Project title: Extending Delauny triangulations for density estimation

Supervisors: Dr.ir. Mart Janssen (Sanquin Blood Supply Foundation)
Dr. Matthijs van Leeuwen (LIACS)

Short description:
A functional approximation of multivariate data has many advantages for simulation as well as for epidemiological studies comparing patient populations and outcomes of clinical studies. Despite the fact that some methods for deriving functional approximations exist, no satisfactory solution has been found for many applications. An elegant and successful attempt using Delauny triangulations was recently published.[1] However, for practical applications an extension of this work is required. For instance, the visualization of a one- or two-dimensional density estimate of multi-dimensional data requires a projection of multi-dimensional objects on a lower dimensional subspace. As the complexity of such projections increases with the dimensionality of the dataset, efficient methods for the calculation of such projections need to be developed. Also, to obtain correct representations the multi-dimensional descriptions obtained by Delaunay triangulations require a constrained solution. In addition, the space enclosed by the data needs local (data driven) extensions. When coding for the visualization and extension of the triangulation method is completed, the overlap between objects from different datasets needs to be assessed.[2] This requires computation of multi-dimensional intersection objects. Again, efficient algorithms need to be developed to allow for a practical implementation.
Once intersectional objects can be calculated, a re-assessment of various epidemiological studies on publicly available and not publicly available datasets will be performed and their outcomes compared to the outcomes obtained with the novel approach.

Required skills
- Programming in R / Python / C++
- Interest in computational geometry
- Interest in optimization of algorithms for multi-dimensional geometrical (linear) objects
- Good communication and presentation skills

References
1. Liu, Y. & Yin, G. Nonparametric Functional Approximation with Delaunay Triangulation. *ArXiv190600350 Cs Stat* (2019). https://export.arxiv.org/pdf/1906.00350
2. Dirksen, H. Sections of the regular simplex - Volume formulas and estimates. *ArXiv150906408 Math* (2016) doi:10.1002/mana201600109. https://arxiv.org/pdf/1509.06408.pdf

For more information, please contact dr.ir. Mart Janssen (m.janssen@sanquin.nl, 06-50 25 14 33).

