

Integrative Biological Chemistry Program Includes The Use Of Informatics Tools, GIS And SAS Software Applications

Malcolm J. D'Souza, Ph.D., Wesley College, USA

Richard J. Kashmar, Ph.D., Wesley College, USA

Kent Hurst, Ph.D., Wesley College, USA

Frank Fiedler, Ph.D., Wesley College, USA

Catherine E. Gross, Wesley College, USA

Jasbir K. Deol, Wesley College, USA

Alora Wilson, Wesley College, USA

ABSTRACT

Wesley College is a private, primarily undergraduate minority-serving institution located in the historic district of Dover, Delaware (DE). The College recently revised its baccalaureate biological chemistry program requirements to include a one-semester Physical Chemistry for the Life Sciences course and project-based experiential learning courses using instrumentation, data-collection, data-storage, statistical-modeling analysis, visualization, and computational techniques. In this revised curriculum, students begin with a traditional set of biology, chemistry, physics, and mathematics major core-requirements, a geographic information systems (GIS) course, a choice of an instrumental analysis course or a statistical analysis systems (SAS) programming course, and then, students can add major-electives that further add depth and value to their future post-graduate specialty areas. Open-sourced georeferenced census, health and health disparity data were coupled with GIS and SAS tools, in a public health surveillance system project, based on US county zip-codes, to develop use-cases for chronic adult obesity where income, poverty status, health insurance coverage, education, and age were categorical variables. Across the 48 contiguous states, obesity rates are found to be directly proportional to high poverty and inversely proportional to median income and educational achievement. For the State of Delaware, age and educational attainment were found to be limiting obesity risk-factors in its adult population. Furthermore, the 2004-2010 obesity trends showed that for two of the less densely populated Delaware counties; Sussex and Kent, the rates of adult obesity were found to be progressing at much higher proportions when compared to the national average.

Keywords: Biological Chemistry; Wesley College; GIS; SAS; Big Data; NIH NIGMS IDeA program; NSF EPSCoR program; NSF S-STEM program; Cannon Scholar program; DE-INBRE; DE-EPSCoR

INTRODUCTION

The American Chemical Society (ACS) defines biological chemistry programs as crossover fields that encompass the understanding of the chemical processes occurring in living systems. Biochemistry or biological chemistry focuses on the study of large bio-macromolecular polymers, while chemical biology utilizes chemistry (organic, physical, and analytical chemistry) tools to evaluate the molecular basis of biological problems. Both programs are inherently interdisciplinary (Cravatt and Gottesfeld, 2010) and require comprehensive training through a broad spectrum of courses in chemistry, biology, physics, computer science, and mathematics. The American Society for Biochemistry and Molecular Biology (ASBMB) lists 77 career options for students trained in the biological sciences (biological chemistry and chemical biology). The ASBMB's career brochure and the ACS biochemistry career website effectively outline the requirements and possible career roadmaps to manage challenges and motivate undergraduates to these fields.

In 2003, in the BIO 2010 Report, the National Research Council (NRC) made several recommendations to reform, strengthen and integrate the biology, chemistry, physics, mathematics, and computer science coursework taken at the undergraduate level in any of the biological sciences programs. In response new undergraduate curricular materials have been developed and implemented (Rylands et al., 2013; Jungck et al., 2010; Matthews et al., 2010; Pursell, 2009; Owen and Breyer, 2005; Steen, 2005; Slonczewski and Marusak, 2004).

The NRC BIO 2010 Report and other like-minded education models (Smith and Dragojlovic, 2013; Brownell et al. 2012; Canaria et al. 2012; Thiry et al. 2011) strongly encourage the exposure of undergraduate students to independent research and experiential forms of learning that promote the oral and written presentation of scientific results. Higher education literature on pedagogical practices has firmly demonstrated that actively engaging students in undergraduate research (Chang et al., 2014; Finley and McNair, 2013; Tinto, 2012; D'Souza and Wang, 2012; D'Souza et al. 2011; Kuh, 2008) is a high-impact method for adding value and achieving excellence in liberal-arts education. Additionally in recent years, hands-on student-centered project oriented teaching using Informatics tools (Hersh et al. 2014; Longenecker et al., 2012; D'Souza and Gerges, 2010; Wold, 1995), Geographical Information Systems (GIS) spatial analysis programs (Kolvoord et al., 2014; Lee and Wong, 2001; Malczewski, 1999) and Statistical Analysis System (SAS) training (Hatcher and O'Rourke, 2014; Sall et al., 2012; Dowdy et al., 2011) has become a subject of considerable interest.

In the biological chemistry related fields, the United States (US) Department of Labor Bureau of Labor Statistics published reports that the employment outlook for its graduates is projected to grow 19% faster than the average for all occupations from 2012 to 2022. The 2013 median annual wages for new graduates from such majors was reported to be \$91,640. Significant improvements in job prospects in these fields are being realized along the East Cost (Ainsworth, 2012; Fiegerman, 2011). Furthermore, the adoption and application of information technology tools will drive future approaches in the clinical, bioenvironmental, biomedical, and public health fields (Hwang and Curl, 2014; Barh et al., 2013; Matthews-Juarez, 2013; Kesh and Ramanujan, 2010).

Wesley College (Wesley) is a primarily undergraduate institution that is fully accredited by the Middle States Commission on Higher Education (MSCHE). The 2013 total academic year (AY) student population was 1600 students. The College's Departments of Science and Mathematics (STEM) offer baccalaureate programs in biology, biological chemistry, environmental science, environmental policy, medical technology, and mathematics. In AY 2013 there were 110 STEM majors. Housed in Cannon Hall, the STEM programs provide a learning environment where hands-on experience through undergraduate research is continually incorporated (D'Souza and Kroen et al., 2015; D'Souza and Wang, 2012; D'Souza et al., 2011).

Delaware is also home to an increasing number of financial institutions, insurance, chemical, pharmaceutical, and medicinal (biotech) industries. Employment concentrations in the financial fields began with the 1981 Financial Center Development Act and today, the health care industry and the financial service industry are two of the prominent private-sector employers in Delaware. The Delaware Department of Labor reports that 18 of the 25 fastest growing occupations are jobs in the scientific and technical workforce. According to the Delaware Economic Development Office 2012 Data Book, "Delaware ranks first in industry investment in research and development, and has the fourth highest concentration of scientists and engineers in the United States."

In addition, the Delaware chemistry and biochemistry programs are also identified as one of the top five most common educational areas for green technology occupations said to be emerging in the State (Brown and Ratledge, 2012). The new employment demands in such applied science occupations and the aging workforce create a critical need for college graduates trained in science and technology to fill these jobs (DE-STEM Council Report, 2012-2013). Also, the 2013 Beyond Borders Biotechnology Industry Report showed a pronounced regional increase in healthcare, biochemist, biotech, bioenvironmental science, chemistry, database administrator, statisticians, and pharmaceutical job postings, and a June 4, 2013 Wilmington News Journal article, indicated that the highest proportion of these jobs as a fraction of job openings through 2018 is in the Delmarva (Delaware, Maryland, Virginia) region. The occupations just listed, for the most part, are also technical areas of great national need and all of the aforementioned industries rely heavily on data mining and analysis.

From the time that John Snow solved the source of the London cholera outbreak in 1854, scientists and practitioners in various fields have recognized the power of geographic data visualization (Brundson and Comber,

2015; Koch, 2004). With the evolution of cheaper, more powerful, and more distributed computing platforms, geographic information systems technology has taken major steps forward in its accessibility, functionality, and classroom utility. Additionally data intensive scientific analyses (Park and VanRoekel, 2013; Ailamaki et al., 2010) that merge visualization and workflow technologies have become especially important aspects to sort large quantities of available biological data; Big Data (United States Census Bureau, University of Wisconsin Population Health Institute). Mature spatial analysis GIS applications on the Environmental Systems Research Institutes (ESRIs) ArcGIS suites and SAS techniques developed by the SAS Institute are found to be useful in integrating optimal search strategies and are powerful evidence-based practice tools in the domain of public health (Ogden et al., 2014; Duncan et al., 2014; Musa et al., 2013; Hawthorne and Kwan, 2012; Moore et al., 2008; Ricketts, 2003; Lee and Wong, 2001; Clarke et al., 1996).

For example, in the US, the overweight and obesity epidemic has been dramatic and has caused substantial economic impacts (Imes and Burke, 2014; Ogden et al. 2014; Slack et al., 2014). Researchers have used GIS-visualization mapping technology to evaluate locational factors, environmental attributes, and trends in both childhood (Duncan et al., 2014; Wridt, 2010; Lytle, 2009) and adult obesity rates (Clift et al., 2014; Giang et al., 2008). GIS technologies when coupled with SAS tools provide transformative high-performance advanced analytics to sort through large volumes of obesity datasets to process and uncover key relationships that help forecast and refine predictive models (Ogden et al., 2014; Moore et al., 2008).

In Delaware, the adult and childhood overweight and obesity prevalence has also greatly affected healthcare within the State (Gupta, 2014; Chang et al., 2014; Xu et al., 2013). A recent Wesley College obesity study (D'Souza and Walls et al., 2015) determined that 29.5% of the Wesley undergraduates are overweight and 19.8% are obese. At the undergraduate level, research on both the causes and prevention of obesity are valuable and can be a catalyst for interdisciplinary collaborations.

THE WESLEY COLLEGE BIOLOGICAL CHEMISTRY PROGRAM

The National Science Foundation (NSF) and the National Institutes of Health (NIH) have developed special federal-state partnership programs for US states and territories that have historically received less than 15% of the total federal research and development (R&D) funding. The NSF Experimental Program to Stimulate Competitive Research (NSF-EPSCoR) program and the NIH Institutional Development Award (IDeA) from the National Institute of General Medical Sciences (NIGMS) program helps these states and territories establish partnerships with government, higher education, and industry partners, in order to develop their research bases and improve the quality of training to promote scientific progress at their colleges and universities. In Delaware, the lead institution on these federal awards is the University of Delaware (UD). Wesley College, Delaware State University (DSU), and the Delaware Technical and Community College (DelTech) System are the other Delaware higher-education institutions that partner with UD on the NSF-EPSCoR and NIH-NIGMS IDeA Networks of Biomedical Excellence (NIH-NIGMS-INBRE) programs. In Delaware the EPSCoR/INBRE programs are referred to as the DE-EPSCoR (IIA-1301765) and the DE-INBRE (P20GM103446) programs. The Christiana Care Health System (CCHS) and the Nemours/A.I duPont Hospital for Children are the medical partners on the DE-INBRE program.

