

TEACHING A LARGE INTRODUCTORY STATISTICS COURSE: MAKING IT WORK!

JSM 2016
Section on Statistical Education

Chair: Roger Woodard ~ North Carolina State University

Panelists

- Megan Mocko ~ University of Florida
- Stacey Hancock ~ University of California at Irvine
- Christopher Wild ~ University of Auckland
- Brenda Gunderson ~ University of Michigan

Megan Mocko ~ University of Florida

- 49,785 students
- 32,008 undergraduate students
- 16,272 graduate students

- Historically a residential campus but growing online program since 2014

Course

Megan Mocko ~ University of Florida

Fall Offerings	Spring Offerings	Summer Offerings
HYBRID: 46 sections of 40 students – 1920 students total (live/online lectures with in class labs)	HYBRID: 47 sections of 40 students – 1960 students total (live/online lectures with in class labs)	In Class Only: 150 – 250 students Live Lecture
Completely Online: 100 UF Online students	Completely Online: 100 UF Online students	Completely Online: 125 students X 2 (100 students UF residential and 20-30 UF Online)
25 student section for students with registered learning disabilities	25 student section for students with registered learning disabilities	

Staff for Hybrid

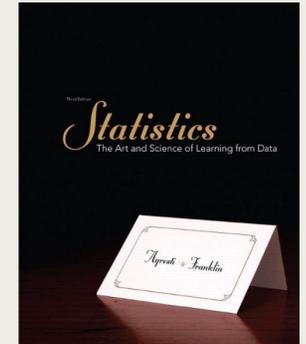
Megan Mocko ~ University of Florida

Course Coordinator – student centered (1 FTE)	Lab Coordinator – TA centered (.5 FTE)	Teaching Assistants (16-18 full time)
Responsible for weekly quizzes	Responsible for labs, lab materials, technology issues	Teach 3 -4 labs per week
Writes, proofs and grades exams	Trains TAs	Holds 2 -3 tutoring hours (For a total of over 40 tutoring hours available for the students.)
Responds to most students emails	Handles any student related problems in the lab	Grades short answer labs
Submits final grades	Coordinates grading of labs	Proctor Exams

Textbook

Megan Mocko ~ University of Florida

- *Statistics: The Art and Science of Learning from Data*
- By Alan Agresti and Christine Franklin
- Used 1st, 2nd, and 3rd editions
- Will use 4th edition in the Fall



Stacey Hancock ~ University of California at Irvine

- ~ 32,000 students
- ~ 25,000 undergraduate students
- ~ 7,000 graduate students



Founded in 1965; one of ten UC campuses

Course

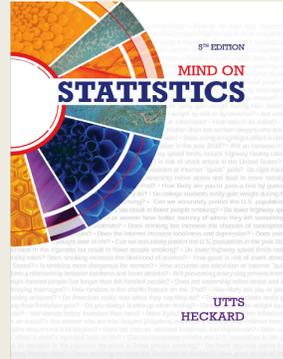
Stacey Hancock ~ University of California at Irvine

- Stats 7: Basic Statistics
- 3 sections per quarter; 220 students per section
- Lecture meets three times per week for 50 minutes
- GTAs run weekly 50 minute discussion sections of 55 students each
- Service course for wide range of undergraduate students: e.g., biological sciences, political science, psychology, business, drama, art

Materials Stacey Hancock ~ University of California at Irvine

Textbook:

- *Mind on Statistics* (5th ed.) by Jessica Utts and Robert Heckard



Audience response system:

- iClicker remotes
- UCI classrooms have base stations
- students use remotes for several courses

Chris Wild ~ University of Auckland, New Zealand

- 40,000 students
- 29,000 undergraduate students
- 11,000 graduate students
- Residential campus

Course Chris Wild ~ University of Auckland

- Stats 10x
- 5,000 students/year (~70% of each cohort)
 - 1st Semester: 2,500 students
 - Class sizes (Lecture Rooms) 300-600
- Reputation for **best organized** and one of the best-taught courses on campus
 - (national teaching award in 2003 helped establish rep)
 - Reputation for high levels of support
 - **All teachers rate very highly** with students
- Big component of departmental budget
 - if this course sneezes the Dept catches cold
- Teaching team are “professionals”
 - If you are not an excellent teacher you won't be let near this course

