ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity

MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #36459

# Prepare Data Science Program Student Outcomes and Curricula for ABET Accreditation

**David Liu**

# Prepare Data Science Program Student Outcomes and Curricula for ABET Accreditation

David Q. Liu
Department of Computer Science
Purdue University
Fort Wayne, Indiana, USA
liud@pfw.edu

## ABSTRACT

ABET CAC (Computing Accreditation Commission) is in the semi-final stage of approving program criteria for Data Science within Computing programs for the first-time. Pilot CAC Data Science accreditation is being planned for the upcoming 2021-2022 accreditation cycle. In the meantime, ABET ANSAC (Applied and Natural Science Accreditation Commission) is also working with American Statistical Association in initiating the Data Science accreditation within Applied and Natural Science programs. This paper describes the ABET General and proposed or potential Data Science Specific Student Outcomes and Curriculum criteria within both Computing Programs and Applied and Natural Science Programs. Based on these criteria, we reviewed our university system's two existing BS Data Science programs: one developed jointly by Computer Science and Statistics programs; another one mainly from a Statistics program and determined whether they satisfy the ABET accreditation requirements for Data Science programs. We investigate how these two programs can be improved to satisfy the ABET Data Science curriculum requirements either as a Computing program or as an Applied and Natural Science program. Recommendations of curriculum changes are made to improve these programs for accreditation. We further describe specially how Data Science programs can work with other programs to incorporate application domains as required by ABET. We expect that this study will serve as a guideline for Data Science programs to develop curricula and seek accreditation.

## KEYWORDS

Data Science, ABET Accreditation, Student Outcomes, Curriculum

## 1 Introduction

Due to an increasing demand for data science related jobs [1], many universities started to offer Data Science degrees in the recent decade [1,2]. According to ABET [3,4], Data science draws on knowledge, skills and abilities from computing, mathematics, and statistics, all applied in the context of domains that make use of data. The pervasive reach and multi-disciplinary nature of data
science causes special challenges in uniting traditionally separate disciplines into a coherent approach that produces ethical and well-trained practitioners. Agreement on standards is an important step in the maturation and establishment of any emerging discipline. The interdisciplinary nature of data science brings challenges in uniting traditionally separate disciplines into a coherent approach that produces ethical and well-trained practitioners. ABET CAC (Computing Accreditation Commission) is in the semi-final stage of approving for the first-time program criteria for Data Science within Computing programs [5]. Pilot CAC Data Science accreditation is being planned for the upcoming 2021-2022 accreditation cycle [3,4]. ABET ANSAC (Applied and Natural Science Accreditation Commission) is also working with American Statistical Association in initiating the Data Science accreditation within Applied and Natural Science programs.

Currently ABET CAC accredits the following programs

- Computer Science and Similarly Named Computing programs
- Cybersecurity and Similarly Named Computing programs
- Information Systems and Similarly Named Computing Programs

ABET ANSAC accredits the following programs

- Construction Management and Similarly Named Programs
- Environmental, Health, And Safety and Similarly Named Programs
- Environmental Science and Similarly Named Programs
- Facility Management and Similarly Named Programs
- Geology, Geological Science and Similarly Named Programs

- Health Physics and Similarly Named Programs
- Industrial Hygiene and Similarly Named Programs
- Safety and Similarly Named Programs
- Surveying, Geomatics and Similarly Named Programs

Each ABET Program Criteria cover students, program education objectives, student outcomes, continuous improvement, curriculum, faculty, facilities, and institutional support. CAC and ANSAC both specify general criteria with their programs and specific criteria for each program. This paper focuses on two very important criteria: student outcomes and curriculum. CAC has proposed Data Science Program criteria in student outcomes. ANSAC is still working on Data Science Program criteria at this time.

In this paper, we review two data science programs. The first program is BS in Data Science (BSDS) jointly by Department of Computer Science and Department of Statistics at a state flagship university. The second program is BS in Data Science and Applied Statistics (BSDSAS), a first Data Science and Statistics program initiated by the Department of Mathematical Science at the flagship university's regional campus.