Extending Delauny triangulations for density estimation

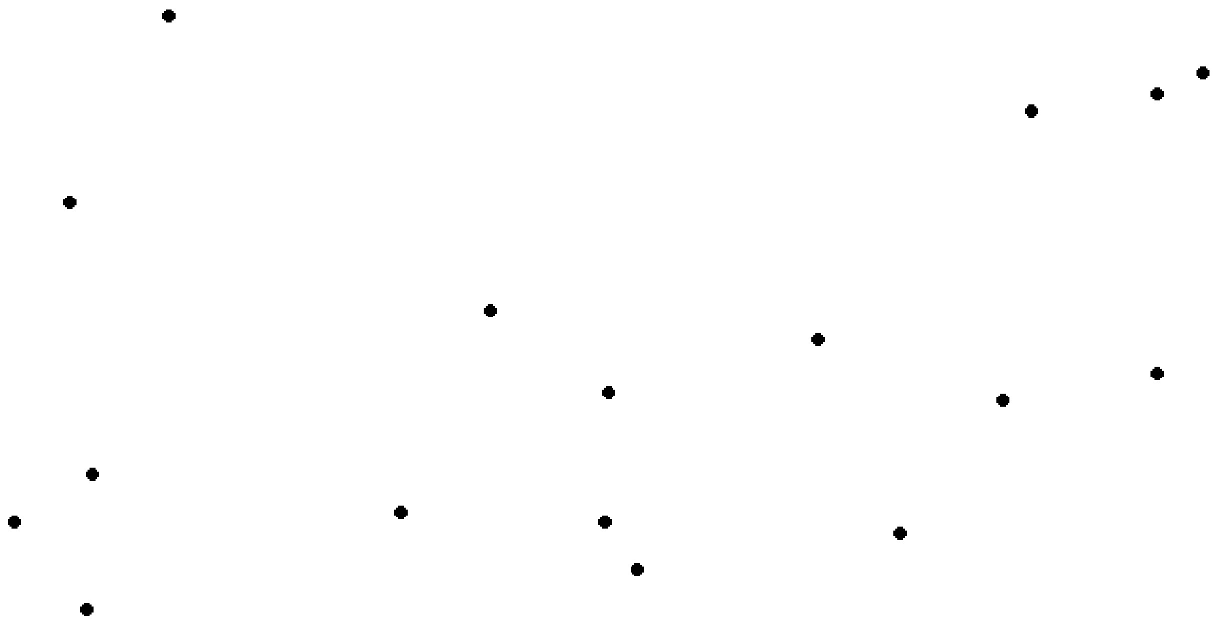
Mart Janssen, 12-01-2021

Sanquin

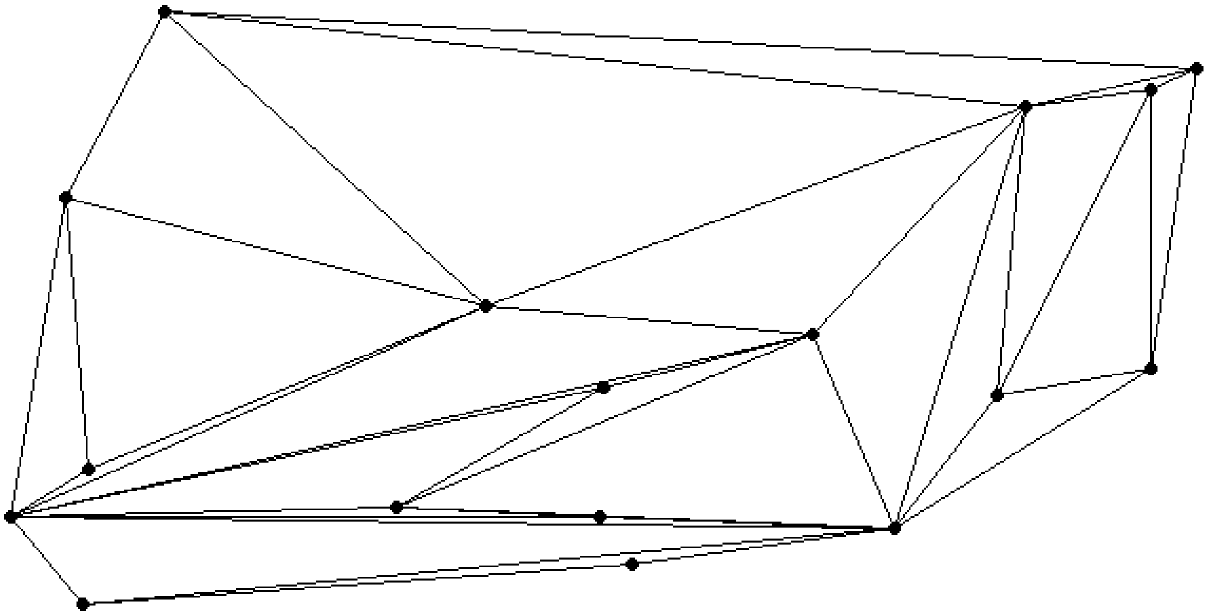
Functional data representations have some useful advantages

- Share data without compromising privacy
- Can be used to generate synthetic data
- Additional means for data analysis

