

BIOMED SCI 552:

STATISTICAL THINKING

LECTURE I

COURSE DETAILS

- August 19th to Sept. 20th, 2024
 - Due to pre-planned travel, there will be no class on August 27th or 29th
 - Suggesting makeup classes on the 24th and 26th of September
 - These lectures will also be recorded
- Office hours are by request
 - Email: Eric.Lofgren@wsu.edu
- Grading:
 - Three problem sets, each worth 13.33% of your grade (40% total)
 - *Participation* is worth 30% of your grade
 - This class calls for active and engaged learning, and you are not guaranteed full participation marks
 - If you need an accommodation in this area, please arrange a meeting
 - A final project, worth 30% of your grade

PROBLEM SETS

- Problem sets will be assigned at the end of a Thursday class, to be due the following Thursday *prior to the start of class*
- Collaboration is encouraged on all problem sets
- While you can work in groups, the work must be individual to you – that is, it should be written in your own words, reflecting your own understanding of the solutions your group arrived at
- If you did work with a group, please list the names of the other students you worked with

THE FINAL

- The final project is to select a topic that is of interest to you, based on your own work, or as part of the material that we've covered.
- Identify a problem
- Discuss the prior research in this area – what is yet to be solved?
- Outline how a statistical approach might aid you in answering it
- All in the form of a 5-page paper
 - One-inch margins, reasonable font sizes, etc. apply.
- The final must be completed independently
- This is due two-weeks after our last class on Thursday, October 10th at 11:59 PM PST.

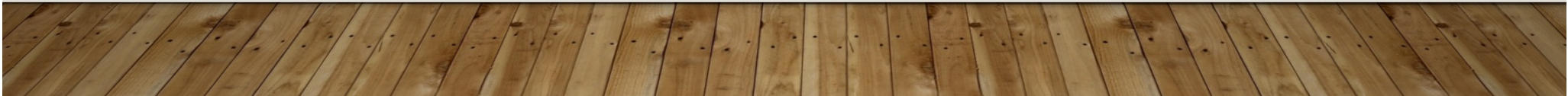
LATE WORK

- Assignments submitted after the beginning of class will receive one-half of the total graded points
- A late *final* will lose 10% off the final grade for each day it is late, beginning on October 11th at 12:01 AM PST.



ACADEMIC HONESTY

- Academic integrity is the cornerstone of higher education. As such, all members of the university community share responsibility for maintaining and promoting the principles of integrity in all activities, including academic integrity and honest scholarship. Academic integrity will be strongly enforced in this course. Students who violate WSU's Academic Integrity Policy (identified in Washington Administrative Code (WAC) 504-26-010(3) and -404) will fail the course, will not have the option to withdraw from the course pending an appeal, and will be reported to the Office of Student Conduct.
- Cheating includes, but is not limited to, plagiarism and unauthorized collaboration as defined in the Standards of Conduct for Students, WAC 504-26-010(3). You need to read and understand all of the definitions of cheating: <http://app.leg.wa.gov/WAC/default.aspx?cite=504-26-010>. If you have any questions about what is and is not allowed in this course, you should ask course instructors before proceeding. If you wish to appeal a faculty member's decision relating to academic integrity, please use the form available at <https://conduct.wsu.edu/>.



ARTIFICIAL INTELLIGENCE

- Artificial intelligence is an evolving field, with a great deal of promise and a great many pitfalls
- At the moment, LLMs *cannot* look up information – they use statistical models to predict what this sort of thing *should* be
- They excel at giving you answers that sound right, but are not
- At the moment, LLMs are extensively trained on data where the authors of that data have not given their consent, been compensated, etc.
- I will not prohibit the usage of AI tools in my class (if for no other reason than this is a rabbit hole of “what counts”), but use with caution

WRITE ME A SPEAKER INTRODUCTION FOR ERIC LOFGREN, THE EPIDEMIOLOGIST. DO NOT USE PLACEHOLDERS, AND USE SPECIFICS FOR WHERE HE WORKS, AND HIS AREAS OF EXPERTISE.

