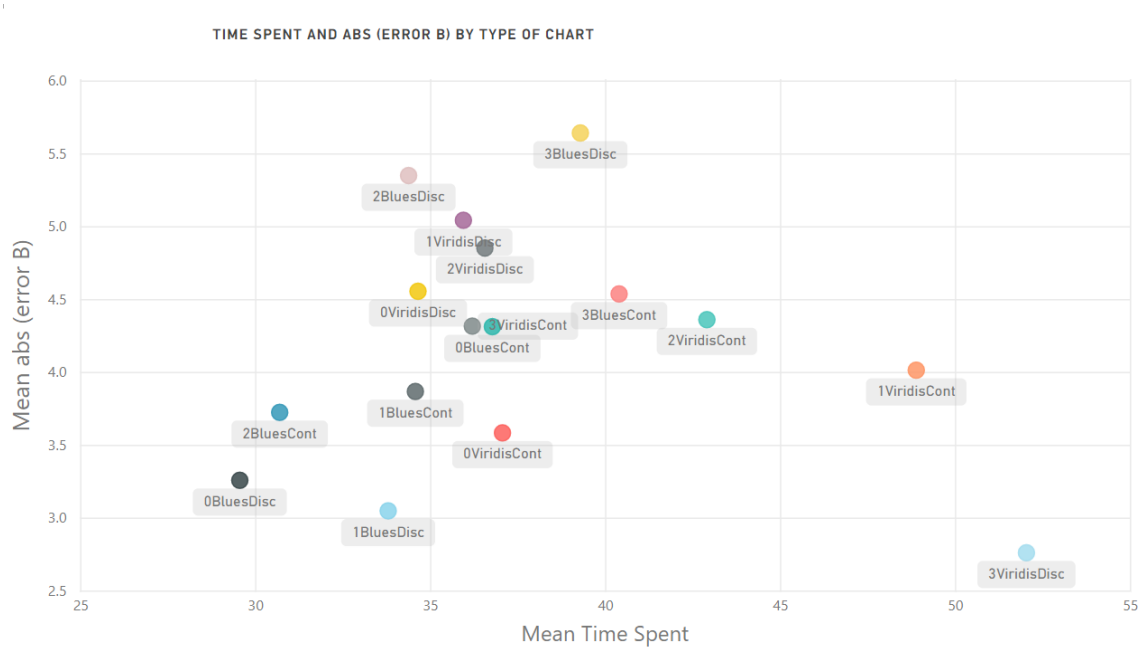


Figure 55. Time spent, Error B distribution by type of chart.

We also found that users tend to subestimate the value of the number of clusters. Like it shown in figure 56. In addition while taking a deep look into the different palette configurations, it can be observed that the sequential palette Blues generates more confusion on the users as shown in Figure 57.

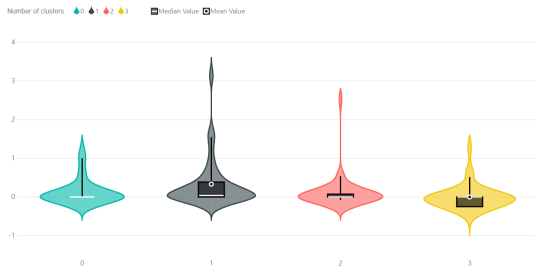


Figure 56. Error A in estimation of Clusters

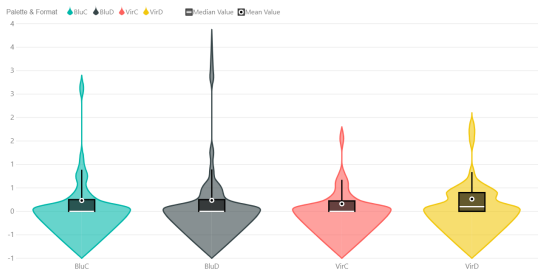


Figure 57. Error B in the different palettes

4.4.3. Dashboard

As a data analyst, I couldn't think about a better way to present the answers of the study besides doing a Dashboard. Although we are not talking business, a Dashboard definitely helps in the process of presenting data through many effective visualizations, and key performance indicators. It will be really useful for the Oral defense indeed.

In addition, thanks to the dashboard, it is fully interactive and a discovery tab, which has been used as an analysis tool to know where to go when trying to find significant statistical results.

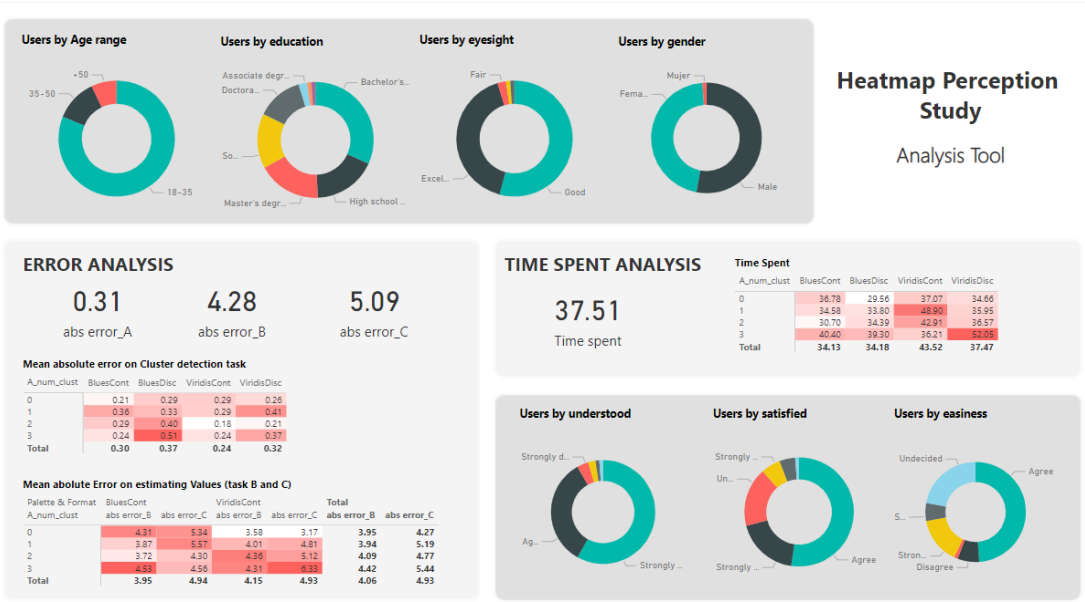


Figure 58. Dashboard used as an analytic tool. Prepared by the author

4.5. Conclusions

The goal of this Final Bachelor Degree Thesis was to assess heatmap visualizations perception, and be able to finally assess on the configuration of the heatmap design for an effective heatmap visualization.

