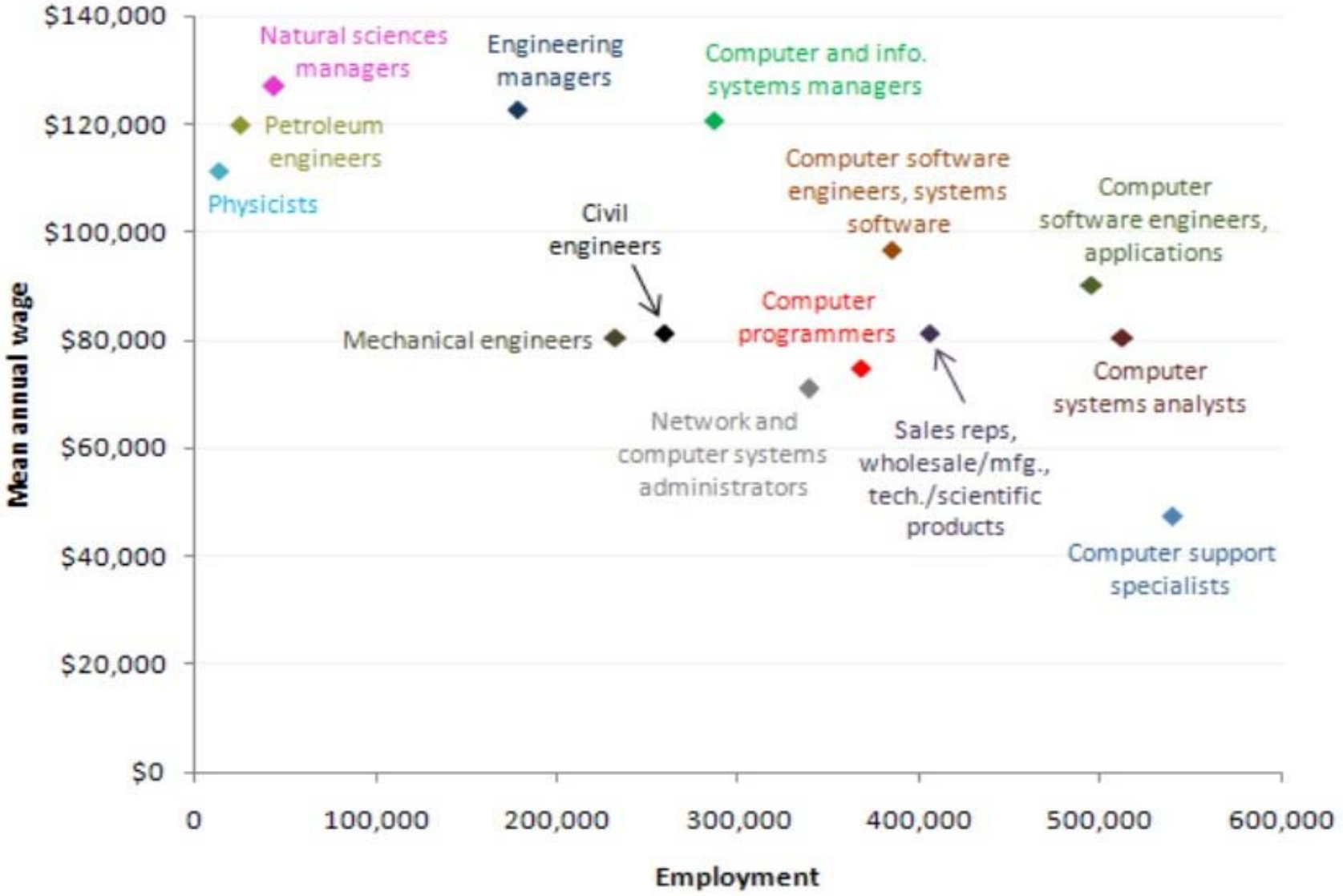
A CAREER IN THE PRIVATE SECTOR: THE SOFTWARE INDUSTRY

