The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: www.tandfonline.com/journals/utas20

A Capstone Course for Undergraduate Statistics Majors

Nicole A. Lazar, Jaxk Reeves & Christine Franklin

To cite this article: Nicole A. Lazar, Jaxk Reeves & Christine Franklin (2011) A Capstone Course for Undergraduate Statistics Majors, The American Statistician, 65:3, 183-189, DOI: 10.1198/tast.2011.10240

To link to this article: https://doi.org/10.1198/tast.2011.10240

Published online: 24 Jan 2012.

A Capstone Course for Undergraduate Statistics Majors

Nicole A. LAZAR, Jaxk REEVES, and Christine FRANKLIN

Many undergraduate statistics students receive limited exposure to real data and the challenges of real data analysis. To help improve our undergraduate program at the University of Georgia, we developed a Statistics Capstone Course. The course has three main threads: (1) teaching advanced/modern statistical methods to undergraduate statistics students; (2) giving these students an intensive, year-long data-analysis experience; and (3) providing the students with an opportunity to improve their written and oral communication skills. In this article, we describe the philosophy behind the Capstone Course, detail its implementation, and informally evaluate the success of our endeavor.

KEY WORDS: Advanced data analysis; Statistical communication skills; Undergraduate consulting experience.

1. INTRODUCTION

In Academic Year 2007–2008, the Department of Statistics at the University of Georgia (UGA) introduced a new Capstone Course for our undergraduate majors and minors. We did this because we believe that it is increasingly important that our graduating students have experience beyond the textbook; that is, they must have experience with analyzing real data. Real datasets are messy; they have missing or problematic observations, and it is not always clear how to analyze them. None of these issues are faced by students who learn all of their data analysis from textbooks, where datasets are clean and analysis is in most instances obvious, since problems are assigned to chapters that focus on particular methods. Furthermore, it is a useful skill to be able to interact with a scientist who has a real problem to solve. This skill can truly be acquired only through experience. Finally, we believe that a capstone course is an appropriate setting in which to give students an opportunity to learn about more advanced statistical methods which will (a) round out their education; (b) help them see that there is more to statistical analysis than t-tests, analysis of variance, and regression; (c) give them more confidence in their own data analysis skills; and (d) make them more attractive to prospective employers. Writ large, the goal of the Capstone Course is to develop *statistical thinking* (Brown and Kass 2009) at the undergraduate level.

Nicole Lazar is Professor (E-mail: *nlazar@stat.uga.edu*), Jaxk Reeves is Associate Professor and Director of the Statistical Consulting Center, and Christine Franklin is Senior Lecturer and Lothar Tresp Honoratus Honors Professor, Department of Statistics, University of Georgia, Athens, GA 30602. The authors thank the participants in the Capstone Course for their enthusiasm; Ana Bargo and Adam Jaeger for their great dedication as TAs for the class; and the support of the Writing Intensive Program at the University of Georgia.

The UGA Task Force on General Education and Student Learning recently made the recommendation: "Encourage all students, and require all students graduating with honors in their major, to complete a substantive capstone experience in their major." This recommendation is particularly sensible for statistics students. The American Statistical Association (ASA) also provides curriculum guidelines for undergraduate programs in the statistical sciences. Some of the recommended skills ASA proposes for a successful statistics undergraduate are: effective technical writing and presentations; teamwork and collaboration; planning for data collection; data management; and familiarity with standard statistical software packages (American Statistical Association 2010). These recommendations have been reflected in much of the recent ongoing discussion about the training of statisticians at all levels, and, indeed, more widely, about the future of our discipline (Brown and Kass 2009). The Capstone Course is being offered (in AY 2011–2012) for the fourth time. We have found it to be an ideal avenue for furthering the recommendations from both UGA and ASA (American Statistical Association 2005, 2010), as well as allowing us to confront head-on many of the challenges inherent in bringing the education of our undergraduate majors "into the modern age of statistics."