In 2008, through impetus from the DE-INBRE program, the Wesley biological chemistry major was implemented to include a comprehensive undergraduate curriculum that emphasized interdisciplinary training in chemistry, biology, physics, and mathematics (D'Souza and Wang, 2012). At the time of implementation, the major's core-requirements included two physical chemistry (I & II) courses and an instrumental analysis course that the students took at DE-INBRE/DE-EPSCoR partner, DSU (D'Souza and Wang, 2012). A distinguishing feature of this program is the emphasis of undergraduate research experiences beginning in the second semester in the freshman year (D'Souza and Wang, 2012; D'Souza et al. 2011). In addition a senior-thesis capstone project is required for all biological chemistry majors. The curriculum provides students with the core knowledge of chemistry, biology, physics, calculus, and statistics courses that are strictly designed to develop analytical, creative, and quantitative-reasoning skills. Hence, the Wesley biological chemistry graduate acquires, articulates, retains, and applies skills and core concepts to effectively communicate, orally and in writing, key scientific findings in the molecular biosciences to lay and professional audiences. In AY 2013, there were 23 biological chemistry majors, which is approximately a fifth of the Wesley-STEM undergraduate population.

THE DIRECTED RESEARCH PROGRAM IN CHEMISTRY

At Wesley, all STEM majors are strongly encouraged to participate in meaningful undergraduate research projects in the DE-INBRE/DE-EPSCoR-sponsored mentored directed research program (D'Souza and Kroen et al., 2015; D'Souza and Wang, 2012; D'Souza et al., 2011), or as interns in industry, or in government internships. The undergraduates can participate in research as (paid) research assistants during the academic year or during an intensive (paid) 10-week internship program. They can choose to work at Wesley or at any partner institution. This affords flexibility and the rigorous preparation for the intensive training required for all graduate and pre-professional programs.

In the Wesley chemistry program, to include undergraduates in research and to develop collaborative partnerships between Wesley and DSU, a collaborative NSF-Major Research Instrumentation (NSF-MRI) grant was obtained (NSF 0520492) for the acquisition of a 300MHz Nuclear Magnetic Resonance (NMR) Spectrometer. As a result interaction amongst the faculty of the two chemistry departments greatly increased and there are additional collaborative joint grant proposal submissions to federal agencies. Wesley students complete undergraduate research projects in the chemistry, biology, and agricultural science departments at DSU, and DSU has undergraduates and graduates work in chemistry at Wesley (D'Souza and Wang, 2012).

Furthermore, the Wesley chemistry faculty obtained an NSF American Recovery and Reinvestment Act (NSF-ARRA) award (NSF 0960503) that provided funds to renovate three undergraduate research laboratories and to upgrade the cyber-infrastructure in Cannon Hall. The renovations and cyber-upgrades reinvigorated the sponsored mentored undergraduate directed-research program and increased the number of undergraduate participants by 40% (D'Souza and Kroen et al., 2015; D'Souza and Wang, 2012; D'Souza et al., 2011).

The Wesley chemistry undergraduate research projects have a two-fold focus: (1) small-molecule synthesis for use in chemometric methods for solution chemical kinetics (D'Souza and Kevill, 2014; D'Souza and Kevill, 2013; Kevill and D'Souza, 2008), and (2) cheminformatics (D'Souza and Barile et al., 2015; D'Souza et al. 2013; D'Souza et al., 2011; D'Souza and AlAbed, 2010; D'Souza and Gerges, 2010; D'Souza et al. 2009; D'Souza and Kyoshi, 2009; D'Souza et al., 2009; D'Souza, 2008; D'Souza, 2007; D'Souza, 2005).

Small-molecule mechanistic projects involving chemometric techniques remain of basic importance in organic chemistry because they study the bond-making and bond-breaking processes of a great many reactions. Such experiments have led to better methods for measuring and understanding the effects of the solvent (medium) in which these reactions take place (D'Souza and Kevill, 2014; D'Souza and Kevill, 2013; Kevill and D'Souza, 2008). Undergraduates trained in such techniques have then completed more sophisticated advanced molecular reaction dynamics and chemical kinetics summer internship projects in the Department of Chemistry & Biochemistry at UD.

The cheminformatics research program utilizes commercial informatics tools (D'Souza, 2008; D'Souza, 2007; D'Souza, 2005) for *in silico* methods in structure-activity relationship studies to teach emerging information technology that is typical in the drug discovery process (D'Souza et al. 2013; D'Souza et al., 2011; D'Souza and AlAbed, 2010; D'Souza and Gerges, 2010; D'Souza et al. 2009; D'Souza and Kyoshi, 2009; D'Souza et al., 2009). The Strategic Pharmaceutical Advisors (SRxA's) *Word on Health* used Wesley's drug development informatics findings (SRxA, 2010) to corroborate the criticism of the FDA from the Government Accountability Office (GAO). Furthermore, students developed a mobile application for smartphones that displayed an in-house designed pesticide database specifically created for the Delaware farming community (D'Souza and Barile et al. 2015). Wesley undergraduates also participate in innovative high-end computational biomedical or bioenvironmental projects at the Center for Bioinformatics & Computational Biology (CBCB) at UD.

Undergraduates involved in this chemistry research program are afforded an opportunity to publish and present data at scientific conferences. To date (March 2015), there have been 48 peer-reviewed publications with 57 undergraduate co-authors and over 250 national and regional conference presentations. Additionally, 74 undergraduates have earned awards and certificates from the American Chemical Society (ACS), the NASA DE-Space Grant program, and the Council of Undergraduate Research (CUR).

In May 2014 and October 2012, *International Innovations*, a magazine devoted to disseminating science, research, and technology information, featured this decade-long transformation that the Wesley undergraduate research program has driven, creating a dynamic and exciting environment where research drives undergraduate teaching for student benefit.

The commitment to effective Wesley practices that help to support and retain students and the successful implementation of multi-tiered mentoring in the directed research program resulted in an NSF Scholarships for STEM (NSF-STEM) award (NSF DUE 1355554 – Wesley College Cannon Scholar Program). The broadest impact of this Cannon Scholar program is an increased ability to seek out, attract, prepare, and graduate Wesley's financially challenged underrepresented minority STEM population.

REVAMP OF THE WESLEY BIOLOGICAL CHEMISTRY PROGRAM

Ten alumni from the biological chemistry program are successfully placed: one to Indiana University, School of Medicine; two to the University of Florida's Graduate School of Biomedical Sciences; one to the Salisbury University Occupational Therapy program; one to the Environmental Soil Chemistry graduate program at UD; one to the graduate Applied Chemistry program at DSU; one to the University of Maryland School of Dentistry; one to the University of Maryland Eastern Shore School of Pharmacy program; one to Temple University's School of Pharmacy; and one joined the Medical Research Unit at QPS, LLC, in Delaware.

The careers chosen by students trained in the Wesley biological chemistry program compelled the faculty to rethink the original two semester physical chemistry (I & II) course requirement (D'Souza and Wang, 2012). This year-long sequence includes a significant amount of material on quantum mechanics, molecular structure, spectroscopy, and statistical mechanics. Such in-depth molecular mechanics and quantum chemical calculations are not relevant for students who enter biomedical, pharmaceutical, or environmental programs, nor is it necessary for entering industries that require a heavier concentration in bio-related methodology and techniques.

Several other private liberal arts institutions in the mid-Atlantic already only require one semester of physical chemistry for students in their biological chemistry or molecular biology programs. Misericordia University, Cedar Crest College, Franklin and Marshall College, Gettysburg College, Moravian College, and Eastern University only require the first semester course of an already existing two-course sequence taken by chemistry majors (the semester covering thermodynamics and kinetics is required; the second semester, covering quantum mechanics, spectroscopy, and molecular structure, is not). McDaniel College, Ursinus College, St. Peter's University, King's College, and West Virginia Wesleyan College have developed their own one-semester course for students in the biological chemistry/molecular biology area. In addition, larger institutions, such as Boston College, Wesleyan University, New Mexico State University, and the University of Buffalo have also developed such one-semester courses. Therefore our replacement of a two-semester sequence with a one-semester course emphasizing those areas of physical chemistry most useful to biological chemistry majors is not uncommon. The 2015 revised program requirements for the Wesley College biological chemistry program are shown in Figure 1.

Figure 1. Wesley College Biological Chemistry program undergraduate degree requirements

<p>CORE CURRICULUM REQUIREMENTS (33 credits)</p> <p>Level One (9 credits) Quantitative Analysis [Calculus I and Statistics substitute for this requirement] Frontiers of Science [Biology I and Chemistry I substitute for this requirement] First-Year Seminar *College Writing *College Writing II</p> <p>Level Two (12 credits) Arts & Culture integrative course Literature and Language integrative course Philosophy & Religion integrative course History & Social Sciences integrative course</p> <p>Level Three (9 credits) – three courses in a defined concentration of Ethical Living or Identifying with Diversity or Personal & Social Responsibility *Level Four (3 credits of capstone course in major) – Research Methods (1 credit) + Experimental and Project Research (2 credits)</p> <p>*MAJOR REQUIREMENTS (69-70 Credit Hours) The Scientific Process (1 credit) Biology I (4 credits) Biology II (4 credits) Anatomy and Physiology II (4 credits) Microbiology (4 credits) Chemistry I (4 credits) Chemistry II (4 credits) Organic Chemistry I (4 credits) Organic Chemistry II (4 credits) Physical Chemistry (4 credits) Instrumental Analysis (4 credits) OR SAS Programming Course (3 credits) Analytical Chemistry (4 credits) Biochemistry (3 credits) Geographic Information Systems (3 credits) Applied Statistics (3 credits) Calculus I (4 credits) Calculus II (4 credits) Physics I (4 credits) Physics II (4 credits)</p> <p>*MAJOR ELECTIVES (12 Credit Hours) Anatomy and Physiology I (4 credits) Medical Microbiology (3 credits) Vertebrate Physiology (4 credits) Cell Biology (4 credits) Special Topic (3 or 4 credits) Animal Behavior (3 credits) Plant Biology (4 credits) Immunology (4 credits) Genetics (4 credits) Ecology (4 credits) Limnology (3 credits) Invertebrate Zoology (4 credits) Plant Physiology (4 credits)</p> <p>FREE ELECTIVES (9-10 credit hours)</p> <p>TOTAL CREDIT HOURS: 124 *A grade of C or better is required</p>

Experimental design, data collection, and interpretation techniques necessary for biochemistry and

molecular biology are covered by comprehensive courses in organic chemistry, microbiology, anatomy and physiology, biochemistry, immunology, cell biology and genetics (Figure 1). These courses are designed to provide the student with a thorough understanding of the cellular and molecular subject material.