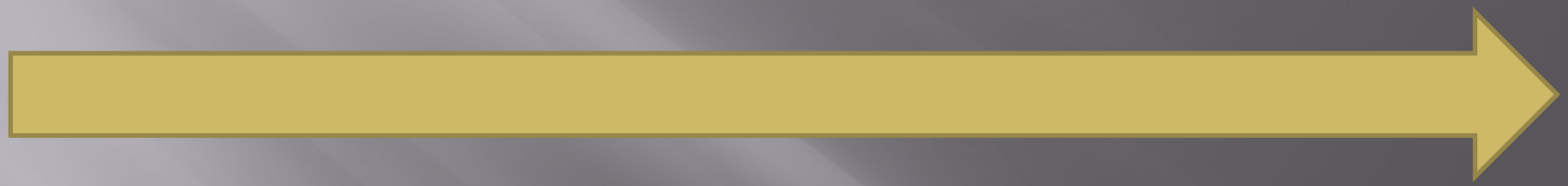
Cathy Maahs-Fladung
Assistant Professor -- STEM
Emma Eccles Jones College of Education and Human Services
cathy.maahs-fladung@usu.edu

Selected science, technology, engineering, and mathematics (STEM) occupations, employment and mean annual wage, May 2009



Source: U.S. Bureau of Labor Statistics

CAREER TIMELINE



1983
1993
SPSS Inc.

1993
2003
SAS Institute

2003
2005
GlaxoSmithKline

2005
2010
East Carolina University

2010
Present
Utah State University

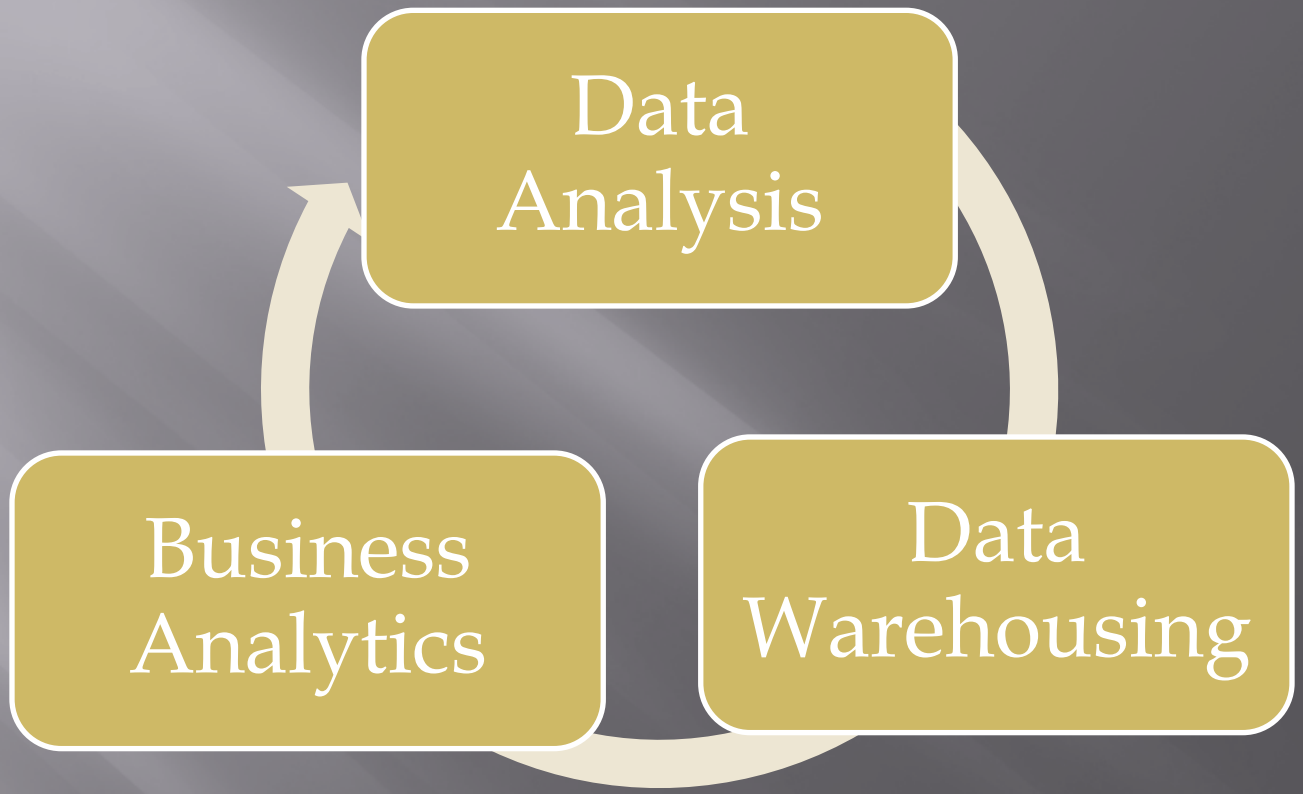
2010
2012
Revolution Analytics
An R Company

