

Informatics 1: Data & Analysis

Lecture 20: Course Review

Ian Stark

School of Informatics
The University of Edinburgh

Tuesday 2 April 2013
Semester 2 Week 11



This is week 11, the last teaching week of Semester 2.

Your final tutorial for Inf1-DA is this week, in which you should receive back your work on the coursework assignment.

This is the last lecture of Inf1-DA.

- Exam arrangements and format
- Summary of course topics
- Guest talk by [Tom Macmichael](#) of [yourtaximeter.com](#)

Time and Place

Informatics 1: Data & Analysis will be assessed by a single two-hour written examination.

Date: Monday 29 April 2013

Time: 1430–1630

Place: St Leonard's Land Games Hall

This information for course code INFR08015 is current at 2013-04-01; please check the link on the Inf1-DA web page nearer to the date to check this and to confirm all of your exams.

Exam Format

As in previous years, the exam will have three compulsory questions.

- Read all questions before beginning the paper
- You don't need to do the questions in order
- Don't assume a question is only using one part of the course
- If you get stuck on one question: don't waste too much time on it; do go on to the next question; and don't give up!

Calculators are permitted, and will be provided at the exam hall. These are a standard scientific model: you can try one out at the ITO if you wish. Questions will only involve simple arithmetic, and you may not find it necessary to use a calculator at all.

Past Exam Papers

Past papers are held online by the University Library, and you can access them through links from the course web page.

These are a good source of revision material, and I strongly recommend you attempt questions from as many of these as you can. However:

- Although the overall format of questions remains similar, the exact topics covered change from year to year.
- Where online “solutions” are provided, they may not always be correct and may not provide information you require. (Most of these were written as guides for external academics reviewing the paper, not as model answers for students.)
- If you are puzzled by a past question, ask on the discussion forum, email your tutor, or email me.
- There have been changes to the course content over the years, so not all past exam questions are relevant.

Which Past Exam Questions?

Past Papers

In each year there are exams from the main and resit diet. The following questions are relevant to the current course syllabus.

Informatics 1B 2005: Questions 1 and 2

Informatics 1B - D&A 2006, 2007: Questions 4 and 5

2008: Everything except 2(c) on [XQuery](#)

2009, 2010, 2011, 2012: All questions

Examinable Material

Unless otherwise specified, all of the following material is examinable:

- Topics covered in lectures
- Directed reading distributed in lectures
- Topics covered in the weekly exercise sheets

Review of Topics

The entity-relationship model, ER diagrams. The relational model, SQL DDL. Translating an ER model into a relational one. Relational algebra, tuple-relational calculus, SQL queries.

Semistructured data models and the XPath data tree. XML documents. Schema languages and DTDs. Relational data converted into XML. XPath as a query language.

Corpora: what they are and how they are made; examples. Annotations and tagging. Concordances, frequencies, n-grams, collocations. Methods for machine translation.

Information retrieval: what it is, evaluating and comparing performance of IR systems; the vector space model and cosine similarity measure.

Data scales, summary statistics, population vs. sample; hypothesis testing and significance; correlation coefficient, χ^2 test.

Some Specific Items

Corpora

In general it is the **principles** of corpora that are examinable, rather than the precise details of individual corpora. Similarly, you should be familiar with the principles underlying POS-tagging and syntactic annotation, but you do not need to know detailed linguistics or specific tag sets.

You should however, be able to give examples of a corpus or a POS tag.

The CQP tool was used in a tutorial, so is examinable — although again for general principles and use, not every detail of syntax.

Statistics

You are not expected to memorize critical value tables; however, you should be able to use one if provided.

You are expected to know the formulas for the various statistics used, and to be able to calculate with them.

Anything Else?

If you have further questions about the course material, past lectures, exercises, the exam, or anything else, please:

- Post a question on the Inf1 discussion forum; *or*
- Ask your course tutor, in person or by email; *or*
- Ask me, in person or by email.

Course Survey

Please complete the online survey for this course. This is anonymous, and I read all the submissions. Ideally, do it immediately after this lecture.

Go to **<http://www.inf.ed.ac.uk/teaching>** and follow the link to the **Semester 2 Course Survey**.