Curricular Analytics Project

Curriculum Report: University of Arizona

Curricular Analytics Project (CAP) Leadership Team

March 11, 2025

CAP Leadership Team

Chaouki T. Abdallah
President
Lebanese American University

Melika Akbarsharifi
Graduate Research Assistant
University of Arizona

Roxana Akbarsharifi
Graduate Research Assistant
University of Arizona

Steven P. Dandaneau
Associate Provost & Executive Director of UERU
Colorado State University

David M. Dueber
Research Assistant Professor, Herb Innovation Center
University of Toledo

Hayden W. Free
Damour Systems, PBC

Gregory L. Heileman
Vice Provost, Undergraduate Education
University of Arizona

Erin Helbig
Damour Systems, PBC

Kristina Manasil
Graduate Research Assistant
University of Arizona

Rick Miranda
Senior Vice President
Colorado State University
Vice Chancellor For Collaboration & Integration
Colorado State University System

Ahmad Slim
Research Scholar
University of New Mexico

Michael D. Toland
Executive Director, Herb Innovation Center
University of Toledo

This research was conducted as a part of the Curricular Analytics Project: 2021-2025, an effort funded by the Ascendium Education Group through a grant provided to the Association for Undergraduate Education at Research Universities (UERU).

Contents

1 Introduction

As a participating member of the Curricular Analytics Project (CAP) research effort, this report was created for the University of Arizona and contains curriculum-focused analyses that utilize the extensive data sets provided by each participating institution. This corpus of data contains a unique collection of higher education information, and the participating institutions are to be commended for the manner in which they worked through the extensive, arduous, and time-consuming data collection and validation tasks associated with this effort. Never before has such an extensive data set been assembled in higher education for the purpose of studying the impact curricula have on student success outcomes, and the results we have been able to obtain, after barely "scratching the surface" with analyses, are significant. Over time, through additional analyses, we expect this work will provide profound insights into the role curricular structures play in shaping student success outcomes.

In this report we consider how the complexity of curricula varies according to academic areas of study, both broadly across the wide variety of fields in higher education, and in some cases more narrowly within the specific disciplines that comprise a field. A fundamental question related to student curricular complexity and success is, are some academic programs inherently more difficult to complete than others? If so, then additional questions immediately follow. For instance, can we quantify the differences between academic program difficulties? The ability to do so would allow us to formalize common beliefs that are often stated without firm factual bases. For example, it is not uncommon for students to hear advice regarding the difficulty of completing programs in particular academic fields; such as, programs in field X are the hardest programs on campus to finish on time. In this report we provide a framework for addressing these questions that uses a formal method for measuring program curricular complexity based on the structural properties of a curriculum (as opposed to instructional properties such as the quality of teaching and support services). A curricular complexity metric is applied to a large corpus of curricular data collected from the wide variety universities involved in this study, allowing us to compare and contrast the structures of academic curricula according to different academic areas of study. We demonstrate that it is indeed possible to quantify the complexity differences that exist between the different fields of academic study in our data set. Furthermore, we show that it is possible to characterize these differences in statistically meaningful ways; that is, in a manner that we believe should prove useful in guiding curricular design and reform efforts aimed at facilitating student success.

This report is organized as follows. Section 2 provides a brief description of the CAP work, including a list of all participating universities, along with their reported student success outcomes. Next, the metrics used to quantify the complexity of the curricula considered in this study are described. After that, the Classification of Instructional Programs (CIP) coding system, used throughout the United States, is described. The CIP coding system is used in this report to distinguish between various academic fields, as well as the disciplines within these fields, when comparing and contrasting the complexities of curricua. Section 3 provides numerous results derived from analyzing the CAP curricular data set. First we observe how the complexity distribution of the data set as a whole resembles the power law distribution, and we describe the implications of this observation. Next we consider the complexity of academic curricula at the field level, and for a few fields we consider the complexity of the disciplines within these fields. It is important to note that this report focuses soley on contexualizing the complexity of academic curricula relative to one another across many different fields of study, and across many institutions, using data from the numerous universities described above. In a future report, we will more formally quantify the impact curricular structures have on student success outcomes. In Section 3.4 we provide what we believe is the most useful portion of this report, namely, detailed analyses of all the curricula at the University of Arizona. These analyses have been provided in a manner that allows you to meaningfully compare the complexity of the curricula at your institution to those at the other institutions in this study, accounting for the field of study associated with each curricula. Specifically, these university-specific analyses are provided at the field level in the form of charts that allow you to compare the curricular complexity of all programs at your institution relative to those at all of the other institutions involved in the CAP study.

A maxim heard throughout higher education is, "faculty own the curriculum," and we concur; faculty, the disciplinary experts, must determine the specific content in their curricula to ensure the quality of the programs they oversee. A corollary to this maxim is also worth noting, namely, "faculty inherit the curriculum." Specifically, the curriculum for a given program has typically evolved over time through many prior generations of faculty. This is important to recognzie because the rationale for prior curriculum changes may have been based upon precepts no longer relevant in the field today, or perhaps they were centered around the expertise of facutly no longer associated with the program. A further complication is the history of curricular changes is often not adaquately documented, and the rationle behind particular curricular modifications is simply lost to time. Because faculty owe it to their students to create curricula that prepare them for the futures they will face, they must periodically review their curricula dispite these difficulties. Furthermore, given the current pace of innovation, the pressures associated with this challenge are unlike that felt by any prior generation of faculty. Faculty today are under pressure to adapt their curricula in response to disruptive technologies that appear at an ever increasing rate. We believe the curricular analytics framework and capabilities described in this report provide faculty and curriculum committees with a valuable toolset that can assist them in this difficult work.

2 Background

The Association for Undergraduate Education at Research Universities (UERU) led the CAP research effort, and involved many of the UERU member institutions. CAP partner universities are distinguished by a commitment to data-informed improvement of undergraduate student success, including a willingness to intentionally assess and improve their curricular structures, a heretofore largely hidden source of barriers to equitable student success. The UERU CAP effort sought comprehensive data on curricular structures, as well as ten years of student success data, from 31 U.S. research universities. At the initiation of the study, at each participating university, 30% or more of the full-time undergraduate students were U.S. Federal Government Pell Grant recipients.

Collectively, the 31 institutions involved in this study are all public universities, except for Clark Atlanta University, a private institution in the historically black colleges and universities (HBCUs) category. A number of the other institutions in the study are Hispanic Serving Institutions (HSIs), flagships, land-grants, and urban-serving research universities, including in the last category, Morgan State University, a public HBCU. Together, these institutions enroll over three-quarters of a million undergraduate students annually and on average achieve 81% first-to-second-year retention rates (with a range of 61% to 94%), and 46% and 64% four- and six-year graduation rates, respectively (within ranges of 21% to 75% and 42% to 88%, respectively). Table 1 provides university details [2]. It is worthwile to note that the 81% retention and 46% and 64% four- and six-year graduation rate outcomes cited above for partcipating institutions are on par with or exceed overall U.S. averages. This is likely a result of research university selectivity as well as a testament to their commitment to college access and equitable graduation success outcomes.

2.1 Curricular Complexity Metrics

The foundation for this reseach effort is the recognition that the curriculum a student must complete in order to earn a degree is the most fundamental element of student success. All students in a particular academic program, regardless of their prior preparation, demographics, participation in extracurricular activieis and high impact practices, etc., must complete the same curriculum in order to receive their degrees. In this sense, a student's ability to progress through a curriculum is the most fundamental measure of success, and therefore all student success interventions should be measured relative to their ability to influence this progress. In short, the curriculum matters, yet it is perhaps one of the most overlooked components of student success.

In this study, a *curriculum* refers to the set of courses (along with the corresponding set of course prerequisites) that, if successfully completed, would allow a student to earn the degree associated with the curriculum. An example electrical engineering curriculum is provided in Figure 1 (a). This curriculum is represented as a graph, where the vertices are the required courses in the curriculum, and the directed edges (arrows) between the vertices correspond to prerequisites. That is, the course on the source end of the directed edge is a prerequisite for the course on the destination end. Directed edges drawn as dashed lines correspond to co-requisites.

The complexity of each curriculum was computed using a unitless graph-theoretic metric imposed by the pre- and co-requisite relationships between the courses in a curriculum.

Institution	Full-Time UGrad	1st-Yr Retention	4-, 6-Yr Grad. Rate
Clark Atlanta University	3,427	71%	34%, 48%
Colorado State University	25,777	86%	47%, 67%
Florida International University	45,688	91%	59%, 79%
Florida State University	32,936	94%	73%, 83%
Georgia Mason University	27,014	85%	48%, 69%
Georgia State University	28,924	78%	34%, 54%
Kent State University	20,418	78%	49%, 65%
Montclair State University	17,290	81%	46%, 64%
Morgan State University	7,609	74%	22%, 42%
New Mexico State University	11,591	72%	32%, 52%
Rutgers University-Newark	7,511	82%	42%, 68%
Temple University	24,016	84%	57%, 75%
University of Arizona	38,751	86%	51%, 66%
University of California, Davis	31,532	92%	69%, 85%
University of California, Irvine	28,662	91%	73%, 86%
University of California, Riverside	22,911	87%	65%, 77%
University of California, San Diego	33,096	93%	75%, 88%
University of Central Florida	58,662	93%	50%, 75%
University of Illinois, Chicago	21,807	78%	40%, 60%
University of Memphis	16,708	75%	31%, 49%
University of New Mexico	16,375	72%	36%, 52%
University of NC, Charlotte	23,641	84%	50%, 68%
University of South Florida	38,047	90%	63%, 75%
University of Texas, Arlington	30,791	74%	35%, 57%
University of Texas, Dallas	21,617	87%	56%, 71%
University of Texas, El Paso	20,165	75%	21%, 46%
University of Texas, San Antonio	29,686	80%	32%, 51%
University of Texas, Tyler	6,971	61%	36%, 47%
University of Toledo	11,965	75%	35%, 58%
Utah State University	24,835	74%	30%, 55%
Washington State University	22,612	81%	43%, 62%
Total	751,035		
Average	24,227	81%	46%, 64%

Table 1: Full-time student enrollments, retention rates, graduation rates for academic year 2022 at the 31 institutions participating in the CAP study. Data extracted from the National Center for Education Statistics IPEDS

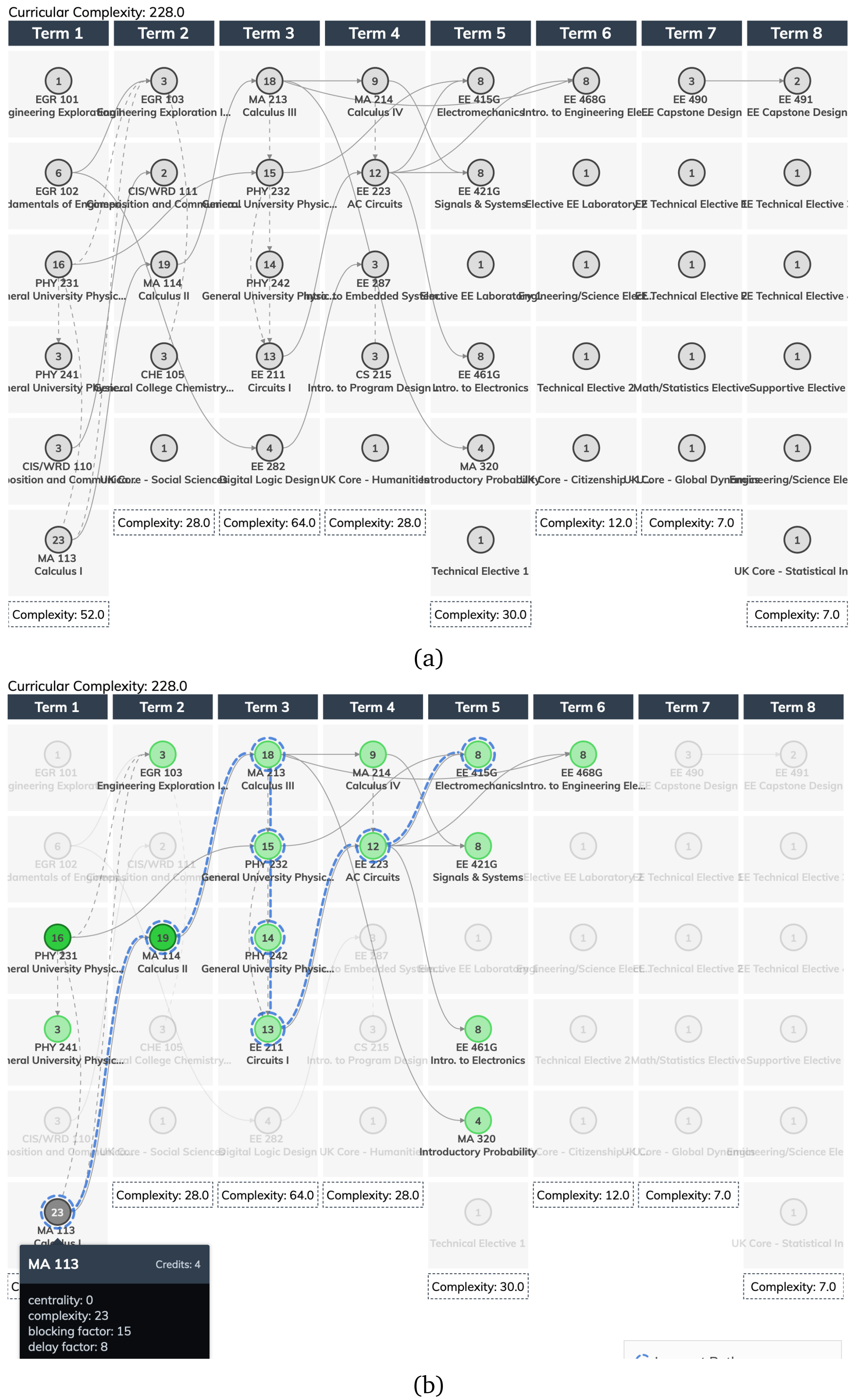


Figure 1: **(a)** An example curriculum, organized as a degree plan over eight terms. The courses in the curriculum are the vertices, and the prerequisites are the directed edges. **(b)** Highlighting the Calculus I course in this curriculum shows Calculus I blocks 15 other courses in the curriculum (shown in green), and the longest path in the curriculum that includes Calculus I (shown as a blue dashed line) has length 8.

This metric, referred to as *structural complexity*, emphasizes that it is the structure of the curriculum we are considering, *not* the academic rigor of particular subject areas. Structural complexity involves two factors. First, each course c in a curriculum is assigned a *blocking factor* which is simply the number of courses a student is precluded from taking, due to pre- and co-requisite constraints, until they have successfully completed course c. In Figure 1 (b), the blocking factor of the Calculus I course is 15. The second factor, called the *delay factor* is determined by the longest pathway in the graph that includes course c. In Figure 1 (b), the delay factor of the Calculus I course is 8. The structural complexity of a course c is determined by adding its blocking and delay factors, and the structural complexity (or more simply complexity) of a curriculum is determined by summing all the course complexities in a curriculum. In Figure 1 (b), the complexity of the Calculus I course is 23, and the complexity of the entire curriculum is 228.

Quantitative research, using a significant data set of curricula and student outcomes, definitively shows that, on average, curricula with higher complexity take more time (i.e., are more difficult) for students to complete [9, 10]. That is, if all other factors are held constant (e.g., quality of instruction, availability of tutoring resources, etc.) then programs with higher curricular complexity will take longer to complete than programs with lower curricular complexity. Furthermore, preliminary results utilizing the student outcomes data collected as a part of the CAP work (not considered in this report) indicate that not only does curricular complexity impact time to degree, but it does so in ways that magnify inequites. This research will be provided in a future report.

In order to consider how the complexities of academic programs vary across and within different fields of study, each academic program considered in this study was classified by the universities involved in the study using Classification of Instructional Programs (CIP) codes. Every postsecondary school in the U.S. receiving federal student financial aid is required to match their academic programs to CIP codes, and to periodically report specific program data to the federal government according to these CIP codes.

2.2 Classification of Instructional Program (CIP) Codes

The National Center for Education Statistics (NCES) within the U.S. Department of Education created the CIP taxonomic coding scheme in order to support data collection and reporting related to fields of study [1]. All institutions in the U.S. providing reports to the NCES via the Integrated Post-secondary Education Data System (IPEDS) must use CIP codes to identify their instructional programs. The CIP codes in this study are based upon the update released by the NCES in 2020 (the CIP system is updated every ten years).

CIP Code Organization. The CIP taxonomy is organized hierarchically into three levels:

1. **Two-digit series (##.).** The most general grouping of related programs, representing a broad field of programs. E.g.,

   04. – Architecture and Related Services

   14. – Engineering

   52. – Business, Management, Marketing, and Related Support Services.

2. **Four-digit series (##.##).** An intermediate grouping of programs having comparable content and objectives. E.g.,

   04.04 – Architecture and Related Services/Environmental Design

   14.09 – Engineering/Computer Engineering

   52.03 – Business, Management, Marketing, and Related Support Services/Accounting and Related Services.

3. **Six-digit series (##.####).** The most detailed program classification within the CIP system, representing specific instructional programs. E.g.,

   04.0401 – Architecture and Related Services/Environmental Design/Architecture

   14.0903 – Engineering/Computer Engineering/Computer Software Engineering

   52.0303 – Business, Management, Marketing, and Related Support Services/Accounting and Related Services/Auditing.

In the two-digit CIP series, there are 48 codes with associated titles. Beneath each of these codes, there are a variable number of four- and six-digit codes. For the purposes of this study, we use the term *field* to refer to each of the two-digit CIP series titles, and the term *discipline* to refer to each of the four-digit CIP series titles. Finally, at a given institution, we assume that if two curricula have been provided with the same six-digit CIP code, they are concentrations (emphases, tracks, etc.) within the same degree program. In which case, we assign the program a curricular complexity that is the average value of its concentrations' curricular complexities.

3 Results

In this portion of the report we provide various analyses related to the complexities of the undergraduate academic programs considered in this study. First, in Section 3.1 we consider the entirety of the curricular data set as a whole, and we provide a number of important observations, as well as cautionary advice, about the distribution of program curricular complexities when considering all fields across all of the institutions in this study. Next, in

Section 3.2, we consider the distribution of curricular complexities within particular fields across all universities. Furthermore, we make the case that at this level of consideration it is possible to make meaningful comparisons regarding the complexities of programs within the same field of study. We believe such comparisons can provide useful information to those involved in curricular design and academic program review. In Section 3.3, we consider the distribution of curricula complexities for the disciplines within particular fields. At this level of consideration, the empirical distribution functions often appear to have a normal (i.e., Gaussian) appearance, potentially making comparisons at the field level even more useful. However, we are data limited at this level detail, and therefore the statistics from these distributions should be treated as unreliable measures. By collecting additional data at the discipline level, it would become possible to make discipline-level comparisons. Finally, in Section 3.4 we provide plots showing the distribution of curricular complexities across all undergraduate programs at your institution. In addition, we provide a useful plot we created that shows how all of the program complexities at your institution within a particular field of study compare those programs at other institutions in the same field of study.

3.1 Complexity Distribution – All Curricula

A total of 3,830 curricula were considered in this study. Table 2 lists the number of undergraduate curricula supplied by each institution in this study, including programs with multiple concentrations. If a program supplied multiple concentrations, the curricular complexities for each concentration was computed, and the program curricular complexity was assigned the average value over all of a program's concentrations.

Figure 2 (a) is a histogram of the curricular complexities associated with all of the programs in the study. Across the institutions in this study, there are many programs at the lower end of the complexity distribution, and significantly fewer programs at the higher end of the complexity distributions. Furthermore, when considering only the complexities of the programs at a single institution, many have a shape very similar to that of Figure 2 (a). That is, all of the institutions in this study have far more lower complexity programs than higher complexity programs. The complexity histogram for your institution is provided in Section 3.4 of this report.

In Figure 2 (b) we provide an empirical probability density function estimation created using the complexities of all programs considered in this study. Specifically, we treat the program complexity values as a random variable x, and we use the *kernel density estimation* (KDE) technique to estimate the probability density function associated with the aforementioned random variable. This methodology uses kernel functions as weights in order to create a smoothed version of the histogram shown above. Specifically, let $x_1, \ldots, x_n$ denote i.i.d. samples drawn from some unknown probability distribution p. Then, the *kernel*

Institution	**# of curricula**
Clark Atlanta University	27
Colorado State University	72
Florida International University	76
Florida State University	100
George Mason University	74
Georgia State University	54
Kent State University	110
Montclair State University	68
Morgan State University	43
New Mexico State University	81
Rutgers University-Newark	34
Temple University	139
The University of Texas at Arlington	63
UC Davis	97
Univeristy of North Carolina at Charlotte	76
University of Arizona	138
University of California Irvine	78
University of California Riverside	94
University of Central Florida	82
University of Illinois Chicago	78
University of Memphis	53
University of New Mexico	86
University of South Florida	90
University of Texas at Dallas	48
University of Texas at El Paso	66
University of Texas at San Antonio	77
University of Texas at Tyler	35
University of Toledo	84
Utah State University	65
Washington State University	115

Table 2: The number of undergraduate curricula supplied by each institution in the study.

