# Investigating Institutional Characteristics Associated with Campus Sexual Assault

Lindsay Saxman
Computational and Data Sciences
George Mason University
Fairfax, Virginia, U.S.
lsaxman@gmu.edu

***Abstract*—Campus sexual assault (CSA) is an incredibly widespread issue on college campuses, but it is often not acknowledged or reported. To ensure safety on campus, college students and employees must have knowledge about campus sexual violence and become familiar with educational and legal resources regarding campus sexual assault. The goal of this project is to help victims, advocates, and researchers understand the current state of sexual violence on college campuses. This project aims to investigate factors surrounding campus sexual assault using exploratory data analysis and visualization. Methodology includes feature engineering, descriptive statistics, statistical hypothesis testing, and data visualization. The scope of this project is limited to exploring the relationships between institutional characteristics, academic achievement, and CSA among United States colleges during 2021. Determining causation and modeling relationships falls outside the scope of this project. Results show an association between institutional characteristics, such as number of NCAA-affiliated sports teams, and CSA, as well as an association between CSA offenses and academic achievement. This project contributes to the body of knowledge of CSA by proving the association between institutional characteristics, academic achievement, and campus sexual assault, although causation was not established. There is still a lack of research, particularly regarding violence against historically marginalized groups, but it is hoped that this project provides valuable insights and reinforces the urgency to address sexual violence on college campuses in the United States.**

## I. Introduction

In the shadows of ivy-covered walls and amid the echoing halls of academia, a sobering reality persists: the prevalence of campus sexual assault (CSA) in the United States. Unseen by some, yet felt by many, this issue cuts through the sanctity of higher education, casting a distressing shadow on the college experience. In recent years, the conversation surrounding sexual assault on campuses has intensified, drawing attention to the urgent need for comprehensive understanding and actionable insights. Through the lens of data science, the goal of this project is not only to quantify but also to comprehend the realities faced by college students, ultimately striving for a safer, more informed collegiate environment.

This paper aims to unveil the intricate relationship between institutional characteristics and the occurrence of campus sexual assault in the United States. By delving into this data-driven exploration, this paper aims to shed light on the prevalence, trends, and contextual factors surrounding this critical issue, to decipher the factors that might contribute to or mitigate the prevalence of CSA, and to illuminate potential pathways toward prevention and support. As we navigate this investigation, the pursuit is not solely confined to uncovering statistical associations but rather to catalyze a broader conversation—one that redefines the role of institutions in fostering safer, more nurturing environments for all students. This project aims to provide actionable insights that catalyze meaningful change and safeguard the sanctity of higher education for generations to come.

## II. Methods

### A. Data

Data were extracted from three sources, all originating from the U.S. Department of Education (DOE). The DOE Campus Safety and Security (CSS) Data Analysis Cutting Tool was used to retrieve data for CSA offenses. Data for institutional characteristics was extracted from the Integrated Postsecondary Education Data System (IPEDS). The DOE's College Scorecard dataset was downloaded to obtain additional institutional characteristic data.

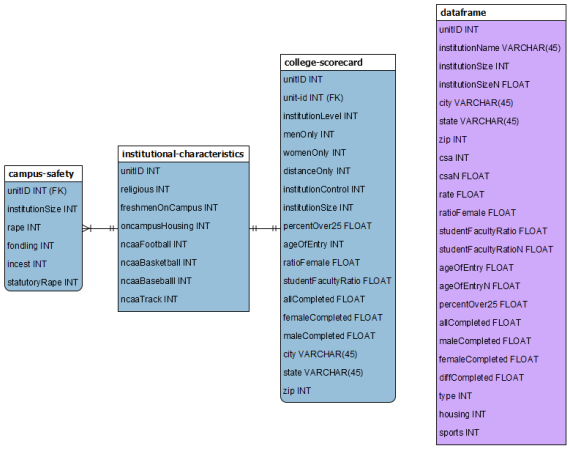


Fig. 1. Entity-relationship diagram of data.

Original CSA data from the CSS Data Analysis Cutting Tool were separated by offense type: rape, incest, fondling, and statutory rape. These four columns were added together to obtain the total number of CSA offenses per institution in 2021, regardless of offense type. In addition, the rate of CSA offenses for every 1,000 students was computed using (1). The average number of CSA offenses per college is 5.38, and the average rate of CSA offenses for every 1,000 students is 1.15.

$$\frac{CSA * 1{,}000}{Institution\ size} \tag{1}$$

| | Mean | Std. deviation | Median | Min. | Max. |
|---|---|---|---|---|---|
| CSA | 5.38 | 20.22 | 1 | 0 | 676 |
| CSA rate | 1.15 | 1.91 | 0.36 | 0 | 19.18 |

IPEDS data included binary institutional characteristics: institution type, religious affiliation, on-campus housing, freshmen housing requirement, and NCAA sports teams affiliations. The institution type and religious affiliation variables were transformed into one column with three levels: public, private (non-religious), and religious. As the pie chart shows, each institution type accounts for roughly one-third of the data, and there are more public institutions than private and religiously affiliated institutions.

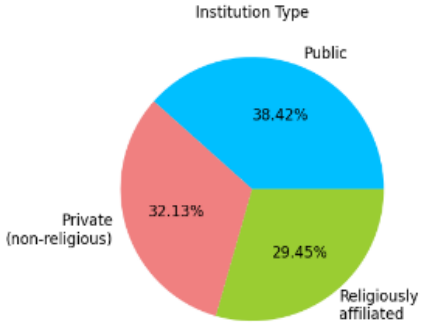


Fig. 2. Distribution of institution type.

| | Count |
|---|---|
| Public | 698 |
| Private (non-religious) | 552 |
| Religious | 542 |

The on-campus housing variable indicates whether or not an institution has on-campus housing, and the freshmen housing requirement variable indicates whether first year (i.e., freshmen) students are required to live on-campus. These variables were transformed into one with three levels: no on-campus housing, freshmen not required to live on-campus, and freshmen required to live on-campus. As the pie chart shows, over three-quarters of institutions provide on-campus housing, and a very small portion of institutions require freshmen students to live on-campus.

