Master of Science in Sustainability Management

SUMA PS5193 Core 3 Credits Area 2 or Elective	e Data Science Concepts for Sustainable Finance
Instructor:	Bruce M. Kahn, PhD. <u>bk2501@columbia.edu</u> 917-509-3376
Office Hours:	2929 Broadway, 5th Floor Wed 4:30-6 pm and by appointment.
Response Policy:	Preferred means of communication is via email and students can expect responses
	typically within 24 hours
Teaching Assistant	TBD
Office Hours:	2929 Broadway, 5th Floor Wed 4:30-6 pm and by appointment.
Response Policy:	Preferred means of communication is via email and students can expect responses
	typically within 24 hours

Course Overview

The course introduces practitioners of sustainability management to the data analysis techniques and statistical methods which are indispensable to their work. The class teaches how to build statistical substantiation and to critically evaluate it in the context of sustainability problems. The statistics topics and examples have been chosen for their special relevance to sustainability problems, including applications in environmental monitoring, impact assessment, and econometric analyses of sustainable development. Students are assumed to have had no previous exposure to statistics.

Learning Objectives

This course demonstrates how to conduct a quantitative analysis of an organization's work processes and operations, resource utilization, and environmental impact necessary to create a rationale for implementing sustainability initiatives. Statistical topics, including probability and random variables, will be discussed in both theory and in their practical applications for sustainability managers. This course will provide students with the skills to conduct regression analysis, to conduct hypothesis and estimation testing, to design surveys, and to prepare statistics packages. These quantitative skills are necessary for a professional manager responsible for the management of people, finances and operations toward sustainability goals.

Readings

Text Books

- Robert M. Leekley, Applied Statistics for Business and Economics, CRC Press, 2009,
- Applied statistics in business and economics. Doane and Seward
- Naked statistics : stripping the dread from the data. Charles Wheelan.
- Data science applied to sustainability analysis. Dunn and Balaprakash



https://www-sciencedirect-com.ezproxy.cul.columbia.edu/book/9780128179765/data-science-app lied-to-sustainability-analysis

• The nature of data : infrastructures, environments, politics. Goldstein and Nost https://ebookcentral.proquest.com/lib/columbia/detail.action?docID=7054866

Readings/Research Articles

The following additional readings will be listed in Courseworks. Any readings whose full-text is not available through the links in Courseworks will be placed on reserve at the Library.

Fuoli, M. (2012). Assessing social responsibility: A quantitative analysis of Appraisal in BP's and IKEA's social reports. Discourse & Communication, 6(1), 55-81. https://doi.org/10.1177/1750481311427788

Samer Abdallah samer.abdallah@elec.qmul.ac.uk & Mark Plumbley (2009) Information dynamics: patterns of expectation and surprise in the perception of music, Connection Science, 21:2-3, 89-117, DOI: 10.1080/09540090902733756

Liang, Y., Lee, S.A. Fear of Autonomous Robots and Artificial Intelligence: Evidence from National Representative Data with Probability Sampling. Int J of Soc Robotics 9, 379–384 (2017). https://doi.org/10.1007/s12369-017-0401-3

Khosravi, Faramarz, Gokhan Izbirak, and Seyed Mahdi Shavarani. "Application of bootstrap re-sampling method in statistical measurement of sustainability." Socio-Economic Planning Sciences 75 (2021): 100781.

Nate Silver: A Climate of Healthy Skepticism. From The Signal and the Noise: Why So Many Predictions Fail — but Some Don't, Nate Silver

Carvalho, Marly M., and Roque Rabechini Jr. "Can project sustainability management impact project success? An empirical study applying a contingent approach." International Journal of Project Management 35.6 (2017): 1120-1132.

Zorio, Ana, Maria A. García-Benau, and Laura Sierra. "Sustainability development and the quality of assurance reports: Empirical evidence." Business strategy and the environment 22.7 (2013): 484-500.

Crook, T. Russell, et al. "Does human capital matter? A meta-analysis of the relationship between human capital and firm performance." Journal of applied psychology 96.3 (2011): 443.

