

The Perceptions of Data Science outputs in a local authority context: the importance of emotion and context



Mohamed Souare and Jo Bates
Information School, University of Sheffield, United Kingdom
Email: chesou88@gmail.com , jo.bates@sheffield.ac.uk

1. INTRODUCTION

- Advanced data visualisation and machine learning techniques are increasingly becoming part of governance and decision-making in local authorities.
- In addition to ethical considerations and limitations raised around the implementation of such techniques, recent research has evidenced that factors such as time and opinions, beliefs, design, literacy and elements of interactivity affect engagement with data through visualisations (Kennedy, Hill, Allen and Kirk, 2016).
- These factors provoke diverse emotions which are a vital component of making sense of data. In other words, “emotion matters in everyday engagement with data” (Kennedy and Hill, 2017).
- In the context of local authorities, it becomes important to understand how emotion shapes employees’ interpretation of data science.
- We argue that the feelings, thinking and ability to interpret data science outputs cannot be ignored in data-driven projects in local government.

2. OBJECTIVES

Investigate how people working in a local authority context engage with data science outputs produced using datasets relevant to their work and which they are familiar with. This is achieved by:

- 1) Producing different outputs from relevant datasets provided by the local authority using different data science methodologies and visualisation design
- 2) Capturing data from a local authority’s employees about their feelings, interpretation and thinking of the outputs generated in objective one
- 3) Identifying and categorising meaningful responses captured in stage two and examine their relation to making sense of the data visualisation

3. METHODS

Phase 1 : Quantitative Phase

- Collection and analysis of relevant datasets on green spaces and deprivation.
- 4 outputs generated using exploratory and spatial techniques in R.



Follow up

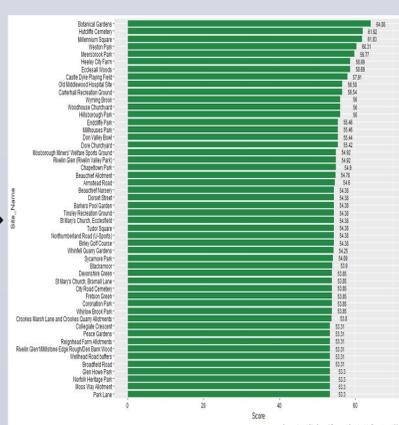
Phase 2 : Qualitative Phase

- Use of outputs from stage 1 as an input to:
- Design the questionnaire
 - Conduct semi-structured interviews with 5 local authority employees working on relevant areas.
 - Qualitative data were analysed using Thematic Analysis.

In phase1, decisions were taken regarding the colour schemes and ordering of entities that aimed to influence the interpretation by others in different directions.

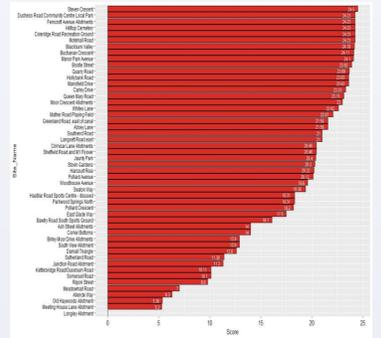
4. RESULTS

Visualisation 1 : Top 50 Green Sites by score



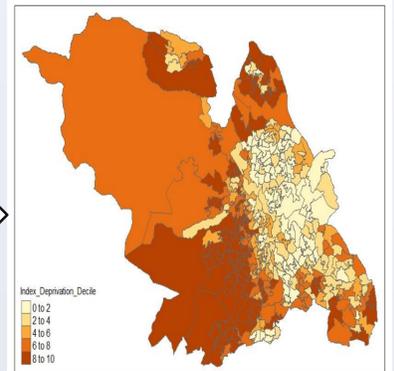
Employees less familiar with the data responded emotionally (surprise) at learning new insight from the viz. As one reacted: **How many of them do you have in total? is it just 50 or more than that? 800. Oh really! Waou! ...that is really interesting**

Visualisation 2: Bottom 50 Green Sites by scores



Emotional response was expressed when a participant related her experience of seeing how data is represented during the annual performance meeting
As she expressed: **“Seeing these plots make feel I am going to lose my job”**

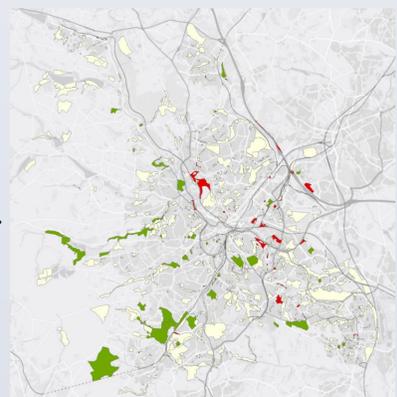
Visualisation 3: Mapping the 2015 Index deprivation decile in the City



- Some employees repetitively **insisted on having further details** on the scores and argued that the graph could be misleading without the wider context.
- Others were **enthusiastic** to bring more context by using their knowledge of the history of the city to explain how deprivation is spread across the local authority.

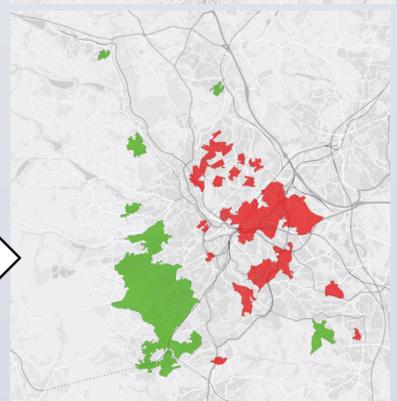
ROUND 2 OF VISUALISATIONS

Visualisation 4: Mapping the Top 50 and Bottom 50 Green sites in the City



- A recurrent question after seeing the 3 visualisations was: where the Top and Bottom 50 green spaces were located on the map?
- This led to build visualisation 4&5
- Most of employees were **pleased** to see that **insights matched with prior knowledge**.

Visualisation 5: Advanced Mapping of the Top 50 and Bottom 50 Green sites in the City



- The perception of the value in the visualisations in relation to:
- Making cases and decision around resources allocation. As one decision maker stated: **“with this, I can clearly prioritize our work”**.
 - Communicating their work to the public as mentioned by the community manager

5. CONCLUSIONS

- **Data and their representations cannot be dissociated from the socio-cultural and contextual factors in which they are generated and interpreted.** Our emotions, the context of the data, the knowledge of the end-user and their perception of the usefulness of the information all influence engagement with data and their visual forms, and resultantly with data science and its outputs.
- **Data and their representations are not neutral, and a variety of affective and contextual factors can influence their production, analysis and interpretation.** As data science techniques are embedded into local authorities, this understanding needs to be integrated in the data science process in order to optimize the benefits of data-driven initiatives for citizens and local governments.
- **As with the private sector, data science problems must be framed strategically in local authorities to bring value and measurable impacts.**

6. REFERENCES

- Kennedy,H & Hill,R.L.(2017). The feeling of numbers: Emotions in everyday engagement with data and their visualisations.SAGE Journals/Sociology,1-9. <https://doi-org.Sheffield.imd.oocl/10.1177/0038038516674675>
- Kennedy,H., Kirk,A., Hill,R & Allen,W. (2016). Engaging with big data visualization: factors that affect engagement and resulting new definitions of effectiveness.First Monday,21(11).Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/6389>