After all the studies conducted, we can now say that the studies and hypothesis conducted lead to a non-statistical difference on the absolute error performed on the tasks evaluated for the different chart configurations conducted.

However, we got different insights on the error and the time spent for the user and number of clusters as shown in Figure 53. In addition we also know that the Blues palette generates more outliers than the divergent palette Virides. In addition we realized how extended this subject can be, so we suggest we continue studying the data gathered in this experiment. Analyzing for instance, the results by the users characteristics, such as the age range, the level of studies, or even segment the users data on their eyesight quality (meaning vision problems). It could also be interesting to take into account the personal evaluation conducted after completing the survey, comparing for instance the results of the users which agreed to understand the tasks to the users that didn't agree.
Moreover, it may also be interesting to study the learning of the users while answering the questions, for instance, some users made more mistakes in the beginning which then eradicate by understanding how the data was designed.

4.5.1. Future work

Due to the lack of time and the very large amount of comparable data in the statistical study, we can extend the study by a thousand different ways. However, we will discuss the future work by two different approaches:

- First, related to the many cases that we have left to study in depth with the data that we currently have.

  Those could be for instance, study the neighborhood of the marked cells to see if it affects the response to have a higher or lower contrast between the colors of the cells, study the possible difference between the errors of the graphs with transposed data and those without, or segment and re-study the errors according to the rating of the view of the users, or the education of these, you could even look if the error in the answers decreases as the survey progresses since users already know the procedure and have found some pattern to answer.

- Second, related to the possible future work either extending the data, gathering more information or conducting more studies on different subjects of heatmap perception.

  Things that come to my mind for example are: you could try with the Viridis reverse palette to see if for the cases where the highest value is darker it is displayed more intuitively and therefore less failures are made. Perhaps we could also evaluate other simple tasks in the graphs such as the detection of outliers or trends, we could even extend the study to the 3d environment as it was originally intended.

4.5.2. Personal conclusions

In this section I will make a personal assessment of the work done, the subject matter, and my personal feelings during the project.

First, I am thrilled to have finally finished the project. It has been lengthened by the obstacles in the development as well as by other external issues. Even so, I am satisfied with everything we have managed to achieve in these 8 months, and the results we have obtained.

Second, I am truly grateful for the opportunity of working on this thesis, which has been as fascinating as complex. And I am also grateful for the director and co-director who have been supporting me from the very beginning and especially in the moments of frustration.

Finally, I may sound a little cocky, but I am proud of myself. I am proud of what I have achieved, of how I have overcome all the obstacles, but most of all from everything I have learned, both personally and professionally. Problem-solving, web development, data visualization or statistics are some of the skills I gathered or improved in the development. Besides, I am proud of the writing, when my thing has never been to write.

5. References

[1] 2 ViRVIG. dirección: https://www.virvig.eu/

[2] AnyChart (n.d.). Heatmap | Chartopedia. [online] AnyChart. Available at: https://www.anychart.com/chartopedia/chart-type/heatmap/

[3] Netek, R.; Pour, T.; Slezáková, R. "*Implementation of Heat Maps in Geographical Information System—Exploratory Study on Traffic Accident".* Open Geosci. 2018, 10, 367–384. https://www.degruyter.com/document/doi/10.1515/geo-2018-0029/html

[4] Finley, K. (2018). WIRED. [online] Wired.com. Available at: https://www.wired.com/

[5] Howe, J. (2006). The Rise of Crowdsourcing. [online] WIRED. Available at: https://www.wired.com/2006/06/crowds/

[6] Estelles-Arolas, E. and Gonzalez-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, [online] 38(2), pp.189–200. Available at: http://www.crowdsourcing-blog.org/wp-content/uploads/2012/02/Towards-an-integrated-crowdsourcing-definition-Estell%C3%A9s-Gonz%C3%A1lez.pdf.

[7] www.cd-genomics.com. (n.d.). Introduction to Heatmap: A Visualized Tool for Microbial Community Composition Analysis - CD Genomics. [online] Available at: https://www.cd-genomics.com/microbioseq/introduction-to-heatmap-a-visualized-tool-for-microbial-community-composition-analysis.html

[8] Davies, J.H. (2013). Global map of solid Earth surface heat flow. *Geochemistry, Geophysics, Geosystems*, 14(10), pp.4608–4622. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/ggge.20271

[9] M. Tory and T. Moller, "Human factors in visualization research," in IEEE Transactions on Visualization and Computer Graphics, vol. 10, no. 1, pp. 72-84, Jan.-Feb. 2004, doi: 10.1109/TVCG.2004.1260759. [online] Available at: https://ieeexplore-ieee-org.recursos.biblioteca.upc.edu/document/4677350

[10] Github. https://github.com

[11] Google Docs. https://www.google.com/intl/en/docs/about/

[12] Salaries at Glassdoor. https://www.glassdoor.es/Sueldos/index.htm

[13] ¿Cuánto cuesta la luz? Endesa. https://www.endesa.com/es/blog/blog-de-endesa/luz/cuanto-cuesta-electricidad

[14] Sustainability – What Is It? Definition, Principles and Examples https://youmatter.world/en/definition/definitions-sustainability-definition-examples-principles/

[15] Fermín Sánchez-Carracedo, David López, Carme Martín, Eva Vidal , Jose Cabré and Joan Climent. "The Sustainability Matrix: A Tool for Integrating and Assessing Sustainability in the Bachelor and Master Theses of Engineering Degrees" 2020, 12(14), 5755; https://doi.org/10.3390/su12145755

[16] Cleveland, W.S. and McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387), pp.531–554.

[17] *M. Tory and T. Moller, "Human factors in visualization research," in IEEE Transactions on Visualization and Computer Graphics, vol. 10, no. 1, pp. 72-84, Jan.-Feb. 2004, doi: 10.1109/TVCG.2004.1260759.* [online] Available at: https://ieeexplore-ieee-org.recursos.biblioteca.upc.edu/stamp/stamp.jsp?tp=&arnumber=1260759

[18]RDlab. (n.d.). *Technical_Report.pdf*. [online] Available at: https://mydisk.cs.upc.edu/s/Ej8SwojwziTZpKN

[19] Jacques Bertin. Sémiologie graphique: *Les diagrammes-les réseaux-les cartes. Technical report*, Gauthier-Villars Mouton & Cie, 1973

[20] Leland Wilkinson. *The grammar of graphics. In the Handbook of computational statistics*, pages 375–414. Springer, 2012.

[21] Andy Kirk. *Data visualization: A handbook for data driven design.* Sage, 2016.

