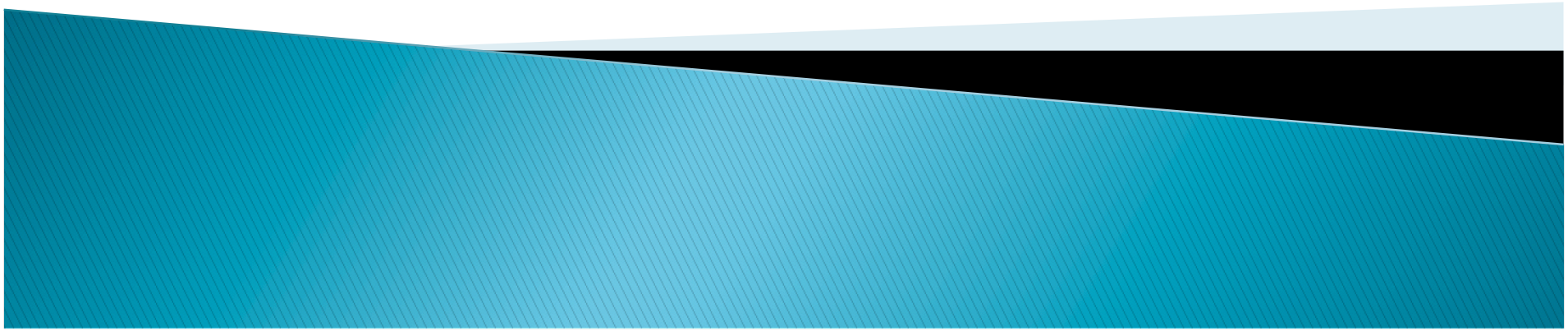
Real Data & Service Learning Projects in Statistics

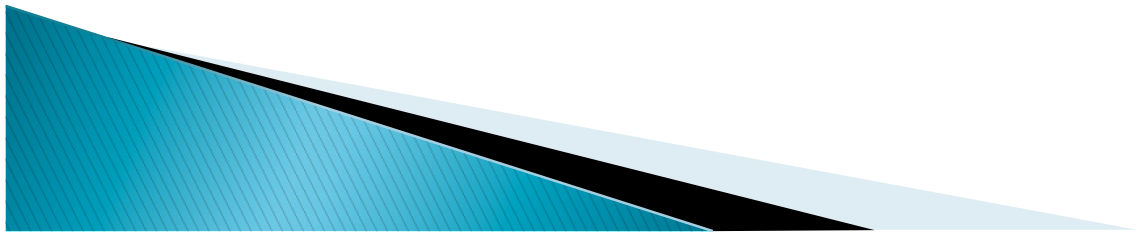
Brad Bailey & Robb Sinn
North Georgia College & State University

Joint Mathematics Meetings
Special Session on Service Learning in Mathematics
January 6, 2011



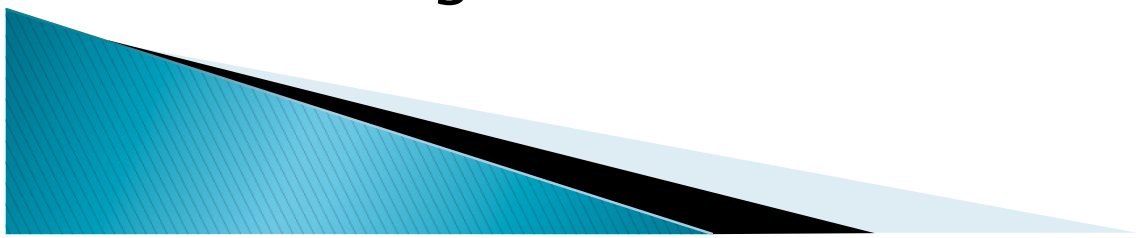
Talk Outline

- Description of the classes (Math 2400H, 3350 & 3390)
- Description of project(s) for
  - Introductory Honors Statistics – Math 2400H
  - Probability and Statistics I – Math 3350
  - Undergraduate Research in Statistics – Math 3390



Some Caveats

- The service-learning in our statistics courses has a **primary focus on learning**, and a secondary focus on service.
  - What does this mean?
    - We run statistical tests that will best align with course goals, not necessarily the tests we would run as researchers working with the same data.
    - We sometimes fail to answer or even address one or more research questions our partners wish to have analyzed.
- Learning outcomes from these projects spill out far beyond the confines of the syllabus and course goals.