Course Chris Wild ~ University of Auckland

- Team teaching for over 25 years
- Uses a **single set of team-developed materials**.
 - Perfect consistency across sections
 - Can focus on innovation and “performance” (not reinventing wheels)
- Weekly team meeting looking back and looking forward
- Strong all-one-team culture built and preserved over 25 years
 - Team gets a lot of fun from working together
 - (core troops out to morning coffee together every morning)
- Very welcoming to newcomers and visitors
- Aim for steady improvement rather than massive leaps (guard against unpredictable effects on numbers)
- Efficiencies gained here allow smaller teaching loads esp to research academics

Course

Chris Wild ~ University of Auckland

■ FutureLearn MOOC: *Data to Insight*

(Another experience to draw on)

- 24 hours' online study materials in 111 "steps":
- 41 videos, 16 exercises,
- 16 quizzes, 8 tests,
- 10+ discussions
 - and blog-discussions on almost every page
- 30+ articles & papers

WEEK 3: RELATIONSHIPS



Introduction

The most useful insights from data usually come from understanding relationships and important changes in them. We will talk about relationships between variables and how to read and interpret the graphs that reveal them.

3.1

INTRODUCTION TO RELATIONSHIPS VIDEO (02:52)

Personnel

Chris Wild ~ University of Auckland

- Lecturers (~5 in 1st semester)
- Administrator
- Help-room tutors
- Computer Lab tutors
- Student markers (large numbers)

Materials

Chris Wild ~ University of Auckland

- Workbook (updated yearly)
- Lectures – all recorded live and videos posted online
- Canvas site using most features
- Assignments & Lab sheets

Brenda Gunderson ~ University of Michigan



- ~ 45,000 students
- ~ 29,000 undergraduate students
- ~ 16,000 graduate and professional students
- Residential campus but growing online programs and learning options:
Digital Education & Innovation (DEI),
Coursera, edX, and more

Course

Brenda Gunderson ~ University of Michigan

- **Stats 250: Introduction to Statistics and Data Analysis**
Covers **descriptive statistics to hypothesis testing**
- Primarily Freshmen/Sophomores
- Psych, Econ, Business, PolySci

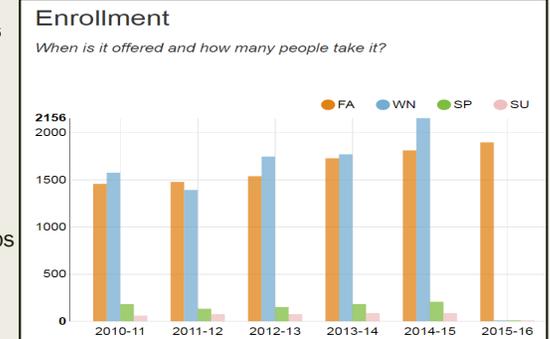
Freshman	13%
Sophomore	45%
Junior	28%
Senior	11%
Other	1%

Psychology BA	22%
Neuroscience BS	13%
Biopsych, Cognit & Neurosci BS	12%
Economics BA	11%
Business Administration BBA	10%
Political Science BA	10%
International Studies BA	8%
English BA	5%
Communication BA	4%
Environment BS	4%

Course

Brenda Gunderson ~ University of Michigan

- **LARGE** = ~ 2100 students
- **Weekly:**
3 hours of lecture,
1.5 hour computer lab
- **Coordinated**
across 6 lectures
of 150 to 400 students,
~ 64 GSI led computer labs
(30 to 35 students).



Materials

Brenda Gunderson ~ University of Michigan

Required:

- Stats 250 **Coursepack**: Lecture Notes and Lab workbook
(Dollar Bill or print/use from Canvas or Open.Michigan)
- Online **homework** tool (Course.Work)
- **iClicker**
- **Calculator**

(Highly) Recommended:

- **ECoach**
- Text **Mind on Statistics** (MOS) 5e

Some "Prepared" Questions to start us out

- If someone were thrust into the position of teaching a 500+ student class what practical considerations would they need to understand?
 - *What course structures have you found to be particularly helpful?*
 - *What pitfalls should you avoid in such a large class?*
- How to handle assessments?
- How to prepare teaching assistants?
- How to incorporate active learning?
- How to get to know students?
- How to address academic integrity?

Some “Prepared” Questions to start us out

- Are any quality checks done for the course?
If so, describe how it works?
- What types of technology do you use?
- What is special about trying to apply the newly revised GAISE guidelines in large classes?
- What is special about applying new environments (face-to-face, MOOC, hybrid, flipped, etc.) in large classes?
- How team teaching can improve the teaching of large courses?
- What does teaching a large number of students inform us about teaching in general? Are there truths to be learned that apply to small classes as well?”