To familiarize students with the more theoretical and mathematical aspects of chemical behavior and energetics, especially as they apply to biological and biochemical phenomena, in AY 2014, Wesley developed a one-semester physical chemistry course (Physical Chemistry for the Life Sciences). Topics covered include thermodynamics and thermochemistry, physical and chemical equilibrium processes (especially those involving biological or biochemical processes), chemical kinetics and transport processes, and basics of quantum theory applied to structures of molecules and spectroscopy. The laboratory experiments covered thermodynamics and thermochemistry of reactions, protein stability and ligand-protein binding, thermodynamics of hydrophobic interactions, ionic strength effects on equilibria, phase equilibria and chemical equilibria, chemical kinetics (including enzyme kinetics), and applications of UV/visible absorption and fluorescence spectroscopy (to study polyene structure, equilibria, and solution interactions).

In addition, the Analytical Chemistry laboratory was modified to include analyses which would be of more interest to all students in the area of biological chemistry, some of which make use of the DE-INBRE/DE-EPSCoR newly-purchased spectroscopic instrumentation from Shimadzu (an RF5301 fluorescence spectrometer and an AA7000 atomic absorption spectrometer). Examples of new experiments include the colorimetric determination of the phenolic content of foods, fluorescence determination of quinine in tonic water or dextromethorphan in cough medications, and the determination of metals in foods or personal care products by atomic absorption spectroscopy.

Students interested in pursuing careers that examine the relationships between humans and the environment, the diversity of life, ecological principles, evolutionary concepts in relation to physiological and biochemical adaptation and/or specialization, and population genetics or behavior, can select courses under the major-electives subheading (Figure 1) to serve their needs. Such well-chosen elective choices round out and add breadth and depth to future post-graduate studies.

Fueled by the increasing ubiquity and importance of Big Data in many industry segments, Wesley believes that it is increasingly important for undergraduate students in all disciplines to have some exposure to data science, statistical inference, and visualization methods and technologies. The mathematics and biological chemistry programs recognized this need to help their graduates stand out in highly competitive landscapes and incorporated statistical analysis, data science, and visualization courses into their curricula. In 2015, the Department of Mathematics began offering an interdisciplinary Informatics minor (15 credit hours in total and consisting of five courses), and the Department of Science modified its biological chemistry major program requirements (69-70 credit hours) to include two project based courses as major requirements: a GIS spatial analysis course and a longitudinal multilevel data analysis SAS course (Figure 1).

PROJECT-BASED EXPERIENTIAL LEARNING USING GIS AND SAS TECHNIQUES

The College offers a senior-level spatial analysis course that emphasizes the use of GIS to visualize geographically referenced data toward the solution of a wide variety of student-developed research projects. Originally designed as an exclusive vehicle for exploring the use of remote sensing data in the assessment and solution of large-scale environmental problems, the decision was made to broaden the appeal of the course to disciplines other than environmental science. By developing and integrating material addressing the emerging importance of Big Data and data science in the exploration and solution of a wide variety of disciplinary areas, this course provided students with an introduction not only to GIS technology and functionality, but also to the particular challenges presented by Big Data and approaches being developed in data science.

In Delaware, UD offers graduate courses on SAS programming in their master's level statistics program, but no Delaware institution of higher education is offering an undergraduate SAS course. Yet, at every one of the (recent) Careers in Mathematics Conferences sponsored by the Eastern Pennsylvania and Delaware Section of the Mathematics Association of America (MAA/EPaDel), the importance of data analysis utilizing SAS in government and corporate jobs has been emphasized. The Department of Mathematics' new interdisciplinary Informatics Minor

comprises a general statistics course, a research methods course, and three programming/data mining courses. A junior level SAS programming course was co-developed by the Department of Mathematics and Ms. Haiyan Weng, of Student Loans, LLC, Delaware.

The data-driven courses in the Informatics Minor include data manipulation utilizing SAS programming and geospatial analysis using the ArcGIS platform. The SAS programming and the ArcGIS site licenses were strictly obtained through external grant funding (from the DE-EPSCoR and the DE-INBRE grant programs).

Geographic Information Systems Course (3 credits):

An important aspect of this newly revised biological chemistry curriculum is the exploration of how the research paradigm has changed with the advent of Big Data. A traditional waterfall research project may have started with a well-formed thesis statement followed by data collection, analysis, and conclusions. Increasingly, though very large data sets often used in these research projects lend themselves to exploratory analysis that yields emergent pattern knowledge that can result in the formation of a thesis. While causation remains to be established, correlation of data through visualization can provide interesting research directions that might not have otherwise been discovered.

For the purposes of undergraduate projects and research, GIS technology and visualization provides students with rapid and more easily interpreted feedback of spatially distributed phenomena than might be achieved with more standard statistical analyses using SPSS (Statistical Package for the Social Sciences) or SAS. Interestingly, data science is developing new programming languages (e.g. R and Python) that enable users to rapidly construct and conduct complex statistical and mathematical analyses the results of which can be integrated into the GIS visualization. With version 10.3 ESRI has begun to release integrated packages of Python code that can be used to conduct complex analyses in the ArcMap environment. Packages of open source R code have been developed for spatial analysis and mapping applications (Brundson and Comber, 2015).

Large volumes of demographic data aggregated at various scales have been available for quite some time (United States Census Bureau, University of Wisconsin Population Health Institute). More recently, though, encouraged by President Obama's 2013 Executive Order and Open Data Policy we have witnessed a dramatic growth in online, machine-readable data availability in areas such as health, energy, education, public safety, finance, and global development (Park and VanRoekel, 2013). The availability of these data provides researchers, professionals, and students with an expanding ocean of data with which to conduct their work.

Musa et al. (2013) provide a brief and informative overview of the field of medical geography, which strives to correlate pathological and geographical factors. With the advent and evolution of statistical and computer sciences and the greater availability of medical and epidemiological data, it gave rise to medical GIS, which has integrated statistical and spatial analysis functions. GIS has been used descriptively in population and public health fields (Ogden et al., 2014; Moore et al., 2008; Ricketts, 2003) to examine healthcare service provision (Hawthorne and Kwan, 2012) and epidemiology (Clarke et al., 1996).

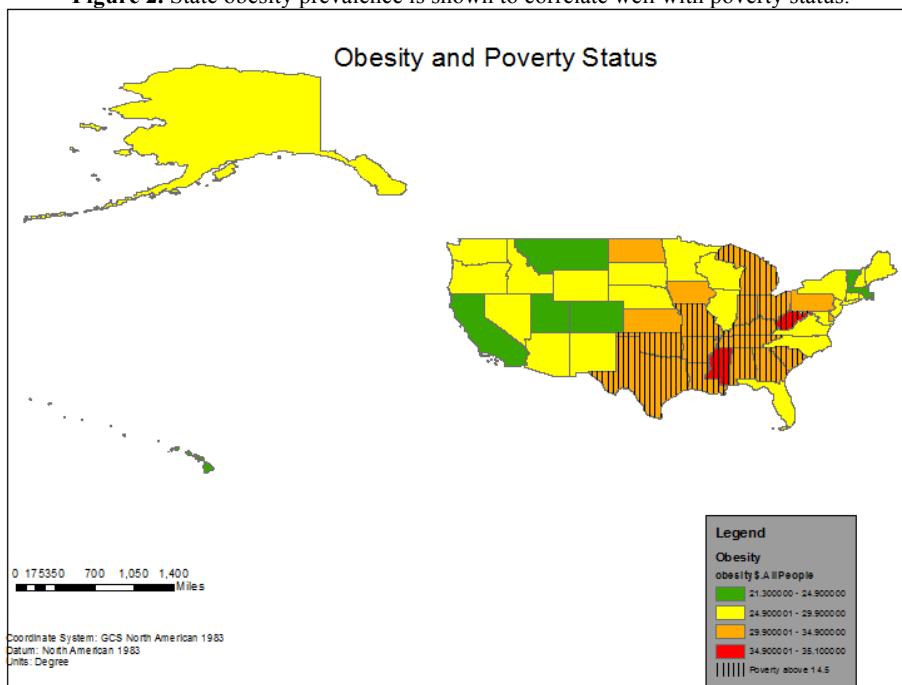
The American Obesity Case Study project outlined below was developed in a senior undergraduate-level spatial analysis class using ESRI ArcGIS Desktop and ArcMap. Most of the students had no prior experience with GIS software and came from a variety of disciplinary backgrounds, so the first six weeks of the semester was devoted to an instructor-facilitated, self-paced, lab-intensive introduction to GIS basics, ArcMap functionality, and a variety of online data sources. The students then completed exercises that demonstrated basic competencies in creating choropleth maps, importing and joining attribute data to geographic base maps, the basics of remote sensing technologies, data, and visualization. A mini-project capped off the first-third of the semester.

