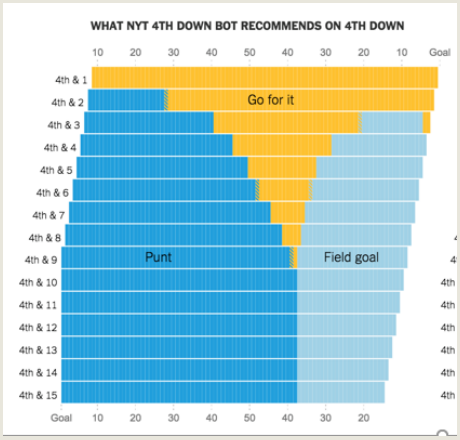
Discovery Seminar – Nerds Win Championships: Statistics in Sports

Instructor

Zachary Binney, PhD
Assistant Professor, Quantitative Theory and Methods
Office: Pierce 234
Email: zbinney@emory.edu
Office hours: By appointment M-F, 9a-5p



Time/Location

M/W, 11:30am-12:45pm, Seney 310
Zoom Address in case we can't meet in person

Contents

Course Description

Course Description

Do you have dreams of athletic glory, but not the body or skills to match? Good news! You don't have to thread dimes to a streaking wideout, hurl an unhittable cutter, nail a dagger 3, or bend it like Beckham to win championships…if you know a little math. Traditionally jocks and nerds are like oil and water, but they're increasingly more like chocolate and peanut butter. The athletes on the field make the plays, but behind the scenes coaches and - increasingly - statistical analysts are setting them up for success. Stats nerds figure out better ways to value and compare players, make smarter in-game decisions, and keep athletes healthy. In this course you'll question and learn from experts in various sports how statistics and geometry can be used to break down a soccer pitch, how to quantify the unquantifiable contributions of an offensive line, and whether LeBron is really better than MJ. You will develop the skills to generate and ask meaningful questions, answer those questions using a mix of quantitative and qualitative evidence, and communicate those answers to diverse audiences. You'll also create your own sports analytics project you can showcase to the public and teams.

Learning Objectives

Discovery Seminars Overall

1. Engage students in inquiry-based learning to begin to ask more meaningful questions, question and examine evidence more rigorously, and use evidence in argument more effectively.
2. Expand students' communication skills.
3. Develop students' information literacy through engagement with the library and other appropriate resources.

Our Course

1. **Improve** our overall problem solving and critical thinking
2. **Learn** how we learn from data asking meaningful questions and rigorously examining relevant evidence in sports
3. **Gain** an understanding of basic statistical/analytical concepts and their applications in sports, including: data acquisition, storage, and access; data analysis; and data visualization
4. **Conduct** a sports statistical analysis using R
5. **Practice** effectively communicating analytical findings, arguments, and conclusions to diverse audiences

Course Materials

Primary Course Resources

The primary materials that contain the information you will be responsible for will be varied. There is no textbook for this course. Instead, blog posts, papers, and book chapters, and even the occasional Twitter thread will be provided to you on Canvas. All will be provided to you freely.

Presentations and Labs

Each unit will contain 1-2 Powerpoint presentations and RMarkdown lab documents. These are all, of course, free and will be posted on the course website.

Computing Needs

A laptop or desktop is required for this course. Internet fast enough for videoconferencing (or at least streaming) is ideal for emergencies, but you can get by without it.

You do need either a Windows, Mac, or Linux machine. Chromebooks (specifically the chrome OS) cannot run the statistical programming language R we will be using.

The library has such laptops available for checkout, or there are other solutions! Please come talk to me if you don't have an appropriate laptop or desktop.

Video in Zoom Sessions

If we are in a synchronous Zoom session I would love to see your faces. But I understand that some of you may not have fast enough computers or internet service, or you may not be comfortable showing off where you are during class on video. So while I *strongly request* you use video during our Zoom sessions, please understand I do not consider it a *requirement*.

Course Structure and Online Sessions

Weekly Expectations and Schedule

The course meetings Mondays and Wednesdays. The core of the course is 10 weeks comprising 5 2-week modules on different statistical topics in sports (SEE Schedule below). Each of these 2-week modules will be structured as follows:

- First Monday: Overview of topic
    - Pre-Class Assignment: Readings
- First Wednesday: Practical R lab
    - Pre-Class Assignment: None
- Second Monday: Article Discussions
    - Pre-Class Assignment: Readings (selected articles)
- Second Wednesday: Guest speaker
    - Pre-Class Assignment: Guest speaker questions

Attendance

Attendance is important in this class, and my hope is >85% attendance on an average day.

