****

Syllabus | EW/MBA257-4 | Spring 2022

People Analytics

**Professor Douglas Guilbeault**

[douglas.guilbeault@haas.berkeley.edu](mailto:douglas.guilbeault@haas.berkeley.edu)

Office: F543

Office Hour: Mondays 3:00-4:00 PM (please confirm by email prior to arrival)

**\****This course is dual-listed between the EWMBA and FTMBA Programs*

**UNITS OF CREDIT**: 2

**MEETING DAY(S)/TIME:** TBD 6:00-9:00PM (Break 7:30-7:50PM).

**CLASS LOCATION:** TBD

[**Privacy and Recording Notice**](https://haas.berkeley.edu/wp-content/uploads/Privacy-and-Recording-Notice-1.pdf)

**PREREQUISITE(S):** None

**CAREER FIELD:** This course will equip students to understand cutting edge developments in data analytics applied to people management ***from a managerial perspective***, and is therefore ideal for anyone seeking a career as a people manager, human resources professional, data scientist, business leader, or CEO. The concepts from this course are relevant to the management of organizations in a wide variety of areas, including business, non-profit, government, law, education, and research.

**CLASS FORMAT (SUMMARY):** *This class will consist of lectures, discussions, weekly readings, group work (e.g., collaborative analyses of case studies), and guest speakers from leading Silicon Valley tech companies.* Every class will end with bullet-point takeaways for students to carry with them in their careers. The final grade will be divided between class attendance and an individual project. In this final project, students will be asked to assume the (hypothetical) role of a People Analytics manager in a large firm and to provide a proposal for how to design/manage the use of an algorithm for an organizational problem. For this final project, students will have the option to specialize in either (1) a design approach, focusing on an in-depth analysis of an algorithmic process from a managerial perspective, or (2) a data science approach, where students will be given data on an algorithmic process to analyze statistically for the purpose of informing a managerial decision.

**ABSTRACT OF COURSE**:

Recent advances in People Analytics – driven by big data, machine learning, and computational social science – have unlocked fundamentally new ways to measure, optimize, and harness the performance of people in organizations, improving everything from hiring and promotion, to collaboration and resource management. However, effectively riding the wave of People Analytics requires confidence with the core concepts of data science as it is applied in managerial contexts. This means knowing how to answer difficult questions such as: *What are the right questions to ask when evaluating algorithms? When can we trust data to help us make decisions? How can we determine that we have the right data for our decision? When do we need more data, and when can additional data cause more harm than good? How can we make the best decision based on limited time and data? What kind of data is most likely to profit an organization? When is human judgment more important than algorithmic predictions? How can we combine human and algorithmic intelligence to improve managerial decisions?* These are some of the practical questions that arise when data is put to action in organizations, and these are precisely the kinds of questions that students in this class will learn how to address.

While each class will discuss data science methods and link to supplementary readings for more technical explorations, the classes themselves will NOT focus in great detail on the technical mechanics of machine learning, artificial intelligence, or advanced statistical modeling. Rather, we will learn how to understand these data science methods as ***managerial tools***, attending to the vital role that human judgments and decisions play in determining the validity and value of data science techniques for people managers. With a unique focus on recent advancements in computational social science, this class will serve both as an effective introduction to People Analytics for those learning for the first time, as well as a powerful new way to think about People Analytics for those who have prior experience with the topic, either through the workplace or related classes.

In this course, you will gain command of the core statistical concepts needed to measure and understand measurements of people in organizations. You will learn how to interpret statistics to make critical recommendations to (or as) senior leaders, and how to tell a compelling story using People Analytics data. This will involve learning the fundamental tools of analytics (descriptive statistics, correlations, etc.), building and testing HR chat bots, reviewing and conducting randomized experiments, interpreting and producing data visualizations, all while accounting for ethical concerns such as data privacy and biases (human and algorithmic). This course will prepare you to present People Analytics outcomes verbally, visually, and in a written form that would be expected in a business environment.

**COURSE THEMES**

Throughout the course, we will engage with four overarching themes.

I. **Understanding algorithms.** What do algorithms do for organizations and what are they incapable of doing? There are many jargony terms – big data, machine learning, deep learning, artificial intelligence, data analytics, natural language processing (to name only a few) – that are often used interchangeably and imprecisely. We refer to these as “**algorithmic methods.**” We will discuss what these terms mean, what types of algorithmic methods exist, and how, broadly speaking, they work. In discussing these methods, we will focus on the difference between an algorithm’s **precision** and **bias**, as well as why statistical bias and human social bias are related but different.