Megan Mocko ~ University of Florida

If someone were thrust into the position of teaching 500+ students what practical considerations would they need to understand? 

- Detailed Syllabus
 - 16 pages long
- Be consistent!
- Double, Triple check everything!
- Be Organized – Tell the students where we are going and how we are going to get there.

Weekly Reminders

Weekly Reminders for Monday, November 16th to Friday, November 20th

Lectures: MWF : Available: 9th period in NRN 137, 7th period in Norman G520 or online

[Quiz 9](#)

- Offered: Monday, November 16nd 8am until Thursday, Nov. 19th 11:59pm
- Three Attempts, highest counts
- Covers Material: Chapter 9 Section 3 – Chapter 9 Section 5, pages 86 - 95 in the lab workbook

Lab: Lab 9: Activity 18 Monday - Friday

- You must attend the section for which you are registered.
- Be on time. You have a 5 minute grace period.
- If one of your lab scores was uploaded incorrectly to Canvas, please talk to your ta in lab. You must bring a copy of the lab worksheet in question. All grades changes must be done by Wed., Dec. 9th. 5pm (the last day of classes).

Weekly Reminders (Part Two)

MM

Suggested HW problems

Ch 9 Statistical Inference: Significance Tests About Hypotheses	
9.1 Steps for Performing a Significance Test	9.1 9.2 9.3 9.4 9.8
9.2 Significance tests About Proportions	9.9 9.11 9.12 9.15 9.17 9.19 9.20 9.21 9.23

Supplemental Videos

Chapter 9 Section 2 (Finding the p-value for a Significance Test for the Population Proportion)

Grades posted in Canvas

Quiz	EC	1	2	3	4	5	6	7				
Lab		1	2	3	4	5	6					
Exam		1										

What course structures have you found to be particularly helpful?

MM

- Instructions in writing
- Predictable pattern each week
- Posting extra help videos



What pitfalls should you avoid in such a large class?

MM

- Be careful what you ask for.
 - “Please let me know if you think of a good memory aid to help remember the difference between Type I and Type II error definitions. I will share them with the class.” **SUCCESS!**
 - “So far, I have received zero questions for the Q and A tomorrow, if I don't receive any there won't be a review.” **Success, yes, manageable, no**

What pitfalls should you avoid in such a large class?

MM

- It helps to be little bit paranoid.
 - “Tell students for two weeks prior to the exam that there will be an extra set of questions on the exam. Explain on the exam that there is an extra set of questions on the exam. 30+ students stopped at “33” the typical number of questions on our exams.” **Not good!**
 - “Do all of the above and add a question on the exam that asks how many questions are on the exam” **Success!!**

How to Handle Assessments?

MM

- **3 multiple choice exams**
 - Check and Re-Check Exams before Printing
 - Have five other people complete and check the exam as well
 - Across 15 - 18 rooms around campus
 - All graduate students help proctor
- **10 (weekly) multiple choice quizzes**
- **10 (weekly) labs**
 - Labs include short answer questions and interpretation statements
 - Active Learning Environments

Exam Nights

MM

- Meet with all graduate students 1.5 hours prior to exam.
- Eat pizza together before going out to exam rooms.
- Explain the procedure verbally and in writing.
- Have a policy for late students to the exam.
- Extra calculators, pencils, etc. for each room.



How to Prepare Teaching Assistants?

MM

- Two hour Training Session at the beginning of the Semester
- Ten one hour training sessions before each lab
- One hour training about entering grades for new graduate students
- One hour "Best Practices" activity in the Spring

"Scripts" for Lab

MM

- **TALK:** *The purpose of today's lab is to talk about calculating and interpreting the confidence interval for the population mean. Suppose that I am interested in determining the average numbers of hours that students spend sleeping. So, I decided to randomly select 100 students at UF and ask them how many hours they slept last night. Out of that sample, the sample average number of hours slept was 7.1.*
- **ASK:** *Does this mean that all UF students sleep 7.1 hours a night on average? No.*
- **TALK:** *We want to determine what the mean number of hours slept for the population of UF students. To do this, we can make a confidence interval for the population mean. This will give us a region in which the population mean is likely to be.*
- **TALK:** *The general format of the confidence is estimator \pm (z or t) stderr .*
- **WRITE:** estimator \pm (z or t) std error }

“Scripts” for lab (con’t)



Part 3 and Part 4 –

- **WALK:** Talk to the groups and see if they have any questions.
- **SHOW:** Now, do this in Minitab. Go to Stat > Basic Statistics > One Sample t. Select the value of C1 as “Samples in Columns” value. Select the Graph options and choose “Boxplot”. The 95% confidence interval is the default setting for Minitab, but if you want to change it, go to Options and enter a new value.
- **TALK:** Now, work through part 5 of the lab worksheet.
- **WALK:** Talk to the groups and see if they have any questions.
- **WAIT:** After they have been working on it for about five minutes.
- **TALK:** There are two statements that are true.
- **DISCUSS:** Go over the answers.