That said, if you have to miss class occasionally, for example for a mental health day or for a longer time if you are quarantined, I completely understand. **And DO NOT COME TO CLASS if you are experiencing any symptoms – even mild symptoms! – of COVID-19, including coughing, loss of taste or smell, fever over 100F, difficulty breathing, congestion, sneezing, or any other respiratory symptoms. Email me before class if you are experiencing these.** I will make accommodations for you to not miss any materials, I promise, but it's much easier to do this ahead of time.

If you are missing for two straight sessions I'll likely check in just to make sure everything is OK.

Try to miss 3 or fewer classes (COVID symptoms excepted) all semester.

And if you anticipate wanting a letter of recommendation, please come to >90% of classes or let me know proactively why you're missing a large number of class sessions.

Assignments and Grading

Pre-Class Reflections (24%; 14 opportunities at 2% each, up to 2 missed assignments)

This is basically your participation grade. In many class sessions (in the 20 sessions across the 5 main units, all sessions *except* the R Practical Lab) you are expected to submit a Reflection response in that day's thread on the Canvas Discussion board. These should be ~2-4 sentences or 50-100 words and show you engaged intellectually with the readings in some way *or* have prepared for our guest speaker. <u>These do not need to be profound insights</u>. If you did the readings they should take you 5-10 minutes to write.

These are due **<u>on the Canvas Discussion board by 8:00am the day of class so I can review them beforehand</u>.**

Each Reflection is graded on a check (2 pts), check-minus (1 pt), or missing (0 pts) scale. If you receive a check-minus, you may attempt one revision within 48 hours. Your two lowest grades will be dropped.

Article Report + Analysis (6%)

Once during the semester you will be asked to identify a reading/analysis for class discussion. Alongside this you will also be asked to write a more detailed (1-2 page) report on the article. More details will be provided on the assignment itself.

Lab Homeworks (30%, 5 at 6% each)

After each practical R lab there will be a homework assignment to practice your skills. They are due at **<u>6:00pm on the days indicated in the course schedule below.</u>**

Project – 40%

The bulk of your grade will come from a sports analytics project you'll complete during the semester. More details are forthcoming, but it will involve a proposal; a presentation; and a written report.

Below is the grading scale for this class. Grades are rounded to the nearest point (89.50% = 90%, 89.49% = 89%) at the end of the semester. I maintain the right to curve the class at my discretion, but it will only be done to *boost* grades (i.e. make each cutoff easier to reach), never the reverse.

A		B			C			D			F
	–	+		–	+		–	+		–	59%-0%
105%-93%	92%–90%	89%–87%	86%–83%	82%–80%	79%–77%	76%–73%	72%–70%	69%–67%	66%–63%	62%–60%	

Questions and Communication

Communication is very important in online classes to foster community and maintain good relations between students and instructors.

Asking for Help

There are 3 options for asking questions in this class, listed in order of preference:

1. **Canvas Discussion board:** If you have a question about a course topic or course logistics, simply create a post here. If it's a question or thought about course material, you may add it to a thread for that day if it already exists.

   This is our preferred communication option. It is where all questions about course material and assignments should go. In short, anything that could conceivably be useful to another student should go here. This reduces duplicated efforts, encourages interacting with your fellow students, and puts everyone on a level playing field. It's win-win-win!

   I understand some students may be shy, but please know: there are *no* stupid questions, and I have a zero tolerance policy for any bullying or mockery.

2. **Office Hours:** You can always ask questions about or for extra help with the course material in office hours, though I might ask you to consider making a Discussions board post with what we talked about if it'll be helpful for others.

   Office hours are also a great option if you have questions about material that extends beyond this course, want to talk about careers, mentoring, academic advice (e.g. you're curious about Quantitative Theory and Methods), or simply want to introduce yourself so I know you better

later for things like letters of recommendation or research opportunities.

I didn't often take advantage of office hours when I was an undergraduate because I didn't really understand what they were for and slightly regretted it later on.

Check out this great video on what office hours are for and common misconceptions! (Though also note this semester I am trying a flexible appointment-only office hours system.)