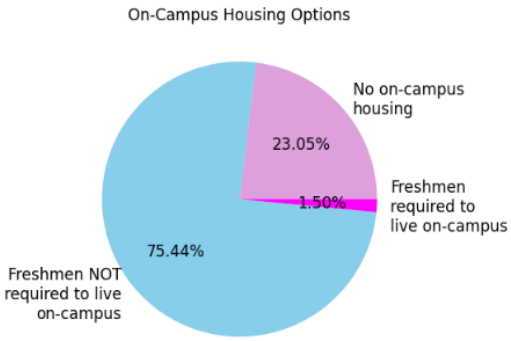


Fig. 3. Distribution of housing type.

| | Count |
|---|---|
| No on-campus housing | 371 |
| Freshmen not required to live on-campus | 1393 |
| Freshmen required to live on-campus | 28 |

There are four variables in the IPEDS dataset indicating whether each institution has an NCAA affiliated sports team for football, basketball, baseball, and track/cross country. Only two of these sports were included in the analyses: football and basketball. As the stacked bar chart shows, more institutions have basketball teams than do not, but less institutions have football teams than do not.

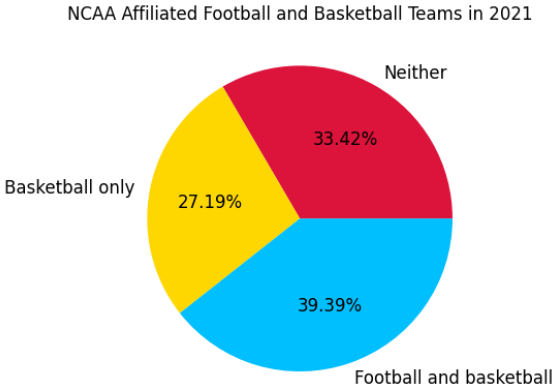


Fig. 4. Distribution of sports team type.

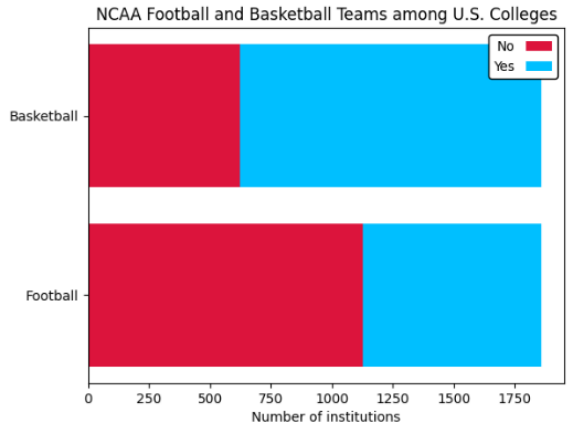


Fig. 5. Distribution of basketball and football teams.

These two variables were transformed into one with three levels: each institution has neither football nor basketball team, only basketball (and no football team), or both football and basketball teams. As the pie chart shows, each level of the sports variable accounts for roughly one-third of the dataset, with more institutions having both football and basketball teams.

| | **Count** |
|---|---|
| Neither team | 557 |
| Basketball team only | 502 |
| Both basketball and football teams | 733 |

The College Scorecard dataset included a categorical indicating institution level and binary columns indicating men-only, women-only, and distance learning institutions. These columns were used to filter out men-only and women-only colleges, colleges that exclusively offer distance education, and less than 4-year colleges. Numerical institutional characteristics were extracted from the College Scorecard dataset. These characteristics include institution size (i.e., total enrollment), ratio of female students (i.e., the number of female students for every 1 enrolled student), average age of entry (in years), percent of students over 25 years old, and student-to-faculty ratio. Academic achievement parameters (i.e., graduation rates) were also obtained from the College Scorecard dataset.

| | **Mean** | **Std. deviation** | **Median** | **Min.** | **Max.** |
|---|---|---|---|---|---|
| Institution size | 5,112 | 8,341 | 2,034 | 2 | 119,248 |
| Female ratio | 0.6 | 0.12 | 0.6 | 0.05 | 0.92 |
| Age of entry | 23.5 | 3.6 | 22.4 | 19.3 | 40.6 |
| % over 25 | 23% | 24% | 15% | 0% | 100% |
| Student-faculty ratio | 14.57 | 5.28 | 14 | 1 | 92 |
| Male graduation rate | 44% | 19% | 44% | 1% | 93% |
| Female graduation rate | 50% | 19% | 51% | 3% | 97% |
| Total graduation rate | 47% | 19% | 48% | 2% | 95% |

### *B. Exploratory Data Analysis*

Summary statistics for each numerical parameter were computed using Python. Distributions of numerical parameters were examined visually using histograms, boxplots, density plots, and quantile-quantile (QQ) plots. Skewness and kurtosis values were also computed for each distribution. The distribution of CSA has a severe right skew (skewness = 23.516) and is taller with more extreme values than the Normal distribution, as the distribution is leptokurtic (kurtosis = 711.943). The distribution of CSA rate per 1,000 students has a less severe right skew (skewness = 3.264) and is less leptokurtic (kurtosis = 14.969) with fewer extreme values than the raw CSA parameter. Visual examination of density histograms () and quantile-quantile plots (), further illustrate the severity of skewness in the CSA parameter.

