Fall 9-14-2020

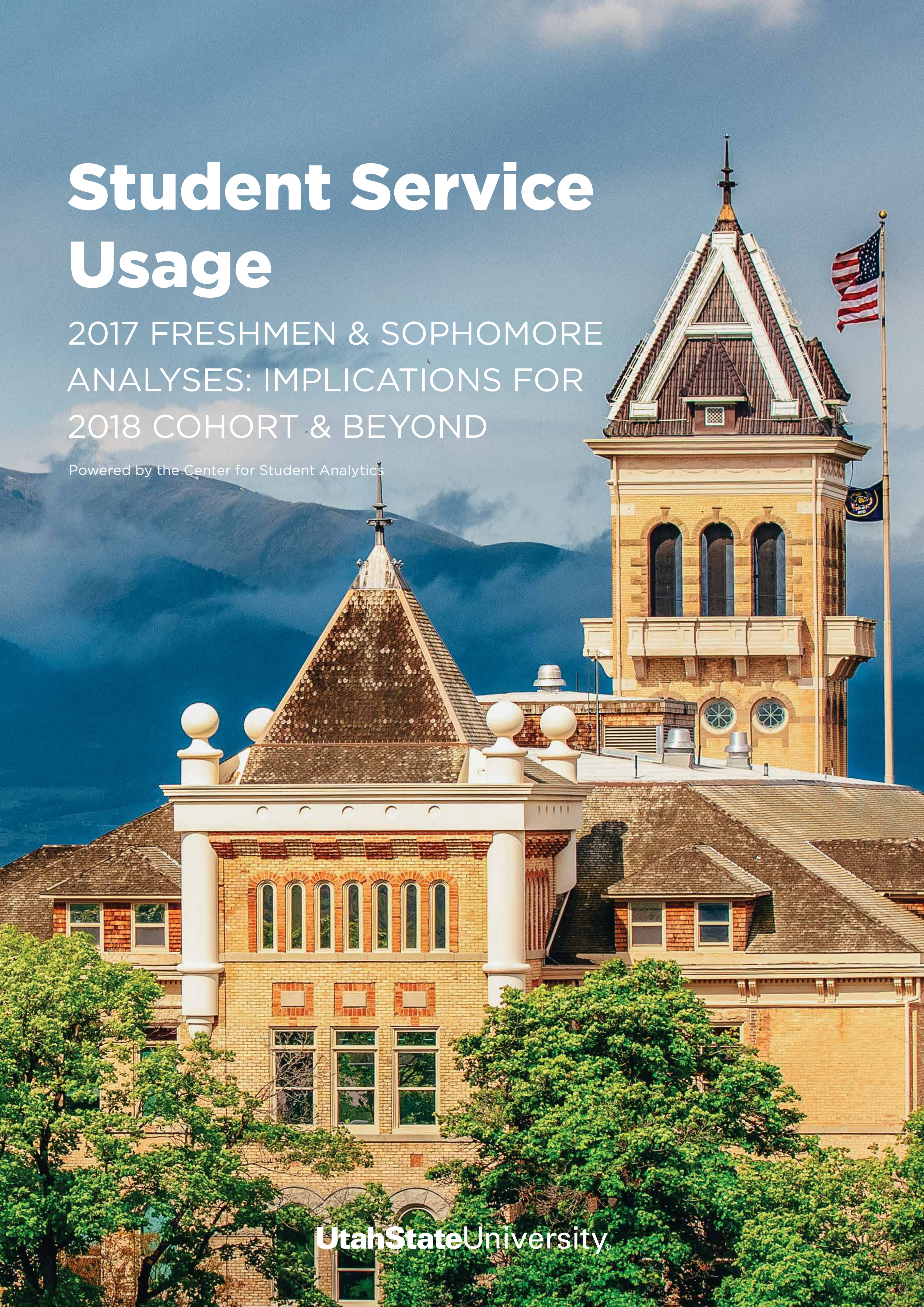
How do USU Students Use University Support Services?

Amanda M. Hagman
Utah State University, amanda.hagman@usu.edu

Meghan Lewis
Utah State University, Meghan.lewis@usu.edu

Kristi Swainston
Utah State University, Kristi.swainston@usu.edu

Mitchell Colver
mitchell.colver@usu.edu



Student Service Usage

2017 FRESHMEN & SOPHOMORE ANALYSES: IMPLICATIONS FOR 2018 COHORT & BEYOND

Powered by the Center for Student Analytics

UtahStateUniversity

How do USU students use University support services?

