Lecture 0: Course Description

Course: Biomedical Data Science

Parisa Rashidi
Fall 2018

THE 5 STAGES OF VACATION GRIEF

1 DENIAL » THE IT'S NOT SO BAD STAGE

Back at your desk. Still feeling the glow of sleeping late and forgetting this place ever existed. Got your coffee—this isn't so bad. After all, you're lucky to have a job in this economy... plus it's dry in the office... not much chance of a bear attacking you.

2 ANGER » THE #@$%*# STAGE

Who put all this work in your inbox? Why are people emailing you about this $#&% now?! That project was approved, then your boss suddenly decides it's not right anymore? And it's due on Monday?! This place is crushing your soul. You. Hate. Everything.

3 BARGAINING » THE I CAN FIX THIS STAGE

Life is what you make it—if you can stay on top of your deadlines and leave the building for lunch—being back at work will be ok. Maybe you'll take a class on time management, find a new podcast.

4 DEPRESSION » THE IT'S ALL MEANINGLESS STAGE

Slept through two alarms. Shit! You hope no one noticed you were late. Ugh, this desk again? Was it always so dark in here? Oh, what does it matter if it's bright or dark? It's all the same. Every day the same as the last. You eat a bagel.

5 ACCEPTANCE » THE TGIF STAGE

Time to put in a request for your next vacation... maybe somewhere sunny like Mexico or Costa Rica.

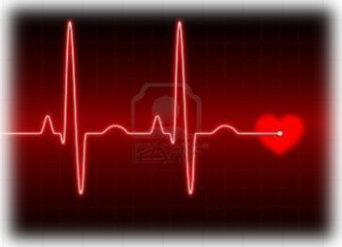
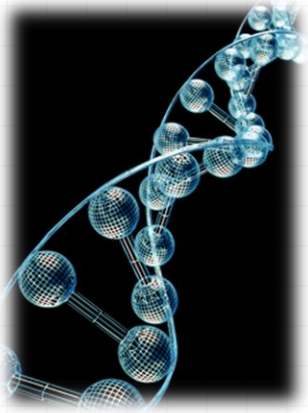
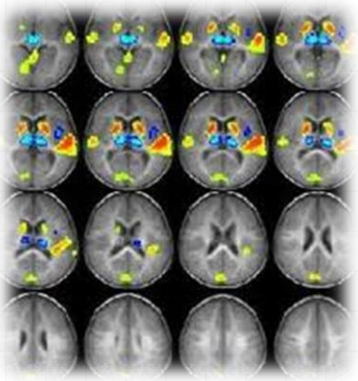
BuzzFeed

Agenda

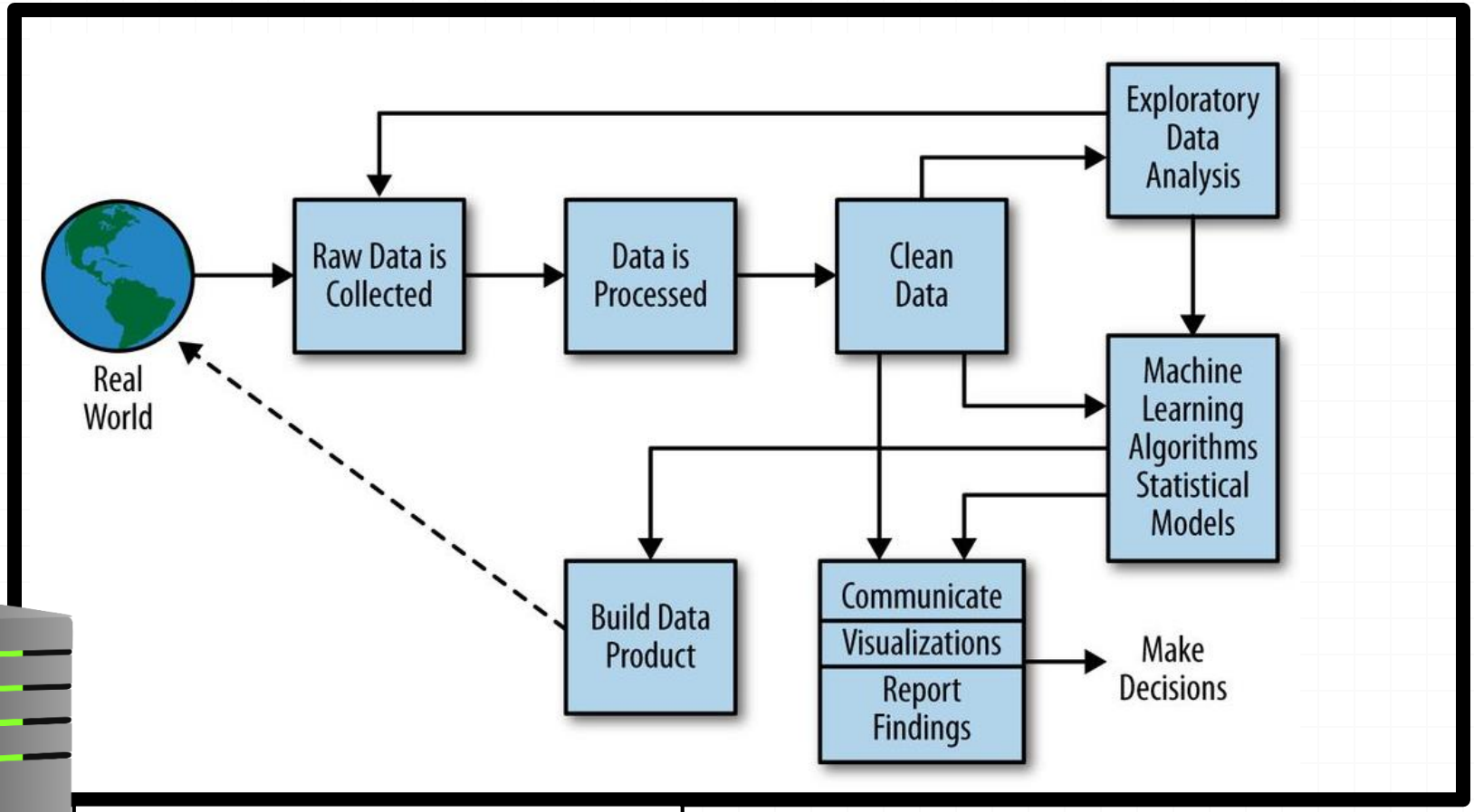
- Short Introduction
  - What you will learn in this course
- Logistics
  - What is expected of you in this course

