

# How Old are Williams Faculty Members?

*Austin Vo '18*

*2016-02-02*

.you need to do CTRL+ALT+L, load all .you need ggplot2

## Introduction

In the **facultyeph**s package, a user will be able to find a variety of public information (age, degree, field of study, etc.) on the Williams College faculty. The package's focus lies on being able to manipulate faculty information primarily based on their age, such as finding out the mean age of Williams College faculty or looking at the oldest and youngest faculty member for each year. I assumed that a faculty member receives a BA when they are 22 and estimated their age accordingly. Although some people get their BA earlier than 22, some entries only have PhD or Masters in their information and should ultimately balance out the assumption of 22 years old. Data has been gathered for the last five years starting with the 2014-2013 academic year and goes back to the 2010-2009 academic year. Unfortunately, the most current academic year does not include each faculty member's year in which they received their undergraduate degree.

It is useful to input this code before using the package: `devtools::load_all(".") library(ggplot2)`

## Background

In order to look at how the **facultyeph**s package works I will use the latest data set that contains Williams College faculty ages, the 2013-2014 academic year.

```
## Loading facultyeph
```

```
## Welcome to the facultyeph package
```

```
## [1] "Name"           "Title"           "Degree"           "Graduation"
## [5] "Age"            "Undergraduate"   "Field"
```

The information provided for each of the faculty member does not only include age but also degree, field of study, graduation year and undergraduate institution. There are approximately 350-400 faculty members each year. From the number of entries starting from catalog09 to catalog13 there seems to be more faculty for each successive year.

A brief description of each variable:

- Name: the name of the faculty member
- Title: the title of the faculty member, most of which include their field of study
- Degree: indicates whether they have a BS, BA, etc.
- Graduation: the year the faculty member graduated from his/her undergraduate institution
- Age: age was calculated by taking their graduation and subtracting it from the appropriate academic year and adding 22
- Undergraduate: the undergraduate institution from which he/she graduated
- Field: the field of study, typically associated with a respective major or concentration

Sometimes a faculty member's title includes more than one field of study or major; in this case, I chose the most prominent field of study that they provide courses in.

Here are a few entries:

```
##                               Name
## 1  *Daniel P. Aalberts
## 2          Sayaka BAe
## 3      Colin C. Adams
## 4 **Jeannie R Albrecht
## 5          *Laylah Ali
## 6 Kristopher D. Allen

##                               Title Degree
## 1                               Professor of Physics      BS
## 2                               Visiting Assistant Professor of Japanese      BA
## 3                               Thomas T. Read Professor of Mathematics      BS
## 4                               Associate Professor of Computer Science      BS
## 5                               Professor of Art      BA
## 6 Clay Artist in Residence in Jazz Activities and Lecturer in Music      BM

## Graduation Age      Undergraduate      Field
## 1      1989  47 MA Institute of Technology      Physics
## 2      1997  39      State University of NY      Japanese
## 3      1978  58 MA Institute of Technology      Mathematics
## 4      2001  35      Gettysburg College Computer Science
## 5      1991  45      Williams College      Art
## 6      1998  38      University of Hartford      Music
```

## Combining Years

Although the data was taken from the .pdfs for each academic year there is still an opportunity to look at the data sets together by using the `rbind()` function as shown below:

```
catalog12.13 <- rbind(catalog12, catalog13)
catalog11.12.13 <- rbind(catalog11, catalog12, catalog13)
dim(catalog12.13)
dim(catalog11.12.13)
```

The dimension output shows that the number of rows increases to 785 and then 1154, respectively. I would then be able to look at certain statistics across different academic years.

## Basics of Facultyeph

Most of the information below is simply utilizing existing R functions with the data that was parsed and organized. Let us start by looking at some basics of the `facultyeph` package. Some of the more obvious functions are `mean()`, `min()`, `max()` and `summary()`.

If we look at the average age of the Williams College faculty it comes out to be:

```
mean(catalog13$Age)
```

```
## [1] 48.98992
```

So the average age of the Williams College faculty in the penultimate academic year is approximately 49 years old.

The range of ages can be found through the `summary()` function and subtracting the max from the min.

```
summary(catalog13$Age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    25.00   38.00   48.00   48.99   59.00   78.00
```

```
summary(catalog12$Age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    27.00   39.00   49.00   49.28   59.00   92.00
```

- $78 - 25 = 53$  years is the range of faculty member in the 2013-2014 year.
- $92 - 27 = 65$  years is the range of faculty member in the 2012-2013 year.

## Youngest or Oldest?

We can find out the age of the youngest and oldest person for each year; above, we see that the the youngest for the '12-'13 and '13-'14 academic year is 27 and 25, respectively. However, this does not allow us to see *who* the youngest and oldest person is. Fortunately, I have created the functions `youngest()` and `oldest()` in order to determine who is the youngest and oldest.