Amanda Hagman

Data Scientist, M.S.

Center for Student Analytics

Meghan Lewis

Data Architect & Analyst

Center for Student Analytics

Kristi Swainston

Data Programmer & Analyst

Registrar's Office

Mitchell Colver

Manager & Analyst, Ph,D.

Center for Student Analytics

KEY INSIGHTS

- Students use services different across their academic career
- Freshmen and sophomores use services at a higher rate than juniors, seniors, and graduate students
- Student performance and academic engagement are among the most important variables in predicting persistence
- Student engagement in student facing programming was predictive of retention
- Meeting with an academic advisor emerged as an important indicator of persistence

UTAH STATE UNIVERSITY'S MISSION

The central mission of USU is to be one of the nation's premier student-centered universities. This is accomplished by fostering academics and offering opportunities to expand students' vision of the world around them through co- and extra-curricular activities. These auxiliary activities are designed to support academic achievement and to engage students in meaningful opportunities to practice and enhance their personal and professional well-being. This analysis describes how students are using co- and extra-curricular services. It investigates the most salient student-facing programming in supporting student retention.

Aggies in the Analyses

Student co-curricular and extra-curricular activities vary across campuses. Students and programming from the Logan main campus are considered in this report. Additional reports can be generate for USU Regional and USU-Eastern campuses that cover student co-curricular and extra-curricular programming and participation at their specific campuses, but is beyond the scope of this project. Non-degree seeking students were excluded from this report.

AGGIE DEMOGRAPHICS

The majority of Logan main campus students are traditional, continuing generation, full-time student in good academic standing. In 201740, there were 14,575 students attended classes on the Logan campus. 23.7% of students are freshmen, 20.9% sophomores, 19.4% are juniors, 25.7% are seniors, and 10.2% are graduate student.

AGGIE PARTICIPATION

On average students participate in 3 (min = 0, max = 13) student services. In 2017 there were 1,999 students that didn't participate in any of the considered activities and there were 223 that participated in all the activities.

A closer look at the data shows variation in participation across students' academic career. Of the students who used all the services, nearly all (86.5%) were freshmen and the rest (13.5%) were sophomores. Non-participation is seen mostly in upper-classmen and graduate students, freshmen only account for 6.5% and sophomores for 12.6% of non-participation. Table 1 shows the differences in usage by class rank across the considered student support services.

Student Service Usage

TABLE 1:

Student service usage for Fall 2017 by class rank.

	Freshman	Sophomore	Junior	Senior	Graduate
Advising	56.2%	58.8%	57.8%	56.1%	0.5%
Math Tutoring	3.9%	4.8%	3.1%	1.6%	0.1%
Stats Tutoring	0.9%	1.7%	1.2%	0.7%	0.0%
Writing Center	33.1%	11.8%	4.6%	3.0%	2.4%
SI	33.8%	19.9%	12.2%	4.1%	0.0%
Connections	42.9%	6.9%	0.7%	0.0%	0.0%
Orientations	45.2%	13.0%	6.3%	1.3%	0.0%
Parent Orientation	31.8%	6.0%	0.8%	0.0%	0.0%
On-Campus Living	47.0%	18.2%	11.1%	9.3%	10.2%
Athletic Events	71.7%	59.4%	53.5%	40.4%	15.1%
USUSA	47.2%	31.1%	22.1%	12.6%	3.5%
Rec Facilities	71.0%	63.9%	62.5%	53.3%	43.9%
Passport	57.8%	15.2%	6.4%	1.3%	4.2%