[22] Colin Ware. *Information visualization: perception for design*. Morgan Kaufmann, 2019.

[23] Tamara Munzner. *Visualization analysis and design*. CRC press, 2014

[24] *Redirect Notice*. Quadri, Ghulam Jilani & Rosen, Paul. (2021). A Survey of Perception-Based Visualization Studies by Task. [online] Available at: https://www.google.com/url?q=https://www.researchgate.net/publication/353284356_A_Survey_of_Perception-Based_Visualization_Studies_by_Task&sa=D&source=docs&ust=1650576476247556&usg=AOvVaw2kqOzJ0Ckiipa1CSJzcp-K

[25] Heer, J. and Bostock, M. (n.d.). *Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design*. [online] Available at: http://vis.stanford.edu/files/2010-MTurk-CHI.pdf

[26]Peer, E., Brandimarte, L., Samat, S. and Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, pp.153–163. https://www.sciencedirect.com/science/article/abs/pii/S0022103116303201

[27] Palan, S. and Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, pp.22–27. https://reader.elsevier.com/reader/sd/pii/S2214635017300989?token=490E4F5A981E135EF9F91B2D31C3931226EFDA4B5F292D61C12DD69D9EEE45501EE425E7868FF6A3DF2105D5EA0D4F4A&originRegion=eu-west-1&originCreation=20220418094031

[28]Peer, E., Brandimarte, L., Samat, S. and Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, pp.153–163.
https://www.sciencedirect.com/science/article/abs/pii/S0022103116303201

[29] Peer, E., Brandimarte, L., Samat, S. and Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, pp.153–163.
https://www.sciencedirect.com/science/article/abs/pii/S0022103116303201

[30] Palan, S. and Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17,p.22–27.
https://reader.elsevier.com/reader/sd/pii/S2214635017300989?token=490E4F5A981E135EF9F91B2D31C3931226EFDA4B5F292D61C12DD69D9EEE45501EE425E7868FF6A3DF2105D5EA0D4F4A&originRegion=eu-west-1&originCreation=20220418094031

[31] Eyal, P., David, R., Andrew, G., Zak, E. and Ekaterina, D. (2021). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3765448

[32] Analytics Vidhya. (2016). *Clustering | Types Of Clustering | Clustering Applications*. [online] Available at:
https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/#:~:text=Clustering%20is%20the%20task%20of,and%20assign%20them%20into%20clusters

[33] Mullin, T. (2020, July 15). DBSCAN Parameter Estimation Using Python. Medium.
https://medium.com/@tarammullin/dbscan-parameter-estimation-ff8330e3a3bd#:~:text=From%20some%20research%20I

[34] hdbscan.readthedocs.io. (n.d.). *Parameter Selection for HDBSCAN* — hdbscan 0.8.1 documentation*. [online] Available at:
https://hdbscan.readthedocs.io/en/latest/parameter_selection.html

[35]scikit-learn. (n.d.). *sklearn.datasets.make_blobs*. [online] Available at:
https://scikit-learn.org/stable/modules/generated/sklearn.datasets.make_blobs.html

[36] Khan Academy. (n.d.). *Perlin noise (article) | Noise*. [online] Available at:
https://www.khanacademy.org/computing/computer-programming/programming-natural-simulations/programming-noise/a/perlin-noise

[37] Scher, Y. (2020). *Playing with Perlin Noise: Generating Realistic Archipelagos*. [online] Medium. Available at:
https://medium.com/@yvanscher/playing-with-perlin-noise-generating-realistic-archipelagos-b59f004d8401#:~:text=In%20the%20python%20noise%20module

[38] Díaz, J., Meruvia-Pastor, O. and Vázquez, P.-P. (n.d.). *Improving Perception Accuracy in Bar Charts with Internal Contrast and Framing Enhancements*. [online] Available at:
https://upcommons.upc.edu/bitstream/handle/2117/122977/3w3m0Ei7hgL1j1ahavAeCv.pdf?sequence=3

[39] flask.palletsprojects.com. (n.d.). *Application Setup — Flask Documentation (2.1.x)*. [online] Available at: https://flask.palletsprojects.com/en/2.1.x/tutorial/factory/

[40] Pallets. (n.d.). *Werkzeug*. [online] Available at: https://palletsprojects.com/p/werkzeug/

[41] Pallets. (n.d.). *Jinja*. [online] Available at: https://palletsprojects.com/p/jinja/

[42] Codeloop. (2020). *Flask WTF Registration Form with SQLAlchemy Codeloop*. [online] Available at: https://codeloop.org/flask-wtf-registration-form-with-sqlalchemy/

[43] Grinberg, M. (n.d.). *The Flask Mega-Tutorial Part III: Web Forms*. [online] blog.miguelgrinberg.com. Available at: https://blog.miguelgrinberg.com/post/the-flask-mega-tutorial-part-iii-web-forms

[44] Fullstackpython.com. (2010). *Flask*. [online] Available at: https://www.fullstackpython.com/flask.html

[45] freeCodeCamp.org (2018). *How to build a web application using Flask and deploy it to the cloud*. [online] freeCodeCamp.org. Available at: https://www.freecodecamp.org/news/how-to-build-a-web-application-using-flask-and-deploy-it-to-the-cloud-3551c985e492/

[46] Otto, M. (2000). *Bootstrap*. [online] Getbootstrap.com. Available at: https://getbootstrap.com/

6. ANNEX

A. 6.1. WEB Design

a. 6.1.1. Index page

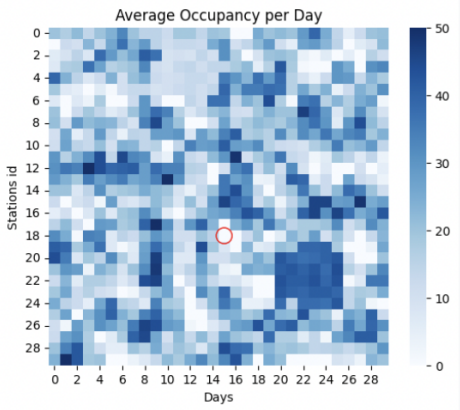
Heatmaps Perception User Study

Welcome!

You are going to participate in a User Study designed to find out how humans perceive data represented through Heatmap Charts.
A Heatmap is a two-dimensional visual representation of data, where values are encoded in colors. Essentially, this chart type is a data table with rows and columns denoting different sets of values or categories.

About the study

To evaluate the effectiveness and characteristics of these visualizations, we will present to you a series of Heatmap charts. They will represent the average number of bicycles in a station (Y axis) for a day number (X axis) from a bike sharing system. As you can see in the next photo.