THE SOFTWARE INDUSTRY

IBM/SPSS

SAS Institute

Revolution Analytics


Data Analysis
Data Warehousing
Business Analytics

IBM/SPSS

- Founders: Norman Nie, Tex Hadley Hall
- Year founded: 1969
- Stanford University, Political Science
- Analysis of Voting Behavior
- Teamed with National Opinion Research Council
- Multiple partners world-wide
- Venture Capital / Traded Publically
- About 500 employees
- Norman leaves for Stanford & eventually Revolution Analytics
- Sold to IBM: 2010

SAS Institute

- Founders: Jim Goodnight, John Sall
- Year founded: 1976
- North Carolina State University, Statistics & Agriculture
- SAS is the largest independent vendor in the "business intelligence" market.
- SAS has offices in 56 countries w/ 400 offices
- SAS has more than 800 alliances globally
- 13,349 total employees

Revolution Analytics

- Founders: David Champagne, Norman Nie, David Rich
- Founded: 2007
- Palo Alto, California – Most employees are virtual
- An open source R company to foster R computing
- "Through our Revolution R products, we aim to make the power of predictive analytics accessible to every type of user & budget. We provide free and premium software and services that bring high-performance, productivity and ease-of-use to R -- enabling statisticians and scientists to derive greater meaning from large sets of critical data in record time.".

What They Have in Common

- Data analytics which they often call business analytics
- Developers who develop statistics and "data mining" tools.
- Full service products including data warehousing
- All started as statistical software companies.
- Main target customers: Business
  - SPSS is more user friendly; more behavioral sciences oriented
  - SAS is and has always been oriented to the scientific community
- Serve academic community as well.
- All are targeting some aspect of the R community as it grows.

Jobs for Statisticians

- Statistical Programmer/Designer – Back-end and Front-end
- Statistical Graphics Programmer
- Statistical Writer
- Quality Assurance
- Statistical Consultant and Trainer
- Human Factors Analyst
- Technical Support/Statistics
- Systems Engineer or Statistical Product Manager – working with Marketing and Sales
- Limited number of post doc positions partnering with universities and/or business

Jobs for Student Statisticians in Training

- Undergraduate Intern
- Graduate intern

What you get if you choose this route

- Remember, you are selling or developing a product…
  - Based on customer needs
  - Based on your insight of the market and customer needs
- The opportunity to work with a lot of bright people who focus on new challenges
- The money to fund new projects
- Collaboration with university, private and public sector
- Often, training by statisticians who are first and foremost in their field
- A collegial environment
- Oh yes, a good salary.

So…what is a job as a statistician like for a major software vendor?