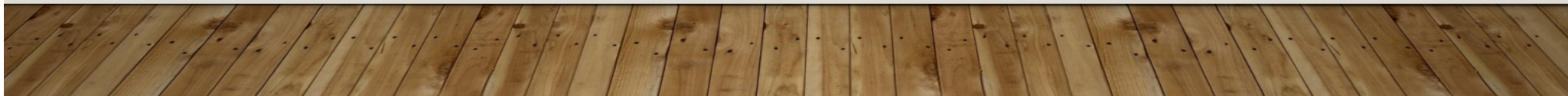
Ladies and gentlemen,

It is my distinct pleasure to introduce our next speaker, Dr. Eric Lofgren. Dr. Lofgren is an esteemed epidemiologist and assistant professor at the School of Public and International Affairs at Virginia Tech. He holds a Ph.D. in Epidemiology from the University of California, Berkeley, where he specialized in the mathematical modeling of infectious diseases.

Dr. Lofgren's research has made significant contributions to our understanding of epidemic dynamics and the global impact of infectious diseases. His expertise in mathematical modeling is crucial for forecasting and managing epidemic outbreaks, with a particular focus on diseases such as influenza and emerging pathogens. His innovative models help guide public health officials and policymakers in crafting effective response strategies.

In addition to his groundbreaking research, Dr. Lofgren is dedicated to bridging the gap between academic research and practical public health applications, ensuring that scientific discoveries are translated into actionable policies. His work not only advances the field of epidemiology but also plays a pivotal role in shaping responses to current and future public health challenges.

Please join me in welcoming Dr. Eric Lofgren.



WRITE ME A SPEAKER INTRODUCTION FOR ERIC LOFGREN, THE EPIDEMIOLOGIST. DO NOT USE PLACEHOLDERS, AND USE SPECIFICS FOR WHERE HE WORKS, AND HIS AREAS OF EXPERTISE.

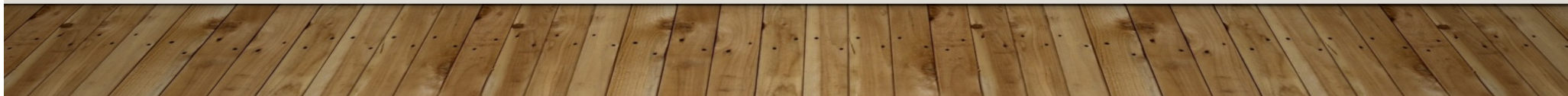
Ladies and gentlemen,

It is my distinct pleasure to introduce our next speaker, Dr. Eric Lofgren. Dr. Lofgren is an esteemed epidemiologist and **assistant professor** at the **School of Public and International Affairs at Virginia Tech**. He holds a **Ph.D. in Epidemiology** from the **University of California, Berkeley**, where he specialized in the **mathematical modeling of infectious diseases**.

Dr. Lofgren's research has made significant contributions to our understanding of epidemic dynamics and the global impact of infectious diseases. His expertise in mathematical modeling is crucial for forecasting and managing epidemic outbreaks, with a particular focus on **diseases such as influenza and emerging pathogens**. His innovative models help guide public health officials and policymakers in crafting effective response strategies.

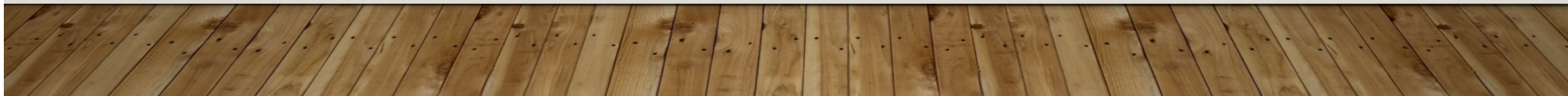
In addition to his groundbreaking research, Dr. Lofgren is dedicated to bridging the gap between academic research and practical public health applications, ensuring that scientific discoveries are translated into actionable policies. His work not only advances the field of epidemiology but also plays a pivotal role in shaping responses to current and future public health challenges.

