

DSCI351-351m-451: Class 09b MidTerm Observation

Profs: R. H. French, L. S. Bruckman, P. Leu, K. Davis, S. Cirlos

TAs: W. Oltjen, K. Hernandez, M. Li, M. Li, D. Colvin

30 October, 2022

Contents

9.3	Class Readings, Assignments, Syllabus Topics	1
9.3.1	Reading, Lab Exercises, SemProjects	1
9.3.2	Textbooks	1
9.3.3	Syllabus	2
9.3.4	For the DSCI 451 students they have an EDA SemProj to do	2
9.4	Results and Observations from Midterm	2
9.4.1	The Curve: DSCI is graded on a curve	4
9.4.2	MidTerm Grades	4
9.4.3	General Observations and Common Mistakes	4
9.4.4	Comments on Grading Philosophy for the Midterm	4
9.5	Links	6

9.3 Class Readings, Assignments, Syllabus Topics

9.3.1 Reading, Lab Exercises, SemProjects

- Readings:
 - For today: R4DS17-21, OIS6.3-6.4
 - For next class: OIS6.3-6.4
- Laboratory Exercises:
 - LE4 : has been turned in
 - LE5 : Given out today.
 - LE5 Due Thursday November 10th
- Office Hours: (Class Canvas Calendar for Zoom Link)
 - Wednesday @ 4:00 PM to 5:00 PM, Will Oltjen
 - Saturday @ 3:00 PM to 4:00 PM, Kristen Hernandez
 - **Office Hours are on Zoom, and recorded**
- Semester Projects
 - DSCI 451 Students Biweekly Update 1 Due
 - DSCI 451 Students
 - * Next **Report Out #1 is Due Friday September 30th**
 - All DSCI 351/351M/451 Students:
 - * **Peer Grading of Report Out #1 is Due October 11th, 2022**
 - Exams
 - * Final: Monday December 19, 2022, 12:00PM - 3:00PM, Nord 356 or remote

9.3.2 Textbooks

- [Peng: R Programming for Data Science](#)

- [Peng: Exploratory Data Analysis with R](#)
- [Open Intro Stats, v4](#)
- [Wickham: R for Data Science](#)
- [Hastie: Intro to Statistical Learning with R, 2nd Ed.](#)

Introduction to R and Data Science

- For R, Coding, Inferential Statistics
 - Peng: R Programming for Data Science
 - Peng: Exploratory Data Analysis with R

Textbooks for this class

- OIS = Diez, Barr, Çetinkaya-Runde: Open Intro Stat v4
- R4DS = Wickham, Grolemund: R for Data Science

Textbooks for DSCI353/353M/453, And in your Repo now

- ISLR = James, Witten, Hastie, Tibshirani: Intro to Statistical Learning with R 2nd Ed.
- ESL = Trevor Hastie, Tibshirani, Friedman: Elements of Statistical Learning
- DLwR = Chollet, Allaire: Deep Learning with R

Magazine Articles about Deep Learning

- DL1 to DL12 are “Deep Learning” articles in 3-readings/2-articles/

9.3.3 Syllabus

9.3.4 For the DSCI 451 students they have an EDA SemProj to do

- Some students didn’t submit their Report out *.mp4 zoom file
 - So we’ll circulate these as just pdfs for peer grading
 - * an extra point for feedback on these

9.4 Results and Observations from Midterm

- Make sure you are loading the packages you are using!
 - Sometimes packages can run in your console
 - But don’t run when you knit your PDF
 - And eventually, you don’t see outputs.
- The `library(tidyverse)` command
 - Is part of your code
 - It has to be in your .Rmd or .R file
- CHECK YOUR SUBMISSIONS
 - We are not responsible for chasing down
 - * Incorrect submissions
 - * Late work
 - * Un-uploaded PDFs
- Explain your answers in detail
 - ‘Yes’ or ‘No’ will not receive full points
 - Comment your code
 - * We can give more CREDIT IF you properly explain your thought process
 - * As opposed to just showing the code
- ?? and ? can be used to access the help documentation
 - This information can be useful in determining
 - What function is from each package

