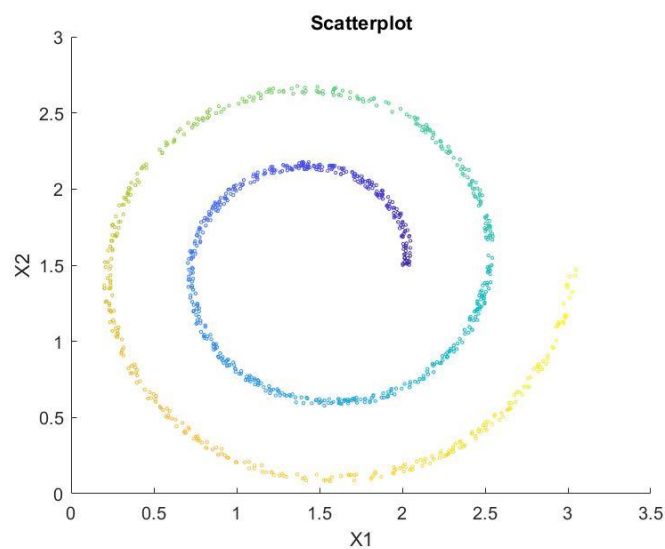


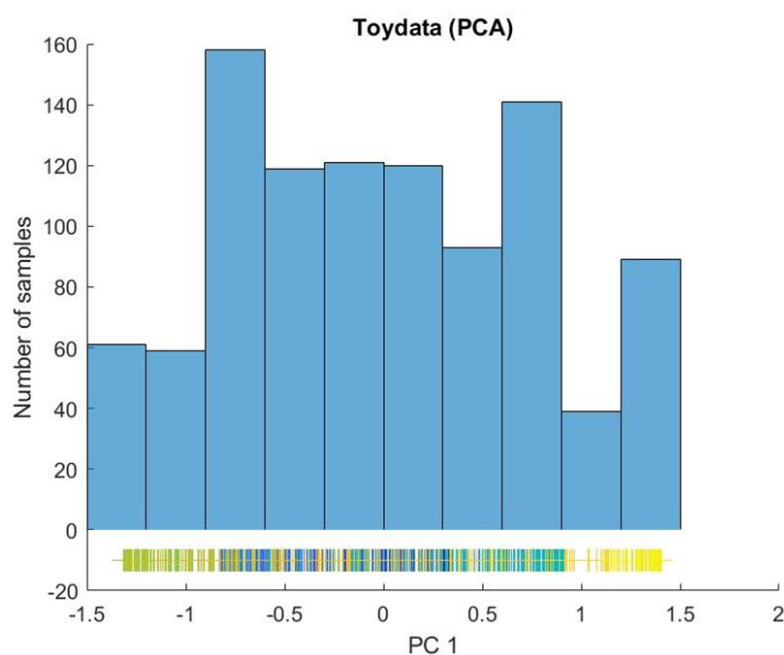
Information visualization exercise 3

Task 1

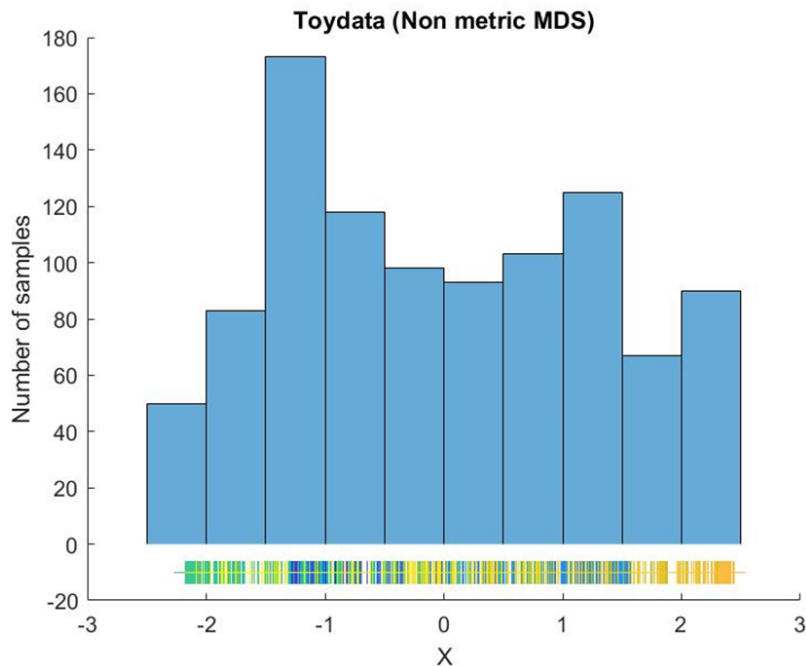
- a) The given dataset 'toydata' shows a two-dimensional spiral in a scatterplot. The color of each point represent the value of the given vector 's'.



