THE DATA MINE: WEEKLY SEMINAR

FALL 2021 SYLLABUS

COURSE INFORMATION

Course number and title: STAT 19000/29000/39000 The Data Mine

Course credit hours: 1 credit hour, so you should expect to spend about 3 hours per week doing work for the class

Course web page: All information will be posted on Brightspace. https://purdue.brightspace.com/
The Data Mine's website can also be helpful: https://datamine.purdue.edu

Course platform (RStudio Server Pro): https://rstudio.scholar.rcac.purdue.edu/

Prerequisites: None for STAT 19000. All students, regardless of background are welcome. Typically, students new to The Data Mine sign up for STAT 19000, students in their second year of The Data Mine sign up for STAT 29000, and students in their third year of The Data Mine sign up for STAT 39000. However, during the first week of the semester only, if a student new to The Data Mine has several years of coding and/or data science experience and would prefer to switch to STAT 29000, we can make adjustments. A placement exam was offered before the semester started, too.

Meeting times: There are officially 4 Monday class times: 8:30 am, 9:30 am, 10:30 am (all in the Hillenbrand Dining Court atrium—no meal swipe required), and 4:30 pm (synchronous online, recorded and posted later). All the information you need to work on the projects each week will be provided online on the Thursday of the previous week, and we encourage you to get a head start on the projects before class time. Dr. Ward does not lecture during the class meetings, but this is a good time to ask questions and get help from Dr. Ward, the T.A.s, and your classmates. Attendance is not required. The T.A.s will have many daytime and evening office hours throughout the week.

COURSE DESCRIPTION

The Data Mine is a supportive environment for students in any major from any background who want to learn some data science skills.   Students will have hands-on experience with computational tools for representing, extracting, manipulating, interpreting, transforming, and visualizing data, especially big data sets, and in effectively communicating insights about data. Topics include: the R environment, Python, visualizing data, UNIX, bash, regular expressions, SQL, XML and scraping data from the internet, as well as selected advanced topics, as time permits.

LEARNING OUTCOMES

By the end of the course, you will be able to:

1. Discover data science and professional development opportunities in order to prepare for a career.
2. Explain the difference between research computing and basic personal computing data science capabilities in order to know which system is appropriate for a data science project.
3. Design efficient search strategies in order to acquire new data science skills.
4. Devise the most appropriate data science strategy in order to answer a research question.
5. Apply data science techniques in order to answer a research question about a big data set.

INFORMATION ABOUT THE DATA MINE STAFF

The Data Mine's staff has grown and will continue to grow this year. For the most up-to-date information about who works in The Data Mine, see:
https://datamine.purdue.edu/about/welcome.html

For the purposes of getting help with this 1-credit seminar class, your most important people are:

- T.A.s: Visit their office hours and use the Piazza site

- General Data Mine staff email: datamine@purdue.edu (read by multiple staff members, this usually gets a faster response than emailing an individual)

- Mr. Kevin Amstutz, Senior Data Scientist and Instruction Specialist, kamstut@purdue.edu or through Piazza

- Dr. Mark Daniel Ward, Director, mdw@purdue.edu or through Piazza (usually faster responding on Piazza than he is by email)

EMAIL POLICY

- For general Data Mine questions, email datamine@purdue.edu
- For questions about how to do the homework, use Piazza or visit office hours.
- For regrade requests, use Gradescope's regrade feature within Brightspace. Regrades should be requested within 1 week of the grade being posted.

Last updated: 6/10/21 12:30 PM

OFFICE HOURS

The office hours schedule is posted here: https://docs.google.com/document/d/1ooJtqOs2U4y-u6c1O51EIKmXWjHNMS-uaZVQkUACRJ4/ Office hours are held in person and/or using Webex. With Webex, it's easy for you to share your screen and talk face-to-face with the TA or staff member.

Piazza is an online discussion board where students can post questions at any time, and Data Mine staff or T.A.s will respond. Piazza is available through Brightspace. There are private and public postings. Last year we had over 11,000 interactions on Piazza, and the typical response time was around 5 minutes.