Olsson, Rickard, (2007), Portfolio performance and environmental risk, No 2007/4, Sustainable Investment and Corporate Governance Working Papers, Sustainable Investment Research Platform, https://EconPapers.repec.org/RePEc:hhb:sicgwp:2007_004.

Semenova, Natalia, and Lars G. Hassel. "Financial outcomes of environmental risk and opportunity for US companies." Sustainable Development 16.3 (2008): 195-212.

Brotcke, L. Time to Assess Bias in Machine Learning Models for Credit Decisions. J. Risk Financial Manag. 2022, 15, 165. <u>https://doi.org/10.3390/jrfm15040165</u>

Nieto-Rodriguez and Vargas. 2023. The Opportunities at the Intersection of AI, Sustainability, and Project Management. Harvard Business Review:2023

Lanza, A, Bernardini, E. and Faiella, E. Mind the gap! Machine learning, ESG metrics and sustainable investment. Banca Italia Occasional Papers. June 2020.

Additional Books of Interest on Reserve:

- The Black Swan: The Impact of the Highly Improbable Nassim Nicholas Taleb
- Moneyball: The Art of Winning an Unfair Game, Michael Lewis
- Freakonomics: A Rogue Economist Explores the Hidden Side of Everything, Steven Levitt and Stephen J. Dubner.
- How to Lie with Statistics, Darrell Huff
- The Signal and the Noise: Why So Many Predictions Fail but Some Don't, Nate Silver.

Assignments and Assessments

The course is based on 200 points.

Homework Assignments: There will be 13 weekly problem-solving assignments each worth 10 points for a total of 130 points.

Participation: Each student will make a presentation on the weeks reading

"sustainability" reading assignments plus their active participation in class worth 10 points

Tests: There will be a take-home midterm exam and a take-home final, each worth 30 points.

The Midterm is due on TBD The Final Exam is due on TBD

I will hold office hours on Wed's from 4:30-5:45 pm at 2929 Broadway, by appointment. Recitation Periods TBD

Grading

The final grade will be calculated as described below:

COLUMBIA UNIVERSITY School of Professional Studies

FINAL GRADING SCALE

Grade	Percentage
A+	98–100 %
А	93–97.9 %
А-	90–92.9 %
B +	87–89.9 %
В	83–86.9 %
B-	80-82.9 %
C+	77–79.9 %
C C-	73–76.9 %
C-	70–72.9 %
D	60–69.9 %
F	59.9% and below

Assignment/Assessment	% Weight	Individual or Group/Team Grade	
Assignment 1-13 10 pts each	65	Individual	
Midterm	15	Individual	
Final Project	15	Individual	
Participation and Individual Presentation	5	Individual	

Course Schedule/Course Calendar

Session 1 Introduction

General Research Methodology: Inductive Method, Hypothetico-Deductive Method, Experimental and Non-Experimental Design, Causal Inference; The Uses of Statistics in Sustainability Studies: Impact Assessment, Monitoring, Auditing, Polling; Using Statistics in Research: Sample vs. Population, Description vs. Inference, Sampling Error and Bias

Readings: Leekley, Chapter 1 and Chapter 2.

Session 2 Describing Data: Tables and Graphs

Measures of Central Tendency: Mean, Median, Mode, Advantages and Disadvantages; Measures of Dispersion: Mean Absolute Deviation, Variance and Standard Deviation, Quantiles and Inter-Quartile Range; Skewness and Kurtosis; Plots: Histogram, Q-Q, ECDF, Box, Scatterplot, Smoothers



Readings: Leekley, Chapter 2

Session 3 Describing Data: Summary Statistics

Measures of Central Tendency: Mean, Median, Mode, Advantages and Disadvantages; Measures of Dispersion: Mean Absolute Deviation, Variance and Standard Deviation, Quantiles and Inter-Quartile Range; Skewness and Kurtosis; Plots: Histogram, Q-Q, ECDF, Box, Scatterplot, Smoothers

