Datasaurus

Ryan Jensen

Datasaurus

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

I've modified the original datasaurs 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