The classes.

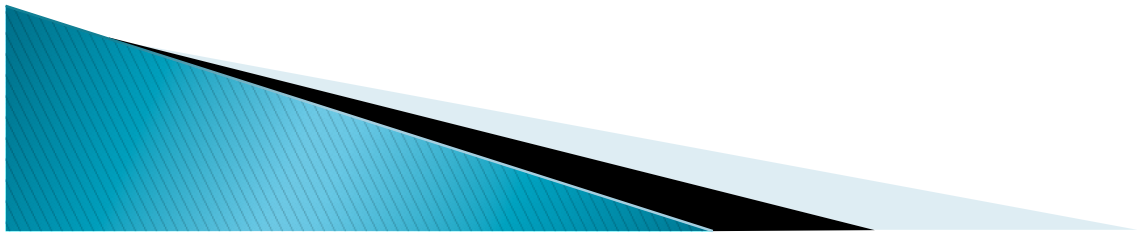
- MATH 2400H
  - An introductory level statistics course for non-technical majors.
  - Populated with business, science & social science majors.
- MATH 3350
  - A first course in advanced probability and statistics.
  - Populated with math, math secondary, and science majors.
- MATH 3390
  - Required prerequisite: strong performance in 2400H or 3350.
    - Enrollment in 3390 is by invitation only.
  - Curriculum topics dictated by needs of research project(s).

MATH 2400H

- Analyzed NSSE (National Survey of Student Engagement) Data for our institution provided by our Office of Institutional Effectiveness.
  - Contained demographic questions: Gender, race, religion, but also Major, Greek, Athlete & Parents level of education.
  - Questions about satisfaction with overall experience, quality of academic advisement, quality of relationship with faculty, administration and other students.
    - These questions were on a five point Likert Scale.
  - Questions about number of hours spent studying, socializing, working.
- The data was “scrubbed” of any FERPA sensitive information.

MATH 2400H

- Student Research Questions:
  - Are any demographics more likely to give higher rating to “Quality of overall experience”?
  - Are any demographics more likely to give higher rating to “Quality of advisement/relationship with faculty, administration, other students”?
  - Are any demographics more likely to spend more time socializing, studying, working, etc?

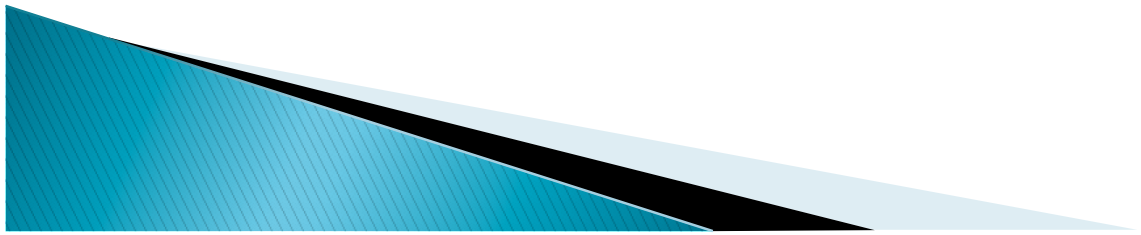


MATH 2400H

- The students used the following methods:
  - Chi squared tests
    - Rating (1-5) versus demographic variable.
  - Z test for population proportion
    - Regard 4&5 (Good & Excellent) as Success, and test for differences in proportions of Successes between Greeks/Non Greeks, Athletes/Non Athletes, Freshman/Seniors.
  - T tests for population means
    - Regard the rating as quantitative data.
  - One way ANOVA
    - Especially useful for comparing ratings for majors.

