Ben Webb, Kiana Chabot, Lena Martin
SC212A
29 April 2019

# Athletic Status Influence on Roommate Choice

## Introduction

Are college athletes more likely to live with other college athletes than non-athletes? Most members of the sophomore, junior, and senior classes at Colby have the opportunity to become roommates with peers of their choice, and many students spend large amounts of time with these chosen roommates. Athletes on the other hand, cannot select who will be on their team, but do still spend a significant amount of time surrounded by their teammates.

Athletes comprise a unique liminal state on college campuses. Valentine et al. posed the seemingly ironic position of college athletes where despite continuously being in the spotlight, student athletes spend significantly less time in contact with non-athlete peers, and thus tend to feel more isolated than their non-athlete counterparts. Most of the research has been on Division I athletes who spend significantly more time practicing and traveling for their sport. Colby has primarily Division III sports teams but there also are Division I teams. It has also been shown that college students struggle to make friends outside of their athletic teams, in part due to the time requirements of their sport in addition to their academic requirements.

Additionally, psychological theories suggest that behaviors are often acquired through modeling of peers or in observation of peers' behavior (Li & Guo, 2016), so that roommates who spend large amounts of time together may be particularly susceptible to the influence of each others' behavior. There is abundant literature discussing athletes' increased aggression outside sports realms, the association between team sports and increased alcohol consumption, and the

high prevalence of sexual assault perpetration among athletes compared to non-athletes (Harway & Steel, 2015). Roommate groups of one or many deviant athletes may exacerbate deviant behaviors on Colby's campus.

**Methods**

The data used in this observational study was provided in part by the Office of Campus Life at Colby College. At this point, freshmen were excluded from the data set because first-years are assigned roommates rather than choosing their own so any crossover between roommates athlete status is likely by chance. From the Office of Campus Life spreadsheet, information including birth date, class year, gender, hometown, and on campus residence. The hometown statistic was further analyzed by finding the latitude and longitude so that hometown distance from Colby could be calculated. Additionally, age was converted into a continuous numeric variable and is reported as students age as of April 18th. In order to assess athlete status and sports team affiliation, data was pulled from rosters on the Colby College Athletics website. The data that was collected from the Office of Campus Life and the Colby College Athletics website was compiled into a comprehensive spreadsheet.

**Results**

Most sophomore, junior, and senior students at Colby (n = 1288) are non-athletes who do not live with athletes (count = 743, p = 0.577). Fewer of those students are athletes who live with other athletes (count = 240, p = 0.186), and even fewer are athletes who live with non-athletes (count = 157, p = 0.122). The remainder of sophomore, junior, and senior Colby students are non-athletes who live with athletes (count = 148, p = 0.115). Figure 1 shows the proportions of Colby sophomores, juniors, and athletes and non-athletes who do or do not live with athletes.