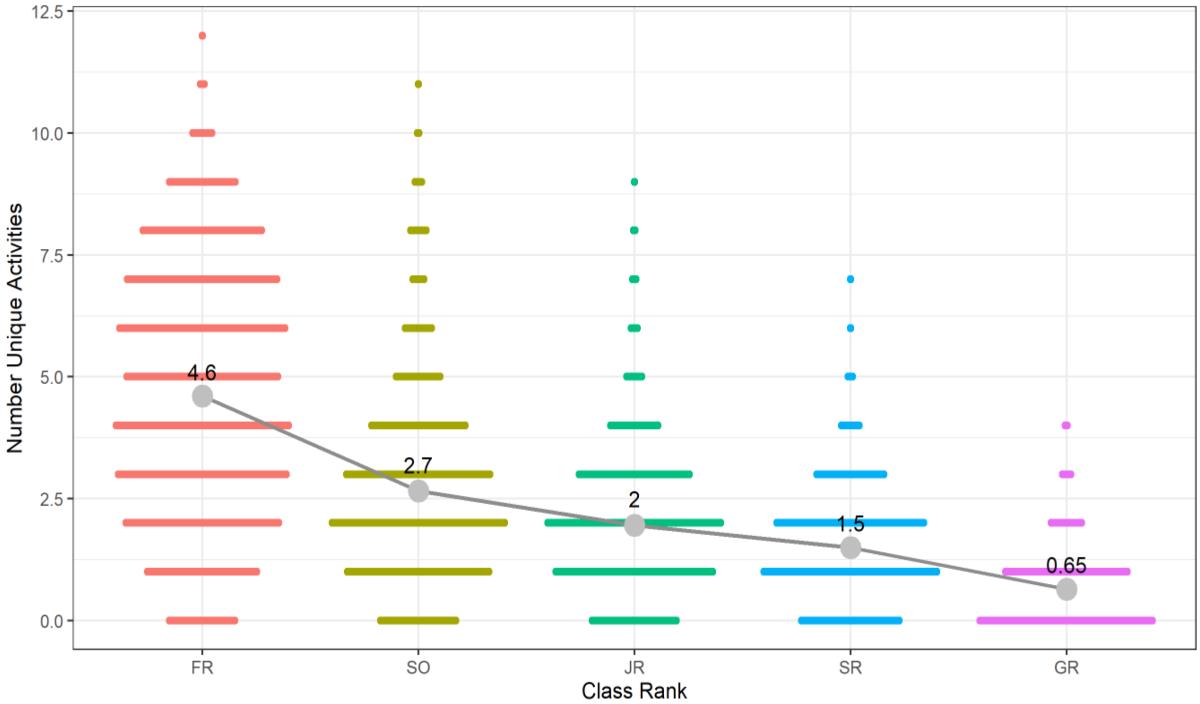


FIGURE 1
Number of unique student services used during Fall 2017 by class rank

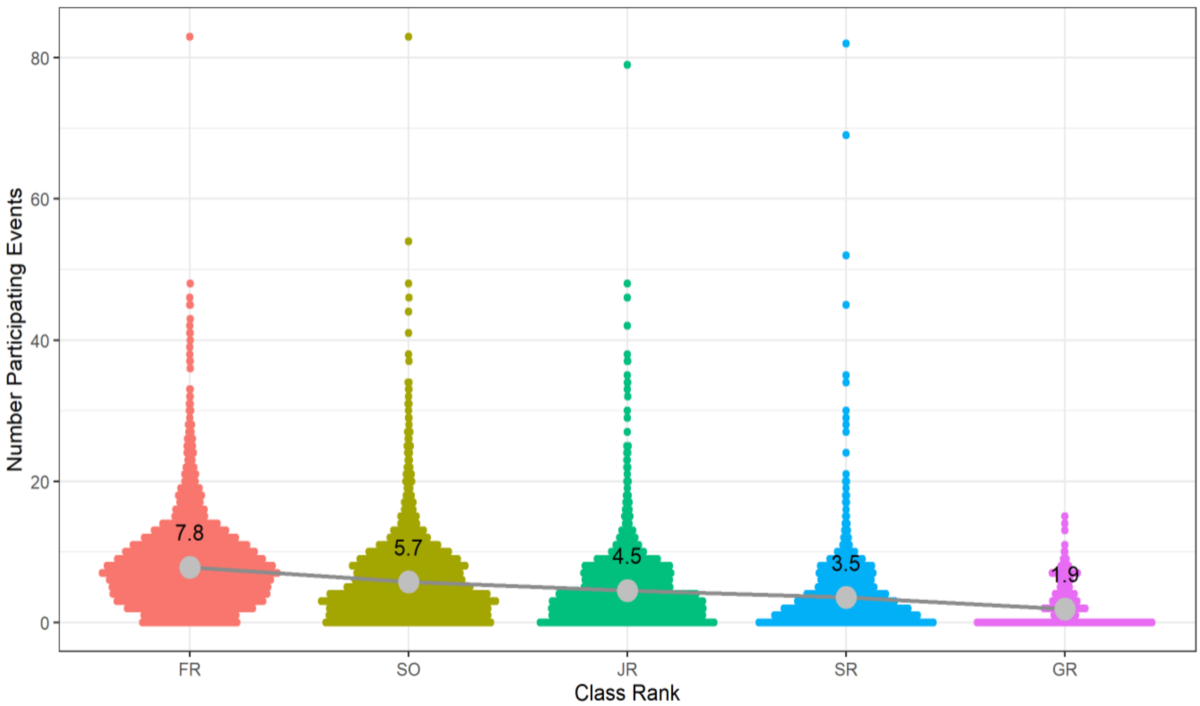


FIGURE 2
Number of unique student participation events in student services used during Fall 2017 by class rank

