

# Associations between Library Usage and Undergraduate Student GPA, 2016-2019

Felichism Kabo, Nicholas Paulson, Doreen Bradley, Ken Varnum, Stephanie Teasley

## Contact Author:

Felichism Kabo  
Institute for Social Research  
University of Michigan  
426 Thompson St.  
Ann Arbor, MI 48104  
Email: [fkabo@umich.edu](mailto:fkabo@umich.edu)  
Phone: 734-936-0479

## Abstract

We present the results of a study of the association between online resource use of licensed content provided by the library and short- and long-term student performance. We capture library usage using EZproxy logs, or more precisely whether an individual has at least one EZproxy session in an academic term. We measure student performance using the grade point average (GPA), specifically semester (short-term) and cumulative (long-term) GPA. Relying on models of information behavior, we generate a theoretical framework that suggests that student performance is a function of factors that apply to all students, such as race and gender (the “fixed” effects). But student performance is also impacted by factors such as academic background (e.g., schools, colleges, etc.) that cluster student behaviors and outcomes, and unobserved, time invariant factors at the student-level such as grit and motivation (the “random effects”). We therefore run panel linear mixed effects regression models of the association between library usage and student performance. The results show that library usage, as measured by access to library-licensed content, is significantly associated with both semester and cumulative GPA. The magnitude of the effect is larger for semester GPA, but also varies depending on if a student resides on- or off-campus. The library usage effect on semester GPA is larger for off-campus students compared to their on-campus peers. The reverse is true for the library usage effect on cumulative GPA as it is larger for on-campus students. This study shows how connecting identifiable library data to other institutional can yield shed important insights into how library usage shapes student outcomes.

## Keywords

EZproxy, GPA, student performance, learning analytics, longitudinal analysis, academic library

## Introduction

The Library Learning Analytics Project (LLAP; <https://libraryanalytics.org/>) is funded by the Institute of Museum and Library Services (IMLS) and examines how libraries impact learning outcomes, specifically in the areas of course instruction, publications, and funded research. Learning in these three areas requires that members of the university community engage in activities such as accessing digital data and publication repositories, conducting literature reviews and managing citations, and creating data management plans. These activities usually entail interacting with the library physically such as attending a library instruction session, or virtually such as when retrieving materials through the library's proxy server. We hypothesize that the degree of use of library resources among individuals is associated with student performance or outcomes. Finally, multi-institutional studies are more likely to enable holistic analysis of complex impacts of the library on learning than are analyses of single institutions. The ability to design and implement studies of this nature has been limited by lack of cross-institutional frameworks to enable the sharing of scripts, protocols, and other resources critical to library learning analysis. LLAP bridges this gap by partnering with 14 diverse institutions, our project advisory group (PAG), in the development of shareable data dictionaries, scripts, and protocols based on principled and inclusive engagement. Our paper reports on analyses we performed on the links between off-campus or off-network electronic usage of library resources and undergraduate academic performance over the short- and long-term. More specifically, we examine this relationship before the COVID-19 pandemic when students had the option of accessing library resources both physically and electronic access.

This work is informed by models of information behavior (Johnson 1997, Wilson 1999, Wilson 2017), and builds on two lines of inquiry: a) research into the associations between college residence and academic performance, and b) work on digital inequalities or the digital divide. Information behavior describes how individuals seek and utilize information (Bates 2017). Information behavior is contingent on factors such as social contexts, socio-demographics, individual expertise, and access to and ease of use of technology (Haglund and Olsson 2008, Niu and Hemminger 2012, Bates 2017). We examine the link from library usage to student outcomes by a) defining library usage in terms of online resource use of licensed content provided by the library, hence an implied need for better digital resources, and b) evaluating the impacts of on-campus residency as a proxy for ease of access to library and other resources and reliable internet.

Research on campus residency has examined the issue of whether there are gains in learning and academic performance from living on- versus off-campus. A study of close to 95,000 first year students in the United States found living on-campus was significantly associated with a range of learning variables even though the residency effect size was small to medium (Graham, Hurtado et al. 2018). An earlier study of first-year students found that the benefits of on-campus residency on academic performance applied to certain racial groups and not to all students. Specifically, Black students who lived on-campus had significantly higher grade point averages (GPAs) than those from the same racial group that lived off-campus (López Turley and Wodtke 2010). Approaching the issue from a different angle, a study of the causal link between campus residency and academic outcomes found living in university-owned housing had a positive association with student retention (Schudde 2011). This finding was in line with prior analysis that established an association between on-campus

living, and academic performance and student retention for first-year students (Huhn 2006). However, an important caveat is that students who were better prepared academically were more likely to live on-campus as opposed to off-campus (Huhn 2006). Most studies of the link between on-campus residence and student persistence have been based on four-year institutions. One exception is a quasi-experimental analysis of community college students that found that living on-campus was associated with a significant increase in upward transfer (to a four-year institution) and, subsequently, bachelor's degree completion rates (Turk and González Canché 2019). Finally, there are exceptions to the general findings on positive associations between on-campus residence and academic outcomes. For example, a study that was conducted at a public four-year university in the Southeast United States that found that commuter or off-campus students had higher GPAs than residential or on-campus students (Simpson and Burnett 2017).