Biomedical Data Science

- Biomedical Data Science
  - Data science, with emphasis on Biomedical data.

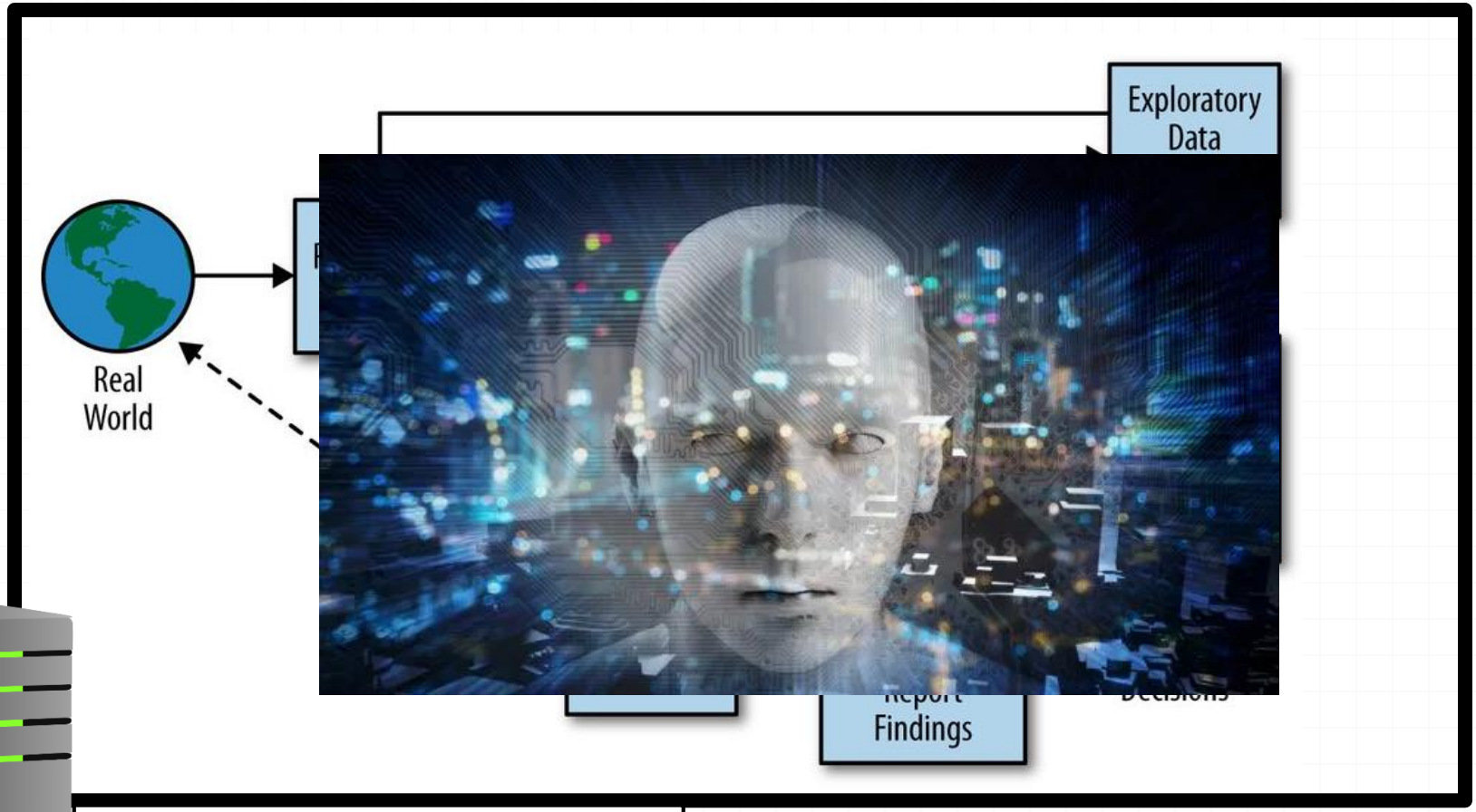




What is Data Science?



High Performance Computing

Beware of the Hype



Anatomy of a Data Scientist

Image credit: The IT staffing company

Data Scientist

- A relatively recent invention, coined by D. J. Patil and Jeff Hammerbacher in 2008
- Data science utilizes ideas from computer science, statistics, mathematics, engineering, physics, operations research, and library science
