The Importance of Communication with Missing Data with Colours of Map for Decision-Making

Merve Polat Kayali*, Ana Basiri**

School of Geographical & Earth Sciences, University of Glasgow

* m.polat-kayali.1@research.gla.ac.uk ** ana.basiri@glasgow.ac.uk

Abstract. Missing data is a growing issue in data-driven studies. This study focuses on measuring the importance of communication with missing data and the perception of missing data with colours to find better solutions to potential problems in research with datasets with missing data. Understanding the significance of the interaction with missingness may help us to improve the visualisation of missing geospatial data. We have created a survey that includes some sentences, terms, graphs, maps, and colours we can encounter in daily life. However, all these expressions have some missingness in measuring participants' perceptions and decision-making mechanisms. One hundred-one respondents from different demographics, all from the UK, participated in the survey. It aims to improve the visualisation of missing geospatial data using the results of this research for future studies. But, in this paper, we specifically focus on the perception of representation of missingness with different colours on maps. We observed some differences depending on the gender, age, and education categories of the participants. In the gender-aged analysis of this study's results, a significant difference was observed, especially in the perception of a specific colour.

Keywords. missing geospatial data, communication with missing data, visualisation of missing data

1. Introduction

All scientific disciplines may experience challenges with missing data (Graham, 2009). Most surveys and a significant number of experimental data have an issue with missing data (Scheffer, 2002). Studies in which missing data are represented by the terms negative space, silence, dark matter, dark data, and a black hole have been examined in different disciplines such as physics, astronomy, genetics, political science, social media, art, psychological science, online shopping, medical science, security of living spaces ((Hand, 2020, Buetow, 2009, Dawid et al., 2006, Huggett, 2020, Lieberman, 2016, Luo et al., 2012, NASA, 2022, Noelle-Neumann, 1977, Smith, 2018, Wong, 2011)). Although they all describe missingness with different terms, the effect of the missingness may cause unintended consequences for all of them. Significant challenges can be posed by missingness in the results of the statistical studies. The potential that the missing data in the observable dataset may significantly impact the procedure or results of the research makes them highly crucial. Because a lack of data in the dataset may be of importance that can affect the result of the whole study in the opposite way, these difficulties caused by missing data may also affect the reliability of the study. A study's statistical significance can be reduced by missing data, which can also lead to skewed predictions and incorrect results (Kang, 2013). Even though we cannot see or record these data, they can significantly influence our choices and actions. The consequences may be fatal when we are unaware of the missingness (Hand, 2020). Therefore, to minimise the negative effects of missing data and contribute to the development of new imputation methods, it may be necessary to examine in more detail how the missing data should be represented in the best way. However, since little research has been observed to find the best way to represent missing geospatial data, this paper focuses specifically on the perception of colour representation of

Published in "Proceedings of the 18th International Conference on Location Based Services (LBS 2023)", edited by Haosheng Huang, Nico Van de Weghe and Georg Gartner, LBS 2023, 20-22 November 2023 Ghent, Belgium.
Page 204
This contribution underwent single-blind peer review based on the paper.
https://doi.org/10.34726/5752 | © Authors 2023. CC BY 4.0 License.

missing data. This indicates a need to understand the various perceptions of visualising missing data using colours on maps. In addition to the terms representing missingness in different fields, colours can represent missingness in some areas. The effective use of colours to represent the data helps to show the patterns of the data and their relationship with each other (Brewer, 1994). In cartography visualisation, colours can be defined as the most effective graphic variable that can increase communication, transmit information, and enable seeing a large amount of information on a single map without negatively affecting the intelligibility of the map (Štěrba and Bláha, 2015). The colours used in map visualisations represent an item or specification on the ground, for example, where water is usually represented by blue (Rosenberg, 2019). Although colours were generally used for decoration on maps in ancient times and real maps were left uncoloured, today, colours are used to convey statistical information about many data (Rother). People's perception of colours is very different from each other and is difficult to measure (Bláha and Štěrba, 2014). The perception of colours can cause people to feel different emotions, just as red is usually a warning by representing a danger (Chesneau et al., 2005). As a result, decision mechanisms may be affected. However, colours may not always represent data or information, but in some cases a deficiency situation.