One of the strengths of such a course is that it can cover a wide variety of topics. Since most of our students come to the Capstone Course after sequences in which they have learned the basic methods of regression and ANOVA, the instructors can concentrate on other techniques, according to their preferences, the preferences of the class, and the particular projects on which the students will be working. Possibilities include: time series, data mining, factor analysis, classification and regression trees, smoothing, bootstrap, and categorical data analysis. The idea is not to go deeply into any one topic; rather, we aim to cover many different topics over the course of the year, spending at most two or three lectures on each (Nolan and Temple-Lang 2009). In addition, the course includes units on effective written and oral communication, and on how to make a poster presentation. The remainder of the class periods are used for student presentations of their datasets and analysis.

A goal of the Capstone Course is to expose students to the wealth of data available from the many disciplines at the University. We actively seek suitable datasets from our university colleagues for analysis by the Capstone students and encourage these researchers from other disciplines (the "clients") to interact with the students and to provide guest lectures. This supports another recommendation from the UGA Task Force on General Education and Student Learning: "Modify undergraduate major requirements to enhance the ability of students to take courses outside of their discipline and recommend that departments engaged in undergraduate education examine their major requirements to ensure that they are providing their students

 DOI: 10.1198/tast.2011.10240

with a broad, general education." We have interpreted this recommendation to expand students' education by providing more interdisciplinary study.

The course has now been completed by students thrice, in AYs 2007–2008, 2008–2009, and 2010–2011; in AY 2009–2010, the course was canceled by upper administration due to low enrollment brought on by a small cohort with scheduling conflicts. Capstone is now mandatory for all of our students; it is not possible at this time for a student to repeat the course, although there may be benefits to doing so. In line with the philosophy outlined above, and as we will detail in the later sections, students have worked on a variety of data analysis projects for researchers around campus, and in one case from the Centers for Disease Control (CDC) in Atlanta, GA. In AY 2008–2009, the students also had the opportunity to design, administer, and analyze a survey for a local nonprofit organization, Nuci's Space. For AY 2010–2011, the students again were in the fortunate position of being able to design a study, and to collect and analyze the resulting data, this time for UGA's Transit Authority. The emphasis throughout is on *real-world problem solving* (Brown and Kass 2009; Legler et al. 2010)—with all the messiness, ambiguity, and frustration that implies.

2. COMPONENTS OF THE COURSE

2.1 Team Teaching

The Capstone Course is taught by a team of two faculty members and one teaching assistant (TA). Part of the philosophy of the course is that students should have the chance to work closely with and to learn from research-active, tenure-stream faculty, and if possible those should be senior faculty. Accordingly, the course has been taught so far by Associate or Full Professors in the department. The course is a year-long course, with the same two faculty members, TA, and students being involved both semesters. As we have implemented it, the course credit is 2 hours per semester, meaning that it is scheduled to meet for two 50-minute periods each week of a 15-week semester. The two faculty members each receive one semester course credit for their year-long involvement.

2.1.1 Lectures

In the Fall semester, there is a somewhat heavier lecture load than in the Spring. This is split between the two instructors. Each year, we decide in advance what topics, in general, we wish to cover, and which instructor will be responsible for delivering which units. In Spring, the student groups give several presentations; if there is also an additional project such as the Nuci's Space survey or the UGA Transit Authority study, class times are used for discussion and presentation of those. Hence, the actual lecturing responsibility on any one faculty member is rather small. It has been our policy that whenever possible both instructors as well as the TA attend all lectures; in this way the students also gain the benefit of seeing when we disagree or when one instructor has a different perspective or experience than the other. This is also helpful to emphasize the idea that statisticians may reasonably disagree on how to analyze a dataset, thereby reinforcing the notion that in real life, unlike in textbooks, there is no one "right answer."

2.1.2 Project Supervision

Project supervision from the statistical perspective is split between the two instructors, with the TA acting as a "floater" who provides supervision to all groups on an as-needed basis. Typically, we have four or five groups of two to three students each, so that each instructor is responsible for two or three groups. These seem to be optimal choices for us. With our current resources, for example, we cannot effectively supervise more than five or six groups. And we've found that the group dynamics with two or three students encourage all to be active participants. In the Spring semester, one of the weekly class periods is devoted to lectures or group presentations; during the other, the instructors meet with their groups. Students are expected to come to these weekly meetings prepared with new analysis results, updates, and questions. That is, they are expected to be active participants, and not simply passive recipients of whatever wisdom or advice the faculty member/TA chooses to impart. The groups also receive supervision on the scientific aspects of their project from the client who provided the data, although this tends to be more sporadic. We were encouraged to see that a similar model has evolved at St. Olaf College's Center for Interdisciplinary Research (Legler et al. 2010), which shares many of the aims of our Capstone Course, albeit carried out on a much larger scale.