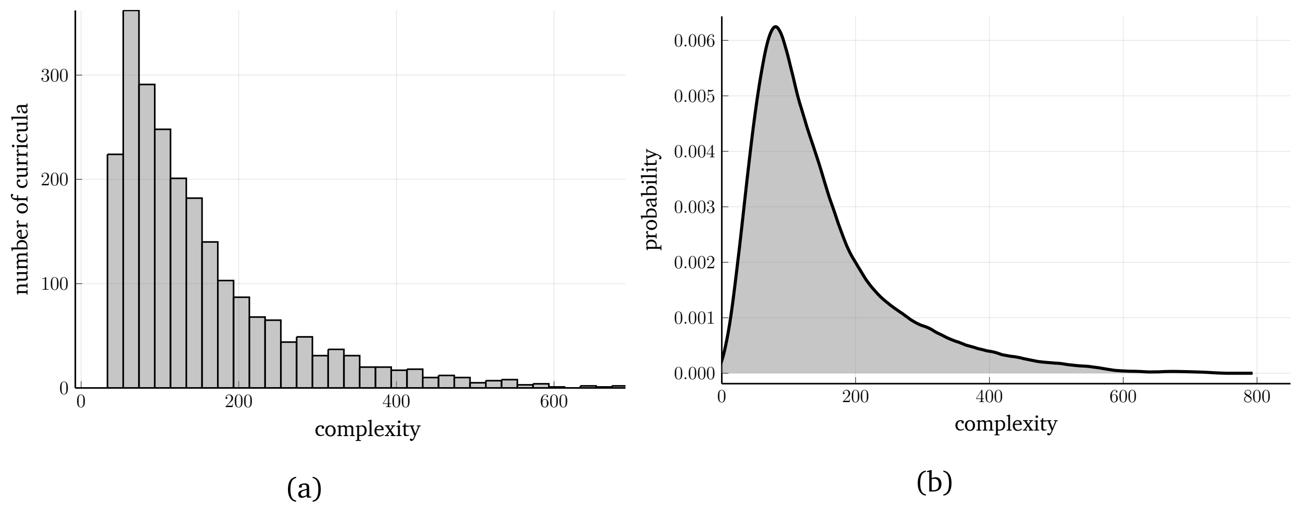


Figure 2: Histogram of the complexities of all programs considered in this study. **(a)** Bins correspond to the total number of curricula within a particular range of complexity values. **(b)** An empirical probability distribution function created from the program complexities using kernel density estimation techniques.

density estimator for p is given by

$$\hat{p}(x) = \frac{1}{nh} \sum_{i=1}^{n} K_n \left(\frac{x - x_i}{h} \right),$$

where K is a non-negative function known as the *kernel*, and h is a smoothing parameter known as the *width* [14].

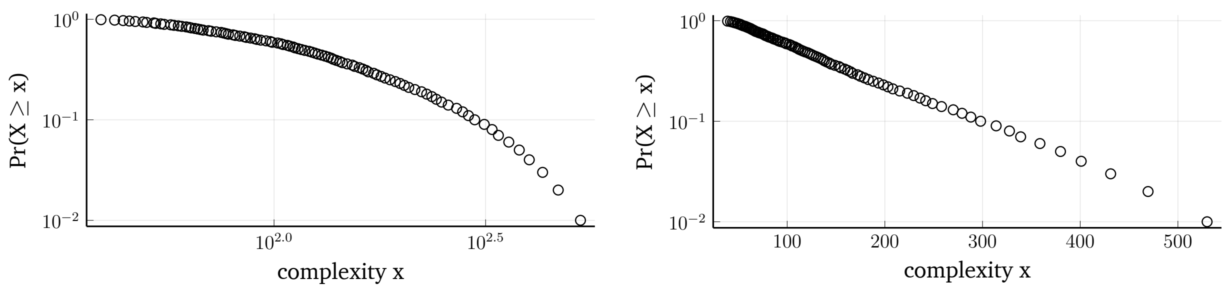
The shape of the distribution shown in Figure 2 (b) resembles that of the *power law*, a distribution with very important and unique properties that shows up when studying a wide variety of both natural and man-made phenomena [8, 11, 12]. These distributions are characterized by long tails, i.e., most samples will cluster around a smaller value, with occasional “black swan” events leading to samples in the tail of the distribution that are far away from this cluster. A power law relationship simply means the probability varies as a power of the argument x, i.e.,

$$p(x) = Cx^{-\alpha},$$

with exponent $\alpha > 1$ (referred to as the scale parameter), where C is a constant, and it is assumed $x > x_{min} > 0$; that is, the power law relationship only holds above the value x_{min}. An important and unique characteristic of the power law distribution is its *scale invariance*; that is, scaling x by a constant factor a will not change the fundamental shape of the distribution,

$$p(ax) = C(ax)^{-\alpha} = a^{-\alpha} p(x) \propto x^{-\alpha}.$$

This agrees with the observation that the complexity distributions at all of the individual universities in this study resemble to some extent Figure 2 (b).



(a) (b)

Figure 3: The empirical distribution (as a complementary CDF) $Pr(X \geq x)$ for the entire data set plotted on (a) a doubly logarithmic scale, and (b) a semi-logarithmic logarithmic scale.

Given that

$$\log(p(x)) = \alpha \log x + C,$$

where C is a constant, it follows that one way of detecting the possible presence of a power law distribution is to plot samples drawn from a distribution on a log-log scale. If the these samples fall on a line, there is evidence (a necessary but not sufficient condition) of the power law distribution. Furthermore, the slope of this line provides an estimate of the value α. A log-log plot of the entire curricular data set is shown in Figure 3 (a), and a best fit line to these data points has slope -1.08. However, we should mention that the detection of the power law in empirical data is a notoriously difficult problem, due to many factors that influence real-world data sets [6]. For instance, the fact that the data points in Figure 3 (a) deviate from a line in the extreme tail of the empirical distribution may be attributed to the fact that undergraduate curricula in the United States are limited in scale. In particular, undergraduate curricula must fit into a four-year time-frame, and are generally limited to 120 credit hours. Thus, the complexity values in the curricular data set are naturally truncated at values above roughly 700. Indeed, the shape of Figure 3 (a) is similar to those of truncated power law distributions [3, 7].

An alternative distribution that may also fit our curricular data set is given by the *exponential distribution*,

$$p(x) = \lambda e^{-\lambda x},$$

where $\lambda > 0$ is referred to as the rate parameter. Because

$$\log(p(x)) = -\lambda x + C,$$

data drawn from an exponential distribution should appear as a straight line on a semi-logarithmic plot. In Figure 3 (b) we show the entire curricular data set plotted on a semi-logarithmic scale, and we observe that the data points fall on a nearly straight line.

Thus, there is evidence that the complexity of the curricular data in this study is either distributed according to a truncated power law distribution or an exponential distribution (and additional research will consider this in more detail). For the purposes of this report, however, either distribution leads to the same conclusion. Specifically, both the mean of the exponential distribution, and the mean of the power law distribution (if it exists), are highly sensitive to outliers or extreme values, meaning even a single very large value can significantly shift the calculated mean, making it more susceptible to changes in the data compared to other measures like the median. That is, due to the skewed nature of these distributions, the mean is heavily influenced by extreme values, leading to a larger difference between the mean and median. For this reason, the median is often considered a more robust measure of central tendency in these cases, as it is less affected by outliers.

Why does this matter? Given the longs tails associated with the distributions described above, the average curricular complexity value is easily skewed by a small number of highly complex curricula. Thus, **any statements made regarding the average curricular complexity value at a particular institution, or across any collection of institutions, when all fields of study are considered, should be considered highly unreliable**; such comparisons are ill-advised. However, we will show in the next section that comparisons between programs within a given field of study do make sense, and therefore they can be useful in guiding curricular design and reform efforts.

3.2 Complexity Distribution – By Field

There are 36 unique two-digit CIP codes associated with the academic programs in this study, each corresponding to a different academic field, as discussed earlier. In Figure 4 we provide a histogram that shows the number of curricula in the study that fall under each two-digit CIP series. The eleven fields with the largest number of curricula in this study (in descending order) include:

1. Business, Management, Marketing, etc (244 programs, CIP Code 52)
2. Visual and Performing Arts (208 programs, CIP Code 50)
3. Engineering (190 programs, CIP Code 14)
4. Social Sciences (161 programs, CIP Code 45)
5. Education (148 programs, CIP Code 13)
6. Biological and Biomedical Sciences (137 programs, CIP Code 26)
7. Foreign Languages, Literatures, and Linguistics (135 programs, CIP Code 16)
8. Physical Sciences (108 programs, CIP Code 40)
9. Area, Ethnic, Cultural, and Gender Studies (94 programs, CIP Code 05)
10. Communication, Journalism, and Related Programs (73 programs, CIP Code 09)

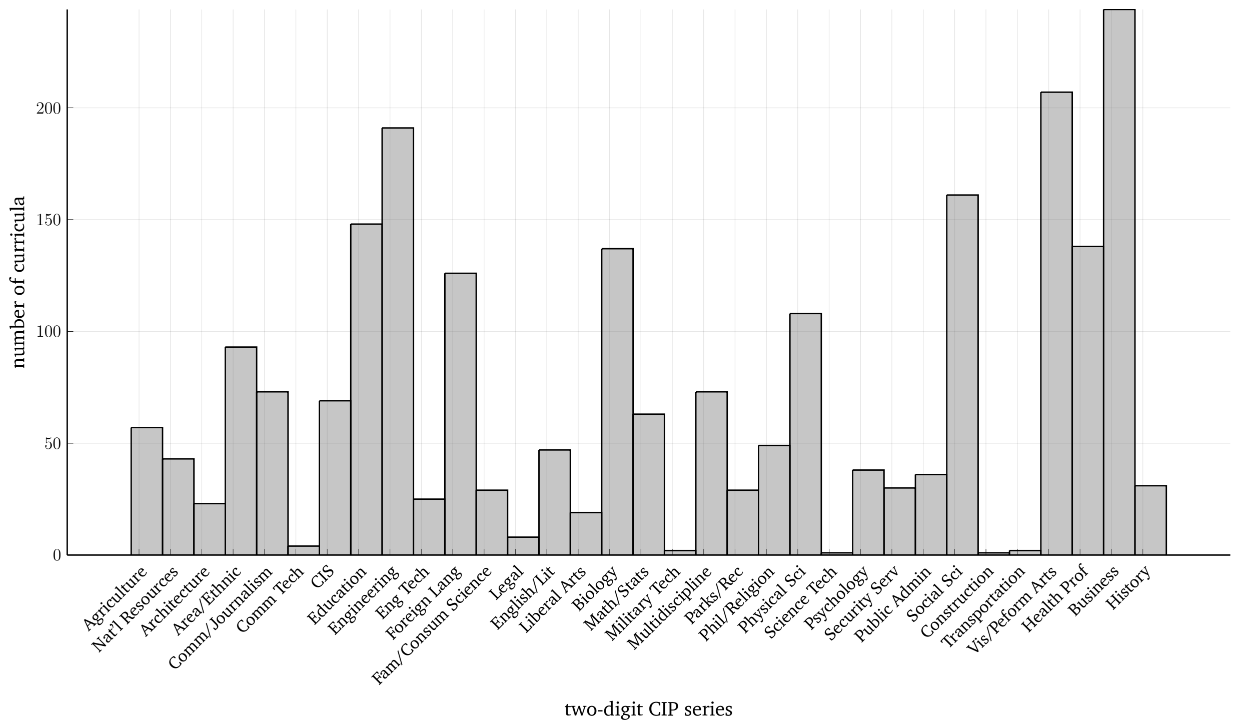


Figure 4: Number of curricula by academic field (i.e., two-digit CIP series).

11. Multi/Interdisciplinary Studies (73 programs, CIP Code 30)

Notice in Figure 4 that there are also a number of fields with very few curricula in the study data set. Specifically, the following six fields contain fewer than nineteen curricula in our study:

1. Legal Professions and Studies (8 programs, CIP Code 22)
2. Communications Technologies/Technicians and Support Services (4 programs, CIP Code 10)
3. Military Technologies (2 programs CIP Code 29)
4. Transportation and Materials Moving (2 programs, CIP Code 49)
5. Science Technologies/Technicians (1 program, CIP Code 41)
6. Construction Trades (1 program, CIP Code 46)

These fields of study are therefore *not* included in the field-level analyses that follow.

Next, let us treat the curricular complexities of all programs sharing a given two-digit CIP code as a random variable. This will allow us to estimate the probability density functions associated with the curricular complexities in each discipline. These estimates provide a principled means for comparing the complexities of academic programs, both within a field, as well as across different fields. Potentially allowing us to answer questions such as,"Is my biology program more complex than the biology programs at other similar institutions?" and "From a curricular perspective, is biology (when considered across multiple institutions) a more complex field of study than business, and if so, by how much?"

Below we provide figures showing the complexity distributions for all of the fields associated with the programs in our study, excluding the fields containing fewer than nineteen curricula. In Figures 5–31, the left hand side shows the raw data for a particular field, represented as a histogram, and the right hand side shows a KDE estimate of the empirical distribution for the field. In order to facilitate comparisons, the plots in each of these figures use the same x-axis scale.

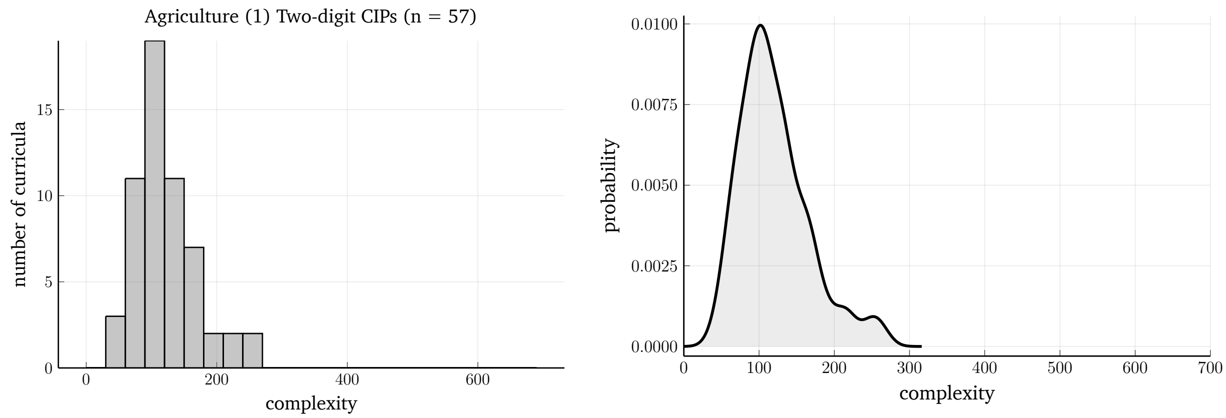


Figure 5: Two-digit CIP series 1, *Agricultural/Animal/Plant/Veterinary Science and Related Fields*.

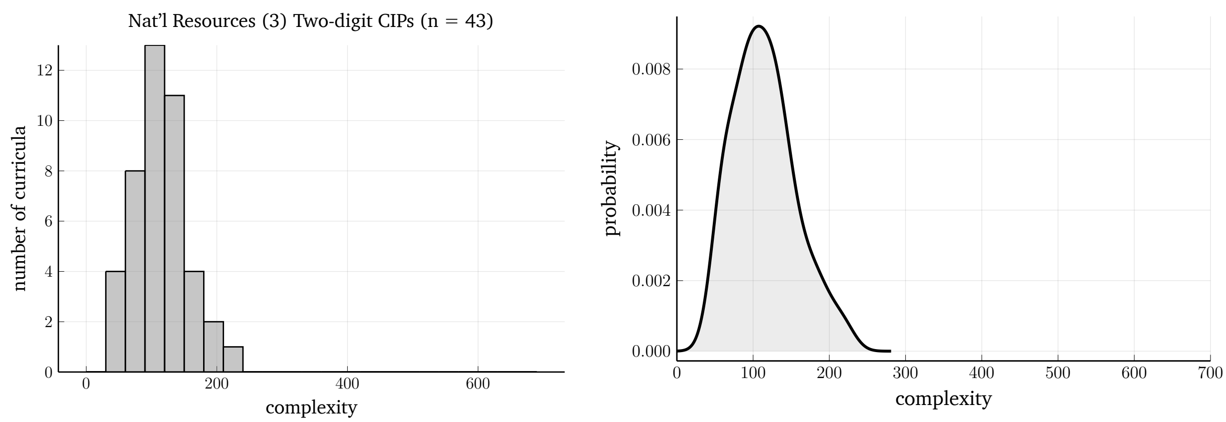


Figure 6: Two-digit CIP series 3, *Natural Resources and Conservation*.

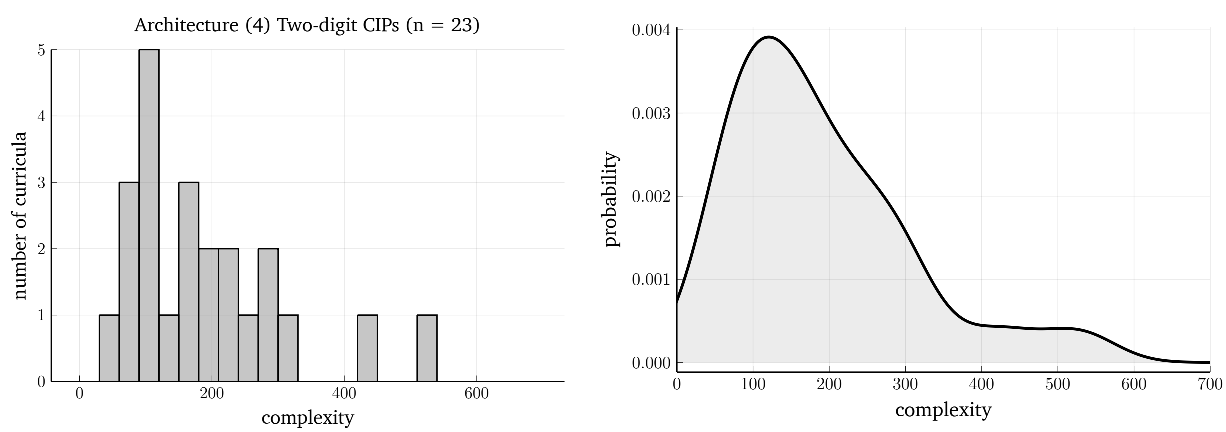


Figure 7: Two-digit CIP series 4, *Architecture and Related Services*.

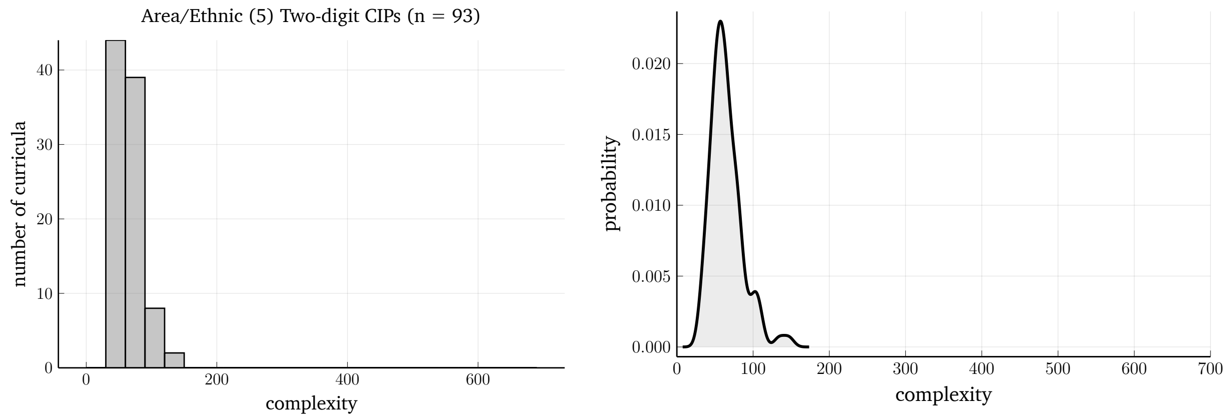


Figure 8: Two-digit CIP series 5, *Area, Ethnic, Cultural, Gender, and Group Studies*.

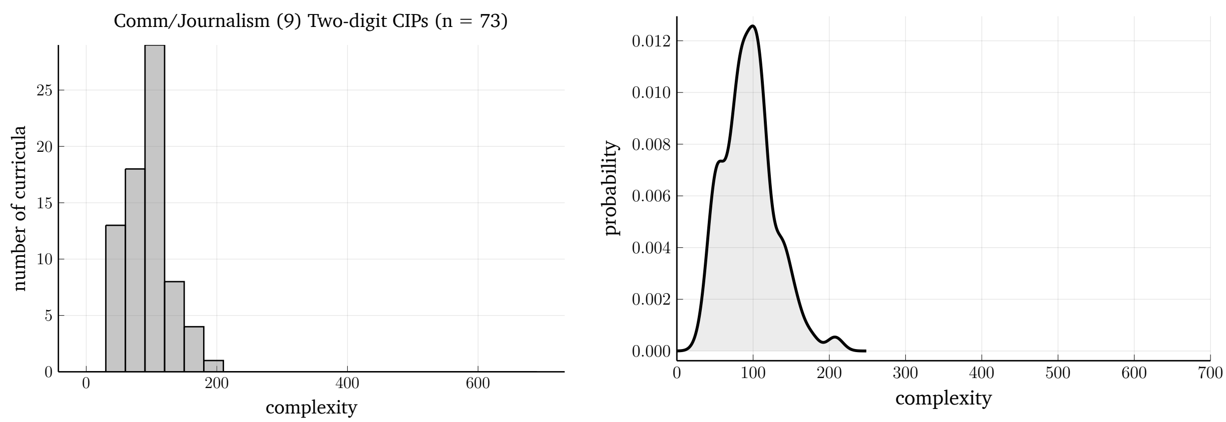


Figure 9: Two-digit CIP series 9, *Communication, Journalism, and Related Programs*.

