Summary

This paper goes over the field of data science. Donoho goes over the field of statistics and talks about how this field has given rise to modern day data science, going over influential papers and looking at coursework in new data science programs. He finishes by defining two areas of data science, the lesser data science and the greater data science.

Reaction

A neat paper that goes over data science. Some parts of this paper seemed difficult for me to read, just how it was laid out or the verbiage maybe. Not too sure how useful this paper will be but might be nice information to have later on.

Questions

What are the negatives of rebranding statistics as data science if what we are doing doesn’t change? Sure, it would be very annoying to the current statisticians, but I can’t think of anything else.

I think in data science there is not really an emphasis on large data systems and scalability today. Is it not really that useful for data science anymore or why have they been phased out.