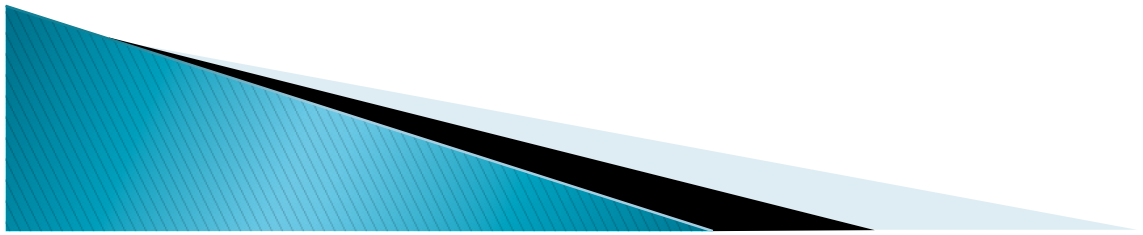
MATH 2400H In Fall 2010

- Analyzed data on student motivation provided by education researchers on over 200 students.
  - Contains responses to 35 question Motivation survey and data on class, course, gender and course grade.
  - Performing many of the same types of analyses as previous project.
  - Student names removed, course changed to single letter code.



MATH 2400H

- The students learned to run hypothesis tests in SPSS.
- They arranged the data in Excel.
- Also used TI Calculators.

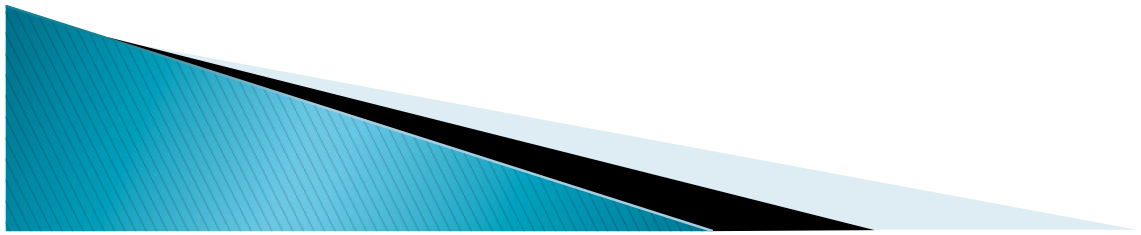


Results

- First, good news that disappointed the students
  - There were very few significant differences between the demographics/majors.
  - The students view of hypothesis testing led them to hope to find differences.
- Second, the students said that they appreciated the experience of working with real data and addressing real world questions.

MATH 3350

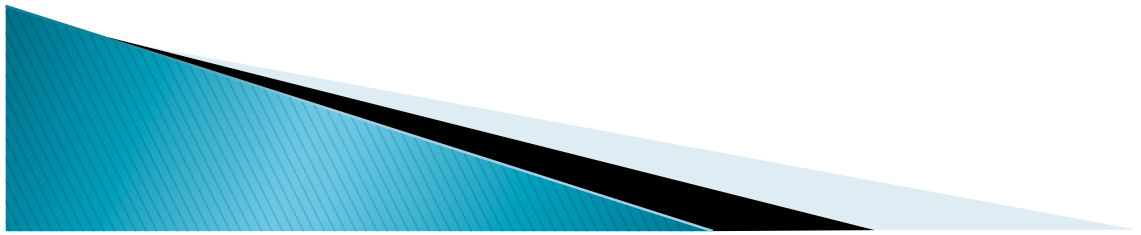
- INTRO
  - North Georgia's new student orientation
    - Five summer sessions hosted 775 freshmen-to-be
- Survey
  - 30 5-point Likert scale questions
  - 10 demographics variables
- Consulting Firm
  - Prob-Stats 3350 formed 6 consulting teams
  - Entered, analyzed and presented results



