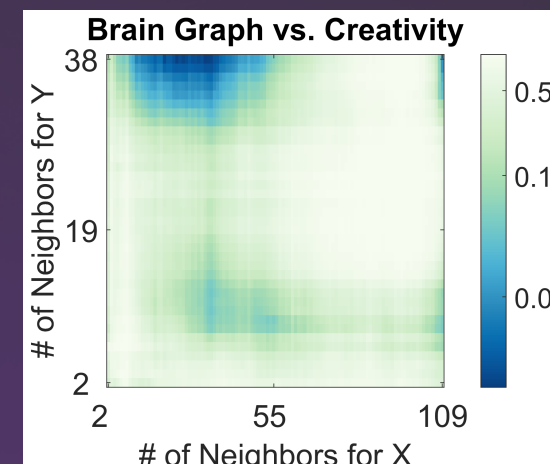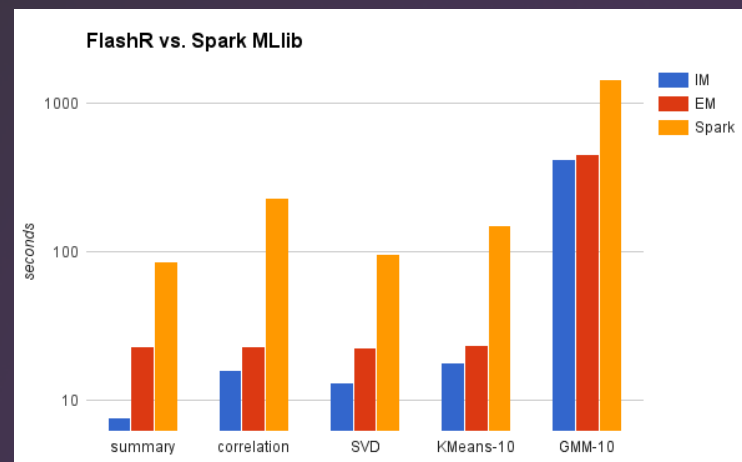# JHU SIMPLEX: Enabling Terascale Neuroscience for Everyone

(1) FlashX for Data Science
- Added a variety of generalized matrix operations to FlashX
- This enables many basic data science routines to be written in native R code but scale to arbitrarily big data on a single machine