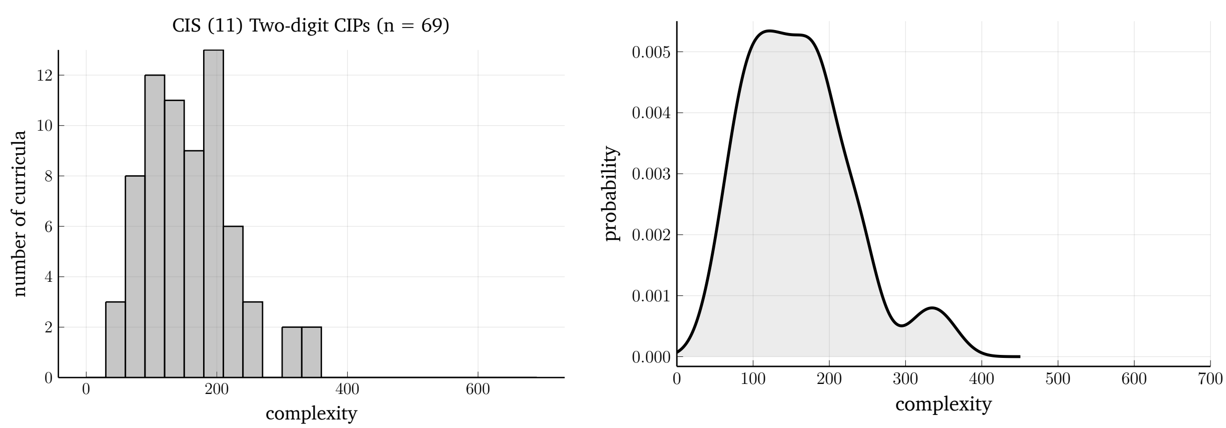


Figure 10: Two-digit CIP series 11, *Computer and Information Sciences and Support Services*.

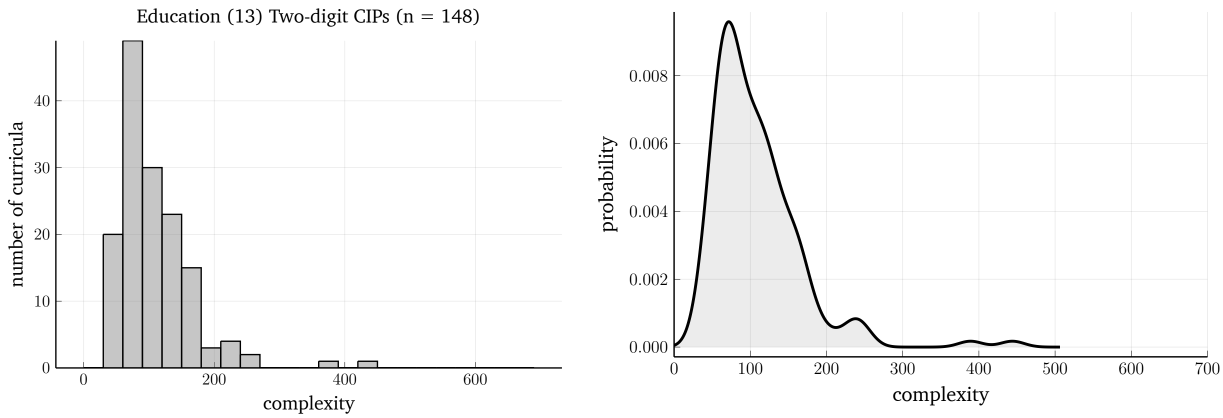


Figure 11: Two-digit CIP series 13, *Education*.

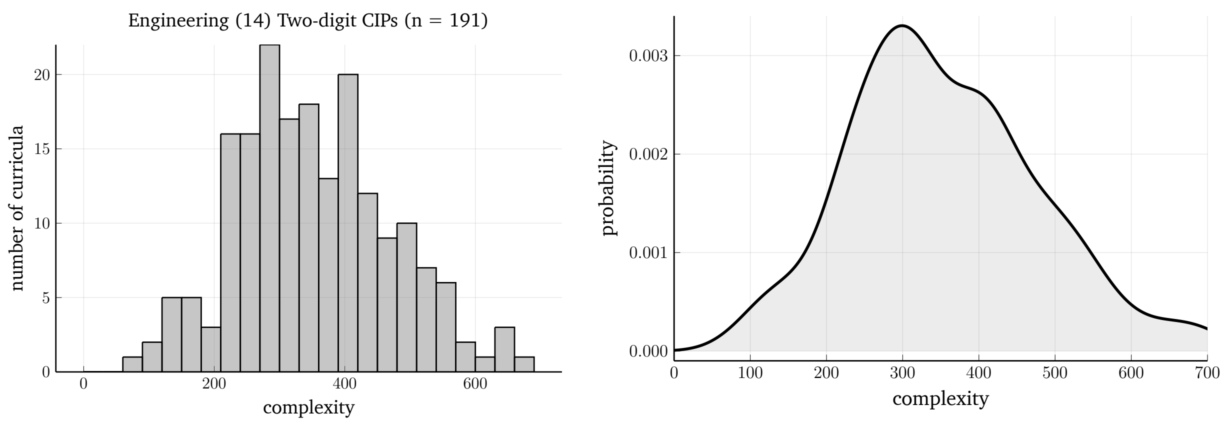


Figure 12: Two-digit CIP series 14, *Engineering*.

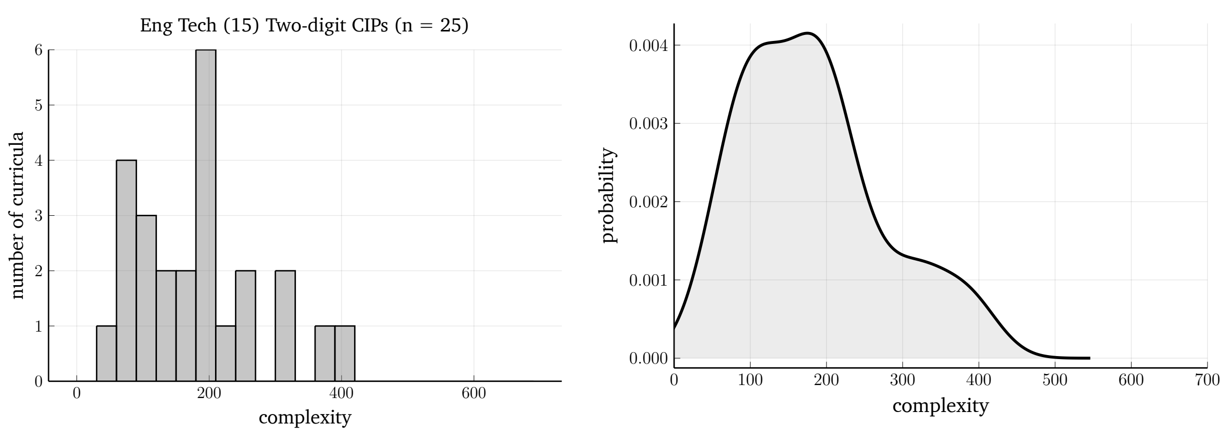


Figure 13: Two-digit CIP series 15, *Engineering/Engineering-related Technologies/Technicians*.

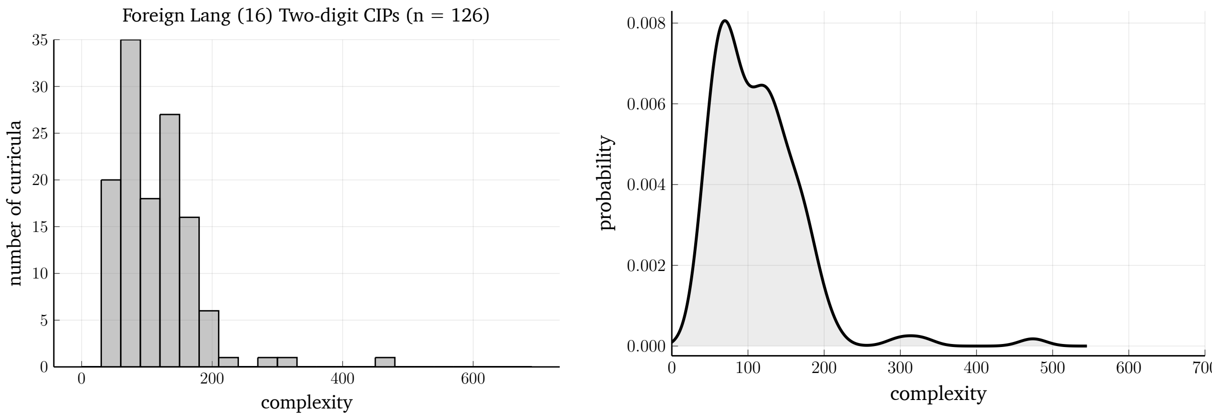


Figure 14: Two-digit CIP series 16, *Foreign Languages, Literatures, and Linguistics*.

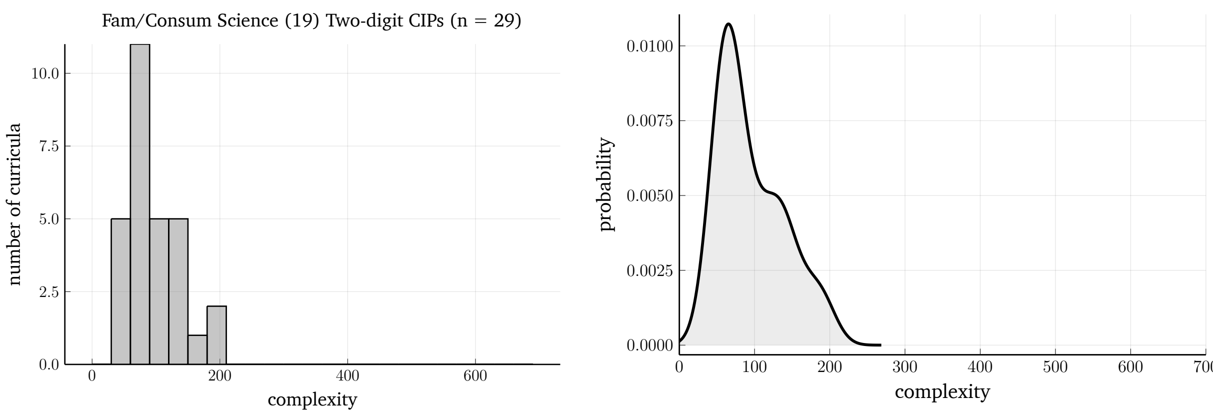


Figure 15: Two-digit CIP series 19, *Family and Consumer Sciences/Human Sciences*.

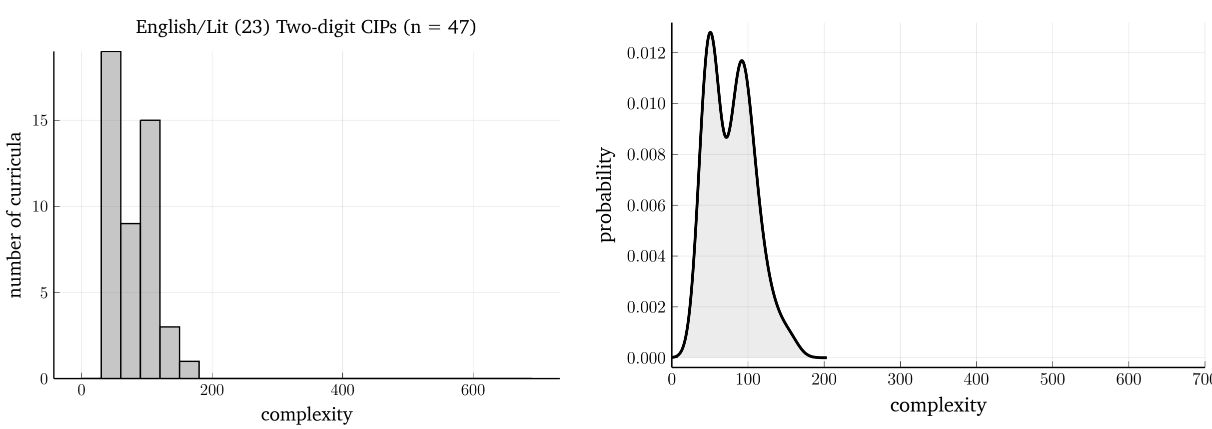


Figure 16: Two-digit CIP series 23, *English Language and Literature/Letters*.

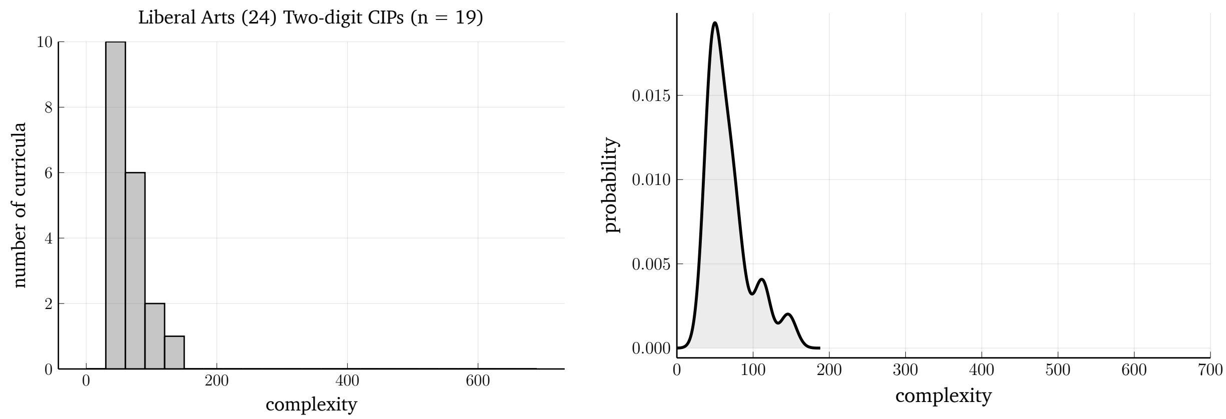


Figure 17: Two-digit CIP series 24, *Liberal Arts and Sciences, General Studies and Humanities*.

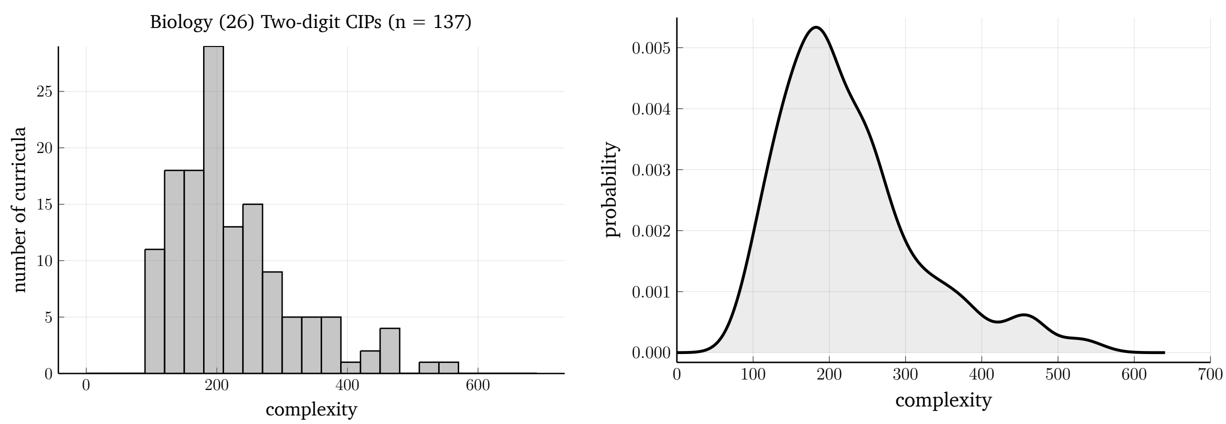


Figure 18: Two-digit CIP series 26, *Biological and Biomedical Sciences*.

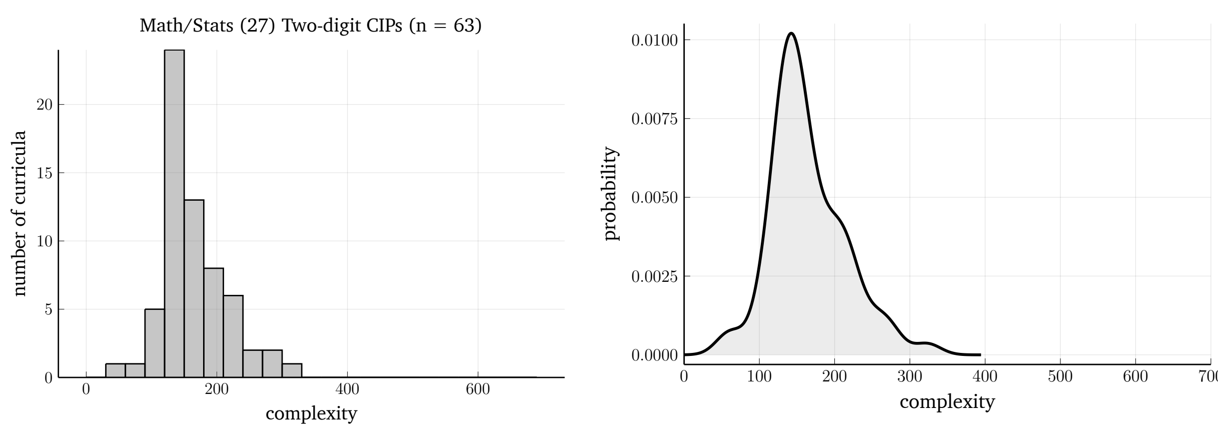


Figure 19: Two-digit CIP series 27, *Mathematics and Statistics*.

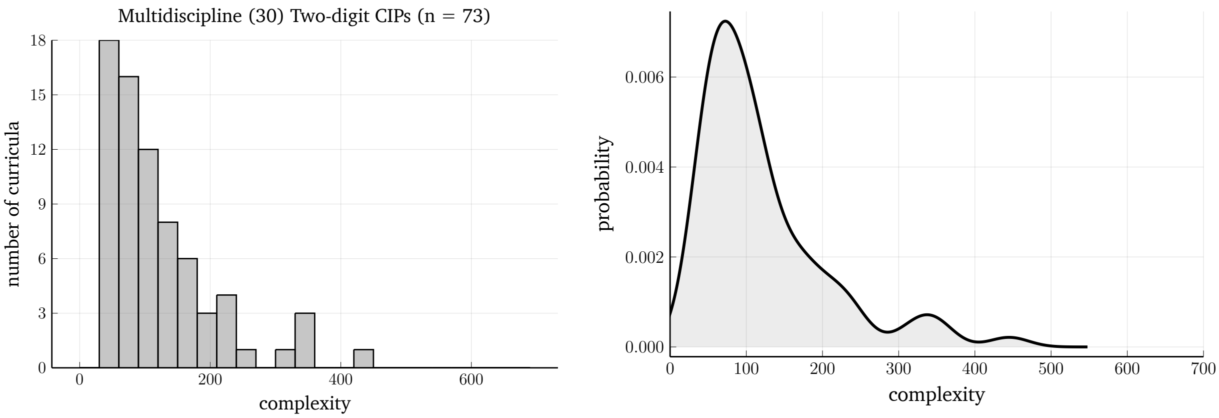


Figure 20: Two-digit CIP series 30, *Multi/Interdisciplinary Studies*.

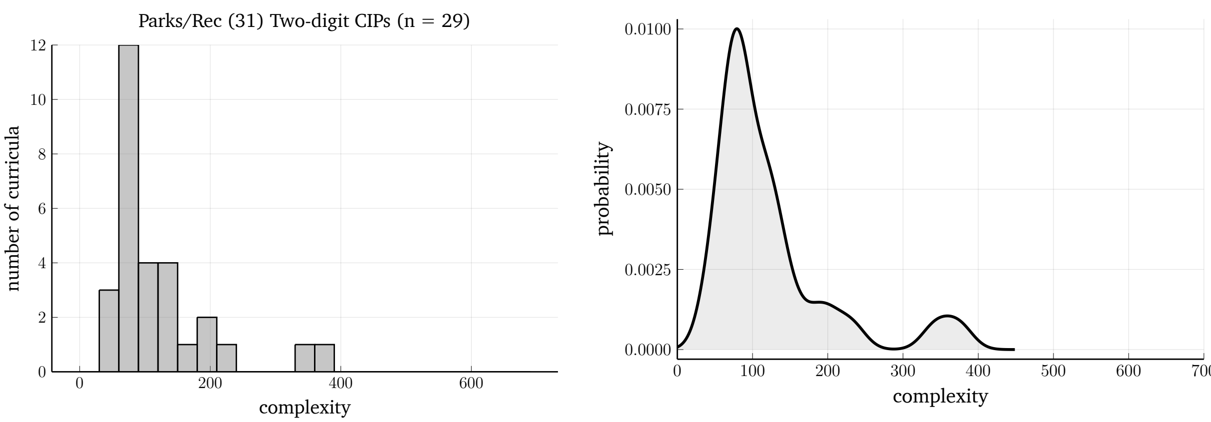


Figure 21: Two-digit CIP series 31, *Parks, Recreation, Leisure, Fitness, and Kinesiology*.

