



A Survey of Statistical Capstone Projects

Susan E. Martonosi & Talithia D. Williams

To cite this article: Susan E. Martonosi & Talithia D. Williams (2016) A Survey of Statistical Capstone Projects, Journal of Statistics Education, 24:3, 127-135, DOI: 10.1080/10691898.2016.1257927

To link to this article: <https://doi.org/10.1080/10691898.2016.1257927>



© 2016 The Author(s). Published with license by American Statistical Association© Susan E. Martonosi and Talithia D. Williams



Accepted author version posted online: 11 Nov 2016.
Published online: 11 Nov 2016.



Submit your article to this journal [↗](#)



Article views: 3402



View Crossmark data [↗](#)

A Survey of Statistical Capstone Projects

Susan E. Martonosi and Talithia D. Williams

Department of Mathematics, Harvey Mudd College, Claremont, CA

ABSTRACT

In this article, we highlight the advantages of incorporating a statistical capstone experience in the undergraduate curriculum, where students perform an in-depth analysis of real-world data. Capstone experiences develop statistical thinking by allowing students to engage in a consulting-like experience that requires skills outside the scope of traditional courses: defining a complex problem, analyzing data, building a strong team, and communicating effectively. We describe the pedagogical benefits as they relate to improved student outcomes and prospective job and graduate school placement, and we classify statistical capstones into four groups: standalone capstone projects, statistical consultancies, capstone projects embedded in an advanced statistics methodology course, and instruction-focused capstone courses. This article serves as a guide for educators seeking to implement an enriching capstone experience in their undergraduate mathematics or statistics curriculum to better prepare students for industrial and academic careers in data science.

KEYWORDS

Capstone; Statistical consultancy; Statistics projects

1. Introduction

In recent years, we have seen an increase in the number of undergraduate degrees in statistics as society responds to the need for a data-analytics trained workforce (Wasserstein 2015). “Big data” now permeates not only just business, but also academic research in many disciplines such as genomics, social sciences, computer science, and more (Manyika et al. 2011). Today’s statisticians are integral to the emerging information era and must acquire cross-disciplinary expertise within industry, medicine, and the many other sectors that depend on data analytics for decision making (Wild and Pfannkuch 1999; Gibbs and Reid 2009). Statistics has emerged as a distinct discipline and no longer sits on the mathematical sideline for use by a few highly trained specialists.

Hardin et al. (2015) emphasized the importance of students learning the full process of statistical analysis, which includes forming a research question, designing an experiment and collecting data, preparing that data for analysis, iteratively choosing an appropriate model for analyzing the data, and communicating the results to stakeholders. In fact, Leman et al. (2015) argued that statisticians need to integrate the qualitative-quantitative-qualitative layers of a problem: first, developing qualitative understanding of the problem being solved and its context (i.e., a *real* problem to be tackled); next, performing quantitative analysis; and finally, communicating qualitative results to stakeholders and decision makers. Today’s statisticians must be just as adept at using computational techniques to apply theory to the real world as they are at communicating ideas and results across disciplines (Horton 2015). At a bare minimum, an undergraduate statistics program should include collection of real data,

computing, synthesis of statistical methodologies, and communication of results.

While the statistics community has been quick to respond to the demands for increased statistical literacy, we have often neglected the importance of incorporating a unifying experience that bridges statistical concepts across courses and disciplines. Even students who have done well in their courses might not have had enough experience with real datasets to be able to solve the sorts of problems typically seen in industry or research (Young et al. 2015). The skills needed to be successful in data science, such as parallel computing and distributed data management, are often missing from the typical undergraduate curriculum (Davidian 2013). Therefore, to prepare today’s students for tomorrow’s opportunities, we must train them in “mapping a scientific question into a statistical approach and developing understanding, experience, and intuition of when and how to use statistical methodology in the scientific context” (Nolan and Temple Lang 2009, p. 117). There is a national need to incorporate inquiry-based learning outside of the classroom while elevating the interdisciplinary research experience, moving away from statisticians acting as statistical consultants and instead becoming integral to a research experience (Landes 2009; Legler et al. 2010). Statistical capstone experiences can play a pivotal role in bringing rich research experiences into the classroom.

A capstone experience is broadly defined as a long-term, multilayered project that culminates a student’s academic experience. Capstone experiences encourage students to think critically by solving challenging problems, often developing skills such as teamwork and oral communication along the way. Durel (1993) noted that the capstone course experience

“integrates a body of fragmented knowledge into a relatively unified whole,” allowing students to “disengage... from the undergraduate status and reemerge ... as graduates prepared to assess critically and act responsibly in civil society” (Durel 1993, p. 223). Statistical capstone experiences reinforce a culture of interdisciplinary collaboration and increase statistical maturity among students. A primary goal of the capstone experience is to develop statistical thinking and a statistical perspective in students, so that they can best contribute to multidisciplinary situations (Brown and Kass 2009). Many capstone experiences already incorporate key aspects of a modern day statistics undergraduate curriculum, including planning and managing scientific studies, computer science and data management, and communication skills (Higgins 1999). Additionally, the American Statistical Association (ASA) Committee on Professional Ethics recommends a capstone course in an undergraduate statistics major that includes one or more projects and exposure to ASA ethical guidelines (Committee on Professional Ethics 1999; Tarpey et al. 2002; Cohen 2014).