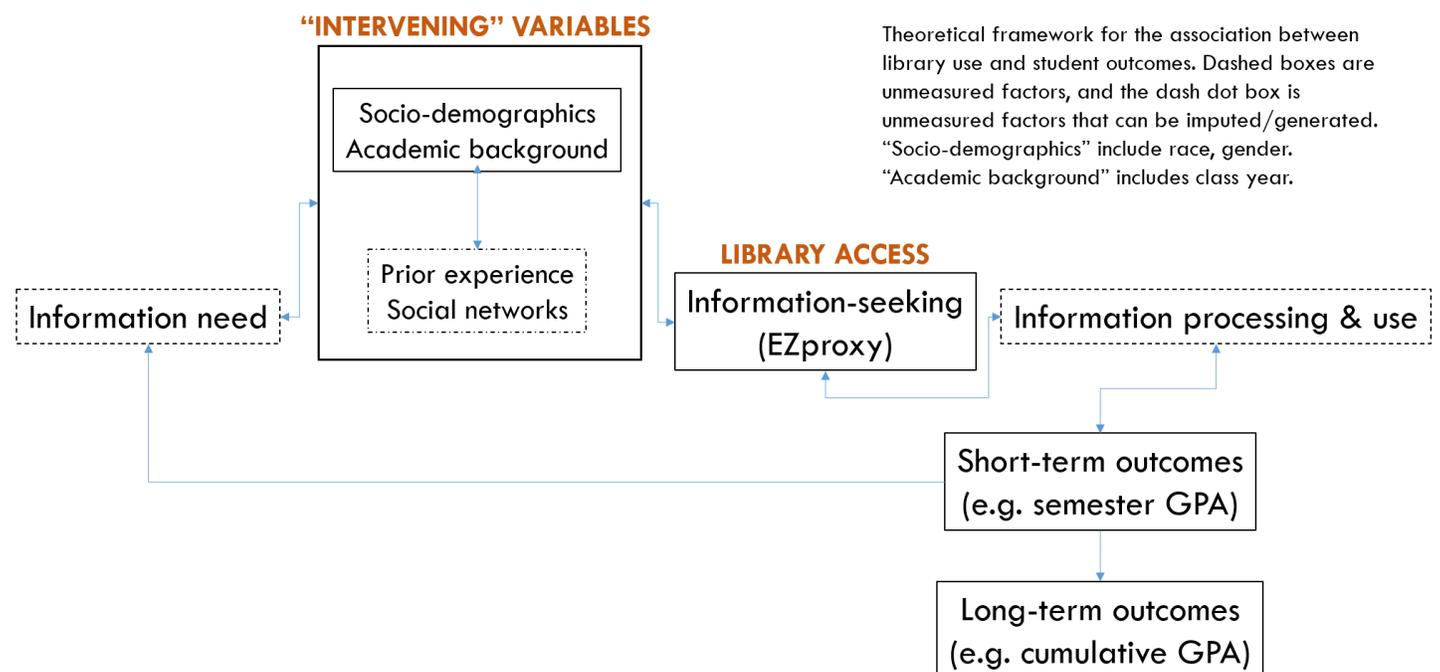
There are well-documented digital disparities in American K-16 education that are shaped by demographic, geographic, and economic factors. The term “digital divide” is commonly used to describe the gap between those privileged to benefit from the internet and other information and communication technologies (ICT), and those who are not. Digital inequalities and disparities impact a broad range of life opportunities and outcomes beyond education, such as economic activity and even health care (Robinson, Cotten et al. 2015, Zhang, Pérez-Stable et al. 2017). In education, digital inequalities and disparities are a life-course issue and affect disadvantaged students from early in the K-16 pipeline (Cleary, Pierce et al. 2006, Jackson, von Eye et al. 2006), all the way through college (Farrell 2005, Jones, Johnson-Yale et al. 2009). The increasing use of technology inside and outside the classroom has significant implications for the digital divide and its impact on student performance. Importantly, some groups of students are systematically more likely to experience digital disparities than others. For example, in 2015 higher percentages of students who were White (66%) used the internet at home compared to Black (53%), Hispanic (52%), and American Indian/Alaska Native (49%) students (KewalRamani, Zhang et al. 2018). In fact, American Indian/Alaska Native students are more likely than other racial groups to have no internet access, or to have only dial-up internet access at home (Musu 2018). Moreover, the interaction of demography and geography serves to disadvantage some students further still. Thus, while 18% of all students in remote rural areas did not have internet access or had only dial-up access in 2015, a much larger percentage of Black (41%) students in remote rural areas did not have internet access compared to White (13%) and Asian (11%) students. Having no or low-bandwidth internet is detrimental for any form of online learning as, for starters, students cannot participate in classes offered via video conferencing systems which rely on high-speed internet (Correia 2020). The COVID-19 pandemic has deepened or worsened the effects of the digital divide, such as for rural students (Lai and Widmar 2021). A group that has been especially impacted by the pandemic is students of color who, as noted earlier, are more likely to lack access to reliable broadband internet and even computers. For minority students, the pandemic has exacerbated existing educational disparities, and has likely widened the achievement gap for students of low socioeconomic status (Kuhfeld, Soland et al. 2020, Reza 2020).

In the United States, the effects of the pandemic on the digital divide are being felt up and down the entire K-16 pipeline, including in higher education. There were varied institutional responses across the American higher education landscape which should, rather perversely, present opportunities for “quasi-experimental” observations of the impacts of the digital divide on amplifying disparities in

student performance. For example, where many colleges and universities stipulated that students residing on-campus leave these residences, some made allowances for students who could not return home and thus enabled them to still have access to reliable broadband internet via the institution (Day, Chang et al. 2021). What was fairly universal, however, was the extent and speed with which university libraries adapted to offering primarily online resources (Day, Chang et al. 2021), which can only meaningfully be accessed via reliable internet connections. Thus, not only were students no longer able to access the library’s physical collections, but they also no longer had access to the library as a study space including for group or collaborative activities (Mehta and Wang 2020, Day, Chang et al. 2021). By examining how “regular” (pre-pandemic) electronic library usage is associated with academic performance, this study may thus help us better understand the likely impacts of the worsening of the digital divide during the pandemic. Our assumption is that the magnitude of any relationship between electronic library usage and academic performance has only amplified in the aftermath of the pandemic.

## Methods

This paper focuses on the association between off-campus or off-network electronic library resource use captured by EZproxy event logs, and short- and long-term academic performance, specifically semester and cumulative GPA, respectively. EZproxy is proxy server software that is widely used by many academic libraries to give authenticated off-campus users access to electronic resources licensed by the library as if they were on campus. Stated simply, after authenticating to a campus system, off-campus users receive an on-campus IP address and are then considered to be a member of the campus community by the information provider. The study sample is all undergraduate students enrolled at the University of Michigan (U-M) in the six-semester period from fall 2016 through winter 2019 (or September 2016 through April 2019).



**Figure 1.** Theoretical framework for associations between library usage and student outcomes adapted from Wilson (1999) and Johnson (1997) models of information behavior.

Building on models of information seeking behavior, we developed a theoretical framework (Figure 1) that correlates student performance with library usage as captured by EZproxy sessions, controlling for factors like socio-demographics and academic background (Johnson 1997, Wilson 1999).

We cleaned and normalized raw, unstructured EZproxy logs using Python scripts and regular expressions. We then entered the data into a relational database using structured query language (SQL) scripts. Over 80% of the EZproxy data have strong university identifiers which facilitate merges with other administrative data, such as course instruction and student data. It is critical to note that EZproxy logs available to the study A) did not include any on-campus usage and B) did not include anyone who used the campus VPN. Using SQL and R scripts, we merged the data and exported the resultant data set into Stata 16 for modeling and analysis (StataCorp 2019).

