





# funkyheatmap: Visualising data frames with mixed data types

Robrecht Cannoodt\* <sup>1,2,3</sup>, Louise Deconinck\* <sup>1,2</sup>, Artuur Couckuyt\* <sup>1,2</sup>, Nikolay S. Markov\* <sup>4</sup>, Luke Zappia <sup>3,5,6</sup>, Malte D. Luecken <sup>5,7</sup>, Marta Interlandi <sup>8</sup>, Yvan Saeys† <sup>1,2</sup>, and Wouter Saelens† <sup>9,10</sup>

**1** Data Mining and Modelling for Biomedicine group, VIB Center for Inflammation Research, Ghent, Belgium **2** Department of Applied Mathematics, Computer Science, and Statistics, Ghent University, Belgium **3** Data Intuitive, Lebbeke, Belgium **4** Division of Pulmonary and Critical Care Medicine, Northwestern University, Chicago, USA **5** Institute of Computational Biology, Computational Health Center, Helmholtz Munich, Munich, Germany **6** Department of Mathematics, School of Computation, Information and Technology, Technical University of Munich, Munich, Germany **7** Institute of Lung Health and Immunity (LHI), Helmholtz Munich, Comprehensive Pneumology Center (CPC-M), Germany; Member of the German Center for Lung Research (DZL) **8** Institute of Medical Informatics, University of Muenster, Muenster, Germany **9** Laboratory of Myeloid Cell Biology in tissue homeostasis and regeneration, VIB Center for Inflammation Research, Ghent, Belgium **10** Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium

DOI: [10.21105/joss.07698](https://doi.org/10.21105/joss.07698)

## Software

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

Editor: Julia Romanowska 

## Reviewers:

- [@rhenkin](#)
- [@meenakshi-kushwaha](#)

Submitted: 14 December 2024

Published: 11 April 2025

## License

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

\*: Shared first authors, †: Shared last authors.

## Summary

The {funkyheatmap} package offers a flexible and user-friendly solution for visualising data frames containing a mixture of categorical, proportional, and text-based data. It simplifies the creation of informative and visually appealing heatmaps while providing extensive customization options to tailor the output. This tool is especially valuable in research settings for summarising and communicating complex results, such as those encountered in benchmarking studies.

The package is available on [CRAN](#), [PyPI](#), and [npm](#). For a showcase of examples and an overview of the different implementations, visit the project website [funkyheatmap.github.io](https://funkyheatmap.github.io).

## Statement of need

Data visualisation is fundamental to exploratory data analysis and communicating findings. Table-creating packages, such as gt ([Iannone et al., 2024](#)), allow the user to create beautiful hierarchical tables, but these packages are limited to showing textual data. Heatmap-creating packages such as pheatmap ([Kolde, 2019](#)), superheat ([Barter & Yu, 2017](#)), and Complex-Heatmap ([Gu et al., 2016](#); [Gu, 2022](#)) require the central visualisation to remain a traditional heatmap and allow additional annotations on the sides of this heatmap. Incorporating multiple data types is possible, but far from seamless. While more powerful tools exist, e.g., ggplot2 ([Wickham, 2009](#)), Matplotlib ([Hunter, 2007](#)), and D3.js ([Bostock et al., 2011](#)), they often require complex scripting to generate comprehensive visualisations for data frames containing a mix of data types.

{funkyheatmap} addresses this challenge by:

- **Seamless Handling of Mixed Data:** Automates the selection of appropriate visualisations (rectangles, bars, pie charts, text) based on data type and allows inclusion of various data types in the same visualisation.

- **Customization:** Provides granular control over colours, groupings, geometries, and annotations for tailored results.
- **Accessibility:** Offers a simplified interface for basic use and detailed documentation for advanced customization.