The recent paper by Horton and Hardin (2015) provides an outstanding summary of pedagogical advances in statistics curricula. However, we see an opportunity to expand on its treatment of the capstone experience, which is the purpose of this article. We highlight the pedagogical benefits of statistical capstones and best practices for incorporating them into the statistics curriculum. Our objective for this article is to provide a useful reference for educators considering using statistical capstones in their curriculum. In Section 2, we summarize the benefits of statistical capstone experiences. In Section 3, we survey types of statistical capstones described in the literature and categorize these experiences into four groups: standalone capstone projects, statistical consultancies, capstone projects embedded in an advanced statistics methodology course, and instruction-focused capstone courses. We also summarize other considerations that apply to any statistical capstone experience. We close the article in Section 4 with some concluding thoughts.

2. Pedagogical Benefits of Statistical Capstones

There are many reasons why we should expect statistical capstones to be beneficial to students.

Khachatryan (2015) pointed out that similar to Bloom’s Taxonomy (Bloom et al. 1956), statistical case studies and projects touch upon *application* (e.g., figuring out that Holt’s method for data smoothing can also be used for imputation), *synthesis* (e.g., doing all steps of a statistical analysis from data acquisition and cleaning to model interpretation), and *evaluation* (e.g., comparing and contrasting model results and recommendations.) Moreover, hands-on projects like capstones provide an opportunity to expose students to advanced topics, help them to see the value of learning more statistics, and build their confidence as statisticians (Hillmer 1996; Nolan and Temple Lang 2015). In this section, we summarize some of the documented benefits of capstones as they pertain to student academic performance, professional training, and the student experience.

2.1. Student Performance Improvements

Statistics practitioners have long lauded capstone experiences as beneficial in helping students retain statistical ideas and material while providing opportunities to reinforce key concepts and assess student learning (American Statistical Association 2014). Various researchers have documented the beneficial effects of the capstone experience on student outcomes. They most notably include

- Improvement in student test scores, student attitudes, and the percentage of students earning a C or higher in a controlled study of the effect of a semester-long experiential learning statistics project compared with those in a traditional lecture-based course.
- Improved retention of statistical concepts, integration of statistical techniques, and integration of multiple datasets and outliers.
- Higher retention rates among students (in the same course) who completed 10-week team projects versus those who did not.
- Improved technical proficiency in statistical problem solving from students in consulting capstone experiences (Keeler and Steinhorst 1995; Kvam 2000; Hakeem 2001; Berenson et al. 2008; Kim et al. 2014; Khachatryan 2015).

Additionally, by having students present their work externally, Lazar et al. (2011) found that the quality of the overall work improved, and the statistical analyses were more robust. This eventually led to adding more applied courses and courses for nonmajors to the curriculum, thereby drawing more students into the major. Thus, evidence suggests that capstones solidify students’ understanding of statistical concepts.

2.2. Job and Graduate School Training Benefits

An analysis of job descriptions posted by employers looking for statisticians at the bachelor’s level found that in addition to specific statistical methods, employers desired job applicants to have a prior hands-on statistical experience and an awareness of the limitations of statistical methods. Equally as important, employers desired nonstatistical competencies such as the ability to explain statistical analyses and insights to nonstatisticians (Ritter et al. 2001). Graduate schools have likewise reaped the benefit of incorporating statistical capstones in the undergraduate experience. Legler et al. (2010) found a 10-fold increase in the number of students going on to graduate school over a 5-year period compared to the previous decade. In a survey of University of Dayton alumni zero to seven years post-graduation who participated in a client-based capstone project in operations management, over 90% of respondents described the capstone project as being more valuable than other courses they had taken (Gorman 2011).

We see several benefits of a statistical capstone experience for both graduate school and job possibilities, including

- The opportunity to apply statistics to real problems.
- Familiarity with the data—collection process, database management, and expertise in a statistical software.
- Effective communication skills and the ability to work in teams and manage people.

- Improved professional skills such as collaboration and written/oral communication.
- The ability to communicate results to a client and communicate technical concepts to a nontechnical audience (Bryce et al. 2001; Ritter et al. 2001; Lazar et al. 2011; Khachatryan 2015).

While the above skills can be strengthened in a hands-on capstone project, they tend to be neglected in a traditional classroom setting and are most notably the qualities that distinguish a practitioner from a theoretician (Higgins 1999). We found that many job interviewers focus primarily on asking students about their capstone experience, often allowing them to discuss their role in the project's completion (Malone et al. 2013). Given the high demand for talented data scientists and statisticians, a capstone experience can be extremely helpful in securing a job or gaining an edge when applying to graduate school.

2.3. Impact on the Student Experience

Not only do statistical capstone experiences improve student outcomes and career opportunities, but also allow students to develop an excitement and passion for statistics that extends beyond individual courses. Several articles in the literature on capstone programs describe how their students reported the experience. In our survey of these papers, we found that students

- Gained an appreciation of the entire process of statistical analysis including messy datasets having missing or redundant data.
- Learned to appreciate the broader statistical research endeavor beyond simple textbook exercises.
- Experienced creative and iterative model building and validation.
- Became more involved with the course content by designing experiments, collecting and analyzing their own data, and synthesizing ideas into a final report.
- Experienced overall excitement, enthusiasm, and diligence.
- Gained an improved understanding of the connections between courses and how to identify a faculty mentor.
- Understood the impact that statistics (and they) can have on an organization's work (Johnson and Dasgupta 2005; Lazar et al. 2011; Kim et al. 2014; Khachatryan 2015).

While this list is by no means exhaustive, it does suggest that the intangible benefits of incorporating a capstone experience can lead to increased student satisfaction and an appreciation of the entire undergraduate experience.