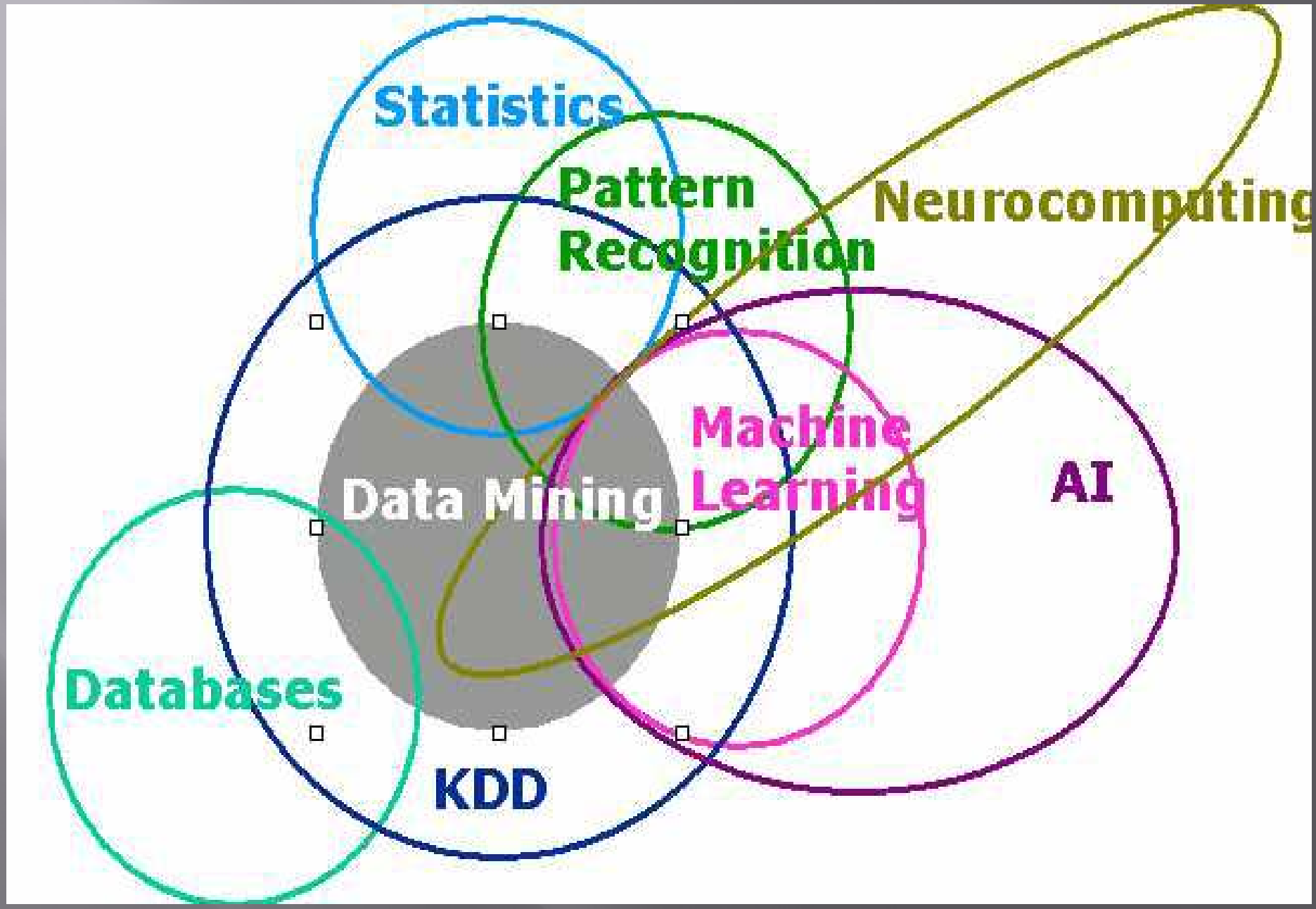
- I began my career at SPSS as a Statistical Writer after obtaining my master's degree at the University of Illinois
- I led the quality assurance team for SPSS's first PC product (the first of it's kind).
- Next, I managed Technical Support for Statistics
- Then, consultant and trainer as well as participating in the early development of decision trees and neural networks for SPSS – (this was prior to SAS's interest in data mining)

So…what is a job as a statistician like for a major software vendor?