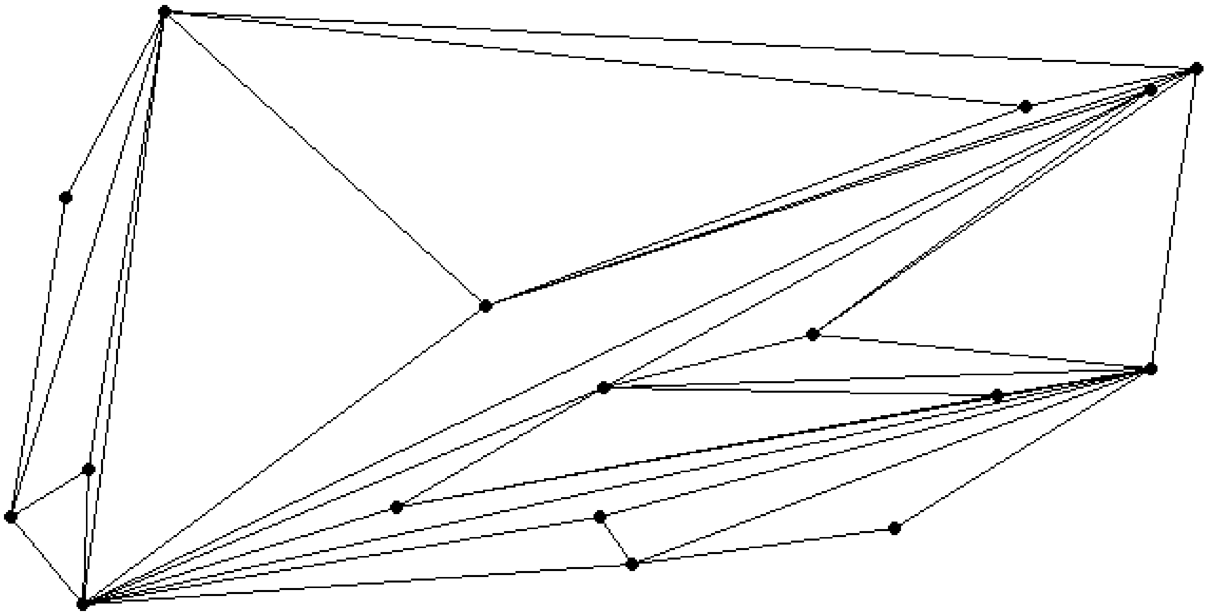
The most simple approach is linear interpolation



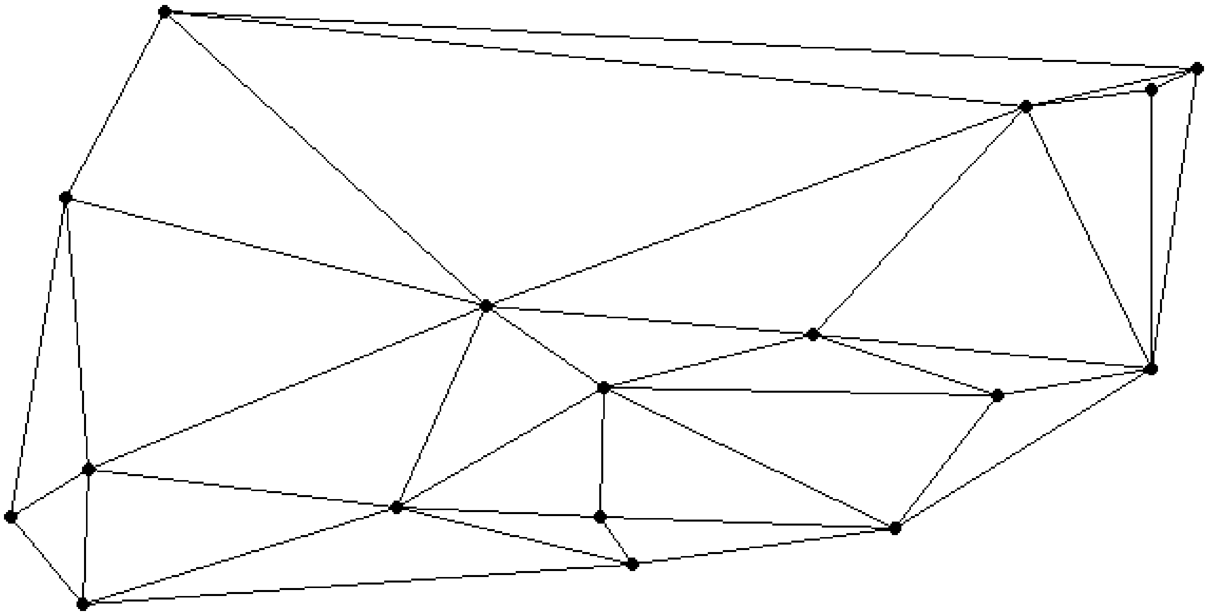
The most simple approach is linear interpolation