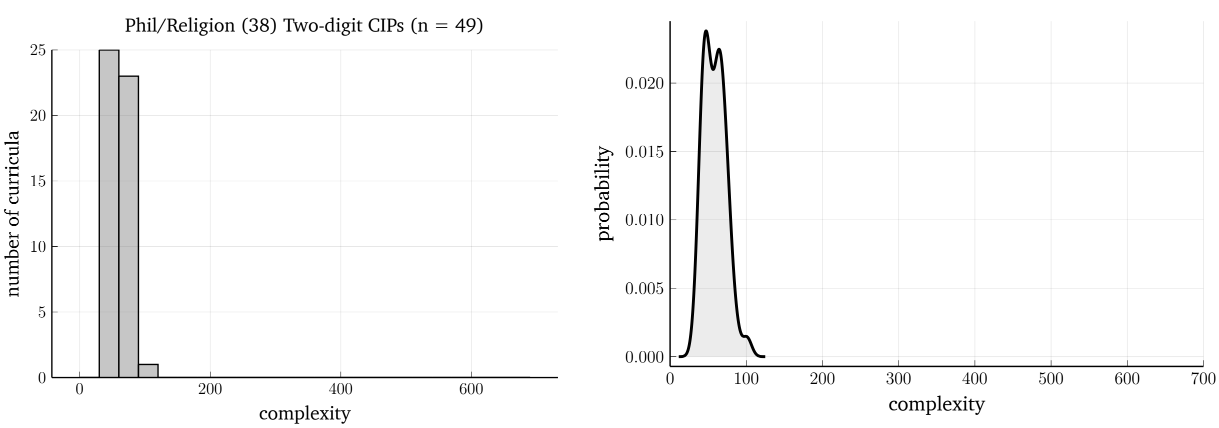


Figure 22: Two-digit CIP series 38, *Philosophy and Religious Studies*.

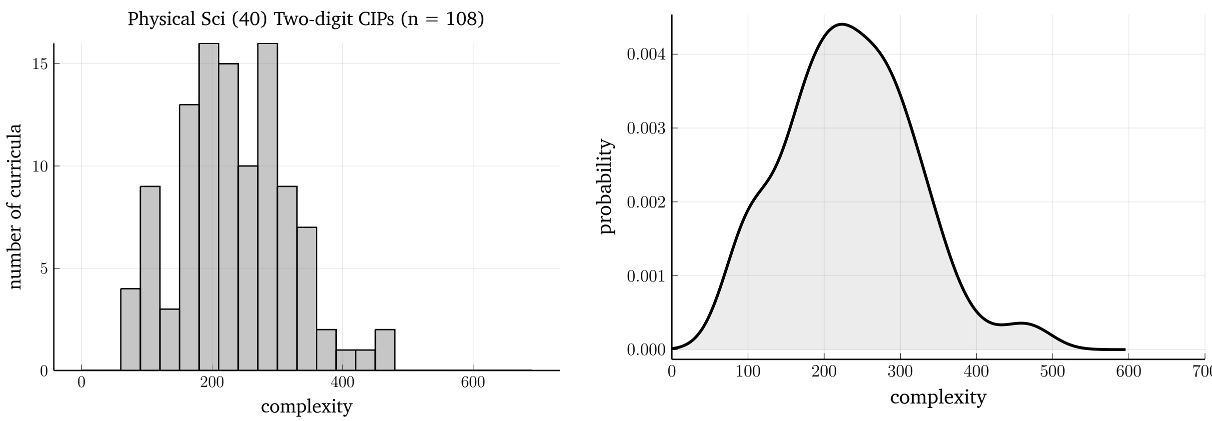


Figure 23: Two-digit CIP series 40, *Physical Sciences*.

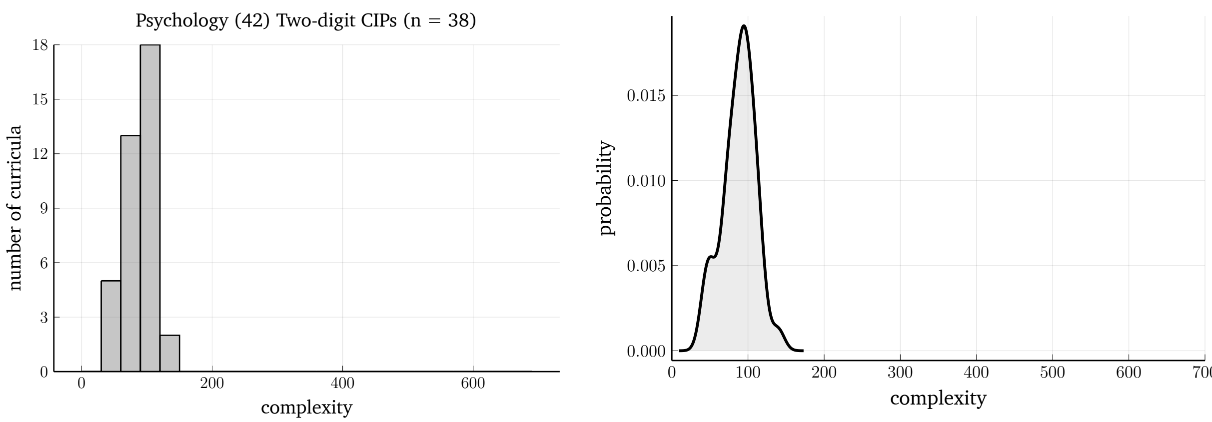


Figure 24: Two-digit CIP series 42, *Psychology*.

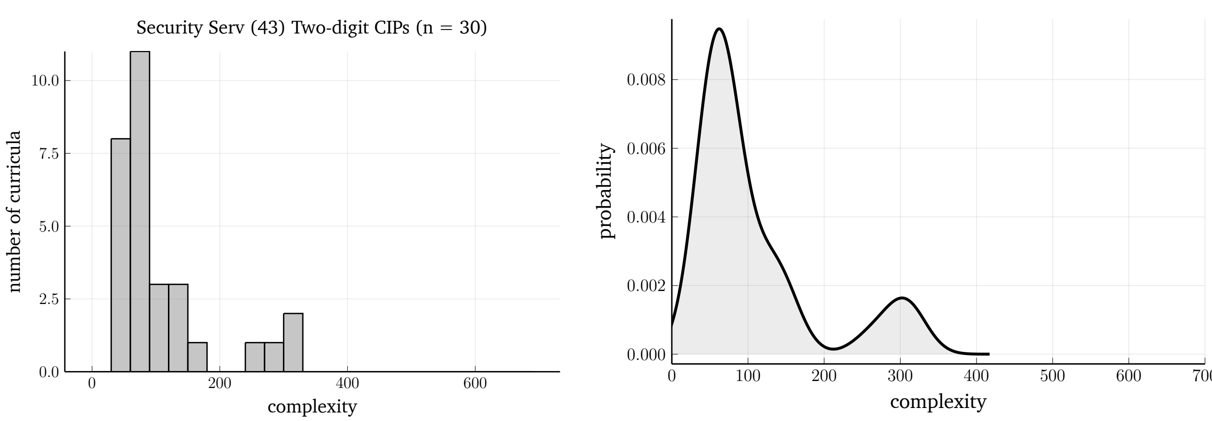


Figure 25: Two-digit CIP series 43, *Homeland Security, Law Enforcement, Firefighting and Related Protective Services*.

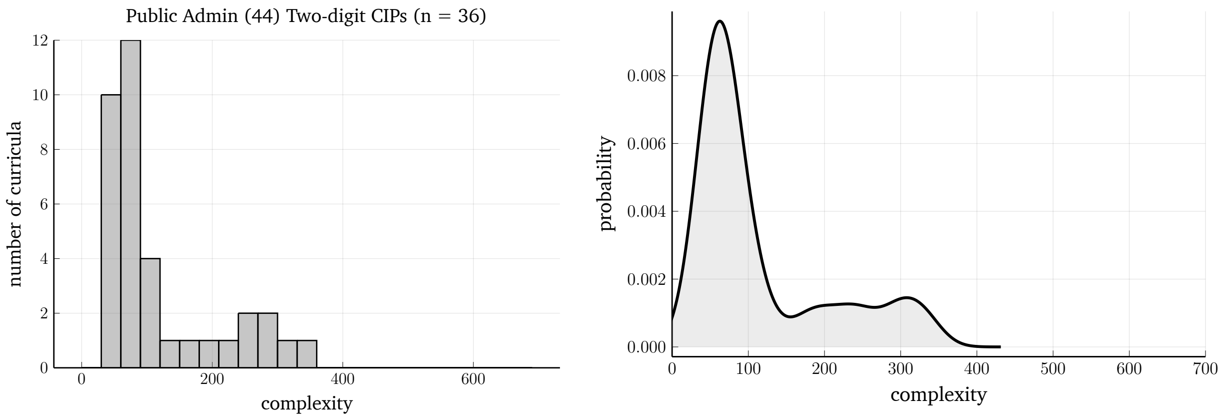


Figure 26: Two-digit CIP series 44, *Public Administration and Social Service Professions*.

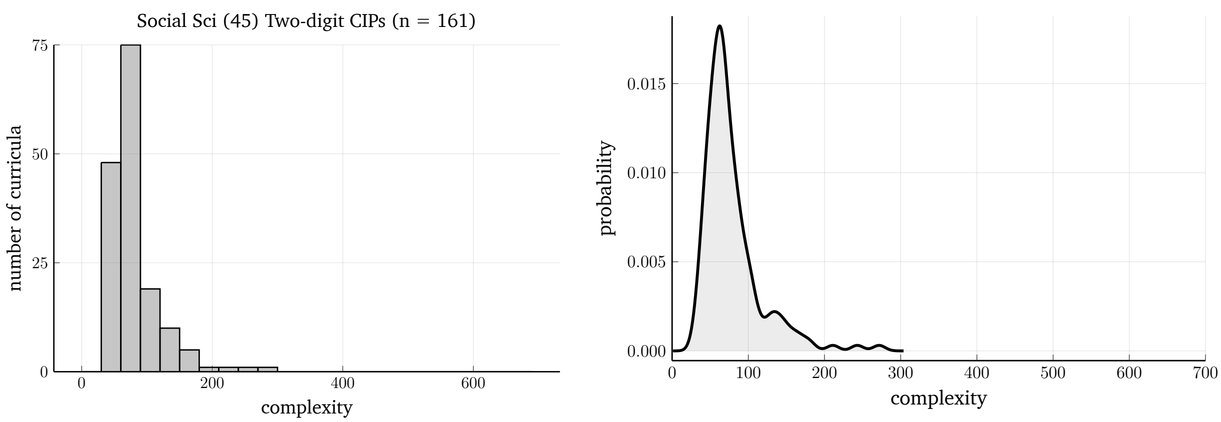


Figure 27: Two-digit CIP series 45, *Social Sciences*.

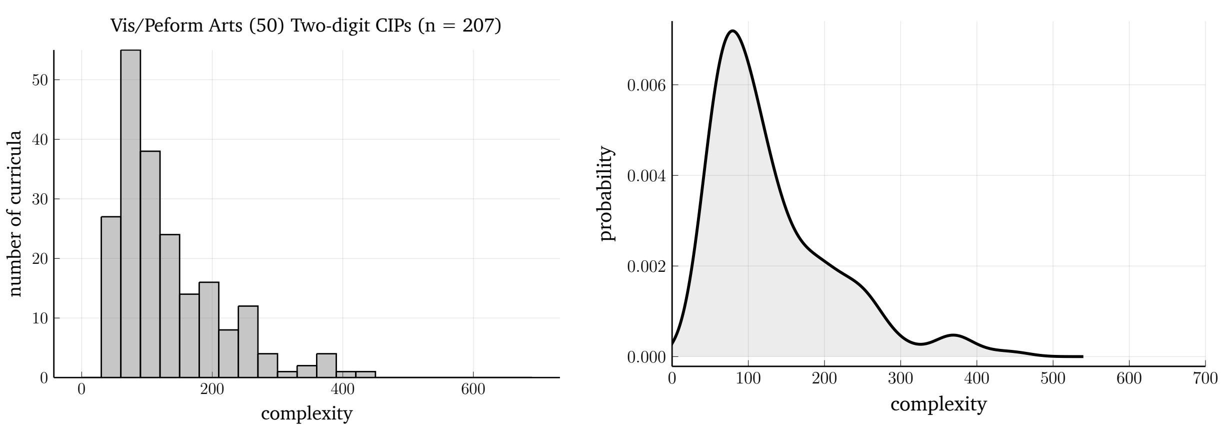


Figure 28: Two-digit CIP series 50, *Visual and Performing Arts*.

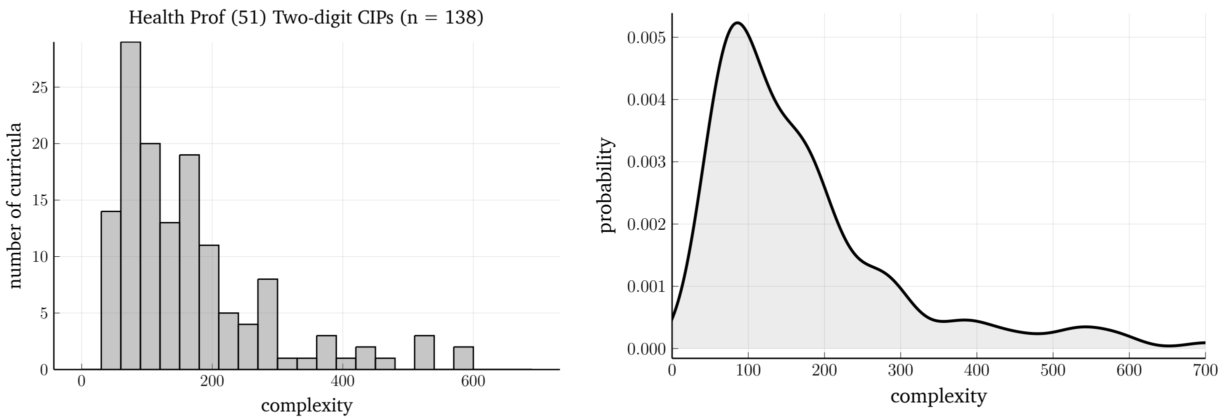


Figure 29: Two-digit CIP series 51, *Health Professions and Related Programs*.

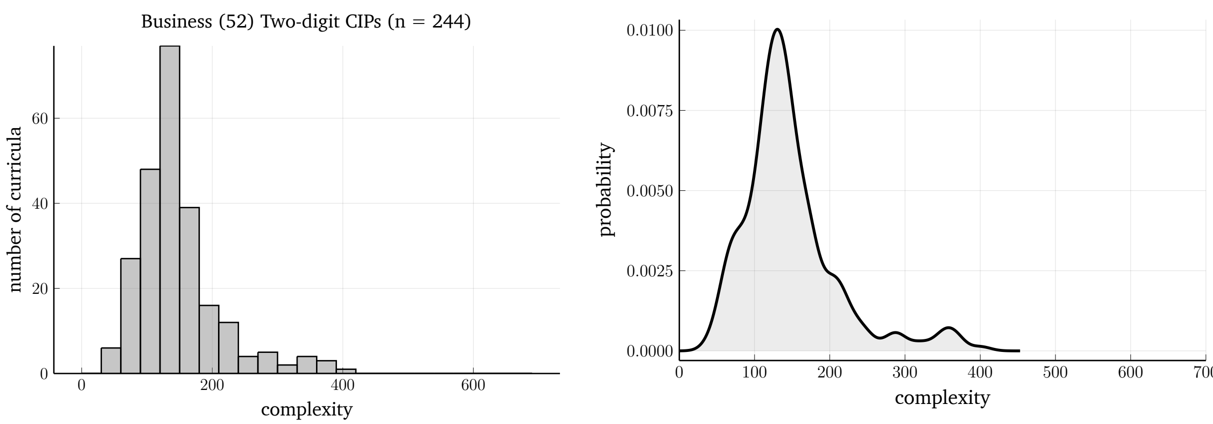


Figure 30: Two-digit CIP series 52, *Business, Management, Marketing, and Related Support Services*.

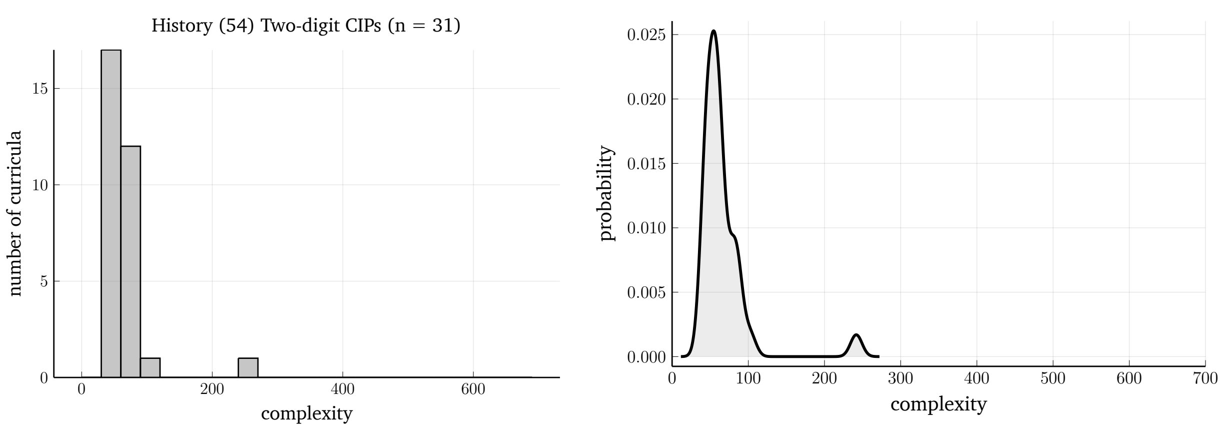


Figure 31: Two-digit CIP series 54, *History*.

In each of the fields considered in the following Figures 5–31, the empirical density functions have a clearly identifiable dominant mode (i.e., a prominent peak). This mode provides an indication of the central tendency of the complexity of the curricula in each field, indicating where the complexity values are most concentrated in a particular field. Many of the distributions have an additional smaller mode. These may be due to the differences in curricular complexity associated with the particular disciplines within a field, a subject considered in more detail in Section 3.3. The important point is, the field-level distributions tend to have shapes that do *not* indicate power law behavior. Indeed, many of the emprical distributions resemble that of a gamma distribution, and more research is being conducted in order to better characterize these distributions. **We believe these empirical distributions support the ability to make qualitatvie comparisons among the complexity differences of programs within particular fields of study.** Thus, in Section 3.4 we utilize these empirical distributions in order to compare the complexity of the curricula at your institution, on a field-by-field bases, to those at all of the other institutions in this study.

To better understand how the curricular complexities found in individual fields contribute to the overall distribution of complexities across all fields, from the figures above, we selected the densities associated with eleven different academic fields, and plotted them all together, as shown in Figure 32. This figure clearly shows the stark complexity differences between the curricula in particular disciplines. Notice that many disciplines have more prominent peaks in their complexity distributions than others, but the distribution for engineering is somewhat flatter, demonstrating a large variance across this field. Figure 32 also makes clearly evident the underlying structures that lead to the long-tailed distribution of data shown in Figure 2, and also explains a possible generative model for the power law distribution when considering all fields. Specifically, the power law distribution often emerges when mixing data sources having a range of variances [13].

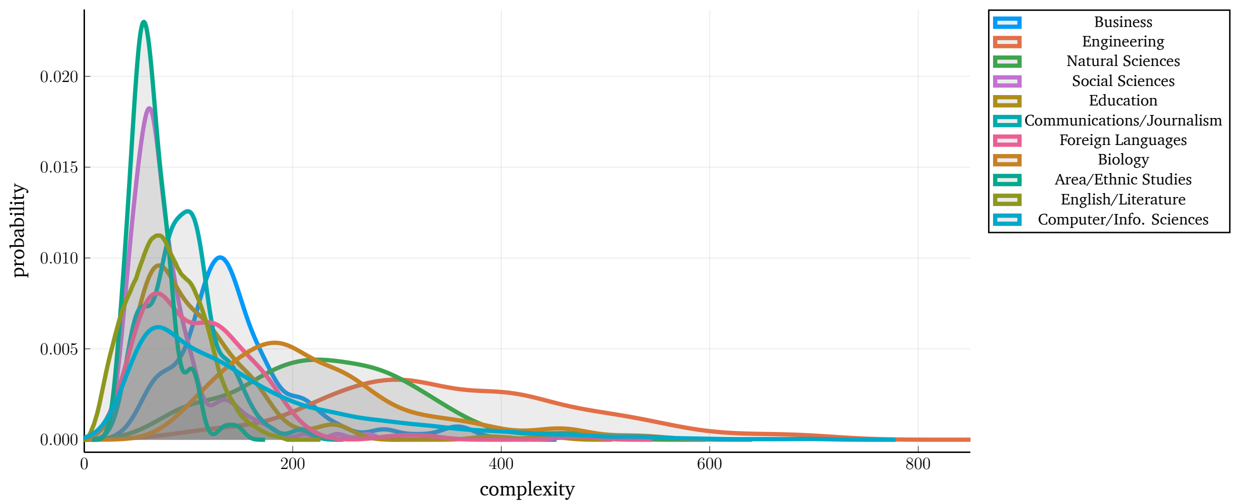


Figure 32: The KDEs for the complexity distributions associated with eleven different fields in the data set.