Readings: Leekley, Chapters 3,

Session 4 Basic Probability

The Origins of Probability Theory; Events; The Laws of Probability; Probability of A or B; Conditional Probability; Joint Probability; Bayes' Rule; Permutations and Combinations, Discrete vs. Continuous, Category vs. Ordered vs. Quantitative; Expected Value, Variance; Discrete Random Variables: Binomial, Poisson, Hypergeometrics; Continuous Random Variables: Normal, Chi-Squared, Exponential

Readings: Leekley, Chapter 4

Session 5 Probability Distributions

The Origins of Probability Theory; Events; The Laws of Probability; Probability of A or B; Conditional Probability; Joint Probability; Bayes' Rule; Permutations and Combinations, Discrete vs. Continuous, Category vs. Ordered vs. Quantitative; Expected Value, Variance; Discrete Random Variables: Binomial, Poisson, Hypergeometrics; Continuous Random Variables: Normal, Chi-Squared, Exponential

Readings: Leekley, Chapter 5.

Session 6 Sampling and Sampling Distributions

Random sampling, stratified sampling, cluster sampling, the t-table, Environmental Sampling, Surveys and experiments; Experimental design; Constructing Samples; Constructing indices and scales; Examples of bad survey questions; Replication in natural vs social sciences.

Readings: Leekly Chapter 6.

Session 7 Estimation and Confidence Intervals

Point and interval estimators, estimate of proportion, populations mean

Readings: Leekley, Chapter 7,

Midterm Due: Distributed on March 6th and due on March 20th.

COLUMBIA UNIVERSITY School of Professional Studies

Review Eccles paper and develop research proposal

Session 8 Hypothesis Testing

Independence of Observations Central Limit Theorem Sampling Distributions Tests for distribution (Kolmogorov-Smirnov, Q-Q tests). The one-sample t-test for a population mean; One-sample Chi-squared test for population variance; Two-sample t and z tests for population mean; two-sample z test for population variance

Readings: Leekley, Chapter 8

Session 9 Hypothesis Testing

Independence of Observations Central Limit Theorem Sampling Distributions Tests for distribution (Kolmogorov-Smirnov, Q-Q tests). The one-sample t-test for a population mean; One-sample Chi-squared test for population variance; Two-sample t and z tests for population mean; two-sample z test for population variance

Readings: Leekley, Chapter 9

Session 10 Hypothesis Testing

Independence of Observations Central Limit Theorem Sampling Distributions Tests for distribution (Kolmogorov-Smirnov, Q-Q tests). The one-sample t-test for a population mean; One-sample Chi-squared test for population variance; Two-sample t and z tests for population mean; two-sample z test for population variance

Readings: Leekley, Chapter 10

Session 11 Hypothesis Testing

Independence of Observations Central Limit Theorem Sampling Distributions Tests for distribution (Kolmogorov-Smirnov, Q-Q tests). The one-sample t-test for a population mean; One-sample Chi-squared test for population variance; Two-sample t and z tests for population mean; two-sample z test for population variance

Readings: Leekley, Chapters 11

Session 12 Regression Analysis

Covariance and Correlation, Spearman Rank Correlation, Correlation Tests; Scatterplot and Univariate Regression. Regression Error, Coefficient of Determination; Assumptions of the Linear Regression Model; Multivariate regression, Hypothesis Tests about Coefficients and the Model; Specification; Missing Data; Heteroschedasticity; Discrete Dependent Variables

Readings: Leekley, Chapters 12

Session 13: Climate Change and Statistics

Readings: Chapter 12 of The Signal and the noise, Nate Silver.

Session 14 Multiple Regression Analysis

Covariance and Correlation, Spearman Rank Correlation, Correlation Tests; Scatterplot and Univariate Regression. Regression Error, Coefficient of Determination; Assumptions of the Linear Regression Model; Multivariate regression, Hypothesis Tests about Coefficients and the Model; Specification; Missing Data; Heteroschedasticity; Discrete Dependent Variables

Readings: Leekley Chapter 13

Session 15: Time Series Analysis

Exploiting patterns over time, basic components of a time series, seasonal variation, the long-term trends, the business cycles, forecasting.