Please join me in welcoming Dr. Eric Lofgren.



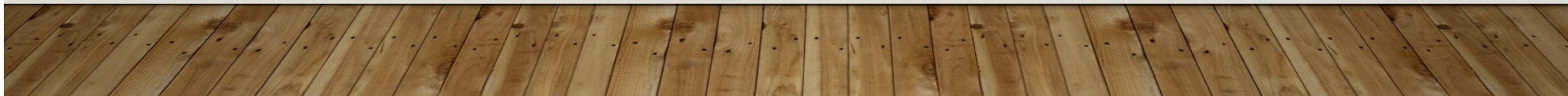
WRITE ME A SPEAKER INTRODUCTION FOR ERIC LOFGREN, THE EPIDEMIOLOGIST. DO NOT USE PLACEHOLDERS, AND USE SPECIFICS FOR WHERE HE WORKS, AND HIS AREAS OF EXPERTISE.

It is our pleasure to introduce Dr. Eric Lofgren, a renowned epidemiologist at the Washington State University Paul G. Allen School for Global Health, where he serves as an Associate Professor. Dr. Lofgren received his PhD in Epidemiology from the University of California, Davis, and has since become a leading expert in the field of zoonotic diseases, with a particular focus on the dynamics of disease transmission at the animal-human interface. His research has taken him to various parts of the world, from studying the epidemiology of influenza in Southeast Asia to investigating the risk factors of disease transmission in African wildlife. Today, Dr. Lofgren will share his insights on the complexities of emerging infectious diseases and global health security, drawing on his extensive experience in the field and his passion for improving health outcomes worldwide. Please join us in welcoming Dr. Eric Lofgren.



WRITE ME A SPEAKER INTRODUCTION FOR ERIC LOFGREN, THE EPIDEMIOLOGIST. DO NOT USE PLACEHOLDERS, AND USE SPECIFICS FOR WHERE HE WORKS, AND HIS AREAS OF EXPERTISE.

It is our pleasure to introduce Dr. Eric Lofgren, a renowned epidemiologist at the **Washington State University Paul G. Allen School for Global Health**, where he serves as an **Associate Professor**. Dr. Lofgren received his **PhD in Epidemiology** from the **University of California, Davis**, and has since become a **leading expert in the field of zoonotic diseases**, with a particular focus on the dynamics of disease transmission at the animal-human interface. His research has taken him to various parts of the world, from studying the epidemiology of **influenza in Southeast Asia** to investigating the risk factors of **disease transmission in African wildlife**. Today, Dr. Lofgren will share his insights on the complexities of emerging infectious diseases and global health security, drawing on his extensive experience in the field and his passion for improving health outcomes worldwide. Please join us in welcoming Dr. Eric Lofgren.

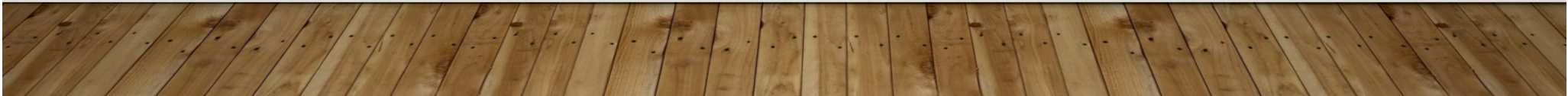


LAUREN'S PROMISE

- On October 22, 2018, Lauren McCluskey, 21 years old, was murdered by a man she briefly dated on the University of Utah campus, where she was a student. Lauren was raised in Pullman, Washington. Together with her parents, who are professors at WSU, this university community stands firmly behind Lauren's Promise: **WSU will listen and facilitate support and reporting options if someone is threatening you.**
- WSU prohibits discrimination and harassment. This includes discriminatory harassment, hate crimes, sexual discrimination, sex-based harassment, stalking, dating violence, domestic violence, sexual assault, and all types of sexual violence.
- If you are in immediate danger, call 911.
- If you have experienced or have witnessed discriminatory behavior, you can contact the WSU Compliance and Civil Rights (CCR) and/or the [WSU Title IX Coordinator](#). CCR can provide information on reporting options, including confidential resources available to you, and facilitate supportive measures. To contact CCR:
 - Online: [Online Reporting Form](#)
 - Email: ccr@wsu.edu
 - Phone: 509-335-8288
- For more information, see the WSU [Policy Prohibiting Discrimination and Harassment](#) (Executive Policy 15), WSU Standards of Conduct for Students ([Chapter 504-26 WAC](#)), and the [WSU Notice of Nondiscrimination](#).