Slides (optional)



2

Suppose that I was interested in finding out what proportion of UF students planned on going on a vacation over Spring Break. I randomly asked 100 students and asked them if they plan on going on a vacation.

Suppose that my sample mean was 0.30. If you took a random sample of 100 UF students, would you get the same value?
No, not likely

What about if your friend took another random sample of 100 students as well, would it be the same?
No, not likely

In order to study this, we need to study the sampling distribution of the sample proportion.

What is the distribution?
Normal, under certain conditions

What are the conditions?
- Both np and n(1-p) are greater than or equal to 15.
- Original Population is Normal.

What is the mean?
p

What is the standard deviation?
 $\sqrt{\frac{p(1-p)}{n}}$

TA evaluation visits



- Once per year.
- Rated on following classroom procedure.
- Engaging the class.
- Rotating between groups of students.
- Explanation of the material.

STA 2023 TA Classroom Visit Form	TA Visited _____ Date: _____ Evaluator: _____
	COMMENTS
Lab starts on time.	
Attendance taken and old worksheets returned during first 5 minutes.	
Course announcements displayed and any important ones emphasized.	
Late students were handled appropriately.	
Lab purpose and background material explained appropriately.	
TA is prepared to teach the activity.	

What Types of Technology do You Use?

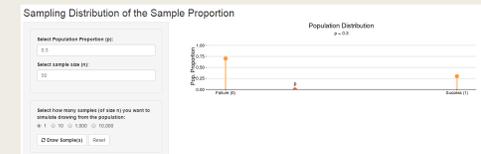
- Canvas Course Management System
– Email based on grade of assignment



- Minitab



- Applets



What is special about trying to apply the newly revised GAISE guidelines in large classes?

MM

- Change is very hard!
- Two new emphases
 - Give students experience with multivariable thinking.
 - I added an example to the lecture last year.
 - This Fall, we updated the example and will re-enforce it with quiz and exam questions.
 - Teach statistics as an investigative process of problem-solving and decision-making.
 - Challenge the status quo. This will be more difficult.

Also, what does teaching a large number of students inform us about teaching in general?

MM

- Format of assessments (electronic versus paper)
- Multiple Attempts for Assignments (actually practice more)
- Feedback (although I can give more feedback on written returned paper assignments, the automatic feedback for large classes is helpful here as well.)

Stacey Hancock ~ University of California, Irvine

How to get to know students?

SH

- Beginning quarter survey
- Office hour log
- Active learning – move around the room during clicker questions or group work

How to address academic integrity?

SH

- Detailed expectations and definitions in syllabus
 - what is academic dishonesty?
 - what are the consequences?
 - why are we here?
 - focus on respect, community, pride, honesty
- Open dialogue at the beginning of the quarter
 - clicker questions

How to address academic integrity?

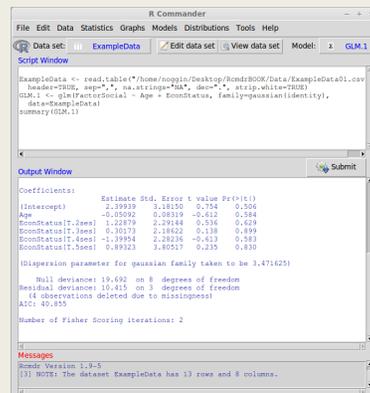
SH

- Different versions of exams
 - No take-home exams
- Randomized seating chart
- Check IDs
- Homework is only a small portion of the grade

What types of technology do you use?