http://flashx.io/

FlashR vs. Spark MLlib

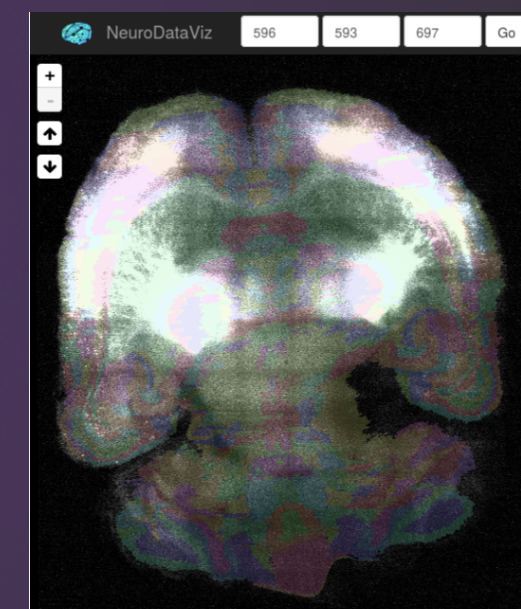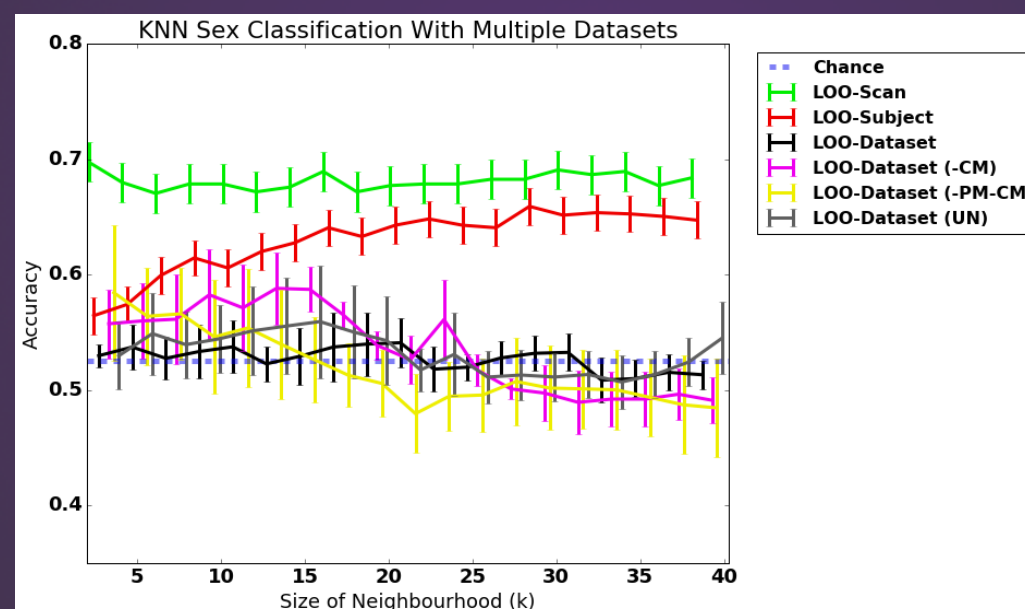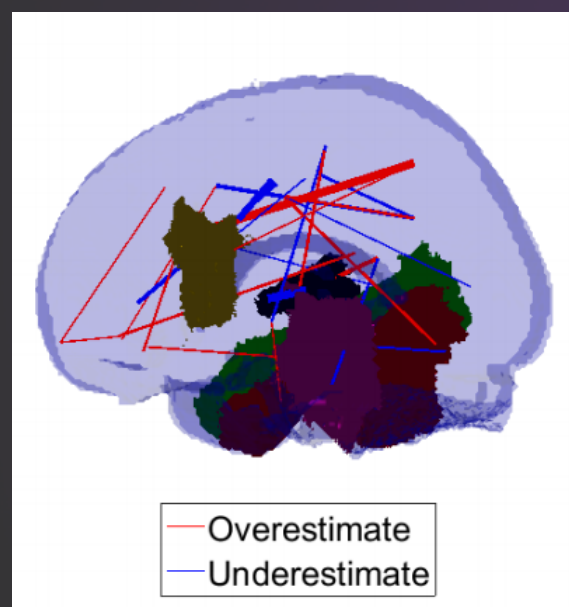
Brain Graph vs. Creativity

(2) MGC 2-Sample Testing
- Extended previous 2-sample testing (dcorr) by sparsifying graphs
- We now have theory and methods (in MATLAB and R) that statistically dominates the previous best method
- Use in several novel neuroscience applications to discover the scales of dependency

https://github.com/jovo/MGC

Overestimate
Underestimate

KNN Sex Classification With Multiple Datasets

NeuroDataViz

(3) Law of Large Graphs
- Proved spectral regularization is more efficient than naive estimate of of average graph
- Proved robust variant is even more efficient in the presence of outliers
- Demonstrated on real data to discover false positives and negatives in previous estimates of mean connectome

https://github.com/jhu-graphstat/LLG

(4) MR Batch Effect
- Using data data and pipeline from GRAPHS, discovered the existence of batch effects
- Simple linear approaches to removing batch effects failed
- Extending now to nonlinear models

http://m2g.io

(5) CLARITY
- Multimodal LDDMM for registering CLARITY to Atlas and other CLARITY brains
- Distributed ROI histogram Web-service
- Ingested 12 ~1TB CLARITY brains and registered them each to Allen Atlas

https://github.com/neurodata/ndreg

In summary, we provide a suite of tools and techniques which extend the boundaries of scientific discovery