

## Seminar in Data Analytics Schedule

*\*Schedule is subject to instructor modification. Updates will be shared in the online copy of the syllabus.*

<b>Dates</b>	<b>Topics</b>	<b>Assignments</b>
Week 1 Jan 21	Introductions, research questions, and early question development	Initial self interest and skill assessment (s) Identify some potential sources, and datasets
Week 2 Jan 28	Writing professionally I, breaking down a complex project, brainstorming project proposals.	<b>Classes canceled for MLK Celebration and Events</b> Problem statement and project proposals due. Upload to discussion group (TBA) by Wednesday evening. Reviews posted by Friday evening. (s)
Week 3 Feb 4	Project proposals, abstracts, and lit reviews, how to peer edit	Revised Project proposals (5%) and Annotated bibliographies (8-12 references) (s)
Week 4 Feb 11	Building domain knowledge	Domain Review due (5%)
Week 5 Feb 18	The intersection of design and data; research design, data availability, and data gathering	Skill review due (5%)
Week 6 Feb 25	Ethical principles of data handling, Data management + manipulation	Ethical statement and protocols (5%) IRB certification, application, or approval (s)
Week 7 Mar 4	Analysis choice - defining characteristics of data and question of interest	Research Design (5%)
Week 8 Mar 11	Early visualization and storytelling <i>March 15, Physics seminar on IBM and border security data analysis</i>	Early Results (5%) Data visualization workshopping (s) Initial data, code, and visuals pushed to GitHub (s)
Spring Break Mar 18	<i>Optional: Women in Analytics conference, March 21-22</i>	
Week 9 Mar 25	Refining the project, making final choices	Keep working on final projects
Week 10 Apr 1	Depiction of the problem, conveying results of key tests; Mad Writing	<b>Dr. Supp away all week at Data Science conference in Boulder, CO. Available via email or skype. Dr. Brady.</b>
Week 11 Apr 8	Completing a final product for professional and scholarly audiences	Final paper due for peer review process (30%) Executive summary (5%) <b><i>Data Fest (Apr 12-14)!</i></b>
Week 12 Apr 15	Replication and Code Review	Manuscript and data/code review - practice and discuss (s)

Week 13 Apr 22	The policy implications of results	Peer code review completed (5%)
Week 14 Apr 29	Peer review and final presentations	Presentations (10%)
Week 15 May 6	Final meeting wrap-up	Final paper revisions completed and final version turned in to instructor (10%)
<b>CONGRATULATIONS SENIORS !!</b>		

# Seminar in Data Analytics

**DA 401**; Spring 2019  
Denison University  
Burton Morgan 405 (Monday, 1:30-4:20p)

**Instructor:** Dr. Sarah Supp

**Office:** Burton Morgan 409    **Email:** [supps@denison.edu](mailto:supps@denison.edu)    **Phone:** 740-587-5048

**Prerequisites:** For Senior Data Analytics majors only. DA 301, DA 350, CS 181, Math 220, a disciplinary research methods course, and completion of an approved DA summer experience.

**Office Hours:** Mon and Thurs 9:30-11:30 am; or by appointment

## Overview

The seminar is the culmination of the Data Analytics major, a showcase for the problem-driven display of analytic, statistical, programming, and presentational skills through an independent research project. The seminar may build from the summer experience, or investigate a new problem with appropriate data that is acquired or generated. In all cases, students' individual projects will build upon their entire skill set and domain concentration in a complete research project that synthesizes and hones pre-existing skills, develops new project specific techniques, generates deeper domain knowledge, and professionally shares the results through written, visual, and oral communication. Since the topics will vary widely based on student choice, class sessions will resemble workshopping and brainstorming sessions commonplace in research hubs, where peers provide assistance and feedback as projects develop. Students must present well, but also engage with their peers in substantive, constructive ways which maximizes the collaborative nature of research.

## Course Outcomes

*At the end of the course you should have increased confidence and ability to:*

1. Execute an independent *Data Analytics* project with minimal supervision demonstrating the complete research cycle: question development, domain research, hypothesizing, skill development, data gathering/cleaning, and analysis
2. Use self-assessment to identify, and then fill, gaps in domain and skill knowledge needed to complete the project which you are proposing
3. Evaluate analytical and statistical approaches to determine which is the best to provide an accurate and complete answer for the question at hand.
4. Identify key structural elements of an academic research paper, and apply the knowledge created in your project to this format to communicate your results and their relevance.
5. Translate technical research and results professionally and appropriately to technical and non-technical audiences, with an emphasis on written, including oral and visual, formats
6. Provide, seek, and receive peer or supervisory feedback to iteratively work through and improve your data analysis process
7. Manage your time, learning, and tasks efficiently, and apply this management process confidently to the semester project, and in the future, to longer, more complex projects

