UNIVERSITY OF LEEDS

School of Computing
Faculty of Engineering

Assessed Coursework Header/Feedback Sheet

This form should be completed and submitted with each piece of assessed coursework.

My signature here confirms acceptance of the declaration of academic integrity printed overleaf.	
Student Name:	Adam Brown sc20asb, Ynyr Evans sc20ye, Macsen Hue sc20hw, Karl Zhu bn22kz
Signature:	
SID Number:	2 0 1 4 0 8 8 6 4

You must make sure that you have read the declaration of academic integrity.

To be completed by the student submitting work
(please make sure that you complete ALL fields)

Module Code C O M P 3 7 3 6

Module Title Information Visualization

Coursework Number 2

Name of Lecturer Marking Work Roy Ruddle ****You must fill this in****

Deadline Date 15/12/2022 **Date Handed In** 14/12/2022

Feedback section to be completed by the marker

Section	Marks Available	Marks Awarded	Comments on sections
Total			
Late Penalty (marks to be deducted)			

General Comments on Work

Declaration of academic integrity

Further information about plagiarism, fraudulent or fabricated coursework and malpractice in University assessments is available at:
http://www.leeds.ac.uk/aaandr/cpff.htm

I am aware that the University defines plagiarism as **presenting someone else's work, in whole or in part, as your own.** Work means any intellectual output, and typically includes text, data, images, sound or performance.

*(On the understanding that other members of the group have made contributions to the attached submission,)** I promise that in the attached submission I have not presented anyone else's work, in whole or in part, as my own and I have not colluded with others in the preparation of this work. Where I have taken advantage of the work of others, I have given full acknowledgement. I have not resubmitted my own work or part thereof without specific written permission to do so from the University staff concerned when any of this work has been or is being submitted for marks or credits even if in a different module or for a different qualification or completed prior to entry to the University. I have read and understood the University's published rules on plagiarism and also any more detailed rules specified at School or module level. I know that if I commit plagiarism I can be expelled from the University and that it is my responsibility to be aware of the University's regulations on plagiarism and their importance.

I re-confirm my consent to the University copying and distributing any or all of my work in any form and using third parties (who may be based outside the EU/EEA) to monitor breaches of regulations, to verify whether my work contains plagiarised material, and for quality assurance purposes.

I confirm that I have declared all mitigating circumstances that may be relevant to the assessment of this piece of work and that I wish to have taken into account. I am aware of the University's policy on mitigation and the School's procedures for the submission of statements and evidence of mitigation. I am aware of the penalties imposed for the late submission of coursework.

**For the submission of group work*

Penalties for late submission

University rules on penalties for late submission of coursework require 5% of the total marks available to be deducted for each calendar day that passes after the date of required submission. The deduction will be applied to the grade/mark for the coursework component concerned before any conflation with other grades/marks to give the overall result for the module. If coursework is not submitted by the end of 14 calendar days following the prescribed deadline, a grade/mark of zero should be returned for that component.

Coursework 2: Evaluation exercise

Introduction

The aim of our experiment is to compare the compare the effectiveness of bar charts against maps for the visualisation of plastic pollution. The data for each visualisation will be randomly generated. This includes producing different data for each question, regardless of if that specific question has been shown for a different chart type.

We are going to measure the effectiveness of both chart types by measuring how quickly and accurately participants can comprehend the data presented to them. Our experiment is designed to ask a question related to each visualisation and to measure two metrics. Firstly, if the answer is correct, to quantify accuracy, and the response time to quantify comprehension speed. These metrics will be collated from all participants to determine how effective each chart type is.

The task will be executed by gathering a group of 10 participants and displaying 10 maps and 10 bar charts to each participant. Each visualisation will be displayed one by one, with the participant having to answer a multiple-choice question in a graphical user interface about the current visualisation. Both the correctness and response time of the participant's answer will be recorded. Each participant will perform this action for all 20 charts. The experiment should take each participant no longer than 15 minutes to complete, as there is only one question per visualisation and the questions themselves are straightforward to answer.