3.3 Complexity Distribution – By Discipline

For those academic fields where the curricular complexities have large variability, it makes to consider whether or not this variability can be attributed to the particular disciplines within these fields. For instance, the variances in the KDEs for the Engineering field (Figure 12 (b)), the Biological Sciences field (Figure 18 (b)), the Physical Sciences field (Figure 23 (b)), and the Health Professions field (Figure 29 (b)), appear to be larger than in most of the other disciplines.

In order to better understand the large complexity variance in our data set for the engineering field, we next consider the complexities of disciplines within the engineering field, according to their four-digit CIP codes. Figure 33 shows the KDEs associated with the six different engineering disciplines, constructed in the same manner as described above. From this figure, it is easy to see how the combination of engineering discipline complexities produces the widely dispersed complexity distribution for the entire engineering field shown in Figure 12 (b). Furthermore, in this data set the specific engineering disciplines are much more defined. For instance, the Systems/Industrial Engineering programs tend to cluster around a particular complexity value that is quite different from that of the Chemical Engineering programs. Similarly, the Civil Engineering programs appear to be less complex than the Mechanical Engineering programs. Each of these engineering disciplines had 15–25 programs in our data set. Thus, at the moment we are reluctant to say more about these particular distributions of engineering disciplines. Our intention is to extend this research in order to collect additional data for the purpose of performing more detailed discipline-specific analyses.

Figure 34 shows four KDEs for disciplines within the Natural Sciences field. Again, evidence suggests there may be quantifiable differences amoung the curricular complexities of the

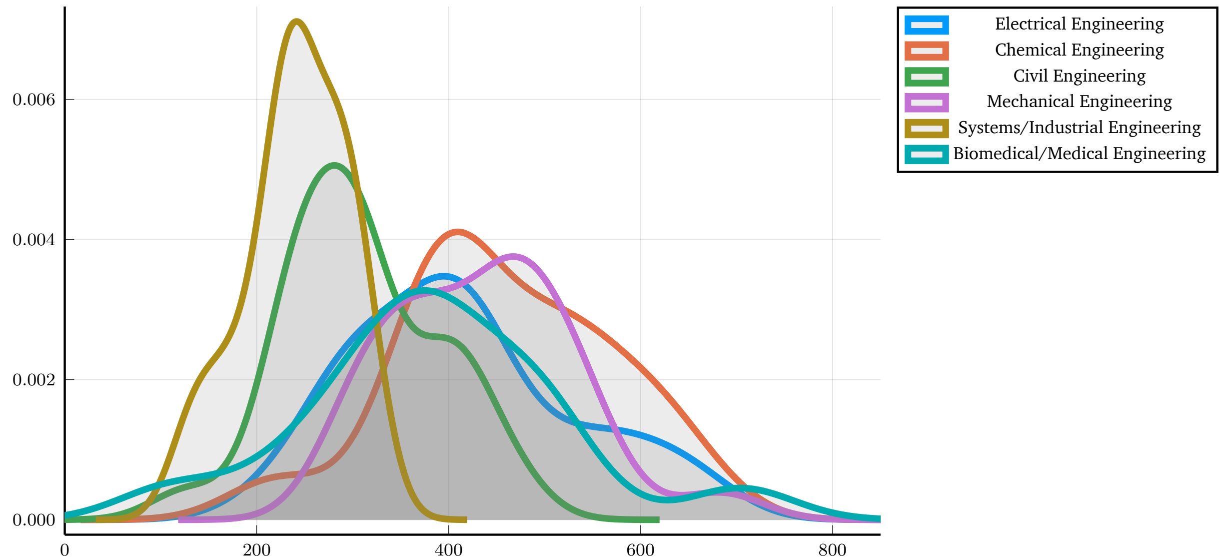


Figure 33: The KDEs for the complexity distributions associated with six different engineering disciplines in the data set.

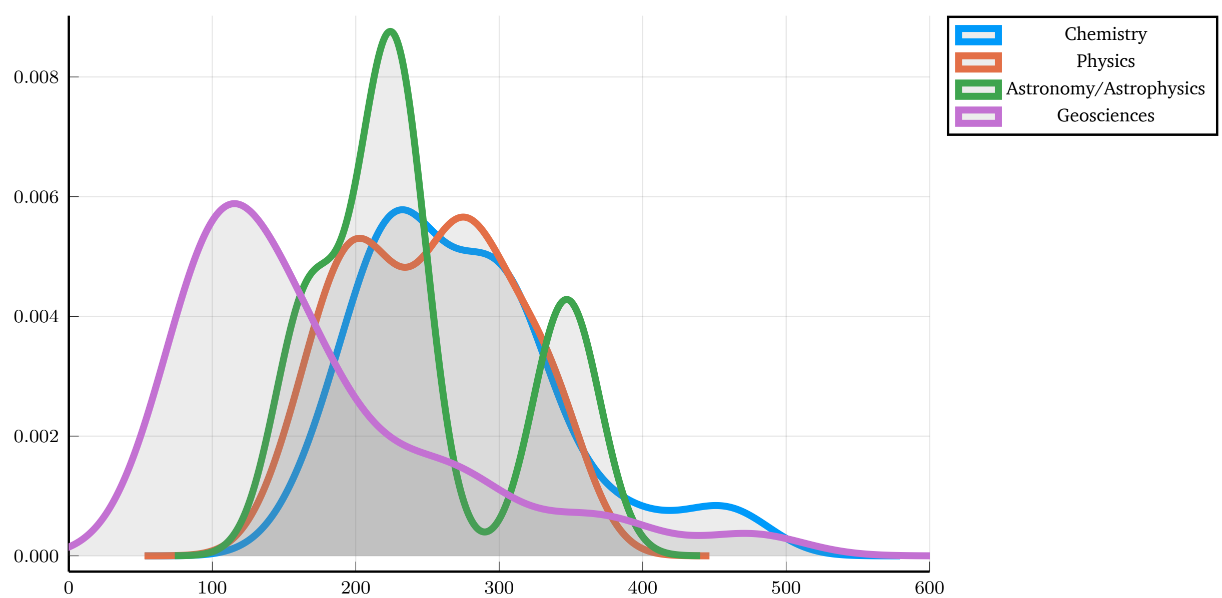


Figure 34: The KDEs for the complexity distributions associated with four different natural sciences disciplines in the data set.

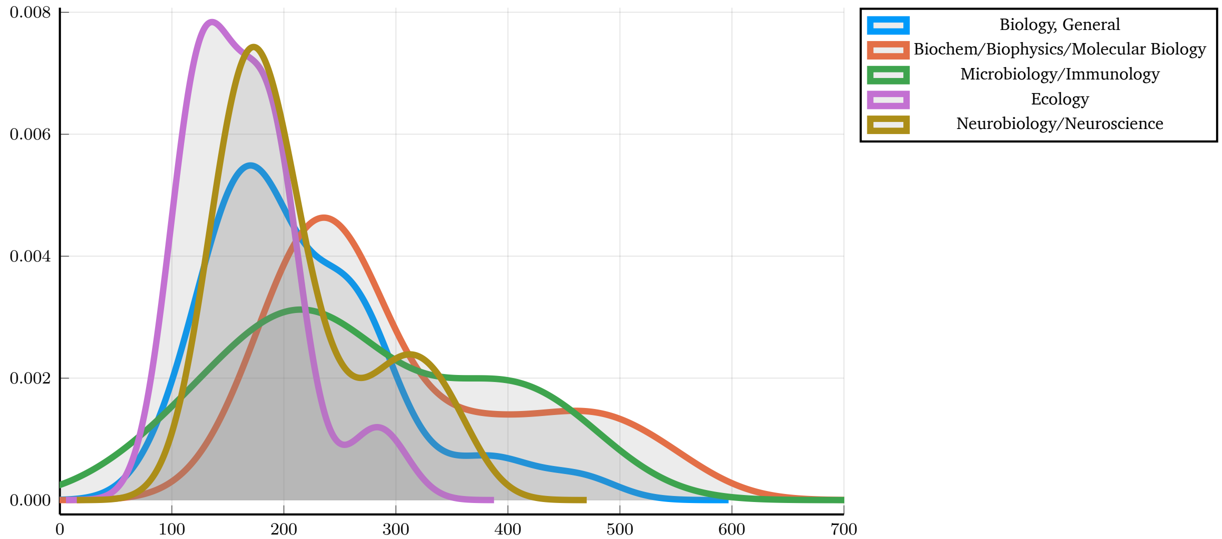


Figure 35: The KDEs for the complexity distributions associated with five different biological and biomedical sciences disciplines in the data set.

disciplines in this field. For instance, it appears the Geosciences discipline is less complex than the Physics discipline. Of course, the aformentioned data limitations apply to these disciplines as well.

Figure 35 shows five KEDs for disciplines within the Biological and Biomedical Sciences field. Preliminary evidence indicates complexity similarities between Ecology and Neurobiology, with large variance in the Microbiology and Immunology disciplines.

Figure 36 shows four KEDs for disciplines within the Health Professions and Related Clinical Sciences field. In this figure, the biomodal nature of the Health and Medical Administration KDE is interesting to note, but we speculate this may be modulated through additional data collection. The wide variance of the Pubic Health discipline also requires additional investigation.

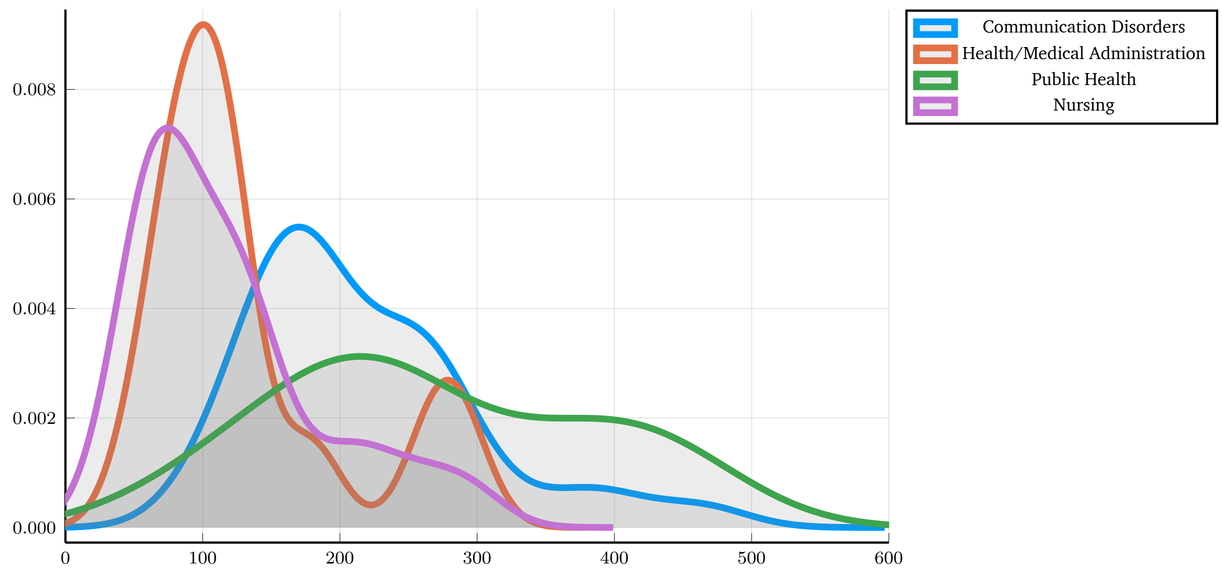


Figure 36: The KDEs for the complexity distributions associated with four different health professions and related clinical sciences disciplines in the data set.

3.4 University of Arizona Curricula

The program complexity histogram for the University of Arizona is provided in Figure 37. A list of the particular academic programs that appear in this histogram, along with their CIP codes, and their individual curricular complexities, is provided in Table 3.

Program Name	**CIP code**	**Complexity**
Aerospace Engineering	140201	431
Electrical and Computer Engineering - Computer Engineering	141099	336
Materials Science Engineering	143101	206
Software Engineering	140903	202
Engineering Management	151501	185
Industrial Engineering	143501	210
Mechanical Engineering	141901	357
Mining Engineering - Sustainable Development Track	142101	223
Optical Sciences and Engineering	140501	274
Performance - Composition	500903	155
Studio of Art	500702	72
Systems Engineering	142701	177
Theatre Arts	500501	78
Architectural Engineering	140401	360
Accounting	520301	58
Africana Studies	50201	65
Agribusiness Economics and Management - Agribusiness Managem...	10101	81
Agricultural Systems Management	10199	66
Agricultural Technology Management and Education - Agricultu...	10106	76
American Indian Studies	50202	67
Animal Sciences - Animal Industry	10901	129
Anthropology	450201	76

Applied Biotechnology - Food and Beverage Fermentation	261201	158
Applied Computing - Applied AI	110104	86
Applied Humanities - Business Administration	240103	68
Applied Physics	141201	215
Applied Science - Administration of Justice	419999	88
Arabic	161101	164
Architecture	40902	293
Art History	500703	79
Art and Visual Culture Education - Community Museums	131302	131
Astronomy	400201	225
Biochemistry	260202	230
Bioinformatics	261103	144
Biology - Biomedical Sciences	260101	173
Biosystems Analytics and Technology	151199	78
Biosystems Engineering	144501	150
Business Administration	520201	66
Business Economics	520601	82
Business Management	521001	68
Care Health and Society	511504	53
Chemical Engineering	140701	484
Chemistry	400501	231
Civil Engineering - (Bachelor of Science in Civil Engineerin...	140801	307
Classics - Classical Languages	161200	81
Communication	90101	95
Computer Science - (Bachelor of Arts)	110701	176
Creative Writing	231302	120
Criminal Justice Studies	430301	55
Cyber Operations - Cyber Engineering	290207	88
Dance	500301	113
Deaf Studies - Educational Interpreting	139999	130
Design Arts and Practice	500102	114
Early Childhood Education	131209	81
East Asian Studies - Chinese Culture	50104	85
Ecology and Evolutionary Biology	261310	136
Economics	450601	79
Elementary Education - Bilingual	131202	123
Emergency Medical Services	510904	72
English	230101	155
Entrepreneurship	520701	60
Environmental Engineering BS	141401	282
Environmental Sciences - Soil	30104	127
Environmental Studies	30103	67
Environmental and Water Resource Economics	450602	77
Family Studies and Human Development	190701	63
Fashion Industry Science and Tech	190901	61
Film and Television	500601	168
Finance	520801	93
Food Safety	11099	164
Food Studies	190599	65
French - French and Francophone Studies	160901	126
Game Design and Development	500411	76

Gender and Women’s Studies - Chicana / Latina Studies	50207	69
General Studies - Sports and Society	300000	52
Geographic Information Systems Technology	450702	62
Geography - Geographic Info Science	450701	61
Geosciences - Earth Ocean and Climate Emphasis	400601	98
German Studies	160501	115
Global Studies	302001	101
Government and Public Service	451001	97
History	540101	60
Human Services	440000	70
Hydrology and Atmospheric Sciences - Atmospheric Sciences - ...	400605	199
Information Science - Data Science	110101	80
Information Science and eSociety	110801	59
Intelligence and Info Ops - Information Warfare	290299	58
Italian - Italian Studies	160902	123
Journalism - Broadcast Journalism	90401	150
Judaic Studies	380206	66
Landscape Architecture	303301	84
Latin American Studies	50107	60
Law	229999	71
Linguistics	160102	75
Learning and Leadership - Community Education	130607	50
Live and Immersive Arts	500706	77
Management Info Systems	521201	58
Marketing	521401	58
Mathematics - Applied	270101	136
Medicine - Basic Medical Sciences	511201	138
Mexican American Studies	50203	68
Microbiology	260502	147
Middle Eastern and North African Studies	50108	67
Mild Moderate Disabilities	131001	129
Molecular and Cellular Biology - Education and Communication	260406	202
Music Education	131312	183
Music	500901	140
Musical Theatre	500509	449
Natural Resources - Conservation Biology	30101	108
Neuroscience and Cognitive Science - Cognition	261501	167
Nursing - Conventional	513801	143
Nutrition and Food Systems	301901	189
Operations Management	520205	63
Organizational Leadership and Regional Commerce - Organizati...	520213	69
Personal and Family Financial Planning	190401	94
Pharmaceutical Sciences	512010	209
Philosophy Politics Economics and Law	380199	64
Physics	400801	278
Physiology and Medical Sciences - Exercise and Extreme Physi...	260901	178
Plant Sciences	11101	136
Precision Nutrition and Wellness	190504	168
Prof and Technical Writing	231303	89
Psychological Sciences	422704	117
Psychology	420101	94

Public Health - Environ and Occupational Health	512201	114
Public Management and Policy - Public Admin and Management	440401	71
Rehabilitation Studies Services	511599	52
Religious Studies	380201	60
Retailing and Consumer Sciences	520212	100
Russian - Culture Focused	160402	77
Sociology	451101	67
Spanish - General	160905	150
Speech - Language and Hearing Science	510201	111
Statistics and Data Science	270503	159
Studies of Global Media	90102	59
Sustainable Plant Systems - Agronomy	19999	111
Urban and Regional Development	40301	58
Veterinary Science - Applied Animal Behavior	511104	178
Public Health - Wellness and Health Promo Prac - Aging and P..	510001	76
World Literature	160104	90

Table 3: All undergraduate curricula at the University of Arizona, along with their CIP codes, and computed curricular complexities.

In this table, the average curricular complexity for all programs sharing the same six-digit CIP code are displayed; however, the program name accompanying that average is simply the name of the first curriculum encountered among all those sharing a six-digit CIP code. Again, it is assumed the programs sharing a six-digit CIP code at an instition are concentrations belonging to the same major. There are also instances in the data where to progams with different degree types share a six-digit CIP code. For instance, a BA in Psychology and a BS in Psychology may share the same six-digit CIP code at an institution.

The plots provided next in this section of the report all have the form shown in Figure 38. This plot involves overlaying three distinct visual elements that should help you better appreciate the distribution of curricular complexities associated with a given field of study, and how the programs at your institution fit within that distribution. The three components of this plot (referred to as a Box Scatter Plot) include:

1. **Box-and-Whisker Diagram** – composed of the the dark lines in this plot.
2. **Scatter plot** – consisting of all round data points shown in the plot (both red and gray).
3. **Empirical distribution** – computed using all data points, and reflected about the vertical axis of the box-and-whisker plot.

These plots should be interpreted as follows. First the box-and-whisker diagram (also known as a boxplot) summarizes the data within a given field of study (two-digit CIP category), in this case Engineering. The box itself spans the lower (Q1) and upper (Q3) quartiles of the data set, and the distance between these two is known as the *interquartile range*

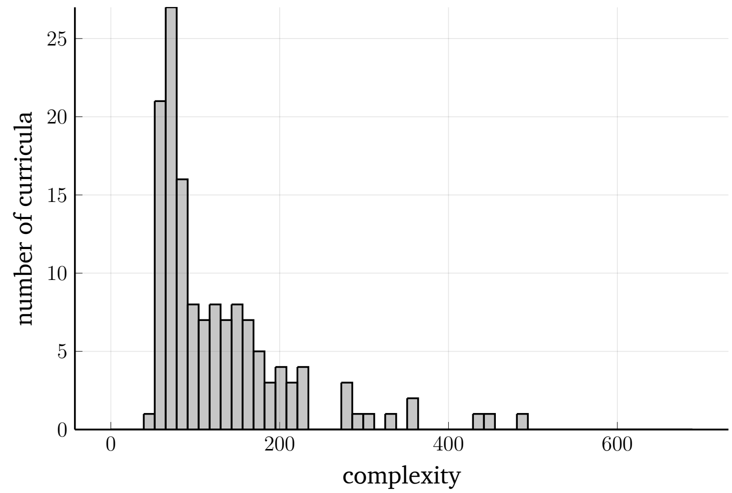


Figure 37: A histogram of the program complexities for all undergraduate programs at the the University of Arizona.

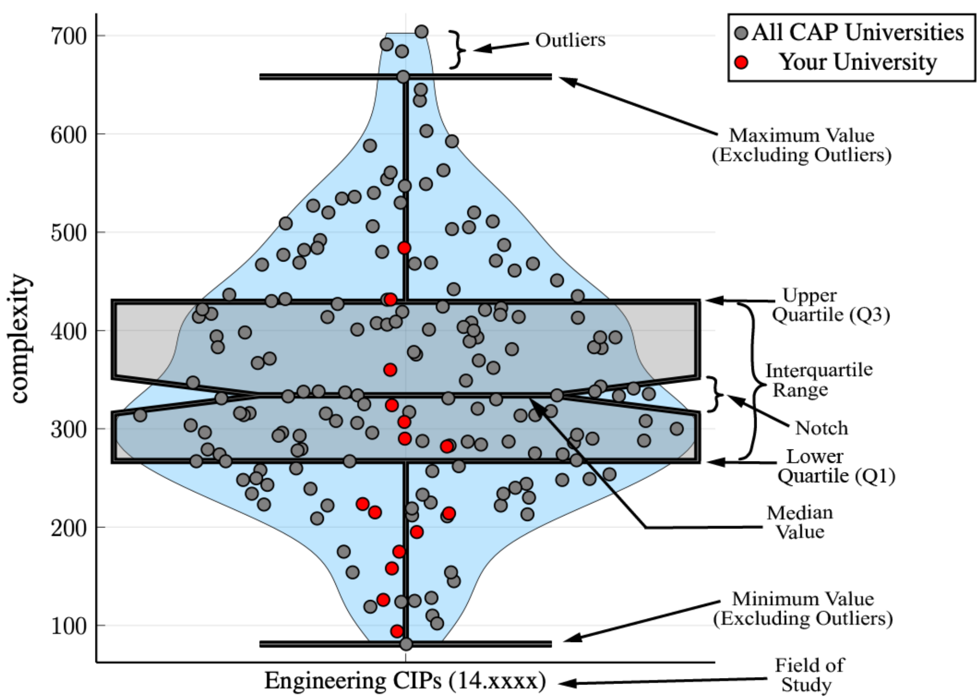


Figure 38: Example plot detailing the distribution of curricular complexity for all programs in a given field of study at a particular university.