2.1.3 Dedicated Teaching Assistant

We feel that continuity in the instructional staff is important to the success of the course; this includes using the same TA throughout the year. During the second and third years of the course, thanks to participation in and support from UGA's Writing Intensive Program (WIP), we were able to have not just the same TA for the whole year, but also that TA was dedicated entirely to our course, an assignment of 18 hours per week, with no other departmental duties, and was trained to devise and grade writing assignments. In addition, our TA for AY 2010–2011 was a former Capstone student.

There are many advantages to having a dedicated TA. Since the TA is working only on this one course, whereas TAs in our department typically are assistants for two classes, he or she has more time (18 hours with WIP support versus 9 without) to devote to helping the students on their projects, and in taking on the secondary supervising role. Also, we have found that the Capstone students tend to form a bond with the TA beyond that formed with the faculty members. This gives them an extra sounding board and helps them to feel more a part of the department. The dedicated and consistent TA can also help smooth over difficulties that arise within groups; some students might be hesitant to bring these problems before a faculty supervisor.

2.2 Advanced/Modern Topics

One of the aims of the Capstone Course is to give our undergraduates exposure to topics that they might not otherwise encounter, since we have neither the resources nor the student demand to offer a wide variety of narrow, specialized courses at the undergraduate level. This can have immediate practical benefits in giving the teams the tools they will need to analyze their datasets. Beyond that, the course simply broadens the students'

statistical horizons, which can have the added value of prompting thoughts of graduate school. Students have reported to us informally, for example, that some of the advanced topics they saw during the Capstone year convinced them that they didn't know as much as they thought they did about statistics; that this intrigued and excited them; and that it spurred them to want to learn more. The introduction of modern topics also reinforces the idea that statistics is an interesting, dynamic discipline in its own right (Nolan and Temple-Lang 2009). Ideally, the topics covered will vary from year to year with different pairs of instructors, keeping the course exciting and current. Note that we are not aiming for expertise or depth; rather, the idea is simply to introduce students to a variety of techniques that are available, beyond the familiar linear model.

Related to this, students in the Capstone Course also gain exposure to multiple programming languages. Most of our students learn SAS in their various applied classes. Capstone, in addition, introduces them at least to R/Splus and Matlab, showing them more flexible programming-oriented approaches to data analysis. Again, as with the specialized topics, we are not expecting the students to become expert R or Matlab users, aiming instead for basic awareness and familiarity. We have experimented with ways of introducing these new languages—formal in-class tutorial with instructor and TA, or informally via examples in handouts and as projects require. In practice, if students do move away from SAS for their projects, it will be to one of R or Matlab, not both. Some clients come in requesting analyses in a certain language, so that they can distribute code more widely among their colleagues; more and more we are finding clients requesting analyses in R. Other data are particularly amenable to analysis in a specific language. Students learn from this that there are multiple platforms, that they must be able to accommodate client requests on this front, and that it is helpful for them to know multiple statistical software packages.

2.3 Real Data Analysis

An important aspect of the course is that the students are working with *real* data, usually provided by researchers on campus, although we have also accepted or solicited projects from off-campus. The importance of working with real data has been widely recognized (see, among numerous others, Bryce et al. 2001; Tarpey et al. 2002; American Statistical Association 2005, 2010; Gibbs and Reid 2009). Before the school year starts, we contact colleagues from the UGA community who might have data suitable for a team of undergraduates. Each is asked to give a lecture to the class about the data and the scientific questions of interest. Since we arrange for more potential projects than there are student groups, we emphasize to each prospective client that his or her project might not be selected by any group. Also, we emphasize that undergraduate students will be analyzing the data over a school year, so "mission critical" projects, either in terms of timing, since the course lasts a year, or in terms of impact on future research, should not in general be presented. It's quite possible—although we have yet to experience this ourselves—that a team will simply not be able to make any progress beyond that which the researchers would have been able to accomplish on their own.