II. **Applying algorithms.** How do algorithmic methods inform organizational processes? We will begin by examining discrete events: hiring, performance evaluation and departure. Later we will engage with more complex organizational processes including culture, interpersonal interaction, social networks, collective intelligence, and team performance. We engage with some of the most cutting-edge algorithmic research addressing these organizational processes.

III. **Designing and evaluating experiments.** How can people managers run experiments to study and optimize organizational behavior and performance? When should people managers consider using experiments rather than correlational algorithmic approaches, and vice versa? We will discuss cutting edge techniques for experimental design in People Analytics involving both individuals and social networks, with special attention to work on collective intelligence in organizations. We will discuss how to interpret experimental effects, and we will review leading techniques currently being deployed by organizations to run both offline and online experiments.

IV. **The Ethics of Analytics.** What are the ethical and societal implications of data analytics in people management? A major theme in this course is that algorithms are managerial tools, not substitutes for human management – and this is especially clear in considering the vital role that carefully designed experiments continue to play in analytics. It is key for managers to understand how data science tools work; but that does not mean they should cede judgement to these tools. An important reason is that algorithms in the workplace have ethical and societal implications that go well beyond the scope of the algorithms and the intentions of their programmers. Aspects of these ethical problems are technical in nature, and we will discuss these technical features. But ultimately, the more important aspects of this problem are human and social in nature. We will discuss these ethical implications specifically when thinking about what “algorithmic fairness” is and how it relates to inequality and diversity across a wide range of organizational contexts.

**BIOGRAPHICAL SKETCH:**

Douglas Guilbeault is an Assistant Professor in the Management of Organizations at the Haas Business School, and is a leading expert in People Analytics and Computational Social Science. His publications have appeared in top scientific journals, including *Nature Communications*, *The Proceedings of the National Academy of the Sciences*, and *Management Science*; and his studies have received top research awards from numerous organizations, including the *International Conference on Computational Social Science*, the *International Communication Association*, and the *Cognitive Science Societ*y. Douglas was among the first to expose the use of algorithmic bots on social media to manipulate elections around the world, and his exposés have appeared in prominent venues including the *Atlantic* and *Wired*. He has also served as a People Analytics consultant for various organizations, including Google and Facebook. Recently, he was the winner of Stanford’s “Art of Science” competition for his piece, “Changing Views in Data Science of 50 Years.”

**RECOMMENDED (BUT NOT REQUIRED) TEXTS[[1]](#footnote-1):**

Agrawal, Gans, and Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press.

Mitchell. 2019. *Artificial Intelligence: A Guide for Thinking Humans*. Farrar Pub.

Salganik. 2018. *Bit by Bit: Social Research in the Digital Age*. Princeton University Press.

# CLASS FORMAT (STRUCTURE AND PREPARATION)[[2]](#footnote-2)

Each class will, most commonly, consists of lecture, discussion, in-class exercise, or a visit from a guest speaker. In-class exercises will be done in small groups and will typically include either interpreting data analysis results or designing a data analytic process. Readings include a combination of cases, popular press and scientific articles.

The number of readings per session is relatively small. But many of the readings are academic papers, some of which are challenging to read. These represent cutting-edge research at the intersection of algorithmic methods and the study of organizations and human decision-making. Your exposure to these recent innovations will inform you of the various people management possibilities—some of which have yet to be implemented in real life organizations—afforded by new algorithmic methods.

It is important that you read these articles through the **lens of a consumer** (as opposed to a producer) of data analysis. **It is not important that you understand the statistical or algorithmic mechanics in great detail**. You may skip mathematical definitions if you find them difficult. It is important, however, that you understand why a particular algorithm was applied. As a future user of algorithmic methods (as managerial tools) it is imperative that you understand what the algorithms do, even if you do not fully understand how they do it. You will need to evaluate their effectiveness, limitations, and biases.

In the session descriptions below, you will find a few session-specific guidelines and questions that will help you prepare for class discussion.

# CLASS DISCUSSION

Class discussion will be a mix of lecture, discussion, and exercise. The quality of the experience depends on student participation: the more students engage in productive discussion, the more rewarding the experience for all involved.