(IQR). The line in the middle of the box denotes the median value of the data set, which is roughly 340 in this example. Thus, 50% of the data lies below this line, with complexity scores less than 340, and the other 50% lies above the line, with complexity scores greater than 340. Furthermore, the data points below the box constitute 25% of the data set corresponding to the lowest complexity scores, and the data points above the box constitute the 25% of the data corresponding to the highest complexity scores. Thus, the box itself contains the middle 50% of the complexity scores in the data set. The whiskers extending from the box show range of the data, from minimum to maximum complexity scores, where the data points above the maximum value are considered outliers. More specifically, the whiskers extend to the farthest data point that is within 1.5 times the IQR, with data points outside this range considered outliers. Finally the notch in the box indicates the most likely values of the median value. The size of the notch is directly proportional to the IQR, and inversely proportional to the square root of the number of samples in the data set. The notch itself provides an approximate 95% confidence interval for the median of the entire population of programs with this two-digit CIP. That is, it provides a rough estimate of the confidence we should have when using the sample median as an estimate of the population median. Thus, the notches in these plots are useful for comparing the samples drawn from different fields of study. If the notches from two different fields of study do *not* overlap on the complexity axis, it is an indication the median complexities values for these fields of study are different. In a few fields, the width of the notch exceeds the Q1 value and/or the Q3 value. This can occur if there are a limited number of data values for that particular field or if the data values themselves vary widely. In the CAP data set, this behavior seems to be largely do to the former. In any event, the plots with this feature remain valid, even if they are somewhat unsightly.

Next, the blue outlined shape on this plot is an empirical probability distribution function obtained by applying kernel density estimation techniques to the sample data. If this shape is cut in half along the central axis of the boxplot, and then the left half is rotated (clockwise) by 90 degrees, you will obtain the shape of the empirical probability distribution function for the data set. This portion of the plot is useful for determining the number of modes that might exist in the distribution of the actual population. For instance, in the plot shown in Figure 38 there appears to be two modes, one with a peak at approximately 300 complexity points, and another at approximately 400 complexity points.

Finally, the data points themselves are plotted, using a function that randomly scatters them about the central axis of the box plot according to the empirical probability distribution function. If these points were not scattered in this fashion, they would plot one on top of the other along the central axis. In other words the placement of these data points along the horizontal dimension has no meaning, other than to make them visible. Specifically, by scattering them, it is easy to compare the red data points, corresponding to the programs at your institution, to those belonging to the other institutions in this study.

The following figures contain the field-specific placements of the programs at the University of Arizona, relative to the complexity of the academic programs at the other institutions in this study.

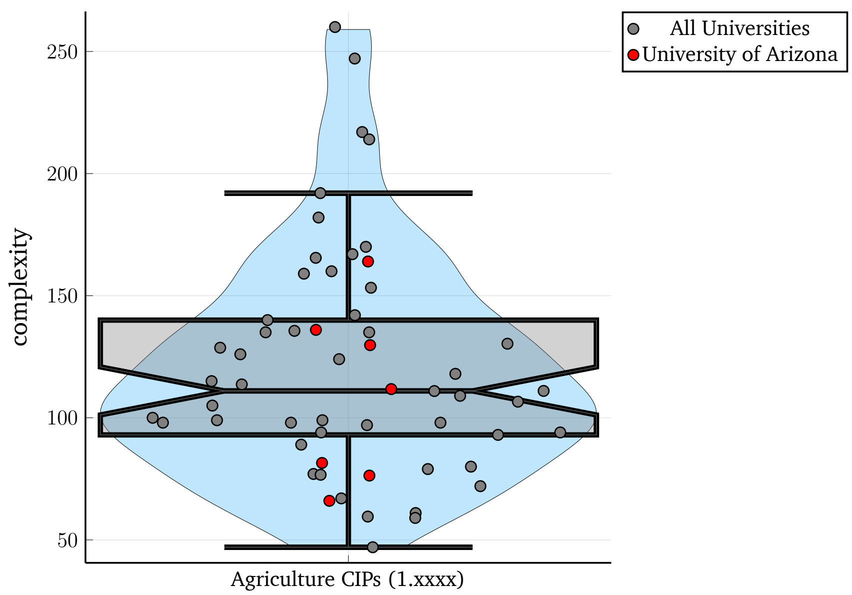


Figure 39: Box scatter plot for two-digit CIP series 1, *Agricultural/Animal/Plant/Veterinary Science and Related Fields*. The red dots correspond to programs at the University of Arizona, including Agribusiness Economics and Management - Agribusiness Managem... (10101), complexity = 81; Agricultural Systems Management (10199), complexity = 66; Agricultural Technology Management and Education - Agricultu... (10106), complexity = 76; Animal Sciences - Animal Industry (10901), complexity = 129; Food Safety (11099), complexity = 164; Plant Sciences (11101), complexity = 136; Sustainable Plant Systems - Agronomy (19999), complexity = 111.

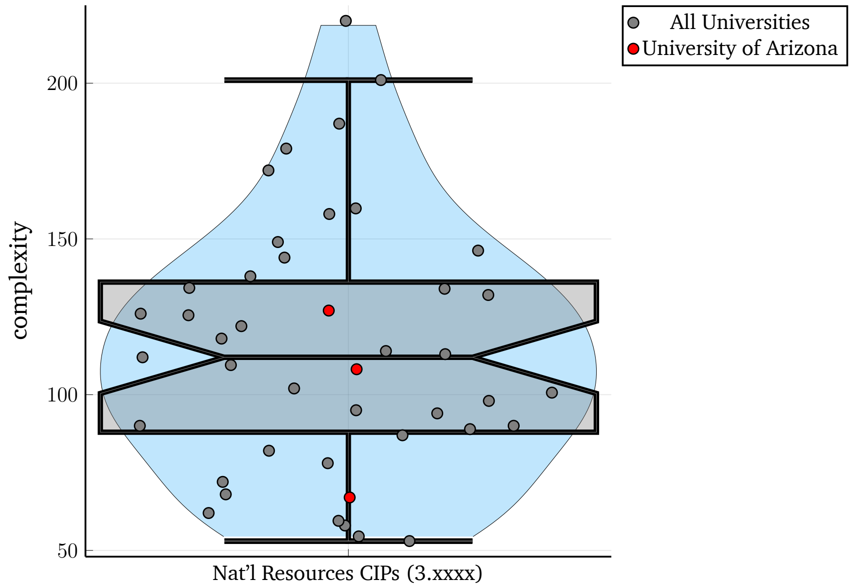


Figure 40: Box scatter plot for two-digit CIP series 3, *Natural Resources and Conservation*. The red dots correspond to programs at the University of Arizona, including Environmental Sciences - Soil (30104), complexity = 127; Environmental Studies (30103), complexity = 67; Natural Resources - Conservation Biology (30101), complexity = 108.

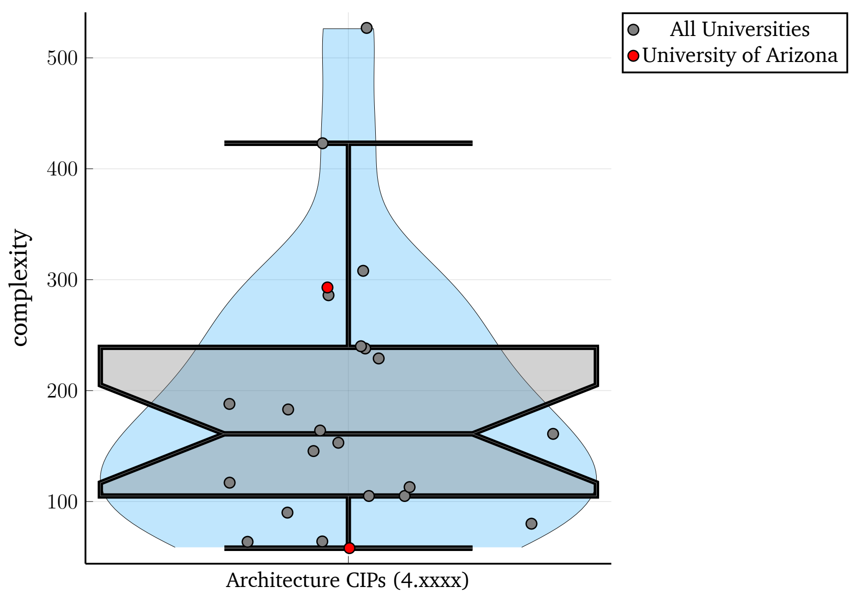


Figure 41: Box scatter plot for two-digit CIP series 4, *Architecture and Related Services*. The red dots correspond to programs at the University of Arizona, including Architecture (40902), complexity = 293; Urban and Regional Development (40301), complexity = 58.

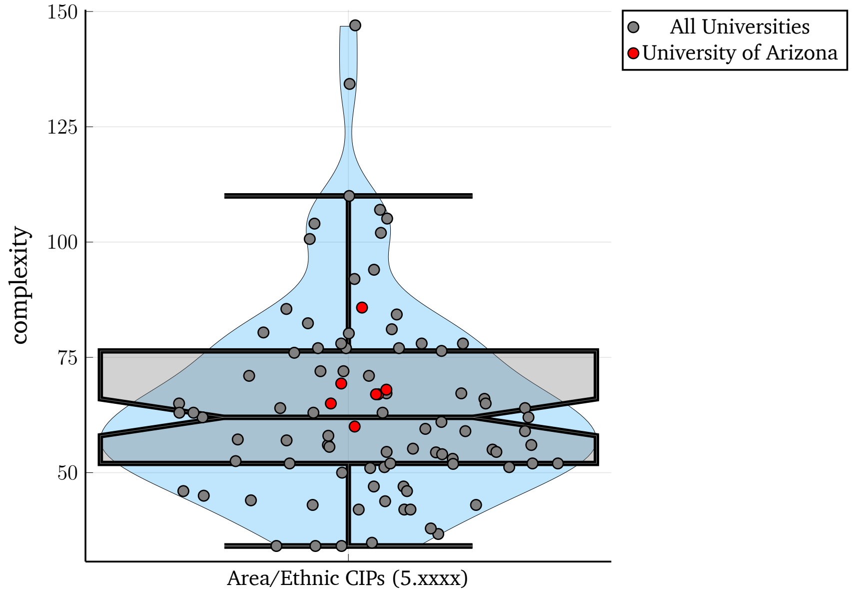


Figure 42: Box scatter plot for two-digit CIP series 5, *Area, Ethnic, Cultural, Gender, and Group Studies*. The red dots correspond to programs at the University of Arizona, including Africana Studies (50201), complexity = 65; American Indian Studies (50202), complexity = 67; East Asian Studies - Chinese Culture (50104), complexity = 85; Gender and Women's Studies - Chicana / Latina Studies (50207), complexity = 69; Latin American Studies (50107), complexity = 60; Mexican American Studies (50203), complexity = 68; Middle Eastern and North African Studies (50108), complexity = 67.

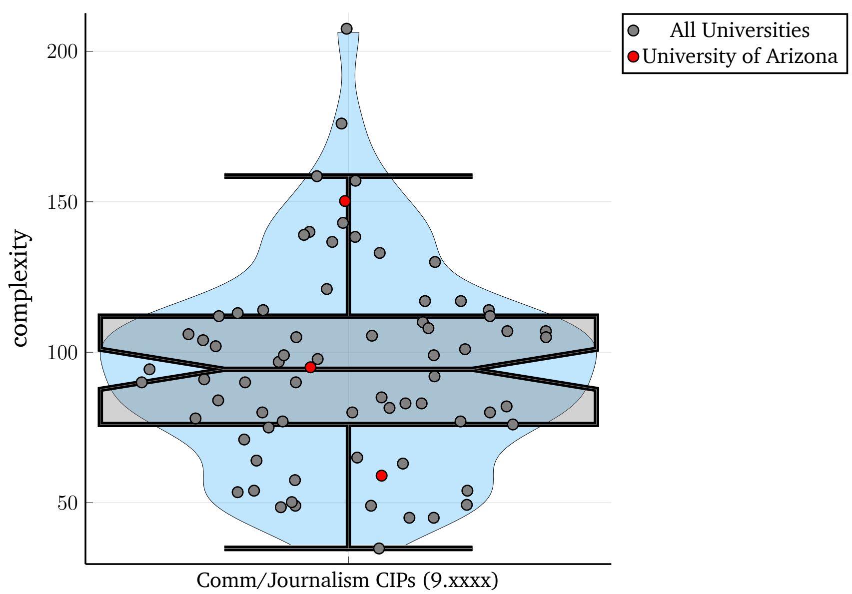


Figure 43: Box scatter plot for two-digit CIP series 9, *Communication, Journalism, and Related Programs*. The red dots correspond to programs at the University of Arizona, including Communication (90101), complexity = 95; Journalism - Broadcast Journalism (90401), complexity = 150; Studies of Global Media (90102), complexity = 59.

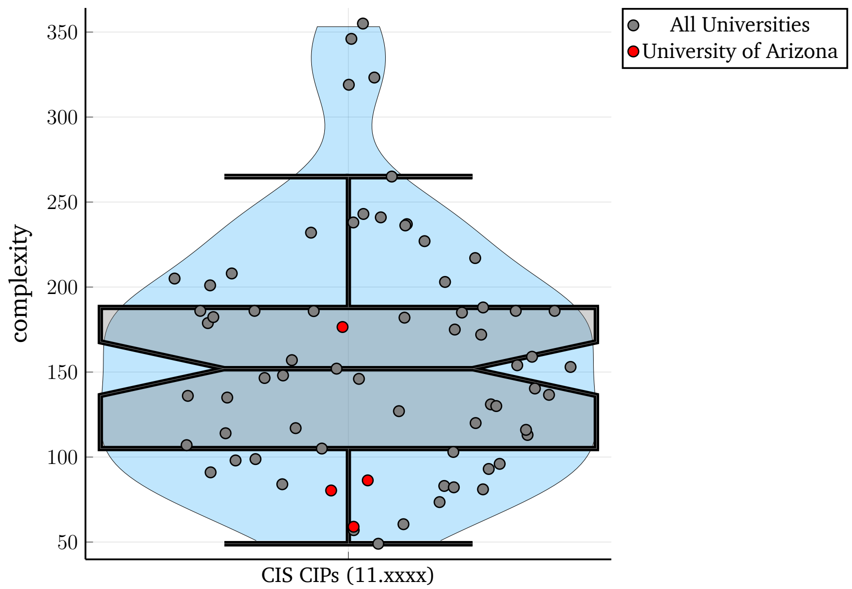


Figure 44: Box scatter plot for two-digit CIP series 11, *Computer and Information Sciences and Support Services*. The red dots correspond to programs at the University of Arizona, including Applied Computing - Applied AI (110104), complexity = 86; Computer Science - (Bachelor of Arts) (110701), complexity = 176; Information Science - Data Science (110101), complexity = 80; Information Science and eSociety (110801), complexity = 59.

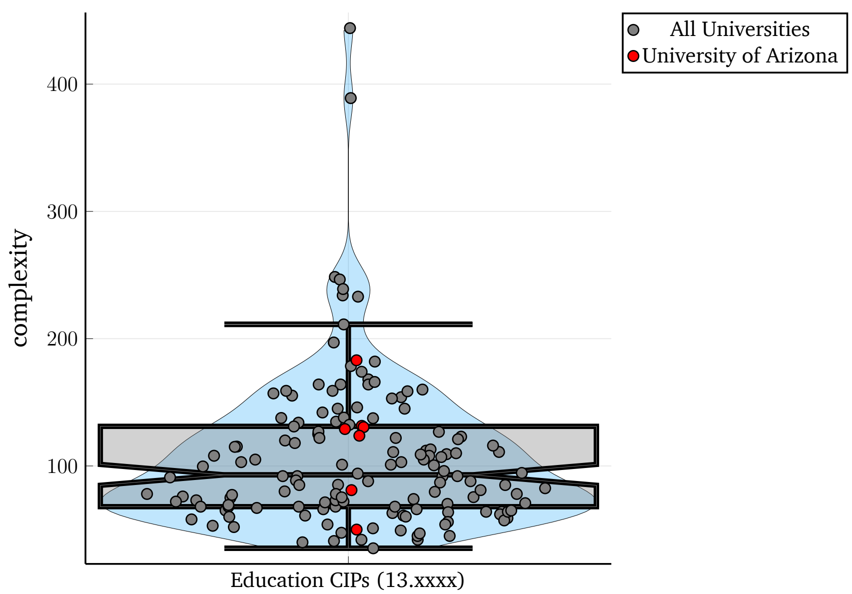


Figure 45: Box scatter plot for two-digit CIP series 13, *Education*. The red dots correspond to programs at the University of Arizona, including Art and Visual Culture Education - Community Museums (131302), complexity = 131; Deaf Studies - Educational Interpreting (139999), complexity = 130; Early Childhood Education (131209), complexity = 81; Elementary Education - Bilingual (131202), complexity = 123; Learning and Leadership - Community Education (130607), complexity = 50; Mild Moderate Disabilities (131001), complexity = 129; Music Education (131312), complexity = 183.

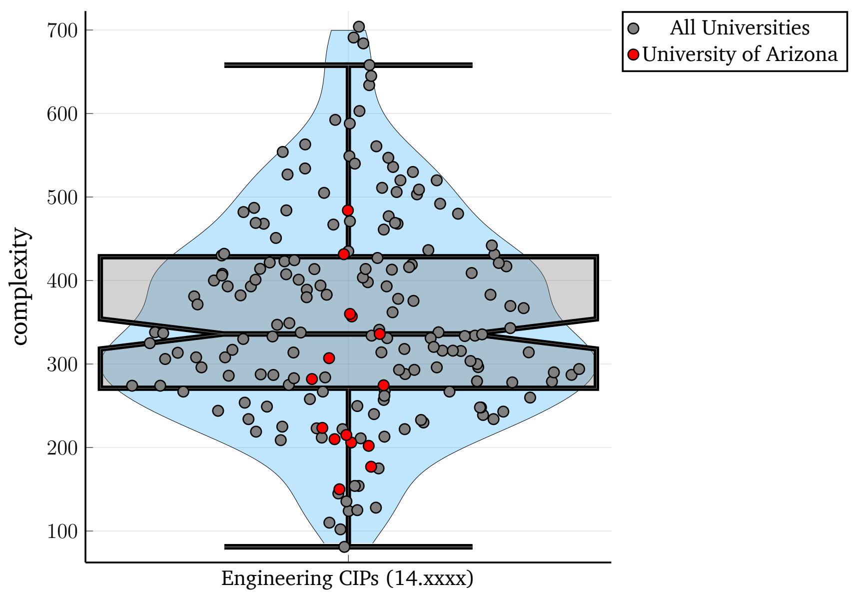


Figure 46: Box scatter plot for two-digit CIP series 14, *Engineering*. The red dots correspond to programs at the University of Arizona, including Aerospace Engineering (140201), complexity = 431; Electrical and Computer Engineering - Computer Engineering (141099), complexity = 336; Materials Science Engineering (143101), complexity = 206; Software Engineering (140903), complexity = 202; Industrial Engineering (143501), complexity = 210; Mechanical Engineering (141901), complexity = 357; Mining Engineering - Sustainable Development Track (142101), complexity = 223; Optical Sciences and Engineering (140501), complexity = 274; Systems Engineering (142701), complexity = 177; Architectural Engineering (140401), complexity = 360; Applied Physics (141201), complexity = 215; Biosystems Engineering (144501), complexity = 150; Chemical Engineering (140701), complexity = 484; Civil Engineering - (Bachelor of Science in Civil Engineerin... (140801), complexity = 307; Environmental Engineering BS (141401), complexity = 282.