ON EMAIL...

- Faculty are very busy people, and their inboxes are usually flooded
- I have 300+ unread emails on an ordinary Monday
- If I haven't answered you, it is 100% not because I hate you, think your email is dumb, or am silently judging you – it's because I looked at it, didn't get a chance to answer, and it got buried
- *Please* if I have not responded to an email from you about the course, email me *again*
- I will never be upset about a gentle reminder or nudging something urgent to the top of the pile



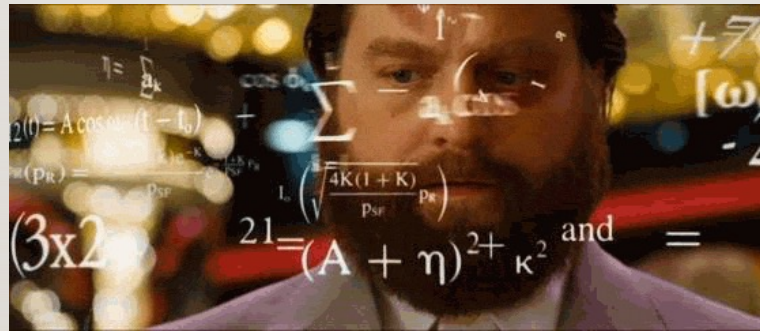
WHAT IS “STATISTICAL THINKING”?

- "The great body of physical science ... [is] only accessible and only thinkable to those who have had a sound training in mathematical analysis, and the time may not be very remote when it will be understood that for complete initiation as an efficient citizen ... it is necessary to be able to compute, to think in averages and maxima and minima, as it is now to be able to read and write." – H.G. Wells



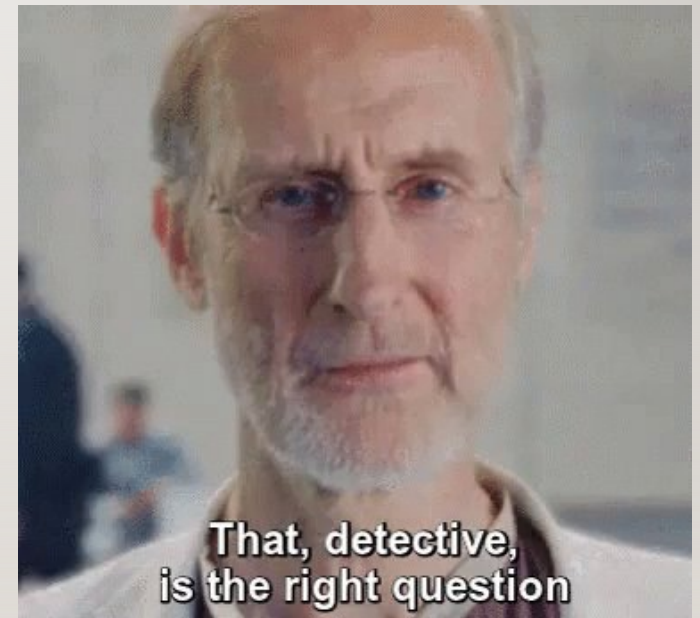
WHAT IS “STATISTICAL THINKING”?

- Statistical thinking is both a toolset and a mentality for looking at the world in a way where we are trying to understand both the world's underlying processes – and that there is uncertainty surrounding those processes
- It is not just *doing* statistics, but understanding why, and approaching the design of experiments in a way that will yield meaningful data



WHY THIS CLASS?

