Comparison of the Trends of the Proportion of Women Enrolled in Biology and Life Science, Computers and Mathematics, Engineering, Health, and Physical Science Majors at the University of Florida

## Natalie Skala

2 December 2019

# Motivation and Background

As a woman in a STEM field, I didn’t feel the effects of the gender gap in STEM until my junior year of college, when I had several job interviews each week. I was interviewed time and time again by male interviewers who hire predominantly male employees which made female representation in companies very important to me. I asked my female friends in other STEM fields for their opinions and found that my engineering friends seemed to feel there was significantly more males than females in their engineering specific classes compared to their other classes. This made me wonder if there was a statistically significant difference in proportions for females in each STEM field. I found data from Fivethirtyeight that I could compare to the University of Florida. My goal was to find the University of Florida enrollment data, compare the female proportion for each STEM field to see how the proportion of females has changed over time, and how the University of Florida proportions compare to the other colleges.

# Aims

1 Compare the proportion of females enrolled in Biology and Life Science, Computers and Mathematics, Engineering, Health, and Physical Science majors at the University of Florida from 2014-2018

2 Compare the total proportion of females enrolled in Biology and Life Science, Computers and Mathematics, Engineering, Health, and Physical Science majors at the University of Florida from 2010-2012 to the total proportion of recent graduated females from 2010-2012 in Biology and Life Science, Computers and Mathematics, Engineering, Health, and Physical Science majors

# Data Sources

**Aim 1:**

Fivethirtyeight previously created a dataset and wrote an article about “The Economic Guide to Picking a College Major.” When I searched through their data, I found a women in STEM dataset to help me guide me in categorizing data. The data can be found at [github.com/fivethirtyeight/data/tree/master/college-majors](https://github.com/fivethirtyeight/data/tree/master/college-majors). All the Fivethirtyeight data is from American Community Survey 2010-2012 Public Use Microdata Series. This survey is produced by the Census Bureau. Fivethirtyeight created this dataset including totals for recently graduated undergraduate degree seeking students in Biology and Life Science, Computers and Mathematics, Engineering, Health, and Physical Science majors. To create my dataset for the University of Florida, I used the “Headcount by Major Table” on the website ir.aa.ufl.edu/uffacts/enrollment-1/ and compared each major in each category in the Fivethirtyeight dataset to create my dataset. I first filtered by undergraduate (degree-seeking) students only. Using only the chosen majors, I entered the totals for each category of major from the years 2014-2018 and then recorded the total number of female students enrolled in each category.

I manipulated the University of Florida dataset and manually entered into a spreadsheet the following variables:

* Year
* Category Category of major
* Total Total number of students enrolled
* Women Total number of women enrolled
* Proportion Calculated as Women/Total

**Aim 2:**

Since the Fivethirtyeight dataset was for total recent graduates from 2010-2012, I went back to the University of Florida dataset and recorded the totals from 2010-2012. I again filtered by undergraduate(degree-seeking) students, used only the chosen majors, entered the totals from 2010-2012 in each category of major, and entered the total number of female students enrolled in each category. To compare my totals, I created a pivot table in Excel of the Fivethirtyeight data of each Major Category as the rows and the columns as the sum of the number of students enrolled and the sum of the total number of women enrolled.

In both the Pivot Table and the new University of Florida data set I entered the following variables:

* Category Category of major
* Total Total number of students enrolled
* Women Total number of women enrolled
* Proportion Calculated as Women/Total

# Data Analysis

First, we note that this data is observational in nature and there is no way to randomly select units to receive a treatment. This is demographic data on two populations: all recent college graduates from 2010-2012, and all students enrolled at the University of Florida. The data for each year recorded in the file is a census count. No random samples were drawn; sampling was not needed, as the entire populations were counted. Therefore, it isn’t necessary to do inference based on this data. For Aim 1, the main part of the analysis will be a graph with five lines for each major category over the years 2014 to 2018.