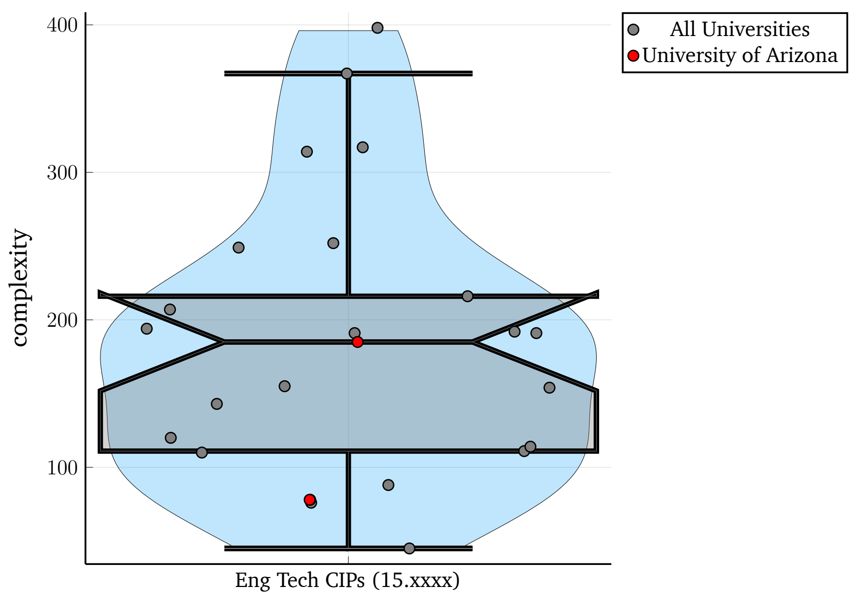


Figure 47: Box scatter plot for two-digit CIP series 15, *Engineering/Engineering-related Technologies/Technicians*. The red dots correspond to programs at the University of Arizona, including Engineering Management (151501), complexity = 185; Biosystems Analytics and Technology (151199), complexity = 78.

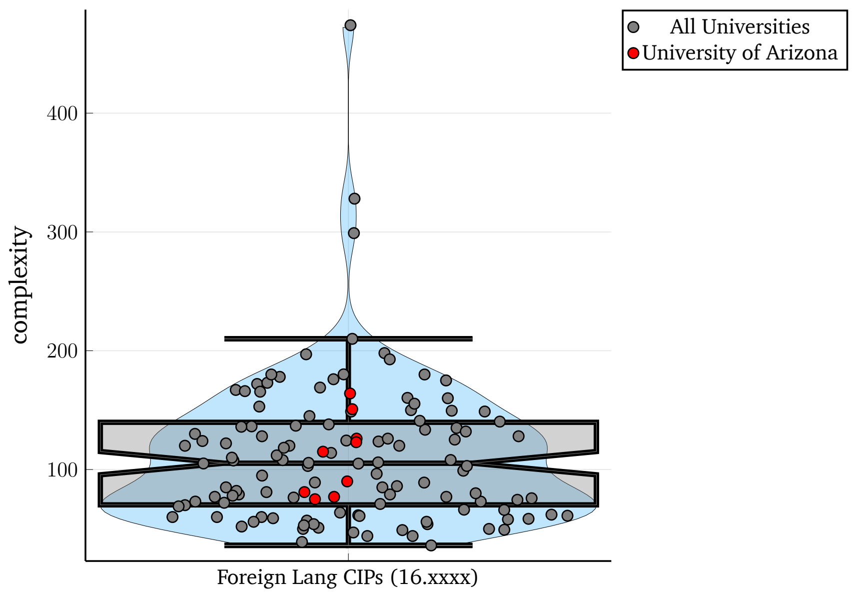


Figure 48: Box scatter plot for two-digit CIP series 16, *Foreign Languages, Literatures, and Linguistics*. The red dots correspond to programs at the University of Arizona, including Arabic (161101), complexity = 164; Classics - Classical Languages (161200), complexity = 81; French - French and Francophone Studies (160901), complexity = 126; German Studies (160501), complexity = 115; Italian - Italian Studies (160902), complexity = 123; Linguistics (160102), complexity = 75; Russian - Culture Focused (160402), complexity = 77; Spanish - General (160905), complexity = 150; World Literature (160104), complexity = 90.

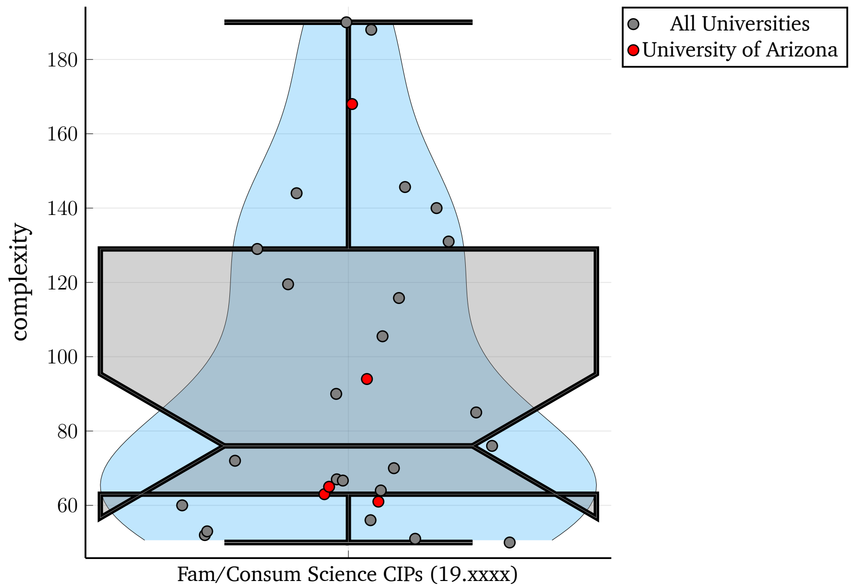


Figure 49: Box scatter plot for two-digit CIP series 19, *Family and Consumer Sciences/Human Sciences*. The red dots correspond to programs at the University of Arizona, including Family Studies and Human Development (190701), complexity = 63; Fashion Industry Science and Tech (190901), complexity = 61; Food Studies (190599), complexity = 65; Personal and Family Financial Planning (190401), complexity = 94; Precision Nutrition and Wellness (190504), complexity = 168.

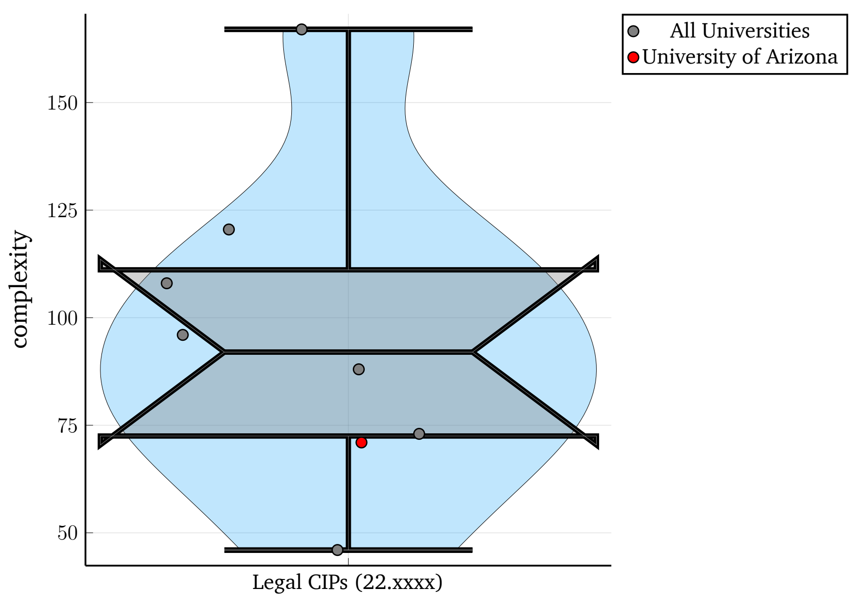


Figure 50: Box scatter plot for two-digit CIP series 22, *Legal Professions and Studies*. The red dots correspond to programs at the University of Arizona, including Law (229999), complexity = 71.

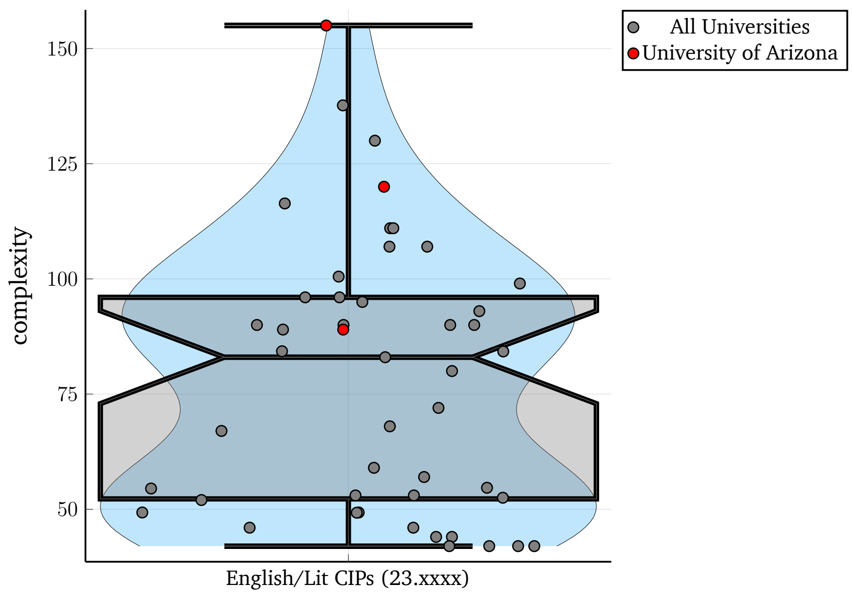


Figure 51: Box scatter plot for two-digit CIP series 23, *English Language and Literature/Letters*. The red dots correspond to programs at the University of Arizona, including Creative Writing (231302), complexity = 120; English (230101), complexity = 155; Prof and Technical Writing (231303), complexity = 89.

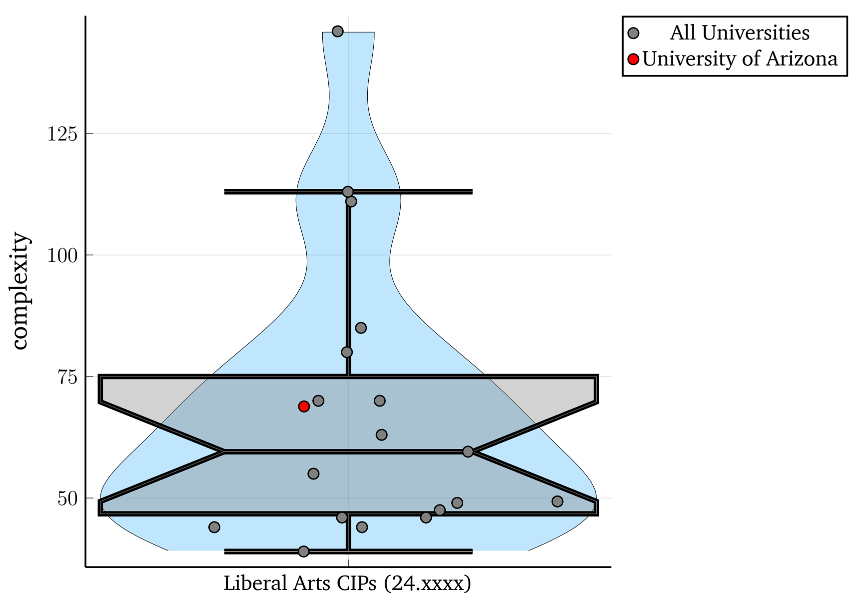


Figure 52: Box scatter plot for two-digit CIP series 24, *Liberal Arts and Sciences, General Studies and Humanities*. The red dots correspond to programs at the University of Arizona, including Applied Humanities - Business Administration (240103), complexity = 68.

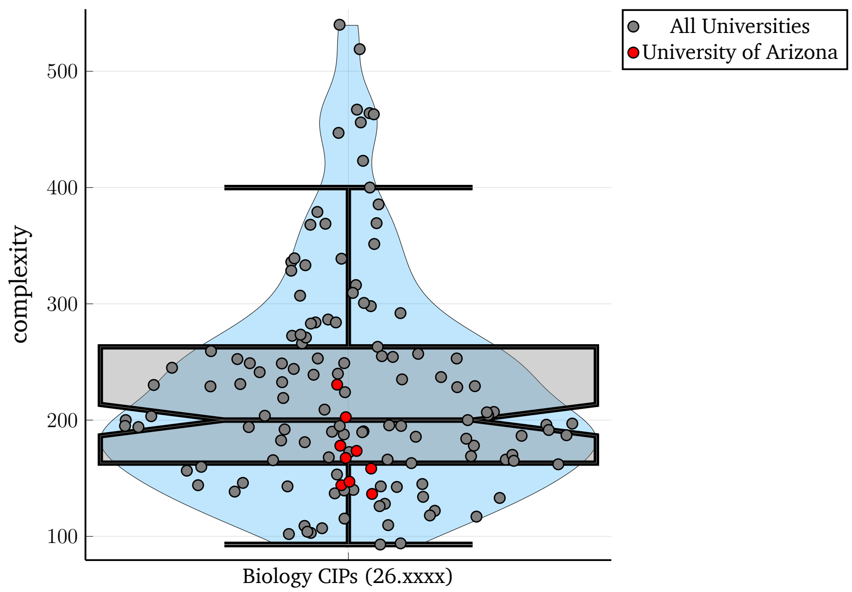


Figure 53: Box scatter plot for two-digit CIP series 26, *Biological and Biomedical Sciences*. The red dots correspond to programs at the University of Arizona, including Applied Biotechnology - Food and Beverage Fermentation (261201), complexity = 158; Biochemistry (260202), complexity = 230; Bioinformatics (261103), complexity = 144; Biology - Biomedical Sciences (260101), complexity = 173; Ecology and Evolutionary Biology (261310), complexity = 136; Microbiology (260502), complexity = 147; Molecular and Cellular Biology - Education and Communication (260406), complexity = 202; Neuroscience and Cognitive Science - Cognition (261501), complexity = 167; Physiology and Medical Sciences - Exercise and Extreme Physi... (260901), complexity = 178.

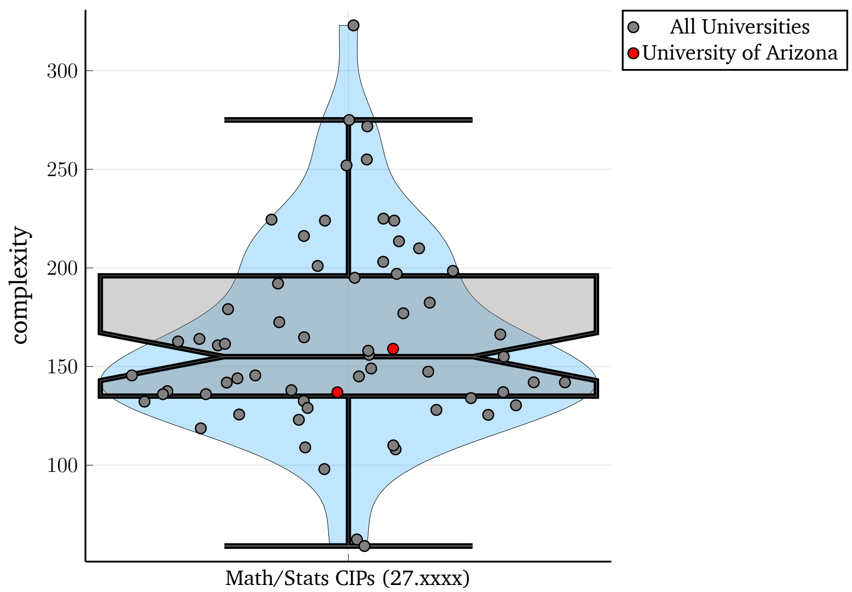


Figure 54: Box scatter plot for two-digit CIP series 27, *Mathematics and Statistics*. The red dots correspond to programs at the University of Arizona, including Mathematics - Applied (270101), complexity = 136; Statistics and Data Science (270503), complexity = 159.

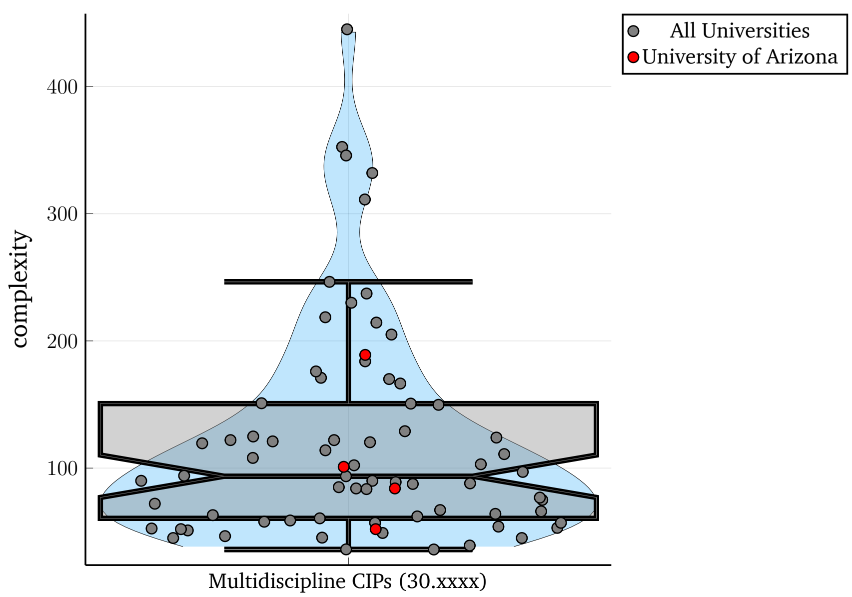


Figure 55: Box scatter plot for two-digit CIP series 30, *Multi/Interdisciplinary Studies*. The red dots correspond to programs at the University of Arizona, including General Studies - Sports and Society (300000), complexity = 52; Global Studies (302001), complexity = 101; Landscape Architecture (303301), complexity = 84; Nutrition and Food Systems (301901), complexity = 189.

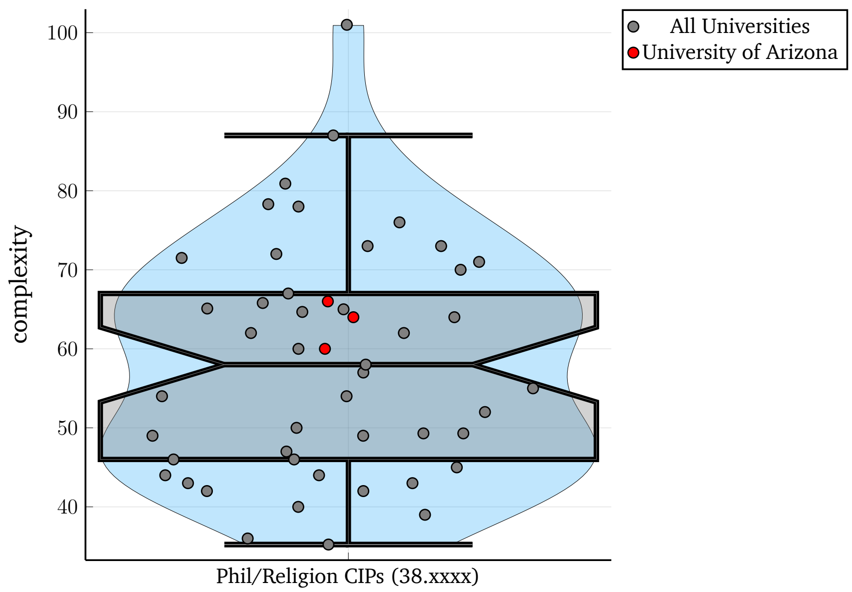


Figure 56: Box scatter plot for two-digit CIP series 38, *Philosophy and Religious Studies*. The red dots correspond to programs at the University of Arizona, including Judaic Studies (380206), complexity = 66; Philosophy Politics Economics and Law (380199), complexity = 64; Religious Studies (380201), complexity = 60.

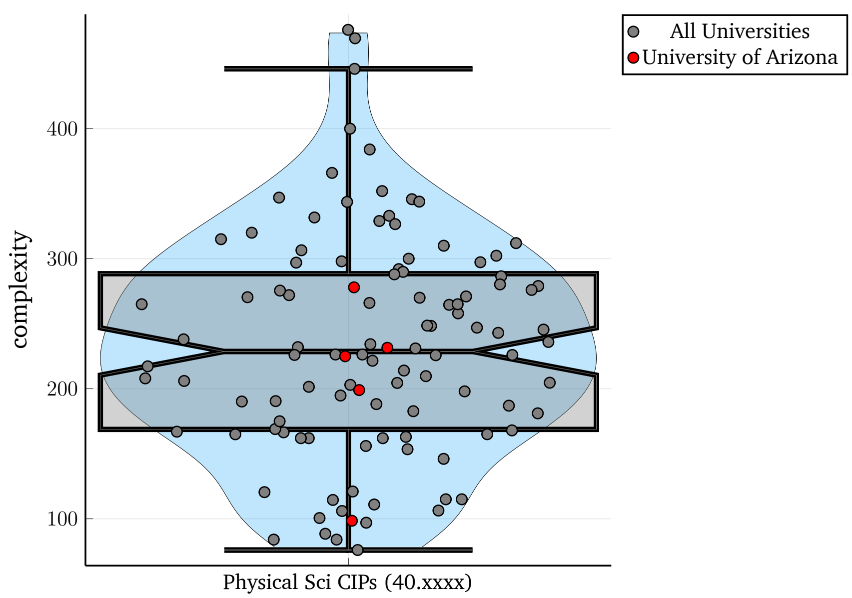


Figure 57: Box scatter plot for two-digit CIP series 40, *Physical Sciences*. The red dots correspond to programs at the University of Arizona, including Astronomy (400201), complexity = 225; Chemistry (400501), complexity = 231; Geosciences - Earth Ocean and Climate Emphasis (400601), complexity = 98; Hydrology and Atmospheric Sciences - Atmospheric Sciences - ... (400605), complexity = 199; Physics (400801), complexity = 278.