Readings: Leekley Chapter 14

Session 16: Final Exam Presentations

The Final Exam is due on May 8th.

All readings and assignments should be completed prior to the class in which it is going to be discussed. For example, Chapter 1 should be read before the second class session on January 24th.

COLUMBIA UNIVERSITY School of Professional Studies

Session	Date	Description	Textbook Readings (Leekley)	Assignment(s) Due	Industry Readings	Things to Focus on in the Industry Readings	Data Science Concepts Readings	Nature of Data Infrastru ctures
1		Overview: Putting the						
2		Introduction to Data: Types of Data and How to Describe It (Part 1)	Chapter 1: Introduction to Statistics Chapter 2: Describing Data (Tables and Graphs) [Section 2.1]		Naked Statistics: What's the Point? [Chapter 1] and Review Mind the Gap. Banca Italia, 2020		Chapter 1	Chapter 1
3		Introduction to Data: Types of Data and How to Describe It (Part 2) Summary Statistics	Chapter 2: Describing Data (Tables and Graphs) [Sections 2.2-2.3] Chapter 3: Describing Data (Summary Statistics)	2.2, 2.4, and 2.6	Neked Statistics: Descriptive Statistics: Who was the best basepall player of all time? [Chapter 2] Harvard Rusiness Review: The Opportunities at the Intersection of AI, Sustainability, and Project Management		Chapter 2	Chapter 2
4		Fundamentals of Data Science: Probability (Part 1)	Chapter 4: Basic Probability		55-81. https://doi.org/10.1177/1750481311427788	Types of data, summary statistics, and frequency	Chapter 3	Chapter 3
5		Fundamentals of Data Science: Probability (Part 2)	Chapter 5: Probability Distributions		Information dynamics: patterns of expectation and surprise in the perception of <u>music</u> , Connection Science, 21:2-3, 89-117, DOI:	Visualizations and ways of portraying data	Chapter 4	Chapter 4
6		Sampling (Part 1)	Chapter 6: Sampling and Sampling Distributions		Evidence from National Representative Data with Probability Sampling. Int J of Soc Robotics 9, 379–384 (2017). https://doi.org/10.1007/s12369-017-0401-3		Chapter 5	Chapter 5
7		Fundamentals of Data Analysis: Statistical Sampling (Part 2)	Chapter 7: Estimation and Confidence Intervals		of bootstrap re- <u>sampling</u> method in statistical measurement of sustainability." Socio-Economic Planning Sciences 75 (2021): 100781.	Differences in data, methods, and data manipulation	Chapter 6	Chapter 6
8		Fundamentals of Data Analysis: Hypothesis		7.2, 7.4, 7.6, 7.8, and 7.10	Nate Silver: A Climate of Healthy Skepticism		Chapter 7	Chapter 7
		SPRING BREAK						
9			Chapter 8: Tests of Hypotheses (One				Chapter 8	Chapter 8
10	27-Mar	Data Analyses: Two- Sample Tests	Chapter 9: Tests of Hypotheses (Two Sample Tests)	8.2, 8.4, 8.6, 8.10, and 8.14	Carvalho, Marly M., and Poque Rabechni Jr. "Can <u>project sustainability</u> <u>management</u> impact project success? An empirical study applying a contingent approach." International Journal of Project Management 35.6	Hypothesis testing (T-statistics) and statistical	Chapter 9	Chapter 9
11		of-Fit	Chapter 10: Tests of Hypotheses (Contingency and Goodness-of-Fit)	9.2, 9.4, 9.8, 9.10, and 9.12	Zorio, Ana, Maria A. García Benau, and Laura Sierra. "Sustainability. development and the quality of assurance reports: Empirical evidence." Business strategy and the environment 22.7 (2013): 494-500.	Hypothesis testing (Chi-square tests) and statistical		10
12	10-Apr	Data Analyses: ANOVA and Tests of Variances	Chapter 11: Tests of Hypotheses (ANOVA and Tests of Variances)		Crook, T. Russell, et al. "Does human capital matter? A meta-analysis of the relationship between human capital and firm performance ." Journal of	Hypothesis testing (Chi-square tests),		Chapter 11
13			Chapter 12: Simple Regression and Correlation		Olsson, Rickard, (2007), <u>Portfolio performance</u> and <u>environmental risk</u>, No 2007/4, Sustainable Investment and Corporate Governance Working Papers, Sustainable Investment Research Platform,	-	Chapter 12	12
14		Data Analyses: Multiple Regression	Chapter 13: Multiple Regression		environmental risk and opportunity for US companies." Sustainable	Regression analysis	Chapter 13	Chapter 13
15	, in the second s	Time Series	Ch. 14 Time Series			Interpreting and interrogating data		Chapter 14
16	8-May	Student Final Presentations and Wrap-Up: Asking the Right Questions and Finding		Final Presentation and Paper, Plus Chapter 14. Q 14.4,				Chapter 15