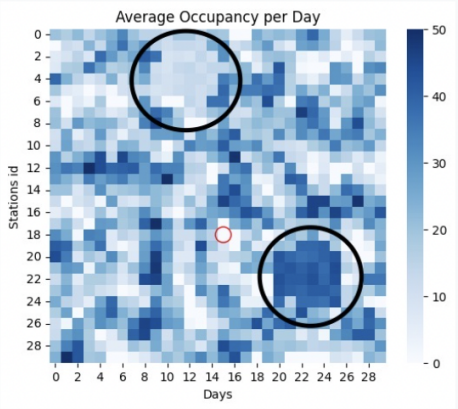
* Note that the color palettes can differ from this one.

The tasks

You will conduct 3 different tasks for each one.

1. Count the number of clusters represented.
2. Guess the value of the cell marked in red.
3. Click over a cell that you think contains a value we indicate.

Referring to a cluster as a group or region of the heatmap where several neighboring cells have similar occupancy (not necessarily the same).
For instance, in the Heatmap shown before, you can see two of these groups, now marked in the next Figure.



This study is structured as follows

1. Demographic Survey. First, you will be asked to provide some personal information (age, gender, quality of eyesight, etc.). We do not ask for your identity, so you will remain anonymous.
2. Introductory Survey. This first introductory survey is a short example to get familiar with the tasks and evaluate your suitability to proceed with the test.
3. Survey. Having passed the introductory task, you will begin the main survey, where you will have to answer the same tasks as the introductory survey but now with open values for the answers.
4. Personal Evaluation Survey. Finally, a personal evaluation survey, to gather additional comments and opinions about the study, if any.

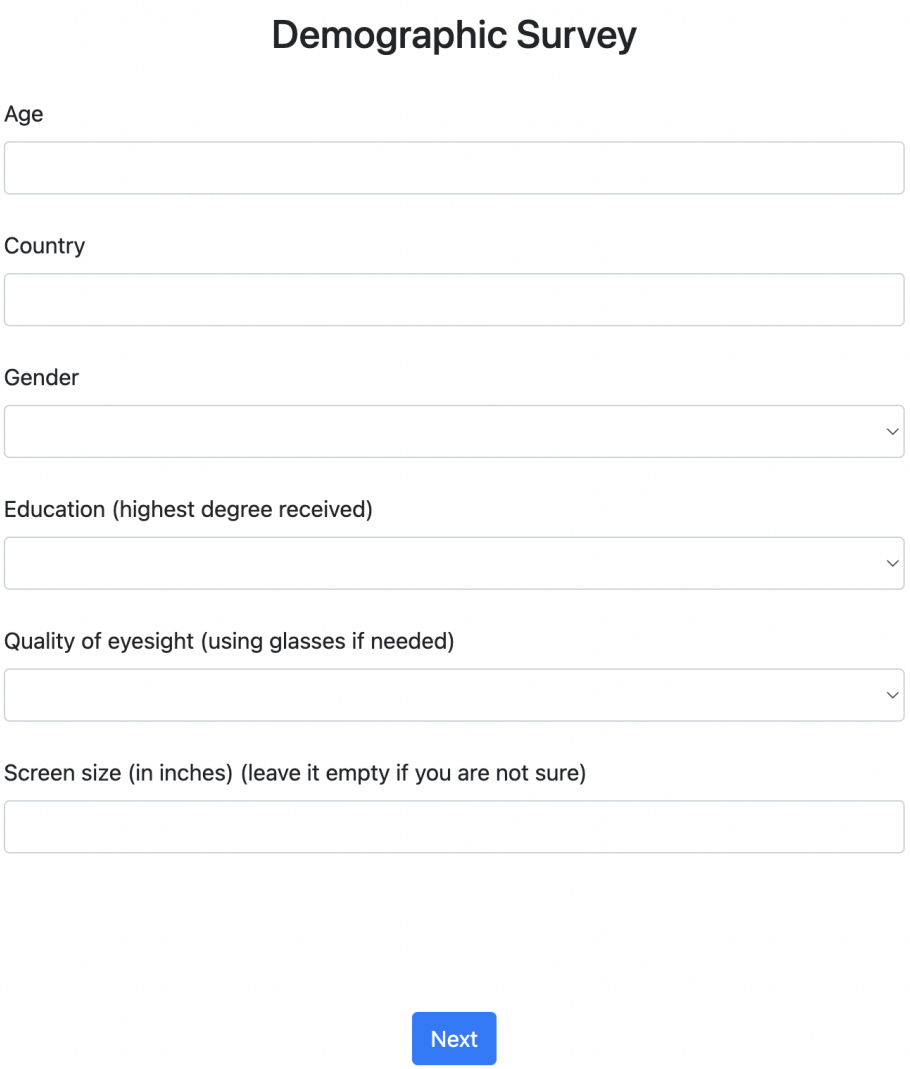
Further considerations

- As a perceptual study, you should not use rulers or any other external tool to measure length.
- The average time to complete the study is about 20 minutes (25 - 35 seconds for each chart). However, you can spend as much time as you need to complete it.
- The data collected will only be used for academic purposes.
- Due to the nature of the study, it does not result in risks to human health.

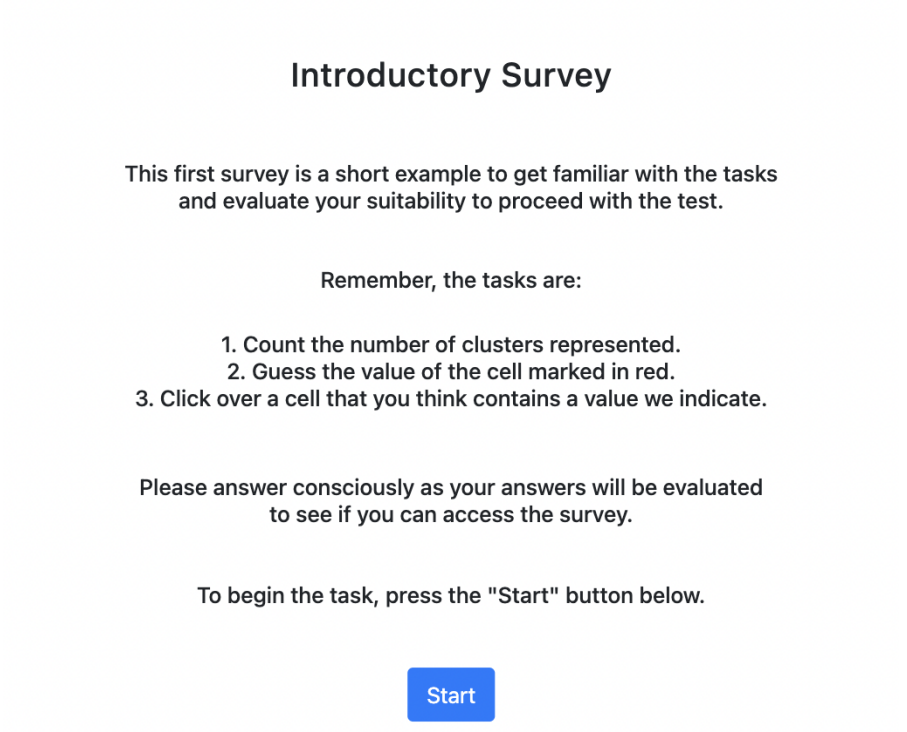
☐ I accept the terms of the study stated above

