# Some Advantages of Utilizing Docker for Data Science

Using Docker for Data Science research and development holds the promise of greatly expanded productivity. Docker lets one create containers of work that can easily be shared amongst multiple data scientists. A scientist can create a self-contained container with various code and methods, as well as data (although data could be retrieved from outside the container if so set up) that someone needs to analyze a problem, and the container can then be shared to various team members or others. These others would not have to do set up to get themselves going, and all would have same starting point.

The containers can also allow for iteration type testing. For example, maybe multiple containers are set up with the same data, but with different code and analytical methodologies. Container 1 contains the first iteration of work, container two, contains the second and so forth. It could make easily run comparisons between multiple ideas on how to analyze a problem.

One last item to be mentioned, though we are clearly just scratching the surface of their usefulness, is Docker containers can help speed long running intensive analysis. If complex problem analysis can be broken into multiple containers, that can be run in parallel, the docker containers can be on different processors and could therefore much more quickly returns results.