The course schedule below includes questions that will help you prepare for class discussion. Thinking about these questions in advance will provide you with the foundations to engage productively in class discussion.

While you are expected to be an active participant throughout the class, please note that the frequency (i.e., the quantity) of your contributions in class is not the only criterion for effective class participation. It is the *quality* of your participation that is most important. Criteria that are useful in measuring effective class participation include:

1. Is the participant a good listener?
2. Are the points that are made relevant to the discussion? Are they linked to the comments of others?
3. Do the comments show evidence of analysis of the issue at hand and an understanding of the assigned reading/s?
4. Do comments lead to a clearer statement of the concepts being covered and the problems being addressed?

Ultimately, a good comment is one that pushes the discussion forward. Please note:

* There are no “right” or “wrong” answers in this class. Rather, good arguments are ones that rest on sound and consistent assumptions.
* Because of the varied backgrounds in the class, many of you will have important contributions to make based on your personal and professional experience. You are encouraged to bring these experiences to bear on class discussion.

# 

# INDIVIDUAL PROJECT

The final individual project is due on **April 15th, 2022, 11:59 pm.** Please submit the project via Bcourses at “Final Project” (under “Assignments”).

In the final project you will be asked to assume the (hypothetical) role of a People Analytics manager in a large firm and provide a proposal for an algorithmic process in response to a specific organizational problem. You will have the option to focus either on the design side or the data analytics side. *More details about the assignment will be provided via bCourses.*

# BEHAVIORAL EXPECTATIONS

Every faculty member has somewhat different expectations as to class operations and norms for individual behavior. Here are the behavioral expectations for this class:

1. **Attendance is mandatory**. Excused absences will not affect your grade. Any unexcused absence will negatively impact your participation grade.
2. Please note that if you need to miss class you must notify me by e-mail **before** the session or within one hour after class ends.
3. As a courtesy to the instructor, your classmates, and guest speakers, please do your best to arrive on time. If there is a reason why you must leave the class early please let me know in advance.
4. Please let me know **before class** if an emergency has made it impossible for you to be prepared adequately for class discussion.
5. All slides will be made available via bCourses. All in-class sessions will be recorded and will be made available via bCourses shortly after.
6. Social media is an increasingly common part of our lives. **However, it is critically important that anything a guest says in the class stays in the class.** Should you feel compelled to blog, Tweet, Instagram, etc., it should only be to thank our guests for their generosity in helping with the education process.

# GRADING

Your course grade will be based on the final project and class participation. These will be weighted as follows:

Final project 25% (April 15th, 2022)

Midterm: 25% (February 24th, 2022)

Class participation/attendance: 50%

Class participation will be evaluated along the following dimensions: (1) Did the student attend all classes unless excused? (2) Did the student come to class prepared, as indicated by providing contributions to class discussion that reflect clear and thoughtful engagement with the readings and lecture material? (3) Did the student engage productively and collaboratively in group activities? (4) Did the student ask thoughtful and respectful questions of the guest speakers? For each class, the professor will make note of everyone who attended in full, and the professor will also keep notes on the quality of everyone’s participation. Please feel free to write the professor at any time if you would like further clarification on how participation will be assessed.

Please note that to pass the course, both class participation and final project requirements must be fulfilled. **In particular, inadequate class participation cannot be made up by outstanding performance on the final project**. Students who miss class due to unexcused reasons cannot make up for this absence.

# STUDENTS WITH DOCUMENTED DISABILITIES

Students who may need an academic accommodation based on the impact of a disability should initiate the request for assistance with Berkeley’s Disabled Student’s Program. Please also contact the instructor to inform him of this request and the nature of your needs so that appropriate assistance can be developed.

# COURSE SCHEDULE

**Class 1: January 20th, 2022**

**Topic: The Data Science Revolution**

In this introductory session, we will try to move beyond the buzz and figure out what types of algorithmic methods exist and what they do (as well as what they cannot and/or are not designed to do). Drawing on Agrawal et al, we will think about these methods through the prism of **prediction** and **judgment**. We will focus on the difference between **Precision** and **Bias**. We will discuss the statistical concepts that will be used throughout the course, and we will discuss the sources of bias (both in statistics and human judgment) that can lead algorithms astray, setting the stage for future classes on the promises and perils of algorithms in people management. Throughout the class, emphasis will be made on the key role played by human judgment in determining the validity and value of algorithms.