The theoretical framework shown in Figure 1 suggests that student outcomes are a function of factors such as race and gender that apply to all the students in the study (“fixed effects”), and factors such as academic units or schools that cluster student behaviors and outcomes (“random effects”). We also account for student random effects for unobserved, time invariant factors, such as motivation or grit. Thus, we ran panel linear mixed effects regression models of the association between library usage (having one or more EZproxy sessions in an academic term) and student GPA, contingent on students being enrolled in at least four semesters in the period Fall 2016 – Winter 2019.

### *Variables*

The two dependent variables are short- and long-term student outcomes, specifically semester GPA (“SEM\_GPA”) and cumulative GPA (“CUM\_GPA”), respectively. SEM\_GPA is a continuous 0 – 4.4 scale while CUM\_GPA is on a continuous 0 – 4.314 scale. The independent variable “Ever EZproxy Session in Term” is coded one if a student is associated with one or more EZproxy sessions during an academic term and is coded zero otherwise.

We also account or control for potential “intervening” variables as follows. The variable “On-campus Residence” is coded one if a student was residing in a university residence, and zero otherwise. The variable “High School GPA” is on a continuous 0 – 4 scale and captures a student’s academic performance before matriculation at the university. Gender is captured by the dichotomous variable “GENDER” (1 = Female, 2 = Male). Note that the learning analytics dataset used for the study does not account for non-binary options. The effects of race, first generation status, family income, and class level were captured using the categorical variables “RACE” (1 = White, 2 = Asian, 3 = Black, 4 = Hispanic, 5 = Two or More, 6 = Other, 7 = Not Indicated), “FIRST GENERATION” (1 = First Gen, 2 = Not First Gen, 3 = Don’t Know), “FAMILY INCOME” (1 = More than \$100,000, 2 = Less than \$25,000, 3 = \$25,000 - \$49,999, 4 = \$50,000 - \$74,999, 5 = \$75,000 - \$99,999, 6 = Don’t Know, 7 = Missing), and “CLASS LEVEL” (1 = Freshman, 2 = Sophomore, 3 = Junior, 4 = Senior), respectively.

### *Statistical Modeling*

We ran panel linear mixed-effects (LME) regression models with random effects for individuals and by school or academic unit (see Table A.7 in the appendix for a list of the 15 schools that undergraduate students were affiliated with). LME models have both fixed effects, which are directly estimated and

are analogous to standard regression coefficients, and random effects, which in our case take the form of random intercepts. The fixed effects in our LME models correspond to the variables in the previous section, while the random effects account for the fact that student behaviors and outcomes may be, instead of being uniform across all undergraduates, grouped by academic units as these may map onto disciplinary and organizational boundaries within the university. The random effects also enable us to account for unobserved, time invariant individual-level factors, such as motivation or grit. Table A.7 in the appendix shows that there are notable differences across schools with respect to the percentage of students that have at least one EZproxy session during an academic term. After each LME model a likelihood-ratio was run comparing this model with a one-level ordinary linear regression. This test was highly significant for each of the LME models in our study, supporting the decision to use the LME model.

## Results

### *Descriptive statistics*

We were able to identify 49-58% of all undergraduates as having one or more EZproxy sessions during an academic term over the course of the six academic terms from fall 2016 to winter 2019 (Table 1). For example, in winter 2019 we identified about 57% of the undergraduate population as having had at least one EZproxy session over the course of the semester.

<b>Academic Term</b>	<b>Number of Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	28,682	16,605	58%
WN 2017	27,408	13,434	49%
FA 2017	29,161	16,034	55%
WN 2018	27,852	14,855	53%
FA 2018	29,726	16,191	54%
WN 2019	28,355	16,299	57%

However, there are some notable demographic differences as shown in Table 2 below which shows the descriptive statistics on EZproxy usage across demographic categories for winter 2019 (see the appendix for descriptive statistics on all semesters). Off-campus students are more likely to have at least one EZproxy session in the academic term than are on-campus students. This makes sense because students who are on-campus are more likely to access electronic library resources on the university's network, in which case authentication is not required. Recall that we are only able to identify students in the EZproxy logs when they had strong identifiers, which happens when authentication is required such as when a student accesses electronic library resources outside the university's network e.g., from an off-campus residence, coffee shop, etc. There is a notable gender difference with nearly two-thirds of females showing up in the EZproxy logs compared to half of all males. This is even though more males (69%) than females (66%) resided off-campus in winter 2019. Note that the likelihood of having at least one EZproxy session increases with each class level. A plausible explanation could be that this is because students are more likely to move or reside off-campus as they progress from freshman to seniors. However, a factor that weakens this argument is that at

U-M it is not compulsory or mandatory for freshmen and sophomores to live on-campus as is the case in some colleges and universities. An alternative explanation is that lower-level classes are less research-intensive and where there are research and writing projects, they may not need library-provided resources but can rather be accomplished through open-web, non-licensed, materials.

**Table 2.** Percentage of students associated with EZproxy sessions by socio-demographics and academic background, Winter 2019

<b>Variable</b>	<b>Category</b>	<b>Number of Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
First Gen Status	Don't Know	47	32	68%
	First Gen	3,890	2,310	59%
	Not First Gen	24,418	13,957	57%
Family Income	Less than \$25,000	1,507	923	61%
	\$25,000 - \$49,999	2,212	1,269	57%
	\$50,000 - \$74,999	2,009	1,217	61%
	\$75,000 - \$99,999	2,074	1,213	58%
	More than \$100,000	13,951	7,892	57%
	Don't Know	515	278	54%
	Missing Income Information	6,087	3,507	58%
Class Level	Freshman	2,557	1,300	51%
	Sophomore	6,397	3,373	53%
	Junior	7,132	4,114	58%
	Senior	12,269	7,512	61%
Race	Asian	5,829	3,137	54%
	Black	1,268	766	60%
	Hispanic	1,899	1,099	58%
	White	16,604	9,738	59%
	2 or More	1,302	745	57%
	Other	46	22	48%
	Not Indic	1,407	792	56%
Gender	Female	14,204	9,219	65%
	Male	14,151	7,080	50%
Residency	On-campus	9,261	4,540	49%
	Off-campus	19,110	11,765	62%
Academic Unit	Undergrad Music, Thtre & Dance	717	515	72%
	Undergraduate Architecture	181	124	69%
	Undergraduate Art and Design	524	381	73%
	Undergraduate Business Admin	1,799	740	41%
	Undergraduate Dental Hygiene	101	70	69%
	Undergraduate Education	126	54	43%
	Undergraduate Engineering	6,313	2,847	45%
	Undergraduate Information	260	122	47%
	Undergraduate Joined Deg Prog	10	7	70%
	Undergraduate Kinesiology	954	678	71%
	Undergraduate L S & A	16,409	10,030	61%
	Undergraduate Nursing	607	475	78%
	Undergraduate Pharmacy	55	36	65%