STUDENT SERVICES & CLASS RANK

Of the considered co- and extra-curricular activities there is differences in the number of services accessed by class rank, see Figure 1. This is partially understood in the targeted use of student services towards underclassmen. Connections and orientation are freshmen services. Passport is marketed to freshmen.

Supplemental instruction opportunities are available mostly in freshmen and sophomore level courses. Comparing trends only between juniors, seniors, and graduate students, there is a clear decline in participation in the considered co- and extra-curriculars as students advance in their academic careers from 2 for juniors to 1.5 for seniors to 0.65 for graduate students.

There is a similar trend towards less participation when number of participation events are considered across the academic career (Figure2). Since orientation and Connections account for a single "participation event", Figure 2 may act as a better reflection of student co- and extra-curricular activities across the academic career. Freshmen account for the majority of participation events, followed by sophomores, then juniors, and so on. During fall 2017, freshmen had an average number of participations of 7.9. The level of participation drops to 3.5 by the time students are seniors, 1.9 for graduate students.

This downward trend in service use aligns with availability and marketing of student co- and extra-curricular activities considered in this report. As such, it is likely that student services may have a stronger impact on student retention for students earlier in their academic careers, compared with upper-classmen. For this reason, the following methodologies and analyses contain only freshmen and sophomores.

Methodology

POPULATION

Traditional students, defined as FG or NF admits who are under the age of 25 at admission, are the target population to understand the impact of student co- and extra-curricular activities on student retention. Table 2 contains demographics of freshmen and sophomore cohorts.

Gini index data from the national census was also included to investigate the impact of community characteristics on college success. The following community level data was added to student level data with high school zip code as the cross-walk:

- Accessibility to internet
- Percent community education attainment
- Percent community race and ethnicity
- Percent community age and gender

Student Demographics

TABLE 2:

Student demographics by class rank for 2017 fall students

	Freshman (3,458)	**Sophomore (3,046)**
Age	19.1 (1.5)	20.7 (2.5)
Gender: Female	52.0%	50.7%
Race: White	91.3%	91.4%
GPA	2.6 (1.3)	2.7 (1.3)
First Generation	16%	13.7%
Time-Status		
Half-Time	4.1%	5.4%
3/4-Time	7.3%	8.5%
Full-Time	88.5%	86.0%

CO- AND EXTRA-CURRICULAR ACTIVITIES

Student participation data is collected by various organizations across campus. Starting in fall 2017, cross-talk between participation data sets became more feasible as university-wide data analytic initiatives were implemented. Data from card swipe, tutoring, advising, etc. repositories found a home at the Center for Student Analytics.

For this analysis student participation is viewed by term, accounting for fall 2017 and spring 2018 in separate variables. Data was collected across the following co- and extra-curricular activities:

Co-Curriculars

- Advising
- Connections
- Orientation Modules
- Parent Orientation
- On-campus Living
- Writing Center
- Math & Statistics Tutoring
- Service Learning Courses
- Supplemental Instructions

Extra-curricular

- Athletics
- USUSA

- Recreational Facilities
- Student Involvement and Leadership
- Passport

The data from these activities can be view in two manners. First, students who have any record of participation in activity are considered participants. This is accurate for activities that only have 1 possible level of participation; Connections, orientation modules, parent orientation, on-campus living fall into this category. Other activities can also be viewed in this same manor, all levels of participation are viewed equally. The second view of participation by level. Advising, writing center, tutoring, supplemental instructions, athletics, etc. all have variable levels of participation. It is likely that level of participation may influence retention differently.