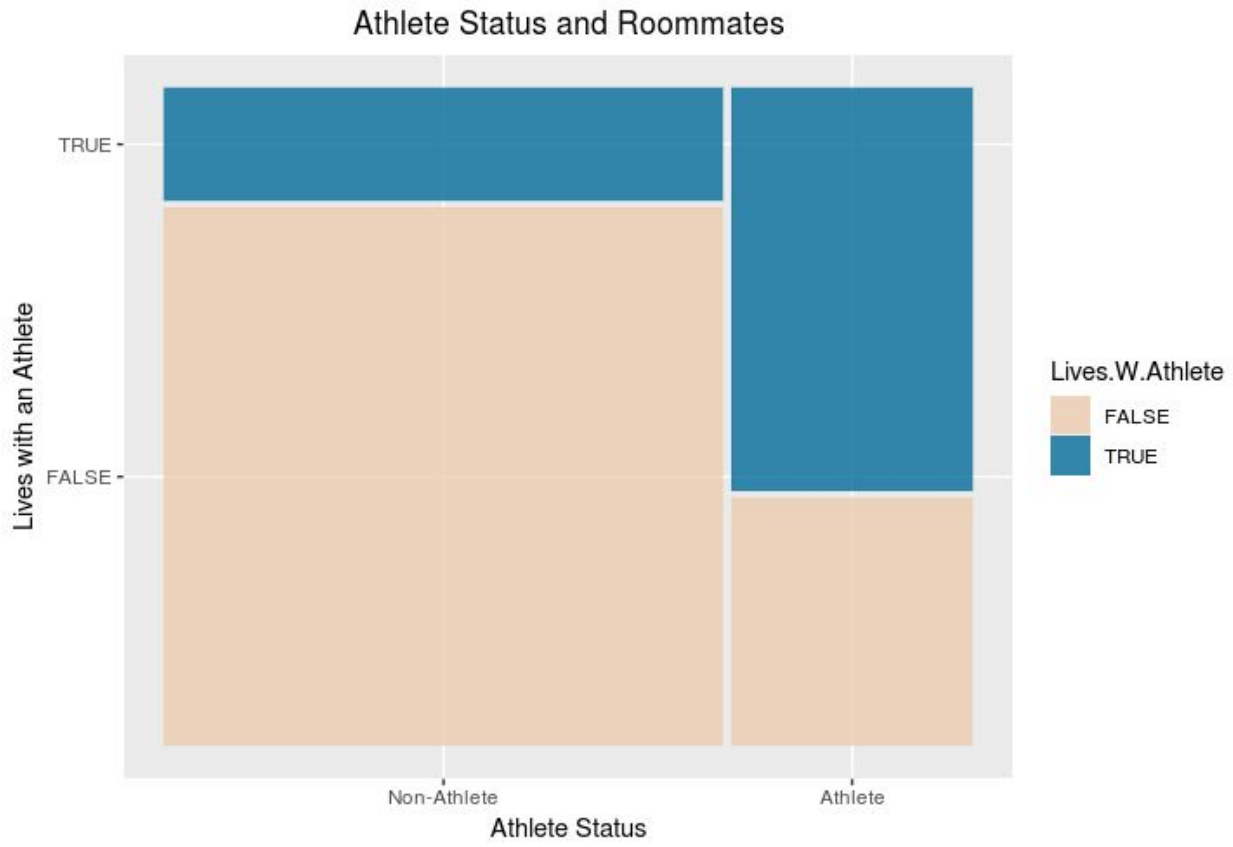



*Figure 1. Proportions of athletes and non-athletes who do or do not live with athletes.*

We compared the proportion of athletes who live with other athletes to the proportion of non-athletes who live with other athletes. We found significant evidence that the proportion of athletes who live with other athletes is greater than the proportion of non-athletes who live with athletes (p-value = 0.000).

Furthermore, we looked at the dynamics of living with teammates. We found significant evidence that athletes are likely to live with other members of their team over other athletes (p-value = 0.05). We did not include non-athletes in this analysis, because they cannot live with teammates.

The locations of sophomore, junior, and senior Colby students' hometowns are a mean distance of 2028 km from Colby. Figure 2 shows distributions of hometown proximities to Colby for non-athletes ($\mu$ = 2339) and athletes ($\mu$ = 1306). We have significant evidence that mean proximity of non-athletes' hometowns to Colby is different than the mean proximity of athletes' hometowns to Colby (p-value = 0.000).