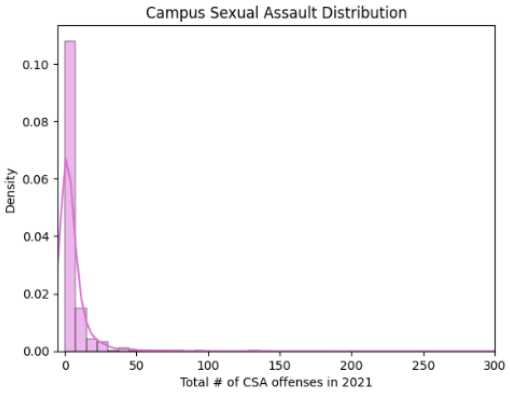


Fig. 6. Distribution of CSA parameter.

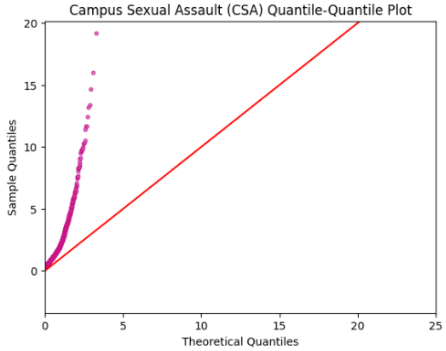


Fig. 7. Quantile-quantile plot of CSA parameter

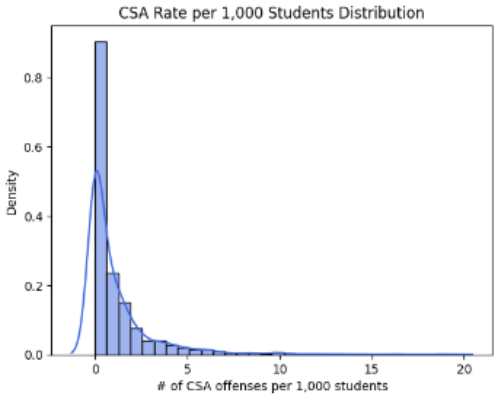


Fig. 8. Distribution of CSA rate parameter.

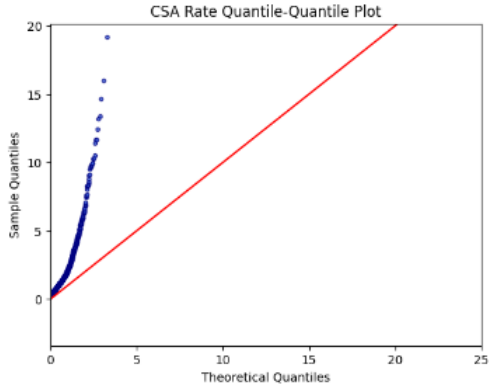


Fig. 9. Quantile-quantile plot of CSA rate parameter.

Institution size and student-to-faculty ratio distributions are both leptokurtic with a severe right skew. The distributions for average age of entry and percent of students over 25 years old have a moderate right skew, and are both platykurtic, indicating they are flatter with fewer extreme values than the Normal distribution. The institutional characteristic with the least amount of skew is the ratio of female students; the skewness value of -0.444 indicates this distribution is slightly left-tailed, which is also unique to this variable. The only other variable with a left-tailed distribution is the female graduation rate parameter (skewness = -0.096). The distribution of male graduation rates is slightly right-tailed (skewness = 0.181), and the distribution of graduation rates for all students is symmetric (skewness = 0.015).

Next, categorical parameters were summarized with bar charts and a contingency table.

| **Housing** | **Sports** | **Private** | **Public** | **Religious** | **All** |
|---|---|---|---|---|---|
| Freshmen not required to live on-campus | Basketball team only | 110 | 154 | 221 | 485 |
| Freshmen not required to live on-campus | Football and basketball team | 125 | 325 | 256 | 706 |
| Freshmen not required to live on-campus | Neither | 73 | 85 | 44 | 202 |
| Freshmen required to live on-campus | Basketball team only | 2 | 0 | 0 | 2 |
| Freshmen required to live on-campus | Football and basketball team | 19 | 2 | 5 | 26 |
| No on-campus housing | Basketball team only | 2 | 12 | 1 | 15 |
| No on-campus housing | Football and basketball team | 1 | 0 | 0 | 1 |
| No on-campus housing | Neither | 220 | 120 | 15 | 355 |
| **All** | | 552 | 698 | 542 | 1792 |

The frequency table shows that out of all combinations of three levels of each of the three categorical parameters, the most frequent combination is public colleges with both football and basketball teams, which do not require freshmen to live on-campus. All colleges that require freshmen to live on-campus have both a football and a basketball team, except for two colleges that have a basketball team only. Most colleges with no on-campus housing have neither a basketball team nor a football team.

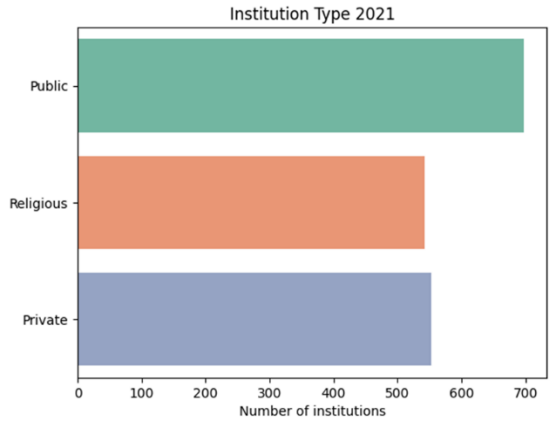


Fig. 10. Distribution of institution types.

As shown in the bar chart of institution type, the most frequent institution type in the dataset is public, followed by private (non-religious).

