KTH Stockholm
CSC :: CST
Visualization, Autumn 2017, DD2257
Tino Weinkauf

Homework assignment No. 02

Due September 14, 2017

Download the zip-File with the code and place the contents in the same folder as the folder from the introductory coding tutorial. Follow the same process to activate the `IVW_MODULE_DD2257LAB1` module in CMake and build Inviwo.

For each of the tasks in this assignment there exists a workspace which can be found under *File* → *Example Workspaces* → *DD2257Lab0* in the Inviwo application or in the `dd2257lab1/data/workspaces` folder. For each task you will modify one Inviwo processor. The skeleton code for these along with additional instructions as comments and code snippets are all located directly in the `dd2257lab1/`. Files in other location do not need to be modified.

Task 2.1: Scatter Plot **5 P**

Create a scatter plot for two selected dimensions of a given data set. Your task is to modify the skeleton (in `scatterplot.cpp`) for the scatter plot processor so that it places data points and coordinate axes based on data, resulting in a similar image to the one in Figure 1.

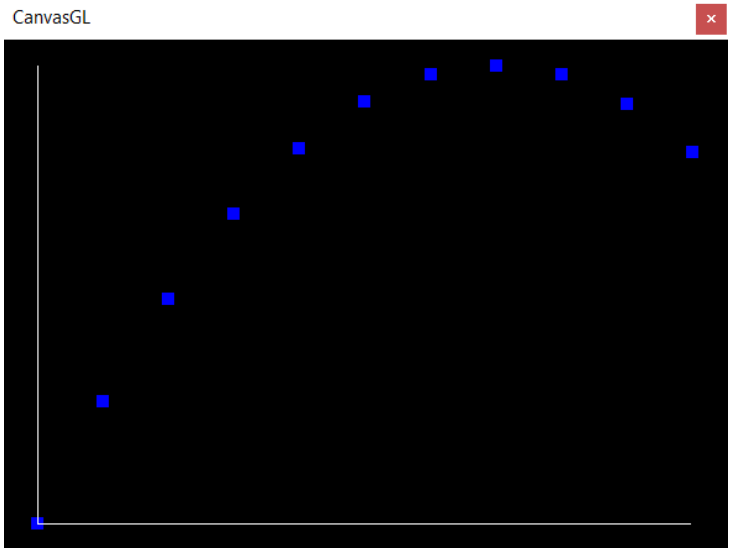

Figure 1: Result after loading the *Anscome's Quartet* data set and selecting *x2* and *y2* as the dimensions to visualize.

The network in the example workspace `scatterplot.inv` loads a given CSV file and creates an Inviwo `DataFrame` from it. CSV means *comma separated values*. CSV files are text files where each line represents a different observation point. Within a line, the dimensions are separated by commas. Here is an part of the *cars.csv* data set example:

```
MPG,Cylinders,Displacement,Horsepower,Weight,Acceleration,Year,Origin
18,8,307,130,3504,12,1970,1
15,8,350,165,3693,11.5,1970,1
```

The file starts with a header line, which describes each dimension. This is the only line in which you have alphabetic characters. All following lines contain only numbers and commas or another separation character.

This assignment comes with three data sets. The *cars.csv* data set and two sets of data sets where each have the same statistical properties, i.e. the same mean and standard deviation, but are clearly different and visually distinct: the Anscome's Quartet and the Datasaurus Dozen (see `http://www.autodeskresearch.com/publications/samestats` for more details). These are combined in one file each, with two columns belonging to one data set each. Use these to test your code. However, we may also load other data sets during the interview (with a different number of dimensions and observation points). You can change the loaded data set by specifying its Path in the properties of the *CSV Source Processor*. For files with more than two columns, you can choose which columns or dimensions to display within the properties at the *Scatter Plot Processor*.

The processor is supposed to create two meshes, on point mesh for the data points and one line mesh for the coordinate axes. The skeleton code for the processor already creates one line and one point. Note that Renderer following the Scatter plot processor displays only a certain part of the domain.