- b) The is mixed up, so PCA is not suitable for this dataset. function Matlab function 'pca' calculate all principle components for the given dataset 'toydata'. The first column of the calculated matrix shows the first principle component. The data has to be centered because PCA calculate a covariance matrix for the eigendecomposition. The density and color of the representation is mixed up, so PCA is not suitable for this dataset.



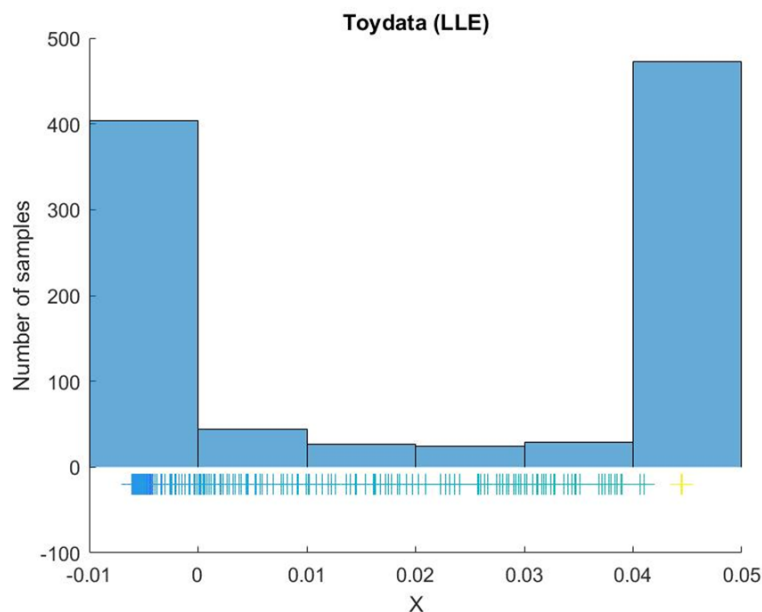
- c) 'mdscale' is an inbuilt Matlab function as well and calculate the non-metric MDS. For calculating the dissimilarities the function 'pdist' has been used on the zscores of the data. The color and density of the representation is mixed up, in this way non metric MDS is not suitable for this data set.



- d) For performing the LLE the function 'lle' of the Matlab Toolbox for Dimensionality Reduction has been used.

The toolbox can be obtained from: <http://homepage.tudelft.nl/19j49> .

The colors of the datapoint is on order, the density is quite shifted near to zero and to the 'borders', but the representation of the datapoints to each other is good.

