BADM 372 Applied Analytics

BADM 372

2023-01-16

Contents

Chapter 1

About this course

This website serves as headquarters for **BADM 372 Applied Analytics**.

Content here will be updated with any changes made during the semester, so if at any point you are told there was a change in the schedule or an assignment, you can come here to get the updated version.

Also, this website has benefited greatly from lots of free, readily available resources posted on the web and we leverage these extensively. I would encourage you to review these resources in your analytics journey. Some that we specifically use with great frequency are these (**and I say a loud THANK YOU to the authors!**):

- R for Data Science
- An Introduction to Statistical Learning with Applications in R
- Data Science in a Box
- stackoverflow.com, for example

Chapter 2

Syllabus

Instructor: Tobin Turner

Office Hours: mutually convenient time arranged by email e-mail: jtturner@presby.edu

2.1 Course Objectives and Learning Outcomes

This course is designed to introduce to data science. Students will apply statistical knowledge and techniques to both business and non-business contexts.

At the end of this course students should be able to:

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

- gain insight from data
- gain insight from data, reproducibly
- gain insight from data, reproducibly, using modern programming tools and techniques
- gain insight from data, reproducibly and collaboratively, using modern programming tools and techniques
- gain insight from data, reproducibly (with literate programming and version control) and collaboratively, using modern programming tools and techniques
- communicate results effectively

This course will be focused on both understanding and applying key business analytical concepts. Although the text serves as a useful foundation for the concepts covered in the class, simple memorization of the material in the text will not be sufficient. Class participation, discussion, and application are critical.

2.2   Text and Resources

- This course website (primary resource)
- R for Data Science
- An Introduction to Statistical Learning with Applications in R
- Data Science in a Box
- stackoverflow.com, for example
- Other free, publicly available datasets and publications.

2.3   Performance Evaluation (Grading)

- Labs, Quizzes and Assignments - 40%
- Exam 1 - 20%
- Exam 2 - 20%
- Final Exam - 20%

2.3.1   Missed Exams/Assignments = ZERO

Arrangements for missed or late assignments must be made **PRIOR** to the exam or due date or the student may receive a ZERO grade for the exam or assignment. See attendance and quiz sections below for more details.

- **Missed Quizzes and Assignments cannot be made up later. Be present.**

2.3.2   Exams

Exams will cover assigned chapters in the textbook, other assigned readings, lectures, class exercises, class discussions, videos, and guest speakers. I will typically allocate time prior to each exam to clearly identify the body of knowledge each test will cover and to answer questions about the format and objectives of the exam.

2.3.3   Labs/Quizzes – DON'T MISS CLASS

- The average of all labs, quizzes and assignments will comprise the Quizzes and Assignments - 40% portion of your final grade
- Quizzes and Assignments are designed to prepare you for your exams and to ensure you stay up with the course material
- **Missed Quizzes and Assignments cannot be made up later. Be present.**

Quizzes rule. **LISTEN.** - **Missed Quizzes and Assignments cannot be made up later. Be present.**

2.3.4 Final Average

- Final Average Grade

    - 90-100 A
    - 88-89 B+
    - 82-87 B+
    - 80-81 B-
    - 78-79 C+
    - 72-77 C+
    - 70-71 C-
    - 60-69 D
    - 59 and below F

2.4 Class Participation:

I will frequently give readings or assignments for you to complete prior to the next class meeting. I expect you to fully engage the material: answer questions, pose questions, provide insightful observations. Keep in mind that quality is an important component in "participation." Periodic cold calls will take place. I will also put students in the "hot seat" on occasion. In these class sessions, I may select a random group of students to lead us in the discussion and debate. Because the selection of participants will not be announced until class begins, everyone will be expected to prepare for the discussion. Reading the assigned chapters and articles are the best way to prepare for the discussion. If you have concerns about being called on in class, please see me to discuss. The purpose of the "hot seat" is not to stress or embarrass students, but to encourage students to actively engage the material.

2.5 Phones

Phones are not allowed to be used in class without the instructor's prior consent. If you have a need of a phone during class please let me know before class. Unauthorized use of electronic devices may result in the lowering of the grade or dismissal from the class. **I mean this.**

The phone thing? I mean this.