Day:Date	Foundation	Practicum	Reading	Due
w01a:Tu:8/30/22	ODS Tool Chain	R, Rstudio, Git		
w01b:Th:9/1/22	Setup ODS Tool Chain	Bash, Git, Slack, Agile	PRP4-33	LE1
w02a:Tu:9/6/22	Bash-Git-Knuth-Lit.Prog.	RIntroR	PRP35-64	451 Update1
w02b:Th:9/8/22	What is Data Science	OIS:Intro2R	OIS1,2	
w02Pr:Fr:9/9/22			PRP65-93	
w03a:Tu:9/13/22	Data Intro	Data Analytic Style	PRP94-116	LE2 LE1 Due
w03b:Th:9/15/22	Rand. Var. Normal Dist.	Git, Rmds, Loops	OIS4	
w04a:Tu:9/20/22	Tidy Check Explore	Tidy CapMinder	EDA1-31	
w04b:Th:9/22/22	Inference, DSCI Process	Other Distrib. 7 ways	R4DS1-3	LE3 LE2 Due
w04Pr:Fr:9/23/22			EDA32-58	451 Update2
w05a:Tu:9/27/22	OIS4 Rand. Var.	EDA of PET Degr.	OIS5	
w05b:Th:9/29/22	OIS5 Found. of Infer.	Multivar Corr. Plot	R4DS4-6	
w05Pr:Fr:9/30/22				451 RepOut1
w06a:Tu:10/4/22	Pred., Algorithm, Model		R4DS7-8	
w06b:Th:10/6/22	Summ. Stats & Vis.	Anscombe's Quartets	R4DS9-16	LE4 LE3 Due
w06Pr:Fr:10/7/22				451 Update3
w07a:Tu:10/11/22	Midterm Rev. Tidy Data	Correl Plots Summ Stats	OIS6.1-2	PeerRv1 Due
w07b:Th:10/13/22	HypoTest, Infer. Recap	Penguin EDA, Sampling		
w08a:Tu:10/18/22	MIDTERM	EXAM		
w08b:Th:10/20/22	Programming & Coding	Code Packaging		LE4 Due
w08Pr:Fr:10/21/22				451 Update4
Tu:10/24,25	CWRU	FALL BREAK	R4DS17-21	
w09b:Th:10/27/22	Cat. Inf. 1 & 2 propor.	Indep. Test, 2-way tables	OIS6.3-4	LE5
w09Pr:Fr:10/28/22				451 RepOut2
w10a:Tu:11/1/22	Goodness of Fit, χ^2 test	t-tests 1&2 means	OIS7.1-4	451 Update5
w10b:Th:11/3/22	Num. Infer, Cont. Tables	Stat. Power		
w10Pr:Fr:11/4/22				
w11a:Tu:11/8/22	Sample & Effect Size	Stat. Power GGmap	OIS8	PeerRv2 Due
w11b:Th:11/10/22	Inf. 4 Regr, Test & Train	Curse of Dimen.	ISLR1,2.1,2	LE6 LE5 Due
w12a:Tu:11/15/22	Lin. Regr. Part 1	Residuals	OIS9	
w12b:Th:11/17/22	Lin. Regr. Part 2	Regr. Diagnostics		
w12Pr:Fr:11/18/22				451 Update6
w13a:Tu:11/22/22	Mult. Lin. Regr.	Var. & Mod. Selec.,	ISLR3.1	LE7 LE6 due
w13b:Th:11/24/22	Log. Regr.	GIS Trends	ISLR3.2	
w13Pr:Fr:11/25/22				451 RepOut3
w14a:Tu:11/23/22	Classificat., Sup. Lrning	Caret, Broom 4 modeling	ISLR4.1-3	
Th,Fr:11/24,25	THANKSGIVING	Vacation		
w15a:Tu:11/29/22	Big Data Analytics	Clustering		PeerRv3 Due
w15b:Th:12/1/22		Dist. Comp., Hadoop		
w15SPr:Fr:12/2/22		Read Article by	Mirletz, 2015	
w16a:Tu:12/6/22	Final Exam Review			LE7 due
w15b:Th:12/8/22				
Friday 12/12	SemProj	Final Report		SemProj4 due
Monday 12/19	FINAL EXAM	12:00-3:00pm	Nord 356	or remote