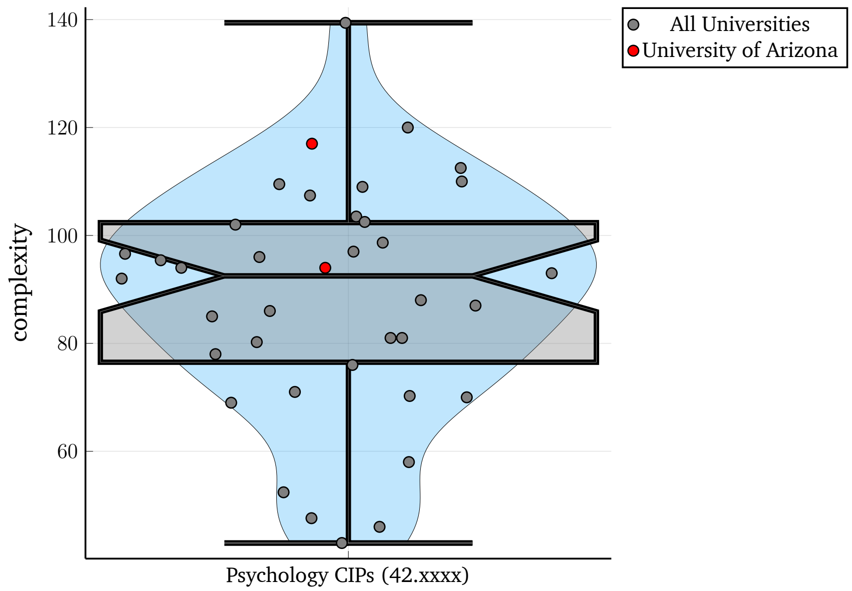


Figure 58: Box scatter plot for two-digit CIP series 42, *Psychology*. The red dots correspond to programs at the University of Arizona, including Psychological Sciences (422704), complexity = 117; Psychology (420101), complexity = 94.

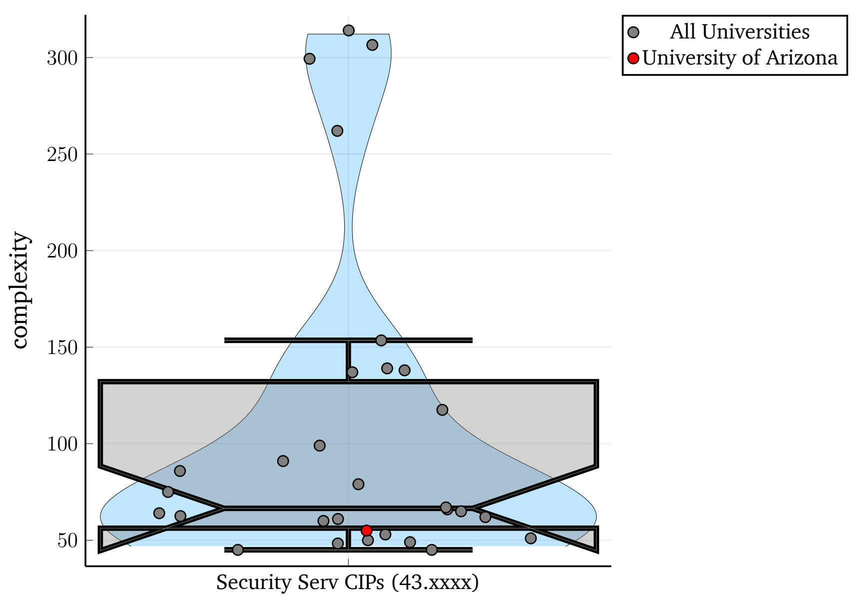


Figure 59: Box scatter plot for two-digit CIP series 43, *Homeland Security, Law Enforcement, Firefighting and Related Protective Services*. The red dots correspond to programs at the University of Arizona, including Criminal Justice Studies (430301), complexity = 55.

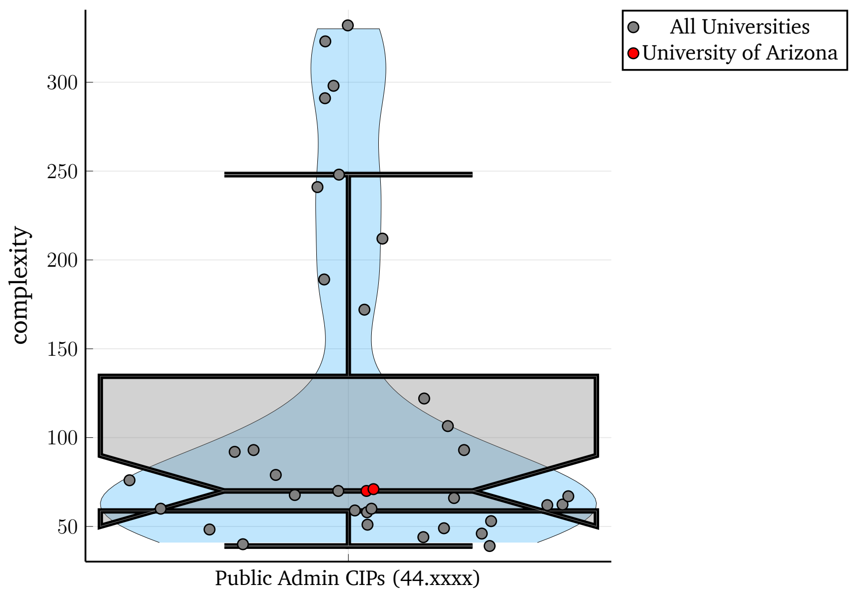


Figure 60: Box scatter plot for two-digit CIP series 44, *Public Administration and Social Service Professions*. The red dots correspond to programs at the University of Arizona, including Human Services (440000), complexity = 70; Public Management and Policy - Public Admin and Management (440401), complexity = 71.

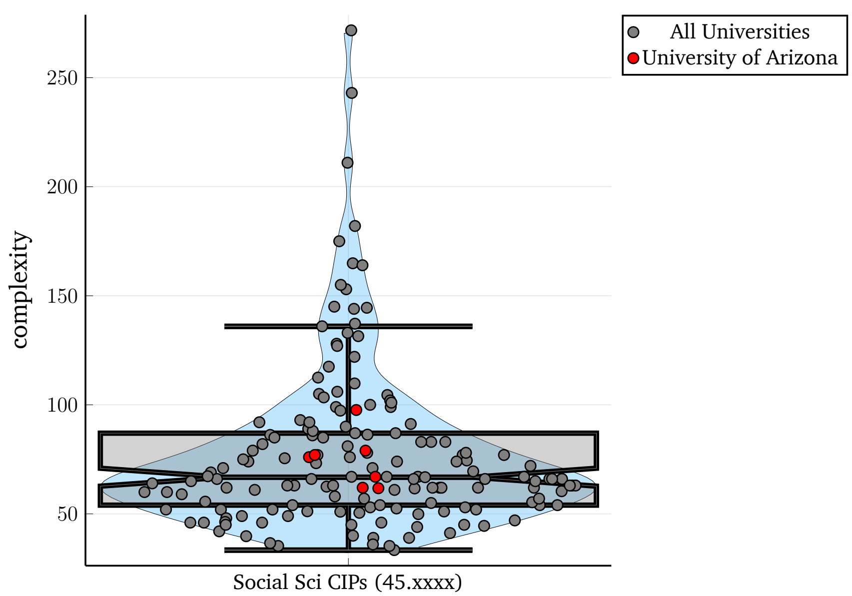


Figure 61: Box scatter plot for two-digit CIP series 45, *Social Sciences*. The red dots correspond to programs at the University of Arizona, including Anthropology (450201), complexity = 76; Economics (450601), complexity = 79; Environmental and Water Resource Economics (450602), complexity = 77; Geographic Information Systems Technology (450702), complexity = 62; Geography - Geographic Info Science (450701), complexity = 61; Government and Public Service (451001), complexity = 97; Sociology (451101), complexity = 67.

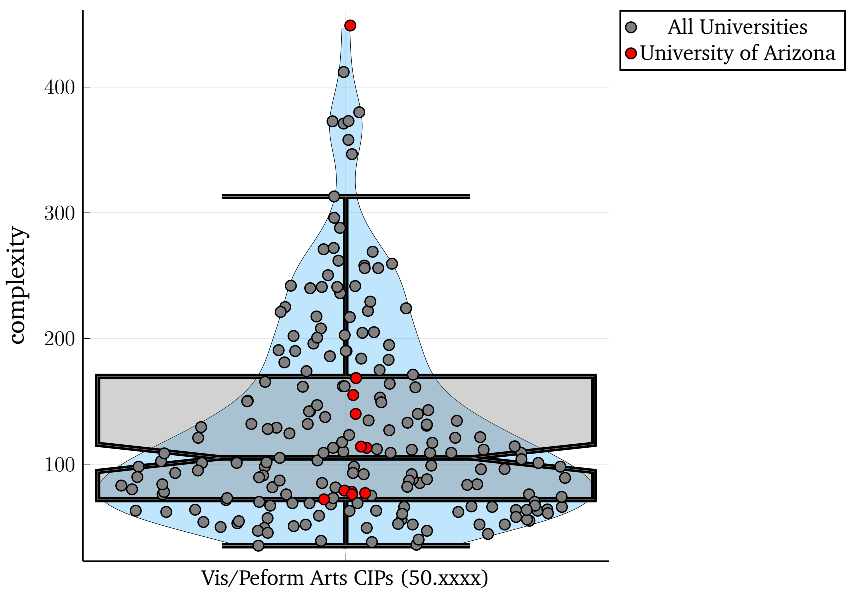


Figure 62: Box scatter plot for two-digit CIP series 50, *Visual and Performing Arts*. The red dots correspond to programs at the University of Arizona, including Performance - Composition (500903), complexity = 155; Studio of Art (500702), complexity = 72; Theatre Arts (500501), complexity = 78; Art History (500703), complexity = 79; Dance (500301), complexity = 113; Design Arts and Practice (500102), complexity = 114; Film and Television (500601), complexity = 168; Game Design and Development (500411), complexity = 76; Live and Immersive Arts (500706), complexity = 77; Music (500901), complexity = 140; Musical Theatre (500509), complexity = 449.

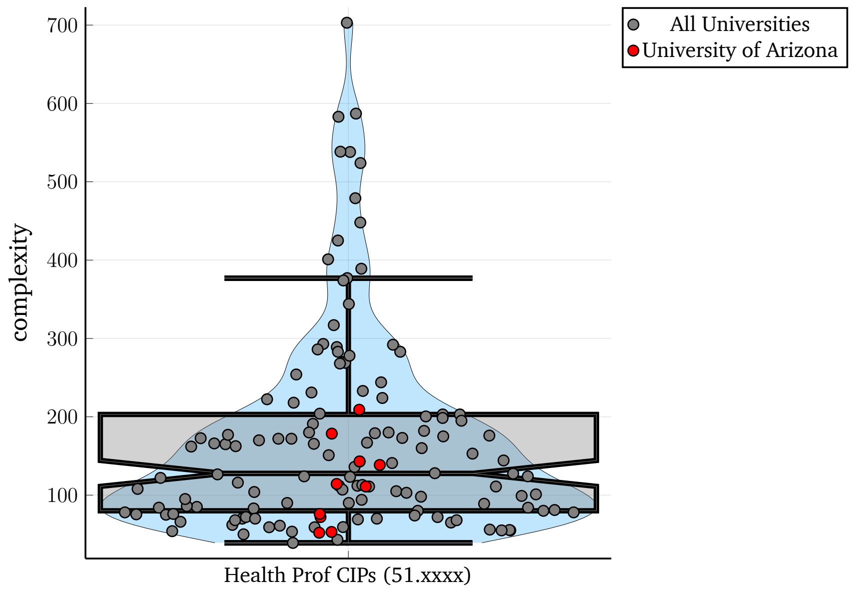


Figure 63: Box scatter plot for two-digit CIP series 51, *Health Professions and Related Programs*. The red dots correspond to programs at the University of Arizona, including Care Health and Society (511504), complexity = 53; Emergency Medical Services (510904), complexity = 72; Medicine - Basic Medical Sciences (511201), complexity = 138; Nursing - Conventional (513801), complexity = 143; Pharmaceutical Sciences (512010), complexity = 209; Public Health - Environ and Occupational Health (512201), complexity = 114; Rehabilitation Studies Services (511599), complexity = 52; Speech - Language and Hearing Science (510201), complexity = 111; Veterinary Science - Applied Animal Behavior (511104), complexity = 178; Public Health - Wellness and Health Promo Prac - Aging and P.. (510001), complexity = 76.

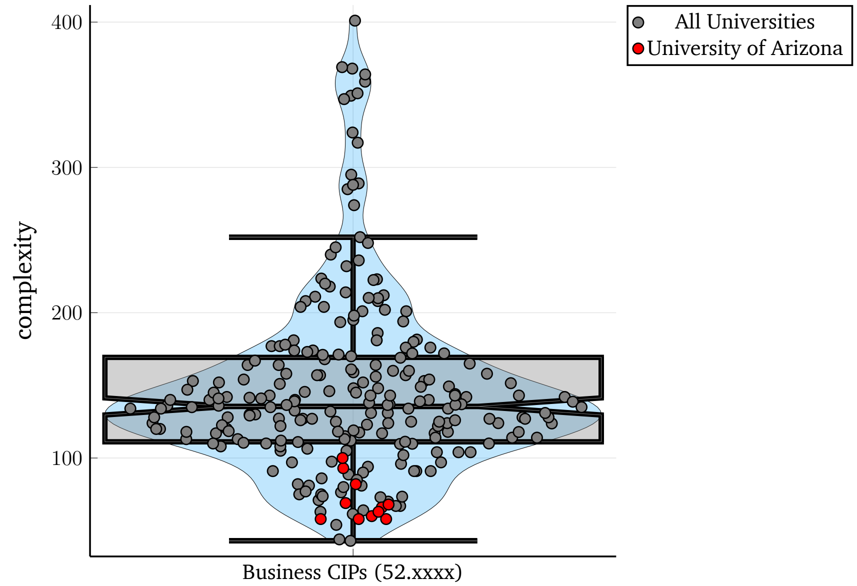


Figure 64: Box scatter plot for two-digit CIP series 52, *Business, Management, Marketing, and Related Support Services*. The red dots correspond to programs at the University of Arizona, including Accounting (520301), complexity = 58; Business Administration (520201), complexity = 66; Business Economics (520601), complexity = 82; Business Management (521001), complexity = 68; Entrepreneurship (520701), complexity = 60; Finance (520801), complexity = 93; Management Info Systems (521201), complexity = 58; Marketing (521401), complexity = 58; Operations Management (520205), complexity = 63; Organizational Leadership and Regional Commerce - Organizati... (520213), complexity = 69; Retailing and Consumer Sciences (520212), complexity = 100.

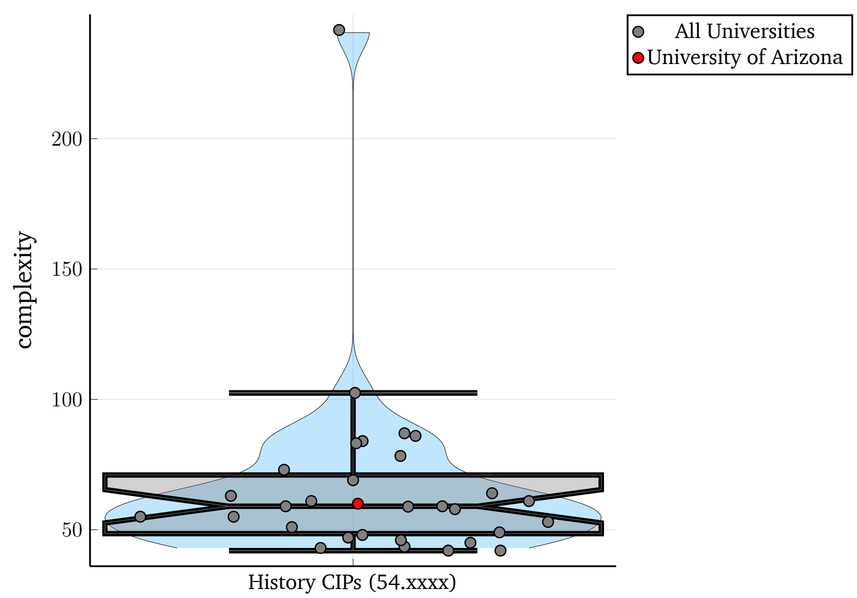


Figure 65: Box scatter plot for two-digit CIP series 54, *History*. The red dots correspond to programs at the University of Arizona, including History (540101), complexity = 60.

4 Conclusions: How to Interpret Results

The results contained in this report are the first of their kind. Previous studies were comparatively preliminary and undertaken with respect to smaller data sets. This study, in contrast, is construed at sufficient scale and undertaken with sufficient analytical rigor as to form a firm basis from which to draw nationally and statistically significant conclusions.

We underscore this point not out of hubris but for contrary reasons. We urge readers to interpret the results of this research with caution, and to utilize them with care. The composition of a curriculum at a particular university is influenced by many factors, including accreditation requirements, state law, a university's academic and administrative policies, curriculum committees, the customs within a discipline, and lately, by intrusive state- and national-level political agendas. The terrain one must navigate to facilitate curricular change is often a minefield of anecdotes with years of precedence, as well as firmly held convictions and entrenched opinions. The results of this research should not be used to bludgeon, but to illuminate pathways through this challenging terrain. Critical debate and reasoned dialog have been the hallmark of higher education for centuries, and we hope this work advances the goal of allowing faculty to apply the same scrutiny to curricula as they do to the grand challenges in their own disciplines. We mean for these reports to serve first and foremost as a spur to further campus reflection and analysis. There is much more to learn, and greater refinement and validation to be had.

We thank our university partners for the benefit of their expertise and their ongoing participation in this enterprise, and we look forward to supporting University Partners in this work and in various efforts to engage faculty and staff in review of their unit's curricular structures for which they are content experts and for which they are chiefly responsible.

As we have mentioned, curricular structures are inherited from generations of colleagues. As a result, there is often relatively little intentionality evident in them. Moreover, never previously have those of us with responsibility for curricular oversight had the benefit of the Curricular Analytics Toolkit and as robust and sound comparative analyses of curricular structures as provided here. In other words, only now, and with the help of this report, do we have the means to tackle the discriminatory barriers in what were heretofore hidden factors assumed to be benign. If we use this research and the tools that made it possible to their best consequence, we may simultaneously improve the quality of our degree programs as well as reduce inequities.

Programs deemed unnecessarily complex are amenable to reform in the direction of greater efficiency. We can in this way improve student time-to-degree, integrate high-impact practices, and reduce cost of attendance. For programs deemed insufficiently scaffolded, there are opportunities to add structure and enhance learning outcomes. Faculty committees at the unit, college, and university level would also, no doubt, better appreciate the student experience via collaboration with academic advisors, student affairs professionals, students and alumni themselves. Academic leaders might better attract prospective students to their majors and minors and facilitate student success.

Our vision is that every unit in every institution of higher learning would regularly use these tools and avail themselves of an expanding context for apples-to-apples comparison, and that students would ultimately learn more and better, and face fewer barriers to success, as a result of data-informed and thus more effectively constructed curricular structures. Our vision is consistent with that first announced in 1998 in the Boyer Report [5] and rearticulated for today in the Boyer 2030 Report [4]. It is, in truth, consistent with the long tradition of democratic educational theory and practice, embodied in the U.S. context by the likes of John Dewey, Ernest L. Boyer, and diverse UERU colleagues coast-to-coast.

References

[1] The Classification of Instructional Programs. National Center for Education Statistics, `https://nces.ed.gov/ipeds/cipcode`, 2020.

[2] U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2024.

[3] I. B. Aban, M. M. Meerschaert, and A. K. Panorska. Parameter estimation for the truncated pareto distribution. *Journal of the American Statistical Association*, 101(473):270–277, 2006.

[4] Boyer 2030 Commission. The equity-excellence imperative: A 2030 blueprint for undergraduate education at U.S. research universities, 2022.

[5] Boyer Commission on Educating Undergraduates in the Research University. Reinventing undergraduate education: A blueprint for America's research universities, 1998.

[6] A. D. Broido and A. Clauset. Scale-free networks are rare. *Nature Communications*, 10(1017):270–277, 2019. `https://doi.org/10.1038/s41467-019008746-5`.

[7] S. M. Burroughs and S. F. Tebbens. Upper-truncated power laws in natural systems. *Pure and Applied Geophysics*, 158:741–757, 2001.

[8] A. Clauset, C. R. Shaliz, and M. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[9] S. Ganesan, A. Lionelle, C. Gill, and C. Brodley. Does reducing curricular complexity impact student success in computer science? In *Special Interest Group, Computer Science Education (SIGCSE) Technical Symposium*, Pittsburgh, PA, Feb.–March 2025. Association for Computing Machinery.

[10] G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman. Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations. *www.arXiv.org*, arXiv:1811.09676 [cs.CY], 2018. `arxiv.org/abs/1811.09676`.

[11] M. Mitzenmacher. A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1(2):226–251, 2004.

[12] M. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[13] M. Patriarca, E. Heinsalu, L. Marzola, A. Chakraborti, and K. Kaski. Power laws in statistical mixtures. In *Proceedings of the European Conference on Complex Systems*, pages 271–282. Springer Proceedings in Complexity, 2016.

[14] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, Boca Raton, FL, 1998.