EXTREME VALUES IN CO- & EXTRA-CURRICULAR ACTIVITIES

Some activities can be used in a manner that skews the data significantly. For example, some students will exercise in a rec facility, return later to play basketball, return again to shower, each day. These students could have participation levels around 240 unique semesterly visits. To account for this reality student participation in continuous activities is binned and capped in the following ways:

Recreational Facilities, Math & Stat Tutoring: There are 7 participation levels for rec facilities, math and stat tutoring. Data is binned by 5 visits, ie. students with 1-5 visits received a value of 1. Individuals with more than 30 visits received a score of 7.

Writing Center Visits: Over 90% of students who used the writing center had fewer than 10 visits. The maximum visits in the 201740 semester was 52 visits during the semester. To reduce this variability, writing center visits were capped at 10 visits.

STUDENT CANVAS USAGE

Student number of days with Canvas access indicates how often students log into their Canvas account during the semester. In fall 2017 there were minimum of 6 and a maximum of 110. In spring 2018 there is a minimum of 0 and maximum of 119. Both semesters are negatively skewed.

RETENTION TO THE FOLLOWING FALL SEMESTER

Student retention is used as the outcome variable for the analysis. Retention is measured using data from USU and the National Clearninghouse data. Triangulating between the USU and National Clearninghouse data enabled the data to be classified as retained (appeared in USU data), excused leave of absence (LOAtransferred (students appearing at another university in the National Clearninghouse), and dropped (not appearing in the National Clearinghouse). Students with an LOA were classified for the analysis as retained because they show intention to return. This simplification left a final classification of retained, transferred, and dropped.

Student Retention

TABLE 3:
Student demographics by class rank for 2017 fall students

	Freshman (3,458)	Sophomore (3,046)
Retained	1,519	2,700
LOA	398	91
Transfer	150	242
Dropped	291	904

PERSISTENCE TO SPRING 2018

Persistence to spring 2018 was not considered in this analysis. Transferring and dropping out were rare event which produce model with poor classification. Future analysis can investigate persistence as more historical and accurate student participation data becomes available.

Analytic Approach

PREDICTIVE MODELING & VARIABLE SELECTION

Random forest is a machine learning technique that can be used to predict and classify outcomes. This technique is capable of handling large predictor variable sets and accounts for correlations and interactions between variables. Random forest produce (1) a predictive model and (2) a list of important variables within the prediction. In this analysis random forest was used to predict classification of "retained" or "dropped" students and to produce variables important this classification.

The data was divided into training (70% of the sample) and testing data (30% of the data). The random forest included student demographics, canvas activity, home-community census variables, and participation data. The random forest was trained on the training data set and then applied to the testing data set.

The quality of a random forest is measured by sensitivity (true positive rate; i.e. correct classification of dropped students) and specificity (true negative rate; i.e. correct classification of retained students), with favor given to high sensitivity.

The model was applied to the sophomore 201740 cohort to see how well the model held true for sophomores. Finally the model developed for Fall 2017 freshmen was applied to Fall 2018 freshmen to anticipate Fall 2019 freshmen retention.

VARIABLE INTERPRETATION

Random forests are power predictors for large datasets; however, they are not concerned with the interpretation of variables within the model. In other words, the prediction is the intended outcome, not interpretation of variables within the model. To have a better understanding of how variables are associated with each other and the outcome, variables identified as important from the random forest were integrated into a logistic regression for variable interpretation. The logistic regression was conducted in R using the package glm2. Model significance is measured through a chi^2 test, p-values less than 0.05 for variables within the model, and McFadden's pseudo R^2 to estimate the effect size.

Results

PREDICTIVE MODEL AND VARIABLE SELECTION

201740 Freshmen Cohort. The random forest had a low sensitivity for classifying dropped students, 55.8%. The random forest had a similar classification capacity as chance for dropped students. The model has good specificity at 3.9% due mostly to correct classification of retained students. Over all the model had a balanced accuracy of 69.16%.

The variables identified as most important are listed below:

- End of term GPA
- Days of Canvas use Spring 2018
- Meeting with an advisor in Spring 2018
- Number of visits with an advisor Spring 2018
- End of term earned hours
- High school GPA
- Registered hours
- Race: Euro-American
- Number recreational facility use Spring 2018
- Days of Canvas use Fall 2017
- Number of athletic events Fall 2017
- Home-community proportion of high school or more graduates
- Living on campus
- Gini Index
- Home-community proportion between 20 and 34 years old
- Home-community proportion under 20 years old
- Home-community proportion earning $100,000 or more

2ND-YEAR STUDENTS FALL 2017

The same algorithm was used on second-year students from Fall 2017, the model sensitivity was similar. Correct classified for dropped students was near chance, 54.3%.