Undergraduate Public Health	157	116	74%
Undergraduate Public Policy	142	104	73%

---

Finally, there are noteworthy differences between academic units. Additional work would be needed to clarify the factors that account for these differences. For example, 45% of engineering undergraduates had at least one EZproxy session compared to 73% of art and design undergraduates even though both academic units are co-located at the university. A potential explanation could be that these differences reflect disciplinary differences (STEM versus arts and humanities). Another plausible explanation could be that the differences reflect gaps in technological expertise between the two groups of students, with engineering students being more likely to access electronic library resources using the university’s virtual private networks (VPNs) which bypasses the authentication process on the library’s proxy server. We should also keep in mind factors such as the interplay between residency and socioeconomic statuses. It is more expensive to live on- rather than off-campus, implying that students in the former group may tend to be from better off families. For example, 78% of nursing undergraduates had at least one EZproxy session compared to 41% of business administration undergraduates. Tabulations of residency for the two academic units showed that 32% of business undergraduates resided on-campus in winter 2019, compared to 20% of nursing undergraduates. Similarly, tabulations of family income for the two academic units showed that 58% of business undergraduates had a family income of more than \$100,000, compared to 48% of nursing undergraduates. These findings suggest that library usage data have the potential to reveal existing disparities and inequalities and could therefore help libraries make significant analytical contributions of interest to their parent institutions.

### *Regression models*

The results from the regression modeling are summarized in Tables 3 (semester GPA) and 4 (cumulative GPA). The regression models showed positive and statistically significant associations between having at least one EZproxy session in an academic term and student GPA both in the short- and long-term, controlling for “intervening” variables including residency, race, gender, high school GPA, family income, first generation status, and class status. In the short-term, having at least one EZproxy session *during an academic term* was correlated with a student having a semester GPA that is 0.14 higher compared to students who we could not identify as having had any EZproxy sessions (model 1). The size of this effect ( $\beta = 0.138$ ) is larger than that of residing on-campus ( $\beta = 0.097$ ). To examine further the impact of campus residency considering the link between authentication requirements and a student’s detectability in the EZproxy logs, we also ran separate models for on-campus (model 2) and off-campus (model 3) students. Having at least one EZproxy session has a larger effect for off-campus students ( $\beta = 0.171$ ) than it does for on-campus students ( $\beta = 0.084$ ). That is, having at least one EZproxy session in an academic term is correlated with an off-campus student having a semester GPA that is 0.17 higher compared to off-campus students without an identifiable EZproxy session. In comparison, having at least one EZproxy session in an academic term is correlated with an on-campus student having a semester GPA that is 0.09 higher compared to on-campus students without a detectable EZproxy session. With respect to the other “intervening” variables, it is noteworthy that the gender gap with respect to GPA (females have higher GPAs) is smaller for on-campus students compared to their off-campus peers. It is also interesting that, the

small sizes of the effects notwithstanding, the first-generation disadvantage (non-first-generation students have higher GPAs) is larger for on-campus students relative to off-campus ones.

<b>Table 3.</b> Panel LME Regressions on Semester GPA, FA 2016 - WN 2019 (4 or More Semesters)			
VARIABLES	(1)	(2)	(3)
	SEM_GPA	SEM_GPA	SEM_GPA
Ever EZproxy Session in Term	0.138*** (0.00304)	0.0837*** (0.00415)	0.171*** (0.00419)
On-campus Residence	0.0967*** (0.00471)		
High School GPA	0.0273*** (0.00194)	0.0435*** (0.00345)	0.0211*** (0.00235)
<i>GENDER (Reference = Female)</i>			
Male	-0.0908*** (0.00529)	-0.0616*** (0.00662)	-0.108*** (0.00685)
<i>RACE (reference = White)</i>			
Asian	0.0499*** (0.00660)	0.0534*** (0.00838)	0.0404*** (0.00851)
Black	-0.376*** (0.0128)	-0.374*** (0.0145)	-0.400*** (0.0181)
Hispanic	-0.164*** (0.0107)	-0.181*** (0.0126)	-0.143*** (0.0145)
2 or More	-0.101*** (0.0126)	-0.0812*** (0.0150)	-0.121*** (0.0167)
Other	-0.239*** (0.0631)	-0.209** (0.0781)	-0.255** (0.0784)
Not Indic	-0.00568 (0.0121)	0.0168 (0.0160)	-0.0188 (0.0155)
<i>FIRST GENERATION (reference = First Gen)</i>			
Not First Gen	0.119*** (0.00851)	0.138*** (0.0106)	0.112*** (0.0112)
Don't Know	-0.166** (0.0525)	-0.0157 (0.0845)	-0.202** (0.0640)
<i>FAMILY INCOME (reference = More than \$100,000)</i>			
Less than \$25,000	-0.150*** (0.0127)	-0.129*** (0.0159)	-0.166*** (0.0167)
\$25,000 - \$49,999	-0.101*** (0.0106)	-0.115*** (0.0131)	-0.102*** (0.0141)
\$50,000 - \$74,999	-0.0557*** (0.0104)	-0.0719*** (0.0133)	-0.0581*** (0.0134)
\$75,000 - \$99,999	-0.0545*** (0.0100)	-0.0528*** (0.0129)	-0.0572*** (0.0128)
Don't Know	-0.0505* (0.0196)	-0.0385 (0.0238)	-0.0688** (0.0260)
Missing Income Information	-0.00505 (0.00652)	-0.0117 (0.00827)	-0.00127 (0.00831)