MATH 3350

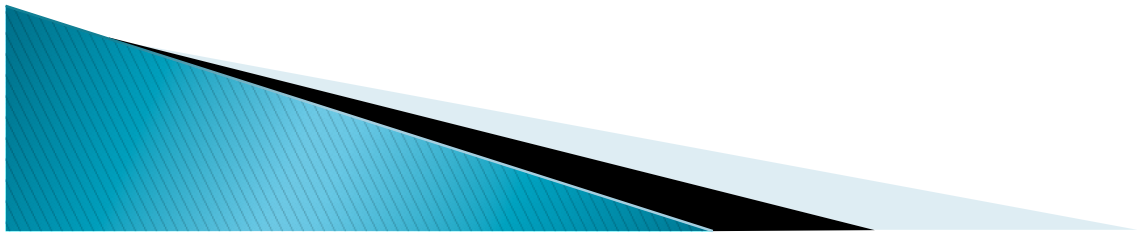
- Ordinal Scale Development
  - Researcher Choice
    - Questions were grouped (by me) into 6 constructs
      - Regression and ANOVA now possible
      - Chi-Square used for categorical comparisons
- Scales Measured Effectiveness in 6 areas:

1. Advising
2. Academic Information
3. Orientation Activities
4. Expectations
5. Pre-INTRO Information/web site
6. Did student feel welcome? Did INTRO confirm college choice?



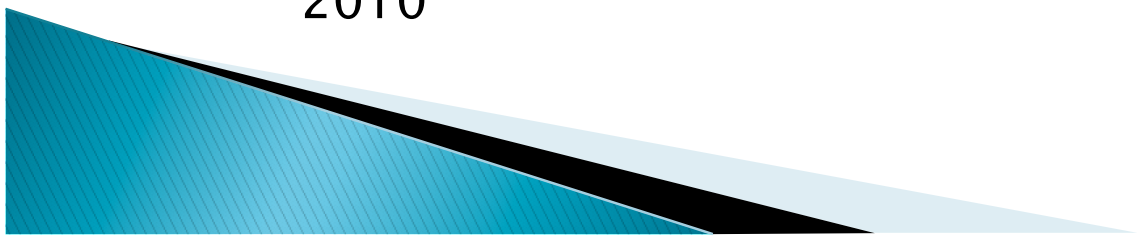
Statistical Tasks

- ANOVA's
  - Compared effectiveness of orientation across demographic categories
  - Compared effectiveness across the five different INTRO sessions
- Multiple Hierarchical Regression
  - What were the most significant predictors of desired outcomes such as "felt welcome at NGCSU"?
  - How can we improve academic advisement and course registration for all freshmen?



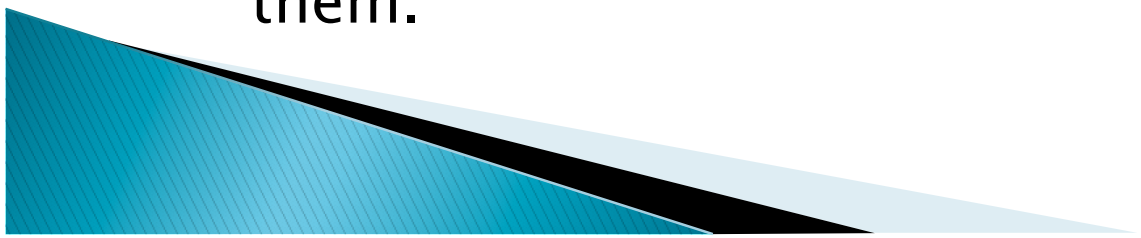
MATH 3350

- Outcomes
  - 17 guests attended our presentation
    - Vice President for Student Affairs
    - Staff Members from Student Affairs and:
      - Recreational Sports
      - Advising Center
      - Residence Life
    - Assistant Vice President for Academic Affairs
    - Our Dean (Dr. Bodri) and our Dept. Head (Dr. Cruthirds)
  - Student work was praised and highly valued
    - Dean Bodri “ordered” Brad and I to proceed with an undergraduate statistics research seminar course for Spring 2010