Students then formed small project groups according to their disciplinary interests and developed project plans around selected research questions. A series of interim deliverables due over the remainder of the semester (e.g., candidate data sources, literature review, and methodology overviews) was defined to insure that the groups would remain on-track. The final deliverable was assigned to be a presentation-grade poster to be presented to the class on the night of the final examination.

While small assignments were made in order to cover other aspects of GIS relevant to their various disciplines, the remainder of the semester was primarily devoted to the group research projects. The instructor was available in-class and online to provide support and guidance for the project teams.

In their American Obesity Case Study, three biological chemistry majors mined the 2013-2014 state-and-county-level demographic data from the United States Census Bureau and the 2013-2014 obesity ranking data from the Centers for Disease Control (state-level for the 50 states) and County Health Rankings and Roadmaps (county-level for the contiguous 48 American states). This was done to explore any correlations between obesity, social class and other socioeconomic characteristics of the people affected by the ongoing obesity epidemic. An example of the data visualizations produced during this phase of the project is shown in Figure 2.

Figure 2. State obesity prevalence is shown to correlate well with poverty status.



Results of a quick, state-level selection-by-attribute analysis are reported in Table 1 listing states with obese populations classified according to poverty level, health insurance status, income, educational level, and age. For the twenty states shown there was a demonstrable correlation between obesity and social class and/or socioeconomic status. Two states, West Virginia (WV) and Mississippi (MS), have obese populations in all of the five categorical classifications. The other 18 other states had average obesity prevalence values that correlated well with obesity in at least one of the five classifications. Out of these 18 states, 3 states, Arkansas (AR), Georgia (GA), and South Carolina (SC) have high poverty levels, low rates of health insurance coverage, low household income, low educational attainment and obesity for populations with a median age over 35 years-of-age. There are 7 states (Alabama, Indiana, Kentucky, Michigan, Missouri, Ohio, and Tennessee) with high poverty levels, low household income, low educational attainment, and populations whose age over 35 years old are overweight, but high rates of health insurance coverage. This indicates that there is may be a weak correlation between health insurance and obesity. For Delaware the data suggest that educational achievement and age may be prevailing risk factors in determining obesity.

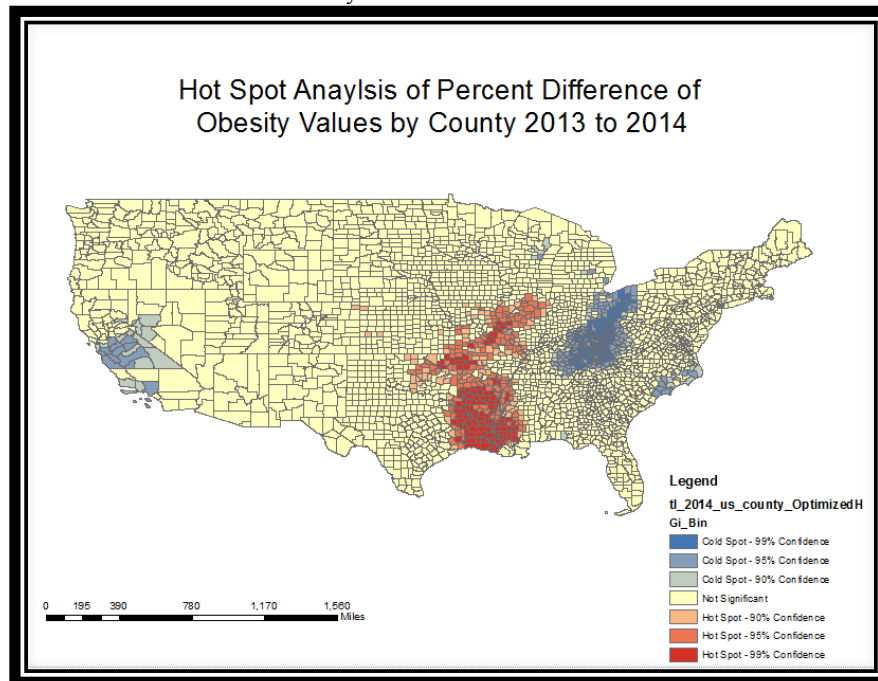
Table 1. Obesity (XX indicates correlation across all variables) correlates with social class and/or socioeconomic status in twenty states (2013 data).

US State		Obesity + Poverty	Obesity + No Health Insurance	Obesity + Income	Obesity + Educational Attainment	Obesity + Age
Alabama	AL	X		X	X	X
Arkansas	AR	X	X	X	X	X
Deleware	DE				X	X
Georgia	GA	X	X	X	X	X
Indiana	IN	X		X	X	X
Iowa	IA				X	X
Kansas	KS					X
Kentucky	KY	X		X	X	X
Louisiana	LA		X	X	X	X
Michigan	MI	X		X	X	X
Mississippi	MS	XX	XX	XX	XX	XX
Missouri	MO	X		X	X	X
North Dakota	ND					X
Ohio	OH	X		X	X	X
Oklahoma	OK	X	X		X	X
Pennsylvania	PA				X	X
South Carolina	SC	X	X	X	X	X
Tennessee	TN	X		X	X	X
Texas	TX	X	X		X	
West Virginia	WV	XX	XX	XX	XX	XX

In a second, more detailed, analysis, the students used county-level demographic and obesity data. The characteristics analyzed were; income, poverty status, health insurance coverage, education, and age. After joining these data by GEOID, a unique state/county code, they performed cluster analyses that provided them with statistically significant geographical concentrations of obesity (or lack thereof).

The initial analysis (Figure 3) showed statistically significant obesity hot spots (groupings of counties with similarly high incidences of obesity) in red and cold spots (groupings of counties with similarly low incidences of obesity) in blue across the contiguous 48 states.

Figure 3. The map highlights counties where obesity levels advanced (red) above the Center for Disease Control (CDC) national average (28.6%), and where obesity levels fell (blue) fell below that level. The yellow areas show no significant change in obesity levels from 2013 to 2014.



Census attribute data joined to the county basemap were then subjected to individual cluster analyses and visually compared to the obesity prevalence layer.

A cluster analysis of educational attainment (individuals over 25 years-of-age with a bachelors or higher degree) was performed. High educational attainment clusters are shown in black; low educational attainment clusters are shown in blue (Figure 4). Selecting counties by attribute, a separate map layer showing counties reporting obesity levels greater than the national average of 28.6% reported by the Centers for Disease Control (CDC) was superimposed. These counties are shown outlined in red. Visual inspection of this superposition indicates that the prevalence of obesity is generally inversely proportional to educational attainment.

In similar fashion analyses were conducted on the remaining census county-level attribute data and their relationships to the obesity prevalence data:

- Median age (Figure 5)
- Median household income (Figure 6)
- Households with income below the national poverty level (Figure 7)
- Individuals without health insurance (Figure 8)

Figure 4. A cluster analysis of educational attainment (individuals over 25 with a bachelor’s degree or higher) shows areas where there are high levels of educational attainment (black) and area with low educational attainment percentages (blue) compared to the national average of educational attainment.

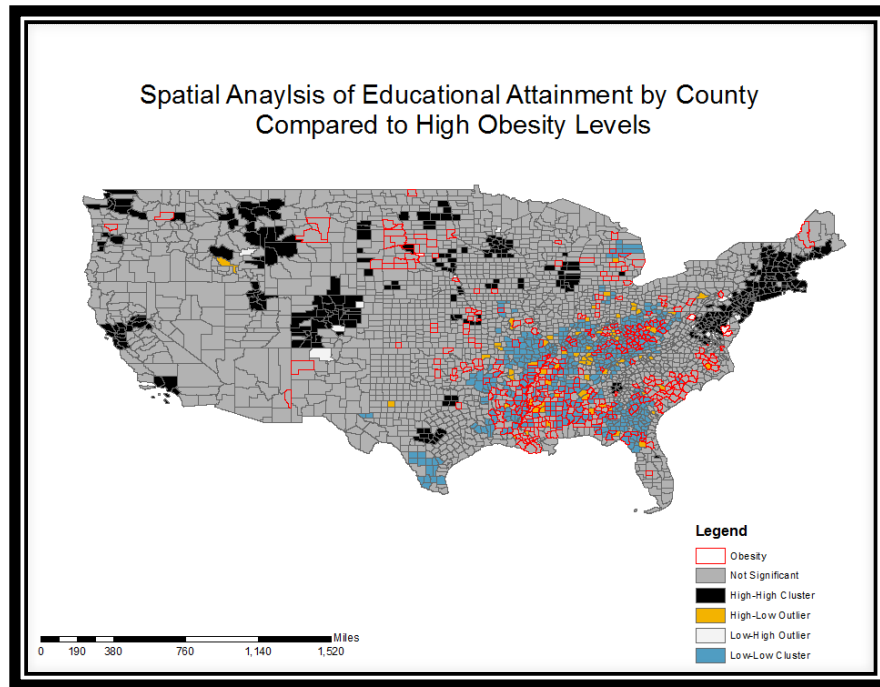


Figure 5. A cluster analysis of median age shows areas where the population is older (black) compared to younger populations (blue). According to the CDC people 35 years or older are most likely to be obese.

