

COMP4462: Data Visualization

Final Examination – Part 2

Spring 2020

Instructor: Huamin Qu

Monday, 25 May 2020, 3:10PM - 4:30PM

Student Name: _____

Student ID: _____

Instructions:

- 1. This is an open-book, open-notes examination. Part 2 has totally 4 pages and 3 questions.*
- 2. Write your name and student ID on each page*

4. Visualization Design [15 marks]

With the development of computer vision techniques, we can derive humans' emotion data from a video. Suppose you are given the emotion data of all students in COMP4462, which is collected from a ZOOM meeting video. The data record the emotion of each student at each second. Due to the model constraints, emotions are expressed by categorical attributes of five basic types, i.e., Happy, Surprised, Angry, Neutral, and Sad. For instance, one data entry is:

Timestamp	People	Emotion
01:35	Alice	Happy
	Bob	Surprised

(a) Please design a visualization scheme to show an overview of the emotional evolution of all the students during the video. [5 marks]

(b) Please design a visualization to display the aggregated emotional transitions of students in the video. Your design should be able to show the total number of transitions between two emotions (e.g., 50 students change their emotions from Happy to Surprised; 30 students change from Surprised to Happy; 35 students change their emotions from Neutral to Sad). [10 marks]

5. Visualization Design [20 marks]

Search data may reflect the life for millions of people. Given a search database consisting of the search queries related to the health issues, you need to develop a search monitoring visualization system for all the countries in the world. Specifically, you are provided with data extracted by a data processing engine which monitors the search queries and produces at regular intervals the following information:

- Date (e.g., 25/10/2014)
- Health issues (e.g., Flu, Obesity, Depression, Heart disease, ...)
- Country (e.g., China, USA, Japan)
- Volume (search times) (e.g., 1356 for Flu, 256 for Obesity, 45 for Depression)

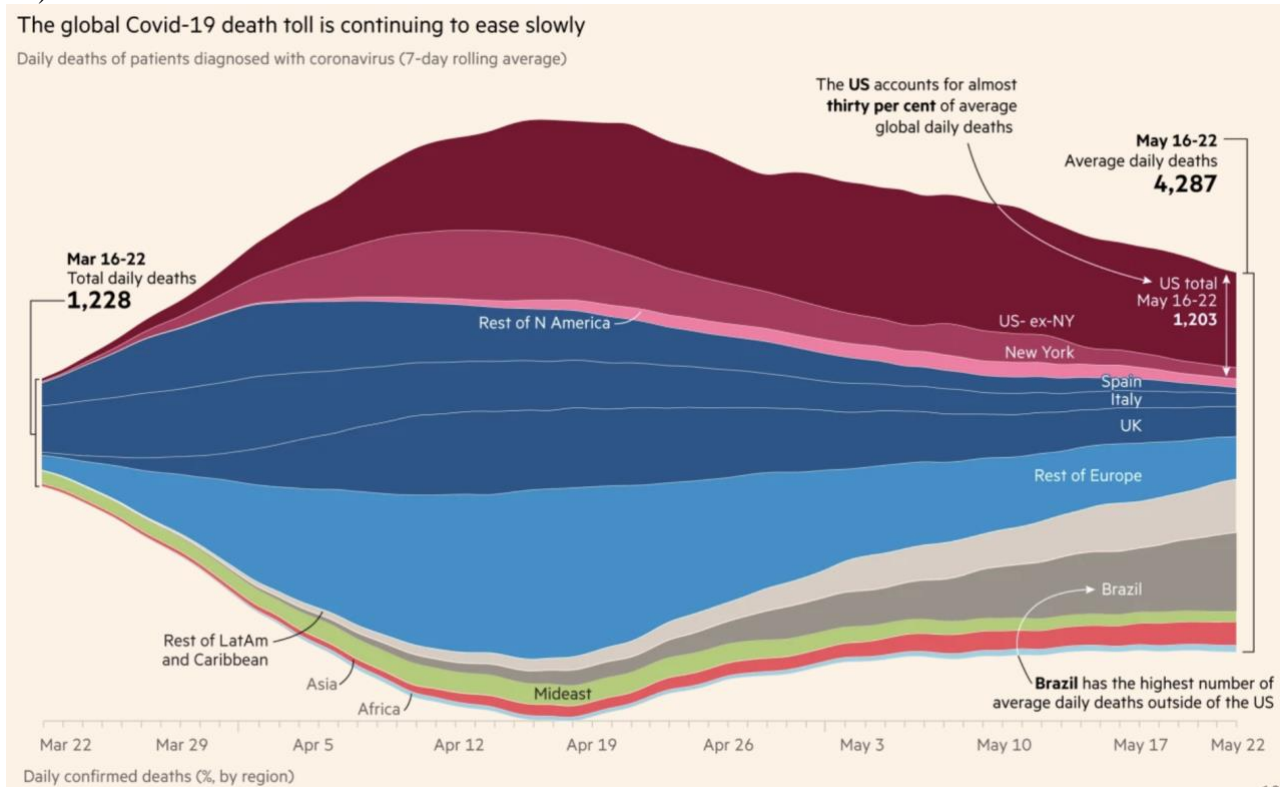
(a) You are asked to design a spatio-temporal visualization system to show the search trend about each health issue in different countries. You need to show both the spatial and temporal attributes of the data, and your solution can consist of multiple visualizations. Briefly justify your design choices. [10 marks]