<i>CLASS LEVEL (reference = Freshman)</i>			
Sophomore	0.0176*** (0.00498)	0.0237*** (0.00455)	0.0184 (0.0229)
Junior	0.0326*** (0.00605)	0.00259 (0.00680)	0.0704** (0.0229)
Senior	0.0815*** (0.00662)	0.0403*** (0.0112)	0.116*** (0.0230)
Constant	3.207*** (0.0357)	3.242*** (0.0444)	3.174*** (0.0448)
Observations	151,049	53,896	97,153
Cohort	AllUGrads	AUG-OnCam	AUG-OffCam
Standard errors in parentheses			
*** p<0.001, ** p<0.01, * p<0.05			

Having at least one EZproxy session in an academic term has smaller effects in the long-term than it does in the short-term. Model 4 shows that having at least one EZproxy session in an academic term was correlated with a student having a cumulative GPA that is 0.02 higher than that for a student with no detectable EZproxy session. To further examine the effect of being on- or off-campus, we ran separate models for on- (model 5) and off-campus (model 6) students which show differences between the two groups of students but in ways that are contrary to semester GPA. Having at least one EZproxy session in an academic term has an effect that is larger in magnitude for on-campus students compared to their off-campus peers, the caveat being that the magnitude of both effects is very small. Note that, like semester GPA, the female advantage in cumulative GPA was smaller for on-campus students relative to off-campus students. In a similar vein, the first-generation disadvantage with respect to cumulative GPA is larger for students who are on-campus compared to those that are off-campus, keeping in mind the small sizes of the effects.

<b>Table 4.</b> Panel LME Regressions on Cumulative GPA, FA 2016 - WN 2019 (4 or More Semesters)			
VARIABLES	(4) CUM_GPA	(5) CUM_GPA	(6) CUM_GPA
Ever EZproxy Session in Term	0.0201*** (0.000896)	0.0242*** (0.00190)	0.0144*** (0.000871)
On-campus Residence	0.0216*** (0.00149)		
High School GPA	0.0222*** (0.00162)	0.0364*** (0.00313)	0.0141*** (0.00182)
<i>GENDER (Reference = Female)</i>			
Male	-0.0735*** (0.00447)	-0.0573*** (0.00603)	-0.0841*** (0.00528)
<i>RACE (reference = White)</i>			
Asian	0.0655*** (0.00558)	0.0654*** (0.00763)	0.0559*** (0.00658)
Black	-0.330***	-0.328***	-0.364***

	(0.0108)	(0.0134)	(0.0139)
Hispanic	-0.157***	-0.168***	-0.150***
	(0.00904)	(0.0116)	(0.0112)
2 or More	-0.0769***	-0.0648***	-0.0885***
	(0.0107)	(0.0137)	(0.0129)
Other	-0.197***	-0.159*	-0.192**
	(0.0549)	(0.0717)	(0.0611)
Not Indic	0.0120	0.0296*	-0.00456
	(0.0102)	(0.0145)	(0.0121)
<i>FIRST GENERATION (reference = First Gen)</i>			
Not First Gen	0.105***	0.118***	0.102***
	(0.00721)	(0.00971)	(0.00867)
Don't Know	-0.209***	-0.0901	-0.233***
	(0.0451)	(0.0786)	(0.0503)
<i>FAMILY INCOME (reference = More than \$100,000)</i>			
Less than \$25,000	-0.113***	-0.101***	-0.126***
	(0.0107)	(0.0147)	(0.0129)
\$25,000 - \$49,999	-0.0806***	-0.0963***	-0.0837***
	(0.00901)	(0.0120)	(0.0109)
\$50,000 - \$74,999	-0.0342***	-0.0543***	-0.0364***
	(0.00883)	(0.0122)	(0.0104)
\$75,000 - \$99,999	-0.0438***	-0.0416***	-0.0460***
	(0.00850)	(0.0118)	(0.00990)
Don't Know	-0.0326*	-0.0317	-0.0454*
	(0.0165)	(0.0217)	(0.0200)
Missing Income Information	-0.00391	-0.00968	-0.00127
	(0.00553)	(0.00753)	(0.00644)
<i>CLASS LEVEL (reference = Freshman)</i>			
Sophomore	-0.00343*	-0.00615**	0.00310
	(0.00150)	(0.00209)	(0.00513)
Junior	-0.00137	-0.0235***	0.0217***
	(0.00187)	(0.00322)	(0.00517)
Senior	0.0241***	-0.0114*	0.0483***
	(0.00209)	(0.00538)	(0.00520)
Constant	3.430***	3.376***	3.456***
	(0.0275)	(0.0358)	(0.0304)
Observations	151,049	53,896	97,153
Cohort	AllUGrads	AUG-OnCam	AUG-OffCam
Standard errors in parentheses			
*** p<0.001, ** p<0.01, * p<0.05			

The study findings suggest that using library resources has differential impacts on students' academic performance contingent on demographic and socioeconomic factors. These impacts were larger in magnitude for short-term rather than long-term performance. For example, for semester GPA, first generation students had a lower GPA (-0.119) than non-first-generation students. Further, males had

a lower semester GPA (-0.091) than females. Thus, the impacts of gender and first-generation status on semester GPA were smaller in magnitude than the impact of having at least one EZproxy session during an academic term.

## **Study Limitations**

It is difficult to disentangle students who are off campus and not using the VPN and those who access electronic resources via the proxy server. Undoubtedly there are economic, technical, and experiential causes to these differences, however they were not explored in this study. There are a host of socioeconomic factors which could affect student use of the library proxy server. Another factor which may account for some observed differences are in the varying nature and demands of curricula across programs and colleges. Certain programs have required library and research components at different points in their programs. For example, some may have library research needs in first year experience courses, whereas other programs may be more library-intensive in the third or fourth year.

## **Discussion**

Because library data are often not integrated into other university data, there are major obstacles to demonstrating the complexity of the value of academic library usage to students who use these resources. We show how merging library usage and student outcome data at one institution yields valuable insights on the value of the academic library. Understanding patterns of off-campus use of library resources offers an additional point of insight into potential gaps in use by certain groups of students. For example, as noted earlier, living off campus can be associated with lower academic success and retention for certain groups of students. If students in particular programs tend to live off campus, yet their programs are library-research intensive, what could this mean for those students? For example, 4 out of 5 of undergraduate nursing students live off campus, yet we know the nursing program integrates the library heavily in its curriculum. We could explore off campus use by students in this program to potentially identify students at risk of falling through the cracks or to provide indicators to faculty advisors if a student's GPA in research-intensive courses falls below a certain threshold. As additional data from other library services may be added in the future, libraries can develop models to explore other questions around library usage, student success, and curricular integration. Through all this work, one of our goals is to eliminate educational disparities where we can. Library usage data adds depth of perspective of the student experience, and student engagement broadly, during the undergraduate years. and can be a valuable addition to institutions of higher education as they continue to make data-informed decisions that improve undergraduate education. Further, in the process of doing this work we have created shareable scripts and tools that could be used to replicate our work in other institutional settings.