3. Designing a Statistical Capstone

After providing a clear motivation for hands-on projects in statistics education, we can now describe several of the design choices an institution can make when establishing a statistical capstone program. Sound practical and theoretical principles can be used to guide the design and development of statistical capstones that promote innovative learning. In our review of the literature referenced in this article, four main types

of statistical capstone experiences emerged: a standalone capstone project, a statistical consultancy, a capstone project within a statistical methodology course, and an instructional capstone course. These are not mutually exclusive categorizations, as field-based standalone capstone projects and statistical consultancies are often organized around a course. Moreover, other disciplines related to statistics use similar capstone experiences for educating their undergraduates; see, for example, a survey of field projects in operations research and management science by Martonosi (2012). However, these categories allow us to highlight key pedagogical differences in capstone experiences. What is common to these different types of capstones is that they help students grow as practitioners through their use of real data and emphasis on professional skill development.

3.1. Standalone Capstone Project

In a standalone capstone project, students complete a statistical project from start to finish. The projects are usually proposed by an interested party or client who has a real research or business problem to solve. The students are actively involved in formulating the problem in consultation with the client, selecting the appropriate solution approach, collecting and/or cleaning the data, developing and validating the model, analyzing the data, and communicating the results back to the stakeholder. In this way, the standalone capstone project is an authentic reflection of the sorts of activities a professional statistician would engage in both industry and academia.

Capstone projects are typically one or two semesters long and completed in teams of two to four students supervised by a professor. For example, at St. Olaf College, the teams of three to four students are supervised by both a statistics professor and by a professor in the application domain of the project. Students from any major can apply to participate in the program, provided they have taken a prerequisite course in applied regression. The project lasts for a full academic year, occasionally extending into the following summer (Legler et al. 2010). Miami University offers several major and minor options in statistics and related fields (Miami University 2016). Their one-semester capstone is required for the statistics major but is an elective for seniors in other statistical programs. Students are assigned to teams to balance out majors and prior experience in statistics and programming. One student is selected to be team leader, a position that rotates through the members of the team. The instructor acts as a senior consultant who directs the team to advanced methods as needed (Smucker and Bailer 2015). At the School of Computing and Management Sciences at Sheffield Hallam University in the United Kingdom, the 3-month long standalone capstone project serves as preparation for an off-site capstone experience, in which third-year students spend a full year in industry. Students work in teams of four, guided by a faculty supervisor whose role is to intervene only when the team appears to be heading off-course (Rangecroft and Wallace 1998). Carnegie Mellon University offers a semester-long capstone course by invitation only to its best students to overcome the challenges of scale in their large statistics program (Nugent 2012).

In many of these projects, there is an external client to whom the students report. In most cases, the projects come

from researchers in other departments at the university. Occasionally, local government agencies, nonprofit organizations, or companies serve as clients. The capstone program at Harvey Mudd College leverages its alumni base to recruit clients from across the country (Borrelli 2010). When choosing a potential client, Legler et al. (2010) advised that a good problem must be of interest to the researcher, not yet solved, but still solvable. They typically start a project with data in-hand, but occasionally involve students in the data collection. Moreover, they advise that while the data should require some cleaning so that students have a realistic experience, the data cleaning should not be too time consuming. They also recommend that a project consist of multiple phases so that it can be deemed successful even if not every stage is completed. Lazar et al. (2011) recruited more projects than they needed for the number of students they had and allowed the students to pick the projects they were interested in. They advise their clients to avoid entrusting a project on their organization's critical path to the students, both due to the long duration of the project and due to the uncertainty of success, advice which Borrelli (2010) echoed. Smucker and Bailer (2015) pointed out that the instructor responsible for vetting potential projects needs to be confident that the project is of the appropriate level of difficulty for the students without presolving it.

Because one of the objectives served by a capstone project is effective communication with a client (who might not have a statistical background), all of these programs require students to interact directly with the client on a regular basis. Therefore, the client must be willing to allocate time to the project; Borrelli (2010) recommended a few hours per week. For local clients, these interactions can take place in face-to-face meetings, and for remote clients, they can take place by telephone or video conferencing such as Skype or the occasional site visit. In the early stages of the project, the purpose of these interactions is for the students to listen to the client's needs and understand the project's context. In the intermediate stages of the project, the communication can take an iterative form in which the students present intermediate results, and the client provides feedback to help clarify the project's needs. In the final stages of the project, students prepare oral and written reports to communicate the results of the project and make recommendations to the client. Because effective communication is a two-way street, Nolan and Temple Lang (2015) recommended identifying clients who are enthusiastic about the project and can communicate well with the students.

For one implementation model of a client-based capstone, the reader is referred to Harvey Mudd College (2015). This handbook contains a complete set of guidelines for students, including advice for professional and effective communication with clients, workflow management, technical writing, and protecting client confidentiality.

In a departure from the client-focused model, some schools have research-based individual capstone experiences. Leman et al. (2015) described a semester-long research-focused capstone experience at Virginia Tech, similar to a thesis experience. However, advising such research experiences is time intensive for the faculty member. For instance, until 2015, Duke University expected students to complete their capstone research project individually, working directly with a faculty

member, with the goal being publishable research. Their students worked on the project over two semesters, and met regularly as a class to make presentations and discuss their progress. Each faculty member invested 1–1.5 hr per week supervising each student (Malone et al. 2013). Due to the difficulty of advising a large number of students in this manner, Duke now offers individual research as an elective in the major, and instead requires students to take a one-semester course in statistical consulting for their capstone experience (Duke University 2016).