The rest of the paper is organized as follows. In Section 2, one of important and major development of Data Science is discussed. Section 3 focuses on the ABET Criteria in Student Outcomes. The Sections 4 summarizes the general program curriculum criteria for CAC and ANSAC respectively. CAC Data Science program specific curriculum criteria is discussed in Section 5. In Section 6, we review both BSDS and BSDSAS curricula and make recommendations of improvement to satisfy ABET CAC or ANSAC requirement. Section 7 describes how Data Science programs can work with other programs to incorporate application domains as required by ABET. Finally, Section 5 summarizes the paper and gives recommendations for future work.

## 2 Define Data Science

After the National Science Foundation organize a workshop [6] to gather diverse perspectives on Data Science in 2015, the Park City Math Institute 2016 developed interdisciplinary curriculum guidelines for undergraduate programs in data science [7]. The base assertion of the report and proposed curriculum is that data is the core: "The recursive data cycle of obtaining, wrangling, curating, managing and processing data, exploring data, defining questions, performing analyses, and communicating the results lies at the core of the data science experience." The Park City Report further include an outline of the Data Science Major:

1. Introduction to data Science
    a. Introduction to Data Science I
    b. Introduction to Data Science II
2. Mathematical foundations
    a. Mathematics for Data Science I
    b. Mathematics for Data Science II
3. Computational thinking
    a. Algorithms and Software Foundations
    b. Data Curation—Databases and Data Management
4. Statistical thinking
    a. Introduction to Statistical Models
    b. Statistical and Machine Learning
5. Course in an outside discipline

American Statistical Association adopted the curriculum outline in the Park City Report in 2017 [7]. In January 2021, ACM Data Science Task Force competed its final report on Computing Competencies for Undergraduate Data Science Curricula [7] and considered the Park City Report's report as a base for Data Science Major course outline.

## 3 Student Outcomes

### 3.1 CAC Data Science Program Student Outcomes

ABET CAC proposed DS Student Outcomes for review and piloting during the 2021-2022 Accreditation Cycle as follows

1. Analyze a complex computing problem and to apply principles of computing and other relevant disciplines to identify solutions.
2. Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.
3. Communicate effectively in a variety of professional contexts.

4. Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.
5. Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.
6. Apply theory, techniques, and tools throughout the data analysis lifecycle and employ the resulting knowledge to satisfy stakeholders' needs. [DS]

Currently this state flagship university only has been accredited in all its programs in engineering, engineering technology including Computer and Information Technology. However, its Computer Science program is not accredited by ABET CAC. Neither CS or DS list its student outcomes [8]. We assume that both CS and DS will adopt ABET CAC Student Outcomes for its program when applying for ABET Accreditation.

## 3.2 Applied and Nature Science Program Student Outcomes

ABET ANSAC is starting on Data Science program criteria. The following are the Student Outcomes for all applied and nature science programs:

1. An ability to identify, formulate, and solve broadly defined technical or scientific problems by applying knowledge of mathematics and science and/or technical topics to areas relevant to the discipline.
2. An ability to formulate or design a system, process, procedure or program to meet desired needs.
3. An ability to develop and conduct experiments or test hypotheses, analyze and interpret data and use scientific judgment to draw conclusions.
4. An ability to communicate effectively with a range of audiences.
5. An ability to understand ethical and professional responsibilities and the impact of technical and/or scientific solutions in global, economic, environmental, and societal contexts.
6. An ability to function effectively on teams that establish goals, plan tasks, meet deadlines, and analyze risk and uncertainty.

In the future, ANSAC may add some specific student outcomes for Data Science Programs.

The above-mentioned region campus' BS in CS is accredited by ABET CAC for many years. However, its BS in Data Science and Applied Statistics [9] was initially designed as BS in Applied Statistics and focused heavily on statistics components. It is likely that BS in DSAS program will seek accreditation rat ABET ANSAC It lists its Student Learning Outcomes as

- Students who complete the DSAS degree should have sufficient preparation to extract meaning from and interpret real world data and to effectively communicate the results.