- Career
  - Hiring platforms Glassdoor and Indeed both list over 100,000 open data science jobs in January 2017

Davenport T, Patil DJ. Data scientist: The sexiest job of the 21st century. Harvard Business Review. 2012, Oct [cited 2017 Jul 12]. Available from: https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century

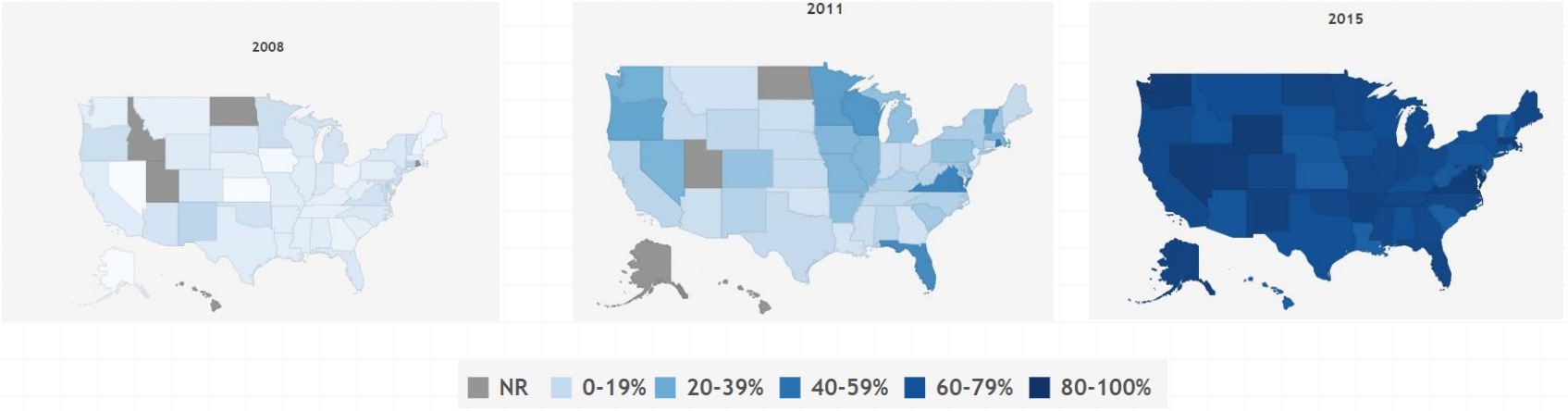
Why Biomedical Data Science?