## **Future Research**

There are several future directions pointed to by this research, for example: What is the effect of course selection on the need to use library-licensed resources? How does level of study correlate to use of licensed resources and to academic outcomes?

## References

- Bates, M. J. (2017). Information Behavior. Encyclopedia of Library and Information Science, Fourth Edition, CRC Press.
- Cleary, P. F., G. Pierce and E. M. Trauth (2006). "Closing the digital divide: understanding racial, ethnic, social class, gender and geographic disparities in Internet use among school age children in the United States." Universal Access in the Information Society **4**(4): 354-373.
- Correia, A.-P. (2020). "Healing the Digital Divide During the COVID-19 Pandemic." Quarterly Review of Distance Education **21**(1): 13-21.
- Day, T., I. C. C. Chang, C. K. L. Chung, W. E. Doolittle, J. Housel and P. N. McDaniel (2021). "The Immediate Impact of COVID-19 on Postsecondary Teaching and Learning." The Professional Geographer **73**(1): 1-13.
- Farrell, E. F. (2005). "Among Freshmen, a Growing Digital Divide." The Chronicle of Higher Education **51**(22): A32.
- Graham, P. A., S. S. Hurtado and R. M. Gonyea (2018). "The Benefits of Living on Campus: Do Residence Halls Provide Distinctive Environments of Engagement?" Journal of Student Affairs Research and Practice **55**(3): 255-269.
- Haglund, L. and P. Olsson (2008). "The Impact on University Libraries of Changes in Information Behavior Among Academic Researchers: A Multiple Case Study." The Journal of Academic Librarianship **34**(1): 52-59.
- Huhn, C. (2006). The 'Housing Effect' on First-Year Outcomes. Madison, WI, Academic Planning and Analysis, Office of the Provost, University of Wisconsin-Madison.
- Jackson, L. A., A. von Eye, F. A. Biocca, G. Barbatsis, Y. Zhao and H. E. Fitzgerald (2006). "Does home internet use influence the academic performance of low-income children?" Developmental Psychology **42**(3): 429-435.
- Johnson, J. D. (1997). Cancer-related information seeking. Cresskill, N.J. :, Hampton Press.
- Jones, S., C. Johnson-Yale, S. Millermaier and F. S. Pérez (2009). "U.S. College Students' Internet Use: Race, Gender and Digital Divides." Journal of Computer-Mediated Communication **14**(2): 244-264.
- KewalRamani, A., J. Zhang, X. Wang, A. Rathbun, L. Corcoran, M. Diliberti and J. Zhang (2018). Student Access to Digital Learning Resources outside of the Classroom. NCES 2017-098. Washington, DC, U.S. Department of Education.
- Kuhfeld, M., J. Soland, B. Tarasawa, A. Johnson, E. Ruzek and J. Liu (2020). "Projecting the Potential Impact of COVID-19 School Closures on Academic Achievement." Educational Researcher **49**(8): 549-565.
- Lai, J. and N. O. Widmar (2021). "Revisiting the Digital Divide in the COVID-19 Era." Applied Economic Perspectives and Policy **43**(1): 458-464.
- López Turley, R. N. and G. Wodtke (2010). "College Residence and Academic Performance: Who Benefits From Living on Campus?" Urban Education **45**(4): 506-532.
- Mehta, D. and X. Wang (2020). "COVID-19 and digital library services – a case study of a university library." Digital Library Perspectives **36**(4): 351-363.
- Musu, L. (2018). "The Digital Divide: Differences in Home Internet Access." NCES Blog <https://nces-ed.gov.proxy.lib.umich.edu/blogs/nces/post/the-digital-divide-differences-in-home-internet-access> 2021.

- Niu, X. and B. M. Hemminger (2012). "A study of factors that affect the information-seeking behavior of academic scientists." Journal of the American Society for Information Science and Technology **63**(2): 336-353.
- Reza, F. (2020). "COVID-19 and Disparities in Education: Collective Responsibility Can Address Inequities." Knowledge Cultures **8**(3): 68-75.
- Robinson, L., S. R. Cotten, H. Ono, A. Quan-Haase, G. Mesch, W. Chen, J. Schulz, T. M. Hale and M. J. Stern (2015). "Digital inequalities and why they matter." Information, Communication & Society **18**(5): 569-582.
- Schudde, L. T. (2011). "The Causal Effect of Campus Residency on College Student Retention." Review of higher education **34**(4): 581-610.
- Simpson, D. B. and D. Burnett (2017). "Commuters Versus Residents: The Effects of Living Arrangement and Student Engagement on Academic Performance." Journal of College Student Retention: Research, Theory & Practice **21**(3): 286-304.
- StataCorp (2019). Stata Statistical Software: Release 16. College Station, TX, StataCorp LLC.
- Turk, J. M. and M. S. González Canché (2019). "On-Campus Housing's Impact on Degree Completion and Upward Transfer in the Community College Sector: A Comprehensive Quasi-Experimental Analysis." The Journal of Higher Education **90**(2): 244-271.
- Wilson, T. D. (1999). "Models in information behaviour research." Journal of Documentation **55**(3): 249-270.
- Wilson, T. D. (2017). Information Behavior Models. Encyclopedia of Library and Information Science, Fourth Edition, CRC Press.
- Zhang, X., E. J. Pérez-Stable, P. E. Bourne, E. Peprah, O. K. Duru, N. Breen, D. Berrigan, F. Wood, J. S. Jackson, D. W. S. Wong and J. Denny (2017). "Big data science: Opportunities and challenges to address minority health and health disparities in the 21st century." Ethnicity & disease **27**(2): 95-106.

## APPENDIX

Tables A.1 – A.7 show the percentages of students who had at least one EZproxy session in an academic term by various sociodemographic and academic factors.