# Class Preparation Questions:

# What kinds of prediction problems typically exist in organizations?

# Where do you see the greatest benefit of algorithmic approaches to addressing these prediction problems?

# Can all prediction problems in organizations be solved by algorithms? Where is human judgement needed?

# Think of one type of organizational challenge that algorithmic methods are well-suited to address. Why are these methods appropriate? What are potential limitations of these methods?

# Required Readings:

Agrawal and Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. pp. 1-30.

Kahneman, Daniel. 2011. “Chapter 21 – Intuitions vs. Formulas.” *Thinking Fast & Slow*. Pg. 222-234.

# Recommended Materials:

Davenport and Ronanki. 2018. [3 Things AI Can Already Do for Your company](https://hbr.org/2018/01/artificial-intelligence-for-the-real-world). *Harvard Business Review*.

Dastin. 2018. [Amazon Scraps secret AI recruiting tool that showed bias against women](https://in.reuters.com/article/amazon-com-jobs-automation/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idINKCN1MK0AH). *Reuters Technology News*.

Luo, Qin, Fang, and Qu. 2020. Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions. *Journal of Marketing*, 1-19.

# Analytics Tutorial:

# [Custom Tutorial on Running Regressions for People Analytics in Python.](https://github.com/drguilbe/PeopleAnalytics2021/blob/main/Introduction%20to%20Regression%20Techniques%20in%20Analytics.ipynb)

**Class 2: January 27th, 2022**

**Topic: Intro to AI and Machine Learning, with Business Applications**

The growth of AI applications in organizations is exponential. This class will provide an accessible introduction to the core AI methods that are revolutionizing people analytics. We will review examples of particularly useful applications of AI to enhance firm performance, and we will also discuss limitations in the use of AI for people analytics applications, setting the stage for Class 7: Algorithmic Bias.

# Class Preparation Questions:

# What is automation? Which aspects of people management can and should be automated? Which aspects cannot and should not be automated?

# How important is it for a people manager to be able to know how an algorithm works in order to make good decisions about the use and interpretation of an algorithm?

# What happens when an algorithm is very powerful and predictive, and yet impossible for anyone to interpret? Should managers rely on these kinds of algorithms? Why and why not?

# What kinds of questions can people managers ask about data and algorithms to make sure they are used responsibly and in ways that enhance an organization?

# Required Readings:

Mitchell. 2019. “Chapter One: The Roots of Artificial Intelligence.” *Artificial Intelligence: A guide for thinking humans*. pg. 17–34. **(\*emphasis on pages 24 –32**)

Mitchell. 2019. “Chapter Two: Neural Networks and the Ascent of Machine Learning.” *Artificial Intelligence: A guide for thinking humans*. pg. 35–42.

Mitchell. 2019. “Chapter Four: Who, What, When, Where, Why.” *Artificial Intelligence: A guide for thinking humans*. pg. 68–80.

# Recommended Materials:

# A technical summary of the basics of neural networks: <http://neuralnetworksanddeeplearning.com/chap1.html>

# Stateoftheart.ai provides interactive visualizations of the linkages between AI approaches, both in terms of theories and concepts, and also publications. This platform also links to a number of potentially relevant datasets. See here to explore: <https://www.stateoftheart.ai/>

**Analytics Tutorial**:

*Note: Keep Python’s Jupyter Notebook’s open before clicking on these links so that the links below show you interactive code in your web browser.*

[Custom Intro Tutorial on Building and Applying a Random Forest Classifier in Python.](https://github.com/drguilbe/PeopleAnalytics2021/blob/main/Intro%20to%20Random%20Forest%20Classifiers.ipynb)

[Custom Intro Tutorial on Deep Learning in Python.](https://github.com/drguilbe/PeopleAnalytics2021/blob/main/Intro%20to%20Deep%20Learning.ipynb)

**Class 3: February 3rd, 2022**

**Topic: Applications of Algorithms to HR**

Can algorithms replace humans in making HR decisions (e.g., hiring, promotion, retention)? For example, the decision to hire is a very complex one. As Nobel Prize Laureate Michael Spence observed almost half a century ago, to hire someone is often like purchasing a lottery ticket. Can machines reduce this uncertainty? Human recruiters often take into consideration a variety of dimensions in deciding who to recruit. The reading from Youyou et al. (2015) will focus on one such dimension: personality. The reading from Guenole & Feinzig (2020) will focus on IBM’s experience and perspective on using algorithms to predict /inform a much broader array of dimensions in hiring practices.