Although we have used the categorization “standalone capstone project” to differentiate these projects from capstone experiences occurring primarily in a classroom environment, many schools couple the standalone capstone project with occasional lectures, workshops, and seminars on professional skill development. These occasional classroom activities help students learn important skills needed to manage long-term, client-focused projects. At the University of Georgia, the capstone is a year-long course with the fall semester focused on learning advanced material, hearing client presentations, and developing professional skills; the spring semester is focused on the project work (Lazar et al. 2011). St. Olaf College offers a Research Skills Seminar Series that includes statistical consulting communication skill development, peer-reviewed mock consulting practice, research record-keeping, career preparation, and community building (Legler et al. 2010). They and the University of Miami also include discussions on the ethical practice of statistics (Smucker and Bailer 2015). At Harvey Mudd College, we have developed short, improvisational skits in which the faculty advisors role play different team scenarios (such as a teleconference or site visit with the client) and ask the student audience to critique the behaviors of the characters in the skit.

3.2. Statistical Consultancy

In lieu of a standalone capstone project, many universities treat student participation in on-campus statistical consulting centers as a capstone experience. In such consultancies, clients (such as students or faculty at the same university, or local nonprofit organizations) approach the center with research questions requiring statistical advice ranging from experimental design to model selection to data analysis. The students work on these questions for varying periods of time, from short, on-the-spot consultations to projects lasting several months. New models for statistical consultancies recommend *collaborations*, in which the statistician is fully involved in the project from start to finish rather than serving only as an expert for hire (Davidian 2013; Vance 2015). Collaborations offer students even greater opportunities to synthesize knowledge across disciplines and statistical methodologies; participate in experimental design, data collection, and data cleaning; and develop communication skills through interactions with their collaborators.

At some universities, student consultants are paid to staff the consultancy (Kim et al. 2014), whereas at others, they receive academic credit (Jersky 2002; Mackisack and Petocz 2002). Winona State University has a hybrid approach, where students enroll in a statistical consulting course in the fall, and top

students in that course are invited to work as paid consultants in the spring (Hooks and Malone 2012). Purdue University's STATCOM (Statistics in the Community), is a pro bono statistical consulting service that is fully run by students in the statistics department (Gunaratna et al. 2007; STATCOM 2013). While participation in STATCOM is voluntary and uncompensated, it is viewed as a valuable co-curricular activity that prepares students well for statistical practice.

As in the case of the standalone capstone project, students working in consultancies often receive some classroom training in statistical consulting to address effective communication techniques (including writing reports), project management, and ethics in statistical practice. Videos, role-playing, and mock consulting sessions can provide feedback to students on how well they ask questions of and listen to the needs of the client (Legler et al. 2010; Hooks and Malone 2012; Vance 2015). Additionally, the classroom component can serve to teach students advanced statistical methodology on an as-needed basis as the projects require.

Students participating in consultancies are not always required to be statistics majors, as long as they have completed the necessary prerequisite course work. The prerequisite course could be as simple as introductory statistics (Jersky 2002), with some schools also requiring an additional course in regression (Hooks and Malone 2012). Jersky (2002) reported benefits to heterogeneity of the student consultants including nonmajors developing statistical skills which help them get jobs, early statistics majors getting motivated by the projects and invested in their education, and senior statistics majors deepening their understanding of statistics by mentoring the younger students.

Student consultants often work in teams of two to four students directed by a faculty member. At Penn State, the two students on a team are designated as either the communication lead, responsible for all interactions with the client, or the technical lead, responsible for the accuracy of the statistical work (Penn State University 2016). At Truman State University, teams of four students are led by a faculty advisor and a "senior consultant," who is a student having at least 1 year of consulting experience; the senior consultant is responsible for project management and mentoring the other team members (Kim et al. 2014).

One challenge of requiring statistical consulting as a capstone experience is recruiting a sufficient number of clients having problems of an appropriate level of difficulty and time scope, a task which typically falls on the instructor (Jersky 2002; Mackisack and Petocz 2002; Kim et al. 2014). Hooks and Malone (2012) reported an added challenge to be ensuring a constant demand for projects over the semester, which can be mitigated by soliciting projects in advance if the clients' needs are not time sensitive. In addition to recruiting projects and supervising the students, the instructor or supervisor has many other administrative duties. These include handling the accounting if the client pays for the consulting services, ensuring that the students adhere to any confidentiality requirements, and assisting with the writing of the final report. In some cases, the faculty advisor might even take over the project if the required statistical methodology is too advanced or if political obstacles arise (such as the client making unreasonable demands on the students' time) that the students are not

equipped to handle (Jersky 2002; Mackisack and Petocz 2002; Hooks and Malone 2012; Kim et al. 2014). It is this latter responsibility that distinguishes statistical consultancies from other capstone experiences found in the literature: When a university establishes a statistical consultancy center, clients expect professional caliber statistical expertise, and the faculty supervisor is often on the hook to provide this.

3.3. Project Within a Statistical Methodology Course

In many instances found in the literature, a capstone-like experience is embedded in a statistical methodology course, such as linear models or time series analysis. Students complete projects designed to vertically integrate the statistical process but that are specifically relevant to the particular methodology being studied. Although such projects lack the important step of choosing which methodology to use, they typically involve data collection and cleaning, model development and validation, analysis and communication of results. In some instances, the students may also be responsible for experimental design. Implementation of such a project can be at the sole discretion of the instructor, or there can be a department-wide commitment to including applied courses throughout the curriculum. For example, at Macalester College, certain upper-level statistics courses that require a statistical project are designated as capstone courses that fulfill the capstone requirement of the major (Malone et al. 2013).