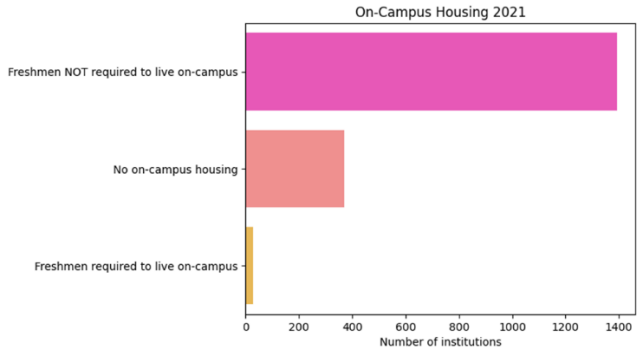


Fig. 11. Distribution of housing type.

As shown in the housing bar chart (11), there are very few colleges that require freshmen to live on-campus, and most colleges offer on-campus housing.

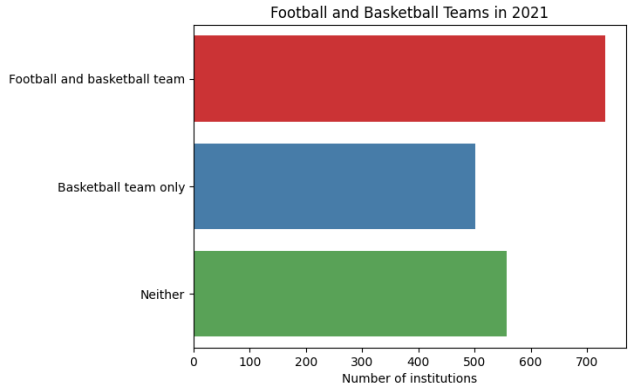


Fig. 12. Distribution of sports team type.

The sports type bar chart (12) shows that roughly the same number of colleges have a basketball team only, or neither team. More colleges have both a football and basketball team than do not.

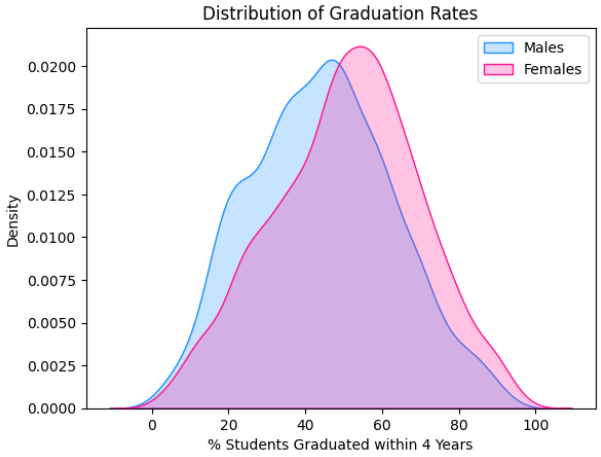


Fig. 13. Distribution of male and female graduation rates.

Examination of male and female on-time graduation rates showed that females tend to graduate at a higher rate than males, as shown in the density plot (13). To prove that the mean graduation rate for males is significantly different from the mean graduation rate for females, a paired t-test was performed using R. The t-test output (14) yielded a t-statistic large in magnitude with a statistically significant p-value ($p<0.05$). The mean difference between male and female graduation rates is -6%. Thus, it can be concluded that male and female graduation rates are significantly different, with females graduating at a higher rate than males, as shown by the negative mean difference.

```
        Paired t-test

data:  df$maleCompleted and df$femaleCompleted
t = -42.714, df = 1791, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -0.06691514 -0.06103990
sample estimates:
mean difference
    -0.06397752
```

Fig. 14. Paired t-test output.

The relationships between two variables were examined next (bivariate analysis). The age parameters average age of entry and percent of students over 25 years old were plotted against each other (). In general, as average age of entry increases, the percent of students over 25 years old also increases, and vice versa. Logically, this makes sense since both of these parameters measure age.

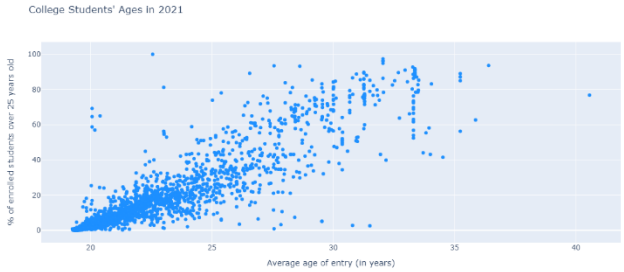


Fig. 15. Average age of entry vs. percent of students over 25 years old.

The relationship between ratio of female students and female graduation rates was examined next. The trend line on the scatterplot (16) shows that in general, higher ratios of female students are associated with lower female graduation rates.

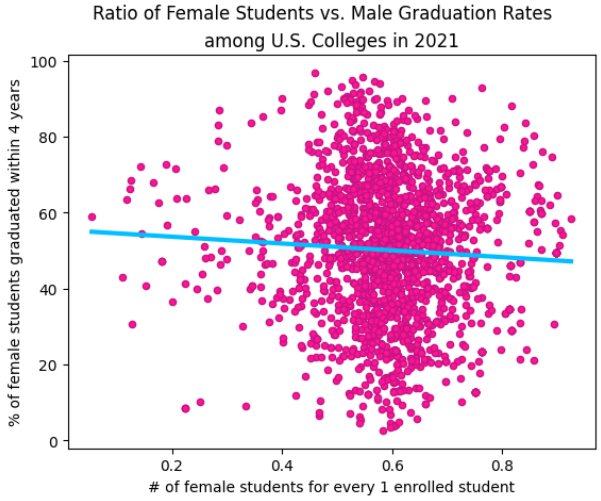


Fig. 16. Ratio of female students vs. female graduation rates.