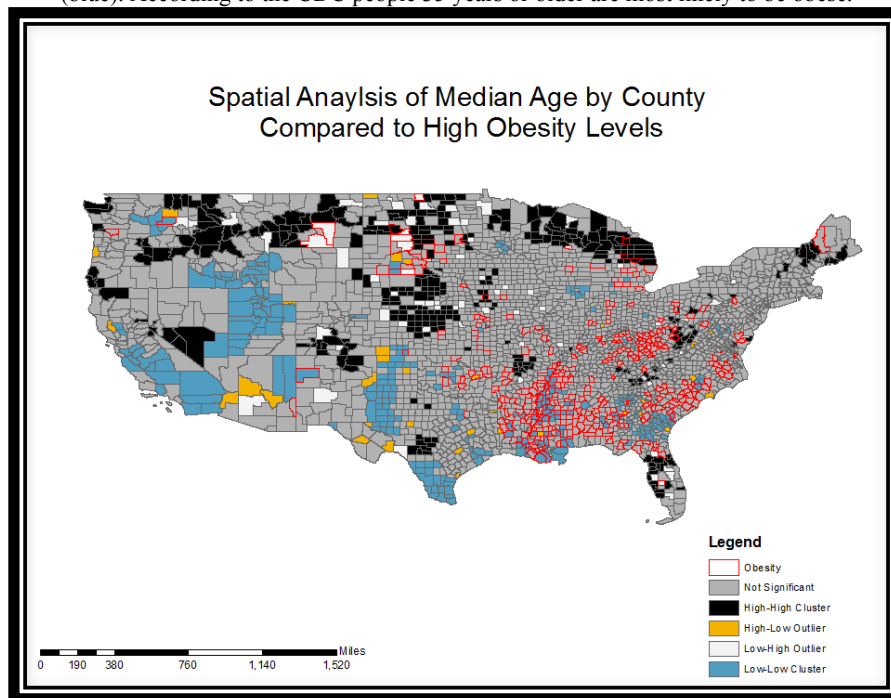


Figure 6. A cluster analysis of median income by household shows areas of high income (black) and areas with low income (blue). In the US, the median household income is \$50,000.

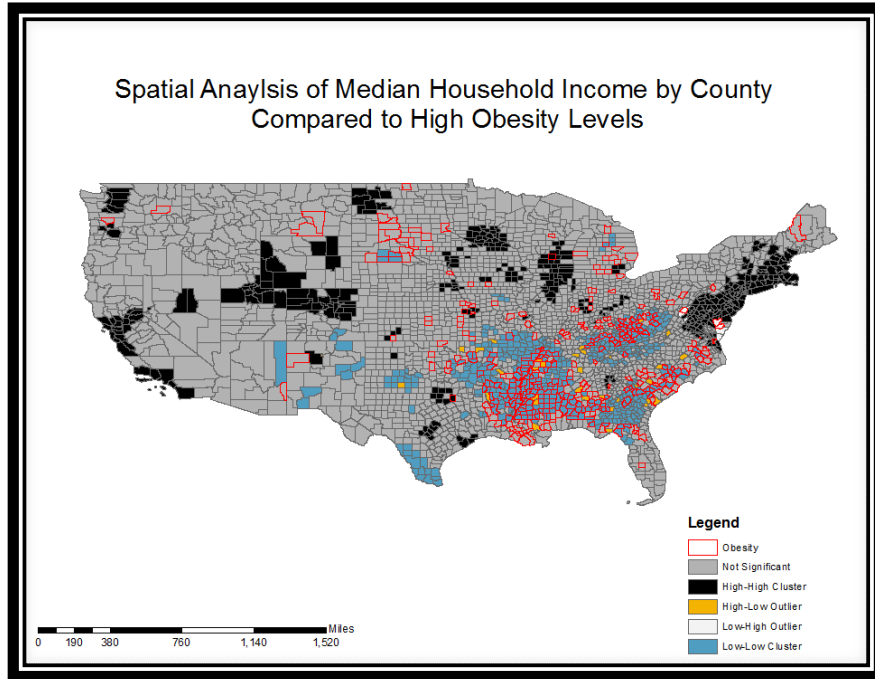


Figure 7. A cluster analysis shows areas where there is high poverty (black) and area with low poverty (blue). According to the CDC 14.5 % of the national population falls at or below the official poverty level.

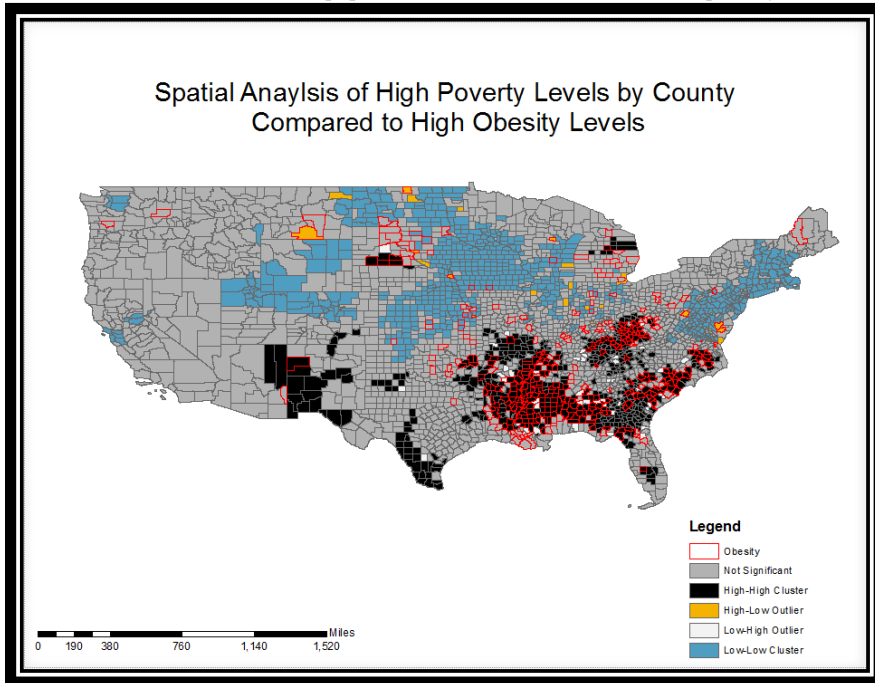
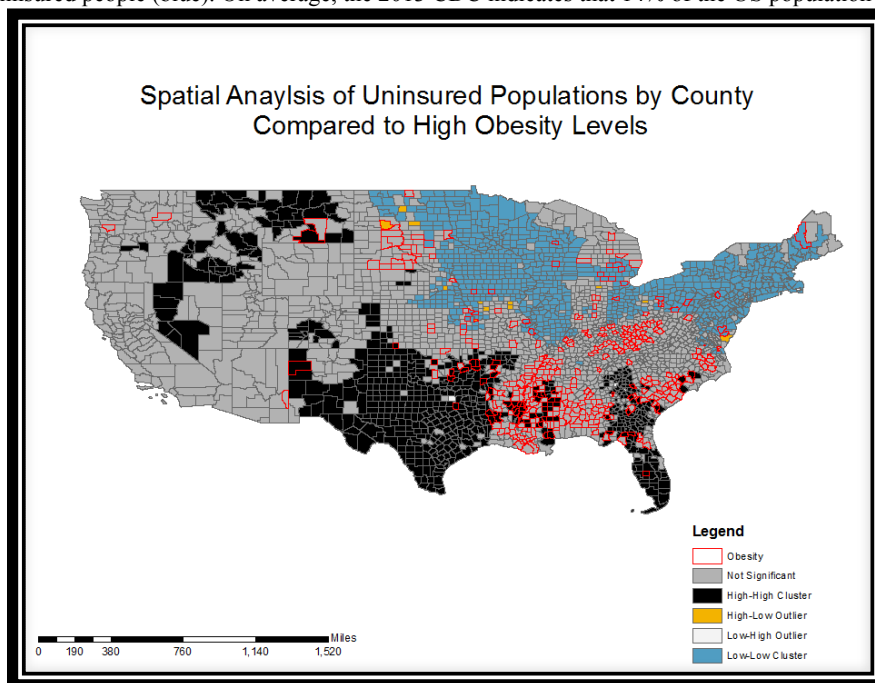


Figure 8. A cluster analysis shows areas where there are high levels of medically uninsured people (black) and areas with low levels of uninsured people (blue). On average, the 2013 CDC indicates that 14% of the US population is uninsured.



From these analyses, the students concluded that:

- Age has little relation to obesity prevalence even though obesity is higher among middle age adults, 40-59 years old (Figure 5).
- The prevalence of obesity is inversely proportional to median income (Figure 6).
- High poverty levels correlate directly to the prevalence of obesity (Figure 7).
- There is little to no correlation between obesity and the proportion of the population that is medically uninsured (Figure 8)

In retrospect, it is not necessary for the analyses using GIS analysis and visualization to be terribly complex for them to yield insightful results for inquisitive students. The benefits from introducing GIS to undergraduate students in an interdisciplinary context are manifold. The most obvious among these is the experience in working with GIS software, a skill that is increasingly attractive both in industry and academia. As GIS applications become more and more web-based, integrating them into coursework across the curriculum will become easier and more attractive.

As noted above, there is an increasing volume of georeferenced data available spanning a wide variety of disciplines. With rudimentary and easily acquired GIS skills an enormous amount of attribute data (e.g., demographic data) can be registered on underlying geographical maps. Students producing their first choropleth maps from these data are awestruck at the ease with which the maps are produced and manipulated and how compelling are the stories they can tell. Students can use the skills they develop in formal GIS courses to augment their coursework in other disciplines.

Working with GIS software also provides students an entry into the world of Big Data and data science. As much of the attribute data is not clean, students have to develop skills in examining (large) datasets for missing or extraneous data and using various tools (e.g., R and Python, in addition to Microsoft Excel) to scrub their data before it is usable by GIS applications. With the increasing integration of mathematical and statistical modeling functionality into GIS applications, more and more of this work can be done in a single application (using SAS).

SAS Programming Course (3 credits):

Rather than using SAS for new, authentic research, this course was designed for students to learn data manipulation and representation using data collected from prior undergraduate (experimental) research projects. This allows them to focus on the data manipulation aspect using SAS as part of the research process. The main goal of the course is to influence the utilization of SAS programming as one tool available for advanced research analytics. The course also attempts to teach clear and rational competencies common to the corporate environment – viz. reading and managing data, analyzing and evaluating data, and presenting data to evaluate options; form accurate conclusions; and to make final decisions. While many corporations use Microsoft (MS) PowerPoint in conjunction with MS-Excel for that last step, this course uses the SAS output delivery system (ODS) for final reporting.

SAS is software written for users for reading, transforming, and organizing data while maintaining the flexibility of writing their own customized code. Users can quickly create a few lines of code that read data from a file, manage variables and observations, and create a table for reporting. For the instructor, this creates an advantage when teaching students without much formal training in computer programming.

