

# Datasaurus

Ryan Jensen

## Datasaurus

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