Method

Participants

For this evaluation coursework, we had a total of 10 participants. Each of the participants were given a participant information sheet that outlined the experiment, and were informed that by participating in the experiment, they are giving informed consent. All participants were voluntary in the process, at no stage were there any incentives offered or provided to them.

Materials

To create the visualisations, the matplotlib library for python was used for all bar charts produced. Whilst the geopandas library was used to create all map chart visualisations. We also used python's dateTime library to utilise timestamps to record response times of the participants to each question.

The Tkinter library in python was used to provide the GUI for each question. Python's random library was used to create all the random data used for the visualisations. A static random seed value was also used to ensure the reproducibility of the random data. There was also an initial terminal before the GUI was loaded that explained the procedure and details of the experiment to participants was also made in Python 3.

When the question is first displayed to the user, the timestamp is stored as a variable. Upon the user answering the question, the overall time in milliseconds is calculated, based on the difference between the earlier stored timestamp and the current timestamp.

The users' response time and if their answers were correct or not is then stored within a .csv file.

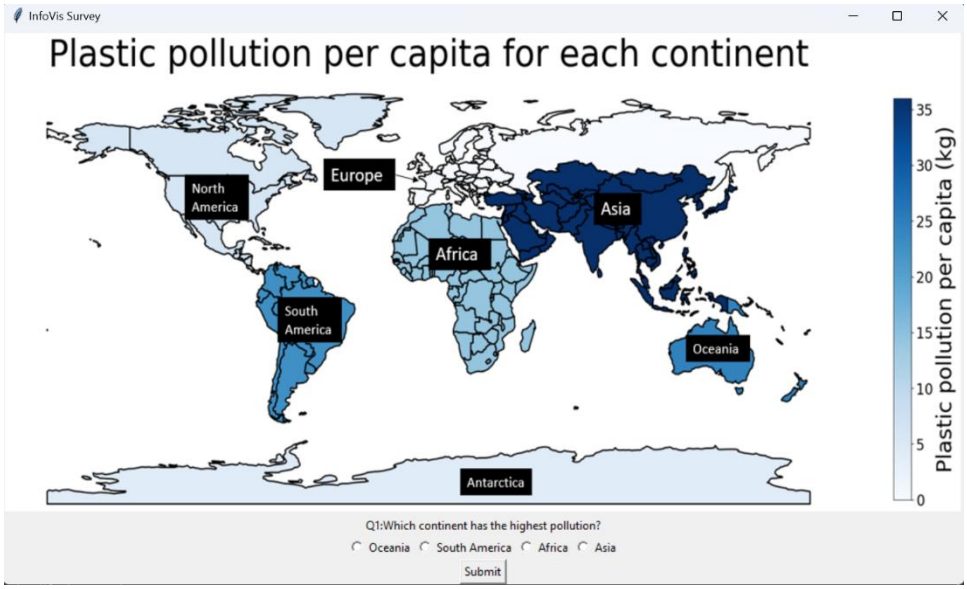
Procedure

The experiment will start by running the python application on a laptop. Once it starts running a message appears on the terminal informing the participant about what the experiment will entail, what metrics are being recorded, and instructions for how to answer the questions.

The user presses the enter key in the terminal to ensure they understand the procedure. This helps ensure that response times for the first few questions won't be unnecessarily long if the user is confused about what to do.

Each participant will then be presented with 20 visualisations and multiple-choice question regarding that data, which will be shown one by one. There is a short one second delay between displaying each visualisation.

To answer each question the user will have to comprehend what information the visualisation is showing and answer a question based on the visualisation in question. The GUI provided shows the visualisation, 4 options and an 'enter' button. A sample screenshot of the GUI is provided below:



To answer the question the participant must press an option, then press the 'enter' button in the GUI. Both the data and related question for each visualisation will remain constant for every participant, in order to achieve fairer results.

After a participant completes all these questions, we will store the metrics (the correctness and response time). After we have repeated this procedure to collect metrics from 10 participants, the metrics will be loaded into Excel for data analysis.