This SAS programming course was developed in modules that are based on the basic analytic work-flow. Modules are repeated with more complex content throughout the semester. First, students learn how to get data into SAS. Initially this is limited to standard text files, such as comma-separated values. Over the course of the semester, this module is reiterated by introducing more advanced importing of delimited files or MS-Excel files, and finally, extracting data from a SQL database. In a second module, students learn how to explore, modify, manage, and analyze their data. Among other topics, this involves validating data, filtering variables or observations, learning how to read data efficiently by use of options, and using the ODS for simple PDF (Portable Document Format) output. Emphasis is placed on testing code on a subset of observations which is especially useful when working with large data sets. The third module introduces custom tables and reports. Students also learn how to create custom formats to present data in particular ways. For all modules, students are taught how basic SAS macro programming can help make code more readable and maintainable.

As part of the course requirements students complete two projects that revisit data already analyzed in other Wesley College research projects. The first of these projects uses the 2014 County Health Rankings Trend Data to analyze obesity rates over time in all three Delaware counties. This is then compared to the State-wide data and the total U.S population. This data was (already) part of the GIS project mentioned above, so students were aware of the expectations prior to delving into this project. The focus in the SAS course was on efficient data management and presentation. In order to create the graphs of obesity rates (Figures 9 and 10), two biological chemistry seniors read the raw data set and filtered it by using an adult obesity variable within the county geographic area observations. They calculated the percent of obese adults and created a time series using ODS. Special emphasis was put on labeling and formatting the graph appropriately using options and custom formats.

Figure 9. The 2004-2010 obesity trends observed for the three Delaware counties are shown. The Delaware data were obtained from the County Health Rankings & Roadmaps 2014 website.

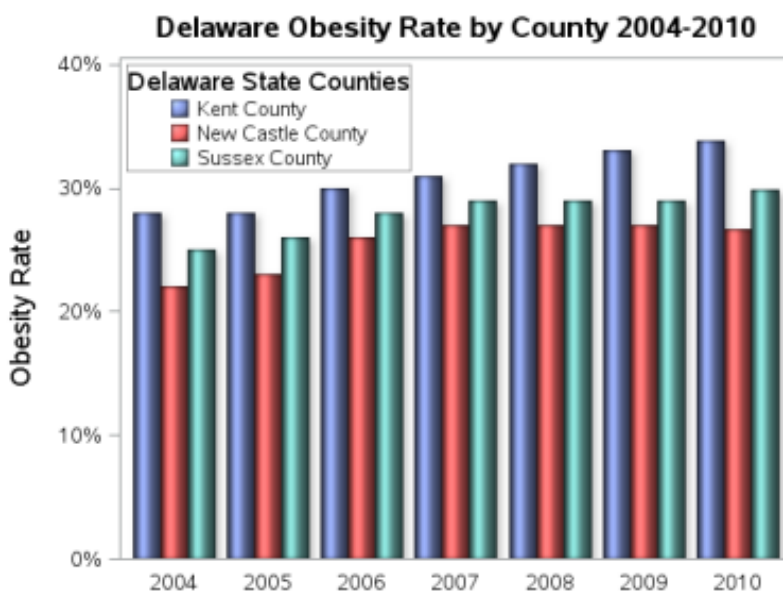
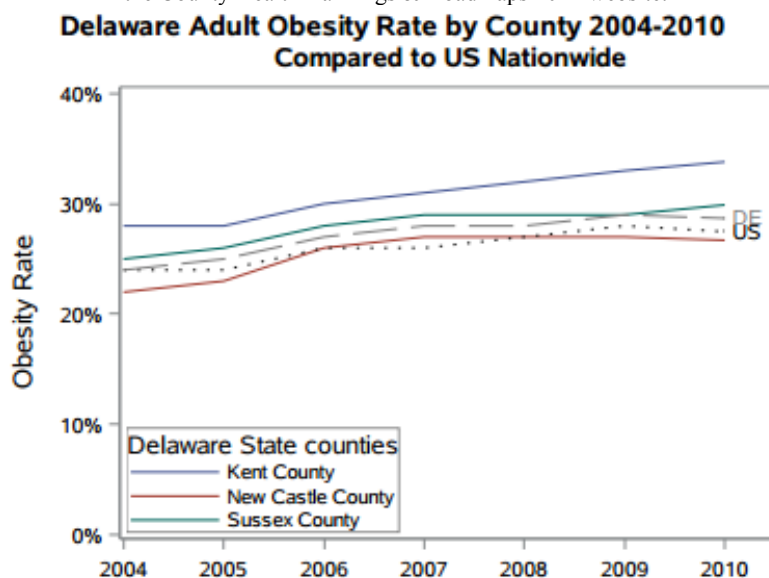


Figure 10. The Delaware 2004-2010 obesity trends are compared to national trends. All of the obesity data were obtained from the County Health Rankings & Roadmaps 2014 website.



The 2004-2010 Delaware trends shown in Figures 9 and 10 are consistent with previously published studies on the lifestyle choices and underlying risk factors for many chronic conditions which contribute to poor health, reduced quality of life, and higher health care costs for the two southern Delaware counties, Kent and Sussex counties (D’Souza and Walls et al., 2015; Gupta, 2014; Chang et al., 2014; Xu et al., 2013).

ACKNOWLEDGEMENTS

The authors acknowledge support from an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences (NIGMS) at the National Institutes of Health (NIH grant no. P20GM103446, Delaware INBRE program); a National Science Foundation (NSF) Experimental Program to Stimulate Competitive Research

grant IIA-1301765 (Delaware EPSCoR program); an NSF American Recovery and Reinvestment Act (ARRA) grant (NSF ARI R2 0960503); an NSF S-STEM grant (NSF DUE 1355554 - the Wesley College Cannon Scholar program); and the State of Delaware. The DE-INBRE and DE-EPSCoR grants were obtained through the leadership of the University of Delaware and the authors sincerely appreciate their efforts. Dr. Frank Fiedler would like to thank Ms. Haiyan Weng, College Ave Student Loans, LLC, for using her extensive experience in modeling, underwriting strategy, and risk management with SAS to collaborate and create a SAS course that is suitable for undergraduates.

AUTHOR INFORMATION

Malcolm J. D'Souza is Professor of Chemistry at Wesley College, in Dover, Delaware. He is the Principal Investigator on the Wesley College DE-INBRE, DE-EPSCoR, NSF ARRA, and Cannon Scholar (NSF S-STEM) programs. In 2009, to mark the occasion of its 50th anniversary, the College of Liberal Arts and Sciences, Northern Illinois University (NIU), recognized the accomplishments of Dr. D'Souza as one of its 50 most distinguished alumni of the college. In 2012, he was awarded the American Chemical Society's (ACS) E. Emmett Reid Award, which recognizes high-quality teaching in chemistry at small colleges in the ACS Mid-Atlantic region. In addition to his research in physical organic chemistry, he also has projects, presentations, and publications in the areas of chemometrics, STEM education, weight-loss studies, cheminformatics and in the design of commercial databases that assist in the development of new pharmaceutical and agricultural products. Dr. D'Souza also serves as the Associate Dean of Interdisciplinary/Collaborative Sponsored Research. E-Mail: malcolm.dsouza@wesley.edu (Corresponding author).

Richard J. Kashmar is Professor of Chemistry and Physics at Wesley College. A graduate of Carnegie Mellon University, he received a Ph.D. in Physical Chemistry from the University of Pittsburgh. Dr. Kashmar's teaching responsibilities include the Chemistry I and II courses, as well as Physical Chemistry for Life Sciences and Analytical Chemistry. He has done research work involving the statistical comparison of student performance in the traditional freshman-sophomore chemistry sequence vs. the nontraditional sequence used at Wesley. He has presented his results at an American Chemical Society (ACS) National Meeting in 2009, at the Middle Atlantic Regional ACS Meeting in 2011, and at the 22nd Biennial Conference on Chemical Education in 2012 (where he also organized and moderated a symposium on nontraditional chemistry sequences). E-Mail: richard.kashmar@wesley.edu.

Kent Hurst serves as Director of Environmental Science Graduate Program at Wesley College. He is a graduate of the University of Texas Arlington with a doctorate in urban planning and public policy. He also holds advanced degrees in physics and business. His dissertation research investigated the role of local and regional planning in the mitigation of climate change. His current interests include the political economy of climate change and its effects on mitigation and adaptation efforts, especially in coastal areas. E-Mail: kent.hurst@wesley.edu.

Frank Fiedler is Associate Professor of Mathematics at Wesley College. He is a graduate of Dresden University of Technology with minors in Mechanical Engineering, Electrical Engineering, and Physics. Dr. Fiedler received his Ph.D. in Mathematics from the University of Delaware. His research interests lie mainly in the mathematics of digital communication. E-Mail: frank.fiedler@wesley.edu.

Catherine E. Gross is a senior in the Biological Chemistry program at Wesley College. In AY 2014-2015, Catherine received a Cannon Scholar award (through an NSF DUE S-STEM program). She has presented her chemistry undergraduate research work at several national and regional scientific conferences and in 2013; one her posters was competitively selected to be showcased at Congress during the Council of Undergraduate Research (CUR) 2013 *Posters on the Hill* event. On graduation, Catherine will enroll in the Wesley College Master of Arts in Teaching (MAT) program to become a prospective teacher who will serve in schools with diverse and/or high-needs populations in STEM majors. E-Mail: catherine.gross@email.wesley.edu.

Jasbir K. Deol is a senior in the Biological Chemistry program at Wesley College. In AY 2014-2015, Jasbir received a Cannon Scholar award (through an NSF DUE S-STEM program). She has presented her chemistry undergraduate research work at several national and regional scientific conferences. In her junior year, she earned

poster presentation awards from both the National and the Delaware American Chemical Society (ACS) sections. In 2015, Jasbir received the Delaware American Chemical Society (DE-ACS) undergraduate award and the National ACS Undergraduate Environmental Chemistry Award for outstanding achievements in chemistry. On graduation, Jasbir will enter the Temple University School of Pharmacy program. E-Mail: jabir.deol@email.wesley.edu.