- An essential element of becoming an independent scientist is being able to articulate not only *what* you did, but *why*
- In your defense, etc. “Because [My PI] told me to...” is not really a sufficient answer
- I have seen grant proposals, etc. dinged for being vague about statistics and taking a “we’ll sort that out later” approach
- If you think about statistics from the outset of your research, you are likely going to generate better data that will be easier to analyze, and give you a clearer picture of what you’re doing
 - It’ll also help you ask the right question

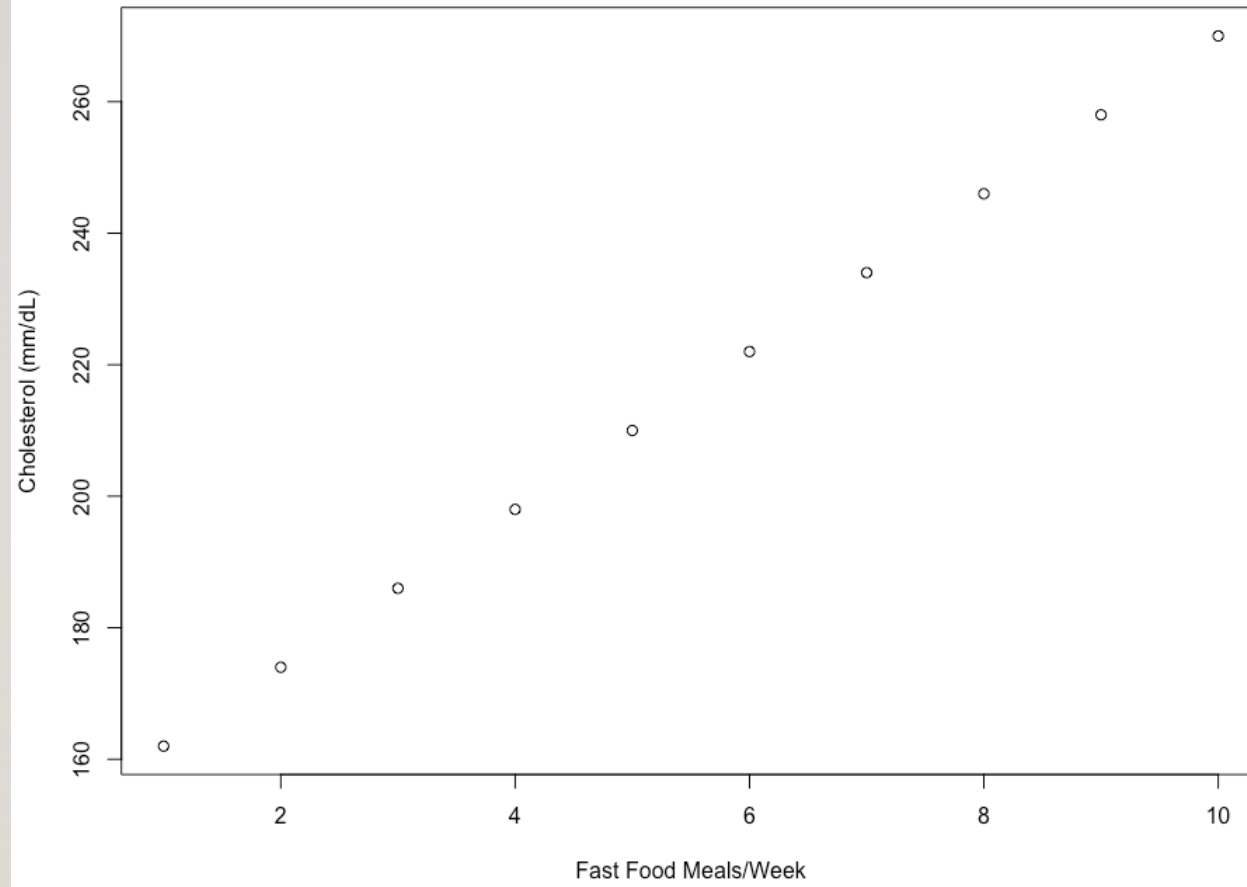


WHAT THIS CLASS ISN'T

- A “Stats” class
 - A single credit class isn't nearly enough to teach you the mechanics of statistical analysis in the biological sciences
 - Arguably, a *series* of classes isn't necessarily enough

WHY DO WE NEED STATISTICS AT ALL?