Start

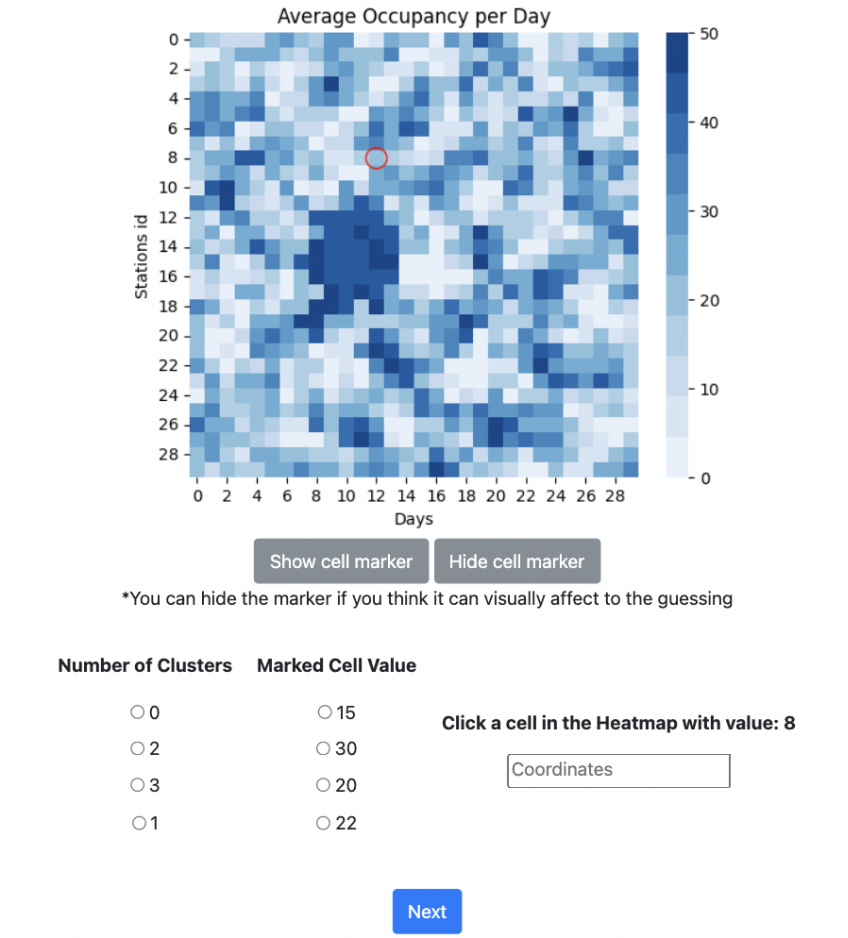
b. 6.2. Demographic Survey



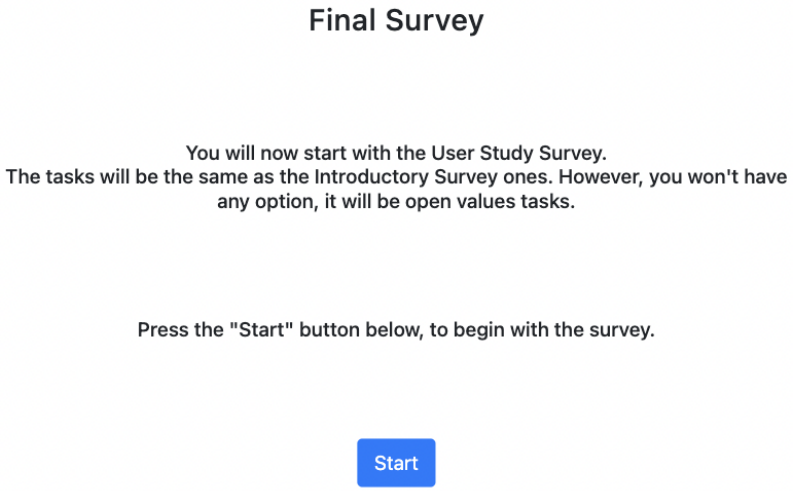
c. 6.2.1. Introduction to the Introductory Survey



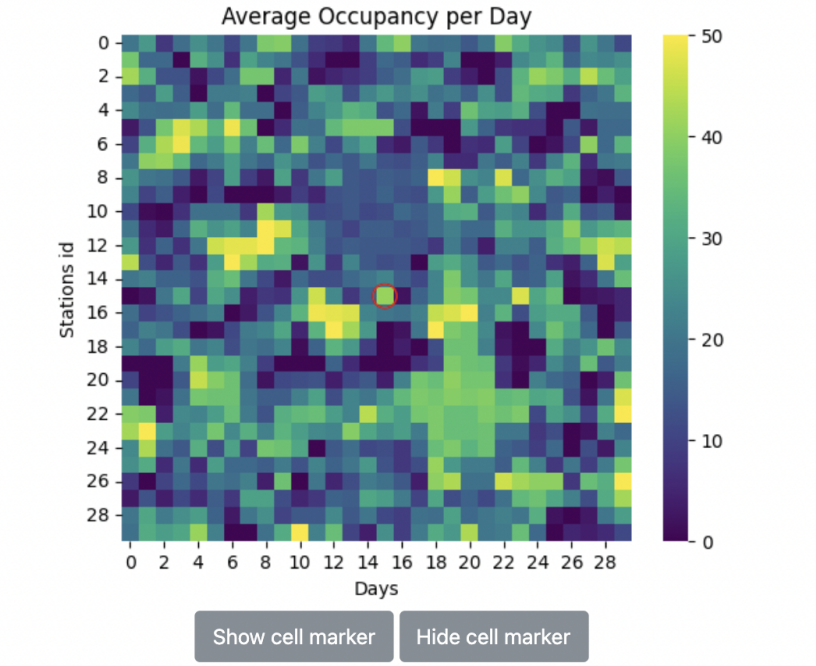
d. 6.2.2. Example of a Task in the Introductory Survey



e. 6.2.3. Introduction to the Main Survey



f. 6.2.4. Example of a Task in the Main Survey



Num of Clusters:

Marked Cell Value:

Click a cell with value 45:

Coordinates

Next

g. 6.2.5. Personal Evaluation Survey

Personal Evaluation

I have understood the different tasks to perform...