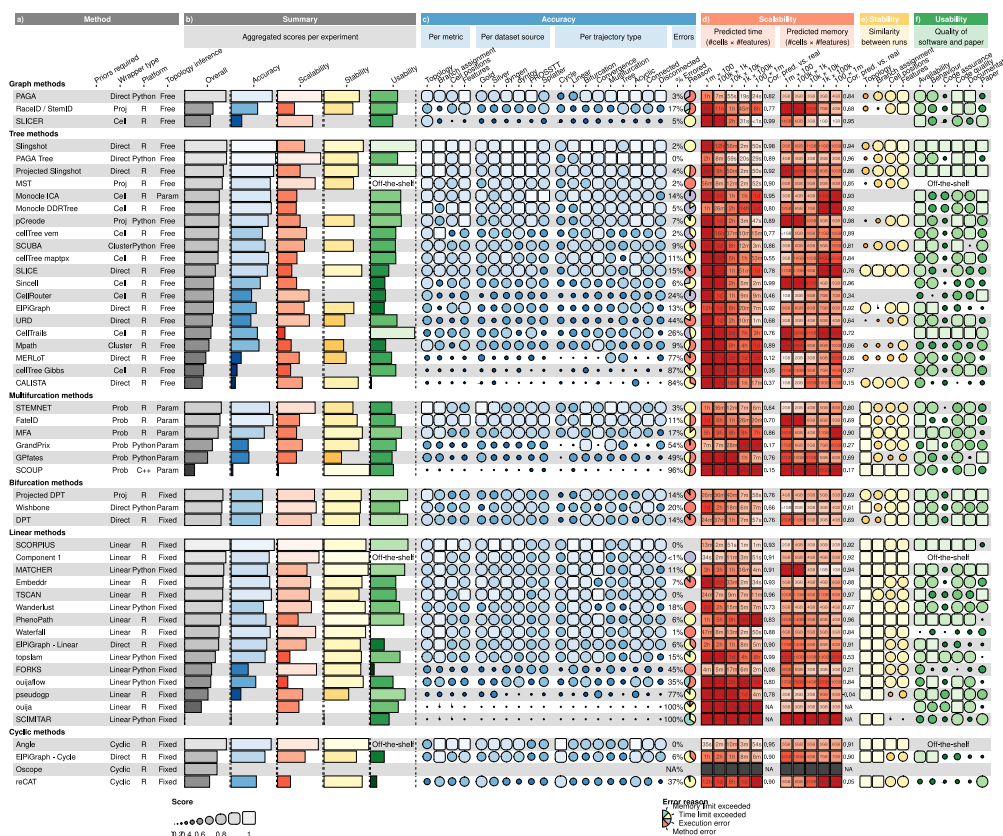
{funkyheatmap} has proven its utility in benchmarking studies within single-cell omics (Li et al., 2023; Luecken et al., 2022, 2024; Saelens et al., 2019; Sang-Aram et al., 2024; Yan & Sun, 2023) and its applications extend to diverse fields where visualisation of mixed data types is needed.

## Functionality

Figure 1 showcases the functionality of {funkyheatmap}, namely:

- **Diverse Geometries:** Supports a range of geometries (rectangles, bars, pie charts, text, images) to effectively represent different data types.
- **Hierarchical Categorical Grouping:** Facilitates the organisation of rows and columns into semantic groups with distinct colour palettes.

Besides this, the package includes comprehensive documentation, vignettes, and a test suite.



**Figure 1:** An example of a {funkyheatmap} visualisation using data from a benchmarking study of trajectory inference methods (Saelens et al., 2019).

See the following [Table](#) for more information regarding the recommended geom for different types of data.

Data type	Example	Recommended geom
Numerical data	Scores from 0 to 1	funkyrect
Aggregated data	The mean of scores	bar
Measurement data	3MB or 4h	rect + text overlay
Categorical data	R or Python	text or image
Proportional data	80% success, 10% OOM, 10% failed	pie

The table presents the suggested visualisation methods (geoms) based on the data type of the columns. These recommendations provide a starting point for users to select the most appropriate visual representation for their specific data.

### Example usage

In order to produce a {funkyheatmap} visualisation, you need to provide the data in the form of a dataframe, which must contain a column named `id`. If you provide no other information, a basic visualisation will be provided, but customization is possible by providing additional information, such as a `column_info` dataframe which details how the columns in the dataframe get translated into different geoms, or a `row_groups` dataframe which allows you to group rows in the visualisation.

As an example, [Figure 2](#) shows a visualisation of the `mtcars` dataset. For an in-depth explanation on how to fine-tune this figure and the different data structures involved, please see the [Getting started](#) article in the package documentation.

