Project Protocol

Contents

1 Title

Assessing the effectiveness of different visualisation methods for Australian spatial data.

2 Project team roles & responsibilities

Researcher	Role	Affiliation
Stephanie Kobakian	Masters Student	Faculty of Science and Engineering (QUT)
Prof. Kerrie Mengersen	Principal Supervisor	Faculty of Science and Engineering (QUT)
Dr. Earl Duncan	Associate Supervisor	Faculty of Science and Engineering (QUT)
Prof. Dianne Cook	External Supervisor	Department of Econometrics and Business Statistics, Monash University

3 Background information

3.1 Project outline

This project will test the effectiveness of two types of spatial displays: a choropleth map, and a hexagon map. It will contrast the effectiveness of communicating information when each geographic region is represented by a hexagon, rather than its geographical shape. The purpose of these displays is to convey the spatial distribution of the disease occurrence, or incidence. This can allow for detecting hot spots corresponding to outbreaks, or spatial trends, for example, indicating occurrence is related to latitude or even rural vs urban differences. Effectiveness of the display will be measured by accurate and efficient perception of these patterns.

3.2 Introduction/background information

Dorling (2011) suggests alternative map displays that aim for ‘equitable representation design’ present each member of a society equally. They are considered more ‘socially just’ but also help to present an honest display of the experiences of the whole population (Dent 1972). This work aims to determine if using a coloured hexagon display is a more effective method of communicating the spatial relationship between a set of areas on a map.

3.3 Rationale/justification

This work is motivated by the newly developed Australian Cancer Atlas where information regarding cancer incidence and mortality is provided to the community through a web page. We are interested to know whether providing the spatial distribution in a hexagon map is more effective for communication than the traditional choropleth map. The responses of participants will be contrasted when shown the same information in both a geographic and hexagon map display and asked to find the same real distribution in a collection of maps arranged in a line up. The rate of participants who pick the maps in the line ups that display the true spatial relationship will be used to contrast the effectiveness of the displays. We hope that this will show a hexagon map is an effective method for communicating spatial relationships.

4 Study objectives

4.1 Hypotheses

Participants will pick the map of a true distribution in a line up of null maps more often when the information is presented in a hexagon map, as opposed to a geographic map line up.

4.2 Research questions/ aims

1. Measure the accuracy in spotting a real spatial distribution from a line-up of null maps
2. Contrast the accuracy when participants pick from geographic or hexagon maps

4.3 Outcome measures

Experiment information collected:

- Choice of plot in each line up
- Reason for choice of plot
- Time taken to answer
- Perceived difficulty of question

Factors:

- Plot type (Hexagon/ Geographic)
- Spatial dependence structure (Normal errors / Chi-squared errors)

We will record demographic information to consider the impacts of:

- gender (female / male / other),
- degree education level achieved (high school / bachelors / masters / doctorate / other),
- age range (18-25 / 25-35 / 35-45 / 45-55 / 55+ / other)
- lived in Australia (Yes / No / Other)

5 Study design

Each participant will first see three test displays, they will then undertake an online survey linked to them through the online crowdsource platform Figure-Eight. Within the experiment: A line up protocol will be used to arrange 20 maps into one display. Each display will be created by a combination of plot type, error distribution and spatial dependence. We will contrast the sets of different plot designs, hexagon tilemap and geography in the lineups created using the same data, and same null positions within the lineup. We will compare the length of time taken and the accuracy with which participants report their choice.

The possible levels of the factors are:

- Plot type: Geography, Hexagons
- Error distribution: Normal, Chi-Square
- Spatial Dependence: Rurality level, North to South, East to West

Allowing each factor combination to be considered will provide 12 (2x2x3) line up displays to be seen by all participants.

6 Study population

6.1 Participants

We will aim to survey 100 participants. Participants will be approached via their dashboard through their Figure-Eight accounts.

6.2 Inclusions and exclusion criteria

The platform will advertise this survey to participants that fulfill the following: - level 2 or level 3 on the Figure-Eight Platform. - at least 18 years old - trained and qualified to annotating images.

6.3 Recruitment strategies, timeframe

Recruitment will take place via advertising on the Figure-Eight dashboards. The Figure-Eight platform will advertise this as a "job" to potential participants, they will not be under any obligation to complete this task. We will have the opportunity to wait until enough participants have completed the survey and provided their judgements.

6.4 Consent approaches

Participants will be shown instructions which will provide sufficient information about the task when they select to do the survey. These instructions will explain the survey to participants, and give them an example of the questions they will answer. Contact information for the researchers will be made available to the participants. Participants will be asked to provide consent via a checkbox question. They will then continue to the survey, and complete the test questions. This page will also explain the use and storage of the data they contribute. The data will be the answers they provide, which will be publicly available on github at https://github.com/srkobakian/experiment.

6.5 Participant withdrawal

Participants will be able to withdraw by leaving the survey at any time. XXX will they still be paid?

7 Procedures

7.1 Screening of participants

Participants will need to have achieved contributor level two on the Figure-Eight platofrm, this is a smaller group of more experienced, higher accuracy contributors. Test questions are the most important quality control mechanism in the Figure-Eight Platform. Participants will complete them to practice and prepare for the experiment questions, they will need to reach an accuracy threshold we will set, to be allowed to participate in answering the experiment survey questions.