Table 1. Summary of Capstone projects. These are the projects that students actually chose to work on; others were presented but not selected by any group.

Title	Client department
Comparison of Latin American and Indian art sales	School of Business
Analysis of single-subject fMRI data	Psychology
Study of the Indian art market	School of Business
Study of PBDE concentration in the body	Centers for Disease Control
The development of visual attention	Psychology
Mortality rates of dogs	Genetics
Analysis of fMRI data from multiple subject groups	Psychology
Analysis of tense usage by Spanish speakers	Linguistics
Imaging seizures at the brain-scale in zebrafish larvae	Bioengineering and Mathematics
Strategies to alleviate stuttering in children	Communication Science
Water usage of greenhouse hydrangeas	Horticulture
Sound reduction in Spanish speakers	Linguistics
Effect of pressure oscillations on temperature	Agricultural Engineering
Temperature change in Georgia	Agricultural Engineering

In the first three years of the course, we were able to find a wide range of real data projects for the students; these came from developmental psychology, cognitive neuroscience (multiple projects), art (multiple projects), linguistics (multiple projects), health statistics, biomedical engineering, genetics, horticulture, communication, and climatology. A complete list of the projects, along with the client departments, is in Table 1. For the second run of the course, the students designed a survey on medical insurance needs of musicians for Nuci's Space, an Athens, GA, nonprofit organization that serves the local music community. Working with the directors of Nuci's Space, the students designed a survey to assess the need and desire of local musicians for health and dental insurance. The students then set up a MySpace page to contact Athens-area bands and musicians to make them aware of the survey, which was administered via SurveyMonkey. After collecting the survey data, the students analyzed the results and presented a final written report to the director of Nuci's Space. All students worked on this project over the course of the year, with some students being on teams that required more effort in the Fall (e.g., building the sampling frame) and others in the Spring (e.g., analyzing the data).

In AY 2010–2011, in addition to their main projects, students also worked on a project for UGA's Transit Authority to (a) estimate the number of passenger miles per year on the campus bus system and (b) estimate how many UGA-affiliated passengers use the municipal bus system (on which students, faculty, and staff ride free; the University pays the city for this use) for rides that take place solely on campus. The students devised sampling schemes and collected the data during Fall semester, and analyzed the results, writing them up as final reports, in the Spring.

We think that there are obvious advantages to this approach. Students learn, first of all, that they can have a real impact on someone else's work—this is both a great responsibility and a positive opportunity. Second, they see that real life doesn't work like a textbook. Data are messy, or incomplete; it's not always clear what the right analysis is—or even that there is such a

"right" analysis. Third, they learn about the interpersonal side of doing statistics in a consulting environment. For instance, clients can change their minds about what they want, or might not know exactly what they want; they may inadvertently conceal possibly relevant information, which comes to light only after the fact. Some clients want a lot of communication with the group, others don't. Students learn to accommodate a variety of client styles.

2.4 Writing Intensive

According to Bryce et al. (2001), effective written communication is an important aspect of interacting with scientists, yet it is often overlooked in the training of statisticians. We agree with this, and so in the second year of the Capstone Course, we participated for the first time in UGA's Writing Intensive Program; that participation continued in AY 2010–2011. This program trains a graduate assistant to provide detailed feedback on student writing, to help devise "low- and medium-stakes" writing assignments, and to generally aid faculty who wish to make writing a more integral part of their classes.

Offering the course as writing intensive means that we can scatter smaller writing assignments throughout the year, so-called "low-stakes" writing. These are exercises that don't have a large impact on the final grade, that generally don't take a long time, but that afford the students regular writing practice. In addition, the writing intensive structure gives us a framework for helping the students produce their final project reports in an organized manner. In the Spring semester, we have a relatively rigid schedule of deadlines for both students and instructors. The groups are required to submit draft pieces of their reports at set times; the instructors and the TA then have one week to read and comment on the drafts, which are then subject to revisions by the students. All of the commenting and revision are done electronically.