- In 1993, I joined SAS in Cary, North Carolina
- 1.5 years in Statistical Technical Support supporting IML, CALIS, Time Series
- Worked with development to develop data, text mining and campaign management tools
- Director, Worldwide Data Mining Consulting and Training
- Worked to John Sall to develop data mining procedures for SAS/JMP

DATA MINING

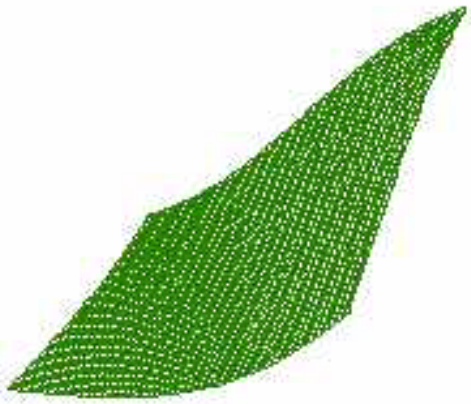
Supervised and Unsupervised Classification


Statistics
Pattern Recognition
Neurocomputing
Machine Learning
AI
Data Mining
Databases
KDD

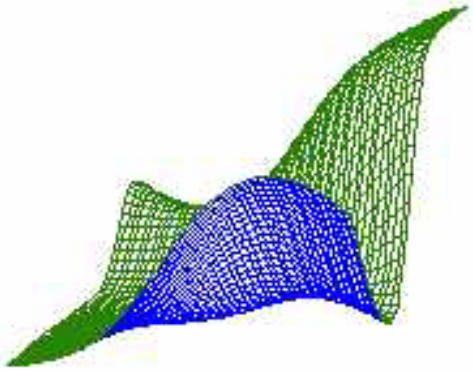
Problem Translation

- Millions of records, transaction data, many missing values
- Supervised and Unsupervised Analysis
  - Predictive Modeling (supervised classification)
  - Unsupervised Classification
    - Cluster Analysis
    - Association
    - Rule based Analysis

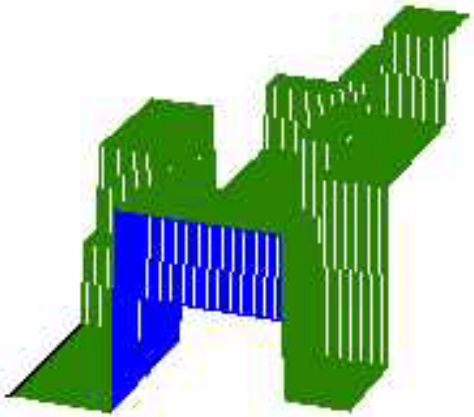
Modeling Methods



Generalized
Linear Models



Neural
Networks



Decision
Trees

Type of Analysis	Supervised Analysis	Unsupervised Analysis	Vertical Market (Examples)
Classification	Segmentation	Segmentation	Database Research
	Response Analysis		Marketing Analysis
	Attrition/Retention		Banking/Credit Card
	Charges		Network Analysis
	Fraud Detection		Risk Management
	Campaign Analysis		
Prediction	Attrition/Retention	Not feasible	
	Fraud Detection		
	Campaign Analysis		
	Scoring		
Exploration	Segmentation	Segmentation	
	Profiling	Profiling	
	Attrition/Retention		
	Fraud Detection		

Beyond the Software World

- GlaxoSmithKline
  - Principal Applications Developer – Client Side
  - Clinical Drug Development
- East Carolina University
  - Associate Director, Institutional Planning, Assessment and Research
  - Adjunct Professor, Psychology
  - Adjunct Professor, School of Education