HOW TO SUCCEED IN THIS COURSE

If you would like to be a successful Data Mine student:

- Be excited to challenge yourself and learn impressive new skills. Don't get discouraged if something is difficult—you're here because you want to learn, not because you already know everything!
- Start on the weekly projects on or before Mondays so that you have plenty of time to get help from your classmates, TAs, and Data Mine staff. Don't wait until the due date to start!
- Remember that Data Mine staff and TAs are excited to work with you! Take advantage of us as resources.
- Network! Get to know your classmates, even if you don't see them in an actual classroom. You are all part of The Data Mine because you share interests and goals. You have over 800 potential new friends!
- Use "The Examples Book" with lots of explanations and examples to get you started. Google, Stack Overflow, etc. are all great, but "The Examples Book" has been carefully put together to be the most useful to you. https://the-examples-book.com
- Expect to spend approximately 3 hours per week on the projects. Some might take less time, and occasionally some might take more.
- Don't forget about the syllabus quiz, academic integrity quiz, and outside event reflections. They all contribute to your grade and are part of the course for a reason.
- If you get behind or feel overwhelmed about this course or anything else, please talk to us!
- Stay on top of deadlines. Announcements will also be sent out every Monday morning, but you should keep a copy of the course schedule where you see it easily.
- Read your emails!

REQUIRED MATERIALS

- A laptop so that you can easily work with others. Having audio/video capabilities is useful.
- Brightspace course page.
- Access to Scholar.

- "The Examples Book": https://the-examples-book.com
- Good internet connection.

COURSE LOGISTICS

See the schedule and later parts of the syllabus for more details, but here is an overview of how the course works:

In the first week of the beginning of the semester, you will have some "housekeeping" tasks to do, which include taking the Syllabus quiz, Academic Integrity quiz, and Check-in #1.

Generally, every week from the very beginning of the semester, you will have your new projects released on a Thursday, and they are due 8 days later on the Friday at 11:55 pm Purdue West Lafayette (Eastern) time. You will need to do 3 Outside Event reflections (due September, October, and November).

We will have 13 weekly projects available, but we only count your best 10. This means you could miss up to 3 projects due to illness or other reasons, and it won't hurt your grade. We suggest trying to do as many projects as possible so that you can keep up with the material. The projects are much less stressful if they aren't done at the last minute, and it is possible that our systems will be stressed if you wait until Friday night causing unexpected behavior and long wait times. Try to start your projects on or before Monday each week to leave yourself time to ask questions.

ASSIGNMENTS AND GRADES

Projects (best 10 out of Projects #1-13)	86%
Outside event reflections (3 total)	12%
Academic Integrity Quiz	1%
Syllabus Quiz	1%
Total	100%

In this class grades reflect your achievement throughout the semester in the various course components listed above. Your grades will be maintained in Brightspace. This course will follow the 90-80-70-60 grading scale for A, B, C, D cut-offs. If you earn a 90.000 in the class, for example, that is a solid A. +/- grades will be given at the instructor's discretion below these cut-offs. If you earn an 89.11 in the class, for example, this may be an A- or a B+.

PROJECTS

- The projects will help you achieve Learning Outcomes #2-5.
- Each weekly programming project is worth 10 points.
- There will be 13 projects available over the semester, and your best 10 will count.
- The 3 project grades that are dropped could be from illnesses, absences, travel, family emergencies, or simply low scores. No excuses necessary.

- No late work will be accepted, even if you are having technical difficulties, so do not work at the last minute.
- There are many opportunities to get help throughout the week, either through Piazza or office hours. We're waiting for you! Ask questions!
- Follow the instructions for how to submit your projects properly through Gradescope in Brightspace.
- It is ok to get help from others or online, although it is important to document this help in the comment sections of your project submission. You need to say who helped you and how they helped you.
- Each week, the project will be posted on the Thursday before the seminar, the project will be the topic of the seminar and any office hours that week, and then the project will be due by 11:55 pm Eastern time on the following Friday. See the schedule for specific dates.
- If you need to request a regrade on any part of your project, use the regrade request feature inside Gradescope. The regrade request needs to be submitted within one week of the grade being posted (we send an announcement about this).