## Readings

- *The Scientist's Guide to Writing*, Stephen Heard (~\$20 on Amazon)
- Reading will be assigned as appropriate – e.g., regarding a research design that will be used, ethical standards in data presentation, and new statistical and data presentation techniques.
- Selected readings may be made available via NoteBowl or Google Drive
- Self-directed reading relevant to your chosen research topic

## Class Technology (Hardware, Software)

Projects in this course should use the tools that are the best suited to work through your individual data analysis and to produce the product(s) your client has requested. It is strongly suggested that you use workflow and programming tools that will help you to document, manage, and keep track of your changes throughout the semester. You will be required to periodically update your code, visuals, and documentation to [GitHub](#) as a central repository for your instructor to review. Depending on your project and data permissions, you may choose to work on a public repository, or to have your instructor help you set up a private repository.

- **Version Control:** git and GitHub
- **Programming:** whatever tool is right for the job. Examples might include, but are not limited to - R, Python, SQL, Tableau, unix shell, Stata, SPSS, etc...
  - *Note: Your professor is more or less proficient in some of these tools (e.g., R, SQL, etc).*
- **Computers:** Laptops should be sufficient, computer lab or server available if needed

## Online Tools and Class Forum

I will update the syllabus as needed on NoteBowl, and you can use the main forum as a place to share relevant class information with your peers, or to share conceptual questions that may benefit others. Additional information from class time, assignments, and readings may be posted on NoteBowl, as links to a shared Google Drive.

---

## Summary of Key Assignments and Grade Distribution

---

### Individual Research Paper (divided into parts outlined below)

70% (total)

#### Project Proposal(s)

5%

Write two initial project proposals (each <500 words) that clearly establishes a research question, the core of what is to be investigated, the general analytic and design strategy, and the project's purpose, importance, and relevance to both skilled and general audiences. These will be workshopped and reviewed by peers as well as the instructor.

5%

#### Domain Review

An approximately 5 page paper that clearly establishes the driving research question chosen from the workshopping in the previous assignment and makes the case for this question through an early review of key literature in the relevant substantive domain. The process of adding to this body of literature will continue after this assignment so the focus should be on identifying key scholarly sources at the core of the topic, citing correctly, writing a review with a focus on synthesis (not "*author A said... author B said...*"), and identifying the debate/puzzle that the project fits into.

#### Skill Review

5%

A 2-3 page document outlining the skills and techniques commonly used in research and applications of the domain. This paper should also identify scholarly and technical sources for these skills. Finally the paper should identify and discuss strategies to develop skills necessary to complete the project. This will be reviewed and discussed in pairs to help identify skill gaps and strategize opportunities for growth. Pairs will be made with an eye toward skill overlap and possible peer teaching opportunities if possible.

#### Ethical Statement and Protocols

5%

A 1-2 page document that articulates limitations that must be placed on treatment of the data in order to protect individuals involved. Categories of concern involve data gathering procedures, data storage procedures, and reporting of results. We will review IRB and other standards and procedures before completion. As necessary for each project, this statement will serve as the backdrop to secure IRB approval.

#### Research design

5%

A 2-4 page short paper that builds off the skill review and links the research question to a logic of inquiry that involves design, measurement, and statistical choices. Research design involves questions about causality given the level of control involved (experimental vs. non-experimental) and then whether the data involve time, space, and levels as well as its variable structure. Research design statements begin by discussing the needs dictated by the research question and comparison of competing design strategies. If the data are pre-existing, this involves a statement of what can and cannot be said (causally) with the data as well as what an ideal design might entail. It also needs to discuss the statistical analysis choices that need to be made based on the structure of the data and how they are measured. These will be assigned discussants and presented in class. Revisions are expected based upon peer and instructor feedback.

#### Early Results

5%

A 1-2 page summary (not counting visuals) that summarizes early findings of the analytics. The focus here is on an early concrete story and effective written communication of the visual and quantitative data. These will be presented orally and workshopped in class to fine tune the style/voice of the author and identify new questions for the final paper.