Figure 1: DSCI351-351M-451 Syllabus

9.4.1 The Curve: DSCI is graded on a curve

- Results

DSCI 351/451

- A > 93
- B > 84
- C > 75
- D > 65

9.4.2 MidTerm Grades

- The midterm grades
 - Use all the graded submissions of each student
 - * If an assignment or activity
 - has not been graded
 - has not been submitted
 - * It wasn't counted
 - So these grades didn't have "assumed zero points"

9.4.3 General Observations and Common Mistakes

- ALL SUBMISSIONS should be DOUBLE CHECKED for proper submission quality
 - We cannot give points if we don't see reasonable outputs
 - This is applicable for LEs and exams
 - It is the responsibility of the student
 - * to verify that they are submitting the correct file
 - Explain your answers in detail
 - * 'Yes' or 'No' will not receive full points
 - * Comment your code
 - We can give more CREDIT IF you properly explain your thought process
 - As opposed to just showing the code
 - Some questions can be answered by knowing how certain functions work
 - * `?function-name` and `??function-name` can be used
 - to access the help documentation
 - * This information can be useful
 - in determining which function is from a certain package
 - Some students have spent a lot of time on some questions
 - * And couldn't work on the remaining ones
 - * Don't spend too much time on one question
 - * Time management is always important!

9.4.4 Comments on Grading Philosophy for the Midterm

- Deductions where made for
 - improper solutions
 - and missing information.
- Failure to library in the appropriate packages
 - resulted in a point deduction,
 - with a maximum deduction of 0.5.
- Additional points were deducted
 - for not obtaining the output.
- Solutions that
 - did not library the appropriate packages

