UNC, Stat & OR

Data @ Carolina Workshop 2015

Thoughts on Big Data

J. S. (Steve) Marron

<marron@unc.edu>

- Statistics and OR
- Biostatistics
- Computer Science (Adjunct)
- Lineberger Cancer Center (Res. Member)

October 1, 2015

UNC, Stat & OR

Big Picture Research Proposal Idea

Well Understood Concept:

Great science now done by <u>teams</u>

with complementary skill sets

- Biology
- Chemistry
- Engineering
- Quantitative Work

⋮

Common Current Idea:
1 Team member

UNC, Stat & OR

Big Picture Research Proposal Idea

Next Generation of this Concept:

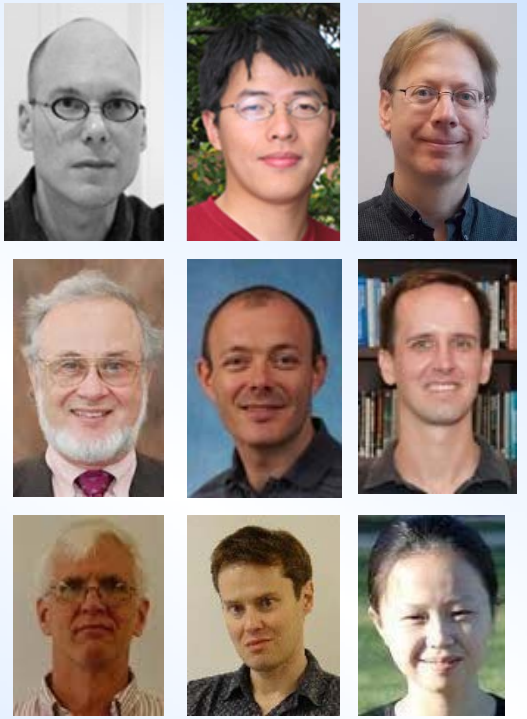
Great *Quantitative Work* needs <u>teams</u>

with complementary skill sets

- Statistics
- Imaging
- Optimization
- Data Base

⋮

→ Big Data



UNC, Stat & OR

My Research

BIG DATA is a Challenge

Perhaps Bigger Challenge: *Complex Data*

Useful approach:

Object Oriented Data Analysis

Focal Point of Interdisciplinary Discussions:

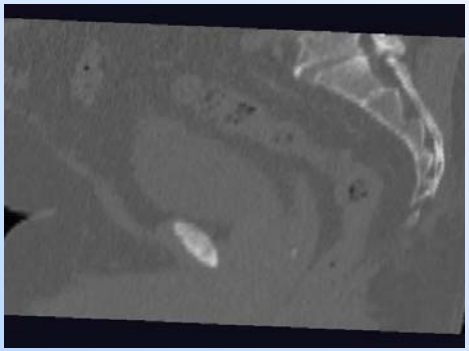
"What should be the data objects?"

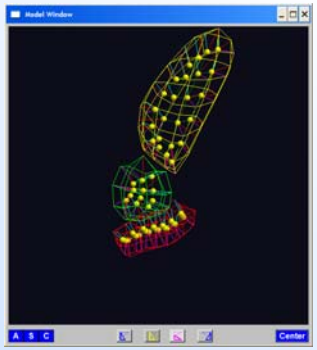
UNC, Stat & OR

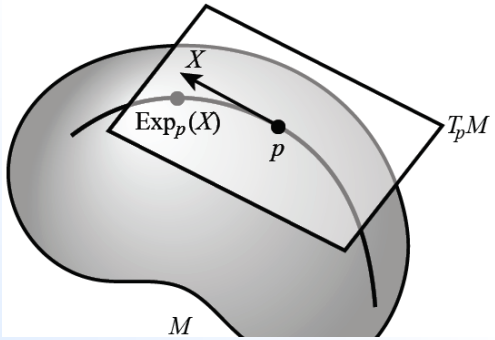
My Research

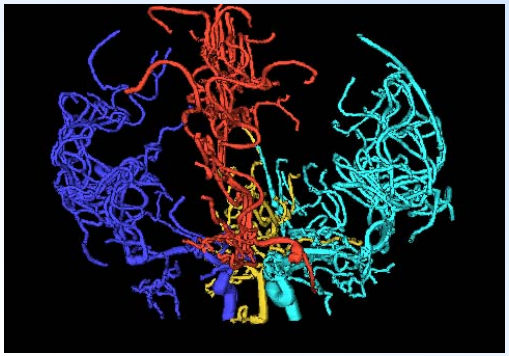
Object Oriented Data Analysis

- Data Visualization
- Complex Data Spaces









- Recent Breakthrough:

  Backwards PCA

UNC, Stat & OR

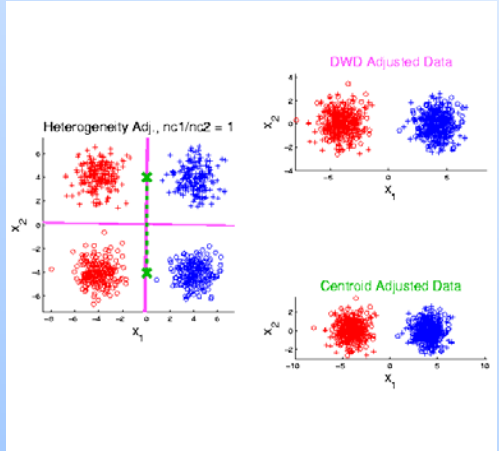
My Research

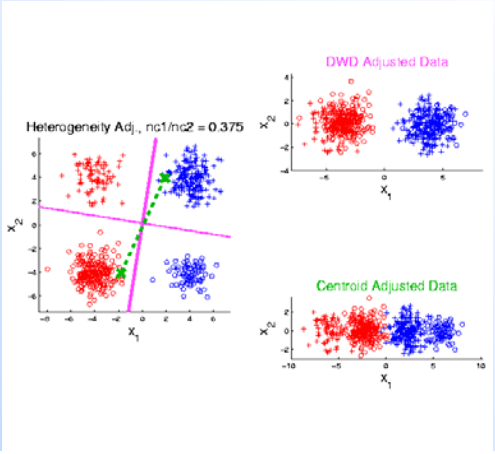
New Trend in Big Data:

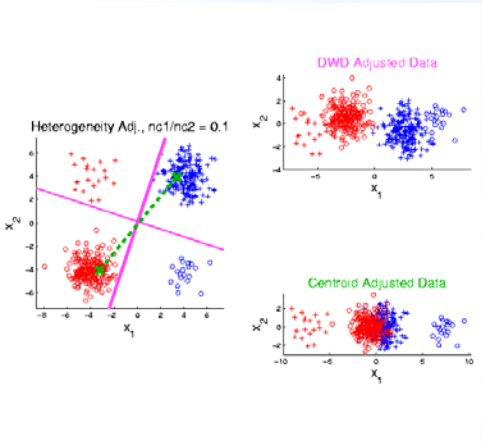
Heterogeneity & *Robustness* Against It

New Conceptual Model:

(Unknown Gaussian Mixtures)







UNC, Stat & OR

What Can UNC Do About Big Data?

- ✓ Nurture Research & Collaboration
- ✓ Get *Serious* About Faculty Retention
- ✓ Treasure Innovation & Creativity