Because such projects are embedded within a methodology course, they typically occur toward the end of a course, after the students have learned sufficient methodology to complete the project. In the statistical capstone at North Carolina State University, a project proposed by the instructor is completed during two or three sessions of class time at the end of a course and requires no work outside of class. Small group discussion culminates in full-class consensus on the experimental design. Next, each group independently conducts the short experiment. Data analysis can be done by each group or by the whole class using data aggregated over all of the teams (McGowan and Woodard 2010). Khachatryan (2015) assigned teams of students to month-long consulting cases toward the end of a time series course. The teams meet weekly or bi-weekly with the instructor, who acts as the client on a business problem. To promote effective teamwork, each student has a prescribed role on the team (project manager, data manager, or data analyst) that interacts in a prescribed way with the other roles. Additionally, the students are told that each student on a team will receive the same grade for the final report. Teams complete a report and give an oral presentation, as well as complete peer evaluations of their teammates. Although these examples above fail to satisfy the "long-term" criterion of a typical capstone project, course projects could be introduced in stages to the students throughout the course. For instance, in the early part of the semester, students could be asked to identify a dataset and perform exploratory data analysis. In the middle of the course, students could apply simple methodology that has already been covered in the course. In the final weeks of the semester, students could expand their model to include more sophisticated methodology and present their results before the class.

There are several advantages of structuring a capstone experience within an existing methodology course. The first is that students can immediately learn how to put to use the methodology seen in the course, and they do not have to wait until their senior year to do so. This can help students get jobs or be accepted to graduate school, as they will have a significant experience with data analysis that they can refer to in their applications. Next, it is an efficient way to give students practice with hands-on data analysis without demanding too much faculty time. For instance, the capstone courses at Macalester replaced an earlier thesis capstone wherein students worked on research individually with a faculty advisor (Malone et al. 2013).

One disadvantage of designating a course project as a capstone is that the projects are limited to the particular methodology used in the course. This prevents students from synthesizing knowledge across many fields of statistics, and reduces their practice of method selection. Another concern is that such a course project might not be viewed as a genuine *capstone* experience to be taken seriously by the students. However, there are ways to mitigate this. For instance, at Carnegie Mellon University, students can elect to go beyond such a capstone project by completing an Honors thesis. At the University of Georgia, students are expected to communicate directly with the client, so they take the project more seriously. And at Macalaster College, students present their course projects at Capstone Day, a college-wide event in April devoted to student presentations (Malone et al. 2013).

3.4. Instructional Capstone Course

The last category of capstone experiences we consider in-depth is an instructional capstone course. Although the previous examples of a dedicated capstone project, a statistical consultancy, and a capstone project within a statistical methodology course included instructional components, we distinguish an instructional capstone course as one in which the primary activity is instruction in elements of statistical practice. Such a course involves regular class meetings, and students typically complete readings or small projects but not an in-depth, hands-on project with a client. Students are taught strategies for effective communication and project management. We outline the structure of a few such courses we found in the literature.

Spurrier (2001) described a one-unit semester-long capstone course at the University of South Carolina which tries to prepare senior students to work as professional statisticians. Class time is spent on lessons on nonstatistical skills such as oral communication, writing (and receiving feedback on writing), and career advice. Additionally, students conduct a series of short experiments over the semester, ranging in complexity, in which students need to apply multiple statistical skills. Most of these experiments involve collection of real data and working in teams. An important component of the course is class discussion at the end of each experiment both to summarize lessons learned and to expose the students to the other approaches used by their classmates. The projects are drawn from a textbook, *The Practice of Statistics: Putting the Pieces Together* (Spurrier 2000), and while fictitious in context, they promote active data collection and analysis. The book also

contains chapters on making presentations, ethical decision making, and techniques for getting a job in statistics (Dobler 2003).

Capstone courses can also be used in an online course setting, for example, by the instructor drawing on completed client projects from their own consulting experiences as the basis for student training in consulting skills. Young et al. (2015) described a two-semester Statistical Consulting Practicum capstone course in an online professional Masters of Applied Statistics program at the Pennsylvania State University. When the program was initially developed, the intention was that the online students would pursue projects at their place of employment where a supervisor or colleague could serve as the client. However, that has not proven possible in practice, so instead, the online students typically analyze previously completed projects from the campus Statistical Consulting Center. The faculty instructor serves as the “client” in summarizing the problem, and the students must question the instructor to get a complete understanding of the problem from which they can perform the data analysis, write a report, and give an oral presentation. Although the students do not generally deal directly with real data, they get practice in communicating with a client to develop their understanding of the problem context.

Other schools find it beneficial to introduce students to applied statistics skills earlier in their education. Taplin (2003) described an introduction to statistical consulting course at Murdoch University in Australia whose prerequisite background is only an introductory statistics course. The course attracts second- or third-year undergraduates, approximately half of which are Math and Statistics majors and half are majors in fields where statistics is commonly used (e.g., biological sciences and social sciences). In the course, students watch and discuss videotaped consulting sessions and do role playing to learn about effective questioning of and communication with clients. They learn how to interact with different personality types, formulate and solve problems, and write reports. The instructors emphasize the importance of reflecting on one’s experiences as a consultant. Students have the opportunity to practice their skills on “clients,” who are students in an introductory statistics course. The purpose of exposing students to statistical practice early is both to attract students into the statistics major as well as provide the context of statistical practice before they learn statistics methodology.

The advantages of such courses are that they do not require more faculty time than a typical statistics course, and they rely on short, predetermined projects which can be reused from year to year. An important disadvantage is that they lack the authenticity that comes from students interfacing directly with clients and collecting and cleaning messy, real-world data.

3.5. Other Considerations

In this section, we outline some additional considerations and challenges that are common to all types of statistical capstones.