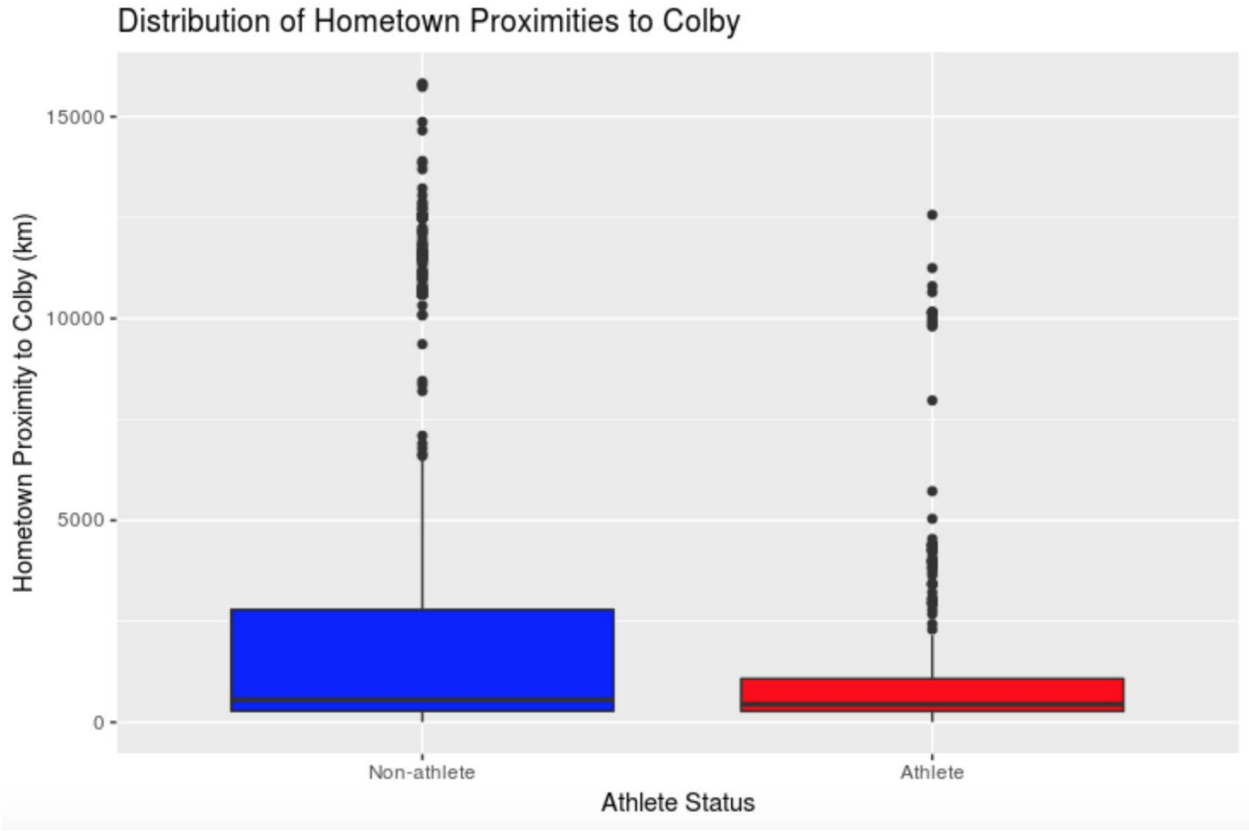



*Figure 2. Distribution of hometown proximities to Colby for athletes and non-athletes.*

Figure 3 shows that there is not a strong association between students' hometown proximities to Colby and their age. The linear regression with confidence interval of both Athlete and non-athletes are also shown. This data shows no clear trends between the age of the student and their distance home.