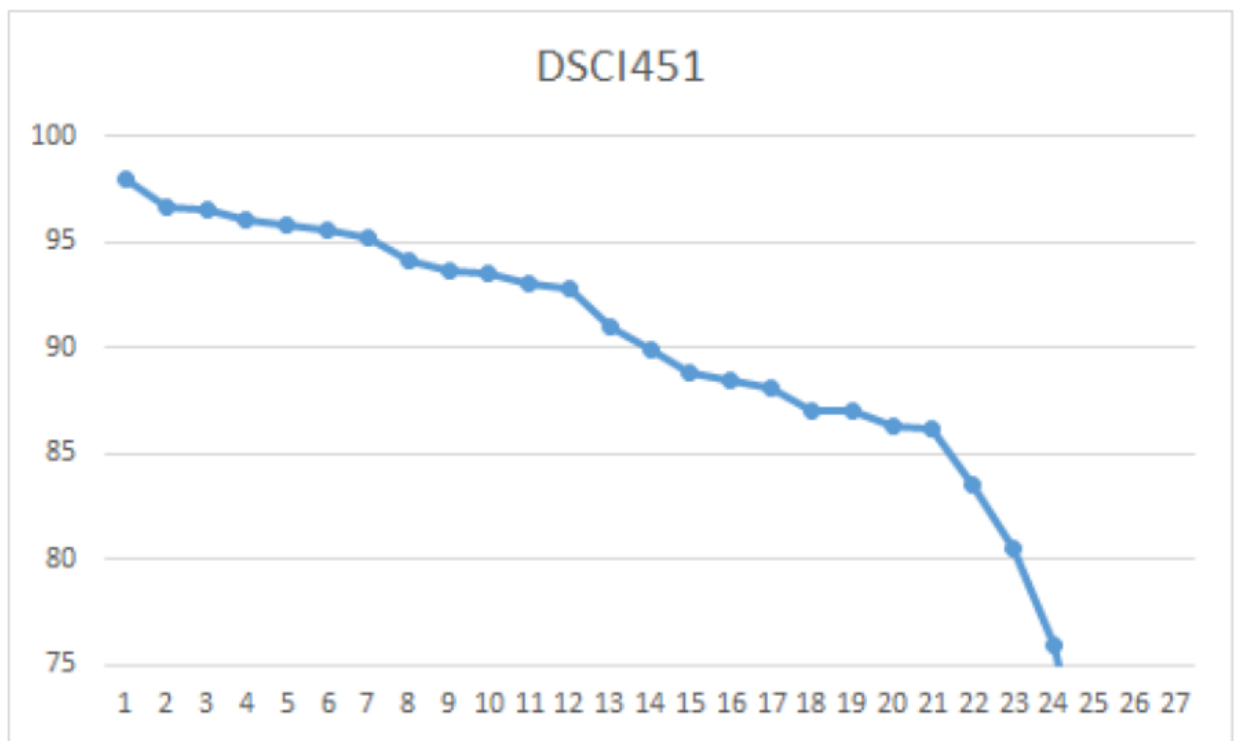
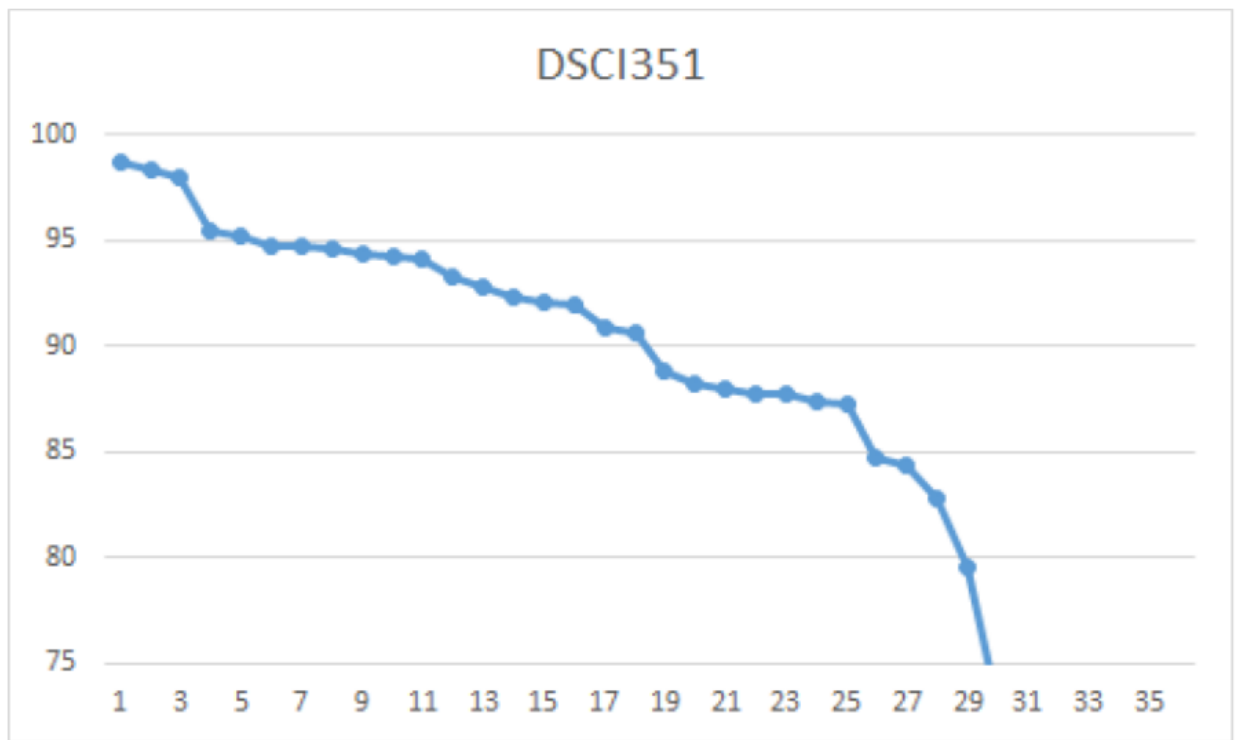


Figure 2: Midterm Grades

- and were incorrect received no credit.
- Partial credit was assigned as needed,
 - especially when a solution was properly documented with comments.

9.5 Links