# Class Preparation Questions:

1. What determines whether a hiring decision was good?
2. How can you evaluate a hiring decision? Which aspects of hiring can and should be automated? Which aspects can and should not be automated?
3. What kind of data would you need to train a hiring algorithm?
4. What is the role of human judgement in hiring?
5. Can you use the Youyou et al procedure to inform hiring decisions? If so, how?

# Required Materials:

Youyou, Kosinski and Stilwell. 2015. [Computer-based personality judgments are more accurate than those made by humans](https://www.pnas.org/content/112/4/1036). *PNAS* 112(4): 1036-1040.

Kahneman, Daniel. 2011. “Chapter 22 – Expert intuition: When can we trust it?” *Thinking Fast & Slow*. Pg. 234-245.

Guenole and Feinzig. 2020. [The Business Case for AI in HR: With Insights and Tips on Getting Started](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjk7YW1-IfuAhWQtZ4KHeMvCqQQFjAAegQIAhAC&url=https%3A%2F%2Fwww.ibm.com%2Fdownloads%2Fcas%2FAGKXJX6M&usg=AOvVaw3w9LpTWqJG6Pu-LmSfgdCN). *IBM Smarter Workforce Institute*.

**Recommended Materials:**

Schmidt and Hunter. 1998. The validity and utility of selection models in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin* 124(2): 262 –274.

Flanigan, Bailey, Paul Gölz, Anupam Gupta, Brett Hennig, and Ariel D. Procaccia. 2021. [Fair Algorithms for Selecting Citizens’ Assemblies](https://www.nature.com/articles/s41586-021-03788-6#Fig1). *Nature* 596(7873): 548–52.

**Analytics Tutorial**:

-[Custom Python Tutorial on How to Build a Simple HR Chat Bot](https://github.com/drguilbe/PeopleAnalytics2021/blob/main/How%20to%20Build%20a%20Chat%20Bot%20(Basic%20Introduction)%20.ipynb).

-The [Viralgorithm](https://www.facebook.com/theviralgorithm/): App for tracking social media success on Facebook

**Class 4:** **February 10th, 2022**

**Topic: Applications of Algorithms to Performance**

**Guest:** Monica Lee, Data Scientist at Facebook.

All organizations, ultimately, want to increase their employees’ performance, both individually and at the collective level. But measuring performance is not straightforward. As Deslauriers et al. demonstrate, performance measures are plagued by inaccuracies and biases, sometimes in very surprising ways. What does that mean for the applications of algorithms to enhance performance?

# Class Preparation Questions:

# What affects individual performance in organizations? How does individual performance differ from team performance?

# How do data scientists leverage machine learning to develop products and solve problems in the tech industry?

# Is performance easier to measure in certain roles or organizations? If so, what kind?

# Are algorithms more or less susceptible than humans to performance bias?

# Required Readings:

Deslauriers, McCarty, Millier, Callaghan, Ketti. 2019. [Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom](https://www.pnas.org/content/116/39/19251). *PNAS* 116(39): 19251-19257

Wilson and Daugherty. 2018. [Collaborative Intelligence: Humans and AI are Joining Forces](https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces)*. Harvard Business Review*.

Luo, Qin, Fang, and Qu. 2020. Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions. *Journal of Marketing*, 1-19. (\*recommended for Class 1; please review this material.)

# Recommended Readings:

Shirado, Hirokazu, and Nicholas A. Christakis. “Locally Noisy Autonomous Agents Improve Global Human Coordination in Network Experiments.” *Nature* 545, no. 7654 (May 17, 2017): 370–74. <https://doi.org/10.1038/nature22332>.

Traeger, et al. 2020. [Vulnerable robots positively shape human conversational dynamics in a human-robot team.](https://www.pnas.org/content/117/12/6370) PNAS 117(12): 6370-6375.

**Case Study**:

Open AI and Impact of COVID on Performance. This case study will be shared before class.