The data file was read into R:

library(readxl)

women\_stem <- read\_excel("R/4211/women-stem.xlsx")

women\_stem$Category <- factor(women\_stem$Category, labels=c("Biology & Life Science","Computers & Mathematics","Engineering", "Health", "Physical Science"))

Year Category Total Women Proportion

*<dbl>* *<fct>* *<dbl>* *<dbl>* *<dbl>*

1 2014 Engineering 6317 1620 0.256

2 2015 Engineering 6430 1700 0.264

3 2016 Engineering 6814 1868 0.274

4 2017 Engineering 6885 1951 0.283

5 2018 Engineering 7002 2018 0.288

6 2014 Health 2600 2169 0.834

7 2015 Health 2647 2230 0.842

8 2016 Health 2670 2262 0.847

9 2017 Health 2743 2313 0.843

10 2018 Health 2988 2514 0.841

# ... with 15 more rows

The proportions for each major category are then compared on a graph from 2014 to 2018.

plot(women\_stem$Year,women\_stem$Proportion, pch=21, bg=c("red","green3","blue", "orange", "purple")

[unclass(women\_stem$Category)], main = "Proportion of Women in STEM at UF, 2014-2018",

xlab = "Year", ylab="Proportion")

lines(women\_stem$Year[women\_stem$Category=="Biology & Life Science"],

women\_stem$Proportion[women\_stem$Category=="Biology & Life Science"], col="red", type="b")

lines(women\_stem$Year[women\_stem$Category=="Computers & Mathematics"],

women\_stem$Proportion[women\_stem$Category=="Computers & Mathematics"], col="green3", type="b")

lines(women\_stem$Year[women\_stem$Category=="Engineering"],

women\_stem$Proportion[women\_stem$Category=="Engineering"], col="blue", type="b")

lines(women\_stem$Year[women\_stem$Category=="Health"],

women\_stem$Proportion[women\_stem$Category=="Health"], col="orange", type="b")

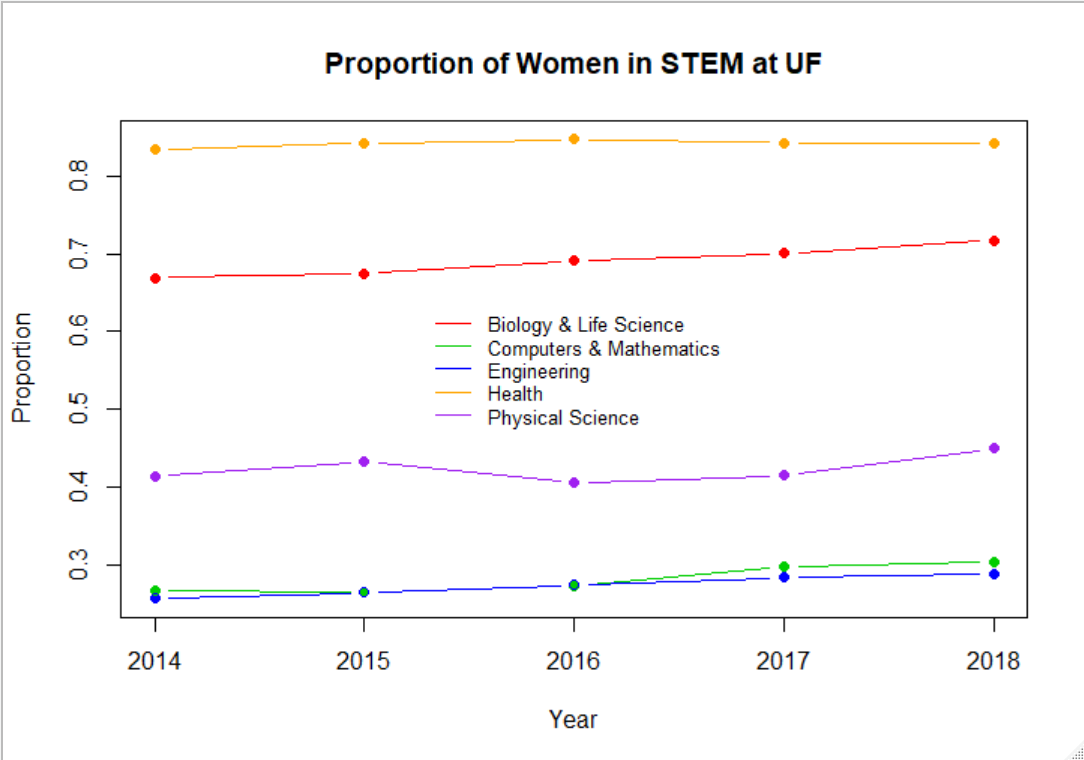
lines(women\_stem$Year[women\_stem$Category=="Physical Science"],