To discover whether this trend is the same for males, the ratio of male students was computing using () and then was plotted against male graduation rates. The trend line on the scatterplot shows that, in general, as the ratio of male students increases, male graduation rates also increase. So, the trend for males and females is different; there is a negative association between female ratios and graduation rates, but a positive association between male ratios and graduation rates.

$$Male\ Ratio = 1 - Female\ Ratio$$

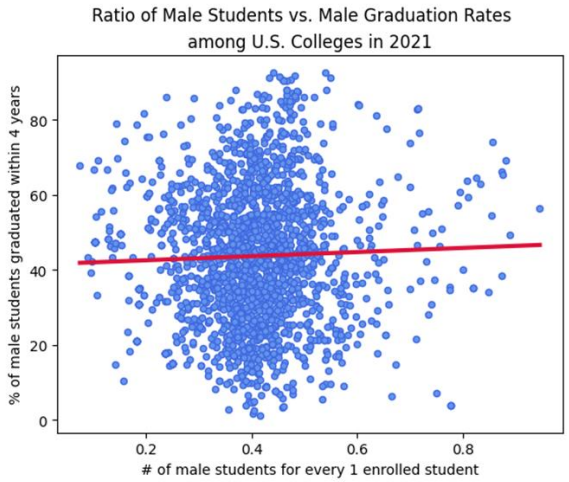


Fig. 17. Ratio of male students vs. male graduation rates.

### C. *Algorithm*

Pearson's correlation coefficients between institutional characteristics were summarized using a correlation heatmap. Pearson's and Spearman's rank correlation coefficients between CSA rate and numerical institutional characteristics were computed using Python. Since the CSA parameter has an extreme skew, the non-parametric Spearman's rank correlation coefficient is more accurate and meaningful than Pearson's coefficient in this case. Point biserial correlation coefficients were computed for categorical institutional characteristics.

An Analysis of Variance (ANOVA) test was performed using R to determine whether mean CSA rate differs between groups of categorical variables (type, housing, and sports). For each of the three variables, the sum of squares, F-statistic, and p-value were computed and displayed in the ANOVA output. Lastly, since the ANOVA output showed there is a significant difference between groups, Tukey's Honest Significant Difference (HSD) test was used to assess the significance of differences between pairs of group means. For each of the three variables, Tukey's HSD output includes three differences between group means, along with a 95% confidence interval and associated p-value.

Simple linear regression was used to determine the relationship between CSA and academic achievement. Two linear equations were obtained using R, one for male graduation rates and one for female graduation rates, along with plots of the regression lines. $R^2$ and adjusted $R^2$ values were used to assess each linear model.

## III. Results

The top 10 colleges for campus sexual assault were examined using a bar chart. Ohio State University has the most CSA offenses in the dataset (676) and the highest rate of CSA: 14.7 CSA offenses for every 1,000 students.

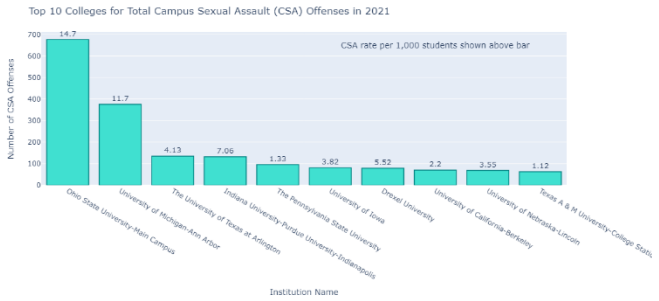


Fig. 18. Top 10 colleges for CSA in 2021.

Next, CSA rates for the 28 colleges that require first-year (freshmen) students to live on-campus were examined (19).

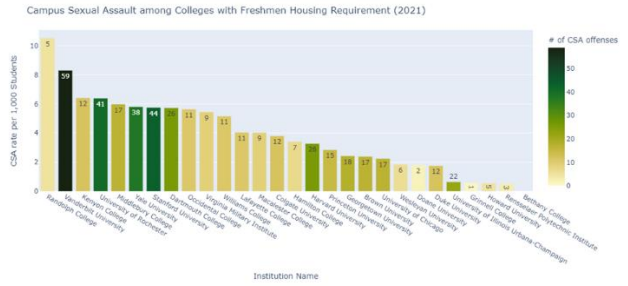


Fig. 19. CSA among colleges that require freshmen to live on-campus.

There are only 8 states that require explicit consent, also known as "yes means yes" consent: California, Florida, Illinois, Minnesota, New Hampshire, New Jersey, Washington, and Wisconsin (Towl & Walker, 2019, p. 40). CSA rate for these 8 schools was examined and visualized using a bar chart. California had the most total CSA offenses of these 8 states in 2021, but Minnesota had the highest rate of CSA per 1,000 students. Besides California and Florida, all of these states have fewer CSA offenses than the state average of 234. Besides Wisconsin and Minnesota, all of these states have lower CSA rates than the state average of 0.82 per 1,000 students.

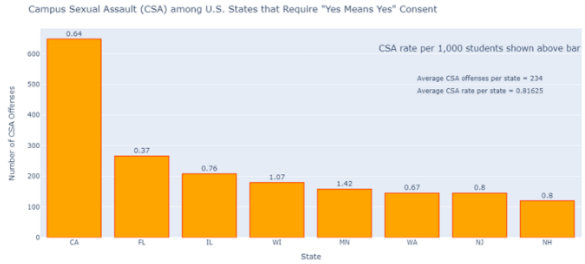


Fig. 20. States that require explicit "yes means yes" consent.

States that require explicit "yes means yes" consent were then visualized on a map with CSA offenses. It appears that there is no relationship between states that require explicit consent and CSA offenses.