# Class 5: February 17th, 2022

**Topic: Natural Language Processing and Business Applications**

One of the most powerful ways in which machine learning has revolutionized people analytics is through the automated analysis of text as a window into personality and culture. In this class, I will provide a high-level overview of the dominant methods of natural language processing currently used in analytics (i.e., word embedding models), and we will discuss how these methods have been used to enhance firm performance and reveal key insights into individual personality and organizational culture.

# Class Preparation Questions:

1. What can language reveal about a person? What can language reveal about a social group, an organization, a culture?
2. How do organizations rely on language? For example, what role does language play in marketing, hiring, innovation, collaboration?
3. Which aspects of language would be most useful for an organization to predict?
4. Some say a picture is worth 1000 words. Others say that 1001 words are worth more than a picture. Where do you fall? Why? What kind of information about an organization is captured in words and not in images; and vice versa?

# Required Readings:

Garg, Schiebinger, Jurafsky and Zou. 2018. [Word embeddings quantify 100 years of gender and ethnic stereotypes](https://www.pnas.org/content/115/16/E3635). PNAS 115 (16) 3635-3644

Lix, Goldberg, Srivastava, Valentine. 2020. [Timing Differences: Discursive Diversity and Team Performance](https://faculty.haas.berkeley.edu/srivastava/papers/Timing_Differences-03_27_20.pdf). *Management Science* (Forthcoming).

**Recommended Materials:**

Eichstaedt, et al. 2018. [Facebook language predicts depression in medical records](https://www.pnas.org/content/115/44/11203). PNAS 115(44): 11203–11208.

Kozlowski, et al. 2019. [The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings](https://journals.sagepub.com/doi/full/10.1177/0003122419877135). *American Sociological Review* **84** (5), 905–49.

**Analytics Tutorial**:

[Custom Python tutorial](https://github.com/drguilbe/PeopleAnalytics2021/blob/main/Intro%20to%20Computational%20Linguistics%20for%20Cultural%20Analytics.ipynb) on using word embedding models for cultural analytics.

Online tutorial for visualizing and analyzing networks of associations among words:

<https://cbail.github.io/textasdata/text-networks/rmarkdown/Text_Networks.html>

# Class 6: February 24th, 2022

**Topic: Cultural Analytics for Organizations**

**Guest:** Mary Kate Stimmler, People Analytics at Google.

Recent years have seen a dramatic increase in the importance that firms put on managing and cultivating culture. Organizations increasingly recognize that culture is a source of competitive advantage, both in how it shapes employees’ behaviors and as a feature that draws, motivates, and retains employees. But culture is notoriously difficult to measure. Nevertheless, recent advances in computational linguistics provide new avenues for studying culture algorithmically.

# Class Preparation Questions:

1. Why is culture difficult to measure in organizations?
2. What kinds of data are amenable to algorithmic analyses of culture?
3. What are surveys’ weaknesses and strengths, relative to the data identified in question 2 above?
4. Can you think of ways to pair surveys with algorithmic methods that would enhance cultural analysis in organizations?

# Required Readings:

Srivastava, Goldberg, Manian and Potts. 2018. [Enculturation Trajectories: Language, Cultural Adaptation, and Individual Outcomes in Organizations](https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2016.2671?journalCode=mnsc). *Management Science* 64(3): 1348–1364

Srivastava, Goldberg. 2017. [Language as a Window into Culture](https://journals.sagepub.com/doi/full/10.1177/0008125617731781). California Management Review 60(1): 56–69.

**Recommended Materials:**

Guilbeault, Baronchelli, Centola. 2021. [Experimental Evidence of Scale-induced Category Convergence across Populations](https://www.nature.com/articles/s41467-020-20037-y). *Nature Communications* 12(327).

# Class 7: March 3rd, 2022

**Topic: Algorithmic Bias**

In this section of the course, we will focus in detail on the issues of bias and algorithmic flaws which have surfaced throughout the class thus far. We will discuss the imperfections of machine learning algorithms that are absolutely essential to understand in order to be effective users and consumers of machine learning analyses.

# Class Preparation Questions:

1. What is bias, in cognitive, social, and statistical terms?
2. How does an algorithm become biased?
3. How can we determine whether or not an algorithm is biased?
4. How, if possible, can algorithms be designed to reduce or avoid bias?
5. Can algorithms be used to reduce human bias, and vice versa?