Course Policies

Participation and Attendance

Students are expected to arrive on time, attend all classes, and to stay until the end of class unless they have notified the instructor at the beginning of the session that they will be leaving early. Each unexcused absence will result in a 1 point deduction from the participation grade.

Late Work

Assignments are due on the dates/times identified. One letter grade will be deducted from any assignment submitted after the due date/time. No assignment will be accepted after the deadline for submitting final grades.

Incompletes

As outlined in the School's grading and academic starts policy, "A grade of 'I' (incomplete) is a temporary grade indicating failure to complete assigned work. The mark is given only upon the request of the student and at the discretion of the instructor. The student and faculty member must sign a completed 'Request for Grade of Incomplete Form' before the final class session. The 'I' must be removed within one year after the end of the semester in which the student received the grade. Students seeking an extension of this time limit must have the approval of the instruction and successfully petition of the director of their program. If no petition is made, of if the petition is unsuccessful, the grade is chanced to an N-Permanent Incomplete- which remains on the student's permanent record.

Citation & Submission

All written assignments must use standard citation format (e.g., MLA, APA, Chicago), cite sources, and be submitted to the course website (not via email).



School and University Policies and Resources

Copyright Policy

Please note—Due to copyright restrictions, online access to this material is limited to instructors and students currently registered for this course. Please be advised that by clicking the link to the electronic materials in this course, you have read and accept the following:

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted materials. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use," that user may be liable for copyright infringement.

Academic Integrity

Columbia University expects its students to act with honesty and propriety at all times and to respect the rights of others. It is fundamental University policy that academic dishonesty in any guise or personal conduct of any sort that disrupts the life of the University or denigrates or endangers members of the University community is unacceptable and will be dealt with severely. It is essential to the academic integrity and vitality of this community that individuals do their own work and properly acknowledge the circumstances, ideas, sources, and assistance upon which that work is based. Academic honesty in class assignments and exams is expected of all students at all times.

SPS holds each member of its community responsible for understanding and abiding by the SPS Academic Integrity and Community Standards posted at

<u>https://sps.columbia.edu/students/student-support/academic-integrity-community-standards</u>. You are required to read these standards within the first few days of class. Ignorance of the School's policy concerning academic dishonesty shall not be a defense in any disciplinary proceedings.

Diversity Statement

It is our intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength and benefit. It is our intent to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture.

Accessibility

Columbia is committed to providing equal access to qualified students with documented disabilities. A student's disability status and reasonable accommodations are individually determined based upon disability documentation and related information gathered through the intake process. For more information regarding this service, please visit the University's Health Services website: <u>https://health.columbia.edu/content/disability-services</u>.

School Policies and Expectations:

Page 9

Accessibility Statement – I want you to succeed in this course. Contact disability@columbia.edu for learning accommodations.

Names/Pronouns

You deserve to be addressed in a manner that reflects your identity. You are welcome to tell me your pronoun(s)and/or name (if different from University records) at any time, either in person or via email.