SH

- R Commander (Rcmdr)
 - Free and open source
 - R library
 - GUI
- iClickers



```
ExampleData <- read.table("/home/soyila/Desktop/SondrBOOK/Data/ExampleData01.csv")
header=c("ID", "Age", "Sex", "Status", "FamilyIncome")
write.csv(header, "header.csv", fileEncoding="UTF-8", as.is=T, append=FALSE)
GLM1 <- glm(FractionCorrect ~ Age + SonStatus, family=gaussian(link=identity),
data=ExampleData)
summary(GLM1)
```

Output Window

```
Coefficients:
(Intercept)      2.39939      3.18155      0.752      0.506
Age             -0.05092      0.08319     -0.412      0.584
SonStatus[T.2sec] 1.22879      2.29144      0.534      0.629
SonStatus[T.3sec] 0.30173      2.18862      0.138      0.899
SonStatus[T.4sec] -1.39954      2.28226     -0.419      0.583
SonStatus[T.5sec] 0.89223      3.80517      0.235      0.830

Dispersion parameter for gaussian family taken to be 3.471625

Null deviance: 19.622  on 8  degrees of freedom
Residual deviance: 10.615  on 3  degrees of freedom
(4 observations deleted due to missingness)
AIC: 40.855

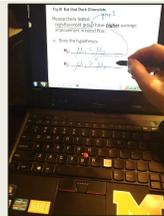
Number of Fisher Scoring iterations: 2

Messages
Rcmdr Version 1.9-5
[1] NOTE: The dataset ExampleData has 13 rows and 8 columns.
```

Brenda Gunderson ~ University of Michigan

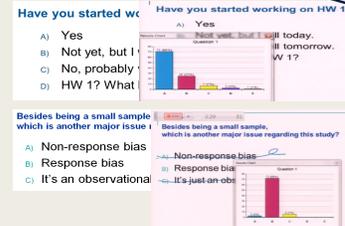
How to Incorporate Active Learning? BG

- **Coursepack:** Lecture Notes and Labs
 - Initially incomplete, students are co-authors work with instructors to complete and create own study guide
 - Instructors use doc cam or tablet to guide



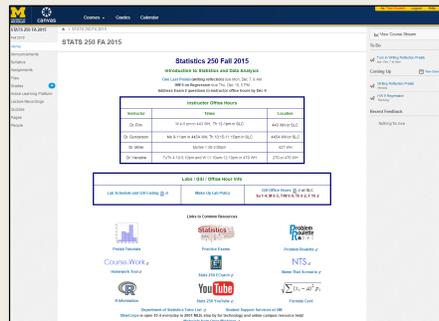
How to Incorporate Active Learning? BG

- **iClickers:** Lectures and Labs
 - Offers Low Stakes Assessments in class
 - Provides useful, actionable data
 - Incentive = cartwheel



What Types of Technology do You Use? BG

- **Canvas Course Management System**
 - Homepage = one stop shopping for resources
 - Announcements
 - Groups
 - Virtual Office Hours
 - Admin GSI helps to maintain site



What Types of Technology do You Use? BG

- **R / R Commander** (since Spring 2015)
 - Drop-In R Install Sessions
 - R Help Email
 - Prelab Videos each week
- **HyFlex Lecture and Make Up Lab Option**
Dr. Jackie Miller (jabmille@umich.edu)

Histograms and Boxplots Prelab

After completing this prelab, you will be able to:

- Open data sets in R Commander
- Create histograms using R
- Create boxplots and side-by-side boxplots using R
- Use R to calculate descriptive statistics for quantitative variables
- Save your graphs as PDF

Video:

Note: This library video makes use of the "cars1" dataset. This is one of the sample data sets included with R, and is also available on Canvas in the Data Sets Folder in Files. Your short prelab assignment has you work with a different data set (StudyHours.Rcmdr), also available on Canvas.

Prelab Assignment (to be submitted via Canvas Assignments):

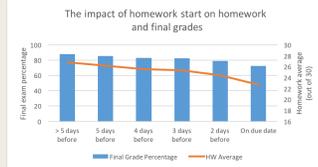
1. Download the "StudyHours.Rcmdr" R data set
2. Use R and create a histogram of the "hours per week studying" (as reported by student) (StudyHrs) variable. Think about it: how would you describe the histogram in terms of shape, center, spread, and any outliers?
3. Add an appropriate title for your graph and include attribution (By your name)
4. Save your graph (with title and named as a .png) and upload it
5. In the comments, provide a sentence that describes the distribution of study hours per week. Your sentence should include some aspect of shape, center, spread, and if any outliers.

Ready for the assignment? [Click here to go to the assignment for this prelab activity.](#)

What Types of Technology do You Use? BG

Online HW Tool = Course.Work

- Open/close at same time
- GSI graded with rubric and feedback given
- No "lost" HW
- Instructors can 'see' any HW
- Lots of data



Question: Before conducting the hypothesis test of interest, you need to determine if all conditions are met. Clearly state the additional required to be able to perform the one sample t test.

