**INM402 / IN3030 Data Visualization Coursework**

Note the instructions below refer to producing data visualization using the standard 'java' flavour of Processing. If anyone would like to use an alternative technology, such as[p5.js](http://p5js.org/), [processing.py](http://py.processing.org/) or [D3](http://d3js.org/) you should **contact me first to approve the technology**. Regardless of the technology you use, the assessment outcomes and marking scheme below will apply to all submissions.

**Aims and Learning Outcomes**

The aim of the assessment is for you to demonstrate the skills required to produce an effective data visualization sketch and evidence that you can identify and apply good practice in visualization design. You should aim to show you can *abstract a data visualization task* and *validate your visualization design*. Specifically, on successful completion of the assessment you will have:

* identified (a) *research question(s)* about a real dataset that can be answered with data visualization;
* built a working *data visualization* that represents a real dataset;
* applied good practice in the *visual encoding* and *interaction* in your design;
* provided a *validation* (justification) of your design;
* provided some *insight* into the data you have visualized (*postgraduate students only*).

**The Task**

The single assessed piece of coursework for this module involves you choosing one or more data sources and creating a sketch in Processing to visualize the data. You will also provide a design justification for your visualization validating the design decisions you made in creating your visualization.

**Choosing a data source**

You may choose any dataset from any source and you are encouraged to look widely for possible sources. A good place to start is the [Guardian Data Blog](http://www.theguardian.com/news/datablog/interactive/2013/jan/14/all-our-datasets-index). There are many hundreds of datasets available so you should spend some time carefully considering the task and data that interest you. The following guidance may help:

* Choose a dataset that is sufficiently complex that it will benefit from exploration though interactive data visualization (e.g. a table of 10 numbers would probably not need anything more than a simple Excel chart)
* Choose a dataset which will help you answer a motivating *research question*. Examples of research questions could be *What are the patterns of criminal behaviour and do they show changes over time?* or *Is there any structure to the text in this collection of documents that suggest they have something in common?*.
* Avoid choosing datasets that are very difficult to parse either because they are too large, too complex or are not in a form amenable to parsing in Processing.
* You may choose to link more than one dataset if that helps you answer your motivating research question. There is no minumum or maxumum number of datasets you should chose.

**Creating your Processing Sketch**

You should create your Processing sketch in much the same way as you have been for the non-assessed data challenges. You will not be assessed on the quality of your code, but rather the design and effectiveness of the resulting visualization. You may use other libraries or PDEs in your sketch (e.g. giCentreUtils, geoMap, ControlP5 etc.) if you wish, but you*must credit any other code that you have not written yourself with clear comments in your code*.

If you find that you cannot implement everything you wish to in your design, your submission may include a description and/or mockup of what you intended. If you do so, your submission should make clear what is implemented and what is a mockup.

**Completing the Justification Document**

You must use the file **courseworkTemplate.doc** as the basis for your justification document. This outlines the structure of your document which must be no longer than 5 pages (postgraduates) or 4 pages (undergraduates) inclusive. Do not change the formatting or margins in order to fit in more text. You should save a local copy of the document with the name **Validation*MySurname*.doc** (substituting *MySurname* with you own).

**Submission Instructions**

You should submit **a single .zip file** file (do not use other archive files such as .rar) containing the following to the *Coursework Submission Area* on Moodle no later than **Sunday 1st May, 2016, 5:00pm**:

1. Your working Processing sketch including all relevant PDE files, the data directory containing the data you are visualizing and any other relevant files, such as font (.vlw) files, additional libraries required to allow your sketch to run (if you use either the giCentreUtils or geoMap libraries provided by the giCentre, there is no need to include these with your submission).
2. A PDF document no longer than *4 pages (undergraduates)* or *5 pages (postgraduates)* using the template given above outlining your design validation and, for postgraduate students only, data insights.

Late submissions will not be marked, so you are very strongly advised to plan submitting your work well in advance of the deadline. Any material in the justification document beyond the page limit will not be marked.

You are encouraged to support each other in terms of general approaches to data visualization and the use of Processing, but you must not discuss or share details of your own coursework with any other students - this is an individual piece of assessment.

You will receive written feedback on your work within 4 weeks of the coursework deadline.

**What do I have to do to pass / get a good mark?**

In marking your work I will be looking for evidence that you have met the aims and outcomes at the top of this page. Specifically, you will get credit for the following:

* identifying (a) research question(s) that can be answered effectively through data visualization taking advantage of both the 'human in the loop' and 'computer in the loop'.
* building a working data visualization sketch that provides some insight into the data you are representing and the research questions(s) you are answering
* demonstrating that you have considered good practice in the design of your data visualization
* providing insight into the data demonstrating that visualization has helped you to generate this insight (*postgraduate students only*)

A submission will **pass** if you have demonstrated a design for a data visualization that answers your research question(s) in a way that would not have otherwise been immediately obvious by looking at your chosen data directly or could have been shown with some simple charts or calculations in spreadsheet such as Excel. Your validation must show some evidence you have considered good practice in the visual encoding and interaction design of your visualization.

For a submission to get marks in the **distinction / first class range** (70% or more), it should at least do everything as above, but also be able to draw on good practice in visualization design from ***literature*** mentioned in the lecture notes or from other external sources. It should demonstrate a clear link between your *task abstraction* and your*validation*. It should demonstrate some sophistication in its design and implementation that clearly goes beyond the examples provided in the lecture materials. It should reveal complex patterns in the data in a manner that is easy to interpret and that directly help in answering your research questions. You should be able to demonstrate how your own design has assisted in answering those questions.

For some inspiration you may wish to look at some of the video summaries of coursework submissions from previous years. **Under no circumstances must you make direct use of any other student's coursework. Any submissions that appear similar to those submitted by others will be investigated for possible academic misconduct.**