women\_stem$Proportion[women\_stem$Category=="Physical Science"], col="purple", type="b")

legend("center",legend=c("Biology & Life Science",

"Computers & Mathematics","Engineering", "Health", "Physical Science"),

col=c("red","green3","blue", "orange", "purple"), lty=1,cex=0.8, box.lty = 0)

The graph shows a clear divide of popularity with more women enrolled in Biology and Life Science and Health majors and fewer women enrolled in Physical Science, Computers and Mathematics, and Engineering majors. The lowest proportions of women are in Computers and Mathematics and Engineering with about 30.4% and 28.8% respectively in 2018. The Biology and Life Science and Health fields are more dominated by women with proportions of about 71.7% and 84.1% respectively in 2018. Interestingly, the only major category close to an even proportion of women and men/other is Physical Science with about 45% women in 2018.

Overall, the proportions are increasing over time. The largest increase of about 4.8% from 2014 to 2018 was in Biology and Life Sciences. Computers and Mathematics also increased about 3.8%, Physical Science increased about 3.5%, Engineering increased about 3.2%, and Health increased by 0.7%.

For Aim 2, the main analysis is comparing two tables with data from the University of Florida and data from Fivethirtyeight to see how the University of Florida proportions in each major category differ from the total proportion.

|  |  |  |  |
| --- | --- | --- | --- |
| University of Florida |  |  |  |
| Category | Total | Women | Proportion |
| Biology & Life Science | 10813 | 6205 | 0.573846 |
| Computers & Mathematics | 1936 | 574 | 0.296488 |
| Engineering | 16197 | 3781 | 0.233438 |
| Health | 9168 | 7557 | 0.82428 |
| Physical Science | 4316 | 2215 | 0.513207 |

|  |  |  |  |
| --- | --- | --- | --- |
| Fivethirtyeight |  |  |  |
| Category | Total | Women | Proportion |
| Biology & Life Science | 453862 | 268943 | 0.592566 |
| Computers & Mathematics | 299008 | 90283 | 0.301942 |
| Engineering | 537583 | 129276 | 0.240476 |
| Health | 463230 | 387713 | 0.836977 |
| Physical Sciences | 185479 | 90089 | 0.48571 |

Overall, the University of Florida has a slightly smaller proportion of women than the total proportion of women in every category except for Physical Sciences with about 2.8% more women. The University of Florida has about 1.9% less women in Biology and Life Science, about 1.3% less women in Health, about 0.7% less women in Engineering, and about 0.5% less women in Computers and Mathematics.

# Discussion Points

* In recent years, there has been a national push to help increase the number of women in STEM and a push at the University of Florida for higher college rankings. From the University of Florida data, this could help explain the increases in proportion of women in each major category from 2014-2018.
* This could also explain why from 2010-2012 the University of Florida had a smaller proportion of women in STEM than the total proportion and, recently, decided to work to increase these proportions.

# References

Casselman, Ben. “The Economic Guide To Picking A College Major.” *Fivethirtyeight*, 12 Sept. 2014, fivethirtyeight.com/features/the-economic-guide-to-picking-a-college-major/.

“Enrollment & Demographics.” *Institutional Planning and Research, University of Florida*, 29 Mar. 2019, ir.aa.ufl.edu/uffacts/enrollment-1/.

Fivethirtyeight. “Fivethirtyeight/Data.” *GitHub*, 9 Feb. 2018, github.com/fivethirtyeight/data/tree/master/college-majors.

Skala, Natalie. “Women in STEM at UF.” *GitHub*, 1 Dec. 2019, github.com/natskala/women\_stem/blob/master/women-stemUF.csv.