It will not be difficult for BS DSAS program to adopt ABET ANSAC Student Outcomes. However, it is very important that the program focus more on graduates' ability to

- understand ethical and professional responsibilities and the impact of technical and/or scientific solutions in global, economic, environmental, and societal contexts. (SO 5)
- function effectively on teams that establish goals, plan tasks, meet deadlines, and analyze risk and uncertainty. (SO 6)

# 4 Program Common Curricula

ABET CAC curriculum requires all computing programs to combine technical, professional, and general education components to prepare students for a career, further study, and lifelong professional development in the computing discipline associated with the program. The curriculum specifies computing topics for all programs and Data Science requirements.

## 4.1 Computing Program Curriculum Requirements

Each program must include mathematics appropriate to the discipline and at least 30 semester credit hours (or equivalent) of up-to-date coverage of fundamental ad advanced computing topics that provide both breadth and depth. All computing programs must include the following topics

1. Techniques, skills, and tools necessary for computing practice.
2. Principles and practices for secure computing.
3. Local and global impacts of computing solutions on individuals, organizations, and society.

### 4.2 Applied and Natural Science Program Curriculum Requirements

Each program must include mathematics appropriate to the discipline and at least 30 semester credit hours (or equivalent) of up-to-date coverage of fundamental ad advanced computing topics that provide both breadth and depth. All computing programs must include the following topics

1. combination of college-level mathematics and sciences (some with laboratory and/or experimental experience) appropriate to the discipline.
2. advanced technical and/or science topics appropriate to the program
3. a general education component that complements the technical and scientific content of the curriculum and is consistent with the program and institution objectives.
4. be prepared for practice in a field of applied or natural sciences through a curriculum culminating in comprehensive projects or experiences based on the cumulative knowledge and skills acquired in earlier course work..

## 5 Data Science Curriculum

### 5.1 Data Science Curriculum as a Computing Program

Data Science and similarly name programs require the following in curriculum

1. At least 45 semester credit hours (or equivalent) of data science course works that must cover
2. Data acquisition
3. Data management
    a. Data preparation and integration
    b. Data analysis
    c. Model development and deployment
    d. Visualization
4. Concepts that span and are applied to the data analysis lifecycle:
    a. Data privacy, governance, and stewardship
    b. Statistics and mathematics
    c. Computing, including substantial coverage of data structures, algorithms, and at least one programming language
5. Advanced data science coursework that provides depth
6. Coverage of at least one application domain area to provide a context for data science activities.
7. A major project that 1) incorporates an application domain area and 2) requires integration and application of knowledge and skills acquired in earlier course work.

### 5.2 Data Science Curriculum in An Applied and Natural Science Program

ANSAC (Applied and Natural Science Accreditation Commission) is still in process of define Data Science program criteria. American Statistical Association is the leading society to define the Data Science criteria under ANSAC. It is expected the curriculum will be very similar to the curriculum criteria under CAC.

## 6 Review Two Data Science Programs

We review two mentioned two programs' curricula and discuss whether a program should apply for CAC or ANSAC accreditation. Further we recommend improvements needed for each program to achiever their accreditation respectively.

### 6.1 Data Science Program Joined by Computer Science and Statistical Programs

Purdue University at West Lafayette, Indiana (PWL) started its join Data Science Program by both Department of Computer Science and Statistics in fall 2017. Table 1 shows the current BS DS curriculum compared with proposed CAC Data Science curriculum, and course mapping to Park City Report requirements. Overall, BS DS curriculum is well designed. We would recommend two improvements for the program:

1. Designs a new course and covers in other courses of principle and practices for secure computing, data privacy, governance, stewardship, and local and global impacts of computing and data science solutions on individual, organization, and society [10].
2. Selects a least one application domain area to provide a context for data science activities.