- A thought experiment: Does the consumption of fast-food cause high cholesterol
 - Measure cholesterol levels in human blood
 - Subjects self report the number of fast food meals they eat
 - Compare the cholesterol level in the blood by how many self reported meals



VARIATION IN THE BIOMEDICAL SCIENCES

- Biomedical science studies living organisms
- There's inherent variation in said organisms
- No two individuals are ever identical, and it's very rare that we can measure every individual we're interested in
- If you look into the history of statistics, a *lot* of it is motivated by biology – the two fields have always been connected



Kathy Collins
Boxing
5' 5" 137 lbs.

Jason Kidd
Basketball
6' 4" 212 lbs.

Annika Sorenstam
Golf
5' 5" 120 lbs.

Carlos Delgado
Baseball
6' 3" 225 lbs.

Vince Muñoz
Handball
5' 7" 165 lbs.



Tabitha Yin
Gymnastics
4' 8" 85 lbs.

Amy Acuff
High Jump
6' 2" 145 lbs.

Jennifer Parilla
Trampoline
5' 1" 120 lbs.

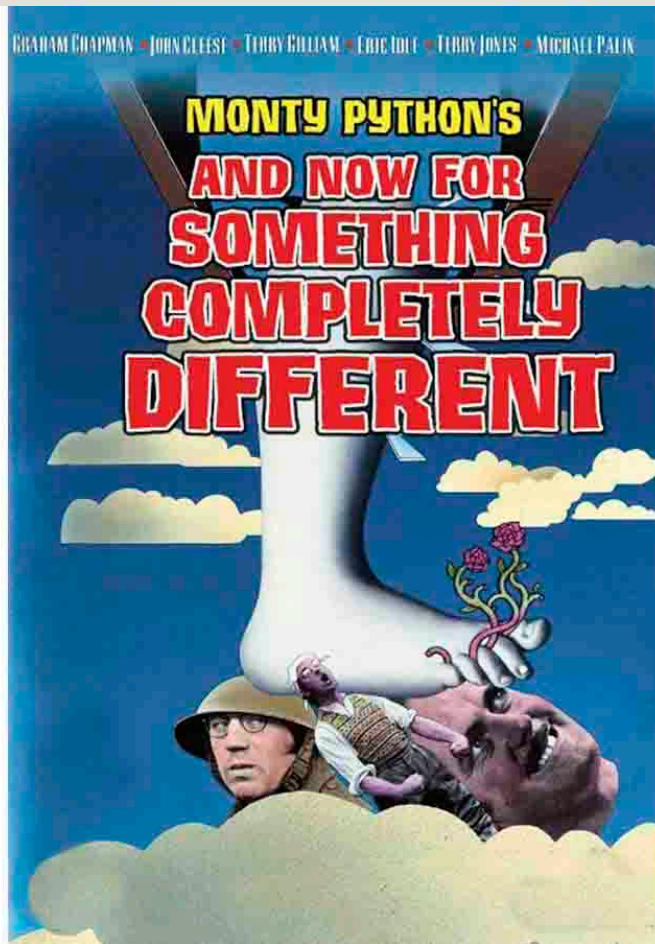
Charles Austin
High Jump
6' 0.5" 170 lbs.

Stacey Bowers
Triple Jump
5' 6" 130 lbs.

Cary Kolat
Wrestling
5' 5" 138 lbs.

"The Athlete" by Howard Schatz and Beverly Ornstein, 2002

QUESTIONS SO FAR?



