

The University of Nottingham

SCHOOL OF COMPUTER SCIENCE

A LEVEL 3 MODULE, SPRING SEMESTER 2016-2017

FUNDAMENTALS OF INFORMATION VISUALIZATION (G53FIV)

Time allowed TWO Hours

Candidates may complete the front cover of their answer book and sign their desk card but must NOT write anything else until the start of the examination period is announced

Answer all FOUR QUESTIONS

Marks available for sections of questions are shown in brackets in the right-hand margin

No calculators are permitted in this examination.

Dictionaries are not allowed with one exception. Those whose first language is not English may use a standard translation dictionary to translate between that language and English provided that neither language is the subject of this examination. Subject specific translation dictionaries are not permitted.

No electronic devices capable of storing and retrieving text, including electronic dictionaries, may be used.

DO NOT turn examination paper over until instructed to do so

ADDITIONAL MATERIAL: None

INFORMATION FOR INVIGILATORS:

Please ensure that **Exam Paper** and **Answer Book** are both submitted.

Question 1: Data, Visual Encoding and Charts (Overall 31 Marks)

Below is a table of student data.

	Name	Age	Grade	Course	Entry Year
Student 1	Tom	20	A	Math	1997
Student 2	Jim	19	B	Math	1998
Student 3	Mary	22	A-	Sport	1995
Student 4	Jane	21	B+	Art	1996

- a. Describe the three basic data types. Assess each column of the table above on the corresponding data type.
(4 Marks)
- b. There are two main ways of presenting multivariate data sets: tables and graphs. Describe the advantages and disadvantages of visualizing data with tables versus graphs?
(4 Marks)
- c. Mapping data attributes to visual features is a fundamental design decision in creating a visualization. Some visual features are better suited to some data types than others.
 - i. List and describe Jaques Bertin's seven visual encoding variables (semiology of graphics).
(4 Marks)
 - ii. For each basic data type in Q1(a), describe respectively one visual encoding that is appropriate and inappropriate to use. Give one example of how each appropriate visual encoding variable is used in a common visualization technique.
(5 Marks)
- d. Compare and contrast two common techniques for visualizing multivariate data: 2D Scatterplots and Parallel coordinates.
 - i. Explain 2D Scatterplots and Parallel coordinates.
(3 Marks)
 - ii. For each technique, identify the strengths and weaknesses in terms of the three following basic visualization tasks: (1) Find value of data case; (2) Find correlation; (3) Find outlier.
(6 Marks)
- e. Describe the limitations of using Bertin's semiology in visualizing multivariate data of over 8 dimensions. List and describe two common methods that are able to visualize high-dimensional multivariate data.
(5 Marks)

Question 2: Visual Perception and Interaction (Overall 21 Marks)

- a. Describe the two stages of perceptual human visual processing model.

(4 Marks)

- b. List a set of ten pre-attentive features. Give two examples of a conjunction of two features that may be pre-attentive.

(5 Marks)

- c. Describe three important Gestalt principles of visual organization and provide examples of how these are used in visualization techniques.

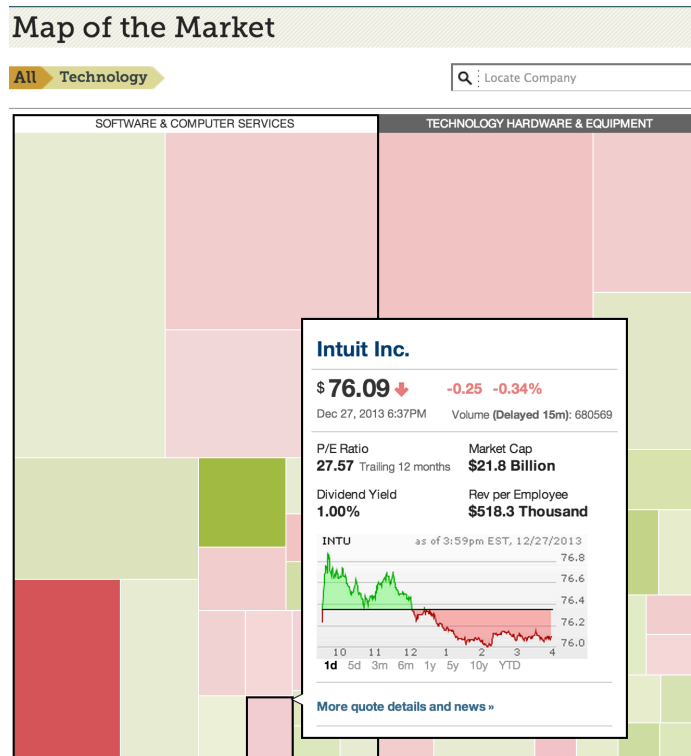
(6 Marks)

- d. Based on user intent, list the 7 categories of interactions for information visualization. Briefly explain three categories of those interactions with examples.

(6 Marks)

Question 3: Tree and Graph (Overall 22 Marks)

Below is a map of the stock market (covered in the G53FIV lectures), visualized using the treemap technique.



- Describe what treemap is. **(2 Marks)**
- List and explain all the visual encodings and their corresponding data types in this treemap visualization. **(5 Marks)**
- Describe the concept of "focus and context". Illustrate the benefits of this using the "map of the stock market" visualization as an example. **(4 Marks)**
- If there are too many companies in the stock market (e.g. >10000), what are the issues using the treemap visualization techniques? Explain the problem, and discuss the basic idea and two common techniques to alleviate this problem. **(5 Marks)**

- e. In addition to treemap, list and briefly describe two other common techniques for visualizing trees. Compare the pros and cons of using treemap against one of the other common tree representations.

(6 Marks)

Question 4: Data Manipulations, Text and Evaluation (Overall 26 Marks)

- a. Describe respectively the data structures for storing the homogeneous data and heterogeneous data in R. Explain the concept of "pipe" in R data processing.
(6 Marks)
- b. List and describe the five most common data manipulation techniques.
(5 Marks)
- c. Explain the challenges of visualizing text. Describe three common methods to visualize text and discuss when to use each.
(5 Marks)
- d. When dealing with text, describe one common algorithm to extract the most informative terms and discuss the advantages and limitations of that approach.
(4 Marks)
- e. Explain in general the methodologies that can be used to evaluate information visualization systems. Discuss the pros and cons of different types of approaches.
(6 Marks)