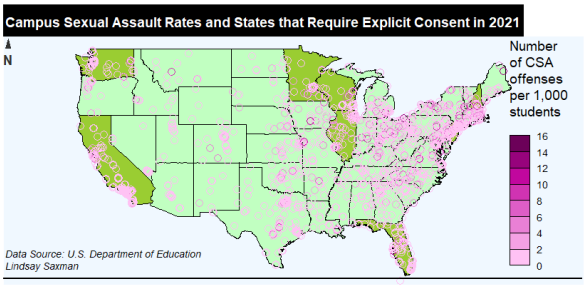


Fig. 21. States that require explicit "yes means yes" consent.

Correlation coefficients for institutional characteristics are shown in the correlation heatmap. The correlation between average age of entry and percent of students over 25 years old has the greatest coefficient (r = 0.88), indicating that there is a strong, positive correlation between these parameters. This is not surprising, considering that as the percent of students over 25 years old increases, average age of entry must also increase, and vice versa. The next largest correlation coefficient is 0.65, which represents the correlation between housing type and sports teams. So, colleges with on-campus housing and the freshmen housing requirement are more likely to have both football and basketball teams. Colleges with no on-campus housing are less likely to have football or basketball teams. The smallest value in the heatmap is -0.69, indicating a strong, negative correlation between housing type and percent of students over 25 years old. There is moderate, negative correlation between housing type and average age of entry as well. This means that, in general, colleges with no on-campus housing are associated with older students, and colleges with on-campus housing, and especially colleges requiring freshmen to live on-campus, are associated with younger students.

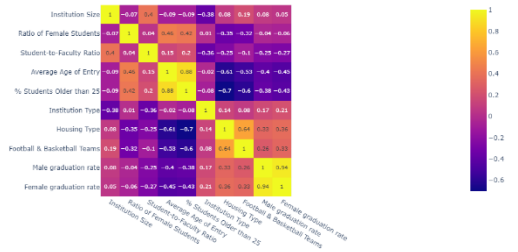


Fig. 22. Correlation heatmap.

Since the CSA data is heavily skewed, the non-parametric Spearman's rank correlation coefficient is more appropriate for the data. Computation of Spearman's rank correlation coefficients between CSA rate and institutional characteristics revealed that there is a moderate, negative association between CSA rate and average age of entry (r = -0.5194), as well as percent of students over 25 (r = -0.5760). This means that as college students' ages increase, the rate of CSA offenses generally decreases. There is a weak, positive association between CSA rate and institution size (r = 0.2930), which makes sense, since as institution size increases, the rate of CSA offenses should generally increase as well. There is a weak, negative association between CSA rate and female ratio (r = -0.2476), as well as between CSA and student-to-faculty ratio (r = -0.1534).

For categorical variables institution type, housing type, and sports teams, point biserial correlation coefficients reveal that the relationship between institution type and CSA rate is weak and positive (r = 0.1221). There is a weak, positive association between housing and CSA rates. So, as housing options increase, CSA offenses generally increase. There is also a weak, positive association between sports and CSA rate (r = 0.3382), so as the number of sports teams increases, CSA rate generally increases as well.

The ANOVA test output (23) showed mean CSA rate differs significantly between institution type groups (public, private, religious), housing type groups (no on-campus housing, freshmen not required to live on-campus, freshmen required to live on-campus), and sports team groups (neither, basketball team only, football and basketball teams).

ANOVA table

| | Sum of Squares | df | *F* | *p* |
|---|---|---|---|---|
| Institution type | 171.610 | 2 | 28.274 | 0 |
| Housing type | 165.090 | 2 | 27.200 | 0 |
| Sports | 251.154 | 2 | 41.379 | 0 |
| Residuals | 5626.505 | 1854 | | |

Fig. 23. ANOVA output.

Post hoc comparisons using Tukey's HSD test indicated that the mean CSA rate was significantly different for all pairs of the housing and sports parameters, but only two of three pairs for the institution type parameter were found to be significant. Tukey HSD output showed that mean CSA rate for public institutions is significantly different for private institutions ( $\bar{x}_{Private} - \bar{x}_{Public} = 0.36$ ) and religious institutions ($\bar{x}_{Religious} - \bar{x}_{Public} = 0.36$), but mean CSA rate does not significantly differ between private (non-religious) and religious institutions ( $\bar{x}_{Religious} - \bar{x}_{Private} = 0.20$; $p = 0.1753$).

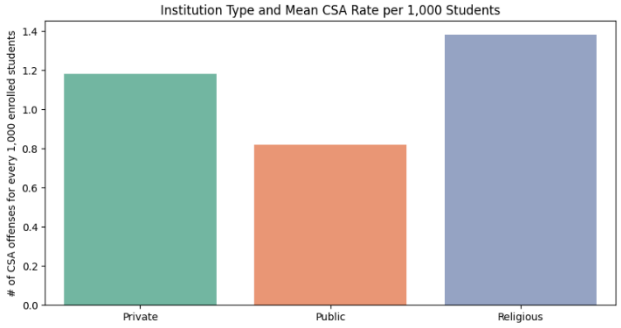


Fig. 24. Mean CSA among institution types.

Mean CSA rate differs significantly between groups of the housing parameter, as shown in the bar chart (25). Tukey's HSD output of the housing variable produced the greatest differences between group means. The greatest difference was between colleges with no on-campus housing and colleges that require freshmen to live on-campus ($\bar{x}_{Freshmen\ required\ on-campus} - \bar{x}_{No\ housing} = 3.79$), followed by the difference between colleges that do require freshmen to live on-campus and colleges that do not require freshmen to live on-campus ( $\bar{x}_{Freshmen\ required\ on-campus} - \bar{x}_{Freshmen\ NOT\ required\ on-campus} = 2.47$ ). There was also a significant difference between colleges with no on-campus housing and colleges with on-campus housing but do not require freshmen to live on-campus ($\bar{x}_{Freshmen\ NOT\ required\ on-campus} - \bar{x}_{No\ housing} = 1.32$).