COMPUTATIONAL TOOLS IN DATA

- Statistics involves coding
- But you can be good at statistics and *bad* at coding
- This is, if you ask some people, the norm
- There are some basics that are good practice for research code
- Three years after you graduate, you don't want to be looking for `analysis_final_final2_edited_finalfinal.R` somewhere on your laptop
- We solve this with something called *version control*

GITHUB

- Github.com is one of the major platforms for version control, and the one we're going to use in this class
- It is built on top of the version control system *git*
- It is *wildly* powerful, and has very sophisticated features, but even a basic use of it will improve your organization, let you undo mistakes, etc.
- Lots of resources for students



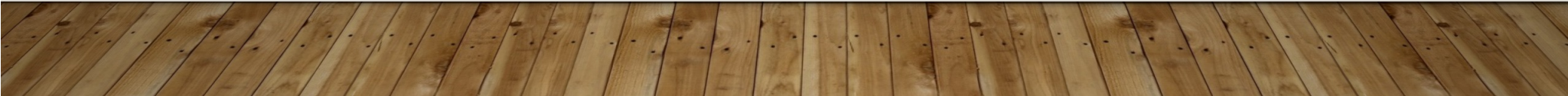
WHY USE VERSION CONTROL?

- Added redundancy for keeping your code safe
 - I used to work at an Apple store a very long time ago, you would be shocked how many times the Genius Bar got asked to try to save someone's dissertation
- Being able to go backwards to fix code, etc.
- Designed around multiple people being able to interact with a code base at the same time
 - No “can you stop editing the document so I can add my section?”, etc.
- Much easier to share than email



ESSENTIAL TERMINOLOGY

- Repository (“Repo”): A collection of code or other files, and the basic organizational unit of GitHub. Think of this like the main directory for a project
- Commit: A single addition of new code, changes, etc. to a repository
- Branch: A sort of side repository that can be worked on without making changes to the main repository, with the intention of them being reincorporated later
- Fork: Creating an entirely new repository based on an original repository where the intention is not necessarily to reincorporate them
- Push: Taking a commit and changing the main repository
- Pull Request: Asking the person who controls a repository to review and incorporate your code



Create Feature Branch
from Master Branch

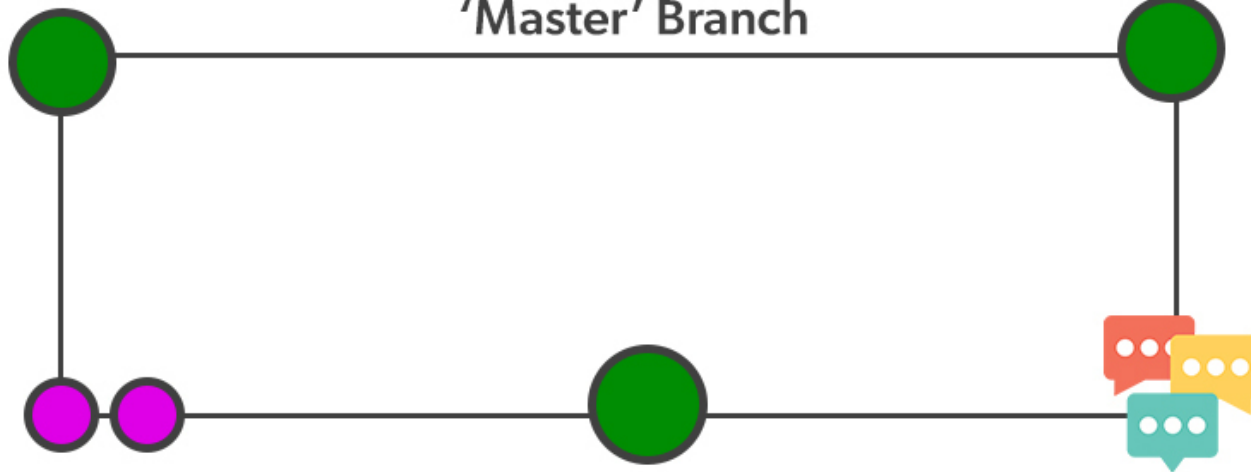
Merge Feature
branch onto Master branch

'Master' Branch

Commit changes
to feature Branch

Submit Pull Request

Discuss the
Proposed Changes



DEMO