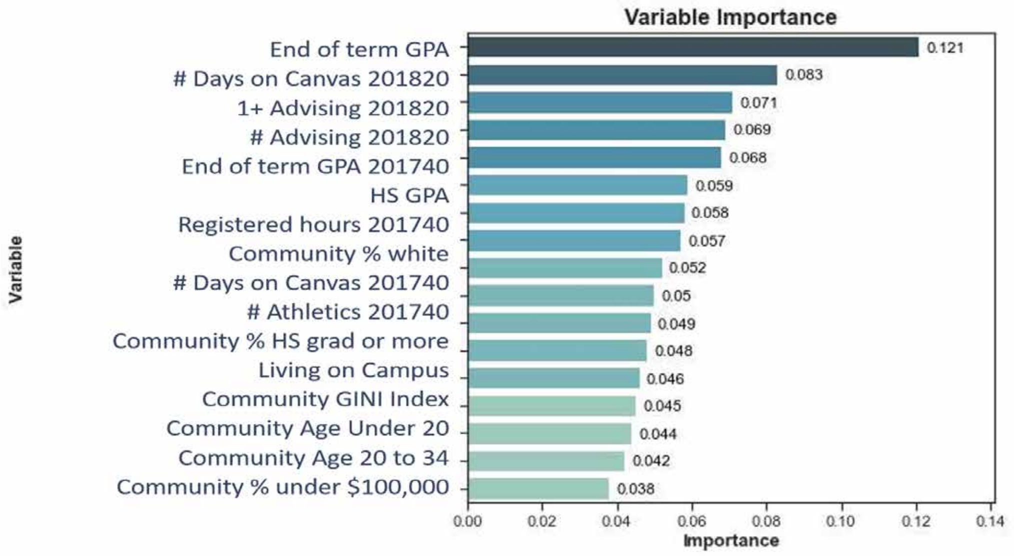


FIGURE 3
Variable importance obtained from the random forest predicting students who drop out of university after their first year.

VARIABLE SELECTION FOR LOGISTIC REGRESSION

Model specificity was low with the random forest, instead of using it for prediction, these models informed a logistic regression. Given that specificity was near chance, variables from the model will only be used as a starting point for the logistic regression. Home-community census variables isolated from the random forest were entered into the logic regression along with student demographic, canvas use, and participation data from the overall dataset. Several variables had high correlation. The most appropriate variable was selected for include in the analysis, while the other was dropped. For example, number of gym visits is highly correlated with any gym participation. While a random forest can hand correlations logistic regressions cannot. Number of gym visits was incorporated into the model, any participation was exclude.

Interactions between high school GPA and canvas activity on overall student involvement, connections, and living on campus.

The logistic regression predicted freshmen who dropped out of college before Fall 2018. Dropped was defined as students who were not registered at USU in Fall 2018 and who were not found through the National Clearninghouse data at another university or who did not have an excused LOA.

The large model that included all considered variables yielded several variable sets that were not statistically significant at the 0.05 level. Non-significant variables were removed stepwise, removing variables with the least impactful estimates first until all remaining variables contributed significantly to the model. The final model can be seen in Table 4.

MODEL FIT STATISTICS

The logistic regression equation accounted for an estimated 25% of the variance in the model using McFadden's psuedo-R2. The area under the curve on the ROC was .75. While the model still contains considerable variability, it has good classification capabilities. Model sensitivity is 63.4% and specificity is 90.2%.

Classification Matrix

TABLE 4:
Actual and predicted classification of freshmen fall 2017 students

	Predicted Retained	Predicted Dropped
Actual Retained	1,865	201
Actual Dropped	52	90

SENSITIVITY:

True positives / (True positives + False Negatives)

SPECIFICITY:

True Negatives / (True negatives + False Positives)

Model Predicting Drop-Out

TABLE 4:

Freshmen who drop after the first year from college were predicted through logistic regression

	B	SE	Wald-Z	p-value	OR	2.5%	97.5%
# Days Canvas 201820	-0.02	<0.01	-6.19	<0.001	0.98	-0.02	-0.01
1+ Advisor Meeting 201820	-0.65	0.21	-3.11	<0.001	0.52	-1.07	-0.25
End of Term GPA	-1.06	0.12	-8.99	<0.001	0.34	-7.30	0.83
# Services Used 201740	-0.17	0.04	-4.04	0.04	0.84	-0.26	-0.09
# Passport Activities	0.39	0.21	1.87	0.06	0.44	-0.06	0.713
Connections	2.91	1.57	1.85	0.06	18.44	-0.05	0.79