Support from WIP brings many benefits, as outlined here and in Section 2.1.3. The training that the TA receives, in particular, has helped us to greatly improve the quality of the final written reports. Even without this infrastructure, though, there are many resources on writing available online through university writing centers, as well as books about writing. These can be used in a Capstone course to either train faculty and TAs, or as direct aids for the undergraduates.

2.5 Responsibility to Clients and Teammates

The Capstone Course provides students with a variety of interactions—between students and instructors/TA, between students and their clients, and among the students working together in a group or groups. The structure of the course emphasizes the responsibility that the students have in all of these interactions. Much of this takes place in the Spring semester, when work on the projects is at its height. Weekly meetings with the project supervisor and frequent in-class oral presentations are two of the tools we use to highlight the importance of responsible, timely performance. A positive side effect of the approach is that it also helps the students to develop communication skills through group work, class presentations, and team work (Tarpey et al. 2002).

2.5.1 Poster Presentation

At the end of the year, we host a poster session. The entire Statistics Department, faculty and graduate students, is invited, as are the clients. The Capstone students are expected to prepare a conference-style poster of their project and to be able to answer questions from any attendee. In our experience, the students take this very seriously; they don't want to be embarrassed in front of the department, nor do they want to embarrass their instructors. Some of the groups have been quite creative in their posters. For instance, in the first year, a group working on a project regarding the Indian art market brought a digital frame that rotated through paintings that they had studied. In the second year, a group that had dynamic data evolving over time prepared a movie of the data that played on continuous loop in front of their poster.

2.5.2 Final In-Class Presentation and Written Report

In addition to the poster session, each group is required to prepare a final in-class presentation and to submit a written report. The in-class presentations are generally for the students, instructors, undergraduate coordinator, and TA only; clients are not currently invited, although we are considering changing this in future offerings of the course, since clients often aren't able to attend the poster session. Each group is permitted half a lecture period to present final results of the analyses and to discuss any challenges or unexpected findings. The written report is sent to the client after it has been approved by the instructors.

2.5.3 Peer Evaluations

As a final check, at the end of each term, students are asked to evaluate the contributions of *all* in their team, including themselves, to the overall progress of the project. This is in the form of a three-point rating scale, where "1" indicates a fully contributing team member and "3" a team member who did not contribute his or her fair share. Usually we are aware of intergroup frictions that may be impeding the work, but we want to give the students the chance to engage in self-criticism, and to give their perspectives on how the team functioned. For the most part, students give the best rating to all members of the team; even when they don't, they tend not to be hypercritical in writing, although we have received occasional negative comments in person about some members of some groups. Interestingly, students are not shy about rating their own contributions negatively, if they feel this is justified. The Associate Editor for this article kindly provided us with a more detailed peer evaluation form that requires the students to evaluate each other along multiple dimensions; we will use this form in the coming academic year.

3. LOGISTICS

A course like this requires much coordination, not only between the instructors, but also between the instructors and the TA, among the student groups, and with the undergraduate coordinator, who advises all of the students and is responsible for assessment of the undergraduate statistics program. In this section, we outline the general scheme we use to keep everything on track, with special attention to timeline.

3.1 Fall Semester

In the Fall, there is more focus on faculty lectures on advanced or modern topics. We have tended to start the semester with review lectures on exploratory data analysis (EDA) and regression, moving from there to generalized linear models. Specialized topics that we have also covered during the "review" have included smoothing, modern regression, and approaches for handling multiple testing problems. We try in these lectures to provide the students with material that may be interesting or relevant for their projects. We also discuss statistical consulting in preparation for the client presentations, and explore websites such as *fivethirtyeight.com*.

Typically, starting in early October, which is about week 6 of the semester, we hold the client presentations. Each client is given a class period to present the project, the data, and the questions of interest, and is also available to answer questions from the students. The students are informed beforehand that they are responsible for forming groups, and each group will eventually be asked to rank their top three projects of those presented. Client presentations take a few weeks, ending in week 8 or 9, after which the groups are required to submit group identities and ranked project selections to the instructors. In the second year of the course, we decided to intervene and rearrange some of the groups, since a couple of students had failed to talk to any classmates about working together. We felt that a group composed solely of students who hadn't even had enough initiative to join a team would not be able to complete a project either! Student interests are diverse enough that in the first two years all groups were able to receive either their first or second choice. In the third year, one group worked on its third choice project.

