

UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2016

IS52032A

Data Journalism and Visualisation

Duration: 2 hours 15 minutes

Date and time:

This paper is in two parts: part A and part B. You should answer ALL questions from part A and TWO questions from part B. Part A carries 40 marks, and each question from part B carries 30 marks. The marks for each part of a question are indicated at the end of the part in [.] brackets.

There are 100 marks available on this paper.

**THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM**

Part A

Question 1 Looking for stories in data is often called interviewing the data. Name three techniques for interviewing data using Excel. [4]

Question 2 In what circumstances would you recommend using a pie chart to represent data? State two design considerations that you must take into account to using them appropriately. [4]

Question 3 What are the three types of quantitative data? Give an example of each type. [4]

Question 4 Give three examples of thematic maps commonly used in online journalism. [4]

Question 5 Given brief definitions of pre-attentive and attentive processing in the perception of graphics. Give an example of each. [4]

Question 6 Data scientist Hadley Wickham called data that was in the correct form for easy processing in spreadsheets and statistical packages ‘tidy data’. List the three rules of tidy data. [4]

Question 7 Describe the differences between a population and a sample. Give two examples of each. [4]

Question 8 Briefly describe forward and reverse geocoding. [4]

Question 9 Describe two advantages that using an SQL database has over using a spreadsheet to store data. [4]

Question 10 What type of relationship does a time-series analysis describe? What is the type of chart most suitable to reflect this relationship? Explain what would be your x and y axes. [4]

Part B

Question 11 Data journalism

- (a) Data journalism is often described in terms of a 5-step workflow. List the 5 steps. [5]
- (b) Describe the typical processes and at least one tool involved at each step. [10]
- (c) What are the three most common types of chart used by journalists to present data in stories. [6]
- (d) What types of data are the three charts in the previous question used to present and give two examples of data that might be displayed using them. [9]

Question 12 Scraping and networks

- (a) What is web scraping? Describe the workflow involved when using web scraping for journalism. [9]
- (b) Give three examples of how scraping is used to produce data sets used in journalism. [6]
- (c) Describe how network graphs can be used in journalism and give two examples of stories or investigations that might use them. [9]
- (d) What is the average path length of a network graph and what does it tell us about a network? [6]

Question 13 Digital mapping

- (a) What are the two common coordinate systems used to represent a point on a map. [6]
- (b) Name three software tools used to create digital maps for web presentation. [6]
- (c) Name four types of visualisation most commonly used by journalists to show geographical data. [8]
- (d) Name three common geographical units of analysis used in the UK and give two examples each of data commonly used by data journalists in each of these units. [10]