2.6   Attendance

You are expected to be regular and punctual in your class attendance. Students are responsible for all the material missed and homework assignments made. If class is missed, notes/homework should be obtained from another student. If I am more than 15 minutes late, class is considered cancelled. No more than 4 absences are allowed during a semester. Exceeding the absence policy may result in receiving an F for the course. The professors roll is the official roll and students not present when roll is taken will be counted as absent. If a student must miss an exam, she or he must work out an agreeable time with the instructor to take the test prior to the exam being given. If a student misses a test due to an emergency, the student must inform the instructor as soon as is possible. In special cases, the instructor may allow the student to take a make-up exam.

2.7   Accommodations

Presbyterian College is committed to providing reasonable accommodations for all students with documented disabilities. If you are seeking academic accommodations under the Americans with Disabilities Act, you must register with the Academic Success Office, located on 5th Avenue (beside Campus Police). To receive these accommodations, please obtain the proper Accommodations Approval Form from that office, and then meet with me at the beginning of the semester to discuss how we may deliver your approved accommodations. I especially encourage you to meet with me well in advance of the actual accommodations being provided, as it may not be feasible to offer immediate accommodations without sufficient advance notice (such as in the case of tests). I can assure you that all discussions will remain confidential. Disability Services information is located at this link http://bit.ly/PCdisabilityservices

Additionally, it is the student's responsibility to give the instructor one week's notice prior to each instance where accommodation will be required.

2.8   Honor Code and Plagiarism:

All assignments/exams must be your own work. Any copying or use of unauthorized assistance will be treated as a violation of PC's Honor Code. If you are unsure of what resources are allowed, please ask. Please note that all text longer than 7 words taken from ANY other source must be placed in quotations and cited. Also, summarizing ANY other source must also be cited. Using ANY other source and showing work to be your own is a violation of plagiarism and the honor code.

2.9   First-Generation Version:

I am a Presby First+ Advocate. I am here to support our current first-generation students. At Presbyterian College, first-generation students are those in which neither parent nor legal guardian graduated from a four-year higher education institution with a bachelor's degree. If you are a first-generation college student, please contact me. For more information about support for first-generation college students on our campus visit our Presby First+ webpage.

2.10   Continuing Advocate Version

I am a Presby First+ Advocate. I am committed to supporting first-generation students at Presbyterian College. At Presbyterian College, first-generation students are those in which neither parent nor legal guardian graduated from a four-year higher education institution with a bachelor's degree. If you are a first-generation college student, please contact me anytime or visit me during my office hours. For more information about support for first-generation college students on our campus visit our Presby First+ webpage.

Chapter 3

Schedule

This is a tentative schedule, and it will change. **BUT** I will do my very best to review this often so that we all stay on the same page and so that you may plan accordingly!

Spring 2023

Date	Topic	Due
Monday, January 9, 2023	Course Intro, R4DS, A1 review; Lab 1	
Wednesday, January 11, 2023	Rmarkdown; ioslides	
Friday, January 13, 2023	Lab 2: Rmarkdown	**Lab 1 Due**
Monday, January 16, 2023	MLK Holiday	
Wednesday, January 18, 2023	Rmarkdown & ggplot	
Friday, January 20, 2023	ggplot	
Monday, January 23, 2023	EDA & ggplot;	**Lab 2 Due**
Wednesday, January 25, 2023	EDA & ggplot	
Friday, January 27, 2023	Lab 3: EDA & ggplot	
Monday, January 30, 2023	Dates and Times	**Lab 3 Due**
Wednesday, February 1, 2023	Dates and Times	
Friday, February 3, 2023	Relational Data	
Monday, February 6, 2023	Review	**Lab 4 Due**
Wednesday, February 8, 2023	** Exam 1**	