During the last week of the semester, each group is required to give a 15-minute project presentation. This presentation should include a short introduction to the scientific background and questions, and simple EDA. The purpose of this early presentation is simply to guarantee, before the start of the Spring semester, that all groups can access their data and carry out basic data manipulations. At this time the groups also submit a short written report that parallels the oral presentation.

At the end of this semester, each student is given an individual letter grade, based on homework, writing, project participation, and peer evaluation.

3.2 Spring Semester

Spring semester is when the serious project work takes place. Lectures are reduced to one per week at most; during the other class period, the groups meet with their supervising instructor or the TA. In addition, there are periodic progress reports, in which each of the groups is required to give an oral presentation to the entire class on the status of their project. These are opportunities to receive feedback from the nonsupervising instructor, who has not been following the project as closely, and from classmates. As the semester proceeds, these presentations are expected to be more in-depth, and more of the data analysis problems are expected to be solved. Now that the course is being offered in a Writing Intensive format (see Section 2.4), there is also a schedule of drafts and revisions put into place early in the semester.

If the students have only their team projects to work on, as was the case in the first year, we continue to assign sporadic homework on the advanced topics that are covered in lecture. In the second and third years, since they were also working on the other surveys, we did not give any additional homework, and used some class periods for discussing the secondary project and its analysis.

The semester ends with a poster presentation to which the entire statistics department is invited, final in-class presentations during the last week of classes, and submission of a written report to be given to the clients.

Just as in Fall, students are given an individual letter grade based on writing, project participation in all of its aspects, and peer evaluation. Recognizing that some datasets are inherently more difficult than others, we try at this stage to adjust expectations when assigning grades. For easier projects, we expect that the students will be able to make greater progress than they would on harder projects. Admittedly this is somewhat subjective, but after nine months of working closely with the groups, we know them and their capabilities quite well.

4. DOES IT WORK?

What does it mean for a course like this to be effective? Some possible answers to this question include: Successful data analyses, resulting in happy, potentially repeat, clients; enthusiastic students; decisions on the part of students to go to graduate school in statistics or related areas, where they might not have otherwise; and acquisition of "soft skills." On all of these dimensions, we deem the Capstone Course to be a success.

Comments from our clients have been overwhelmingly positive, with our faculty colleagues telling us that they would be interested in participating again ("Please think of me next time!" is a common request). They are impressed that undergraduates have been able to do so much with their data. And they ask if they can advertise the course to colleagues in their own departments; limitations on the number of projects we can accept and the types of projects that are suitable has meant that we must be selective and can't advertise broadly, but we take this as a sign that the needs of the clients are being met.

As for the students, of the 36 students who completed the course in the first three years it was offered, four decided to pursue M.S. degrees in statistics immediately upon graduation, of which three had been previously uncertain about their plans; one was accepted to an M.S. program at UGA's business school; one went to a Ph.D. program in psychology; three are currently applying to M.S. programs in biostatistics and bioinformatics; several of the students were in combined B.S./M.S. programs at the time of taking the course. In many of the cases, the students told us that their decision to go to graduate school at all, and to graduate school in statistics or a related area more specifically, was directly due to their experiences in the Capstone Course. Others have told us that they were able to find jobs and be successful in them as a result of their participation in the course, which is at least anecdotal evidence that they have improved their "soft skills" over the year. While it would have been interesting to compare these outcomes to those of students who passed through our program before the Capstone Course was

put into place, we unfortunately are not able to do so, as typically our undergraduate majors have not tended to stay in contact with the department after graduation. The Capstone students, by contrast, have maintained stronger relations with the department post-graduation, perhaps because they felt more a part of the profession.