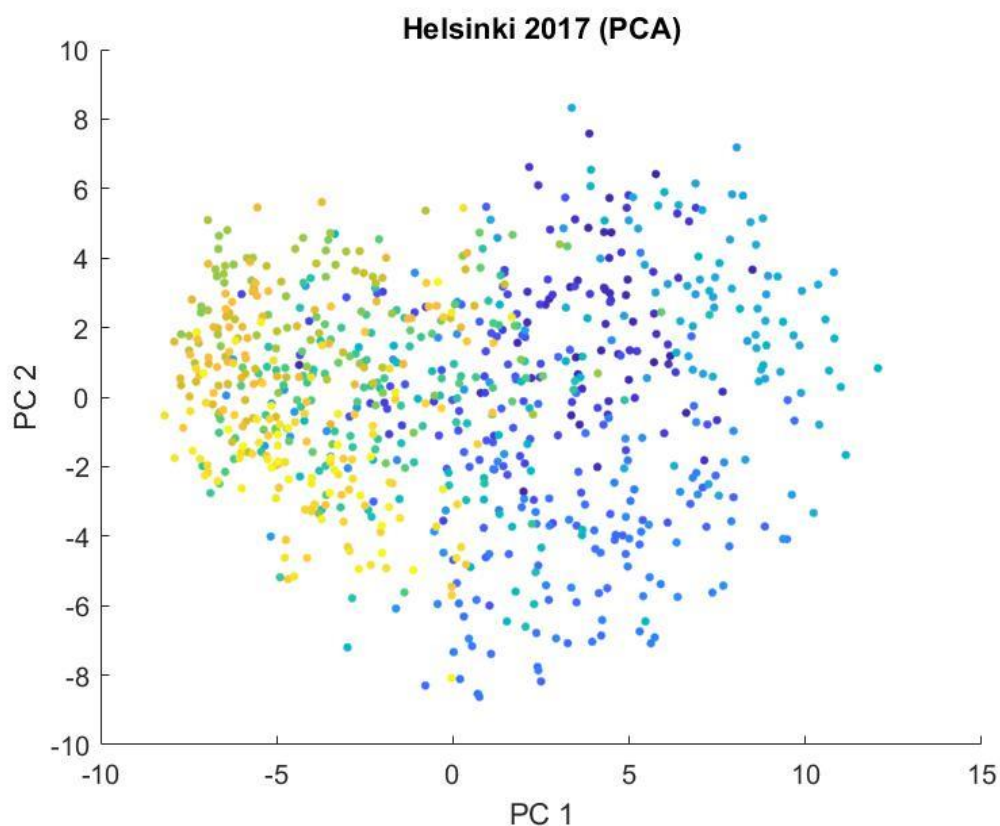


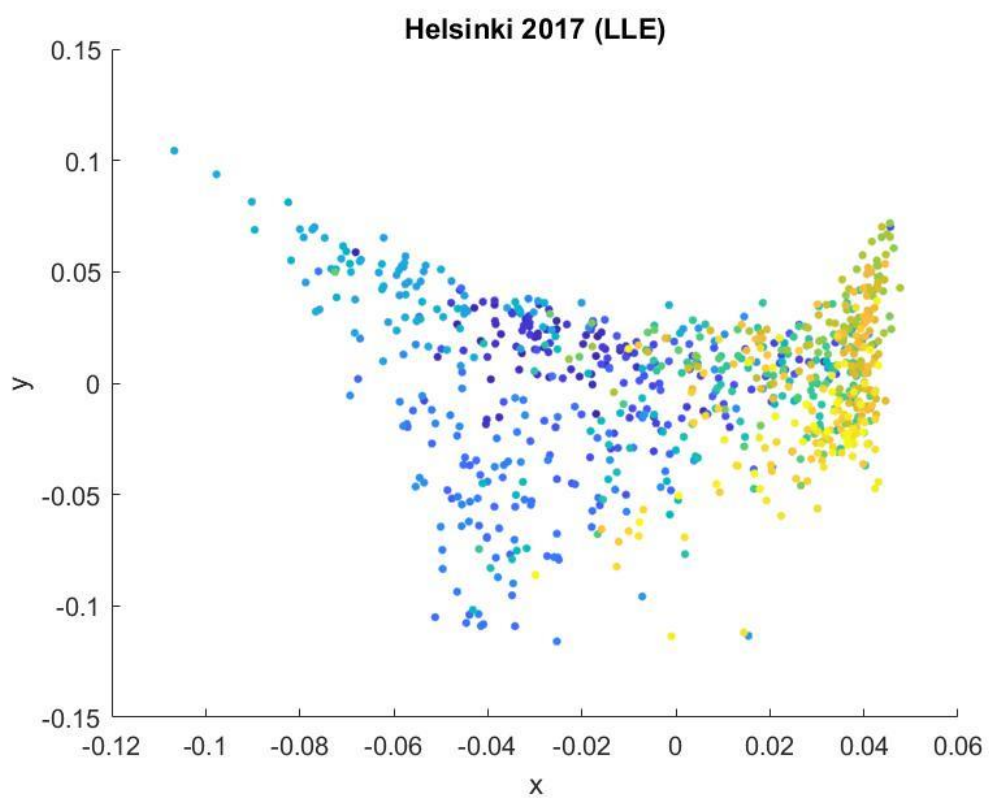
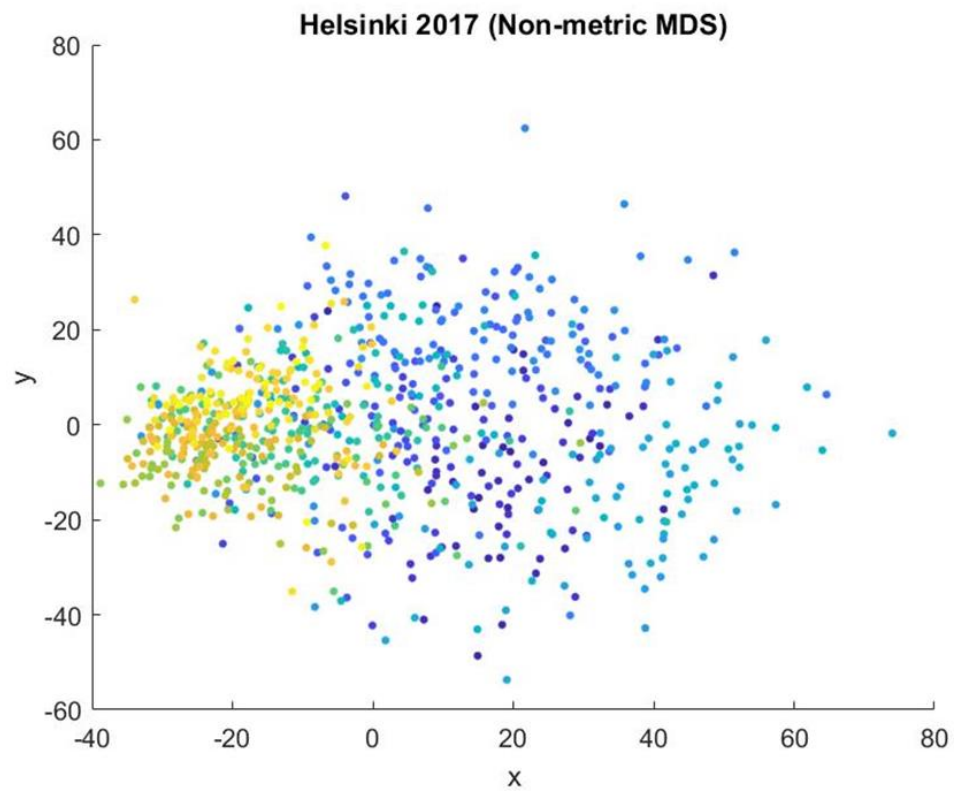
Task 2