Results

A quick note that questions 1-10 are based on maps of the world, UK, Europe, and the US. Questions 11-20 are bar maps displaying the same types of data as in questions 11-20 (e.g., the same countries) but the data itself has been randomised again. Data has been randomised again to eliminate any sort of

participant recall which could lead to the response times for bar chart being quicker than they would be otherwise.

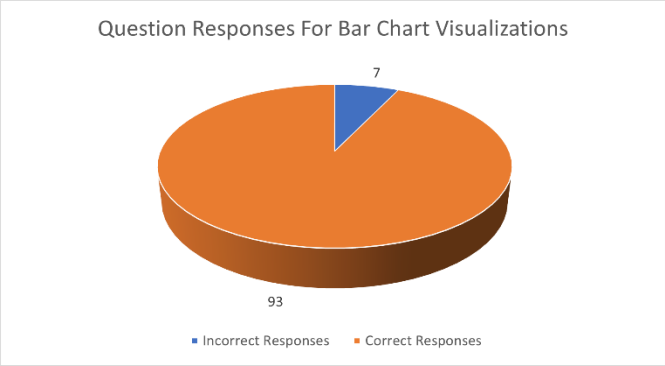


Figure 1: The correctness of responses to each individual question on bar charts

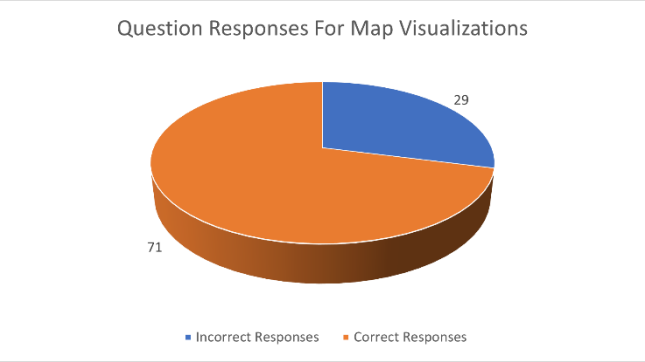


Figure 2: The correctness of responses to each individual question on maps

If we look at the pie charts above, it clearly shows that bar charts are superior to map charts to answer questions about a certain visualisation correctly, producing over 20% more correct responses in our results. This is likely down to bar charts showing the exact values for each data point, where maps only show the hue. Therefore, data differentiation for small differences is far less straightforward for maps, leading to more errors in answers. Whilst both chart types are still effective in accomplishing this task, there is a significant increase in error for map charts.