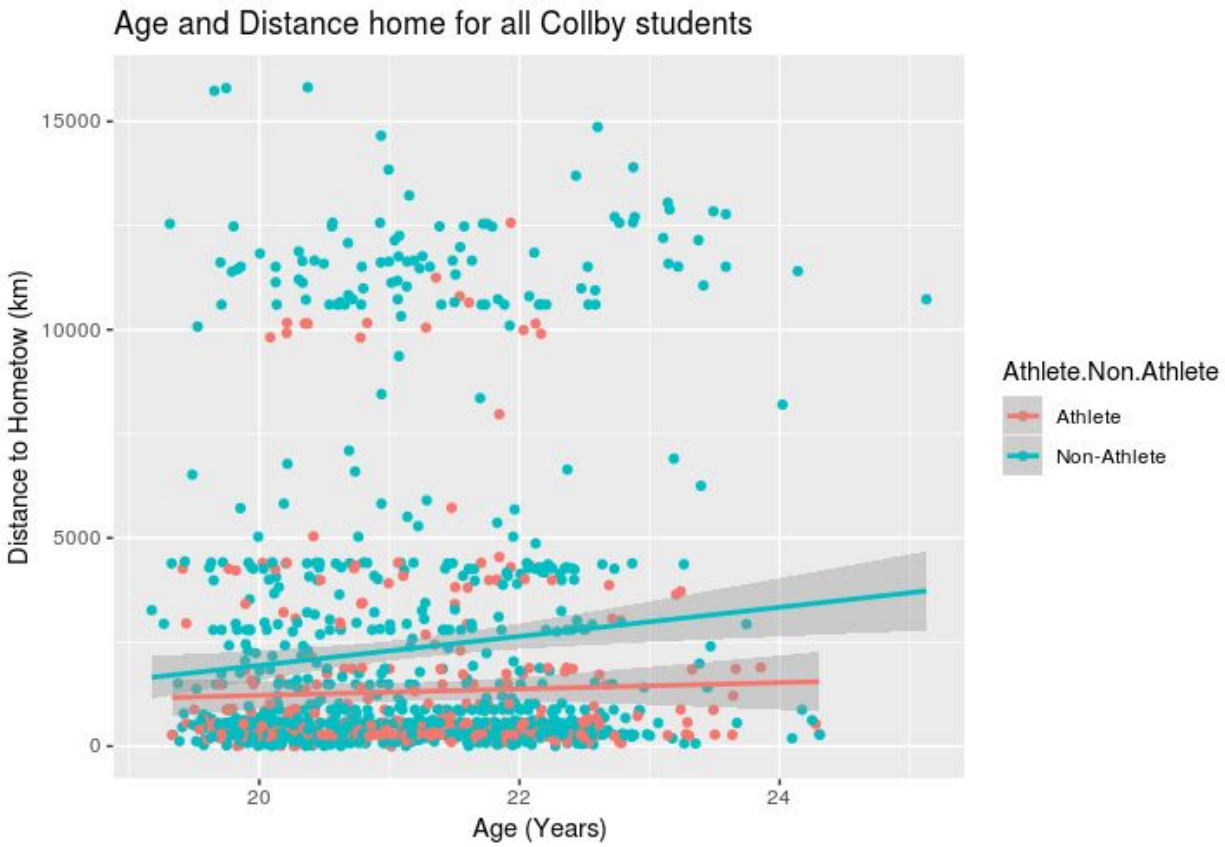



*Figure 3. Correlation between student age and hometown proximities to Colby.*

**Discussion**

We found that the proportion of Colby athletes who have athletes as roommates is greater than the proportion Colby non-athletes who have athletes as roommates. This result is unsurprising because the Colby students in the dataset were able to select their roommates, and athletes have multiple reasons to select each other as roommates (interests in common, similar sleep schedules because of early or late practices, etc.). Non-athletes do not have such natural reasons to select athletes as roommates. We also found that athletes are likely to live with other teammates over others. Finally, we found that the mean proximity of non-athletes' hometowns to Colby is greater than the mean proximity of athletes' hometowns to Colby. It appears that only a

handful of athletes have hometowns greater than 10,000 km from Colby: most of Colby's athletes are Americans and not international students.

Our results encourage a survey of Colby athletes to investigate whether they are suffering from rooming with each other. Are athletes becoming isolated from their non-teammate peers as Valentine et al. suggests? Are roommate groups consisting of teammates more prone to deviant behaviors? Our results also encourage a survey of Colby coaches to investigate why more American students have been recruited for their sports teams than international students.

**Bibliography**

Li, Y. & Guo, G. (2016). Peer Influence on Aggressive Behavior, Smoking, and Sexual Behavior: A Study of Randomly-assigned College Roommates. *Journal of Health and Social Behavior,* 57(3), 297-318.

Harway, M. & Steel, J. (2015). Studying masculinity and sexual assault across organizational culture groups: Understanding perpetrators. *Psychology of Men and Masculinity,* 16(4), 374-378.

Valentine, J. J., & Taub, D. J. (1999). Responding to the developmental needs of student athletes. *Journal of College Counseling*, 2, 164–179.