Answer: We need to assume that the population of all the drill evacuation times (in seconds) has a normal distribution.

Rubric: 0.5 point if user that population needs to have a normal distribution
0.5 point if missing context

Student answer: 0.8 0.5 0.0 1.0 pts

Student Feedback: The additional condition that is required is that the POPULATION of all the drill evacuation times needs to have a normal distribution. Also missing context of the problem (population of fire drill evacuation times)

What Types of Technology do You Use? BG

- Online Applets and GSI created Shiny Apps
- Home-Grown Practice Resources
 - Problem Roulette and Name That Scenario

Question: Select the Type of Sampling Problem: (1) XBAR, Analysis of Variance (2) t, Regression (3) t

Answer: (1) XBAR, Analysis of Variance (2) t, Regression (3) t

Feedback: Please state the Clarity of this problem

Question 1/10: A college administrator in a community high school uses one of four modes of transportation to get to school each day. She thought that half of all students take the bus, and 20% walk, and remaining students are equally split between driving and biking. A chi-square goodness of fit test is used to determine if the data provide evidence to reject the null hypothesis. You choose a random sample instead of a full survey to try and test the hypothesis.

Correct! 1

Correct! 2 The answer is CHI-SQUARE GOODNESS OF FIT TEST.

What Types of Technology do You Use? BG

ECoach ~ personalized tool

Too much material? There are a lot of resources available in Stats 250. I would not advise for students to do them all (i.e. the supplemental homework, textbook problems, etc.). Pick and choose what works for you. Establish this pattern of learning early.

Brenda, here is your GTD (Get Things Done) List for Dec 6-12:

- Check the boxes when you have completed an item. Click Save at the bottom or print a pdf version of this list below.
- After checking that you identified the last **Writing Reflections** **Prereqs**, **celebrate** that there are no more **Prereq assignments**.
- Take resources and check your **gradbook entries**. Apart from get few HW scores, are all entries correct? Report any issues in **gradbook** or **check grades** on Dec 7-8.
- Last Labs** **HW** **Due: 7-8** or **Lab 11 on More Regression** - calculator formulas and tables, and **Mid** **Mid** **Due: 14** as cases and that **Monday**. Watch for announcements about some optional lab reviews to be held from **Tue, Dec 14-15** by a few **CSIs**.
- Use **CourseWork** to work on **required HW** about regression - due **Tue, Dec 10 at 5 PM**. Note: CS1 requires 8 to perform a regression to make this homework grade depends on the homework with the **homework** **lab** **due**. The **Recommended HW** **is** **on** **your** **gradbook**.
- Exam 2 follow-up** - look over that Exam 2 one more time and see if you have any questions you need to address to Dr. Venable, Dr. Park, Dr. Miller, or Dr. Gunderson during office hours by **Wed, Dec 9**.
- Print out and read over the **HW 30 CS1** **required Test Questions** that will be available on **Tue, Dec 8**. These last required questions are available **Tue, Dec 8 at 5 PM**.

What is ECoach? ECoach is a tool to help you do your best in Stats 250. Think of ECoach as your own personal coach. ECoach will offer you strategies and insights.

- The best tools to use to study
- Study hints for exams
- A grade calculator tailored to Stats 250
- Evidence-based tools to help boost your exam performance

Where you are now and where to go from here

- Highly recommended HW
- Prereq assignments
- Prereq assignments

How to Handle Assessments? BG

- 10 (weekly) online HW required (also set of recommended questions)
- 10-12 (weekly) labs
 - Prelab Assignment (due before lab)
 - Warm Up Questions
 - Group ILP (in-lab project often using R/Rcmdr)
 - Cool Down Questions
 - Ticket (out the door)
- 2 semester exams and 1 comprehensive final
 - Paper Exams (not all MC)
 - Exam 101 session to train grad students (monitor Yak Yak exam day/night)
 - Group Grading Session (rubrics)
 - Future = Scanning in graded exams

Page Number	Problem Points	Points	Points Received	Grade
1	1.82	15		
2	2.83	3	22	
4	4.85	4	21	

Where you are now and where to go from here

- Highly recommended HW
- Prereq assignments
- Prereq assignments

that stats exam tho

I need it.

It had a great personality!

what did everyone get for the turkey question on stats?

I put down "I don't know"

8 something I think

Oh sorry

Two Methods for Final Grade: differ in weighting of exams