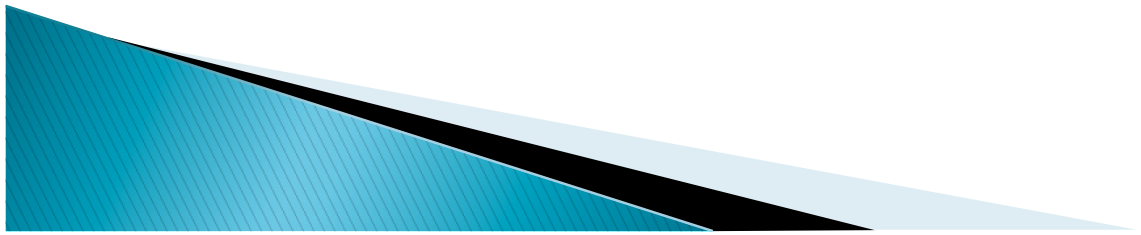
MATH 3390

- Course is driven by research project.
  - Topics have included EAF (Exploratory Factor Analysis), scale reliability and consistency, logistic regression and multiple regression with exploratory, step-wise, hierarchical modeling procedures.
- Course is listed as a 1-hour seminar.
  - Small class size & more informal; like a research group.
  - Students are involved in researching the types of analyses needed & are responsible for performing them.



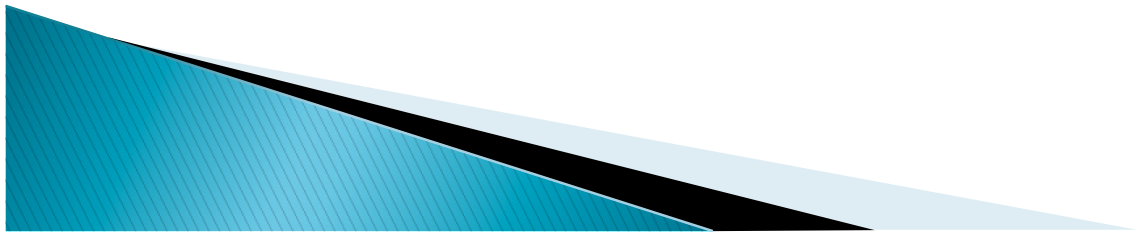
MATH 3390

- NSSE data set.
  - Performed Exploratory Factor Analysis.
    - Checked that mathematically linked variables were also logically linked.
  - Conducted scale reliability analysis.
    - Analyzed Cronbach's alpha and item-deleted alphas for extracted factors.
    - Analyzed KMO-Bartlett measures.
  - Revised the factors based upon reliability measures



MATH 3390

- NSSE factors plus new factors were studied.
  - ANAOVA's, multiple logistic regression and multiple regression (exploratory modeling).
- Reported results back to Office of Academic Affairs.
- A student continued to analyze the data beyond the end of the first semester of 3390.
  - Hopes to publish a peer-reviewed paper based upon her work in Math 3390.



Conclusion

- Students reported:
  - Non-majors felt more confident when facing quantitative research projects in their major coursework.
  - Majors felt the projects were good capstone experiences for the course, combining topics from throughout the semester.
  - All found “large” data sets less daunting after working on these projects using technology like SPSS.
- What we value as instructors:
  - Because they run multiple types of tests in the same project, our students develop a deeper understanding of the similarities and difference between the types of procedures.
- Good news
  - Analyzing data relevant to our campus increases student interest and helps us professionally as well.