3. **Email:** The third option. If it's a question that will be useful to other students, I will ask you to put it on Canvas Discussions instead. Issues appropriate for email include questions specific to your or your group's assignment or project, or other personal matters that need to not be broadcast to the rest of the class (e.g. accommodation requests, absence explanations).

For emails and Discussions board posts, please allow 24 hours for a response from me, or 48 hours on weekends (your fellow students may be quicker, but don't take what they say as 100% correct until I sign off on the response). Please plan accordingly.

A Statement on Empathy

(Text modified from Prof. Andrew Heiss at Georgia State.)

Life is rough right now. None of us is really doing *well*. We're all just hanging on, myself included.

The pandemic has upended all of our lives in various ways, from lost jobs and economic uncertainty to health problems. If it helps, I graduated straight into the Great Recession in 2008 and came through it OK (though many people had a much rougher time). Seriously, I got hired at a consulting firm 2 months before Lehman Brothers collapsed. So you're not doomed.

I'm fully committed to making sure that you learn everything you were hoping to learn from this class! I will make whatever accommodations I can to help you finish your exercises, do well on your projects, and learn and understand the class material. Under ordinary conditions, I am flexible and lenient with grading and course expectations when students face difficult challenges. **Under pandemic conditions, that flexibility and leniency is intensified.**

If you tell me you're having trouble, I will not judge you or think less of you. You never owe me personal information about your health (mental or physical). However, **in order to most effectively help you I DO need to know when something is going on proactively (meaning ASAP).** If you need a special extension on assignments for extenuating circumstances, that's much easier for me to do at the time than 6 weeks later at the end of the semester, when it may be impossible for you to catch up anyway. Does that make sense?

You are also always welcome to talk to me about things that you're going through, though. If I can't help you, I usually know somebody who can – such as Counseling and Career Services. I have particular experience with close family serious sickness and death when I was your age, as well as family mental health and substance abuse issues.

If you need extra help, or if you need more time with something, or if you feel like you're behind or not understanding everything, do not suffer in silence! Talk to me. I will work with you. I promise.

I want you to learn lots of things from this class, but I primarily want you to stay healthy, balanced, and grounded.

Diversity

Diversity and Inclusivity

Oxford College of Emory University's ideals of inclusivity require that we foster an environment where people of diverse backgrounds, identities, abilities, and ideologies are affirmed, respected, and seen as a source of strength; where we strive to learn together, and ultimately thrive communally. If we at all fail to support these ideals, then we encourage discussion towards improvement, and we hope that this statement affirms your right to seek those discussions via dialogue with faculty, staff, and your peers.

I try my best to create a welcoming and inclusive environment for all students regardless of race, gender and gender identity, religion, immigration status, nationality, parental status, or age. If I ever fall short of that, I encourage you to approach me personally after class, during office hours, or via email and help me understand how I can do a better job of that. There are also college-level resources to express concerns if you would prefer not to approach me personally.

Kids in the Classroom

If you're a parent and you ever have a childcare issue and need to bring your child to class, I completely understand and welcome their presence. Fair warning, though: I retain the right to make funny faces at them.

COVID-19 and Other Long-Term Health Issues

If you are placed under quarantine for COVID-19 or are going through any longer-term health or personal issues that may affect your success in the class, please let me know. In addition – or alternatively, if you're more comfortable – I invite you to coordinate with the Advising Support Center (ASC) oxacadsvc@emory.edu. They are very helpful.

Access and Disability Resources

As the instructor of this course I endeavor to provide an inclusive learning environment. I want every student to succeed. The Office of Accessibility Services (OAS) works with students who have disabilities to provide reasonable accommodations. In order to receive consideration for reasonable accommodations, please contact the OAS and complete the registration process. Faculty may not legally provide you with accommodations until an accommodation letter has been processed and discussed with them; accommodations do not start until this point and are not retroactive, so **I ask you to let me know about accommodations ASAP so I can best help you**.

Students registered with OAS who receive a letter outlining specific academic accommodations are invited to immediately coordinate a meeting with their professors to discuss a protocol to implement accommodations that will (or may) be needed over the course of the semester. This meeting should occur **as early in the term as possible**. Contact OAS for more information at (770) 784-4690 or oas_oxford@emory.edu.

Class Session Recording

From time to time, particularly but not limited to if we are on Zoom, class sessions may be recorded for students in the class to refer back to the information, and for enrolled students who are unable to attend live. By attending class either in-person or on Zoom you are consenting to video or audio recordings.