3.5.1. Summer Research Experiences

Of course, the four types of capstones described above are not the only experiences that can provide students the opportunity to synthesize statistical knowledge and solve real-world

problems. For instance, summer research programs or internships offer an immersive experience in which the statistical project is the student's primary responsibility for 8 to 12 weeks during the summer. A variant of this is the Explorations in Statistics Research program at the University of California at Berkeley (Nolan and Temple Lang 2015). In this 1-week summer program, undergraduates work on three 2-day projects in groups of three and are advised by faculty and graduate students. The brief length does not permit deep exposure to any one statistical method or solution approach, but does allow students to experience first hand a breadth of approaches. Moreover, the organizers feel the short format allows students a low-stakes opportunity to explore their interest in statistics before choosing a major or deciding to go to graduate school.

A common source of funding for summer research experiences for undergraduates is the National Science Foundation (NSF). Through their Research Experience for Undergraduates (REU) program, colleges and universities can receive funding to establish a site for several years, and solicit applications from undergraduates across the country. At an REU site, faculty supervise one or more research projects with undergraduates and typically offer mentoring and professional development activities. A full listing of current NSF REU sites in the mathematical sciences, including statistics, is available on the NSF website (National Science Foundation, 2016). The American Statistical Association is itself a recipient of such a grant from 2016 to 2018 to offer REUs in statistics, specifically (American Statistical Association, 2016).

Although such experiences occur outside the regular academic year, a department could choose to allow completion of a summer research experience or internship to satisfy a departmental capstone requirement. In such cases, the department might ask the student to write a report or give a presentation about their work.

3.5.2. Deliverables and Evaluation

Capstone projects are evaluated in many ways. Several articles in the literature describe intermediate writing assignments, such as writing an abstract or a project proposal, to provide students ample opportunity to get feedback and develop their writing skills (Legler et al. 2010; Lazar et al. 2011; Smucker and Bailer 2015). Other deliverables include poster presentations, oral presentations to the client, a final written report, a research log, and peer evaluations. Several articles provide rubrics for grading these deliverables more uniformly (Spurrier 2001; Moore and Kaplan 2015; Smucker and Bailer 2015). Both Macalester College and Harvey Mudd College celebrate the end of the academic year with a day devoted to student project presentations open to the public and clients (Borrelli 2010; Malone et al. 2013).

3.5.3. Faculty Engagement

Faculty teaching credit for supervising capstone projects varies. At St. Olaf College, neither faculty nor students receive credit for participating in the standalone capstone projects (Legler et al. 2010). At the University of Georgia, two senior faculty and one graduate teaching assistant are assigned to supervise all of the capstone teams; they each receive a full semester of course credit for the two-semester commitment. Each

instructor supervises two to three groups of two or three students each, and the teaching assistant provides floating supervision to all teams. At Harvey Mudd College, supervising a capstone clinic team for a year is equivalent to teaching one course. If a particular expertise is not covered by the faculty, adjunct faculty are hired to meet the capstone demand. In a statistical consulting center, sometimes the director of the center receives no teaching credit or compensation beyond the course itself (Hooks and Malone 2012).

3.5.4. Software and Technical Training

Common to all applied statistics projects in the modern era is the use of statistical software. Moreover, employers are looking for students who have experience with statistical software and database management, so it is a useful skill for students to learn (Ritter et al. 2001). By far, the most commonly mentioned statistical tool in the literature is the open-source programming language, R (see, e.g., Lazar et al. 2011; Kim et al. 2014; Nolan and Temple Lang 2015; Smucker and Bailer 2015). The advantages of R are that it is freely available, versatile, and widely used in industry and academia. Moreover, while R can require a steep learning curve, there are tools available that make it more user friendly to novices. RStudio provides several different views of the user's workspace and autofills suggestions for function syntax. Knitr and R Markdown are compilers that can embed R commands in HTML and other document types so that statistical analyses can easily be displayed on webpages. In addition to R, traditional commercial statistical software such as SPSS, Minitab, SAS, S+, and Matlab can also be used (Lazar et al. 2011; Kim et al. 2014).

3.5.5. Challenges

Implementing a statistical capstone is not without its challenges, regardless of the category. Sometimes, a gap exists between what students learned in their prior courses and what they need to do in the capstone, particularly in the areas of programming and advanced statistical methodology (Mackisack and Petocz 2002; Smucker and Bailer 2015). Young et al. (2015) proposed using more case studies in their required and elective courses to close this gap.

Another challenge that is common for any team project is teaching good interpersonal skills (including with the client) and effective leadership (Jersky 2002; Mackisack and Petocz 2002). Rangecroft and Wallace (1998) found that student projects are most effective when there is a clear leader on the team. To encourage all students to contribute equally to a project, Smucker and Bailer (2015) suggested the use of peer evaluations and rotating the teams periodically. They also pointed out that students are unaccustomed to managing multiple projects at the same time within a single class.

Challenges can also arise on the client's side if the client is disinterested, does not make themselves available for regular communication, or does not provide data in a timely fashion (Mackisack and Petocz 2002; Legler et al. 2010). Charging even a modest fee can help keep the client engaged (Haines 1988). Some schools charge fees to cover the operating costs of the consulting center (Rangecroft and Wallace 1998; Jersky 2002), and Harvey Mudd College charges a fixed fee to cover project

costs, travel by the students to the client site, software, administrative salaries, and institutional overhead (Borrelli 2010).

The final challenge is the perennial problem of time constraints. As the popularity of statistics and data science explodes, class sizes increase, and it becomes harder for faculty to supervise these sorts of projects (Nugent 2012). For students, placing a capstone project in the senior year can be overwhelming in combination with job searches and the graduate school application process. For this reason, Macalester College moved its capstone requirement to a statistical methodology course that need not be taken in the student's senior year (Malone et al. 2013).