Alora Wilson is a sophomore in the Biological Chemistry program at Wesley College. As a freshman, Alora was a recipient of the CRC Press *Freshman Chemistry Award*. She has presented a poster of her chemistry undergraduate research efforts at the 2015 National American Chemical Society Meeting, Denver, CO. In Fall 2015, Alora will transfer to the University of Maryland Eastern Shore School of Pharmacy program E-Mail: alora.wilson@email.wesley.edu.

REFERENCES

- Ailamaki, A., Kantere, V., and Dash, D. (2010). Managing Scientific Data. *Communications of the ACM*, 53(6), 68-78.
- Ainsworth, S.J. (2012). Opportunities Along the Atlantic. *Chemical & Engineering News*, 90 (34), 61-63.
- Barh, D., Zambare, V., and Azevedo, V. (Eds.). (2013). *Omics: Applications in Biomedical, Agricultural, and Environmental sciences*. CRC Press.
- Brown, D.T., Ratledge, E.C. (2012). Energy, the Environment, and Delaware Jobs: An Analysis of Delaware's Green Educational Pipeline. Available at: http://128.175.63.72/projects/DOCUMENTS/Educ_pipeline.pdf.
- Brownell, S.E., Kloser, M.J., Fukami, T., and Shavelson, R. (2012). Undergraduate Biology Lab Courses: Comparing the Impact of Traditionally Based "Cookbook" and Authentic Research-Based Courses on Student Lab Experiences. *Journal of College Science Teaching*, 41, 18-27.
- Brunsdon, C. and Comber, L. (2015). *An Introduction to R for Spatial Analysis and Mapping*. London: SAGE Publications Ltd.
- Canaria, J.A., Schoffstall, A.M., Weiss, D.J., Henry, R.M., and Braun-Sand, S.B. (2012). A Model for an Introductory Undergraduate Research Experience. *Journal of Chemical Education*, 89(11), 1371-1377.
- Chang, D.I., Gertel-Rosenberg, A., and Snyder, K. (2014). Accelerating Efforts To Prevent Childhood Obesity: Spreading, Scaling, And Sustaining Healthy Eating And Physical Activity. *Health Affairs*, 33(12), 2207-2213.
- Chang, M.J., Sharkness, J., Hurtado, S., and Newman, C.B. (2014). What Matters in College for Retaining Aspiring Scientists and Engineers from Underrepresented Racial Groups. *Journal of Research in Science Teaching*, 51(5), 555-580.
- Clarke K., S. McLafferty, and B. Tempalski. (1996). On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases*, 2(2), 85-92.
- Clift, K., Scott, L., Johnson, M., and Gonzalez, C. (2014). Leveraging Geographic Information Systems in an Integrated Health Care Delivery Organization. *The Permanente Journal*, 18(2), 71-75.
- Cravatt, B.F., and Gottesfeld, J.M. (2010). Chemical Biology Meets Biological Chemistry Minireview Series. *Journal Biological Chemistry*, 285(15), 11031-11032.
- Duncan, D.T., Regan, S.D., Shelley, D., Day, K., Ruff, R.R., Al-Bayan, M., and Elbel, B. (2014). Application of Global Positioning System Methods for the Study of Obesity and Hypertension Risk Among Low-Income Housing Residents in New York City: a Spatial Feasibility Study. *Geospatial Health*, 9(1), 57-70.
- Dowdy, S., Wearden, S., and Chilko, D. (2011). *Statistics For Research (Vol. 512)*. John Wiley & Sons.
- D'Souza, M.J. (2005). KnowItAll[®] - Software Reviews, *Chemistry World*, 2(9), 70-71.
- D'Souza, M.J. (2007). KnowItAll[®] U System - Software Reviews, *Chemistry World*, 4 (11), 70-72
- D'Souza, M.J. (2008). FDA Consumer Drug database – 2007. HaveItAll - ADME/Tox Experimental Databases Datasheet, Bio-Rad Laboratories, Bulletin # INF-96199, 2008.
- D'Souza, M.J., AlAbed, G.J. (2010). Deficiencies in the Reporting of V_D and $t_{1/2}$ in the FDA Approved Chemotherapy Drug Inserts. *Pharmaceutical Reviews*, 8(1), 1-9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21643531>.
- D'Souza, M.J., AlAbed, G.J., Earley, M., Roberts, N., Gerges, F.J. (2013). Manipulating In-House Designed Drug Databases for the Prediction of pH-Dependent Aqueous Drug Solubility. *American Journal of Health Sciences*, 4(3), 137-150.
- D'Souza, M.J., AlAbed, G.J., Wheatley, J.M., Roberts, N., Veturi, Y., Bi, X., Continisio, C.H. (2011). A Database Developed with Information Extracted from Chemotherapy Drug Package Inserts to Enhance Future Prescriptions. *Conference Computer Vision Pattern Recognition Workshops, IEEE Conference Record #17768*, 2011, 219-226. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25302340>.
- D'Souza, M.J., Barile, B., and Givens, A.F. (2015). Evolution of a Structure-Searchable Database into a Prototype for a High-Fidelity SmartPhone App for 62 Common Pesticides Used in Delaware. *International Conference on*

- Industrial Instrumentation and Control (2015 ICIC)*, IEEE Xplore, Under Review.
- D'Souza, M.J., Dwyer, P., Allison, B.E., Miller J.M., and Drohan, J. (2011). Wesley College Ignites Potential with Undergraduate Student Research Program. *Council of Undergraduate Research Quarterly*, 32(2), 41-45.
- D'Souza, M.J., Gerges, F.J. (2010). Raising a Red Flag: Deficiencies Found in the Reporting of Important Parameters in FDA Approved Drug Profiles. *Pharma IQ*, 2850. Available at: <http://www.pharma-iq.com/article.cfm?externalID=2850>.
- D'Souza, M.J., and Kevill, D. N. (2014). Influence of Sulfur for Oxygen Substitution in the Solvolytic Reactions of Chloroformate Esters and Related Compounds. Review Chapter. *International Journal of Molecular Sciences*, 15, 18310-18332.
- D'Souza, M.J., and Kevill, D.N. (2013). Application of the Grunwald-Winstein Equations to Studies of Solvolytic Reactions of Chloroformate and Fluoroformate Esters. Review Chapter in *Research Recent Developments in Organic Chemistry*, Editor, S. G. Pandalai. Publisher, Transworld Research Network, Trivandrum-695 023 Kerala, India, 13, 1-38. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4215406/>.
- D'Souza, M.J., Kroen, W.K., Stephens, C.B., Kashmar, R.J. (2015). Strategies and Initiatives that Revitalize Wesley College STEM Programs. *Journal of College Teaching and Learning*, 12(3), 195-208.
- D'Souza, M.J., Koyoshi, F., Everett, L.M. (2009). Structure Activity Relationships (SARs) Using a Structurally Diverse Drug Database: Validating Success of Predictor Tools. *Pharmaceutical Reviews*, 2009. Available at: <http://www.pharmainfo.net/reviews/structure-activity-relationships-sars-using-structurally-diverse-drug-database-validating-su>.
- D'Souza, M.J., Koyoshi, F., Everett, L.M. (2009). Structure Activity Relationship (SAR) Patterns Observed Within a Series of Unrelated Common Consumer Drugs. *Bioinformatics, Computational Biology, Genomics, and Cheminformatics*, 2009, 1-6.
- D'Souza, M.J., Kyoshi, F. (2009). Extracting Relevant Information from FDA Drug Files to Create a Structurally Diverse Drug Database Using KnowItAll[®]. *Pharmaceutical Reviews*, 7(3). Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25356090>.
- D'Souza, M.J., Walls, K.-J. E., Rojas, C., Everett, L.M., Wentzien, D.E. (2015). Effect of Gender and Lifestyle Behaviors on BMI Trends in a Sample of the First State's Undergraduate Population. *American Journal of Health Sciences*, 6(1), 59-74.
- D'Souza, M.J., and Wang, Q. (2012). Inter-Institutional Partnerships Propel a Successful Collaborative Undergraduate Degree Program in Chemistry. *Journal of College Teaching and Learning*, 9 (4), 245-252.
- Felder, R. M., and Brent, R. (2005). Understanding Student Differences. *Journal of Engineering Education*, 94(1), 57-72.
- Fiegerman, S. (2011). The Cities with the Jobs of the Future. *MainST*. Available: <http://www.mainstreet.com/slideshow/cities-jobs-future>.
- Finley, A., and McNair, T. (2013). Assessing Underserved Students' Engagement in High-Impact Practices. 2013. Association of American Colleges and Universities. LEAP. Available at: <http://www.aacu.org/assessinghips>
- Giang, T., Karpyn, A., Laurison, H.B., Hillier, A., and Perry, R.D. (2008). Closing the Grocery Gap in Underserved Communities: The Creation of the Pennsylvania Fresh Food Financing Initiative. *Journal of Public Health Management and Practice*, 14(3), 272-279.
- Hatcher, L., and O'Rourke, N. (2014). A Step-By-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling. *SAS Institute*.
- Hawthorne, T. and M. Kwan. (2012). "Using GIS and Perceived Distance to Understand the Unequal Geographies of Healthcare in Lower-Income Urban Neighbourhoods." *Geographical Journal*, 178(1): 18-30.
- Hersh, W.R., Gorman, P.N., Biagioli, F.E., Mohan, V., Gold, J.A., and Mejicano, G.C. (2014). Beyond Information Retrieval and Electronic Health Record Use: Competencies in Clinical Informatics for Medical Education. *Advances in Medical Education and Practice*, 5, 205-212.
- Hwang, D., Curl, S. (2014). The Market for Career Tracks in Undergraduate IS Curricula in the U.S. *Information Systems Education Journal*, 12(3), 4-17.
- Imes, C.C., and Burke, L.E. (2014). The Obesity Epidemic: The USA as a Cautionary Tale for the Rest of the World. *Current Epidemiology Reports*, 1(2), 82-88.
- Jungck, J.R., Gaff, H.D., Fagen, A.P., and Labov, J.B. (2010). "Beyond BIO2010: Celebration and Opportunities" at the Intersection of Mathematics and Biology. *CBE-Life Sciences Education*, 9(3), 143-147.
- Kesh, S., and Ramanujan, S. (2010). Healthcare Information Systems: A Pedagogical Analysis of Current Degree Programs. *Issues in Information Systems*, 11(2), 178-183.
- Kevill, D.N., D'Souza, M.J. (2008). Sixty Years of the Grunwald-Winstein Equation: Development and Recent Applications. *Journal of Chemical Research*, 61-66.
- Koch, T. (2004). The map as Intent: Variations on the Theme of John Snow. *Cartographica: IEEE Xplore - International Conference on Industrial Instrumentation and Control (2015 ICIC)*, 2015, 71-76.