1.1. Research Aims & Questions

We aim to understand if there is a difference between different groups to perceive missingness on a map. Depending on our results about perceiving missingness, we can find the best way to visualise missing geospatial data. However, we need to understand the importance of missing data before visualising missing geospatial data. Communication can be the best way to analyse the significance of the missingness in geospatial data science. To understand the importance of communication with missing data, it is necessary to choose the best way. To observe the impact of communication with missingness and its effect on scientific studies, we decided to observe the impact of missingness for the participants using the sentences, colours, and terms we use in daily life before augmented and virtual visualisation methods. In summary, our main purpose in this study will be to observe how the participants communicate with missing data in colours. As a result, it is thought that by evaluating the significance of communication with missing data, presenting these data to the participants using technologies such as augmented reality and forming a basis for understanding the importance of communication.

In addition to the data sets used in scientific studies, we may experience many missing data in our daily lives, whether we are aware of it or not. These data can affect our decision mechanism depending on their meaning and importance. This paper evaluated the part of preliminary results of the continuing study to find the best way to visualise missingness in a 3D environment. Therefore, we aim to get answers to the survey question that we examine for this paper to the following question:

- Is there any difference between demographics based on the perception of missing data with colour?

2. Method

In our study, we aimed to understand the significance of interacting with missingness and visualising missing data; a survey was created that includes missing data based on different aspects that participants would come across in everyday life. With different questions in the survey, we aimed to measure users' perspectives and decision mechanisms for lack of data or data uncertainty.

We specifically planned to observe if different colours of missing data on maps in different

datasets make sense to participants for this paper. We observe how the missing data on the map in the survey affects the perspectives on the cases and how the importance of the colour or hatches can be perceived according to the types of missing data. Participants were expected to evaluate these deficiencies by reflecting the missing data on the number of data given depending on colour.

It is planned to measure whether the users perceive the colours or hatches as a deficiency and how they evaluate them, and to observe which missing data is more important or unimportant depending on participants' age, gender, and education level, with the results of survey questions. Depending on the idea that visualisation can be one of the effective methods to measure the perception towards missing data, we designed our question with the colours and hatches we used on the map we created. We asked the participants to indicate what the colours and hatch symbols used in different countries on the map mean to them (Figure 1). We designed the answers, so they have two choices, whether a colour represents missing or available data. As a result of this survey, the participants evaluated eight colours (pink, orange, turquoise, black, green, grey, white, cloud blue/grey) and two hatches (dot and line).