**Table A.1.** Percentage of students associated with EZproxy sessions by first gen status, FA16 – WN19

Academic Term	Firstgen Status	Number of Students	Ever EZproxy Session	% ≥ 1 EZproxy
FA 2016	Don't Know	259	171	66%
	First Gen	3,520	2,062	59%
	Not First Gen	24,903	14,372	58%
WN 2017	Don't Know	226	139	62%
	First Gen	3,364	1,664	49%
	Not First Gen	23,818	11,631	49%
FA 2017	Don't Know	92	52	57%
	First Gen	3,753	2,054	55%
	Not First Gen	25,316	13,928	55%
WN 2018	Don't Know	85	42	49%
	First Gen	3,605	2,025	56%
	Not First Gen	24,162	12,788	53%
FA 2018	Don't Know	53	28	53%
	First Gen	4,091	2,308	56%
	Not First Gen	25,582	13,855	54%
WN 2019	Don't Know	47	32	68%
	First Gen	3,890	2,310	59%
	Not First Gen	24,418	13,957	57%

**Table A.2.** Percentage of students associated with EZproxy sessions by on-campus, FA16 – WN19

Academic Term	Residency	Number of Students	Ever EZproxy Session	% ≥ 1 EZproxy
FA 2016	Off-campus	19,130	11,554	60%
	On-campus	9,552	5,051	53%
WN 2017	Off-campus	17,971	10,353	58%
	On-campus	9,437	3,081	33%
FA 2017	Off-campus	19,993	12,049	60%
	On-campus	9,168	3,985	43%
WN 2018	Off-campus	18,793	11,043	59%
	On-campus	9,059	3,812	42%
FA 2018	Off-campus	20,357	12,014	59%
	On-campus	9,386	4,187	45%
WN 2019	Off-campus	19,110	11,765	62%
	On-campus	9,261	4,540	49%

**Table A.3.** Percentage of students associated with EZproxy sessions by gender, FA16 – WN19

<b>Academic Term</b>	<b>Gender</b>	<b>Number of Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	Female	14,296	9,510	67%
	Male	14,386	7,095	49%
WN 2017	Female	13,630	7,817	57%
	Male	13,778	5,617	41%
FA 2017	Female	14,599	9,227	63%
	Male	14,562	6,807	47%
WN 2018	Female	13,910	8,589	62%
	Male	13,942	6,266	45%
FA 2018	Female	14,833	9,304	63%
	Male	14,893	6,887	46%
WN 2019	Female	14,204	9,219	65%
	Male	14,151	7,080	50%

**Table A.4.** Percentage of students associated with EZproxy sessions by class level, FA16 – WN19

<b>Academic Term</b>	<b>Class Level</b>	<b>Number of Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	Freshman	5,665	2,982	53%
	Junior	7,035	3,979	57%
	Senior	9,361	5,920	63%
	Sophomore	6,621	3,724	56%
WN 2017	Freshman	2,727	874	32%
	Junior	6,489	3,291	51%
	Senior	11,896	6,886	58%
	Sophomore	6,296	2,383	38%
FA 2017	Freshman	5,387	2,391	44%
	Junior	7,084	3,918	55%
	Senior	9,647	6,021	62%
	Sophomore	7,043	3,704	53%
WN 2018	Freshman	2,511	1,088	43%
	Junior	6,949	3,785	54%
	Senior	11,985	7,071	59%
	Sophomore	6,407	2,911	45%
FA 2018	Freshman	5,440	2,477	46%
	Junior	7,666	4,257	56%
	Senior	9,663	5,856	61%
	Sophomore	6,957	3,601	52%
WN 2019	Freshman	2,557	1,300	51%
	Junior	7,132	4,114	58%
	Senior	12,269	7,512	61%
	Sophomore	6,397	3,373	53%

**Table A.5.** Percentage of students associated with EZproxy sessions by family income, FA16 – WN19

<b>Academic Term</b>	<b>Family Income</b>	<b>No. Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	\$25,000 - \$49,999	2,073	1,206	58%
	\$50,000 - \$74,999	2,190	1,294	59%
	\$75,000 - \$99,999	2,356	1,372	58%
	Don't Know	935	558	60%
	Less than \$25,000	1,470	896	61%
	Missing Income Information	5,412	3,023	56%
	More than \$100,000	14,246	8,256	58%
WN 2017	\$25,000 - \$49,999	1,973	953	48%
	\$50,000 - \$74,999	2,114	1,069	51%
	\$75,000 - \$99,999	2,249	1,145	51%
	Don't Know	851	435	51%
	Less than \$25,000	1,417	724	51%
	Missing Income Information	5,168	2,425	47%
	More than \$100,000	13,636	6,683	49%
FA 2017	\$25,000 - \$49,999	2,091	1,139	54%
	\$50,000 - \$74,999	2,090	1,134	54%
	\$75,000 - \$99,999	2,210	1,263	57%
	Don't Know	476	263	55%
	Less than \$25,000	1,486	855	58%
	Missing Income Information	6,472	3,631	56%
	More than \$100,000	14,336	7,749	54%
WN 2018	\$25,000 - \$49,999	2,026	1,119	55%
	\$50,000 - \$74,999	2,024	1,123	55%
	\$75,000 - \$99,999	2,080	1,167	56%
	Don't Know	430	222	52%
	Less than \$25,000	1,441	795	55%
	Missing Income Information	6,162	3,334	54%
	More than \$100,000	13,689	7,095	52%
FA 2018	\$25,000 - \$49,999	2,307	1,299	56%
	\$50,000 - \$74,999	2,066	1,150	56%
	\$75,000 - \$99,999	2,161	1,204	56%
	Don't Know	540	285	53%
	Less than \$25,000	1,586	911	57%
	Missing Income Information	6,434	3,582	56%
	More than \$100,000	14,632	7,760	53%
WN 2019	\$25,000 - \$49,999	2,212	1,269	57%
	\$50,000 - \$74,999	2,009	1,217	61%
	\$75,000 - \$99,999	2,074	1,213	58%
	Don't Know	515	278	54%
	Less than \$25,000	1,507	923	61%
	Missing Income Information	6,087	3,507	58%
	More than \$100,000	13,951	7,892	57%