**Required Readings:**

Mitchell. 2019. “Chapter Six: A Closer Look at Machines that Learn.” *Artificial Intelligence: A guide for thinking humans*. pg. 96–116.

Mullainathan. 2019. [Biased Algorithms Are Easier to Fix Than Biased People](https://www.nytimes.com/2019/12/06/business/algorithm-bias-fix.html). *New York Times*. December 6, 2019

**Recommended Materials:**

Snow. 2018. [Amazon’s Face Recognition Falsely Matched 28 Members of Congress with Mugshots.](https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28) ACLU.

Safra et al. 2020. [Tracking historical changes in trustworthiness using machine learning analyses of facial cues in paintings.](https://www.nature.com/articles/s41467-020-18566-7) *Nature Communications* 11(4728).

# Class 8: March 10th, 2022

**Topic:** Diversity and Collective Intelligence

Recent years have seen a growing awareness, by companies and the public alike, of issues related to racial, ethnic and gender diversity and equity in organizations. Meanwhile, research in computational social science has revealed a myriad of ways in which diversity can greatly improve the collective intelligence of teams and organizations. But even the most technologically advanced and politically motivated companies find it difficult to address diversity. Indeed, if it were easy to increase diversity, and if increases in diversity would invariably lead to better firm performance, most companies would have implemented such changes. As we will discuss in class, increasing diversity and implementing equitable policies is a challenging thing to do, though Page offers some compelling recommendations. As Page emphasizes, in order to implement such policies effectively, we need to be able to measure diversity effectively and also to measure the effects of increasing or decreasing it. This raises a number of important questions about the kind and quality of data that is collected for the purposes of Analytics.

# Class Preparation Questions:

1. How would you measure diversity? How does conceptual diversity differ from demographic diversity?
2. How would you measure equity?
3. Imagine a company implementing new policies aimed at increasing diversity. How would you evaluate their effectiveness using algorithmic methods? What data would you rely on?

# Required Readings:

Page. 2019. “Chapter 3, Diversity Bonuses: The Logic” from *The Diversity Bonus, How Great Teams Pay Off in the Knowledge Economy*. Princeton University Press.

Page. 2019. “Chapter 4, Identity Diversity” from *The Diversity Bonus, How Great Teams Pay Off in the Knowledge Economy*. Princeton University Press.

Page. 2019. “Chapter 6, Diversity Bonuses and the Business Case” from *The Diversity Bonus, How Great Teams Pay Off in the Knowledge Economy*. Princeton University Press.

**Recommended Materials:**

Woolley, et al. 2010. [Evidence for a Collective Intelligence Factor in the Performance of Human Groups](https://science.sciencemag.org/content/330/6004/686). *Science* 330(6604): 686-688.

Hostra, et al. 2020. [The Diversity-Innovation Paradox in Science](https://www.pnas.org/content/117/17/9284). *PNAS* 117: 9284–9291.

Guilbeault and Centola. 2018. [Social learning and partisan bias in the interpretation of climate trends](https://www.pnas.org/content/115/39/9714). PNAS 115(39): 9714-9719.

Page. 2019. “Chapter 7, Practice: D&T + D&I” from *The Diversity Bonus, How Great Teams Pay Off in the Knowledge Economy*. Princeton University Press. 220–233.

# Class 9: March 17th, 2022

# Topic: Social Network Analysis

**Guest:** Vlad Barash[[3]](#footnote-3), People Analytics at Graphika[[4]](#footnote-4) Inc.

People in organizations are typically interdependent with each other; in fact, if they weren’t, there would be little reason for the organization to exist. Managing this interdependence and creating positive spillovers between employees is one of the most challenging tasks of organizational leadership. One way to resolve this problem is formal organizational hierarchy. But a significant proportion of the information exchanged between people in organizations flows through informal social networks that span all levels of an organization. In this session, we will discuss how social networks can be measured and how they relate to firm performance. We will focus specifically on the effect of networks on spurring creativity, innovation, and the spread of behavior.

# Class Preparation Questions:

* 1. How are organizations shaped by social networks? What are some of the difficulties in measuring networks and the interdependence of people’s beliefs, values, and actions?
  2. How do social networks shape the spread of ideas and behaviors in organizations and markets?
  3. Firms vary significantly in the centrality of innovation to their competitive advantage: some focus heavily on innovation and others on execution. Should innovation-oriented and execution-oriented companies foster different types of networks?
  4. Can you think of ways to measure networks and interdependencies in organizations? Do they all measure the same thing?