**Appendix**

*Test for a difference in proportions:* athletes who live with other athletes and non-athletes who live with other athletes
Parameter: Difference in proportion of athletes who live with other athletes and proportion of non-athletes who live with other athletes
Hypotheses: $H_o$: $p_A = p_N$; $H_a$: $p_A > p_N$
Conditions:

- Independent groups: yes, athletes and non-athletes
- Large enough sample: yes
    - 397(0.6045) > 10
    - 397(0.3955) > 10
    - 397(0.3955) > 10
    - 397(0.6045) > 10
- SRS: no, entire population

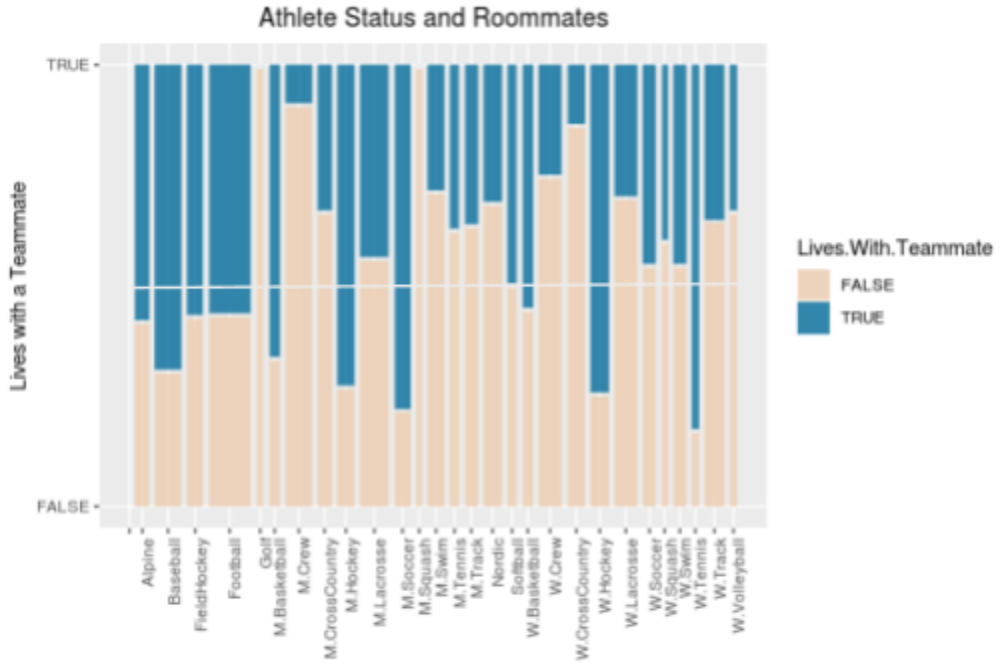
Test statistic: Z = (0.6045 - 0.3955)/sqrt((0.25)(1/397 + 1/397)); Z = 0.209/0.0354886714; Z = 5.8892
P-value: 0.000
Decision: reject the null

$\chi^2$ Test for Association
Difference in proportion of athletes who live with other athletes and proportion of non-athletes who live with other athletes
$H_o$: Roommates choice is completely random; $H_a$: Team is associated with roommate choice
$\chi^2 = 9116.12$
P-value: 0.000

| **Team** | **Observed True** | **Expected True** | **Size of Team** |
|---|---|---|---|
| Alpine | 7 | 0.1025641 | 12 |

| | | | |
|---|---|---|---|
| Baseball | 16 | 0.39316239 | 23 |
| FieldHockey | 8 | 0.14141414 | 14 |
| Football | 21 | 1.03496503 | 37 |
| Golf | 0 | 0.02331002 | 6 |
| M.Basketball | 6 | 0.05594406 | 9 |
| M.Crew | 2 | 0.39316239 | 23 |
| M.CrossCountry | 5 | 0.1025641 | 11 |
| M.Hockey | 11 | 0.16317016 | 15 |
| M.Lacrosse | 11 | 0.46620047 | 25 |
| M.Soccer | 11 | 0.14141414 | 14 |
| M.Squash | 0 | 0.01554002 | 5 |
| M.Swim | 4 | 0.14141414 | 14 |
| M.Tennis | 3 | 0.04351204 | 8 |
| M.Track | 4 | 0.08547009 | 11 |
| Nordic | 5 | 0.18648019 | 16 |
| Softball | 4 | 0.04351204 | 8 |
| W.Basketball | 5 | 0.05594406 | 9 |
| W.Crew | 5 | 0.2952603 | 20 |
| W.CrossCountry | 0 | 0.12121212 | 13 |
| W.Hockey | 12 | 0.18648019 | 16 |
| W.Lacrosse | 6 | 0.2952603 | 20 |
| W.Soccer | 5 | 0.08547009 | 11 |
| W.Squash | 2 | 0.01554002 | 5 |
| W.Swim | 5 | 0.08547009 | 11 |
| W.Tennis | 5 | 0.02331002 | 6 |
| W.Track | 6 | 0.21134421 | 17 |

| | | | |
|---|---|---|---|
| W.Volleyball | 2 | 0.02331002 | 6 |
| Total | 171 | | 1288 |