(b) Now you are also given the top 10 search results returned for each search (e.g., the top 10 webpages returned by the search engine when people in a country searches related to Flu). Please design some text visualizations to summarize and analyze the text data. [10 marks]

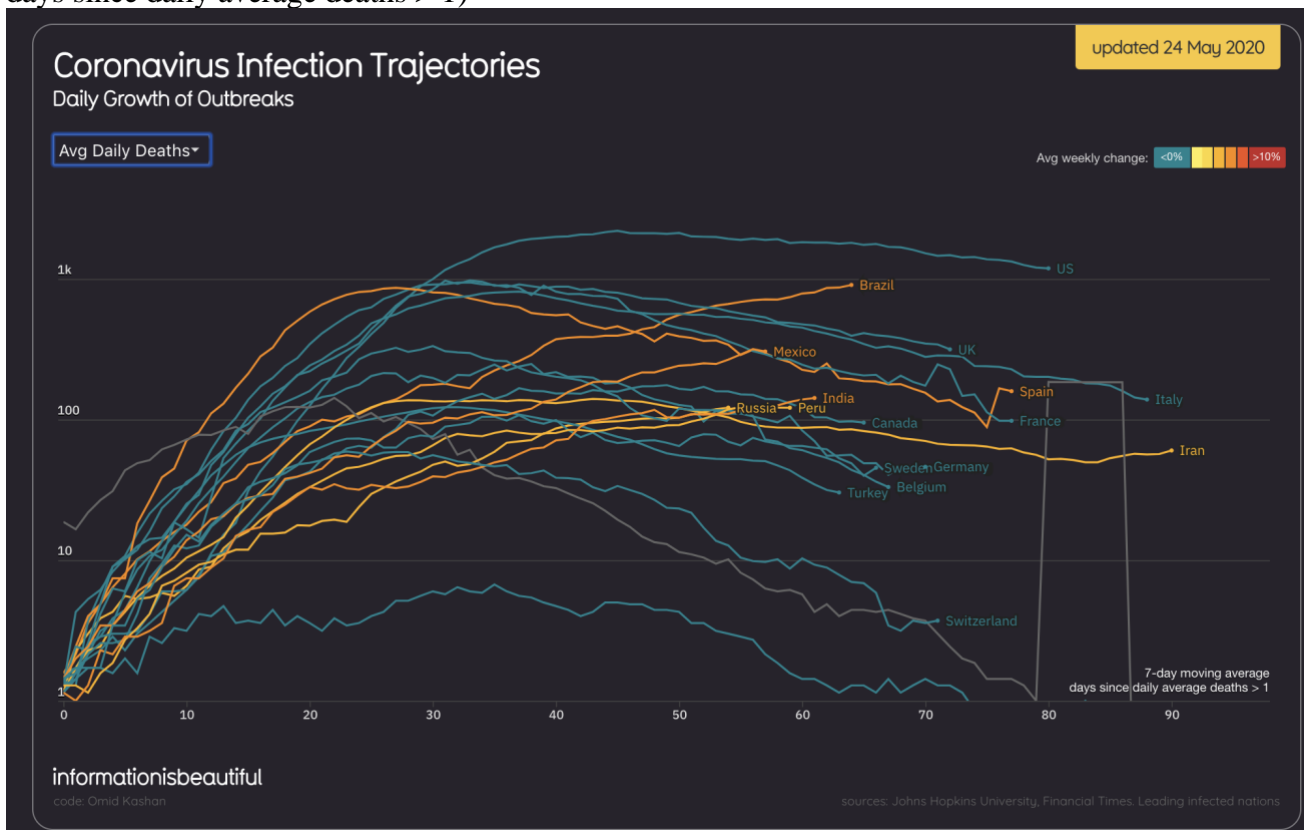
6. Visualization Evaluation [15 marks]

Two different websites use different visualizations to show the daily deaths of COVID-19 patients.

The “Financial Times” website uses the stacked graph. The X-axis is the absolute time (e.g., Mar 22)



The “Information is beautiful” website uses the line charts. The X-axis is the relative time (i.e., days since daily average deaths > 1)



(a) Based on your observation, what are the advantages and disadvantages of these two visualizations respectively? [5 marks]

(b). You are asked to design a controlled experiment to compare the effectiveness of these two visualizations. We hope the evaluation should be quantitative and as rigorous as possible. Please write down your detailed plan to conduct the evaluation. Please write down your hypotheses, the experiment scheme (i.e., within-subject or between-subject), the number of subjects you plan to recruit, the tasks you plan to give, the data you want to collect, and the analysis you want to perform on the collected data. [10 marks]