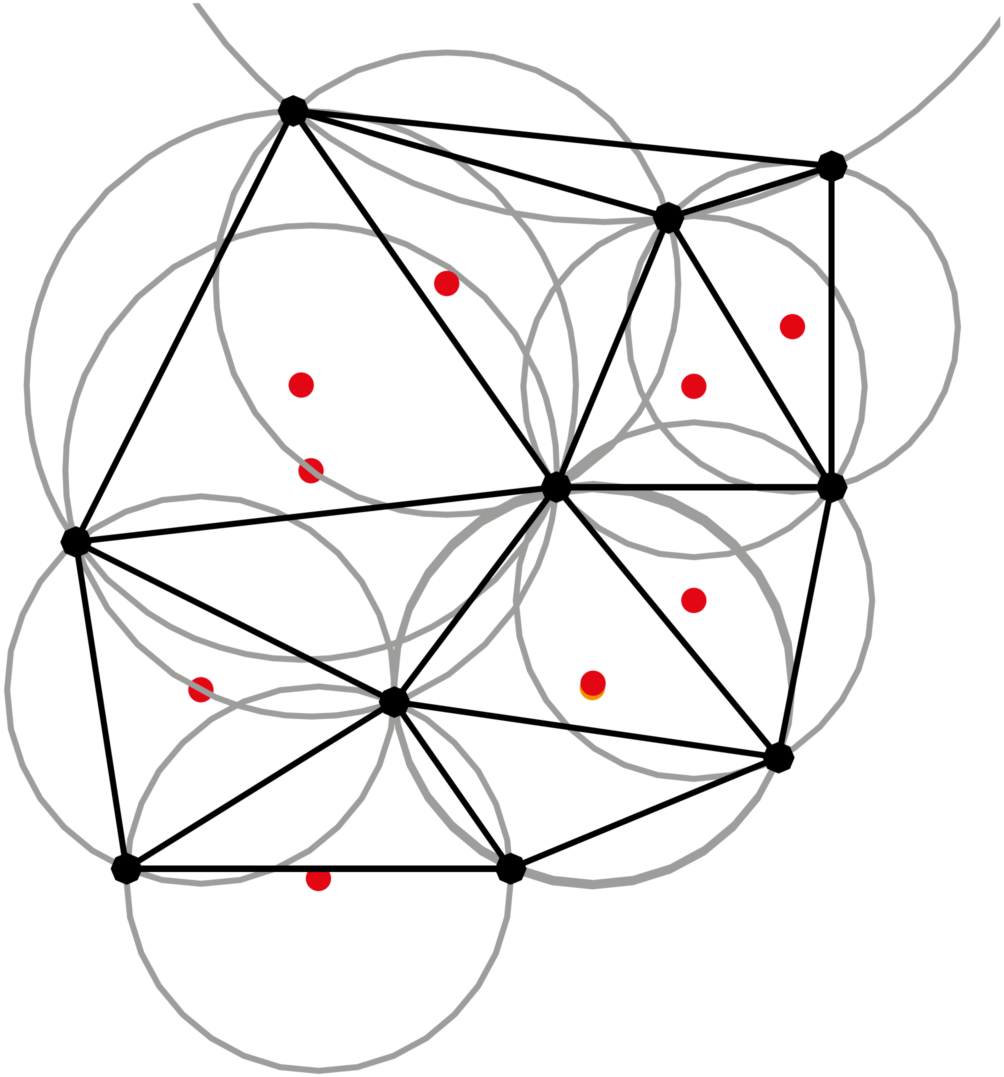
The most simple approach is linear interpolation