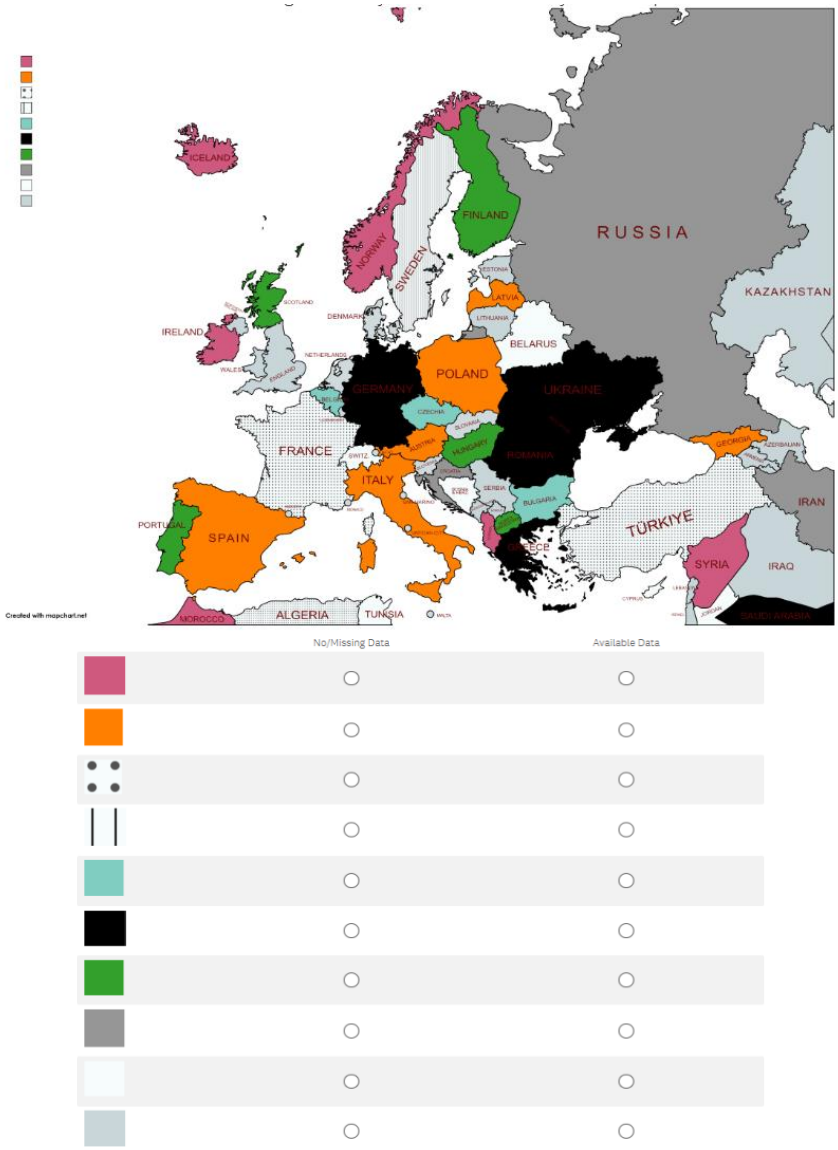

Figure 1 Survey Question – asks participants to indicate whether the colours and patterns used for the map's visualisation represent "available data" or "no/missing data" for them.

3. Results

This study is aimed to observe whether the missing data perception of colours on maps is different between men and women. It was observed that 51 of the survey participants were male, and 50 were female. The age information of the participants was expected in the form of four different categories: "18 - 29", "30 - 44", "45 - 59", and "60+" in the question asked to the participants. According to the initial observations from the survey data, when the participants with different genders perceived different colours on the maps as incomplete data based on their age groups, no significant difference was observed between the male and female participants. As a result, gender-based participants in four different age group categories were divided into two groups, below middle age and over (Table 1), and it was observed whether men and women in these age groups exhibited other behaviours in perceiving the representation of missing data with colours. Our preliminary results show a significant difference between men and women over middle age in perceiving the "black" colour as missing (Figure 2), even though there is no significant difference in perceiving other colours and hatches. We are still in the initial analysis stage of our study, and as the next step, we will observe how the participants with different demographics perceive the missing data from the colours on the maps.

	18 - 44	45+
Female	31	19
Male	16	35

Table 1. Number of Participants

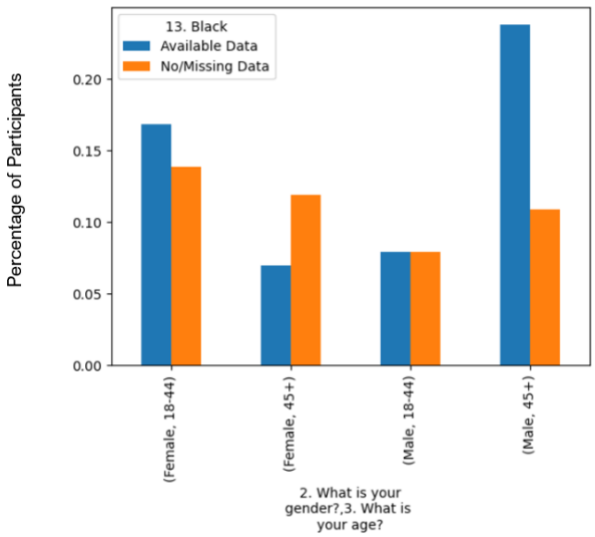

Figure 2. Perceiving “Black” colour on the map

4. Future Works

The purpose of this study is to understand better and examine whether there is any difference in perceiving missing data in colours on the map for different demographics such as age, gender, and education. Findings from this study may have important implications for developing a study to communicate and visualise missing geospatial data. We need to understand the importance of communication with missing data and the importance of missing data and to measure how important missing data is in decision-making. It is planned to determine the effects of missing data based on location, time, colours, and amount separately depending on the age, gender, and location data of the participants as a next step with survey results. Moreover, since the presentation of many colours together on the map designed for the survey may cause different effects on the perception of the participants and may confuse them, it is planned to continue the study to observe the isolated representation of each colour to measure the perception of the participants in a more understandable way that the different colours represent the missing data, as a next step of this study. The results