The functions **youngest** and **oldest** both require a parsed catalog dataset and they act very similarly. In order to determine who is the youngest faculty member for a certain academic year (or more than one academic year) then we would simply type `youngest("catalog year dataset")`.

For example:

```
youngest(catalog13)
```

```
## [1] "Sarah A. Mirseyedi"
```

```
youngest(catalog12)
```

```
## [1] "Daniel Greenberg"
```

Here we can see that Sarah A. Mirsyedi is the youngest person in the '13-'14 academic year and Daniel Greenberg is the youngest person in the '12-'13 academic year.

We can find the youngest person across multiple years.

```
youngest(catalog12.13)
```

```
## [1] "Sarah A. Mirseyedi"
```

Here we see that the youngest person in the years '12-'14 is Sarah A. Mirsyedi, so we can then conclude that Professor Mirsyedi is younger than Professor Greenberg.

The same goes for oldest:

```
oldest(catalog13)
oldest(catalog12)
```

## Field of Study

The **facultyephs** package is also able to look at the field of study that each faculty member is in (or associated department/major). In this case, we can begin to look at relationships between faculty members in different departments. The figures below show histograms of two different fields and its respective age counts. I used the Economics and Philosophy field from the 2013-2014 academic year as examples.

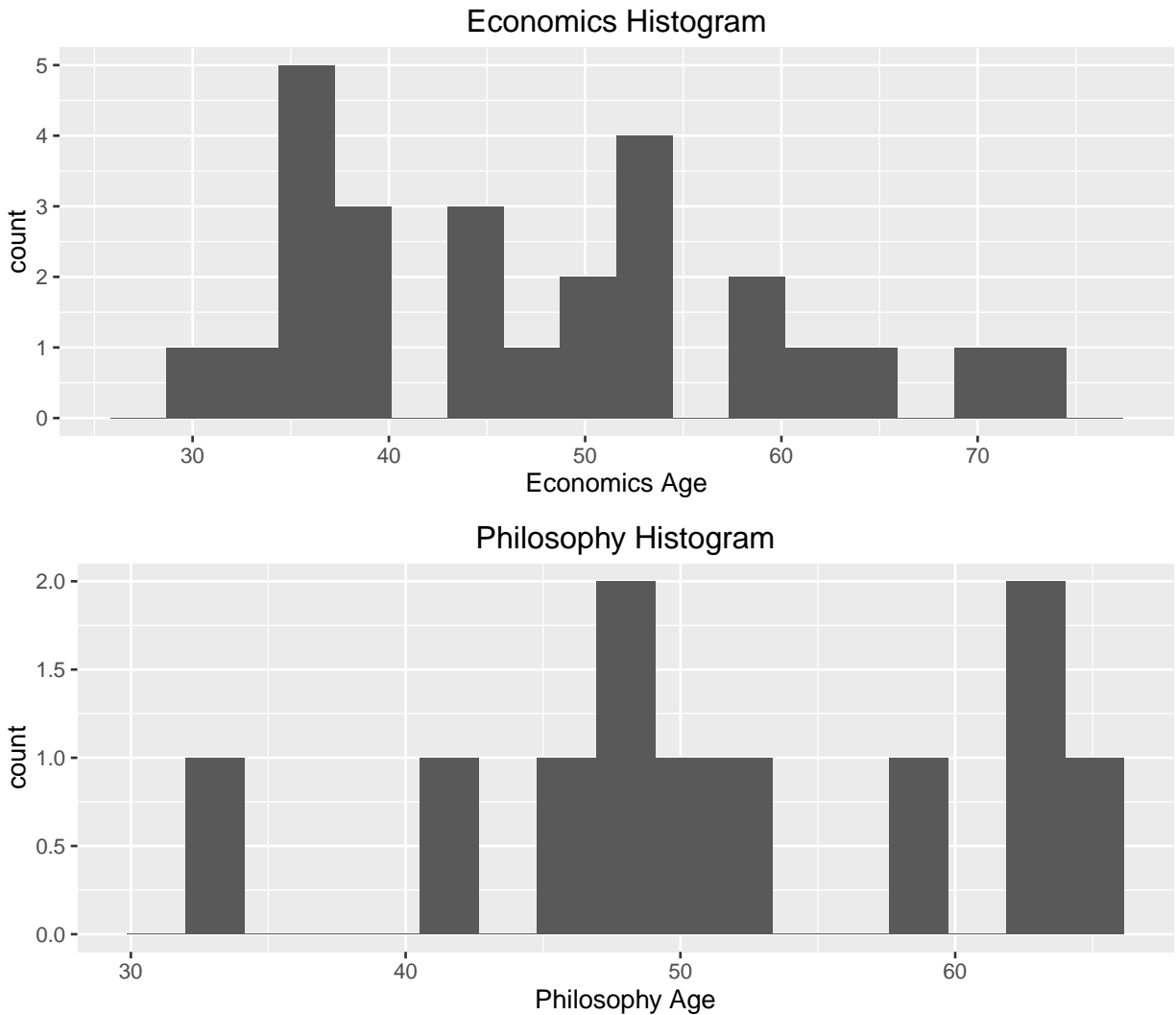


Figure 1: The figure shows a histogram of ages in the Economics and Philosophy department

The figure shows a histogram that indicates most of the Economics faculty members lie in an age range below 55. The histogram for the Philosophy faculty members is relatively evenly distributed, with not much of a concentration at a specific age. However, it would be better to look at density plots to see how each department's age is distributed compared to all of the other departments. Again, I will be using both Economics and Philosophy.