4. Conclusions

As demand for data scientists and statisticians continues to increase, the need to prepare undergraduates to tackle real-world problems will be ever present. Projects that require students to synthesize knowledge across disciplines to analyze real datasets (and likely *very large* datasets) and then communicate the modeling process and results with clients equip them with not just the "why" of mathematical statistics, but also the "how" of applied statistics. Despite the challenges mentioned in the previous section, statistical capstones are likely here to stay. By surveying the literature on statistical capstone experiences, we hope this article has convinced the reader of the value of such experiences, and provided a starting point for considering the design of a new capstone.

Acknowledgments

We would like to thank the editor, associate editor, and two anonymous referees for their helpful feedback on this manuscript.

References

- American Statistical Association (2014), "Curriculum Guidelines for Undergraduate Programs in Statistical Science," [online]. Available at <http://www.amstat.org/education/pdfs/guidelines2014-11-15.pdf>
- American Statistical Association (2016), "ASA Receives Grant to Establish Series of REUs," [online]. Available at <http://magazine.amstat.org/blog/2016/04/01/asareu16/>
- Berenson, M. L., Utts, J., Kinard, K. A., Rumsey, D. J., Jones, A., and Gaines, L. M. (2008), "Assessing Student Retention of Essential Statistical Ideas," *The American Statistician*, 62, 54–61.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (1956), *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*, New York: David McKay Company.
- Borrelli, R. (2010), "The Doctor Is In," *Notices of the American Mathematical Society*, 57, 1127–1130.
- Brown, E. N., and Kass, R. E. (2009), "What Is Statistics?," *The American Statistician*, 63, 105–110.
- 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.
- Cohen, S. (2014), "Ethics for Undergraduates: Workgroup on Undergraduate Education," *American Statistical Association* [online]. Available at <https://www.amstat.org/education/pdfs/EthicsUndergraduates.pdf>
- Committee on Professional Ethics (1999), "Ethical Guidelines for Statistical Practice," *American Statistical Association* [online]. Available at <http://www.amstat.org/about/ethicalguidelines.cfm>
- Davidian, M. (2013), "The International Year of Statistics: A Celebration and A Call to Action," *Journal of the American Statistical Association*, 108, 1141–1146.
- Dobler, C. P. (2003), "The Practice of Statistics: Putting the Pieces Together Book Review," *The American Statistician*, 57, 142–143.
- Duke University (2016), "The Major in Statistical Science," *Department of Statistical Science* [online]. Available at <http://stat.duke.edu/undergraduate-program/major-statistical-science>
- Durel, R. J. (1993), "The Capstone Course: A Right of Passage," *Teaching Sociology*, 21, 223–225.
- Gibbs, A., and Reid, N. (2009), "Comment," *The American Statistician*, 63, 112–113.
- Gorman, M. F. (2011), "Student Reactions to the Field Consulting Capstone Course in Operations Management at the University of Dayton," *Interfaces*, 41, 564–577.
- Gunaratna, N. S., Johnson, C. A., and Stevens, J. R. (2007), "Service-Learning for Graduate Students Through a Student-Run Consulting Program," *Journal of Statistics Education*, 15, 1–21.
- Haines, G. H. (1988), "The Ombudsman: Teaching Entrepreneurship," *Interfaces*, 18(5), 23–30.
- Hakeem, S. A. (2001), "Effect of Experiential Learning in Business Statistics," *Journal of Education for Business*, 77(2), 95–98.
- Hardin, J., Hoerl, R., Horton, N. J., Nolan, D., Baumer, B., Hall-Holt, O., Murrell, P., Peng, R., Roback, P., Temple Lang, D., and Ward, M. D. (2015), "Data Science in Statistics Curricula: Preparing Students To 'Think with Data'," *The American Statistician*, 69, 343–353.
- Harvey Mudd College (2015), "Mathematics Clinic Handbook," *Mathematics Clinic* [online]. Available at <http://www.math.hmc.edu/clinic/handbook>
- Higgins, J. J. (1999), "Nonmathematical Statistics: A New Direction for the Undergraduate Discipline," *The American Statistician*, 53, 1–6.
- Hillmer, S. C. (1996), "A Problem-Solving Approach to Teaching Business Statistics," *The American Statistician*, 50, 249–256.
- Hooks, T., and Malone, C. (2012), "Involving Undergraduates in Statistical Consulting," CAUSEWeb [online]. Available at https://www.causeweb.org/cause/sites/default/files/ecots/ecots12/posters/hooks_slides.pdf
- Horton, N. J. (2015), "Challenges and Opportunities for Statistics and Statistical Education: Looking Back, Looking Forward," *The American Statistician*, 69, 138–145.
- Horton, N. J., and Hardin, J. S. (2015), "Teaching the Next Generation of Statisticians to 'Think with Data': Special Issue on Statistics and the Undergraduate Curriculum," *The American Statistician*, 69, 259–265.
- Jersky, B. (2002), "Statistical Consulting with Undergraduates—A Community Outreach Approach," in *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6)*, Cape Town, South Africa. [online]. Available at: http://icots6.haifa.ac.il/PAPERS/3A1_JERS.PDF
- Johnson, H. D., and Dasgupta, N. (2005), "Traditional versus Non-traditional Teaching: Perspectives of Students in Introductory Statistics Classes," *Journal of Statistics Education*, [online], 13, 2.
- Keeler, C. M., and Steinhorst, R. K. (1995), "Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report from the Field," *Journal of Statistics Education*, [online], 3, 2.
- Khachatryan, D. (2015), "Incorporating Statistical Consulting Case Studies in Introductory Time Series Courses," *The American Statistician*, 69, 387–396.
- Kim, H.-J., Alberts, K. S., and Thatcher, S. (2014), "Teaching Undergraduates Through Statistical Consulting," in *Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9)*, Flagstaff, USA. [online]. Available at: http://iase-web.org/icots/9/proceedings/pdfs/ICOTS9_C195_KIM.pdf
- Kvam, P. H. (2000), "The Effect of Active Learning Methods on Student Retention in Engineering Statistics," *The American Statistician*, 54, 136–140.
- Landes, R. D. (2009), "Passing on the Passion for the Profession," *The American Statistician*, 63, 163–172.
- Lazar, N. A., Reeves, J., and Franklin, C. (2011), "A Capstone Course for Undergraduate Statistics Majors," *The American Statistician*, 65, 183–189.
- 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.
- Leman, S., House, L., and Hoegh, A. (2015), "Developing a New Interdisciplinary Computational Analytics Undergraduate Program: A