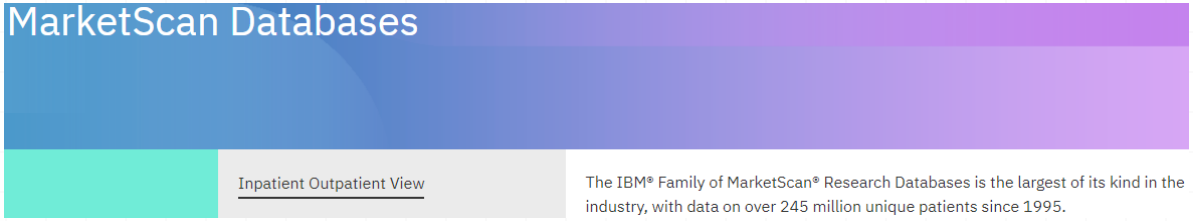
- Data, Data, Data

Clinical Data

- Adoption of Electronic Health record (HER) data

  Percent of non-federal acute care hospitals with EHR adoption





Example: MarketScan dataset
245 million patients

Physiological Data

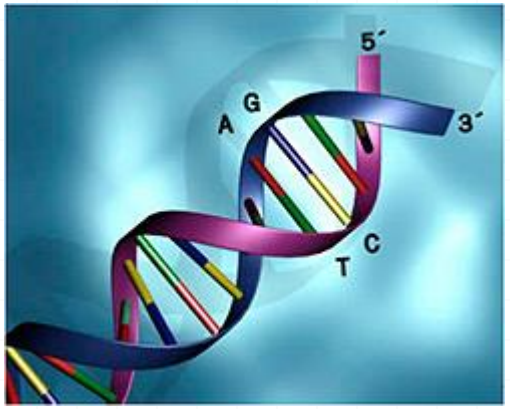
If you use MIMIC data or code in your work, please cite the following publication:

MIMIC-III, a freely accessible critical care database. Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, and Mark RG. Scientific Data (2016). DOI: 10.1038/sdata.2016.35.
Available from: http://www.nature.com/articles/sdata201635

Many Other Sources of Data

Biomedical Images

Wearables



Genomics

Social Media

Lab Tests

Main Topics Covered

- Python Programming
- Data Life Cycle
  - Plan, collect, explore, Quality Assessment, ..
- Biomedical Data: types, storage
- Biomedical Ontologies (e.g. ICD-9, SNOMED, …)
- Visualization and exploratory analysis
- Introduction to Machine learning libraries
- Introduction to Natural Language Processing libraries

Course Objectives

- At the end of the semester, students are expected to be able to:
  - Understand what biomedical data science is,
  - Identify different techniques used to solve biomedical data science problems,
  - Identify when and why a certain library or platform should be used,
  - Demonstrate the ability to apply methods from each of the major domains to solve practical problems.

Course Website

- Everything will be available on Canvas
- Canvas should be considered as the reference
  - All announcements, project postings, schedule changes, ..
  - Check your email!
  - Upload your photo

Class

- We will meet on the following days:
  - Tuesdays
    - One session: 1:55 – 2:45 pm
  - Thursdays
    - First session: 1:55 - 2:45 pm
    - Second session: 3:00 – 3:50 pm

Office Hours

- Office location: NEB 459
- Office hours: by appointment
- E-mail address: parisa.rashidi@ufl.edu
- NOTE: When contacting by email include "Course – 4760/6938" in the subject line to ensure delivery.

Textbook

- *Recommended*, not required
- Your main source: Lecture Notes

Grading

- Final grade is calculated according to:
  - Homework 20%
  - Paper Reading 35%
    - Graduate students: + literature review presentation
  - Exam 45%

Exam

- Exam 1: Thursday, Nov 29
- held during regular class hours

Important Dates

Date	Tuesday Class	Thursday Class
Week 1 (08/20 - 08/24)		Intro
Week 2 (08/27- 08/31)		Paper Reading
Week 3 (09/03 - 09/07)		Paper Reading
Week 4 (09/10 - 09/14)	HW1	Paper Reading
Week 5 (09/17 - 09/21)	HW1 Due	Paper Reading
Week 6 (09/24 - 09/28)	HW2	Paper Reading
Week 7 (10/01 - 10/05)	HW2 Due	Paper Reading
Week 8 (10/08 - 10/12)		Paper Reading
Week 9 (10/15 – 10/19)	HW3	Paper Reading
Week 10 (10/22 – 10/26)		Paper Reading
Week 11 (10/29 – 11/02)	HW3 Due	Paper Reading
Week 12 (11/05 – 11/09)	HW4	Paper Reading
Week 13 (11/12 – 11/16)	HW4 Due	Paper Reading
Week 14 (11/19 – 11/23)		Holiday
Week 15 (11/26 – 11/30)		Exam
Week 16 (12/03 – 12/07)		Reading Day

Quick Poll

- Programming experience
  - Python
  - Matlab
  - R
  - Other
  - No experience

- Introduce Yourself
  - Your name
  - Your major (your research topic?)
  - Why you enrolled in this class