The Statistics department at UGA also routinely conducts exit interviews with the group of graduating majors. Students who have taken the Capstone Course rank it as one of their best, most helpful, experiences in the program. They are unanimously appreciative that the course exists and have been very forthcoming with ways of making it even better. It was students in the first year, for example, who suggested that working on a survey might be a good addition; as happy coincidence the Nuci's Space project presented itself in time for the second year. Student evaluations of the course are also very positive.

In preparing this article, we contacted students who had been through the Capstone Course in the first two years and asked them to provide us with feedback about the experience with now the additional perspective of having been in graduate school or the job market for a year or more. Selected comments, edited slightly for space and in some instances to preserve anonymity of the students:

- The Capstone course was a great experience for me. It gave me the opportunity to approach a real-world statistical issue, and work with professors in a hands-on environment. I often felt that my other statistics courses were good for learning the fundamentals of how to approach a statistical problem, but the problems were often too simplified and did not represent issues that statisticians face in the real world. The Capstone course served as a bridge between academics and a job like consulting where actual data must be collected, analyzed, and presented. This course gave me the confidence that I could handle complex statistical problems.
- I thought that Capstone was a great experience. When I first started taking stat classes back in high school, all of the data that we worked with were clean and ready to analyze. This is often not the case [in the] real world. This class is the first experience many students have with data that "needs work." Often, getting the data into a form that can be analyzed is a bigger problem than actually analyzing the data. The Capstone classes gave me valuable experience working with both clean and messy data. This experience proved invaluable during my internship as a data analyst at one of the local medical centers. I feel as though a good statistics program would not be complete without a similar Capstone class.
- Capstone was a great class! The most beneficial part of Capstone for me was determining how to solve a problem. In other classes, the questions were typically geared for only one way to solve the problem. However, in the real world, there are multiple ways to analyze data. We experienced this in Capstone. In our final projects, we had to think of all the multiple ways to answer a question and then determine which was best. The multiple solutions forced students to use their entire knowledge of statistics, not simply one chapter in one class. This is something I face in the real world now. My boss doesn't tell me what to do in order to find the solution, I'm simply expected to come up with an answer.
- For me the Capstone Course was the most important class I took as an undergraduate statistics major. In class, we typically learn things in units, and when a situation arises on a homework problem or on a test, it is fairly easy to figure out what we need to do. In Capstone, we were put into a unique position, one that centered on something that we had no expertise in, and had to figure out a way to best handle the problem. While we were given advice from our teachers and the leader of the project, a lot of the decision making was left up to us.... In addition, this was the best situation I had to develop my writing in statistical work. This was the most "real-life" experience we had as undergrads.
- I appreciated the range of projects our class worked on—from fMRI to the zebrafish to visual attention, etc. When working on homework/labs in the classroom I always found myself looking at similar "cookie cutter" applications where the need for statistics was obvious. In Capstone I saw the application of statistics in ways that I had not really thought about.

  The biggest advantage was learning how to "write to an audience." I think most statistics majors can explain a hypothesis, the test statistics, level of confidence, and conclusion. However, most students can only explain the experiment using a cookie-cutter approach (somewhat of a methodical step 1, step 2 approach using only statistical terms). I appreciated the time that the instructors took when editing our papers to help explain the methodology better. There are plenty of people who do not know, when reading a statistical result, what the level of confidence is and its implications (there could also be a few stat majors who don't understand what it is either). Often in consulting, finance, wealth management papers are written to senior management and clients who don't know much more than "mean, median, mode, linear regression," so writing simple, clear sentences is necessary. Before Capstone I would try to include lots of big, statistical phrases in my answers/projects/papers. Now when I write a report I keep things as simple and clear as possible.
- Capstone was extremely beneficial to what I am currently doing. Almost immediately after graduating I had an opportunity to work on a project analyzing the effects of climate on viral outbreaks. Having taken Capstone was a huge benefit for this. I am currently working on my third project since Capstone and I find much of what I learned from the capstone course is still applicable to my current projects.

  I do believe the main advantage of the Capstone course is that it exposes undergrads to how actual projects really work. Normally the data we analyze in the undergrad classes are picked to show one specific point, or they are picked to avoid any confounding effects (such as using data for multiple regression where there is no correlation between the explanatory variables). Also the students have an opportunity to look at data and just from preliminary analysis decide what needs to be done based on the client's question, and also the students see and get to try new ways of analysis they would probably not learn in other classes. Finally the big lesson from Capstone is that real data can be messy and most of the statistical tests we learn require distributional assumptions that usually don't hold.