The two dimensional representations of the given candidate's position data in Helsinki 2017 are shown for three different methods PCA, non-metric MDS and LLE. All three representations are very similar to the presented data of Espoo 2017 in the lecture.

For the PCA and the non-metric MDS representations the parties kd and kok (here blue) are located on the right, the parties rkp and sdg (here green) more in the center and the parties vihr and skp (here yellow) left. Compared to each other the PCA and the non-metric MDS are more or less similar. They are based on a similar principle of projecting the data due to its variance in euclidean distances (PCA) or dissimilarities and monotonic transform (Non metric MDS). But the iterative algorithm of non metric MDS is more versatile mapping technique than PCA.

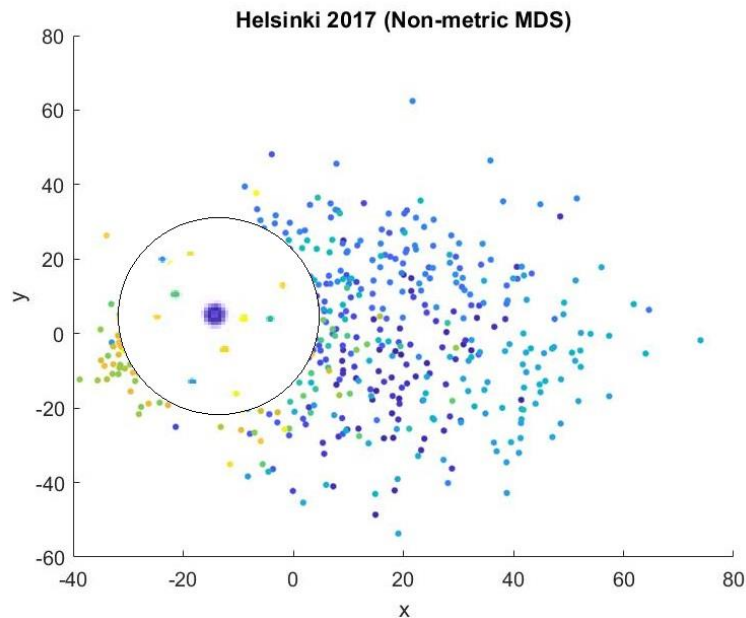
The LLE shows a small and dense band of data points near to zero as a representation (see small y- and x ticks on the axes). The parties in this band are ordered also like in Espoo. LLE is a non-linear technique and finds the nearest neighbors for each point. It computes weights which fit the point as a linear combination of its neighbors best.





Task 3

For a good visualisation of one specific candidate Furnas' Fisheye theory could be used. A left mouse click on the presented dataset of the Helsinki Election 2017 could open a focused circle in the region of interest. In this way two levels of detail can be shown in one scatterplot.



If the mouse is kept on a candidate for a specific time (e.g. 2 seconds) an additional textbox will appear which contains further information of the candidate as a third level of detail. In addition to that, comments of the candidates on the chosen values could be useful for distinguishing among candidates with a similar location in the plot. In this way, the citizen could understand the position of the chosen candidate on their favourite topic more and compare it to other candidates. But in this dataset no comments are available.