Please read the Rules of Zoom Engagement for further advice on participating in any Zoom class sessions.

Academic Integrity

There are two key principles for this course (and for Oxford generally):

- Your work should represent your *own* effort and thought. That means…
    - Do not plagiarize (copy someone else's work or your own work without attribution).
    - Do not turn in things that include no original thoughts of yours or that are entirely someone else's work, even if that work is cited.
        - If this were a pottery class with an assignment to create a coffee mug, it would certainly be OK for you to get the clay, sculpt the mug, put it in a kiln, paint it, and turn it in.
        - It's *also* OK to say "Hey, I found this plain unpainted mug from Bridgette in the kitchen cabinet," paint it yourself, and turn that in while making it clear Bridgette made the original mug. You made a contribution to that mug, building off someone else's work.
        - It is NOT OK to go to the store, buy a pre-made and painted mug, and turn that in as your work (or even while making it clear you just went out and bought it at a store). In this metaphor this is like copying an assignment from a friend.
        - It is also NOT OK to just give me a mug you made for a prior class. You need to at least re-paint it.

- In the case of the group project, anything for which you receive a grade should contain *meaningful intellectual input* from you. That is, you should not receive credit for a group project if your group members did all the hard work.

General Statement: Upon every individual who is a part of Emory University falls the responsibility for maintaining in the life of Emory a standard of unimpeachable honor in all academic work. The Honor Code of Emory College is based on the fundamental assumption that every loyal person of the University not only will conduct his or her own life according to the dictates of the highest honor, but will also refuse to tolerate in others action which would sully the good name of the institution. Academic misconduct is an offense generally defined as any action or inaction which is offensive to the integrity and honesty of the members of the academic community. The typical sanction for a violation of the Emory Honor Code is an F in the course. Any suspected case of academic misconduct will be referred to the Emory Honor Council.

Course Schedule (Fall 2021)

25-Aug	Introduction to and Overview of Stats in Sports	These blog posts (all <5 minutes): Why Sports Analytics Team Uses Part I Team Uses Part II League Uses Media Uses Sports Business Analytics (skim)		Broad applications of sports analytics
30-Aug	Overview of Milestone Project + Begin Introduction to R	Install R and RStudio (complete Tutorial 0) Milestone pre-class materials Guest lecture: Peter McLellan		R scripts and RMarkdown Rectangular data
1-Sep	Introduction to R (main session)	--		Exploring data in R Data manipulation in R
6-Sep	**LABOR DAY! No class, and solidarity forever.**			

8-Sep	Library I: Assessing Source Reliability	Guest lecture: Paige Crowl	Lab HW 1	Critically evaluating sources of data, analyses
13-Sep	Baseball: Overview	IF NOT FAMILIAR WITH BASEBALL: Watch this 3.5 minute video The Sabermetric Manifesto Fangraphs Sabermetric Glossary Library: Getting Started The Many Flavors of DIPS (NOTE: BABIP = "Batting Average on Balls in Play" = how often does a "hit" ball put a batter on base) **OPTIONAL BUT RECOMMENDED**: A Sabermetric Primer (The Athletic) **OPTIONAL:** Pitching and Defense: How Much Control do Hurlers Have (McCracken) **OPTIONAL:** Explore the Fangraphs Library of stats to your heart's content	Reflection	Player valuation Asking analytics questions What makes a good statistic?

15-Sep	Baseball: R Lab			Correlation Simple linear regression
17-Sep			Baseball Article Discussion Choices	
20-Sep	Baseball: Guest Lecture "Cognitive Biases in Baseball and Beyond" Daniel Adler, Assistant GM, Minnesota Twins	Complete Survey (TBA) Do reading OR listen to podcast episode listed in module	Reflection	Cognitive biases

22-Sep	Baseball: Article Discussions	1) The "Season" Barry Bonds Hit 95 Homers 2) Introducing SIERA: Part I 3) Why Baseball Revived a 60-Year Old Strategy Designed to Stop Ted Williams (The Shift) **OPTIONAL ADDITIONAL SHIFT READINGS:** 4) Baseball Therapy: Why the Shift Persists 5) We've Reached Peak Shift	Reflection	TBD

