

Datasaurus

Ryan Jensen

Datasaurus

The datasaurus is a collection of 13 sets of data points. Each set has 5 (almost) identical statistics. Scatter plots reveal that while basic statistics are the same, their look arrangement is very different.

I've modified the original datasaurus collection slightly, and applied persistent homology in order to differentiate some of the data sets. My results are below.

- [Datasaurus Slides](#)
- [Datasaurus to Circle](#)
- [Datasaurus Stats](#)
- [Datasaurus Barcodes](#)