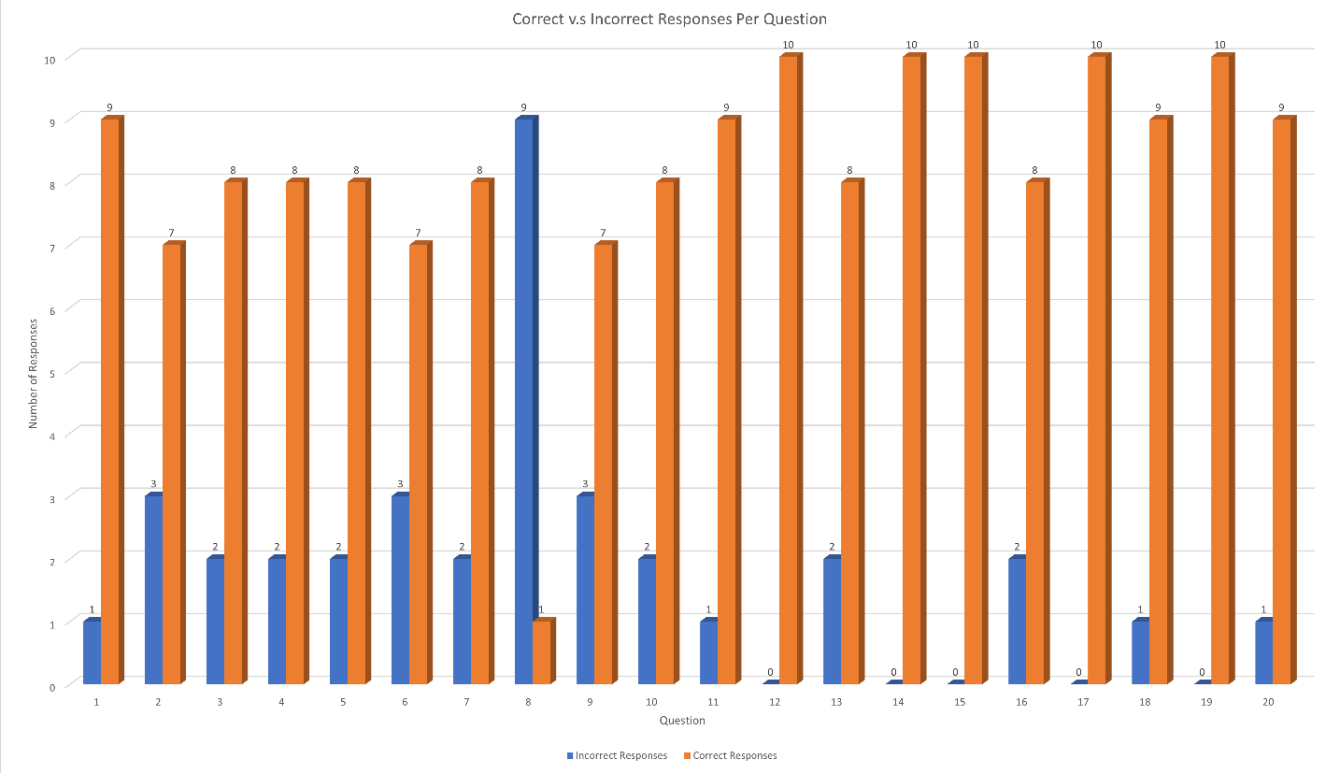


Figure 3: The correctness of responses of all participants for each individual question

The majority of answers given to questions were correct, except for the outlier of Q8. This outlier is caused by the hues on the maps being extremely similar, which has resulted in many respondents getting the question incorrect. This shows that hue is often not as discriminable as other visual encoding methods, as in this instance the majority of participants were unable to correctly read the chart. In addition to this, the regions asking to be compared shared no borders, which made direct comparisons between the hues of these regions more difficult.

The percentage of answers that were correct is also significantly higher for questions 11-20, as these questions are all bar charts. From this we can show that responses are more accurate for bar charts than for the maps. From these results we can conclude that the length of bars is easier to compare than the hues of regions on a map, particularly when data is similar.

