



THE ROBERT DAY SCHOOL
CLAREMONT MCKENNA COLLEGE

**Economics 160– Accounting Data Analytics
Spring 2023 Syllabus**

Prof. George Batta
Office: Bauer 304
Website: www.georgebatta.com

Email: gbatta@cmc.edu
Office Phone: (909)607-7260

Zoom info for hybrid (but please see class attendance policy below):
Meeting ID: 519 755 6999
Link: <https://cmc-its.zoom.us/j/5197556999>

Office Hours: TTh 12:15pm-1:15pm. By appointment over Zoom or in person.

Course	Class Time	Classroom	Final Exam
Econ 160 (Sec. 2)	TTh 11:00a.m.-12:15p.m.	BC24	None
Econ 160 (Sec. 1)	TTh 1:15p.m.-2:30p.m.	BC22	None

Quantitative and Computing Lab (QCL) Course Mentor

Matthew San Luis | MSanLuis23@students.claremontmckenna.edu

“Walk-in” link: Please look for Zoom link in an email from the QCL later this semester.

One-on-one scheduling link: <https://cmc.mywconline.net/>

Required Material (all freely available)

- Wickham, H., and Golemund, G. *R for Data Science*, 2nd edition. You can find a free online version here: <https://r4ds.hadley.nz/>.
- James, G., Witten, D., Hastie, T., Tibshirani, R. *An Introduction to Statistical Learning, with Applications in R*. New York, NY: Springer Science + Business Media, 2nd edition
 - Available at the Huntley Bookstore or your favorite online retailer. You can also find a free PDF here: <https://statlearning.com/>
 - We’ll only cover chapters 2, 4, and 5 in this text.
- Base R Cheat Sheet: <https://github.com/rstudio/cheatsheets/blob/main/base-r.pdf>
- Data Transformation Cheat Sheet: <https://github.com/rstudio/cheatsheets/blob/main/data-transformation.pdf>
- EY Case studies
- Additional required reading, consisting of one academic paper as well as an excerpt from Provost and Fawcett’s *Data Science for Business* text, will be uploaded to Sakai later in the semester.

Catalog Description

This course will introduce students to the use of analytics tools for deriving insights from accounting data and for more effectively performing audits. Companies produce a wealth of data on customers and company performance, and the next generation of accountants needs to be equipped with tools for organizing and analyzing the data to improve company performance and audit its financial accounts. We will explore the topics of data retrieval, cleanup, preprocessing, and validation, before getting into data visualization, internal and external audit analytics, and predictive modeling/machine learning.

Prerequisites

Economics 86 - Accounting for Decision-Making
Economics 120 – Statistics (or equivalent)

Learning Objectives

- Understand and apply the analytics mindset to real-world accounting problems.
- Explore the use of analytics techniques in several areas of accounting, including ratio analysis, cost accounting, internal audit, and financial accounting.
- Become familiar with R, RStudio, and Quarto files, for use in data science.
- Become acquainted with data cleaning, transformation, querying, exploration, visualization, programming basics, databases, and data communication.
- Understand the difference between causal inference and prediction models, and the tools required for both.
- Understand basic principles of statistical learning.

Grading Criteria

Homeworks (5, each worth 4%)	20%
Midterm #1	20%
Midterm #2	30%
Midterm #3	25%
Attendance at speaker presentation	5%

Case studies

The main lessons of the course will be conveyed via case studies encompassing four areas of accounting: financial ratio analysis, cost accounting, internal audit, and financial accounting. With the exception of the first case, cases will extend over four class sessions and will follow this (typical) pattern:

Day 1 – Case introduction, small group question brainstorming, loading dataset into our analysis tool, and initial data analysis.

Day 2 – Lab session, with in-class questions to answer.

Day 3 – Lab session, with in-class questions to answer.

Day 4 – Homework due/ Lab session, with in-class questions to answer, discussion of homework problems, and (twice during the semester) group presentations.

Attendance

Class attendance is important. Many of the sessions will involve working through R-based problems, often collaboratively with other students. While I will offer hybrid sessions for students who are ill, students are expected to regularly attend in-person. Though I won't at all times be monitoring class attendance, repeated in-person absences will eventually result in an email to your academic advisor.

All this in mind, I am inviting a professional speaker this semester. Attendance at this session is **mandatory**. Students will receive full credit for attendance, *unless* they are obviously working on some other personal project (texting, tweeting, etc.) or academic project during the speaker session.

Lecture Recordings

All lecture recordings for the **earlier section** (section 2), will be stored at the link below. This will be posted in the late afternoon or early evening of the lecture date. Again, these are meant to supplement regular, in-person attendance:

<https://claremontmckenna.box.com/s/vp12xib872450zxkxbx401xvvgov89ut8>

Homework

Homeworks are due on the dates listed below. You will typically upload your homework to your Sakai dropbox. You may work in groups of **up to 3 students total** on the assignments. Your homeworks will typically be submitted as Quarto (.qmd) files.

Exams

Exams will be held on the dates listed in the class schedule.

Academic Integrity

I encourage you to read CMC's Statement of Academic Integrity in the online catalogue: <http://catalog.claremontmckenna.edu/content.php?catoid=4&navoid=94>

On the use of AI-driven tools like ChatGPT or GitHub Copilot:

There's a good chance these tools (or derivatives thereof) will be a regular part of your professional workflow. I also believe that any attempt to ban their use outside of the classroom is impossible to enforce. However, to evaluate the quality of these tools' output, you need to first understand the fundamentals of accounting data analytics.