OUTSIDE EVENT REFLECTIONS

- The Outside Event reflections will help you achieve Learning Outcome #1. They are an opportunity for you to learn more about data science applications, career development, and diversity.
- Throughout the semester, The Data Mine will have many special events and speakers, typically happening in person so you can interact with the presenter, but some may be online and possibly recorded.
- These eligible opportunities will be posted on The Data Mine's website and updated frequently (https://docs.google.com/document/d/1tMRrVM08WEqs4E76K9N0Ik6TpytiVBsadUM4ialYqzw/) . Feel free to suggest good events that you hear about, too.
- You are required to attend **3** of these over the semester, with 1 due each month. See the schedule for specific due dates.
- You are welcome to do all 3 reflections early. For example, you could submit all 3 reflections in September.
- **You must submit your outside event reflection within 1 week of attending the event or watching the recording.**
- Follow the instructions on Brightspace for writing and submitting these reflections.
- **At least one of these events should be on the topic of Professional Development. These events will be designated by "PD" next to the event on the schedule.**
- For each of the 3 required events, write a minimum 1-page (double-spaced, 12-pt font) reflection that includes the name of the event and speaker, the time and date of the event, what was discussed at the event, what you learned from it, what new ideas you would like to explore as a result of what you learned at the event, and what question(s) you would like to ask the presenter if you met them at an after-presentation reception. This should **not** be just a list of notes you took from the event—it is a reflection. The header of your reflection should not take up more than 2 lines!
- We read every single reflection! We care about what you write! We have used these connections to provide new opportunities for you, to thank our speakers, and to learn more about what interests you.

LATE POLICY

We generally do NOT accept late work. For the projects, we count only your best 10 out of 13, so that gives you a lot of flexibility. We need to be able to post answer keys for the rest of the class in a timely manner, and we can't do this if we are waiting for other students to turn their work in.

DIVERSITY AND INCLUSION STATEMENT

Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. More details are available on our course Brightspace table of contents, under University Policies.

In The Data Mine, it is very important to us that ALL feel welcomed, supported, challenged, and celebrated. We all bring different strengths and experiences to this learning community, and that is part of the value of being in a community. We can learn from each other. Please speak to any of The Data Mine staff, anonymously if needed, if something has made you uncomfortable so that we can work toward a solution.

NONDISCRIMINATION STATEMENT

Purdue University is committed to maintaining a community which recognizes and values the inherent worth and dignity of every person; fosters tolerance, sensitivity, understanding, and mutual respect among its members; and encourages each individual to strive to reach his or her own potential. In pursuit of its goal of academic excellence, the University seeks to develop and nurture diversity. The University believes that diversity among its many members strengthens the institution, stimulates creativity, promotes the exchange of ideas, and enriches campus life. Link to Purdue's nondiscrimination policy statement.

COURSE SCHEDULE

Week	Date	New project posted	Seminars	Projects due (all at 11:55 pm)
0	Th 8/19	Project #1, Syllabus quiz, and Academic Integrity quiz posted		
1	M 8/23		Seminars for Project #1	
	Th 8/26	Project #2 posted		
	F 8/27			**Syllabus quiz Academic Integrity quiz, Project #1 all due**
2	M 8/30		Seminars for Project #2	
	Th 9/2	Project #3 posted		
	F 9/3			**Project #2 due**
3	M 9/6	No classes (Labor Day)		
4	M 9/13		Seminars for Project #3	
	Th 9/16	Project #4 posted		
	F 9/17			**Project #3 due**
5	M 9/20		Seminars for Project #4	
	Th 9/23	Project #5 posted		
	F 9/24			**Project #4 due & Outside Event #1 due**
6	M 9/27		Seminars for Project #5	
	Th 9/30	Project #6 posted		
	F 10/1			**Project #5 due**
7	M 10/4		Seminars for Project #6	
	Th 10/7	Project #7 posted		
	F 10/8			**Project #6 due**
8	M 10/11	No classes (October Break)		