Contributors that fail quiz mode are not paid and are permanently disqualified from working on your job. Participants will need to reach an accuracy threshold we will set to be allowed to answer the questions in our survey. Test questions will also be included as the job progresses so we can be sure quality remains high.

7.2 Data collection

As the survey is conducted through the Figure-Eight platform, the human data will also be collected in the form of an IP address, Contributor ID, Trust Score and Channel. Some sources believe an IP address is identifiable information, as such it will not be held publicly online when the data is held on github. We will also ask for information regarding gender, education level achieved, and age range to understand perception of different groups.

7.3 Data collection/gathering techniques:

Data will be collected through survey distribution via a web application. Participants will access the application via link from a Figure-Eight "job".

7.4 Impact of, and response to, missing data

If participants want to provide and "I don't know" answer they will be able to select this. If there are more than 50% of questions that are incomplete we will remove this survey and consider this withdrawal.

8 Statistical plan

The design of this study is based on work by Hofmann, Follett, Majumder and Cook (2012).

8.1 Sample size determination and power; or how will determine that data saturation has occurred or data are sufficient for the purpose.

To be able to effectively evaluate the use of a hexagon tilemap we will contrast the proportion of times the real data plot was chosen in a line up of maps. The same data will be used to create 19 null data plots and one real data plot. This data will be displayed to users in both geographic maps and hexagon tile maps.

We will have aim to collect results from 100 participants (n=100), who will have tried to pick the real data in a line up of simulated data.

Let X be the corresponding random variable X, where i.e. X = # times out of n independent repetitions that the data plot is picked from the lineup.

Under the null hypothesis, X has a Binomial distribution: $X\ B_{n,1=m}$. The p-value of a lineup is the probability to have x or more observers picking the data plot, under the assumption that the null hypothesis is true: $p - value = P(X \geq x|H_0) = 1 - B_{n,1/m}(x - 1)$

This sets our Type I error rate, α, at a level of $1/m$. The authors also define how to estimate the power of a lineup: the ratio of correct identifications of the real data, x, out of n viewings.

8.2 Data analysis and Statistical methods

Information will be collected regarding the design (including how will you measure, manipulate and/or analyse the information that you collect/gather, matching and sampling strategies, accounting for potential bias, confounding factors and missing information)

This will be analysed using a generalised linear model, with random effects capturing the differences between the 100 participants. The main effects used will consider the spatial distribution of the errors, gender, degree education level achieved, age range and whether the participant has lived in Australia.

Each lineup will have similar hypotheses: H_0 : there is no real structure visible in the data plot H_1 : there is real structure visible in the data plot

Each lineup is a set of m = 20 plots. One plot contains data with a real structure. The probability that an observer picks the data plot when it is really *not* different from the 19 null plots, is $1/m = 1/20$.

If a participant is able to identify the data plot from the lineup, we reject the null hypothesis. Type I error rate, α, is set to a level of $1/m = 1/20 = 0.05$.

Compare the percentage of times users select the real data in a lineup for the same set of data presented in choropleth maps and hexagon tile maps. Compare percentages of correct responses $\hat{\pi}_C$ and $\hat{\pi}_H$:

Where C denotes the use of choropleth maps and H denotes the use of hexagon tile maps.

The percentage of correct responses where the respondent identified the real data in a choropleth map is denoted by $\hat{\pi}_C = (x_C + 1)/(n_{100} + 1)$, similarly for hexagon tile maps: $\hat{\pi}_H = (x_H + 1)/(n_{100} + 1)$.

9 Outcome measures

An expansion of the various observations, data and measures that will be made during the study and which will contribute to establishing success or otherwise in addressing the research question.

The collection of multiple responses from each person as they complete several different lineup tasks, allows estimates of their perceptual ability and allows them to correct power differences between competing designs: by modeling power using a subject-specific random intercept in a generalized linear model.

10 Data management and record keeping

As the data will be held publicly online, it will be accessible under a license for people to reuse. It will also be held accessible to QUT researchers

10.1 Confidentially and privacy

All data will be deidentified, there is no need to protect confidentiality when the data has been anonymised.

10.2 Data security

All the survey information will be stored on github, a publicly available site, as well as on site at QUT.

10.3 Record retention

The records of data and analysis undertaken will be stored on github as well as on site at QUT.

10.4 Secondary use

The analysis scripts and the data collected will be available for secondary use. Thisn their consent form. will be disclosed to participants

11 Resources

This work will be supported by ACEMS funding - $500 to pay 100 participants equally.

12 Results, outcomes and future plans

- Plans for return of results of research to participants
- Publication plan
- Other potential uses of the data at the end of the project
- Project closure processes
- Plans for sharing and/or future use of data and/or follow-up research

13 References

14 Appendices

- These may include surveys, questionnaires, interview schedules, recruitment materials, manual of operations, etc

Dent, Borden D. 1972. "A Note on the Importance of Shape in Cartogram Communication." *Journal of Geography* 71 (7): 393–401. https://doi.org/10.1080/00221347208981697.

Dorling, Daniel. 2011. "Area Cartograms: Their Use and Creation." In *Concepts and Techniques in Modern Geography (Catmog)*, 59:252–60. https://doi.org/10.1002/9780470979587.ch33.

Hofmann, H., L. Follett, M. Majumder, and D. Cook. 2012. "Graphical Tests for Power Comparison of Competing Designs." *IEEE Transactions on Visualization and Computer Graphics* 18 (12): 2441–8. https://doi.org/10.1109/TVCG.2012.230.