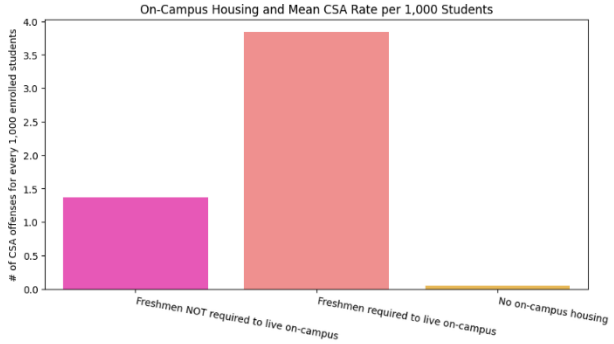


Fig. 25. Mean CSA rate among housing types.

Mean CSA rate is significantly different between groups of the sports parameter, as shown in the bar chart (26). Tukey HSD output for the sports variable showed mean CSA rate differs the most between colleges with both football and basketball teams and colleges with neither team ( $\bar{x}_{Both} - \bar{x}_{Neither} = 1.53$ ). Mean CSA rate also differs between colleges with a basketball team only and colleges with neither team ($\bar{x}_{Basketball\ only} - \bar{x}_{Neither} = 1.04$). Lastly, there is a significant difference in mean CSA between colleges with only basketball teams and colleges with both basketball and football teams ( $\bar{x}_{Both} - \bar{x}_{Basketball\ only} = 0.49$).

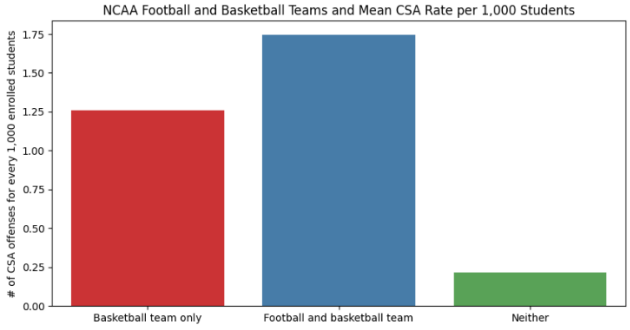


Fig. 26. Mean CSA rate among sports teams types.

Two simple linear equations were obtained to model the relationship between CSA and male and female graduation rates. Both equations were found to be statistically significant ($p<0.05$). The regression output from R for male graduation rates can be observed below (27).

```
MODEL INFO:
Observations: 1861
Dependent Variable: maleCompleted
Type: OLS linear regression

MODEL FIT:
F(1,1859) = 234.47, p = 0.00
R² = 0.11
Adj. R² = 0.11

Standard errors: OLS
-------------------------------------------------
                    Est.   S.E.   t val.      p
----------------- ------ ------ -------- ------
(Intercept)         0.40   0.00    85.24   0.00
rate                0.03   0.00    15.31   0.00
-------------------------------------------------
```

Fig. 27. Simple linear regression output.

The regression output modeling the linear relationship between male graduation rates and CSA (27) indicates that each 1 unit increase in CSA rate is associated with a 3% increase in male graduation rates, and the expected value for male graduation rate is 40% when CSA rate is 0 offenses for every 1,000 students. However, this simple linear model accounts for only 11% of the variation in male graduation rates. This result makes sense with the plot (29), as it is clear that there are many points that fall quite far from the regression line.

The regression output from R for female graduation rates can be observed below (28). The results are similar to those for male graduation rates, with the exception of the intercept coefficient.

```
MODEL INFO:
Observations: 1861
Dependent Variable: femaleCompleted
Type: OLS linear regression

MODEL FIT:
F(1,1859) = 252.66, p = 0.00
R² = 0.12
Adj. R² = 0.12

Standard errors: OLS
-------------------------------------------------
                    Est.   S.E.   t val.      p
----------------- ------ ------ -------- ------
(Intercept)         0.46   0.00    96.77   0.00
rate                0.03   0.00    15.90   0.00
-------------------------------------------------
```

Fig. 28. Example of a figure caption.

The regression output modeling the linear relationship between female graduation rates and CSA (28) indicates that each 1 unit increase in CSA rate is associated with a 3% increase in female graduation rates, and the expected value for female graduation rate is 46% when CSA rate is 0. However, this simple linear model accounts for only 12% of the variation in male graduation rates. This result also makes sense with the plot (29), as it is clear that there are many points that fall quite far from the regression line.

So, the simple linear regression results for male and female graduation rates are almost identical, except the expected graduation rate value when CSA rate is 0 is greater for females, and the equation for female graduation rates fits the data slightly (1%) better than the model for male graduation rates. The scatterplot (29) shown below illustrates the comparison between these two models. The regression lines are parallel, indicating they have the same slope, but they have different y-intercepts, as females tend to graduate at a higher rate than males. These results suggest that CSA prevalence does not have an effect on graduation rates.

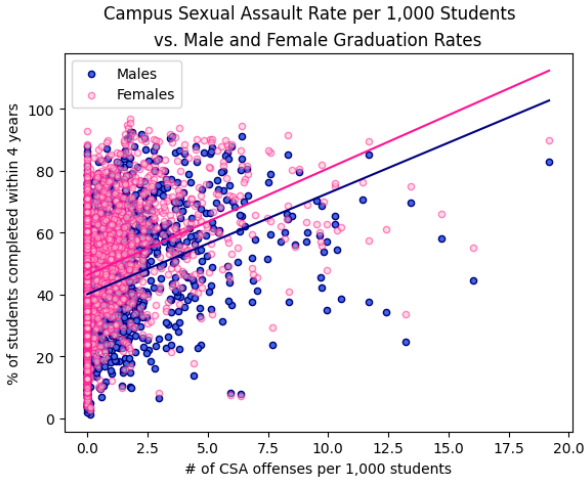


Fig. 29. Campus sexual assault rate vs. male and female graduation rates.