It has been easy to complete the tasks...

I am satisfied with the amount of time spent to complete the tasks...

Additional comments (Optional - max. 1000 characters)

Next

B. 6.3. Heatmap Charts Decision

There was an experiment conducted to have a first idea of which charts are non-clear or creates ambiguity. The charts are marked as follows:

Text Discarded by the director, co-director or the user.
Text Discarded by all the contributors.
Text Confusing display (only marked by the users).
Text Chosen + transposed
Text Chosen

** The charts will be discarded only when they are marked as confusing by the director, co-director and the users.

	Viridis Cont	Viridis Disc	Blues Cont	Blues Disc
0	0heatmap0_Viridis_cont	0heatmap0_Viridis_disc	0heatmap0_Blues_cont	0heatmap0_Blues_disc
	0heatmap1_Viridis_cont	0heatmap1_Viridis_disc	0heatmap1_Blues_cont	0heatmap1_Blues_disc
	0heatmap2_Viridis_cont	0heatmap2_Viridis_disc	0heatmap2_Blues_cont	0heatmap2_Blues_disc
	0heatmap3_Viridis_cont	0heatmap3_Viridis_disc	0heatmap3_Blues_cont	0heatmap3_Blues_disc
	0heatmap4_Viridis_cont	0heatmap4_Viridis_disc	0heatmap4_Blues_cont	0heatmap4_Blues_disc
	0heatmap5_Viridis_cont	0heatmap5_Viridis_disc	0heatmap5_Blues_cont	0heatmap5_Blues_disc
	0heatmap6_Viridis_cont	0heatmap6_Viridis_disc	0heatmap6_Blues_cont	0heatmap6_Blues_disc
	0heatmap7_Viridis_cont	0heatmap7_Viridis_disc	0heatmap7_Blues_cont	0heatmap7_Blues_disc
	0heatmap8_Viridis_cont	0heatmap8_Viridis_disc	0heatmap8_Blues_cont	0heatmap8_Blues_disc
	0heatmap9_Viridis_cont	0heatmap9_Viridis_disc	0heatmap9_Blues_cont	0heatmap9_Blues_disc
1	1heatmap0_Viridis_cont	1heatmap0_Viridis	1heatmap0_Blues	1heatmap0_Blues
	1heatmap1_Viridis_cont	1heatmap1_Viridis	1heatmap1_Blues	1heatmap1_Blues
	1heatmap2_Viridis_cont	1heatmap2_Viridis	1heatmap2_Blues	1heatmap2_Blues
	1heatmap3_Viridis_cont	1heatmap3_Viridis	1heatmap3_Blues	1heatmap3_Blues
	1heatmap4_Viridis_cont	1heatmap4_Viridis	1heatmap4_Blues	1heatmap4_Blues
	1heatmap5_Viridis_cont	1heatmap5_Viridis	1heatmap5_Blues	1heatmap5_Blues
	1heatmap6_Viridis_cont	1heatmap6_Viridis	1heatmap6_Blues	1heatmap6_Blues
	1heatmap7_Viridis_cont	1heatmap7_Viridis	1heatmap7_Blues	1heatmap7_Blues
	1heatmap8_Viridis_cont	1heatmap8_Viridis	1heatmap8_Blues	1heatmap8_Blues
	1heatmap9_Viridis_cont	1heatmap9_Viridis	1heatmap9_Blues	1heatmap9_Blues
2	2heatmap0_Viridis	2heatmap0_Viridis	2heatmap0_Blues	2heatmap0_Blues
	2heatmap1_Viridis	2heatmap1_Viridis	2heatmap1_Blues	2heatmap1_Blues
	2heatmap2_Viridis	2heatmap2_Viridis	2heatmap2_Blues	2heatmap2_Blues
	2heatmap3_Viridis	2heatmap3_Viridis	2heatmap3_Blues	2heatmap3_Blues

	2heatmap4_Viridis	2heatmap4_Viridis	2heatmap4_Blues	2heatmap4_Blues
	2heatmap5_Viridis	2heatmap5_Viridis	2heatmap5_Blues	2heatmap5_Blues
	2heatmap6_Viridis	2heatmap6_Viridis	2heatmap6_Blues	2heatmap6_Blues
	2heatmap7_Viridis	2heatmap7_Viridis	2heatmap7_Blues	2heatmap7_Blues
	2heatmap8_Viridis	2heatmap8_Viridis	2heatmap8_Blues	2heatmap8_Blues
	2heatmap9_Viridis	2heatmap9_Viridis	2heatmap9_Blues	2heatmap9_Blues
3	3heatmap0_Viridis	3heatmap0_Viridis	3heatmap0_Blues	3heatmap0_Blues
	3heatmap1_Viridis	3heatmap1_Viridis	3heatmap1_Blues	3heatmap1_Blues
	3heatmap2_Viridis	3heatmap2_Viridis	3heatmap2_Blues	3heatmap2_Blues
	3heatmap3_Viridis	3heatmap3_Viridis	3heatmap3_Blues	3heatmap3_Blues
	3heatmap4_Viridis	3heatmap4_Viridis	3heatmap4_Blues	3heatmap4_Blues
	3heatmap5_Viridis	3heatmap5_Viridis	3heatmap5_Blues	3heatmap5_Blues
	3heatmap6_Viridis	3heatmap6_Viridis	3heatmap6_Blues	3heatmap6_Blues
	3heatmap7_Viridis	3heatmap7_Viridis	3heatmap7_Blues	3heatmap7_Blues
	3heatmap8_Viridis	3heatmap8_Viridis	3heatmap8_Blues	3heatmap8_Blues
	3heatmap9_Viridis	3heatmap9_Viridis	3heatmap9_Blues	3heatmap9_Blues

C. 6.3.1. Final Heatmaps on the study

35 HEATMAPS + 3 instruction

Introduction:

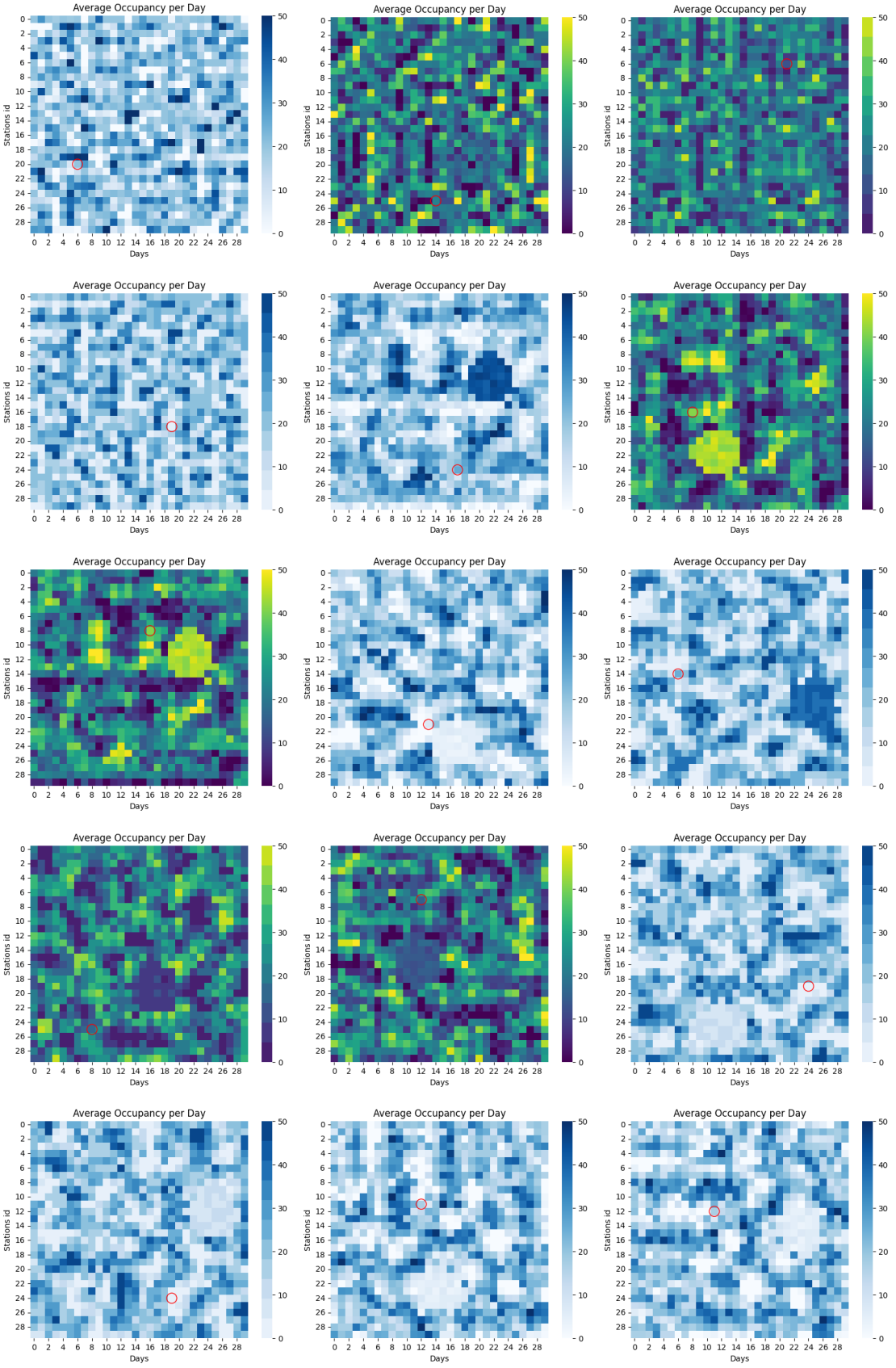
1. 1 cluster Blues disc
2. 3 clusters Viridis cont
3. 2 clusters Blues cont

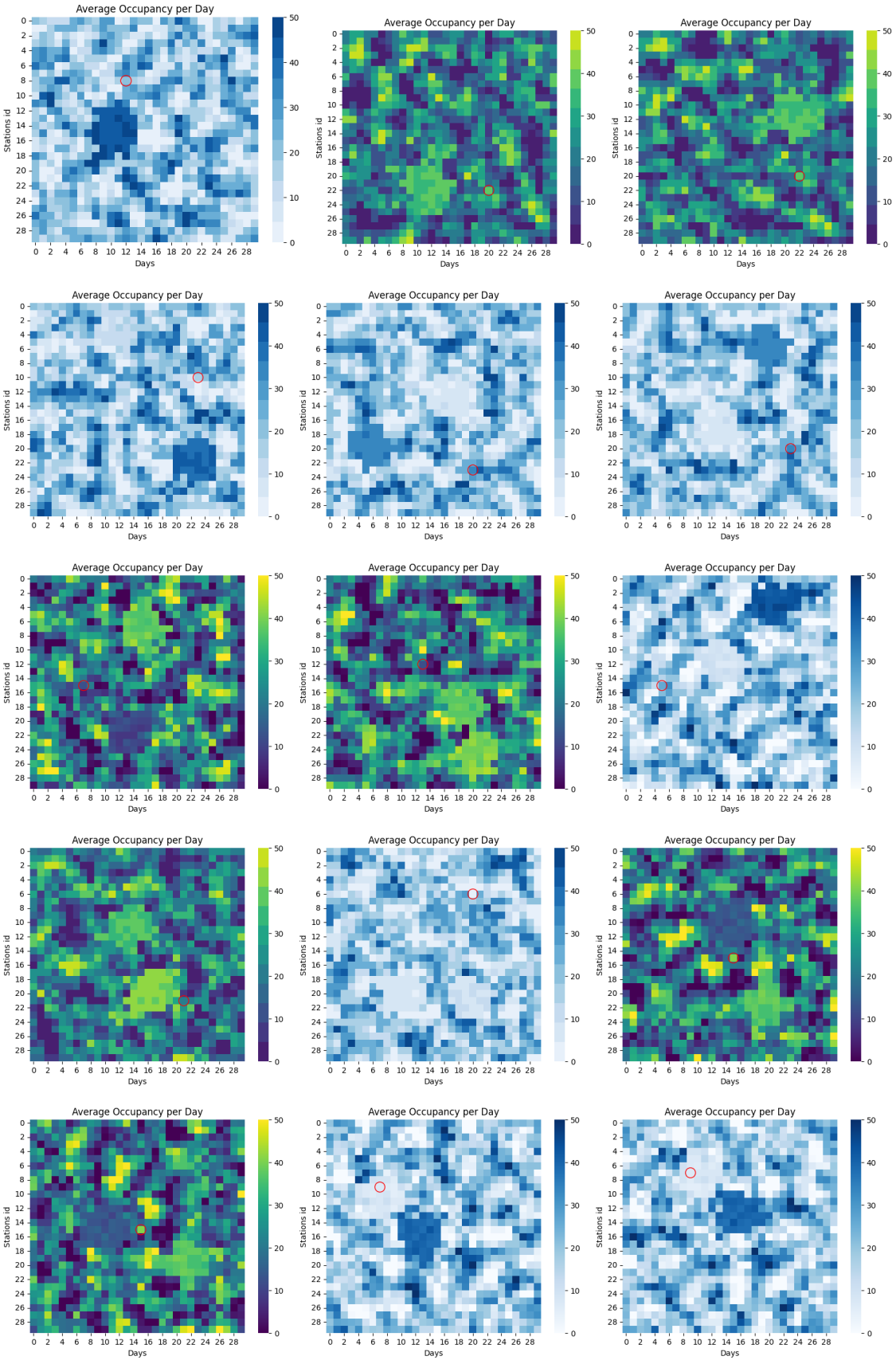
Main survey:

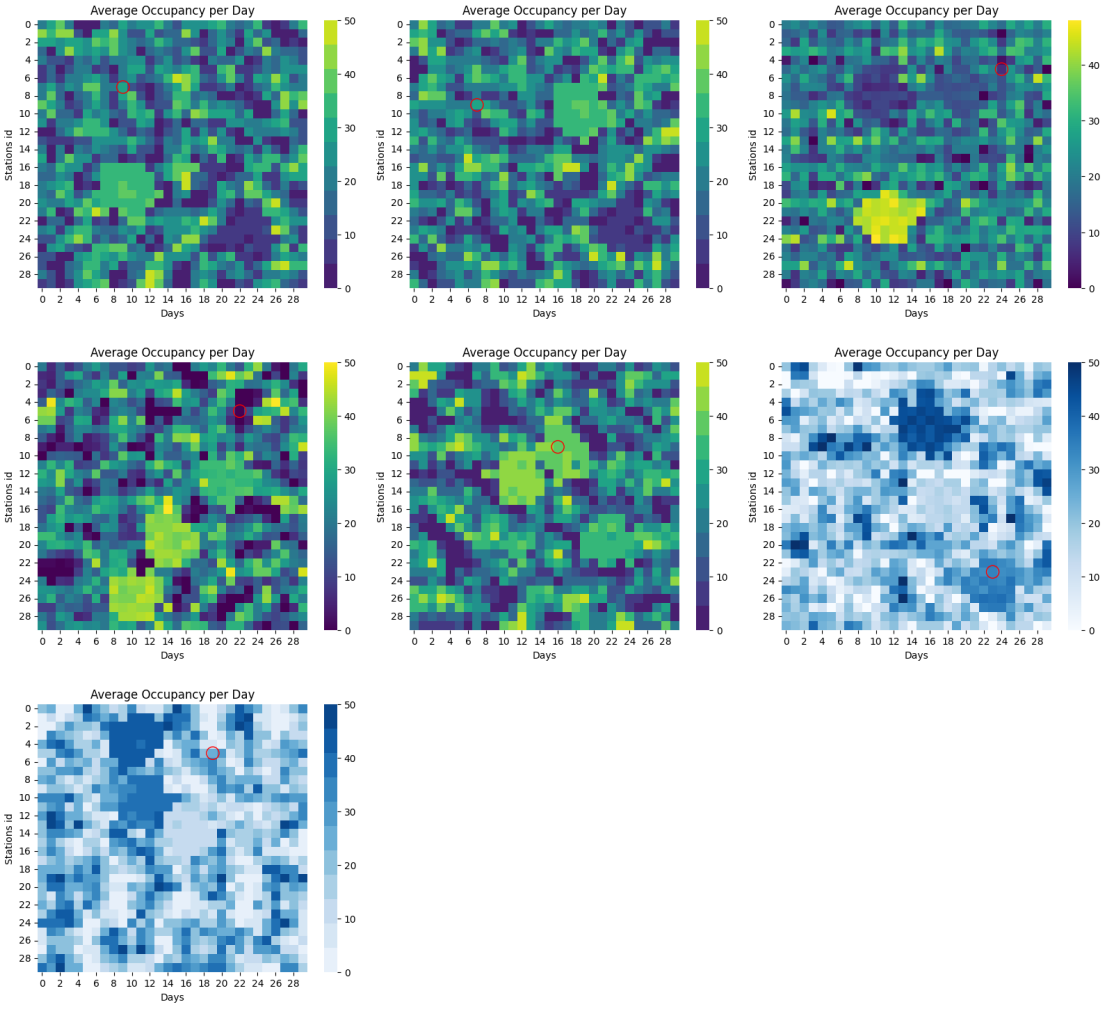
Clusters	Viridis cont	Viridis disc	Blues cont	Blues disc	Viridis cont T	Viridis disc T	Blues cont T	Blues disc T	Validation	Subtotal	Answers
0	1	1	1	1	0	0	0	0	1	**5**	**15**
1	2	2	2	2	1	1	1	1	1	**13**	**39**
2	2	2	2	2	1	1	1	1	1	**13**	**39**
3	1	1	1	1	0	0	0	0	0	**4**	**12**
Subtotal	**6**	**6**	**6**	**6**	**2**	**2**	**2**	**2**	**3**	**35**	
Answers	**18**	**18**	**18**	**18**	**6**	**6**	**6**	**6**	**9**		**105**

VALUES	0-12	12-25	25-38	38-50
20	0	1	0	0
0	1	0	0	0
22	0	1	0	0
10	1	0	0	0
10	1	0	0	0
26	0	0	1	0
17	0	1	0	0
17	0	1	0	0
24	0	1	0	0
0	1	0	0	0
0	1	0	0	0
41	0	0	0	1
41	0	0	0	1
15	0	1	0	0
14	0	1	0	0
38	0	0	0	1
38	0	0	0	1
38	0	0	0	1
26	0	0	1	0
7	1	0	0	0
7	1	0	0	0
34	0	0	1	0
34	0	0	1	0
3	1	0	0	0
41	0	0	0	1
41	0	0	0	1
41	0	0	0	1
12	1	1	0	0
20	0	1	0	0
28	0	0	1	0
28	0	0	1	0
34	0	0	1	0
27	0	0	1	0
6	1	0	0	0
40	0	0	0	1
Total	**10**	**9**	**8**	**9**

6.4. Heatmaps




Average Occupancy per Day
Stations id
Days


Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days
Average Occupancy per Day
Stations id
Days