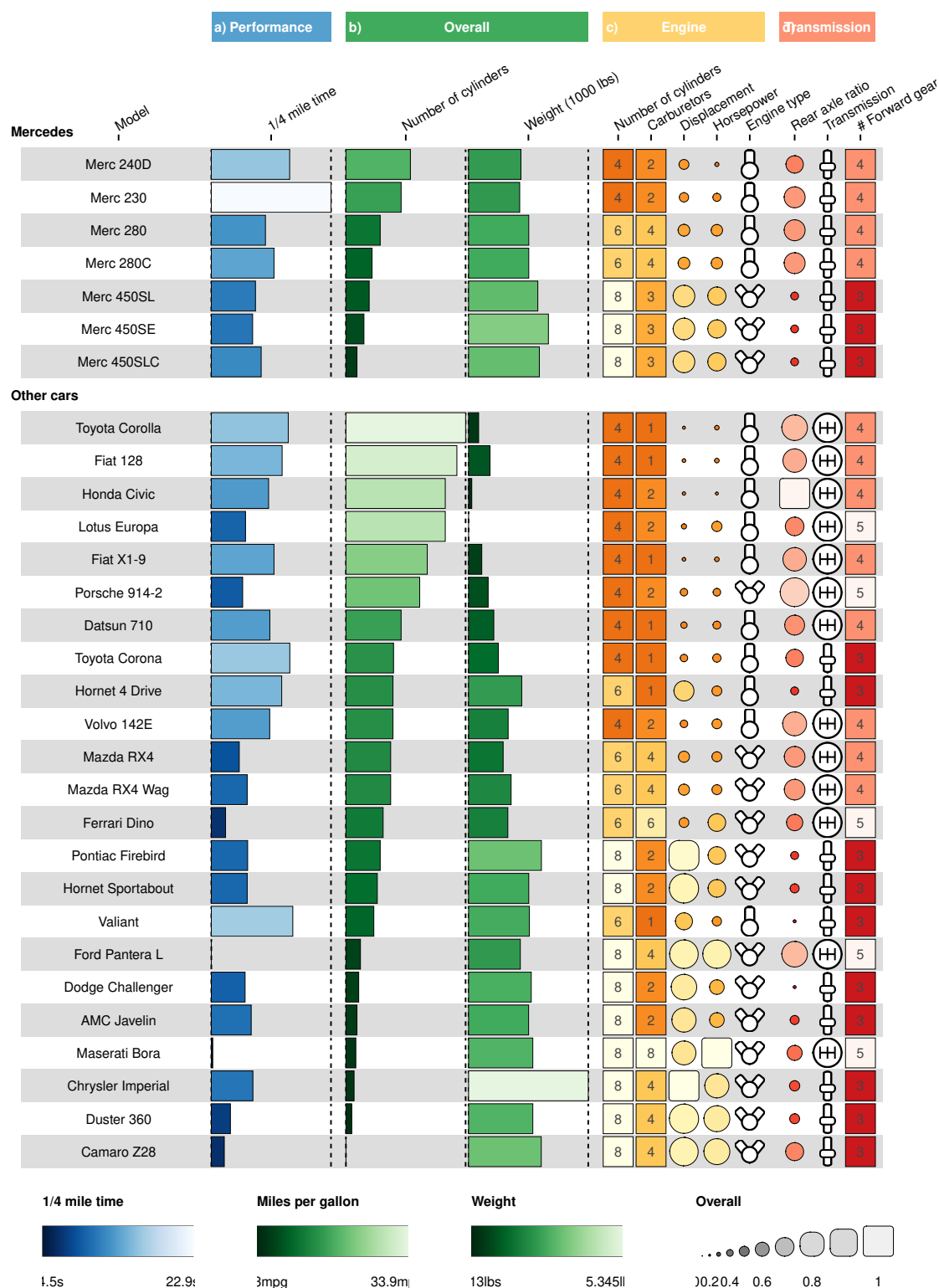


Figure 2: An example of a {funkyheatmap} visualisation using the mtcars dataset.

## Conclusion

{funkyheatmap} streamlines the creation of publication-quality visualisation for mixed data types, empowering researchers and data scientists to communicate their results effectively. The package has been ported to other programming languages, as funkyheatmappy (Python) and funkyheatmapjs (JavaScript). This expands the accessibility of this type of visualisation to non-R users. The JavaScript port incorporates interactive elements, expanding on the visualization's functionality.