27-Sep	Basketball: Overview	(I know this looks like a lot, but most of these are quite short and accessible!) 1) IF NOT FAMILIAR WITH BASKETBALL: Watch this 5-minute video 2) Introduction to Oliver's Four Factors (read through "Measuring Impact on Success" section) 3) An Introduction to Advanced Basketball Statistics: Individual Statistics 4) An Introduction to Advanced Basketball Statistics: Team Statistics 5) *Basketball on Paper* by Dean Oliver, Ch. 24 (pp. 337-42). Available via Emory library. Search for book title, find online access, navigate to it, select chapter 24. 6) 5.5-minute video of Bomani Jones and Pablo Torre discussing Jalen Rose, race, and NBA analytics 7) How Mapping Shots in the NBA Changed it Forever (Goldsberry, FiveThirtyEight) 8) This New NBA Stat Is A Huge Step Forward For Basketball Analysis (Tracking Data, EPVA) **OPTIONAL:** Nylon Calculus 101: Plus-Minus and Adjusted Plus-Minus What is the Best Advanced Statistic for Basketball? NBA Executives Weigh In	Reflection	

29-Sep	Basketball: Overview cont'd		Lab HW 2	
1-Oct			Basketball Article Discussion Choices	
4-Oct	Basketball: R Lab	1) Steph Curry vs. Kevin Durant 2) Nylon Calculus: The Impact of Passing on Offense and Winning 3) MJ vs. LeBron (oh, yes) 4) How Shot-Tracking Is Changing The Way Basketball Players Fix Their Game	Reflection	
6-Oct	Basketball: Guest Lecture		Reflection	
13-Oct	Library II: Finding Data Basketball: R Lab (cont'd)	Review materials in Module prior to class Guest lecture: Paige Crowl		Finding Sports Data
18-Oct	Football: Overview	1) IF NOT FAMILIAR WITH FOOTBALL: Watch this 6-minute video. 2) How One Advanced NFL Statistic is Going Mainstream (EPA)	Reflection	

		3) 2021 NFL Analytics Survey 4) A New Analytics Tool for Fourth Down [and] 2-Point Conversions 5) R for NFL Analysis (just review the plots Mike Lopez created) 6) The 10 commandments of numbers-based football analysis **OPTIONAL**: 7) Peruse the websites of Next Gen Stats, Pro Football Focus (PFF), Pro-Football-Reference (PFR), and/or Football Outsiders 8) Pass Rush and Pass Block Win Rate Metrics Explainer (ESPN) 9) Rethinking Draft Curves		
20-Oct	Football: Guest Lecture		Reflection Lab HW 3	
22-Oct			Football Article Discussion Choices	
25-Oct	Football: Overview (cont'd)			
27-Oct	Football: R Lab			

1-Nov	Football: Article Discussions	1) Rushing Success and Play-Action Passing 2) Justin Herbert and Lamar Jackson Are Crushing It...On Different Downs 3) Adjusted Interceptions 2020 4) When to go for Two, for Real	Reflection	
3-Nov	Basketball: Article Discussions			
8-Nov	Soccer: Overview	1) IF NOT FAMILIAR WITH SOCCER: This 4-minute video 2) xG Explained 3) Soccer is Learning to See the Whole Game 4) The Athletic's Football Analytics Glossary (PDF if link doesn't work - ignore comments that fill last 20 pages) 5) Possession is the Puzzle of Soccer Analytics. These Models Are Trying to Solve It. **OPTIONAL** How Our Club Soccer Predictions Work (FiveThirtyEight)	Reflection Lab HW 4	
10-Nov	Soccer: Guest Lecture Maram AlBaharna, Analytics FC		Reflection	

12-Nov			Soccer Article Discussion Choices	
15-Nov	Soccer: R Lab			
17-Nov	Soccer: Article Discussions	1) How Real Madrid Built the Best Defense in Europe 2) Is Soccer Wrong About Long Shots? 3) On-Ball Value (OBV) in the 2020/21 Premier League	Reflection	
19-Nov			Sports Biz Article Discussion Choices	
22-Nov	Sports Biz: Guest Lecture + Overview Michael Lewis, Emory Marketing Professor		Reflection Lab HW 5	
29-Nov	Sports Biz: Article Discussions	1) https://www.scottkaplan.org/post/understanding-the-economics-of-the-nba-regular-season 2) The European Super League: A SportBusiness Media analysis	Reflection	