- Qualitative-Quantitative-Qualitative Approach,” *The American Statistician*, 69, 397–408.
- Mackisack, M., and Petocz, P. (2002), “Projects for Advanced Undergraduates,” in *Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6)*, Cape Town, South Africa.
- Malone, C., Kaplan, D., Lazar, N., Nugent, R., and Stangl, D. (2013), “The Role and Variety of Undergraduate Statistics Capstones,” *Guidelines for Undergraduate Statistics Programs Webinar Series* [online]. Available at http://www.amstat.org/education/webinars/presentation_slides/Capstones.pptx
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. (2011), “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey & Company [online]. Available at http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
- Martonosi, S. E. (2012), “Project-Based ORMS Education,” in *Wiley Encyclopedia of Operations Research and Management Science*, ed. J. J. Cochran, Hoboken, NJ: Wiley; doi:10.1002/9780470400531.
- McGowan, H., and Woodard, R. (2010), “Supporting Statistical Thinking Through a Capstone Project,” *CAUSEWeb* [online]. Available at <https://www.causeweb.org/cause/webinar/activity/2010-07>
- Miami University (2016), [cited September 22, 2016], “Majors,” *Department of Statistics* [online]. Available at <http://www.miamioh.edu/cas/academics/departments/statistics/academics/majors/index.html>
- Moore, A. A., and Kaplan, J. J. (2015), “Program Assessment for an Undergraduate Statistics Major,” *The American Statistician*, 69, 417–424.
- National Science Foundation (2016), “REU Sites,” [online]. Available at http://www.nsf.gov/crssprgm/reu/list_result.jsp?unitid=5044
- Nolan, D., and Temple Lang, D. (2009), “Comment,” *The American Statistician*, 62, 117–121.
- (2015), “Explorations in Statistics Research: An Approach to Expose Undergraduates to Authentic Data Analysis,” *The American Statistician*, 69, 292–299.
- Nugent, R. (2012), “Maintaining Quality in the Face of Rapid Program Expansion,” *AMSTAT News* [online]. Available at http://magazine.amstat.org/blog/2012/08/01/carnegiemellon7_12/
- Penn State University (2016), “The Statistical Consulting Online Handbook,” *Department of Statistics* [online]. Available at <https://onlinenecourses.science.psu.edu/stat580/process>
- Rangecroft, M., and Wallace, W. (1998), “Group Consultancy, As Easy As Falling Off a Bicycle?,” in *Proceedings of the Fifth International Conference on Teaching Statistics (ICOTS5)*, Singapore. [online]. Available at: https://www.researchgate.net/publication/242222386_GROUP_CONSULTANCY_AS_EASY_AS_FALLING_OFF_A_BICYCLE.
- Ritter, M. A., Starbuck, R. R., and Hogg, R. V. (2001), “Advice from Prospective Employers on Training BS Statisticians,” *The American Statistician*, 55, 14–18.
- Smucker, B. J., and Bailer, A. J. (2015), “Beyond Normal: Preparing Undergraduates for the Work Force in a Statistical Consulting Capstone,” *The American Statistician*, 69, 300–306.
- Spurrier, J. D. (2000), *The Practice of Statistics: Putting the Pieces Together*, Belmont, CA: Duxbury Press.
- Spurrier, J. D. (2001), “A Capstone Course for Undergraduate Statistics Majors,” *Journal of Statistics Education*, [online] 9, 1.
- STATCOM (2013), “Intro to STATCOM Network,” *American Statistical Association*, [online]. Available at <http://www.amstat.org/education/statcom/index.html>
- Taplin, R. H. (2003), “Teaching Statistical Consulting Before Statistical Methodology,” *Australian and New Zealand Journal of Statistics*, 45, 141–152.
- 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*, [online], 10, 2.
- Vance, E. A. (2015), “Recent Developments and Their Implications for the Future of Academic Statistical Consulting Centers,” *The American Statistician*, 69, 127–137.
- Wasserstein, R. (2015), “Communicating the Power and Impact of Our Profession: A Heads up for the Next Executive Directors of the ASA,” *The American Statistician*, 69, 96–99.
- Wild, C., and Pfannkuch, M. (1999), “Statistical Thinking in Empirical Enquiry,” *International Statistical Review*, 67, 223–265.
- Young, D. S., Johnson, G. F., Chow, M., and Rosenberger, J. L. (2015), “The Challenges in Developing an Online Applied Statistics Program: Lessons Learned at Penn State University,” *The American Statistician*, 69, 213–220.