VARIABLE INTERPRETATION FOR LOGISTIC REGRESSION PREDICTED STUDENT DROP OUT

Using the Wald z-score to interpret standardized variable importance found end of term GPA, number of days on Canvas, unique services used, and meeting with an advisor to be the most salient variables in predicting drop out. For every 1 day increase in **Canvas** use during a semester there is a 2% decrease in a students likelihood to drop out. For every 1 letter grade increase in **GPA**, students are 66% less likely to drop out. As students increase their use of unique **university services** there is a 16% decrease in a students likelihood to drop out of university. And, students who meet with an **advisor** are half as likely to drop out as students who do not meet with their advisor.

Connections and **passport** were left in the model because they increase the predictability of the drop-out; however, they only approach statistical significance and should be interpreted with caution. Students who attend connections or use passport appear to be more likely to drop out than freshmen who do not take Connections or who do not use the passport program.

Figure 4 shows the rank order of predicted probability scores by the predicted probability of dropping out. Around -0.75 predicted probability there is a shift toward dropping out.

MODEL TEST WITH SOPHOMORES

Running 201740 sophomores through the same equation yielded good classification for dropout. Misclassfication error was 16.2%, sensitivity was 72.8%, and specificity was 85.0%. Area under the curve was 0.76. Figure 5 depicts the rank order of predicted probability scores by the predicted probability for students who dropped or who are retained. Around -0.75 predicted probability there is a shift toward dropping out.