1-Dec	Project Workday, and/or maybe Sports Biz: Guest Lecturer
6-Dec	Final Project Presentations + Conclusion and Wrap-Up
9-Dec	**Final Project Reports Due!**

Date	**Topics**	**Readings & Pre-Class Prep**	**Assignments Due**	**Key Takeaways**
25-Aug	Introduction to and Overview of Stats in Sports	These blog posts (all <5 minutes): Why Sports Analytics Team Uses Part I Team Uses Part II League Uses Media Uses Sports Business Analytics (skim)		Broad applications of sports analytics
30-Aug	Overview of Milestone Project + Begin Introduction to R	Install R and RStudio (complete Tutorial 0.1) TBA Guest lecture: Peter McLellan		R scripts and RMarkdown Rectangular data
1-Sep	Introduction to R (main session)	--		Exploring data in R Data manipulation in R
6-Sep	**LABOR DAY! No class, and solidarity forever.**			
8-Sep	Library I: Assessing Source Reliability	Guest lecture: Paige Crowl	Lab HW 1	Critically evaluating sources of data, analyses

13-Sep	Baseball: Overview	IF NOT FAMILIAR WITH BASEBALL: Watch this 3.5 minute video The Sabermetric Manifesto Fangraphs Sabermetric Glossary Library: Getting Started The Many Flavors of DIPS (NOTE: BABIP = "Batting Average on Balls in Play" = how often does a "hit" ball put a batter on base) **OPTIONAL BUT RECOMMENDED**: A Sabermetric Primer (The Athletic) **OPTIONAL:** Pitching and Defense: How Much Control do Hurlers Have (McCracken) **OPTIONAL:** Explore the Fangraphs Library of stats to your heart's content	Reflection	Player valuation Asking analytics questions What makes a good statistic?
15-Sep	Baseball: R Lab			Correlation Simple linear regression
17-Sep			Baseball Article Discussion Choices	
20-Sep	Baseball: Guest Lecture		Reflection	Cognitive biases
22-Sep	Baseball: Article Discussions		Reflection	TBD

27-Sep	Basketball: Overview	IF NOT FAMILIAR WITH BASKETBALL: Watch this 5-minute video *The Midrange Theory* by Seth Partnow, Appendix (on Canvas) This 5.5-minute video of Bomani Jones and Pablo Torre discussing Jalen Rose, race, and NBA analytics An Introduction to Advanced Basketball Statistics: Individual Statistics An Introduction to Advanced Basketball Statistics: Team Statistics Introduction to Oliver's Four Factors (through "Measuring Impact on Success" section) How Mapping Shots in the NBA Changed it Forever Nylon Calculus 101: Plus-Minus and Adjusted Plus-Minus This New NBA Stat Is A Huge Step Forward For Basketball Analysis (Tracking Data, EPVA)	Reflection	

		What is the Best Advanced Statistic for Basketball? NBA Executives Weigh In *Basketball on Paper* by Dean Oliver, Ch. 24 (pp. 337-42). Available online via library.		
29-Sep	Basketball: R Lab		Lab HW 2	
1-Oct			Basketball Article Discussion Choices	
4-Oct	Basketball: Article Discussions		Reflection	
6-Oct	Basketball: Guest Lecture		Reflection	
13-Oct	Library II: Finding Data	TBA Guest lecture: Paige Crowl	Lab HW 3	Finding Sports Data
18-Oct	Football: Overview	https://www.thescore.com/nfl/news/2193857/how-one-nfl-advanced-statistic-is-going-mainstream	Reflection	
20-Oct	Football: Guest Lecture		**Project Proposal!**	
22-Oct			Football Article Discussion Choices	
25-Oct	Football: R Lab		Reflection	
27-Oct	Football: Article Discussions		Reflection	
1-Nov			Reflection	
3-Nov			Lab HW 4	

5-Nov			Unit 4 Article Discussion Choices	
8-Nov			Reflection	
10-Nov	Guest Lecture		Reflection	
15-Nov			Reflection	
17-Nov			Lab HW 5	
19-Nov			Unit 5 Article Discussion Choices	
22-Nov			Reflection	
29-Nov	Guest Lecture		Reflection	
1-Dec	Project Workday			
6-Dec	Final Project Presentations + Conclusion and Wrap-Up; **Lab HW 6 Due**			
9-Dec	**Final Project Reports Due!**			