of these studies will form the basis of the next step, the visualisation of missing geospatial data using augmented reality and virtual reality technologies. In addition, by analysing the results of other questions in our survey, we will lay the foundation for finding the most appropriate way to visualise missing geospatial data in a 3D environment using AR technologies.

5. Acknowledgements

Great thanks for the financial support, Republic of Türkiye Ministry of National Education Scholarship Program. The work presented in this paper has been funded by the UK Research and Innovation (UKRI) Future Leaders Fellowship MR/S01795X/2.

References

BLÁHA, J. D. & ŠTĚRBA, Z. 2014. Colour contrast in cartographic works using the principles of Johannes Itten. *The Cartographic Journal,* 51**,** 203-213.

BREWER, C. A. 1994. Color use guidelines for mapping. *Visualization in modern cartography,* 1994**,** 7.

BUETOW, S. A. 2009. Something in nothing: Negative space in the clinician-patient relationship. *The Annals of Family Medicine,* 7**,** 80-83.

CHESNEAU, E., RUAS, A. & BONIN, O. Colour contrasts analysis for a better legibility of graphic signs on risk maps. Proc. 22th International Cartographic Conference (ICC), 2005.

DAWID, A. P., MORTERA, J. & VICARD, P. Representing and solving complex DNA identification cases using Bayesian networks. International Congress Series, 2006. Elsevier, 484-491.

GRAHAM, J. W. 2009. Missing data analysis: Making it work in the real world. *Annual review of psychology,* 60**,** 549-576.

HAND, D. J. 2020. *Dark data: Why what you don't know matters*, Princeton University Press.

HUGGETT, J. 2020. Capturing the silences in digital archaeological knowledge. *Information,* 11**,** 278.

KANG, H. 2013. The prevention and handling of the missing data. *Korean journal of anesthesiology,* 64**,** 402-406.

LIEBERMAN, S. 2016. Missing. *Emotion, Space and Society,* 19**,** 87-93.

LUO, Y., HE, J., ZENG, Z., ZHAO, J., ZHANG, Y. & LAN, C. Study of investigating and improving negative space in an old urban block of Nanning city: Characteristics and formation causes of negative space. 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), 2012. IEEE, 2102-2105.

NASA, O. 2022. *Black Holes* [Online]. NASA. Available: https://science.nasa.gov/astrophysics/focus-areas/black-holes [Accessed 19.03.2022 2022].

NOELLE-NEUMANN, E. 1977. Turbulences in the climate of opinion: Methodological applications of the spiral of silence theory. *Public opinion quarterly,* 41**,** 143-158.

ROSENBERG, M. 2019. The Role of Colors on Maps. *Colors can Represent Boundaries, Elevations, and Bodies of Water [Online] retrieved from https://www. thoughtco. com/colors-on-maps-1435690*.

ROTHER, S. H. E. P. G. *The Geography of Colorants* [Online]. UNIVERSITY OF MICHIGAN LIBRARY. Available: https://apps.lib.umich.edu/online-exhibits/exhibits/show/the-geography-of-colorants [Accessed 2023].

SCHEFFER, J. 2002. Dealing with missing data.

SMITH, H. R. 2018. *What Is a Black Hole?* [Online]. NASA. Available: https://www.nasa.gov/audience/forstudents/k-4/stories/nasa-knows/what-is-a-black-hole-k4.html [Accessed 18.03.2022 2022].

ŠTĚRBA, Z. & BLÁHA, J. D. 2015. The Influence of Colour on the Perception of Cartographic Visualizations. *Color and Image*, 533-38.

WONG, B. 2011. Negative space. *nature methods,* 8**,** 5.