

# LADS 2018 Sarasota

## Computer Science in the Data Science Curriculum

Dennis F.X. Mathaisel  
Professor  
Math/Science Division  
Babson College  
Wellesley, MA



# Babson Computer Science Curriculum

**QTM2630 – Programming with R  
for Business Analytics**

DES3600 - Design and Systems  
Thinking

MIS2645 - Modeling with Excel

MIS3545 - Business Intelligence and  
Data Analytics

MIS3555 - Platforms, Clouds and  
Networks

MIS3560 - The Blockchain: Bitcoin,  
Smart Contracts

MIS3574 - Emerging Technologies

MIS3610 - The Mobile App

MIS3615 - Creating Tech-Savvy  
Entrepreneurs

MIS3620 - Computer and Network  
Security

MIS3625 - Drupal Web  
Programming Essentials

MIS3635 - User Interface Design

**MIS3640 - Problem Solving &  
Software Design**

MIS3645 - Agile Methodology

MIS3655 - Design and Systems  
Thinking

MIS3660 - Prototyping w/It

MIS3690 - Web Technologies

MIS3696 - Drupal Web Programing,  
Bootstrapping Inn



# Data Science v Statistics v Analytics\*

## Data Science

- Mathematics
- Statistics
- Information Science - analysis, collection, classification, manipulation, storage, retrieval, movement, dissemination, and protection of information.
- Computer Science - theory, experimentation, and engineering that form the basics for the design and use of computers.

## Statistics

- Collection, analysis, interpretation, presentation, and organization of data.

## Analytics

- Discovery, interpretation, and communication of meaningful patterns in data.
- Statistics
- Computer Programming
- Operations Research



\* Source: Wikipedia

# 8 Essential Skills for Data Scientists\*

1. **Programming:** a statistical programming language, like R or Python, and a database querying language like SQL.
2. **Basic Statistics:** distributions, statistical tests, estimation/prediction, ...
3. **Data Visualization & Communication:** Visualizing (e.g., ggplot) and communicating data (e.g., apps).
4. **Machine Learning:** k-nearest neighbors, random forests (tree predictors), clustering methods, ...
5. **Data “Munging”:** missing values, outliers, inconsistent string formatting, date/time formatting.
6. **Software Engineering:** design, implementation, testing, and documentation of software.
7. **Thinking Like A Data Scientist:** ability to interact with engineers and product managers.
8. **Basic Multivariable Calculus and Linear Algebra:** they form the basis of the machine learning techniques.



\* Source: Udacity (modified by Mathaisel)