Date	Topic	Due
Friday, February 10, 2023	Day of Celebration	
Monday, February 13, 2023	Tidy data from R4DS; Lab 5	
Wednesday, February 15, 2023	SEDSI	
Friday, February 17, 2023	SEDSI	
Monday, February 20, 2023	Functions; Lab 6	**Lab 5 Due**
Wednesday, February 22, 2023	Functions	
Friday, February 24, 2023	Iteration	
Monday, February 27, 2023	Regression	**Lab 6 Due**
Wednesday, March 1, 2023	Regression (Stepwise)	
Friday, March 3, 2023	** Exam 2**	** Exam 2**
Monday, March 6, 2023	Logistic regression	
Wednesday, March 8, 2023	Logistic regression	
Friday, March 10, 2023	Quiz	**QUIZ**
Monday, March 13, 2023	SPRING BREAK	
Wednesday, March 15, 2023	SPRING BREAK	
Friday, March 17, 2023	SPRING BREAK	**QUIZ**
Monday, March 20, 2023	Tree-based methods	
Wednesday, March 22, 2023	Tree-based methods	
Friday, March 24, 2023		**QUIZ**
Monday, March 27, 2023	Clusters	
Wednesday, March 29, 2023	Clusters	
Friday, March 31, 2023	LAUNCH PROJECT	**QUIZ**
Monday, April 3, 2023	INDEPENDENT PROJECT	
Wednesday, April 5, 2023	INDEPENDENT PROJECT	
Friday, April 7, 2023	Easter Holidays	
Monday, April 10, 2023	Easter Holidays	
Wednesday, April 12, 2023	INDEPENDENT PROJECT	
Friday, April 14, 2023	INDEPENDENT PROJECT	
Monday, April 17, 2023	INDEPENDENT PROJECT	
Wednesday, April 19, 2023	INDEPENDENT PROJECT	

Date	Topic	Due
Friday, April 21, 2023	INDEPENDENT PROJECT	
Monday, April 24, 2023	PRESENTATIONS	
Wednesday, April 26, 2023	PRESENTATIONS	
Friday, April 28, 2023	**LAST DAY OF CLASSES**	

Chapter 4

Lab 1 Excercises

Let's make sure we feel good about BADM 371 material.

All open notes/internet/R4DS/etc., **but all work must be your own**.

Use the starwars data (dplyr package) to answer/do:

1. How many individuals have no missing data?

2. Which variables have the most missing data?
3. Who is the tallest individual? Shortest?
4. How many homeworlds are there?
5. Which homeworld has the most individuals? Fewest?
6. What is the average # of individuals per homeworld?
7. Make a plot of all individuals with mass on the x axis and height on the y axis.
8. Put a best fit line on this plot.
9. Who is the biggest outlier in this data set?
10. Calculate BMI for all these individuals. What is the average BMI for all individuals?
11. Which homeworlds have the greatest percentage of individuals with BMI's greater than the average for each homework? Make sure you are clear about what you do with missing data.

Chapter 5

Lab 2 in Rmarkdown

5.1 Please submit this assignment as a well formatted, easy to read printed, hard-copy R Markdown file.

1. Choose some real data. Explain the data, why it's interesting and use some summary statistics (in line code) to describe this data.

2. Use a code chunk to explore it further – missing data by variable? Interesting stuff?

3. Explore your data further making sure you do use each of these:

- headers
- ordered lists
- links
- an image
- a table

4. Make multiple, correctly-sized plots on a single row (for example):
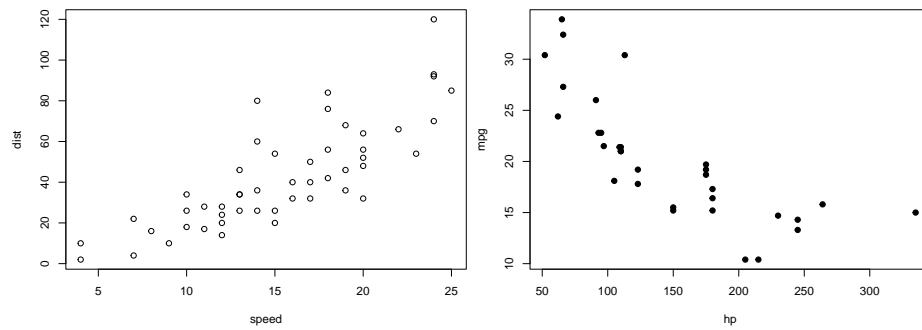
5. Using a well-formatted code chunk answer this question: Which home-worlds have the greatest percentage of individuals with BMI's greater than the average for each homework? Make sure you are clear about what you do with missing data.

6. Something else neat with your data that shows your Rmarkdown skills and why your dataset is of interest.