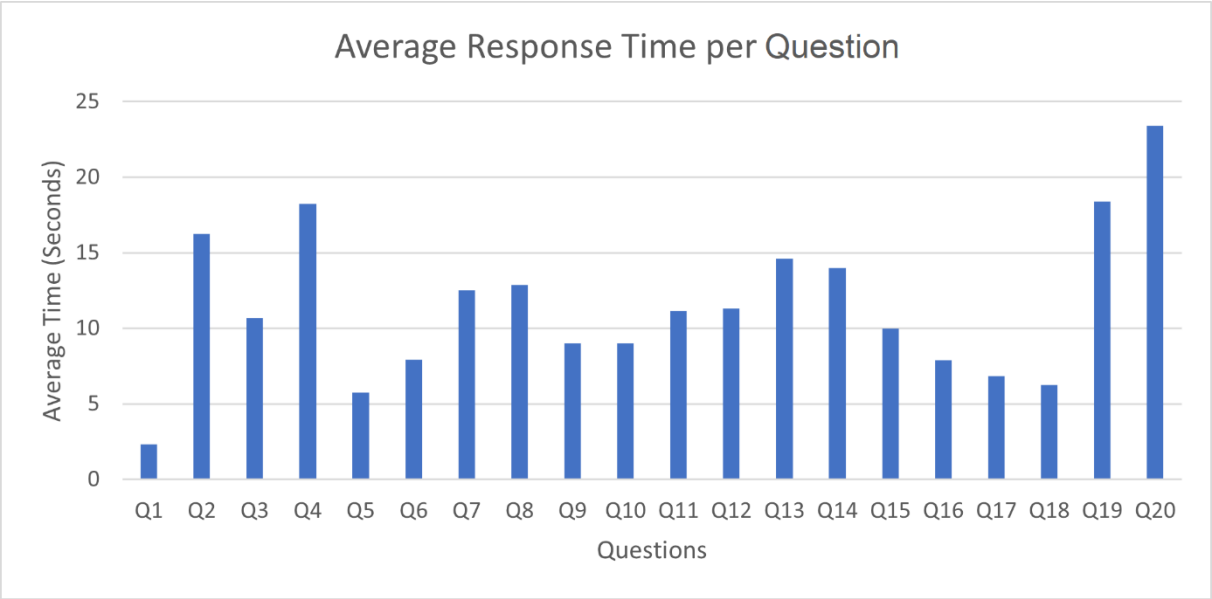


Figure 4: The average response time of each individual questions from all participants.

Generally, the response time for bar charts is slower than the response time for map charts. This suggests that it may be easier to quickly comprehend a map chart in comparison to a bar. It's worth noting that the previously discussed outlier is heavily skewing Q2, and without this outlier it would likely have a similar response time to Q1 due to them both being questions about continents. As continents are large, well known geographic groups they can be comprehended quickly. This result for Q1 illustrates that maps are more suited for areas with fewer and clearer regions/groups (in this case, continents).

The much longer response times for Q19 and Q20 can be explained by the fact that these questions had a very large number of bars, as they were for each state. This takes a lot longer to comprehend than less populated bar charts in Q16, Q17 and Q18. This explains the large jump in response time.

However, it's important to note that Q18's bars are ordered by weight. This made it much easier to spot the highest value and thus was a factor in the very quick response time for this question.

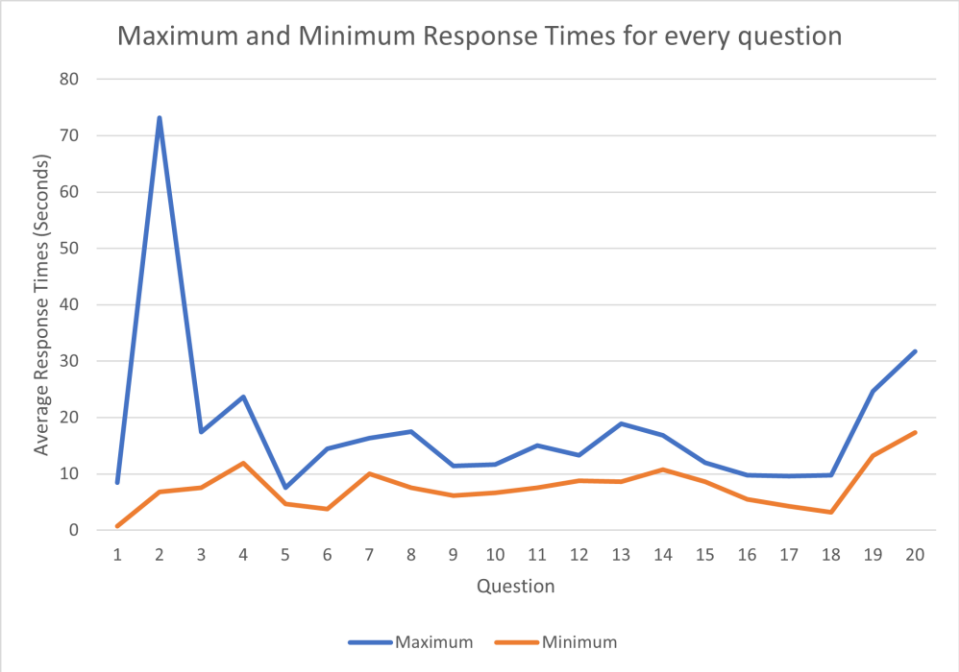


Figure 5: The highest and lowest individual response time for each question.