*Welch Two sample t-test:* proximities of hometowns for athletes and proximities of hometowns for non-athletes

Parameter: difference in mean proximity of hometown to Colby for athletes and mean proximity of hometown to Colby for non-athletes

Hypotheses: $H_o$: $\mu A = \mu N$; $H_a$: $\mu N \neq \mu N$

Conditions:

- Independent groups: yes, athletes and non-athletes
- Large enough sample
    - 900 > 30
    - 388 > 30
- SRS: no, entire population

```
data:  housing.df$DistanceHome by housing.df$Athlete.Non.Athlete
t = 6.2692, df = 1150.8, p-value = 5.122e-10
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  709.7701 1356.4018
sample estimates:
mean in group FALSE  mean in group TRUE
          2339.414           1306.328
```

Test statistic = 6.2692
P-value: 0.000
Decision: reject the null

### Selected Code ###

```
proj.data <- read.csv(file=file.choose()) # Read CSV
proj.data$RoomLocation        <-        paste(proj.data$Building,
gsub("[A-z]","",proj.data$Room)) # Combine Suites
clean.prj  <-  proj.data[order(proj.data$RoomLocation),]  #  Group
by Room
clean.prj  <-  clean.prj[!clean.prj$Class.Year  ==  22,]#  Remove
First Years

athlete.data <- group_by(clean.prj, clean.prj$RoomLocation) %>%
                filter( Athlete.Non.Athlete,  n() > 1  )  %>%
                mutate( Lives.With.Other.Athlete = TRUE ) %>%
                                ungroup()  %>%  select(  Student,
Lives.With.Other.Athlete )

teammate.data   <-   group_by(clean.prj,   clean.prj$RoomLocation,
clean.prj$Sports.team) %>%
                 filter( Athlete.Non.Athlete,  n() > 1  )  %>%
                 mutate( Lives.With.Teammate = TRUE ) %>%
                                 ungroup()  %>%  select(  Student,
Lives.With.Teammate)

for (i in 1:nrow(clean.prj)){
      if   (clean.prj$Student[i]   %in%   athlete.data$Student){
clean.prj$Lives.W.Athlete[i] = TRUE }
  else { clean.prj$Lives.W.Athlete[i] = FALSE }
     if   (clean.prj$Student[i]   %in%   teammate.data$Student){
clean.prj$Lives.With.Teammate[i] = TRUE }
  else { clean.prj$Lives.With.Teammate[i] = FALSE }
}
```

```
write.csv(clean.prj,"project.csv") # Save CSV

getGeoCode <- function(gcStr)
{
  library("RJSONIO") #Load Library
  gcStr <- gsub(' ','%20',gcStr) #Encode URL Parameters
  #Open Connection
                                    connectStr                  <-
paste('https://maps.googleapis.com/maps/api/geocode/json?address
=',gcStr, '&key=’,sep="")
  con <- url(connectStr)
  data.json <- fromJSON(paste(readLines(con), collapse=""))
  close(con)
  #Flatten the received JSON
  data.json <- unlist(data.json)
  lat <- data.json["results.geometry.location.lat"]
  lng <- data.json["results.geometry.location.lng"]
  gcodes <- c(lat, lng)
  names(gcodes) <- c("Lat", "Lng")
  return (gcodes)
}
getGeoCode("BurkeVT")

ggplot(project.data) +
geom_mosaic(aes(x                                                =
product(Lives.W.Athlete,Athlete.Non.Athlete),
fill=Athlete.Non.Athlete), na.rm=TRUE) +
labs(x = "Athlete Status", y = 'Lives with an Athlete', title =
"Athlete Status and Roommates") +
scale_fill_manual(values=wes_palette(n=5, name="Darjeeling2")) +
theme(plot.title = element_text(hjust = 0.5))
```