## 6.2 Data Science Program initiated by Statistics Program

Purdue University Fort Wayne (PFW), a Purdue University's regional campus initially started an undergraduate program in Applied Statistics through its Mathematical Science Department. Later, the program changed to Data Science and Applied Statistics with additional required courses in computer science in fall 2018. Since this program is statistics-heavy, it is expected that BS DSAS will apply for accreditation with ANSAC. Currently ABET ABSAC is still defining its Data Science program criteria. It may not require the coverage of principle and practices for security computing as CAC. However, it is expected that Data Science program specific criteria will be similar between ANSAC and CAC. Table 2 organizes its curriculum related to ABET CAC curriculum, and Park City Report course outline. It appears that BS DSAS program misses some of major 45 semester credits of Data Science courses. We would recommend that BS DSAS program should begin to add the following requirements

1. Adds a course covering data science fundaments, data mining, and data visualization from existing Computer Science program
2. Adds a course covering data privacy, governance, stewardship, and local and global impacts of computing and data science solutions on individual, organization, and society
3. Selects a least one application domain area to provide a context for data science activities.

# 7 Incorporate Application Domain Area

ABET CAC Data Science curriculum especially requires a major project that incorporates an application domain area. Edureka [11] lists the top ten Data Science applications: 1. Fraud and Risk Detection, 2. Healthcare, 3. Internet Search, 4. Targeted Advertising; 5. Website Recommendations; 6. Advanced Image Recognition; 7. Speech Recognition 8. Airline Route Planning; 9. Gaming; 10. Augmented Reality Besides, Data Science has been widely used in both natural and social science theory discovery and analysis, business process improvement, accounting, finance, marketing, banking, agriculture, and pharmaceutical drug discovery, and even government. It appears that many new BS DS programs select Business as an application domain area as in [17]. We should develop more application domain areas.

# 8 Conclusion

The interdisciplinary nature of data science brings challenges in uniting traditionally separate disciplines into a coherent approach that produces ethical and well-trained practitioners. Accreditation process will help in developing agreement on standards. It is very important for every Data Science program to pay a close attention on graduates' ethics and professionalism, and teamwork ability. In this paper, we reviewed both ABET CAC and ANSAC potential program criteria for Data Science and recommend two different Data Science programs' path toward their future accreditation. As Data Science becomes increasingly a major discipline, it is very important to design strong and balanced curricula of both computing science and statistics skills, develop more application domain areas, focus on its impact on individual, organization, and society. In the future, we study and provide guideline for assessment and evaluation processes of Data Science programs

Table 1. PWL BS DS, CAC Curricula and related Park City Report Course Outline

| Course | Title | CAC Knowledge Area | Park City Report | Credits |
|---|---|---|---|---|
| Computer Science Core | | | | |
| CS 18000 | Problem Solving and Object-Oriented Programming | Programming | | 4 |
| CS 18200 | Foundations of Computer Science | Programming | | 3 |
| CS 38003 | Python Programming | Programming language | | 1 |
| CS 25100 | Data Structures & Algorithms | Data Structures & Algorithms | | 3 |
| Statistics Core | | | | |
| STAT 35500 | Statistics for Data Science | Statistics | | 3 |
| STAT 41600 | Probability | Statistics | | 3 |
| STAT 41700 | Statistical Theory | Statistics | | 3 |
| Data Science Core | | | | |
| CS 24200 | Introduction to Data Science | Data Science | | 3 |
| CS 37300 | Data Mining and Machine Learning | Advanced Data Science | | 3 |

| | | | | |
|---|---|---|---|---|
| CS 49000 | Large Scale Data Analytics | Advanced Data Science | | |
| Capstone | | | | |
| CS 49000 | Data Science Capstone | | | 0-3 |
| Computer Science Electives (2 courses) in software engineering, numerical methods, database systems, analysis of algorithms, theory of computation, cryptography, artificial intelligence, web technology, or data visualization | | | | 6 |
| Statistics Elective (1 course) in time series, stochastics processes, statistical programming, and data management, applied regression analysis, statistical quality control, experiment design, sampling and survey techniques, statistical methodology | | | | 3 |
| Ethics Elective (1 course) in Data Science & Society: Ethical, Legal, Social Issues, Ethics for Technology, Engineering, and Design, or Ethics of Data Science | | | | 3 |