# Required Readings:

Leonardi and Contractor. 2018. [Better People Analytics: Measure who they know, not just who they are](https://hbr.org/2018/11/better-people-analytics). Harvard Business Review.

Guilbeault, Becker, and Centola. 2018. [Complex Contagions: A Decade in Review](https://link.springer.com/chapter/10.1007/978-3-319-77332-2_1). In S. Lehmann & Y. Ahn (Eds.), *Spreading Dynamics in Social* *Systems*. Springer Nature. pp. 3-25.

**Recommended Materials:**

Watts. “The ‘New’ Science of Networks.” *Annual Review of Sociology* 30, no. 1 (2004): 243–70. <https://doi.org/10.1146/annurev.soc.30.020404.104342>.

Guilbeault & Centola. 2021. [Topological Measures for Identifying and Predicting the Spread of Complex Contagions](https://www.nature.com/articles/s41467-021-24704-6). *Nature Communications* 12(4430).

Centola, Becker, Brackbill, Baronchelli. 2018. [Experimental evidence for tipping points in social convention](https://science.sciencemag.org/content/360/6393/1116.abstract). *Science* 360(6393): 1116–1119.

**Analytics Tutorials**:

-<http://pablobarbera.com/big-data-upf/html/02a-networks-intro-visualization.html>

-<https://www.jessesadler.com/post/network-analysis-with-r/>

-[1000s of social networks to analyze in the programming language R](https://www.r-bloggers.com/2019/12/a-large-repository-of-networkdata/).

# Class 10: March 31st, 2022

**Topic: Beyond Correlation: Experimental Methods**

While the current and prospective uses of AI and machine learning are promising for firms, they nevertheless face a critical limitation – namely, they almost entirely rely on correlational results. In this class, we review some of the pitfalls of correlational analyses. We discuss the many ways in which correlations can be misleading, and we focus on the gold-standard way to escape these limitations: that is, experimentation. We review cutting edge techniques for experimentation in people analytics, and we discuss both the challenges in running successful experiments, as well as the many payoffs.

1. What are the limits of correlational data? Think of everything you’ve learned so far.
2. Why is correlation not the same as causation? What does correlation provide that causation lacks? What does causation provide that mere correlation lacks?
3. When and how can correlations fail managers? Can these failures be overcome if managers were to obtain causal knowledge?
4. What is the best way for managers to obtain causal knowledge (hint: experiments)? What is an experiment?
5. What is the kind of problem in people analytics that is likely to benefit from experiments? Which problems are unlikely to benefit? What are the limits in running experiments in organizations?

**Required Readings:**

Salganik. 2018. “Running Experiments” from *Bit by Bit: Social Research in the Digital Age*. Princeton University Press. pp. 147 – 178.

**Recommended Materials:**

Salganik. 2018. “Approximating Experiments”, from *Bit by Bit: Social Research in the Digital Age*. Princeton University Press. pp. 50 – 60 (Section 2.4.3).

Salganik. 2018. “Ten Common Characteristics of Big Data” from *Bit by Bit: Social Research in the Digital Age*. Princeton University Press. pp. 23 – 39 (Section 2.3.2 to 2.3.10).

Centola. 2018. Epilogue: Experimental Sociology. *How Behavior Spreads*. Princeton University Press. 179–189.

Meyer, et al. 2019. Objecting to experiments that compare two unobjectionable policies or treatments. PNAS 116(22): 10723 – 10728. <https://www.pnas.org/content/116/22/10723>

1. *Excerpts from these texts will be required reading in class, and all excerpts will be provided for you over bCourses. It will be helpful, though not required, to acquire the full texts.*  [↑](#footnote-ref-1)
2. *During COVID, some classes may be associated with a prerecorded lecture (1-1.5 hours) that will be uploaded a week before class. This will be watched ahead of the corresponding class by the students. When a prerecorded lecture component is provided, the corresponding in-class session will consist of discussion, in-class exercise, or guest speaker, and will run from 6:00pm (PT) to 8:00pm (PT).*  [↑](#footnote-ref-2)
3. <https://graphika.com/team/vladamir-barash/> [↑](#footnote-ref-3)
4. <https://graphika.com/case-studies> [↑](#footnote-ref-4)