**Table A.6.** Percentage of students associated with EZproxy sessions by race, FA16 – WN19

<b>Academic Term</b>	<b>Race</b>	<b>Number of Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	2 or More	1,111	642	58%
	Asian	5,460	3,019	55%
	Black	1,268	730	58%
	Hispanic	1,564	916	59%
	Not Indic	1,483	829	56%
	Other	53	30	57%
	White	17,743	10,439	59%
WN 2017	2 or More	1,084	515	48%
	Asian	5,282	2,425	46%
	Black	1,213	574	47%
	Hispanic	1,500	747	50%
	Not Indic	1,400	712	51%
	Other	53	23	43%
	White	16,876	8,438	50%
FA 2017	2 or More	1,206	631	52%
	Asian	5,685	2,941	52%
	Black	1,291	698	54%
	Hispanic	1,762	955	54%
	Not Indic	1,358	727	54%
	Other	56	29	52%
	White	17,803	10,053	56%
WN 2018	2 or More	1,155	599	52%
	Asian	5,501	2,746	50%
	Black	1,252	683	55%
	Hispanic	1,698	908	53%
	Not Indic	1,268	673	53%
	Other	54	26	48%
	White	16,924	9,220	54%
FA 2018	2 or More	1,346	702	52%
	Asian	6,047	3,063	51%
	Black	1,315	748	57%
	Hispanic	1,972	1,051	53%
	Not Indic	1,472	810	55%
	Other	49	23	47%
	White	17,525	9,794	56%
WN 2019	2 or More	1,302	745	57%
	Asian	5,829	3,137	54%
	Black	1,268	766	60%
	Hispanic	1,899	1,099	58%
	Not Indic	1,407	792	56%
	Other	46	22	48%
	White	16,604	9,738	59%

**Table A.7.** Percentage of students associated with EZproxy sessions by school, FA16 – WN19

<b>Acad. Term</b>	<b>School</b>	<b>No. Students</b>	<b>Ever EZproxy Session</b>	<b>% ≥ 1 EZproxy</b>
FA 2016	Undergrad Music, Thtre & Dance	732	447	61%
	Undergraduate Architecture	145	65	45%
	Undergraduate Art and Design	495	356	72%
	Undergraduate Business Admin	1,673	890	53%
	Undergraduate Dental Hygiene	111	77	69%
	Undergraduate Education	112	66	59%
	Undergraduate Engineering	6,078	2,736	45%
	Undergraduate Information	208	123	59%
	Undergraduate Joined Deg Prog	10	7	70%
	Undergraduate Kinesiology	946	698	74%
	Undergraduate L S & A	17,306	10,395	60%
	Undergraduate Nursing	705	626	89%
	Undergraduate Pharmacy	14	11	79%
	Undergraduate Public Policy	147	108	73%
WN 2017	Undergrad Music, Thtre & Dance	700	402	57%
	Undergraduate Architecture	140	71	51%
	Undergraduate Art and Design	462	249	54%
	Undergraduate Business Admin	1,639	746	46%
	Undergraduate Dental Hygiene	107	63	59%
	Undergraduate Education	112	53	47%
	Undergraduate Engineering	5,909	1,958	33%
	Undergraduate Information	186	89	48%
	Undergraduate Joined Deg Prog	8	5	63%
	Undergraduate Kinesiology	918	576	63%
	Undergraduate L S & A	16,400	8,614	53%
	Undergraduate Nursing	685	512	75%
	Undergraduate Pharmacy	14	7	50%
	Undergraduate Public Policy	128	89	70%
FA 2017	Undergrad Music, Thtre & Dance	747	495	66%
	Undergraduate Architecture	155	82	53%
	Undergraduate Art and Design	497	363	73%
	Undergraduate Business Admin	1,773	869	49%
	Undergraduate Dental Hygiene	112	79	71%
	Undergraduate Education	120	46	38%
	Undergraduate Engineering	6,409	2,666	42%
	Undergraduate Information	253	147	58%
	Undergraduate Joined Deg Prog	12	8	67%
	Undergraduate Kinesiology	976	627	64%
	Undergraduate L S & A	17,160	9,942	58%
	Undergraduate Nursing	667	516	77%
	Undergraduate Pharmacy	42	19	45%
	Undergraduate Public Health	85	72	85%
Undergraduate Public Policy	153	103	67%	
WN 2018	Undergrad Music, Thtre & Dance	715	509	71%
	Undergraduate Architecture	153	107	70%

	Undergraduate Art and Design	481	356	74%
	Undergraduate Business Admin	1,757	760	43%
	Undergraduate Dental Hygiene	109	82	75%
	Undergraduate Education	118	40	34%
	Undergraduate Engineering	6,150	2,571	42%
	Undergraduate Information	214	122	57%
	Undergraduate Joined Deg Prog	12	7	58%
	Undergraduate Kinesiology	951	594	62%
	Undergraduate L S & A	16,294	9,034	55%
	Undergraduate Nursing	636	487	77%
	Undergraduate Pharmacy	42	25	60%
	Undergraduate Public Health	84	72	86%
	Undergraduate Public Policy	136	89	65%
FA 2018	Undergrad Music, Thtre & Dance	743	524	71%
	Undergraduate Architecture	181	119	66%
	Undergraduate Art and Design	556	396	71%
	Undergraduate Business Admin	1,826	753	41%
	Undergraduate Dental Hygiene	103	71	69%
	Undergraduate Education	131	60	46%
	Undergraduate Engineering	6,649	2,755	41%
	Undergraduate Information	302	135	45%
	Undergraduate Joined Deg Prog	11	9	82%
	Undergraduate Kinesiology	962	617	64%
	Undergraduate L S & A	17,262	9,918	57%
	Undergraduate Nursing	632	543	86%
	Undergraduate Pharmacy	56	33	59%
	Undergraduate Public Health	158	130	82%
	Undergraduate Public Policy	154	128	83%
WN 2019	Undergrad Music, Thtre & Dance	717	515	72%
	Undergraduate Architecture	181	124	69%
	Undergraduate Art and Design	524	381	73%
	Undergraduate Business Admin	1,799	740	41%
	Undergraduate Dental Hygiene	101	70	69%
	Undergraduate Education	126	54	43%
	Undergraduate Engineering	6,313	2,847	45%
	Undergraduate Information	260	122	47%
	Undergraduate Joined Deg Prog	10	7	70%
	Undergraduate Kinesiology	954	678	71%
	Undergraduate L S & A	16,409	10,030	61%
	Undergraduate Nursing	607	475	78%
	Undergraduate Pharmacy	55	36	65%
	Undergraduate Public Health	157	116	74%
	Undergraduate Public Policy	142	104	73%