A number of common themes arise from the thoughtful comments provided by the students. First, the course prepares them for the real world and gives them confidence that they will be able to handle data analysis in a more unstructured environment; this is indirect evidence that they have acquired at least some of the necessary soft skills to succeed in their jobs. Second, the course teaches them to think creatively about new problems, ones that don't fit into a standard mold. In this way, the students gain "ownership" of their projects through active learning (Weinberg and Abramowitz 2000). Third, the writing component is beneficial in a number of ways. Indeed, we were surprised to read from several students that they wanted even more writing practice; this seemed to us to be the element of the course that caused the most "pain and misery" in the second year, but in spite of this it was obviously useful.

5. CONCLUSION

We have successfully implemented a Capstone Course at UGA. The course provides our undergraduate majors with an intensive data-analysis experience and exposure to real statistical work. In that sense, it is similar in motivation and implementation to the Center for Interdisciplinary Research at St. Olaf College described by Legler et al. (2010), but run with much less in the way of resources and infrastructure. Once our faculty agreed that such a course would be a beneficial addition to the undergraduate curriculum, it wasn't that difficult to implement. Like most universities, the University of Georgia is trying to expand and deepen interdisciplinary work, even at the undergraduate level. This type of course is very appropriate to that particular educational mission. Many researchers on campuses are eager to receive statistical help; in many instances there is no statistical consulting service, but even when there is, as is the case here at UGA, faculty may not avail themselves of that service for a variety of reasons. A Capstone Course offers these faculty a way of getting some statistical advice, and we have additionally found our friends and colleagues to be very sympathetic to the *philosophy* of the course, namely, training our undergraduates in real data analysis. Support of the local faculty has been key to the success of the course. With some investment of time and energy from statistics department faculty, such Capstone Courses could be put into place on other campuses as well. And, importantly in these times of budgetary constraints, it can be done with little or no external financial support. We believe the effort is worthwhile, and hope we have convinced you to try this yourselves.

[Received December 2010. Revised June 2011.]

REFERENCES

American Statistical Association (2005), *Guidelines for Assessment and Instruction in Statistical Education, College Report*, Alexandria, VA: American Statistical Association. [183,185]

——— (2010), "Curriculum Guidelines for Undergraduate Programs in Statistical Science," available at *http://www.amstat.org/education/curriculumguidelines.cfm*. [183,185]

Brown, E., and Kass, R. (2009), "Statistical Training and Curricular Revision," *The American Statistician*, 63, 105–110. [183,184]

Bryce, G. R., Gould, R., Notz, W. I., and Peck, R. L. (2001), "Curriculum Guidelines for Bachelor of Science Degrees in Statistical Science," *The American Statistician*, 55, 7–13. [185,186]

Gibbs, A., and Reid, N. (2009), Comment on "Statistical Training and Curricular Revision," by E. Brown and R. Kass, *The American Statistician*, 63, 112–113. [185]

Legler, J., Roback, P., Ziegler-Graham, K., Scott, J., Lane-Getaz, S., and Richey, M. (2010), "A Model for an Interdisciplinary Undergraduate Research Program," *The American Statistician*, 64, 59–69. [184,189]

Nolan, D., and Temple-Lang, D. (2009), Comment on "Statistical Training and Curricular Revision," by E. Brown and R. Kass, *The American Statistician*, 63, 117–121. [183,185]

Tarpey, T., Acuna, C., Cobb, G., and De Veaux, R. (2002), "Curriculum Guidelines for Bachelor of Arts Degrees in Statistical Science," *Journal of Statistics Education*, 10. Available at *http://www.amstat.org/PUBLICATIONS/JSE/*. [185,186]

Weinberg, S. L., and Abramowitz, S. K. (2000), "Making General Principles Come Alive in the Classroom Using an Active Case Studies Approach," *Journal of Statistics Education*, 8. Available at *http://www.amstat.org/PUBLICATIONS/JSE/*. [189]