## **Final Paper**

40%

25-30 page paper, not counting references and appendices. Write what you need to write to be holistic, comprehensive, and accurate, but the actual length may also depend on your domain and topic. This clearly builds upon the previous parts to assemble a well-written, coherent, final product, that represents an understanding of the research paper format. The final paper is final report should be highly detailed, documenting and defending the choices you made as well as the implications of those choices (are the outcomes different as a result?). It should include presentation-ready data displays as well as technical appendices (e.g., documenting coding and data transformations) as appropriate. The timing of this product will fall 3-5 weeks before the end of the semester to allow for peer and instructor review of the writing and replication of the code. A round of revisions (modeled after a sort of peer review) will allow authors to “revise and resubmit” which may result in higher grade evaluation of this component.

## **Other project related assignments**

**30% (total)**

### **Executive Summary of Findings (and product if appropriate)**

5%

1 page. One of the hardest tasks to do well is to take a mountain of material and distill it into its essential ingredients. Arguably the most important document you will produce, the “executive summary” includes an overview of the question and its stakes, the logic of inquiry and its limitations, the essential findings, and the implications along with appropriate qualifications. It is not a technical document, though it is also not a simple one. Ideally, this is a one page, highly polished statement that communicates your research to a more general, public audience. If the projects is actually able to be public facing the summary, code, and any related tools, dashboards, data, etc... should all be hosted on the student’s medium of choice.

### **Present Findings**

10%

Think of this as an augmented presentation of the executive summary. Effective presentations do not incorporate every detail of the research, but instead distill essential ingredients. They clearly present the question, logic of inquiry, results, and elaborate implications and appropriate qualifications. Presenters should be ready to elaborate any point with further details and be able to defend their choices. Presentations will occur over the last two weeks of class.

### **Replication and Review**

5%

Toward the final paper deadline each student will be paired with a reviewer and discussant who will be responsible for providing feedback on the draft, working to ensure that the data and code for the project are replicable, and providing initial discussant feedback at the beginning of their partner’s presentation of findings.

### **Seminar participation and Scaffolding tasks**

10%

Much more than class attendance, participation in the seminar involves expectations of professional-level feedback to peers on their research projects. Some of this participation will be in writing, while most of it will be oral in the give and take of discussing ongoing, developing projects. Other required stepping stones (i.e. IRB certification, annotated bibliographies, and initial self interest/skill assessment, data visualization workshopping will all factor into this grade).

## **Course Grading**

A+:	98%+	A:	92%	A-:	90%
B+:	88%	B:	82%	B-:	80%
C+:	78%	C:	72%	C-:	70%
D+:	68%	D:	62%	D-:	60%

F: below 60

## Project Information

The individual research projects are designed to give you the opportunity to explore a topic in your career and domain area of interest, while demonstrating mastery and developing greater depth in independent work through the data analysis cycle. The goals of the semester-long project include learning how to develop your own research question and to manage your work using bigger, open-ended data sets or analysis products, and to give you the opportunity to gain practical experience and expertise in using data analysis tools that you will need in the real world while you still have access to experienced teachers to help you navigate the process. The project will focus on written communication and culminate in a final academic paper.

While some class time will be provided to work on projects, you will need to schedule a significant amount of time to work on your project outside of class. Your instructor and peers will provide feedback at several points throughout the semester.

## Expectations

### Participation

This class requires active and engaged participation. Simply listening and reflecting is not enough. I hope each of you will come prepared with insightful questions regarding concepts, skills, readings, and relevant connections to current events and professional settings. There are two key areas of participation in this course:

#### 1. In-class participation

**“Step up, step back”**: I hope everyone plays a role in leading and driving the discussion with your questions and comments, as well as taking time “step back” to listen, and let others have the opportunity to contribute.

My role is to guide, clarify, and add detail where needed. The participation component of your grade will reflect a broad assessment of your fair, productive, and thoughtful oral and written contributions during class time. Your daily attendance and being on time is assumed. *If your circumstances require you to miss class, be late, or to miss multiple classes in a row for any reason, please speak with me ASAP.* For top marks, be an active participant in class on a regular basis, pose questions and comments that contribute to discussion, remain respectful of others, and practice active listening. Consider the following scale:

- (A) *Were I not in the class, quality of class discussions would be significantly diminished.*
- (B/B+) *Were I not in the class, quality of class discussions would be diminished.*
- (B-/B) *Were I not in the class, quality of class discussions would be slightly diminished.*
- (C) *If I was not in the class, the quality of the course would not be changed.*
- (D or lower) *If I was not in the class, the quality of the course would be improved.*

## 2. Scaffolding Tasks

Throughout the semester, there will be several in and out of class scaffolding activities or assignments that are designed to help prepare for and make progress on your projects. These may receive feedback, but may not always receive a posted score on an A-F grade scale. Examples may include: self assessment of interests/skills, IRB certification, annotated bibliographies, data visualization, workshopping and peer review, etc. Completing these activities to the best of your ability will be wrapped into your course participation score.