There is one “optimal” solution

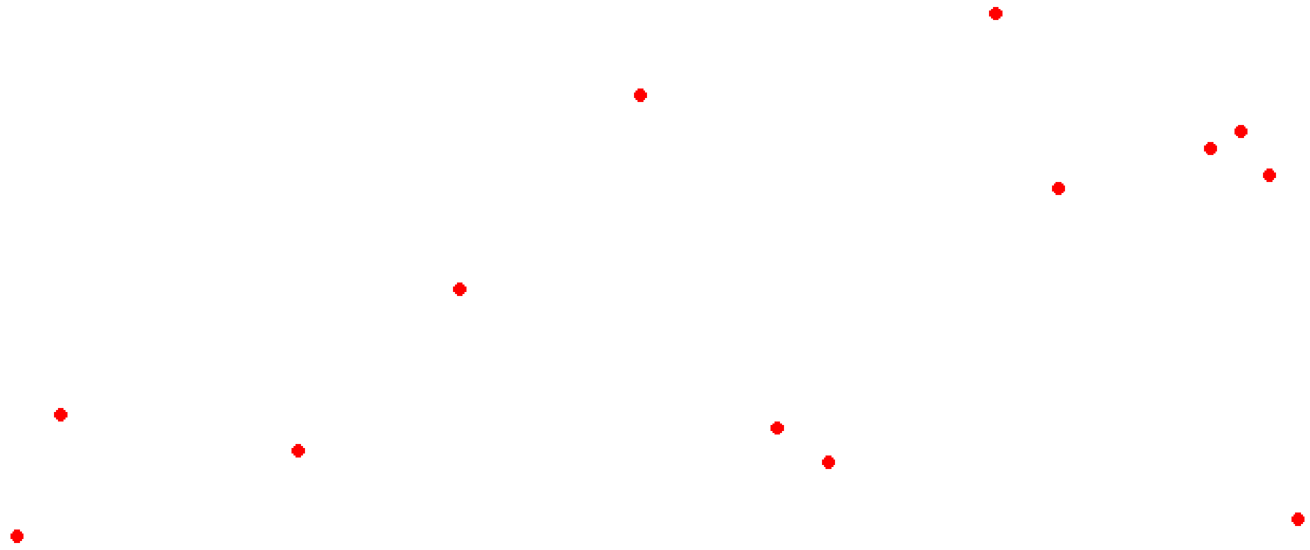


Delaunay triangulation

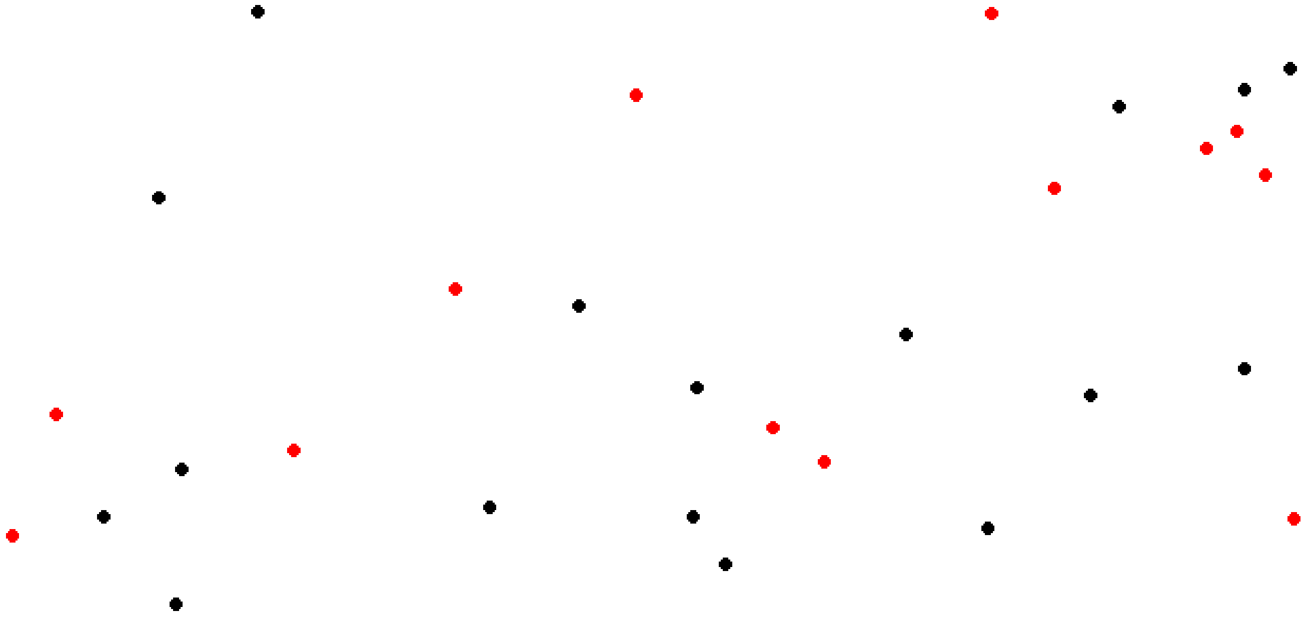


Sanquin

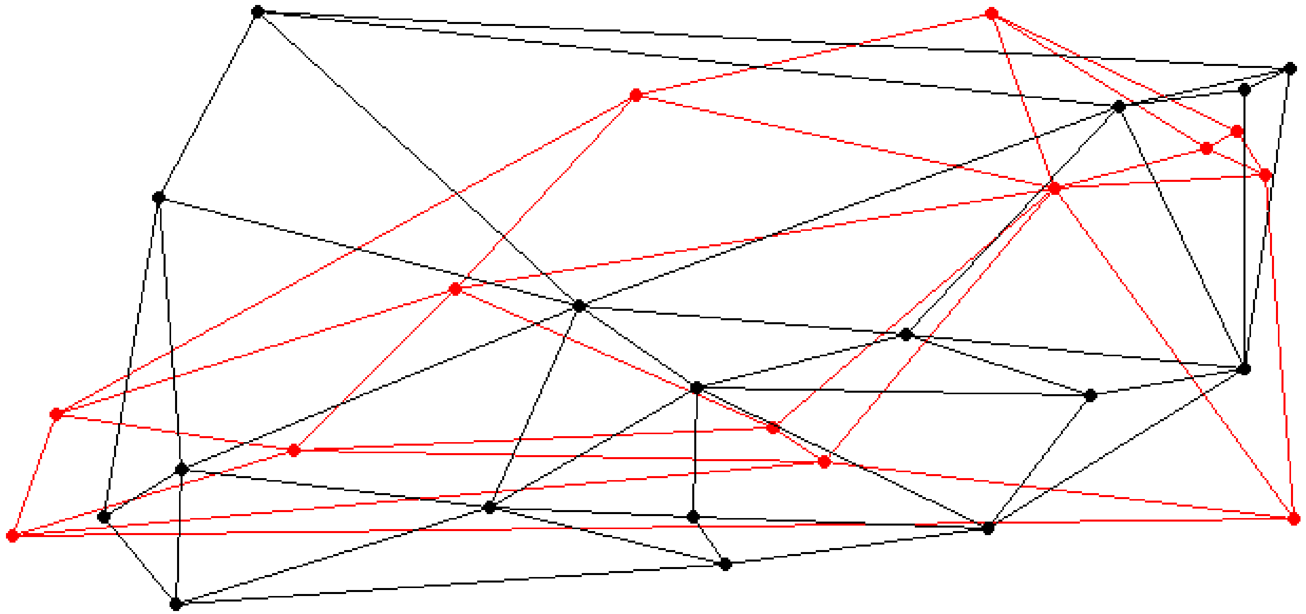
Now you have a second dataset



Which you want to relate to the former…



You could now do inference by mapping the density estimates obtained from the triangulation



Sanquin

What are you going to do in you thesis work?

- Derive and implement efficient calculation of intersections of N-dimensional Delaunay triangulations for many datapoints
- Derive and implement various projections of N-dimensional triangulations on lower dimensions
- Delaunay triangulations provide a tessellation of the complete convex hull, which is sub-optimal for a density representation. These need to be adapted according to the data (e.g. by combining these with kernel density estimates).
- Application of the approach to re-analyse various clinical datasets