As such, use of these tools for homework problems is highly discouraged, for the simple reason that I forbid use of them during in-class exams. The homeworks are designed to develop your understanding of course materials, and that understanding will be evaluated during exams, which form the bulk of your grade in this class.

Grading Disputes

If you believe that there is an error in the scoring of a test, you must submit a **written** request for re-grading within one week from the time that the test is returned. The request must include a copy of the disputed portion of the test, accompanied by a written explanation of why you believe re-grading is appropriate. You should suggest how many points should be awarded.

Statement of Reasonable Accommodations

Your experience in this class is important to me. If you have already established accommodations with Disability Services at CMC, please communicate your approved accommodations to me at your earliest convenience so we can discuss your needs in this course. You can start this conversation by forwarding me your accommodation letter.

If you have not yet established accommodations through Disability Services, but have a temporary health condition or permanent disability (conditions include but are not limited to: mental health, attention-related, learning, vision, hearing, physical or health), you are encouraged to contact Assistant Dean for Disability Services & Academic Success, Kari Rood, at accessibilityservices@cmc.edu to ask questions and/or begin the process. General information and the Request for Accommodations form can be found at the CMC DOS Disability Service's [website](#). Please note that arrangements must be made with advance notice in order to access the reasonable accommodations. You are able to request accommodations from CMC Disability Services at any point in the semester. Be mindful that this process may take some time to complete and accommodations are not retroactive. It is important to Claremont McKenna College to create inclusive and accessible learning environments consistent with federal and state law.

If you are not a CMC student, please connect with the Disability Services Coordinator on your campus regarding a similar process.

Class	Date	Data Analysis Topics	Accounting Topic	Case and/or Dataset	Readings WG: Wickham and Grolemond JWHT: James, Witten, Hastie, Tibshirani	Homework due?
1	17-Jan	Course overview				
2	19-Jan	Data transformation and visualization	Financial Ratio Analysis	Dupont	Read Dupont case study WG: 1 & 3	
3	24-Jan	Data transformation and visualization	Financial Ratio Analysis	Dupont	WG: 4.1-4.2	
4	26-Jan	Data transformation and visualization	Financial Ratio Analysis	Dupont	WG: 4.3-4.4	
5	31-Jan	Data transformation and visualization	Financial Ratio Analysis	Dupont	WG: 2.1-2.4	
6	2-Feb	Data transformation and visualization	Financial Ratio Analysis	Dupont	WG: 2.5-2.6	Yes
7	7-Feb	Relational data and exploratory data analysis	Cost Accounting	IntegrateCo	Read IntegrateCo case study WG: 21.1-21.5	
8	9-Feb	Relational data and exploratory data analysis	Cost Accounting	IntegrateCo	WG: 6.1-6.3	
9	14-Feb	Relational data and exploratory data analysis	Cost Accounting	IntegrateCo	WG: 12.1-12.7	
10	16-Feb	Relational data and exploratory data analysis	Cost Accounting	IntegrateCo		Yes
11	21-Feb	Midterm #1 (in-class)				
12	23-Feb	Data cleaning and querying	Internal Audit	P-card	Read P-card case study WG: 16.1-16.5	
13	28-Feb	Data cleaning and querying	Internal Audit	P-card	WG: 19.1-19.3 & 27.1-27.2	

14	2-Mar	Data cleaning and querying	Internal Audit	P-card	WG: 18	
15	7-Mar	Data cleaning and querying	Internal Audit	P-card		Yes
16	9-Mar	Databases			WG: 22.1-22.5.6	
<i>Spring Break</i>						
17	21-Mar	Iterators and data communication	Financial Accounting	Bank investment portfolios	Read Bank Investment Portfolios case study & prepare questions WG: 27.1-27.2	
18	23-Mar	Iterators and data communication	Financial Accounting	Bank investment portfolios	WG: 12.1-12.2	
19	28-Mar	Iterators and data communication	Financial Accounting	Bank investment portfolios	WG: 12.3-12.6	
20	30-Mar	Iterators and data communication	Financial Accounting	Bank investment portfolios		Yes
21	4-Apr	Intro to statistical learning			Athey (2018) on Sakai JWHT: Ch. 2, pp.15-29	
22	6-Apr	Midterm #2 (in evening, 2 hours, class doesn't meet this date)				
23	11-Apr	Intro to statistical learning			JWHT: Ch. 2, pp. 29-42	

24	13-Apr	Intro to statistical learning & logistic regression			Data Science for Business excerpt on Sakai, pp. 187-192 JWHT, Ch. 4: pp129-135	
25	18-Apr	Logistic regression			JWHT, Ch. 4: pp. 136-139	
26	20-Apr	KNN v. Logit; Decision analytic thinking; and intro to resampling methods			Data Science for Business excerpt on Sakai, pp. 193-196 JWHT, Ch. 5: pp. 197-200 (up to 5.1.1)	
27	25-Apr	Resampling methods and midterm review			JWHT, Ch. 5: pp. 200-208	Yes
28	27-Apr	Midterm #3 (in-class)				
29	2-May	Speaker presentation (alternatively, we'll have an evening Ath session late in semester, in which case, we won't hold class this final week)				