Last updated: 6/10/21 12:30 PM

Week	Date	New project posted	Seminars	Projects due (all at 11:55 pm)
9	M 10/18		Seminars for Project #7	
	Th 10/21	Project #8 posted		
	F 10/22			**Project #7 due & Outside Event #2 due**
10	M 10/25		Seminars for Project #8	
	Th 10/28	Project #9 posted		
	F 10/29			**Project #8**
11	M 11/1		Seminars for Project #9	
	Th 11/4	Project #10 posted		
	F 11/5			**Project #9 due**
12	M 11/8		Seminars for Project #10	
	Th 11/11	Project #11 posted		
	F 11/12			**Project #10 due**
13	M 11/15		Seminars for Project #11	
	Th 11/18	Project #12 posted		
	F 11/19			**Project #11 due & Outside Event #3 due**
14	M 11/22		Seminars for Project #12	
	Th 11/25	Thanksgiving Break (no classes)		
	F 11/26	Thanksgiving Break (no classes)		
15	M 11/29		Seminars for Project #12	
	Th 12/2	Project #13 posted		
	F 12/3			**Project #12 due**
16	M 12/6		Seminars for Project #13	
	F 12/10			**Project #13 due**

ACADEMIC INTEGRITY

In STAT 19000/29000/39000, we encourage students to work together. However, there is a difference between good collaboration and academic misconduct. We expect you to read over this list, and you will be held responsible for violating these rules. We are serious about protecting the hard-working students in this course. We want a grade for STAT 19000/29000/39000 to have value for everyone and to represent what you truly know. We may punish both the student who cheats and the student who allows or enables another student to cheat. Punishment could include receiving a 0 on a project, receiving an F for the course, and/or being reported to the Office of The Dean of Students.

Good Collaboration:
- First try the project yourself, on your own.
- **After** trying the project yourself, then get together with a small group of other students who have also tried the project themselves to discuss ideas for how to do the more difficult problems. Document in the comments section any suggestions you took from your classmates or your TA.
- Finish the project on your own so that what you turn in truly represents your own understanding of the material.
- Look up potential solutions for how to do part of the project online, but document in the comments section where you found the information.
- If the assignment involves writing a long, worded explanation, you may proofread somebody's completed written work and allow them to proofread your work. Do this only after you have both completed your own assignments, though.

Academic Misconduct:
- Divide up the problems among a group. (You do #1, I'll do #2, and he'll do #3: then we'll share our work to get the assignment done more quickly.)
- Attend a group work session without having first worked all of the problems yourself.
- Allowing your partners to do all of the work while you copy answers down, or allowing an unprepared partner to copy your answers.
- Letting another student copy your work or doing the work for them.
- Sharing files or typing on somebody else's computer or in their Scholar account.
- Getting help from a classmate or a TA without documenting that help in the comments section.
- Looking up a potential solution online without documenting that help in the comments section.
- Reading someone else's answers before you have completed your work.
- Have a tutor or TA work though all (or some) of your problems for you.
- Uploading, downloading, or using old course materials from Course Hero, Chegg, or similar sites.
- Using the same outside event reflection (or parts of it) more than once. Using an outside event reflection from a previous semester.
- Using somebody else's outside event reflection rather than attending the event yourself.

Academic integrity is one of the highest values that Purdue University holds. Individuals are encouraged to alert university officials to potential breaches of this value by either emailing or by calling 765-494-8778. While information may be submitted anonymously, the more information that is submitted provides the greatest opportunity for the university to investigate the concern.

The Purdue Honor Pledge "As a boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do. Accountable together - we are Purdue."

Please refer to the student guide for academic integrity for more details.

Purdue's Copyrighted Materials Policy:

Among the materials that may be protected by copyright law are the lectures, notes, and other material presented in class or as part of the course. Always assume the materials presented by an instructor are protected by copyright unless the instructor has stated otherwise. Students enrolled in, and authorized visitors to, Purdue University courses are permitted to take notes, which they may use for individual/group study or for other non-commercial purposes reasonably arising from enrollment in the course or the University generally.