- Kolvoord, R., Charles, M., and Purcell, S. (2014). What Happens After the Professional Development: Case Studies on Implementing GIS in the Classroom. In *Teaching Science and Investigating Environmental Issues with Geospatial Technology* (pp. 303-320). Springer Netherlands.
- Kuh, G.D. (2008). *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*. Washington DC: Association of American Colleges and Universities.
- Lee, J., and Wong, D.W. (2001). *Statistical analysis with ArcView GIS*. John Wiley & Sons.
- Longenecker, B., Campbell, M., Landry, J., Pardue, J.H., and Daigle, R. (2012). A Health Informatics Curriculum Congruent with IS 2010 and IMIA Recommendations for an Undergraduate Degree. *Information Systems Education Journal*, 10(2), 15-32.
- Lytle, L.A. (2009). Examining the Etiology of Childhood Obesity: The IDEA study. *American Journal of Community Psychology*, 44(3-4), 338-349.
- Malczewski, J. (1999). *GIS and Multicriteria Decision Analysis*. John Wiley & Sons.
- Mao, L. (2014). Ranking Academic Impact of GIS Research Organizations in the United States: A Bibliographic Network Analysis Over 20 years. *GIScience & Remote Sensing*, 51(1), 51-62.
- Matthews, K.E., Adams, P., and Goos, M. (2010). Using the Principles of BIO2010 to Develop an Introductory, Interdisciplinary Course for Biology Students. *CBE-Life Sciences Education*, 9(3), 290-297.
- Matthews-Juarez, P. (2013). Developing a Cadre of Transdisciplinary Health Disparities Researchers for the 21st Century. *Journal of Health Care for the Poor and Underserved*, 24(1), 121-128.
- Moore, L.V., Roux, A.V D., and Brines, S. (2008). Comparing Perception-Based and Geographic Information System (GIS)-Based Characterizations of the Local Food Environment. *Journal of Urban Health*, 85(2), 206-216.
- Musa, G.J., Chiang, P.H., Sylk, T., Bavley, R., Keating, W., Lakew, B., Tsou, H.-C., and Hoven, C.W. (2013). Use of GIS Mapping as a Public Health Tool—From Cholera to Cancer. *Health Services Insights*, 6, 111-116.
- National Research Council (US) Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century. *Bio2010: Transforming Undergraduate Education for Future Research Biologists*. Washington (DC): National Academies Press (US); 2003. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK43511/>.
- Ogden, C.L., Carroll, M.D., Kit, B.K., and Flegal, K.M. (2014). Prevalence of Childhood and Adult Obesity in the United States, 2011-2012. *The Journal of the American Medical Association*, 311(8), 806-814.
- Owen, R.L., and Breyer, E.D. (2005). A Participant-Oriented, Research-Based Approach for Design of a Biochemistry Workshop for Faculty at Undergraduate Institutions. *Biochemistry and Molecular Biology Education*, 33(4), 269-273.
- Park, T. and VanRoekel, S. (2013). Landmark Steps to Liberate Open Data. Accessed February 25, 2015 from <http://www.whitehouse.gov/blog/2013/05/09/landmark-steps-liberate-open-data>.
- Pursell, D.P. (2009). Enhancing Interdisciplinary, Mathematics, and Physical Science in an Undergraduate Life Science Program Through Physical Chemistry. *CBE-Life Sciences Education*, 8(1), 15-28.
- Ricketts, T. (2003). Geographic Information Systems and Public Health. *Annual Review of Public Health*, 24: 1-6.
- Rylands, L., Simbag, V., Matthews, K.E., Coady, C., and Belward, S. (2013). Scientists and Mathematicians Collaborating to Build Quantitative Skills in Undergraduate Science. *International Journal of Mathematical Education in Science and Technology*, 44(6), 834-845.
- Sall, J., Lehman, A., Stephens, M.L., and Creighton, L. (2012). *JMP Start Statistics: A Guide to Statistics and Data Analysis Using JMP*. SAS Institute.
- Slack, T., Myers, C.A., Martin, C.K., and Heymsfield, S.B. (2014). The Geographic Concentration of US Adult Obesity Prevalence and Associated Social, Economic, and Environmental Factors. *Obesity*, 22(3), 868-874.
- Slonczewski, J.L., and Marusak, R. (2004). A Response to BIO 2010: Transforming Undergraduate Education for Future Research Biologists, from the Perspective of the Biochemistry and Molecular Biology Major Program at Kenyon College. *Biochemistry and Molecular Biology Education*, 32(3), 151-155.
- Smith, O.L., and Dragojlovic, V. (2013). A Design of an Undergraduate Research Program in Chemistry. *Journal of Laboratory Chemical Education*, 1(2), 25-33.
- Strategic Pharmaceutical Advisors (SRxA's) *Word on Health*. (2010). Available at: <http://srxawordonhealth.com/2010/07/26/fda-criticized-by-the-government-accountability-office/>.
- Steen, L.A. (2005). *Math & Bio 2010. Linking Undergraduate Disciplines*, Washington, DC: The Mathematical Association of America.
- The American Chemical Society (ACS) website: <http://www.acs.org/content/acs/en.html>.
- The ACS Biochemistry website: <http://www.acs.org/content/acs/en/careers/whatchemistsdo/careers/biochemistry.html>.
- The American Society for Biochemistry and Molecular Biology (ASBMB) website: <http://www.asbmb.org/>.
- The ASBMB Career Brochure. Available at: http://www.asbmb.org/uploadedFiles/ProfessionalDevelopment/Professional_Development/ASBMB_2011Caree

- rText_Web_Final.pdf.
- The Beyond Borders Biotechnology Industry 2013 report. Available at:
[http://www.ey.com/Publication/vwLUAssets/Beyond_borders/\\$FILE/Beyond_borders.pdf](http://www.ey.com/Publication/vwLUAssets/Beyond_borders/$FILE/Beyond_borders.pdf).
- The Delaware Department of Labor website: <https://www.delawareworks.com/chooser.php>.
- The Delaware Economic Development Office 2012 Data Book. Available at:
http://dedo.delaware.gov/dedo_pdf/NewsEvents_pdf/publications/DelawareDatabook.pdf
- The Delaware-EPSCoR website: www.epscor.udel.edu.
- The Delaware-INBRE website: www.inbre.udel.edu.
- The Delaware News Journal Article (June 4, 2013). Available at
<http://www.delmarvanow.com/article/20130606/CB01/306060006/STEM-academy-readies-students-high-tech-jobs>.
- The Delaware STEM Council Report, 2012-2013. Available at:
http://delawarestem.org/sites/default/files/Stem_Annual_report_lo1%202.pdf.
- The Environmental Systems Research Institutes (ESRIs): <http://www.esri.com/>.
- The EPSCoR/IDeA website: <http://www.epscorideafoundation.org/>.
- The *International Innovations* October 2012 article. Available at:
http://www.wesley.edu/fileadmin/editors_images/Academics/Biology/p86-88_Malcom_DSouza.pdf.
- The *International Innovations* May 2014 article. Available at:
http://www.wesley.edu/fileadmin/editors_images/Academics/Grants/INBRE/Malcolm_D_Souza_Intl_Innovation_135_Research_Media.pdf.
- The University of Wisconsin Population Health Institute. County Health Rankings & Roadmaps 2014.
www.countyhealthrankings.org.
- The United States Census Bureau website: <http://www.census.gov/>.
- The United States Department of Labor, Bureau of Labor Statistics Occupational Outlook Handbook, 2014-2015 Edition.
Available at: <http://www.bls.gov/oes/current/oes191021.htm>. The Wesley College Cannon Scholar program website: <http://www.wesley.edu/academics/programs/undergraduate-programs/cannon-scholars.html>.
- The Wesley College Core-Curriculum website: <http://www.wesley.edu/academics/programs/core-curriculum.html>.
- The Wesley College EPSCoR website: <http://www.wesley.edu/academics/grants/delaware-epscor.html>.
- The Wesley College INBRE website: <http://www.wesley.edu/academics/grants/inbre.html>.
- The Wesley College Scholars Day website: <http://www.wesley.edu/academics/scholars-day.html>.
- The Wesley College website: www.wesley.edu.
- Thiry, H., Laursen, S.L., and Hunter, A.B. (2011). What Experiences Help Students Become Scientists?: A Comparative Study of Research and Other Sources of Personal and Professional Gains for STEM Undergraduates. *The Journal of Higher Education*, 82(4), 357-388.
- Tinto, V. (2012). *Completing College: Rethinking Institutional Action*. University of Chicago Press.
- Wold, S. (1995). Chemometrics; What Do We Mean With It, And What Do We Want From It? *Chemometrics and Intelligent Laboratory Systems*, 30, 109-115.
- Wridt, P. (2010). A Qualitative GIS Approach to Mapping Urban Neighborhoods with Children to Promote Physical Activity and Child-Friendly Community Planning. *Environment and Planning. B, Planning & Design*, 37(1), 129-147.
- Xu, F., Town, M., Balluz, L.S., Bartoli, W.P., Murphy, W., Chowdhury, P.P., Garvin, W.S., Pierannunzi, C., Zhong, Y., Salandy, S.W., Jones, C.K., and Crawford, C.A. (2013). Surveillance for Certain Health Behaviors among States and Selected Local Areas—United States, 2010. *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, 62(1), 1-247.

NOTES