Informatics and Biomedical Data Science Training at the University of Utah
Incorporating Data Science Modules into Existing Informatics Courses
Just like many other disciplines, what we now call “data science” seems to have always been a part of informatics.

Angst all around
- E.g. “50 Years of Data Science”

Shift of emphasis? Changed opportunities?
- Knowledge engineering ⇔ Discovery from Data
OUR APPROACH AT THE UNIVERSITY OF UTAH

• **Aim 1:** Enhance and Enrich Faculty Development through formal and informal training on data science techniques as well as through pedagogical coaching.

• **Aim 2:** Create Case Study Data Science Modules for use in our courses that apply hands-on data science activities to case studies using real-world data sets.
  – Develop Case Studies models using real-world data sets.
  – Adapt Case Studies to targeted DBMI classes.
5\textsuperscript{2} MODEL OF MODULE DEVELOPMENT

• Clinical context--a description of the case study and its context
• Resources--tools, papers, websites, news stories relevant to the case study
• Pedagogical materials--slides, videos, and demos to teach knowledge related to the case study
• Exercises--guided exercises for students to learn about and apply data science techniques
• Reflection--an assignment for reflecting on the results of the case study

• Domain Knowledge
• Computation
  — Version control; data access, wrangling, transformation, etc
• Visualization
  — Scripting, interactive visual analytics
• Analytics
• Standards
DESIGN DECISIONS

- Python (some R, JavaScript)
- Git
- Jupyter notebooks
- Cloud Computing
  - Centralized resources/minimize concerns with student computers
  - Facilitate remote learning
INFRASTRUCTURE
“ONION MODEL”

• Each Module Facilitates Different Levels of Exploration
  — Use widgets to explore data
  — Guided Scripting
  — Suggested explorations
BIOINFORMATICS EXAMPLE

- Galaxy Workflow
- Jupyter Notebooks within Galaxy
- Command line implementation
EXAMPLE MODULES (WORKS IN PROGRESS)

- NLP (MIMIC)
- ECG (Physionet)
- Flu Transmission (Public Utah Data)
DECART SUMMER SCHOOL
DATA SCIENCE FOR HEALTHCARE

DECART: DATA SCIENCE FOR THE HEALTH SCIENCES

Data, exploration, Computation, and Analytics Real-world Training for the Health Sciences