## IV. Discussion

In conclusion, these analyses shed light on the intricate relationship between campus sexual assault and institutional characteristics among U.S. colleges in 2021. The examination of factors such as average age of entry, student-to-faculty ratio, and female ratio revealed intriguing correlations that warrant further analysis and investigation. The observed associations between institutional characteristics and incidences of CSA underscore the importance of considering broader contextual elements in understanding and addressing this critical issue on U.S. campuses. Moreover, although CSA accounts for very little variation in graduation rates, the analysis did reveal a weak association between campus sexual assault and graduation rates for males and females. Investigating other factors contributing to graduation rates fell outside of the scope of this project, but future work could include more variables to include in the model in order to create a better, more robust model.

Universities with particular demographic compositions or educational structures may demonstrate varying levels of vulnerability to CSA, as well as varying levels of academic achievement. Future work could include demographic data like race and income and more institutional characteristics, such as presence of Greek life, to further understand the context behind CSA and academic achievement. The Department of Education's Campus Safety and Security Tool includes, in addition to on-campus sexual assaults, non-campus and public property sexual assault data, as well as sexual assault offenses occurring in student housing. Future

work could utilize these other datasets of CSA to examine how the setting of the assault (on-campus, non-campus, public property, or student housing) is related to institutional characteristics and academic achievement. Also, the analyses only included 4-year institutions, so future work could include other institutions levels, such as 2-year institutions, to determine if institution level is associated with CSA. Men-only and women-only institutions were excluded in analyses; it might be intriguing to investigate how CSA compares between men-only and women-only colleges.

The most prominent limitation of this study is data availability and confidentiality. Sexual assault data is generally difficult to obtain due to confidentiality laws, and the data that is available is incomplete since sexual assault often goes unreported. Another limitation is depth of analyses; determining correlations, not causation, was the focus of this project. Future work should include causal relationships.

Moving forward, policymakers, university administrators, and advocates can leverage these findings to develop targeted interventions and policies aimed at creating safer campus environments. For example, prevention measures should be focused on colleges that require freshmen students to live on-campus since these schools tend to see more occurrences of CSA. By acknowledging and addressing the influence of institutional characteristics on the prevalence of CSA, proactive measures can be implemented to foster a culture of safety, support, and inclusivity within higher education institutions. Ultimately, this project serves as a foundational step in understanding the nuanced dynamics between institutional characteristics and CSA and encourages a continued exploration of these factors and their role in shaping the safety and well-being of students on college campuses.

## REFERENCES


[1] Anderson, J. C., Peruggia, G., Miller-Walfish, S., Talis, J., Burrell, C., Hayes, M., & Miller, E. (2023). A case series measuring campus and clinic level factors during implementation of a sexual violence prevention intervention in campus health and counseling centers: Does environment matter? Implementation Science Communications, (4). https://doi.org/10.1186%2Fs43058-023-00467-7

[2] Campus sexual violence: Statistics. (n.d.). Rape, Abuse & Incest National Network. Retrieved November 7, 2023, from https://www.rainn.org/statistics/campus-sexual-violence

[3] DeLong, S. M., Graham, L. M., Magee, E. P., Treves-Kagan, S., Gray, C. L., McClay, A. M., Zarnick, S. M., Kupper, L. L., Macy, R. J., Ashley, O. S., Pettifor, A. P. P., Moracco, K. E., & Martin, S. L. (2018). Starting the conversation: Are campus sexual assault policies related to the prevalence of campus sexual assault? Journal of Interpersonal Violence, 33(21). https://doi-org.mutex.gmu.edu/10.1177/0886260518798352

[4] Geppert, A. B., Shah, A. H., & Hirsch, J. S. (2023). "Hardly Able to Move, Much Less Open a Book": A Systematic Review of the Impact of Sexual and Gender-Based Violence Victimization on Educational Trajectories. Trauma, Violence, & Abuse.

[5] Islam, K. N., & Biswas, H. A. (2020, December 26). Mathematical assessment for the dynamical model of sexual violence of women in Bangladesh [Paper presentation]. International Conference on Industrial & Mechanical Engineering and Operations Management, Dhaka, Bangladesh. http://www.ieomsociety.org/imeom/169.pdf

[6] Sexual assault. (2023, May 8). U.S. Department of Justice Office on Violence Against Women. Retrieved November 28, 2023, from https://www.justice.gov/ovw/sexual-assault

[7] Snodgrass, G. M., Rosay, A. B., & Gover, A. R. (2014). Modeling the referral decision in sexual assault cases: An application of random forests. American Journal of Criminal Justice, (39), 267-291.

[8] Towl, G. J., & Walker, T. (2019). Sexual violence on US campuses. In Tackling sexual violence at universities: An international perspective (pp. 38-58). Routledge.

[9] Tredinnick, L. (2020). Sexual assault prevention with student-athletes: Exploring perceptions of the campus climate and awareness of sexual assault policies and resources. Journal of Interpersonal Violence, 37(9-10). https://doi-org.mutex.gmu.edu/10.1177/0886260520967144

[10] An underreported problem: Campus sexual misconduct. (n.d.). American Association of University Women (AAUW). Retrieved November 28, 2023, from https://www.aauw.org/resources/article/underreported-sexual-misconduct/