We can see that there is a large disparity between the maximum and minimum response time for Q2. This is due to a single outlying value, which is shown by the average response time for Q2 (as seen in Figure 4.) not being an outlier, whilst its standard deviation (as shown in Figure 6.) is significantly higher than all other values. This shows that maps can have a significantly wider disparity in response time, due to factors such as some users struggling to discern hue and the users' geographical knowledge varying.

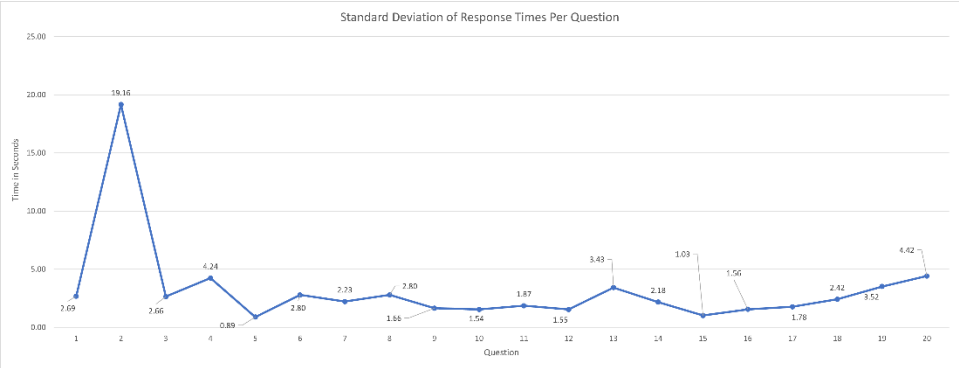


Figure 6: The standard deviation of the response time for each individual question.

Aside from the previously mentioned outlying value influencing Q2, we can see that the next highest standard deviation is Q4. Q4 displayed a colour map of Europe, therefore meaning that some countries being shown were significantly smaller in area than others. This means that users with a poorer knowledge of Europe took far longer in responding than those who had a good geographical knowledge of Europe. This accounts for the standard deviation for Q4 being so generally high, due to a disparity in users' knowledge.

A general trend that we can see for our bar charts (Q11-20) is that questions which have significantly more columns, such as questions 19 and 20, also have larger standard deviations. This shows that bar charts' readability becomes increasingly inconsistent when there are more columns within them.

Conclusions

From this experiment, we've discovered that maps are very useful for getting a good general overview of the data quickly, as shown by their quick response times. However, they aren't as effective for answering specific questions about a visualisation. This is because the hue is not very suitable for showing minute differences in values, especially if the two countries in question aren't bordering each other.

Whereas bar charts are far more effective if you want to answer specific questions about a plastic pollution visualisation, as shown by their far superior accuracy rate for question correctness. This is because it's much easier to spot small differences in values, due to bar length being a more obvious visual encoding method. However, this comes at the expense of taking more time to comprehend and understand the data being shown, which is backed up by the slower question response times in comparison to maps. This is especially true for bar charts with a greater number of columns.

Overall, this report finds that maps are more suited for quicker comprehension, especially for a large number of different variables. However, they struggle to deliver the accuracy that bar charts can provide. This is particularly true for when users have poor geographical knowledge of the areas they're being presented with, and when the values of variables are similar.

Appendix

Participant Information Sheet

Participant Information Sheet: Coursework 2 Evaluation Experiment – COMP3736 Information Visualization

Participating Student:

Address: School of Computing, University of Leeds

We are a group of 4 students in the module COMP3736 Information Visualization here at the University of Leeds. We are interested in how a group of participants evaluate a series of different chart visualizations. If you volunteer for this experiment, you will answer 20 questions in regards to 20 different visualizations. Each question will display either a bar chart or a map visualization, and your task is to answer a question that will test your interpretation of the data that the visualization is presenting.

The experiment will take approximately 30 minutes. Your responses will be recorded, but you will remain anonymous. The research will be reported in our deliverables for the coursework, but no-one should be able to identify you and at no point will your identity be divulged.

By taking part in the experiment, you are indicating informed consent. You are free to withdraw from the experiment at any time.

Finally, please let me know if you have any questions or would like to discuss anything with me.