---

## The Fine Print

### Class Technology

Students are required to provide their own laptops and to install free and open source software on those laptops. Support will be provided by the instructor in the installation of any useful or required software. R can also be accessed via the browser at [r.denison.edu](http://r.denison.edu). If at any time you don't have access to a laptop please contact the instructor and the Data Analytics Program can provide you with a loan from the laptop cart. In class, please use eduroam to connect to the internet instead of Denison Guest.

Please be respectful with your use of laptops and technology in class. I request that you only use them for class related purposes, as I and others may find them distracting (For example, no email or social media should be open in your browser tabs!). Cell phones should be kept silent and put away, and you can expect the same from me.

### Giving and Receiving critical feedback

In this class, you will frequently work with your peers to review, discuss, and provide critical feedback to each other. This peer review process is meant to help you gain practice in the real-world ways that academic research improves and progresses through an iterative cycle of writing, feedback, re-analysis, and re-writing. Similarly, giving and receiving constructive criticism is an important component in other careers as well, and an key skill to develop as a lifelong learner.

Clear communication, being an active and engaged reader and discussant, and providing concrete feedback that is positive and provides a next step will be critical in achieving your end goals for the semester, and in attaining a deep learning experience. I encourage you to approach peer review sessions with an open mind, not to take feedback personally (it's about your project, not *you*) and to be cognizant of how your actions and communication can either help others learn, or create a negative environment that hinders learning and productivity.

### Email and Notebowl

**Read all emails and notebowl posts carefully!** Changes to the syllabus and schedule are likely to happen, but will be communicated in advance as possible in class and notebowl announcements. I will usually use the notebowl feed or email to communicate with the class, so please make sure you are set up to receive email notification from professor and classmate posts. For better or worse, I have to assume you have read all emails. I try to respond within 24hrs, during normal business hours, and expect the same. Emails sent at night (especially before a paper is due) are likely to receive response the next day. NoteBowl will have the reading/assignment schedule and reflect changes to the schedule, reading, and due dates.

## **Absences and emergencies**

Attendance is expected (me too right?). That said, we are adults and life happens. If you have to miss (I hope you don't) I trust it is for a good reason. Given the seminar-based nature of the class, I am sure you will try not to miss unless it is unavoidable. That said, any more than one absence (the equivalent of one week of class) or a pattern of showing up late for class, for unexcused or non-emergency reasons could negatively affect your participation grade, and likely would have consequences in other areas. Of course, in the case where something unexpected does occur (illness, emergencies, etc..) reach out, let me know, and we will work on it. I care about you and all of our well-being. Let me help if I can.

## **Due dates, extensions, penalties**

Given the fast pace and independently-driven nature of the class and that it should be roughly in sync with the other section, deadlines for major assignments will generally be non-negotiable and late assignments may not be accepted. Occasionally we might move a deadline for everyone, but do not expect a deadline to shift individually except in the most extreme circumstances (all students are busy and have multiple deadlines). *Except in the case of a university wide issue, computer based excuses will be hard to defend -- save and backup regularly.* The final project report and presentation are hard deadlines and late materials or presentations will not be accepted.

## **A short note on citations**

Simply, cite your work. Unless otherwise noted, any assignment or presentation that relies on or references (either by paraphrase, direct quote, or data sourcing) the intellectual property of others, should be attributed to that source. In this class I am less concerned with the format but would ask that you choose a style, confirm it with me, and stay consistent. I would like citations to include the year of publication when cited. For example: (authors' last names, year). Citations/sourcing should appear in text (parentheticals, not footnotes) and should also have a references section at the end, with complete entries. If you are looking for a specific method consider the APA styles. I assume that you are familiar with academic citations, and will not spend a significant amount of class time on this. Please ask if you are unsure. Again the emphasis is less of the choice of a specific format, but rather transmissibility and transparency ([https://owl.purdue.edu/owl/research\\_and\\_citation/apa\\_style/apa\\_style\\_introduction.html](https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_style_introduction.html)).