Table 2. PFW BS DSAS. CAC Curricula and related Park City Report Course Outline

| Course | Title | CAC Knowledge Area | Park City Report | Credits |
|---|---|---|---|---|
| Mathematics Core | | | | |
| MA 16500 | Analytic Geometry & Calculus I | Mathematics | | 3 |
| MA 16600 | Analytic Geometry & Calculus II | Mathematics | | 3 |
| MA 26100 | Multivariate Calculus | Mathematics | | 3 |
| MA 35100 | Elementary Linear Algebra | Mathematics | | 3 |
| Computing Core | | | | |
| IST 16000 | Foundation and Role of Information Systems | Information Systems | | 3 |
| IST 27000 | Data And Information Management | Database systems | | 3 |
| CS 29200 | Python Programming for Data Analytics | Programming | | 3 |
| Statistics Core | | | | |
| STAT 49000 | Multivariate Data Analysis | Statistics | | 3 |
| STAT 51100 | Statistics Methods | Statistics | | 3 |
| STAT 51200 | Applied Regression Analysis | Statistics | | |
| STAT 51600 | Basic Probability and Applications | Statistics | | |
| STAT 51700 | Statistical Inference | Statistics | | |
| STAT 51800 | Introduction To Statistical Learning | Statistics | | |
| Capstone | | | | |
| STAT 4xxxx | Data Analysis Capstone | | | 3 |
| Electives (3 courses) in<br>Computer Science: data mining, data visualization, machine learning<br>Statistics: stochastic processes, statistical quality control, experiment design, time series and applications | | | | 9 |

ACKNOWLEDGMENTS

This work is supported in part by the Purdue University Integrative Data Science Initiative, Purdue University Fort Wayne Summer Instructional Development Grant (2021), and the Indiana Data Mine Initiative.

## REFERENCES

[1] LinkedIn. 2019. Emerging Jobs Report. https://business.linkedin.com/content/dam/me/business/en-us/talent-solutions/emerging-jobsreport/Emerging_Jobs_Report_U.S._FINAL.pdf.

[2] Zackary Bennett, Why Colleges Are Offering Data Science Programs, *U.S. News Best Colleges 2021 Guidebook*, U.S. News & World Report L.P.

[3] S. Mathur. The List of all U.S. Colleges with a Data Science Major, 2019 (accessed September 3, 2019).

[4] ABET, Diving into Data Science, September 25, 2020.

[5] ABET, Data Science Program Accreditation, September 25, 2020

[6] CSAB. Data Science Program Criteria Approved for First Reading, August 28, 2020.

[7] Lillian Cassel and Heikki Topi. 2015. Strengthening Data Science Education Through Collaboration. Workshop on Data Science Education, October 2015.

[8] J.R.S. Blair, et al., Establishing ABET Accreditation Criteria for Data Science, The 52nd ACM Technical Symposium on Computer Science Education, SIGCSE '21, March 2021, pages 535--539. ACM, 2021.

[9] R.D. De Veaux et. al. 2017. Curriculum Guidelines for Undergraduate Programs in Data Science. *Annual Review of Statistics and Its Application* 4, 1 (2017), 15–30. https://doi.org/10.1146/annurev-statistics-060116-053930

[10] ACM Data Science Task Force. 2021. Computing Competencies for Undergraduate Data Science Curricula. https://dstf.acm.org/

[11] ABET, Inc. 2020. Criteria for Accrediting Computing Programs. Effective for Review During the 2020-21 Accreditation Cycle. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-computing-programs-2021-2022/.

[12] ABET, Inc. 2020. Criteria for Accrediting Computing Programs. Effective for Review During the 2020-21 Accreditation Cycle. https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-applied-and-natural-science-programs-2021-2022/.

[13] xxx University, BS DS program description, 2021
[14] yyy University, BS DSAS program description, 2021
[15] J.S, Saltz, N.L. Dewar, and R. Heckman, Key Concepts of a Data Science. Ethics Curriculum, The 49th ACM Technical Symposium on Computer Science Education, SIGCSE '18, February 2018, pages 952--957. ACM, 2018.
[16] Edureka, Top 10 Data Science Applications, https://www.edureka.co/blog/data-science-applications/, July 15, 2021
[17] I.B. Hassan, et al., Data Science Curriculum Design: A Case Study, The 52nd ACM Technical Symposium on Computer Science Education, SIGCSE '21, March 2021, pages 529--534. ACM, 2021.