Discrimination

We embrace the diversity of gender, gender identity & expression, sex, sexual orientation, race, ethnicity, national origin, age, religion, disability status, family status, socioeconomic background, and other visible and non-visible identities. Columbia University does not tolerate unlawful discrimination, discriminatory harassment, sexual assault, domestic violence, dating violence, stalking, or sexual exploitation and all such conduct is forbidden by Columbia University Policy

Duty to Report

You deserve a University community free from discrimination, harassment, and gender-based misconduct including sexual harassment, sexual assault, domestic and dating violence, stalking, and sexual exploitation. It is therefore University policy to require Columbia faculty and staff to report to EOAA any instance or allegation of prohibited conduct involving any undergraduate or any graduate student that is disclosed to, observed by, or otherwise known to that employee. This requirement to report is in place to help ensure that students are provided appropriate resources and to allow the University to mitigate harm to our community

Confidential Resources

There are confidential resources on campus who do not have a Duty to Report, including: * Sexual Violence Response & Rape Crisis/Anti-Violence Support Center (SVR) * Ombuds Office * Medical Services * University Counseling and Psychological Services * University Pastoral Counseling * Columbia Office of Disability Services University employees working in a confidential capacity will not report information shared with them.

Inclusion

In the M.S. in Sustainability Management program, faculty and staff are committed to the creation and maintenance of "inclusive learning" spaces – classrooms and other places of learning where you will be treated with respect and dignity, and where all individuals are provided equitable opportunity to participate, contribute, and succeed. All students are welcome regardless of race/ethnicity, gender identities, gender expressions, sexual orientation, socio-economic status, age, disabilities, religion, regional background, Veteran status, citizenship status, nationality and other diverse identities that we each bring to class.

Class Recordings

All or portions of the class may be recorded at the discretion of the Instructor to support your learning. At any point, the Instructor has the right to discontinue the recording if it is deemed to be obstructive to the learning process.



If the recording is posted, it is confidential and it is prohibited to share the recording outside of the class.

SPS Academic Resources

The Division of Student Affairs provides students with academic counseling and support services such as online tutoring and career coaching: <u>https://sps.columbia.edu/students/student-support/student-support-resources</u>.

Columbia University Information Technology

<u>Columbia University Information Technology</u> (CUIT) provides Columbia University students, faculty and staff with central computing and communications services. Students, faculty and staff may access <u>University-provided and</u> <u>discounted software downloads</u>.

Columbia University Library

<u>Columbia's extensive library system</u> ranks in the top five academic libraries in the nation, with many of its services and resources available online.

The Writing Center

The Writing Center provides writing support to undergraduate and graduate students through one-on-one consultations and workshops. They provide support at every stage of your writing, from brainstorming to final drafts. If you would like writing support, please visit the following site to learn about services offered and steps for scheduling an appointment. This resource is open to Columbia graduate students at no additional charge. Visit http://www.college.columbia.edu/core/uwp/writing-center.

Career Design Lab

The Career Design Lab supports current students and alumni with individualized career coaching including career assessment, resume & cover letter writing, agile internship job search strategy, personal branding, interview skills, career transitions, salary negotiations, and much more. Wherever you are in your career journey, the Career Design Lab team is here to support you. Link to <u>https://careerdesignlab.sps.columbia.edu/</u>

Netiquette

Online sessions in this course may be offered through Zoom, accessible through Canvas. A reliable Internet connection and functioning webcam and microphone are required. It is your responsibility to resolve any known technical issues prior to class. Your webcam should remain turned on for the duration of each class, and you should expect to be present the entire time. Avoid distractions and maintain professional etiquette.

Please note: Instructors may use Canvas or Zoom analytics in evaluating your online participation.

More guidance can be found at: https://jolt.merlot.org/vol6no1/mintu-wimsatt_0310.htm

Netiquette is a way of defining professionalism for collaborations and communication that take place in online environments. Here are some Student Guidelines for this class:

• Avoid using offensive language or language that is not appropriate for a professional setting.



- Do not criticize or mock someone's abilities or skills.
- Communicate in a way that is clear, accurate and easy for others to understand.
- Balance collegiality with academic honesty.
- Keep an open-mind and be willing to express your opinion.
- Reflect on your statements and how they might impact others.
- Do not hesitate to ask for feedback.
- When in doubt, always check with your instructor for clarification.