## **Academic Honesty and Integrity**

Academic honesty, the cornerstone of teaching and learning, lays the foundation for lifelong integrity. Academic dishonesty is intellectual theft. It includes, but is not limited to, providing or receiving assistance in a manner not authorized by the instructor in the creation of work to be submitted for evaluation. This standard applies to all work ranging from daily homework assignments to major projects or exams. Students must clearly cite any sources consulted—not only for quoted phrases but also for ideas and information that are not common knowledge. Neither ignorance nor carelessness is an acceptable defense in cases of plagiarism. It is the student's responsibility to follow the appropriate format for citations. Students should ask their instructors for assistance in determining what sorts of materials and assistance are appropriate for assignments and for guidance in citing such materials clearly. *For further information about the Code of Academic Integrity, see <http://denison.edu/academics/curriculum/integrity>.*

## **Disability Accommodation**

Any student who feels they may need an accommodation based on the impact of a disability should contact me privately as soon as possible to discuss his or her specific needs. I rely on the Academic Resource Center (ARC) in 020 Higley to verify the need for reasonable accommodations based on the documentation on file in that office.

## Writing Center

The Writing Center is a free resource available to all Denison students. Student writing consultants from many majors help writers one-on-one in all phases of the writing process, from deciphering the assignment, to discussing ideas, to developing an argument, to finalizing a draft. Because proofreading is a last step in that process, you should leave plenty of time (like at least a week) for getting your ideas right before expecting proofreading help. Consultants also can help writers with personal documents, like job and internship applications. Consultants welcome diversity and are prepared to work with multilingual writers. If needed, Consultants can refer a multilingual writer to Denison's Coordinator of Multilingual Learning for additional support. The Center is located on the fourth floor of Barney-Davis Hall; a satellite location is in the Learning Commons on the entrance level of the Library.

Appointments between 4 pm and 9 pm, Sunday through Thursday, can be made for the Barney-Davis location on the online scheduler at the MyDenison Writing Center website; the library satellite location only is drop-in. Check the website on MyDenison for those hours.

## Multilingual Support

In addition to the academic support services available to all Denison students, students who use English as a second (or third, etc.) language, can meet with Denison's Interim Coordinator of Multilingual Learning, Kalynda Thayer. She offers a variety of support for L2 students, from consulting with you about your written work to helping you devise strategies for developing and effectively using your listening, speaking, reading, and writing skills in English. You can email her at [kalynda.thayer@denison.edu](mailto:kalynda.thayer@denison.edu) to schedule an appointment.

## Reporting Sexual Assault

Essays, journals, and other coursework submitted for this class are generally considered confidential pursuant to the University's student record policies. However, students should be aware that University employees are required by University policy and Title IX guidance to report allegations of discrimination based on gender identity / expression, including sexual misconduct, sexual assault and suspected child abuse/neglect, occurring on campus and / or involving current students at Denison University when they become aware of possible incidents in the course of their employment, including via coursework or advising conversations. There are others on campus to whom you may speak in confidence, including counselors at the Whisler Center for Student Wellness, SHARE advocates, and clergy. More information on Title IX and University policy guidance on gender identity / expression bias and sexual misconduct / assault, including support resources, how to report, and prevention and education efforts, can be found at [denison.edu/titleix](http://denison.edu/titleix); students may also contact Steve Gauger, Campus Title IX Coordinator, in Doane Administration 001, by email at [gaugers@denison.edu](mailto:gaugers@denison.edu), or by phone at 740-587-8660.

## Cautions regarding copyright and licensing

All documents provided to you (i.e. syllabus, assignment prompts, etc.) are the property of the instructor, author, or client. It is a violation of intellectual property to post these online (especially to websites promoting copying/cheating) or to provide them to students not in our class or in future classes. For most of the datasets, sharing them, or the finished reports based on them, will also be a **serious ethical breach**. Your papers are your property, and while you can do with them as you wish, it may be a violation of academic integrity to make them available to others who might use them for plagiarism. Basically, keep course materials and your work to yourself except in the process of classroom editing and peer review.

## **Writing Intensive Seminar “W” General Education Requirement**

Written communication is a critical skills for all data analysts, regardless of your final career path. The writing intensive seminar at Denison encourages students to cultivate the habits of effective writers: (1) Take intellectual risks and see the connection between critical reading, thinking, and writing, (2) Regard writing as a process, and refine your own habits of researching and writing, and give, receive, and implement critical feedback, and (3) Consider your audience and how your writing allows you to join a broader, ongoing, community discussion.

As a Denisonian completing your final W course, you should be able to:

- Craft and support a coherent argument
- Anticipate and meet the needs of your audience
- Gather and synthesize evidence
- Understand writing as a process and apply conventions of style and grammar effectively

Throughout the semester, you will gain instruction, practice, and feedback on written communication through in-class participation and peer review, discussions, and individual practice. Communication is the major way that data analysis efforts become efficient, useful and relevant. While multiple methods of communication are practiced throughout the course, written communication for both technical and non-technical audiences is frequently required and assessed.