

Syllabus - Intro To Data

- **Instructor:** Andy Choens, MSW
- **Course Name - Number:** Intro to Data
 - PBH210 (Intro To Data)
 - PBH211 (Intro To Data Lab)
- **Course Location - Time:** Thursday 17:30 - 19:30
- **Contact Information:**
 - ACPHS Email: Andrew.Choens@acphs.edu
 - Acuitas Email: Andy.Choens@AcuitasHealth.com

Mathematics is the language in which God has written the universe.

– Galileo Galilei

Class Schedule

Description	Date
Week 09	Mar 19, 2020
Week 10	Mar 26, 2020
Week 11	Apr 02, 2020
Week 12	Apr 09, 2020
Week 13	Apr 16, 2020
Week 14	Apr 23, 2020
Reading Day	Apr 29, 2019
Final Exam	Apr 30, 2019

Software

The tools used in this class are Free and Open Source Software (FOSS). We'll discuss what that means in class, but for now it can simply mean you don't have to pay anything to use them. The first lab includes time to install and configure these tools.

Students will use the following software:

Name	Recommended Download Link	License	Cost
R	Download R - CRAN	GPL v2	\$0.00
R Studio	Download RStudio Desktop	AGPL v3	\$0.00

Assignments & Grading

PBH210 Intro To Data

Table 3: Remaining Intro to Data Assignments

Assignments	Due Date
Speaking Truth To a Marmot	2020-04-09
Is Most Published Research Wrong?	
Final Exam	2020-04-30

PBH 211 Intro To Data Lab

Each lab builds on the content of that weeks lecture. Thus, each lab is assigned on the day of the lecture and is due the day before the next lecture. Because I recognize that things happen and that this is a lot to learn in a short amount of time, I will drop your lowest lab grade. But, be aware that you are still responsible for the material on that lab for the midterm/final exam.

Table 4: Intro to Data Labs

Description	Assign Date	Due Date
Lab 10	Mar 26, 2020	Apr 01, 2020
Lab 11	Apr 02, 2020	Apr 08, 2020
Lab 12	Apr 09, 2020	Apr 15, 2020
Lab 13	Apr 16, 2020	Apr 22, 2020
Lab 14	Apr 23, 2020	Apr 29, 2020

Tentative Lecture Plan

Week 09: Mar 19

- **Topics:**
 - Canceled
 - Don't get sick

Week 10: Mar 26

- **Topics:** Deep Breath
 - Group Check In/Challenges:
 - * Zoom!?! (Mea culpa regarding Wednesday, Recording)
 - * Using VMWare
 - * Using Local R
 - * Problems/Concerns?
 - Bias in data/measures:
 - * [Can crime be predicted with AI/ML?](#)
 - * [The Use of Artificial Intelligence in Gauging the Risk of Recidivism](#)

- * [Artwork Personalization at Netflix](#)
- * [Some Viewers Think Netflix Is Targeting Them by Race.](#)
- * Discuss: How are EHRs biased?
- * SpotterEDU
- * People gaming the system (Child Welfare Example)
- * There is a fundamental risk that we embed historical biases into our models simply because our existing data is based on and informed by these biases.
- RMarkdown:
 - * Basic formatting
 - * Code Chunks
 - * Tables (`kable`)

Week 11: Apr 02

- **Topics:**
 - T-Tests and P-Values
 - * How to calculate them in R?
 - * R's model formula.
 - * What do they *really* mean?
 - * When to use, when to be careful.
 - * If not this, then what?
 - Telling stories with data
 - More RMarkdown

Week 12: Apr 09

- **Topics:**
 - Intro to: Linear Regression
 - RMarkdown and Plots

Week 13: Apr 16

- **Topics:**
 - Intro to: Logistic Regression
 - Plotting

Week 14: Apr 23

- **Topics:**
 - More Advanced Regression
 - Final Review

Final Exam: Apr 30

- **Topics:** Final Exam and Pizza