Notes taken in class are, however, generally considered to be "derivative works" of the instructor's presentations and materials, and they are thus subject to the instructor's copyright in such presentations and materials. No individual is permitted to sell or otherwise barter notes, either to other students or to any commercial concern, for a course without the express written permission of the course instructor. To obtain permission to sell or barter notes, the individual wishing to sell or barter the notes must be registered in the course or must be an approved visitor to the class. Course instructors may choose to grant or not grant such permission at their own discretion, and may require a review of the notes prior to their being sold or bartered. If they do grant such permission, they may revoke it at any time, if they so choose.

STUDENTS WITH DISABILITIES

Purdue University strives to make learning experiences as accessible as possible. If you anticipate or experience physical or academic barriers based on disability, you are welcome to let me know so that we can discuss options. You are also encouraged to contact the Disability Resource Center at: **drc@purdue.edu** or by phone: 765-494-1247.

If you have been certified by the Office of the Dean of Students as someone needing a course adaptation or accommodation because of a disability OR if you need special arrangements in case the building must be evacuated, please contact The Data Mine staff during the first week of classes. We are happy to help you.

MENTAL HEALTH AND ADDITIONAL RESOURCES

- **If you find yourself beginning to feel some stress, anxiety and/or feeling slightly overwhelmed, try WellTrack.** Sign in and find information and tools at your fingertips, available to you at any time.
- **If you need support and information about options and resources**, please see the Office of the Dean of Students for drop-in hours (M-F, 8 am- 5 pm).
- **If you're struggling and need mental health services**: Purdue University is committed to advancing the mental health and well-being of its students. If you or someone you know is feeling overwhelmed, depressed, and/or in need of mental health support, services are

available. For help, such individuals should contact [Counseling and Psychological Services (CAPS)](#) at 765-494-6995 during and after hours, on weekends and holidays, or by going to the CAPS office of the second floor of the Purdue University Student Health Center (PUSH) during business hours.

- Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact the Dean of Students for support. Furthermore, please notify Data Mine staff, if you are comfortable in doing so.

- Considering the significant disruptions caused by the current global crisis as it related to COVID-19, students may submit requests for emergency funds. Click [here](#) for details and the request form.

VIOLENT BEHAVIOR POLICY

Purdue University is committed to providing a safe and secure campus environment for members of the university community. Purdue strives to create an educational environment for students and a work environment for employees that promote educational and career goals. Violent Behavior impedes such goals. Therefore, Violent Behavior is prohibited in or on any University Facility or while participating in any university activity.

See the [University's full violent behavior policy](#) for more detail.

ACADEMIC GUIDANCE IN THE EVENT A STUDENT IS QUARANTINED/ISOLATED

If you become quarantined or isolated at any point in time during the semester, in addition to support from the Protect Purdue Health Center, you will also have access to an Academic Case Manager who can provide you academic support during this time. Your Academic Case Manager can be reached at acmq@purdue.edu and will provide you with general guidelines/resources around communicating with your instructors, be available for academic support, and offer suggestions for how to be successful when learning remotely. Importantly, if you find yourself too sick to progress in the course, notify your academic case manager and notify The Data Mine staff (datamine@purdue.edu). We will make arrangements based on your particular situation. The Office of the Dean of Students (odos@purdue.edu) is also available to support you should this situation occur.

ATTENDANCE POLICY DURING COVID-19

While everything we are doing in The Data Mine this semester can be done online, rather than in person, and no part of your seminar grade comes from attendance, we want to remind you of general campus attendance policies during COVID-19. Students should stay home and contact the Protect Purdue Health Center (496-INFO) if they feel ill, have any symptoms associated with COVID-19, or suspect they have been exposed to the virus. In the current context of COVID-19, in-person attendance will not be a factor in the final grades, but the student still needs to inform the instructor of any conflict that can be anticipated and will affect the submission of an assignment. Only the instructor can excuse a

student from a course requirement or responsibility. When conflicts can be anticipated, such as for many University-sponsored activities and religious observations, the student should inform the instructor of the situation as far in advance as possible. For unanticipated or emergency conflict, when advance notification to an instructor is not possible, the student should contact the instructor as soon as possible by email or by phone. When the student is unable to make direct contact with the instructor and is unable to leave word with the instructor's department because of circumstances beyond the student's control, and in cases of bereavement, quarantine, or isolation, the student or the student's representative should contact the Office of the Dean of Students via email or phone at 765-494-1747. Our course Brightspace includes a link on Attendance and Grief Absence policies under the University Policies menu.