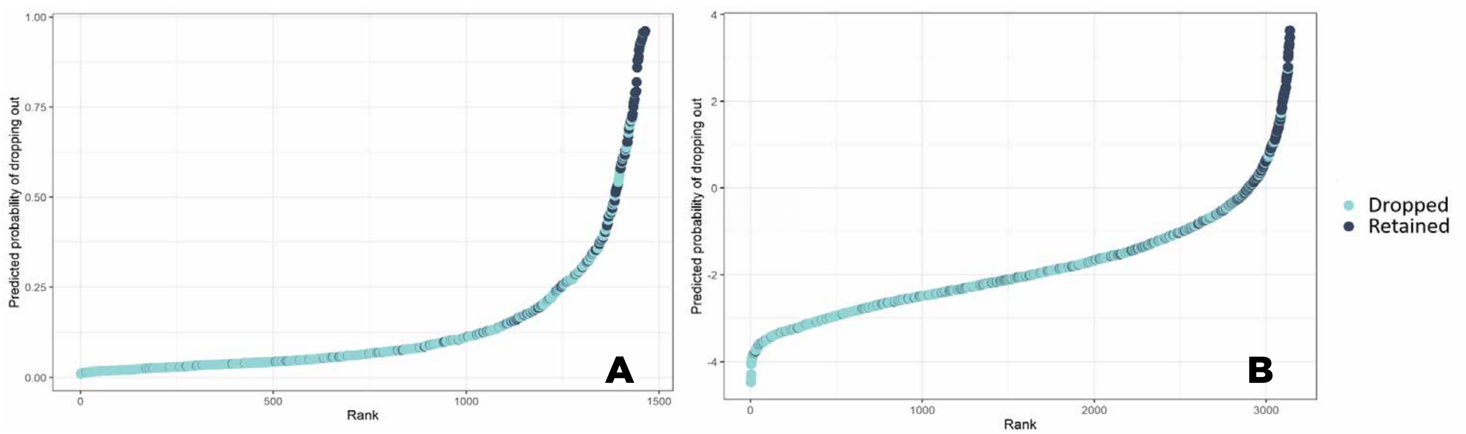
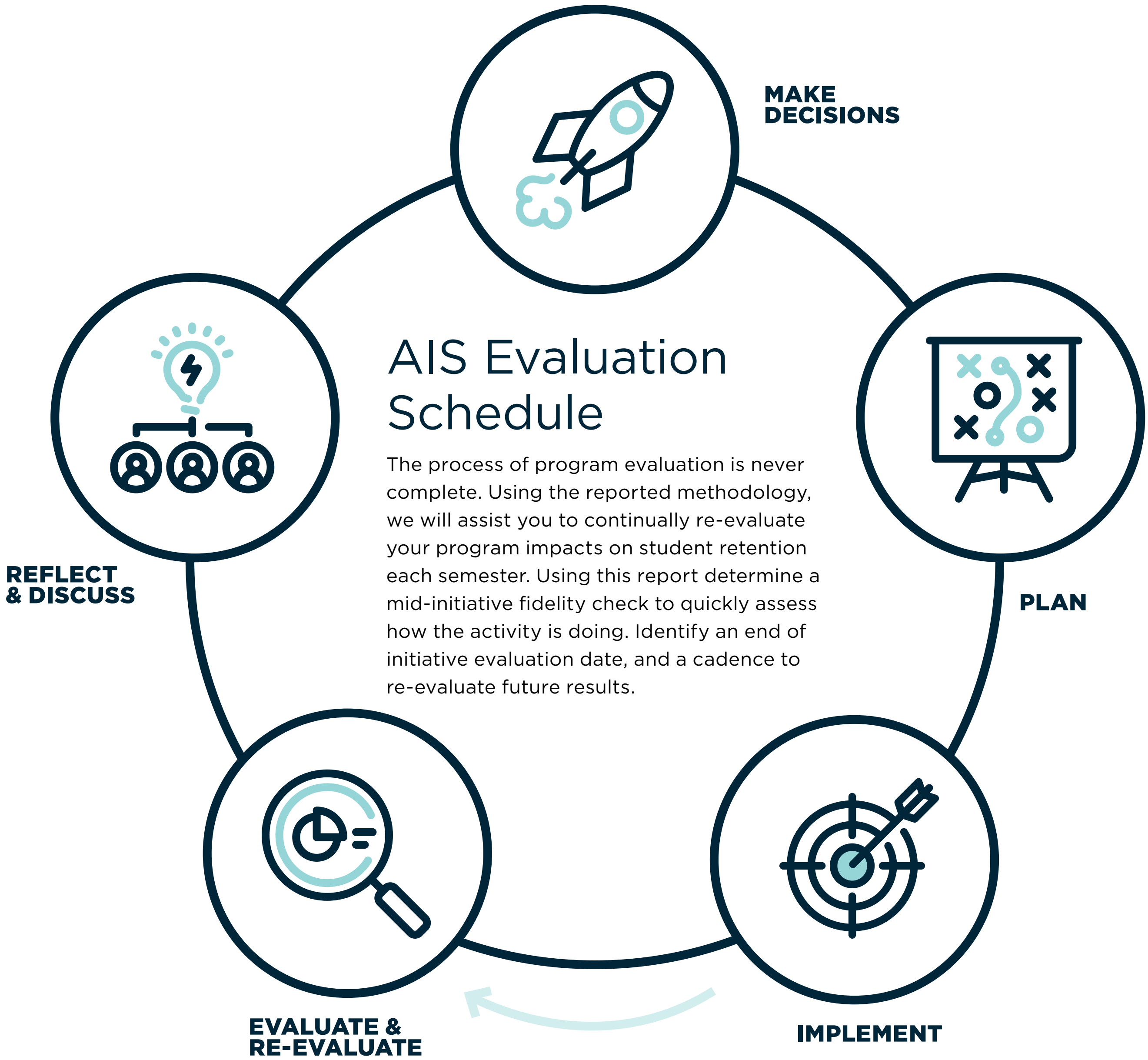


FIGURE 4

Evaluation Schedule

Next Review Date: ______________________________

Midterm Accreditation Check: ______________________



AIS Evaluation Schedule

The process of program evaluation is never complete. Using the reported methodology, we will assist you to continually re-evaluate your program impacts on student retention each semester. Using this report determine a mid-initiative fidelity check to quickly assess how the activity is doing. Identify an end of initiative evaluation date, and a cadence to re-evaluate future results.

EVALUATE & RE-EVALUATE

Get the data to AIS and we can run an evaluation on persistence. For goals that don't include persistence AIS can assist you in finding resources to measure your improvement.

REFLECT & DISCUSS

Consider the report and the evaluators insights to produce discussion within your department.

MAKE DECISIONS

Formulate possible actions to improve your program. Select actions that align with your program goals.

PLAN

Make concrete plans to apply your decisions. Determine the who, where, and when of your actions.

IMPLEMENT

Put your plans into actions. Remember to periodically check the progress of your plans as they are being implemented.