

ggvoronoi: Voronoi Diagrams and Heatmaps with ggplot2

Robert C. Garrett¹, Austin Nar¹, Thomas J. Fisher¹, and Karsten Maurer¹

¹ Miami University, Department of Statistics

DOI: [10.21105/joss.01096](https://doi.org/10.21105/joss.01096)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 16 November 2018

Published: 05 December 2018

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

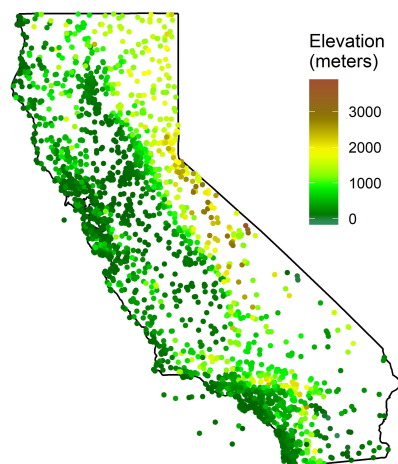
Summary

Spatial data are often gathered at the individual point level, structured as a set of coordinates coupled with a set of measurements. However, attempts to visualize point-level data usually fall into two categories: a scatter plot or some spatial binning technique such as a raster or choropleth diagram of data aggregates. The Voronoi diagram, a mathematical tool that partitions a plane into localized regions around observed points, can be used to create an alternative method for point-level data visualization. A Voronoi diagram is a nearest neighbor diagram: given a set of generator points, the Voronoi diagram creates nearest neighbor cells. Each cell is formed around one generator point, and the set of points closest to that generator is fully contained in its cell. The resulting diagram can be used to find the nearest generator for an arbitrary point or to interpolate the blank space around each generator to create a heatmap.

Voronoi diagrams have a multitude of applications, but `ggvoronoi` implements tools for data visualization and spatial data analysis in R. For data visualization, `ggvoronoi` fits nicely within the `ggplot2` framework, so R users experienced with the popular visualization package can create Voronoi diagrams and heatmaps around any point-level dataset within minutes. For spatial data analysis, R users can create a Voronoi diagram within the `sp` framework as a `SpatialPolygonsDataFrame` object. From there, the user can find the nearest generator point to each location in a new `SpatialPoints` object or use tools, such as those contained in `rgeos`, to analyze the diagram.

`ggvoronoi` depends on R (R Core Team, 2018) as well as a number of R packages. First, the algorithm from `deldir` (Turner, 2018) is used to construct Voronoi cells. The raw output is manipulated with functions from `sp` (R. S. Bivand, Pebesma, & Gomez-Rubio, 2013; E. J. Pebesma & Bivand, 2005), `rgeos` (R. Bivand & Rundel, 2018), and `raster` (Hijmans, 2017) to create spatial objects. Finally, the `ggplot2` (Wickham, 2016) framework is used to visualize the results. The package was built to be used by statisticians and data scientists for spatial data analysis as well as statistics students as a tool to approach data visualization with point-level data. The package is in use in a number of ongoing projects such as extending National Weather Service forecasts to National Climatic Data Center weather stations and interpolating predictions from a Land Use Random Forest model to all of Greater Cincinnati. At Miami University, `ggvoronoi` has been used in the Advanced Data Visualization course as a demonstration of how to visualize point-level spatial data with `ggplot2`. Furthermore, `ggvoronoi` was recently utilized to visually explore the location of train stations in Japan (Chisato, 2018).

Colored Scatter Plot of
Weather Station Elevations in California



Voronoi Diagram Heatmap of
Weather Station Elevations in California

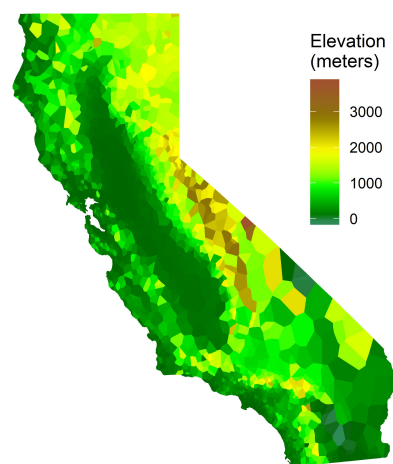


Figure 1: Comparison of a colored scatter plot and Voronoi diagram heatmap. The data containing the National Climatic Data Center weather station coordinates and elevations are included in the package, and the code for this plot is in the package vignette.

Acknowledgements

`ggvoronoi` was inspired by work done at Miami University's annual DataFest. We acknowledge the Center for Analytics and Data Science and the Statistics Department at Miami University for supporting the development of the package and for hosting DataFest with the American Statistical Association.

References

- Bivand, R. S., Pebesma, E., & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, second edition*. Springer, NY. doi:[10.1007/978-1-4614-7618-4](https://doi.org/10.1007/978-1-4614-7618-4)
- Bivand, R., & Rundel, C. (2018). *Rgeos: Interface to geometry engine - open source ('geos')*. Retrieved from <https://CRAN.R-project.org/package=rgeos>
- Chisato. (2018). *Voronoi diagram with ggvoronoi package with train station data*. Retrieved from <https://www.r-bloggers.com/voronoi-diagram-with-ggvoronoi-package-with-train-station-data/>
- Hijmans, R. J. (2017). *Raster: Geographic data analysis and modeling*. Retrieved from <https://CRAN.R-project.org/package=raster>
- Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5(2), 9–13. Retrieved from <https://CRAN.R-project.org/doc/Rnews/>
- R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Turner, R. (2018). *Deldir: Delaunay triangulation and dirichlet (voronoi) tessellation*. Retrieved from <https://CRAN.R-project.org/package=deldir>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. doi:[10.1007/978-0-387-98141-3](https://doi.org/10.1007/978-0-387-98141-3)