

Problem Sheet Three — Practical Assignment

“Parallel Coordinates Plots”

Recommended Hours: 20 hours

Initial Assessment: either practical session in Week 5

Submission Deadline: 12:00 noon, Monday, 20 February 2017

Guest Lecturer: Alfie Abdul Rahman

Lab. supervisors and Assessors: Alfie Abdul Rahman and Mr. Devesh Batra

Summary

You are required to implement one or more parallel coordinates plotting tools to meet the requirements at one of the three levels specified below. The three levels are associated with grades **S-**, **S** and **S+** respectively, and the grading will depend on the quality of your software as well as that of your paper-based report. In other words, attempting Level 3 does not guarantee to gain a grade **S+**, and so on. If there is no submission from a student, the grade **NS** will be given. Nevertheless, the grade of this assignment does not count towards the final mark.

You are allowed to build your tools based on existing examples in the D3.js repository, though you must acknowledge the relevant examples explicitly in your source code and paper-based report. The grading scale has already taken into account the availability of such examples, while recognising the effort needed to adapt them for this assignment. However, no student should make use of others (e.g., online coding service) to complete their assignment, or take programs from another student.

The recommended programming environment is D3.js, for which two lectures are scheduled in Week 3. In addition, six laboratory sessions are scheduled in Weeks 3, 4, and 5. Students will be divided into two groups, and each student is expected to attend 3 laboratory sessions, one in each week. If you wish to use a different programming environment instead of D3.js, please let the course lecturer know in advance.

Students in both groups (Tuesday and Wednesday Sessions) will submit their work to Dr. Alfie Abdul Rahman. Please show your software tools to your assessor during a practical session in Week 5, before submitting the paper-based report on Monday 20 February. During your demo, please let your assessor to have an interim cover page with your name, and the levels that you have attempted. Your assessor will write some brief notes on this cover page. Your paper-based report should include the followings:

- The interim cover page with your assessor’s notes.
- A summary (no more than 1 A4 page), which outlines the level that you have attempted, the existing examples in the D3.js repository that you have adapted, the functionality that you have successfully completed, and the estimated number of hours that you spent on this assignment (for help us balancing the workload in future years).
- One or more screenshots (no more than 2 A4 pages) to demonstrate the visualization results of your work.

- Source code with essential inline documentation.

The assessors will return marked assignment to you in Week 7. You can then remove the interim cover page and submit your marked assignment to the Department using a provided envelope.

LEVEL 1: Simple Adaptation

The software tool at this level should be able to display the tabular data in the spreadsheet of **WorldEconomy2015.xlsx** using a parallel coordinates plot. The table includes seven columns with 19 measured values in each column, that is, seven 19-value data series for the seven variables. The variables are:

- **Region:** Codes for different regions
- **H-Exp:** Household consumption expenditure
- **G-Exp:** General government final consumption expenditure
- **GCF:** Gross capital formation (GCF)
- **Exports:** Exports of goods and services
- **Imports:** Imports of goods and services
- **GDP:** Gross Domestic Product (GDP)

Note that the region codes are not the standard UN codes. The numbers are chosen to provide appropriate grouping of sub-regions in each continent.

Your software must have the following essential functionality:

- Display a line for each of the 19 regions intersecting with the 7 axes at correct places.
- Display labels for each axis.
- Provide a brushing utility.

LEVEL 2: Competent Tool Development

You must complete LEVEL 1 before attempting LEVEL 2.

The software tool at this level should be able to display the tabular data in three **Population** spreadsheets using a parallel coordinates plot. You are required to use **PopulationAmeN.xlsx** (i.e., your Dataset 1) plus two other spreadsheets based on the first letter of your first name and the first letter of your surname. For example, the two letters for John Smith is J and S. The allocation is given in the following table:

1st letter of first name	Dataset 2	1st letter of surname	Dataset 3
A, J, S	PopulationAfrE.xlsx	A, J, S	PopulationAsiC.xlsx
B, K, T	PopulationAfrM.xlsx	B, K, T	PopulationAsiE.xlsx
C, L, U	PopulationAfrN.xlsx	C, L, U	PopulationAsiS.xlsx
D, M, V	PopulationAfrS.xlsx	D, M, V	PopulationAsiSE.xlsx
E, N, W	PopulationAfrW.xlsx	E, N, W	PopulationAsiW.xlsx
F, O, X	PopulationEurE.xlsx	F, O, X	PopulationAmeC.xlsx
G, P, Y	PopulationEurN.xlsx	G, P, Y	PopulationAmeCa.xlsx
H, Q, Z	PopulationEurS.xlsx	H, Q, Z	PopulationAmeS.xlsx
I, R	PopulationEurW.xlsx	I, R	PopulationOcean.xlsx

The datasets are formulated based on the 2015 UN population statistics. Each dataset has 11 columns:

- **Region:** The abbreviation for a region in the world
- **Country or SR:** The name of a country or a special region
- **P-Both:** Population of both genders (in thousands)
- **P-Male:** Male population (in thousands)
- **P-Female:** Female population (in thousands)
- **Rate:** Population growth rate (between 2010 and 2015)
- **RNPI:** Rate of natural population increases (per 1000, between 2010 and 2015)
- **Mdn-Age:** Median age of the population
- **Density:** the number of people per square km
- **Child-D:** Child dependency ratio (ratio of population aged 0-19 per 100 population 20-64)
- **Old-D:** Old-age dependency ratio (ratio of population aged 65+ per 100 population 20-64)

Your software must have the following essential functionality:

- Display a line for each of the countries (or special regions) in your three datasets, intersecting with the 11 axes at correct places.
- Colour-code lines based on their regions (i.e., one colour per dataset).
- Display labels for each axis.
- Provide a brushing utility, and an axis-swapping utility.

LEVEL 3: Problem Solving with a Personal Design

You must complete LEVEL 2 before attempting LEVEL 3.

The software tool at this level should be able to display the tabular data in the spreadsheet of **WorldEconomy1990-2015.xlsx** using a specially-designed parallel coordinates plot. This is a time-varying dataset with six tables. Each table consists of a dataset similar to the one at LEVEL 1, except that they are for 1990, 1995, 2000, 2005, 2010, and 2015 respectively.

There are no specific requirements at this level, except that your software should try to help users visualize the temporal relationships among data values within each column (i.e., the same concept in different years), as well as the relationships among different columns (i.e., different concepts in the same year). For example, you may choose to show 6 parallel coordinates plots simultaneously, or associate each axis with a time series plot or bar chart. Please beware that programming slightly complex interactions may take some time to get them right.

In order to make data at different years comparable, it is recommended to set the value range for each axis based on its global MIN and MAX values (during 1990 ~ 2015).