GENERAL CLASSROOM GUIDANCE REGARDING PROTECT PURDUE

We want to remind you of the general Protect Purdue rules for groups of students. The Protect Purdue Plan, which includes the Protect Purdue Pledge, is campus policy and as such all members of the Purdue community must comply with the required health and safety guidelines. Required behaviors in this class include: staying home and contacting the Protect Purdue Health Center (496-INFO) if you feel ill or know you have been exposed to the virus, wearing a mask in classrooms and campus building, at all times (e.g., no eating/drinking in the classroom), disinfecting desk/workspace prior to and after use, maintaining proper social distancing with peers and instructors (including when entering/exiting classrooms), refraining from moving furniture, avoiding shared use of personal items, maintaining robust hygiene (e.g., handwashing, disposal of tissues) prior to, during and after class, and following all safety directions from the instructor.

Students who are not engaging in these behaviors (e.g., wearing a mask) will be offered the opportunity to comply. If non-compliance continues, possible results include instructors asking the student to leave class and instructors dismissing the whole class. Students who do not comply with the required health behaviors are violating the University Code of Conduct and will be reported to the Dean of Students Office with sanctions ranging from educational requirements to dismissal from the university.

Any student who has substantial reason to believe that another person in a campus room (e.g., classroom) is threatening the safety of others by not complying (e.g., not wearing a mask) may leave the room without consequence. The student is encouraged to report the behavior to and discuss next steps with their instructor. Students also have the option of reporting the behavior to the Office of the Student Rights and Responsibilities. See also Purdue University Bill of Student Rights.

EMERGENCIES

Campus emergencies:
In the event of a major campus emergency, course requirements, deadlines and grading percentages are subject to changes that may be necessitated by a revised semester calendar or other circumstances. Here are ways to get information about changes in this course: Brightspace or by e-mail from Data Mine

staff. General information about a campus emergency can be found on the Purdue website: www.purdue.edu.

Illness and other Student Emergencies:
Students with extended illnesses should contact their instructor as soon as possible so that arrangements can be made for keeping up with the course. Extended absences/illnesses/emergencies should also go through the Office of the Dean of Students.

Hillenbrand Hall emergency plans:
https://www.purdue.edu/ehps/emergency_preparedness/bep/hill-bep.html

EMERGENCY NOTIFICATION PROCEDURES are based on a simple concept – if you hear a fire alarm inside, proceed outside. If you hear a siren outside, proceed inside.
- **Indoor Fire Alarms** mean to stop class or research and immediately **evacuate** the building. Proceed to your Emergency Assembly Area away from building doors. **Remain outside** until police, fire, or other emergency response personnel provide additional guidance or tell you it is safe to leave.
- **All Hazards Outdoor Emergency Warning Sirens** mean to immediately seek shelter (**Shelter in Place**) in a safe location within the closest building. "Shelter in place" means seeking immediate shelter inside a building or University residence. This course of action may need to be taken during a tornado, a civil disturbance including a shooting or release of hazardous materials in the outside air. Once safely inside, find out more details about the emergency*. **Remain in place** until police, fire, or other emergency response personnel provide additional guidance or tell you it is safe to leave.

In both cases, you should seek additional clarifying information by all means possible…Purdue Home page, email alert, TV, radio, etc…review the Purdue Emergency Warning Notification System multi-communication layers at https://www.purdue.edu/securepurdue/

EMERGENCY RESPONSE PROCEDURES:
- Review the **Emergency Procedures Guidelines** https://www.purdue.edu/ehps/emergency_preparedness/flipchart/index.html
- Review the **Building Emergency Plan** (available from the building deputy) for:
    - evacuation routes, exit points, and emergency assembly area
    - when and how to evacuate the building.
    - shelter in place procedures and locations
    - additional building specific procedures and requirements.

MORE INFORMATION
Reference the Emergency Preparedness web site for additional information: http://www.purdue.edu/emergency_preparedness

Last updated: 6/10/21 12:30 PM