Task 2.2: Parallel Coordinates Plot **7 P**

Display all dimensions of a given data set in a parallel coordinates plot such as the one in Figure 2. Use the same order for the axes as it is given in the file (and thus the incoming `DataFrame`). You need to modify `parallelcoordinates.cpp`. The corresponding workspace is `parallelcoordinates.inv`) and has a similar network as the scatter plot workspace.

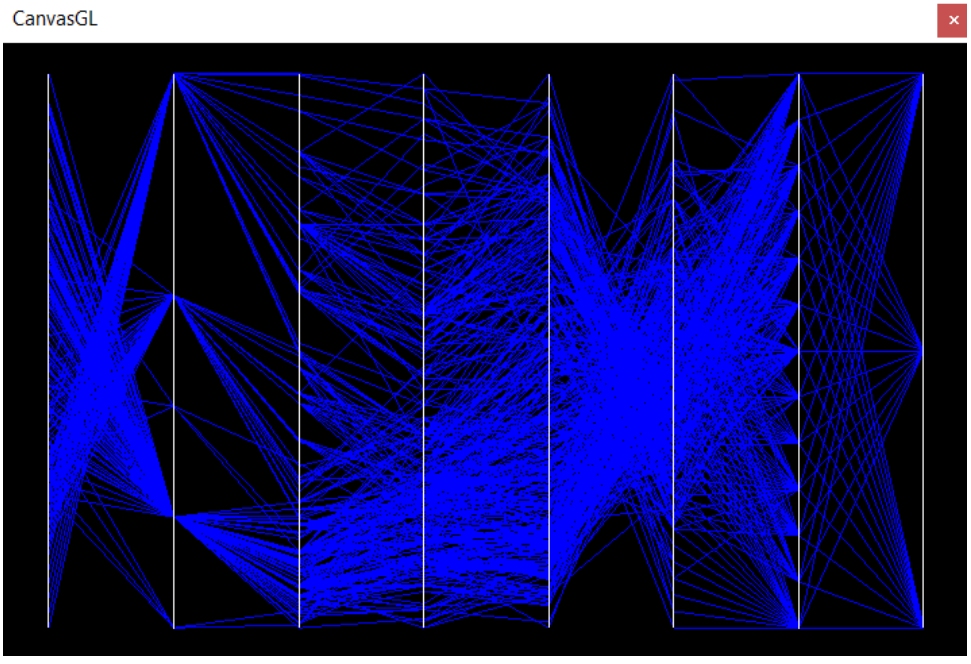

Figure 2: Result after loading the *Cars* data set.

Task 2.3: Dualities for Parallel Coordinates **2+4+4 P**

The `dualities.inv` workspace shows a scatter plot and a corresponding parallel coordinates plot side-by-side. Here, both plots show the same data. Your task is to create data of certain shape in order to investigate dualities (see Figure 3) by extending `generate2ddata.cpp`.

(a) Create data that forms a straight line with a negative slope in the scatter plot.

(b) Create data that forms a circle in the scatter plot.

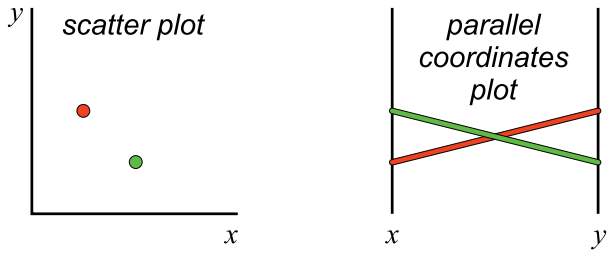

Figure 3: Point-line duality in scatter plot and parallel coordinates plot.

(c) Create data that forms a hyperbola in the scatter plot.
Hint: It is best to use the hyperbolic sine and hyperbolic cosine functions to create the hyperbola as follows:

$$\begin{aligned} x &= a\cosh\mu \\ y &= b\sinh\mu \end{aligned}$$

where a, b allow to adjust the shape of the hyperbola. Good starting values are $a = b = 1$. Furthermore, choose $-3 \leq \mu \leq 3$ with a high number of sample points.