## Acknowledgements

L.Z. acknowledges support from the Bavarian Ministry of Science and the Arts in the framework of the Bavarian Research Association “ForInter” (Interaction of human brain cells). L.D. acknowledges support from the Research Foundation - Flanders (FWO) (1SF3822N). Y.S. acknowledges support from the Flemish Government under the “Onderzoeksprogramma Artificial Intelligence (AI) Vlaanderen” program. A.C. and Y.S. are funded by an FWO TBM grant (Research Foundation – Flanders, T000119N), and Ghent University Special Research Fund (BOF18-GOA-024). W.S. was supported by a Marie Skłodowska-Curie fellowship (101028476).

## Author Contributions

R.C. and W.S. conceived the visualisation method and added the dynbenchmark example. R.C., W.S. and L.D. authored the R version of the package. L.D. and A.C. authored the Python version of the package. N.M. authored the Javascript version of the package. R.C., L.D., L.Z., A.C. and N.M. participated in design meetings and discussions that shaped the software design and functionality of all packages. L.Z., M.I. and M.L. added the scIB example and made indirect code contributions by creating a derived version of the original scripts. Y.S. supervised the original dynbenchmark work and encouraged the creation of the software package.

## References

- Barter, R., & Yu, B. (2017). *Superheat: A graphical tool for exploring complex datasets using heatmaps*. <https://doi.org/10.32614/cran.package.superheat>
- Bostock, M., Ogievetsky, V., & Heer, J. (2011). D<sup>3</sup> data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12), 2301–2309. <https://doi.org/10.1109/TVCG.2011.185>
- Gu, Z. (2022). Complex heatmap visualization. *iMeta*, 1(3), e43. <https://doi.org/10.1002/imt2.43>
- Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btw313>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(03), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Iannone, R., Cheng, J., Schloerke, B., Hughes, E., Lauer, A., Seo, J., Brevoort, K., & Roy, O. (2024). *Gt: Easily create presentation-ready display tables*. <https://doi.org/10.32614/CRAN.package.gt>
- Kolde, R. (2019). *Pheatmap: Pretty heatmaps*. <https://doi.org/10.32614/cran.package.pheatmap>
- Li, H., Zhou, J., Li, Z., Chen, S., Liao, X., Zhang, B., Zhang, R., Wang, Y., Sun, S., &

- Gao, X. (2023). A comprehensive benchmarking with practical guidelines for cellular deconvolution of spatial transcriptomics. *Nature Communications*, 14(1), 1548. <https://doi.org/10.1038/s41467-023-37168-7>
- Luecken, M. D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Müller, M. F., Strobl, D. C., Zappia, L., Dugas, M., Colomé-Tatché, M., & others. (2022). Benchmarking atlas-level data integration in single-cell genomics. *Nature Methods*, 19(1), 41–50. <https://doi.org/10.1038/s41592-021-01336-8>
- Luecken, M. D., Gigante, S., Burkhardt, D. B., Cannoodt, R., Strobl, D. C., Markov, N. S., Zappia, L., Palla, G., Lewis, W., Dimitrov, D., & others. (2024). Defining and benchmarking open problems in single-cell analysis. *Research Square*. <https://doi.org/10.21203/rs.3.rs-4181617/v1>
- Saelens, W., Cannoodt, R., Todorov, H., & Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nature Biotechnology*, 37(5), 547–554. <https://doi.org/10.1038/s41587-019-0071-9>
- Sang-Aram, C., Browaeys, R., Seurinck, R., & Saeys, Y. (2024). Spotless, a reproducible pipeline for benchmarking cell type deconvolution in spatial transcriptomics. *Elife*, 12, RP88431. <https://doi.org/10.1101/2023.03.22.533802>
- Wickham, H. (2009). *ggplot2*. Springer. <https://doi.org/10.1007/978-0-387-98141-3>
- Yan, L., & Sun, X. (2023). Benchmarking and integration of methods for deconvoluting spatial transcriptomic data. *Bioinformatics*, 39(1), btac805. <https://doi.org/10.1093/bioinformatics/btac805>