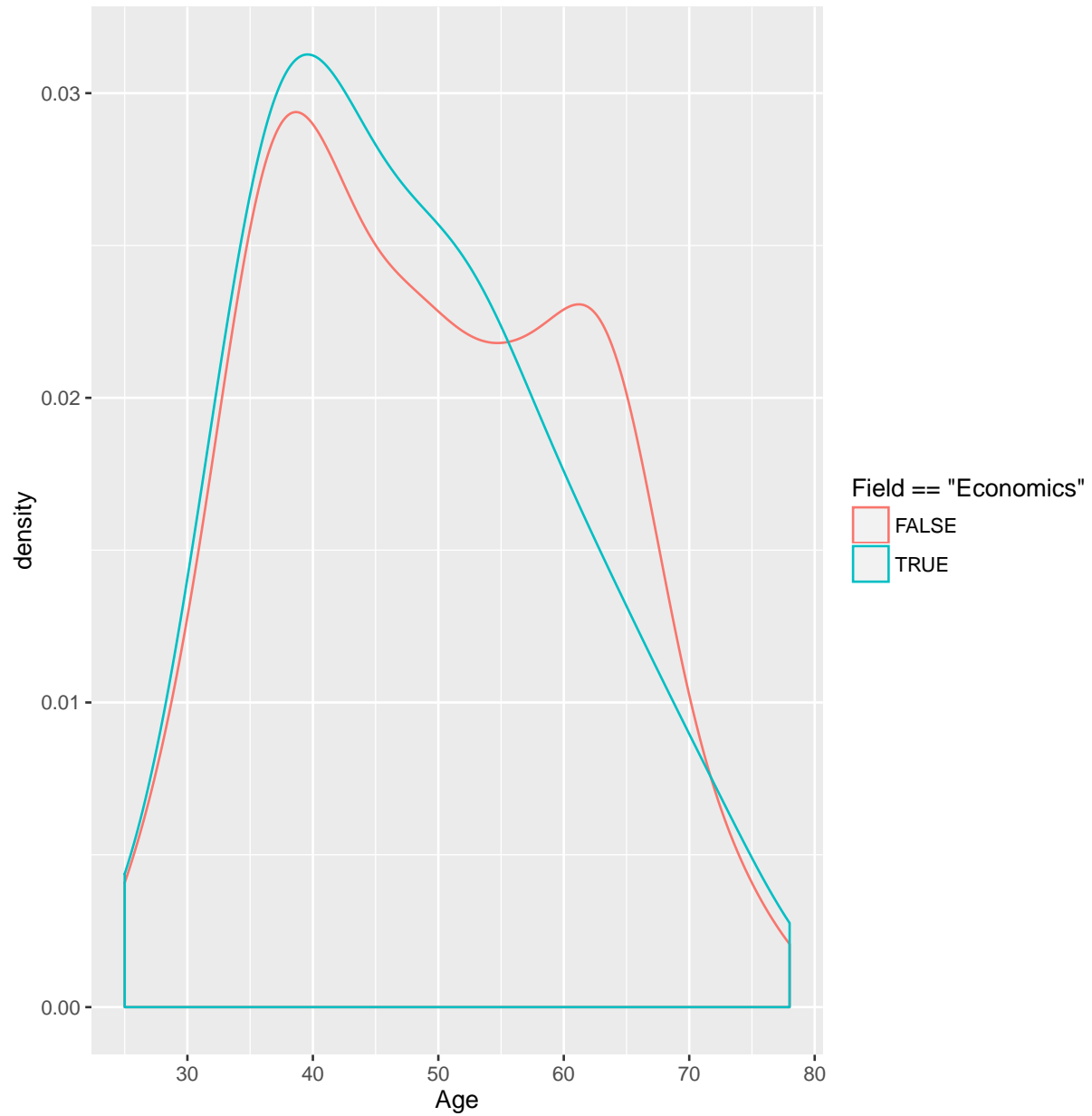


Figure 2: The Economics department seems to have more faculty members between the age of 40 and 45 than the rest of the departments. As shown by the intersection of the red and blue line, it seems that the Economics department has less members above the age of 55 than the other departments.

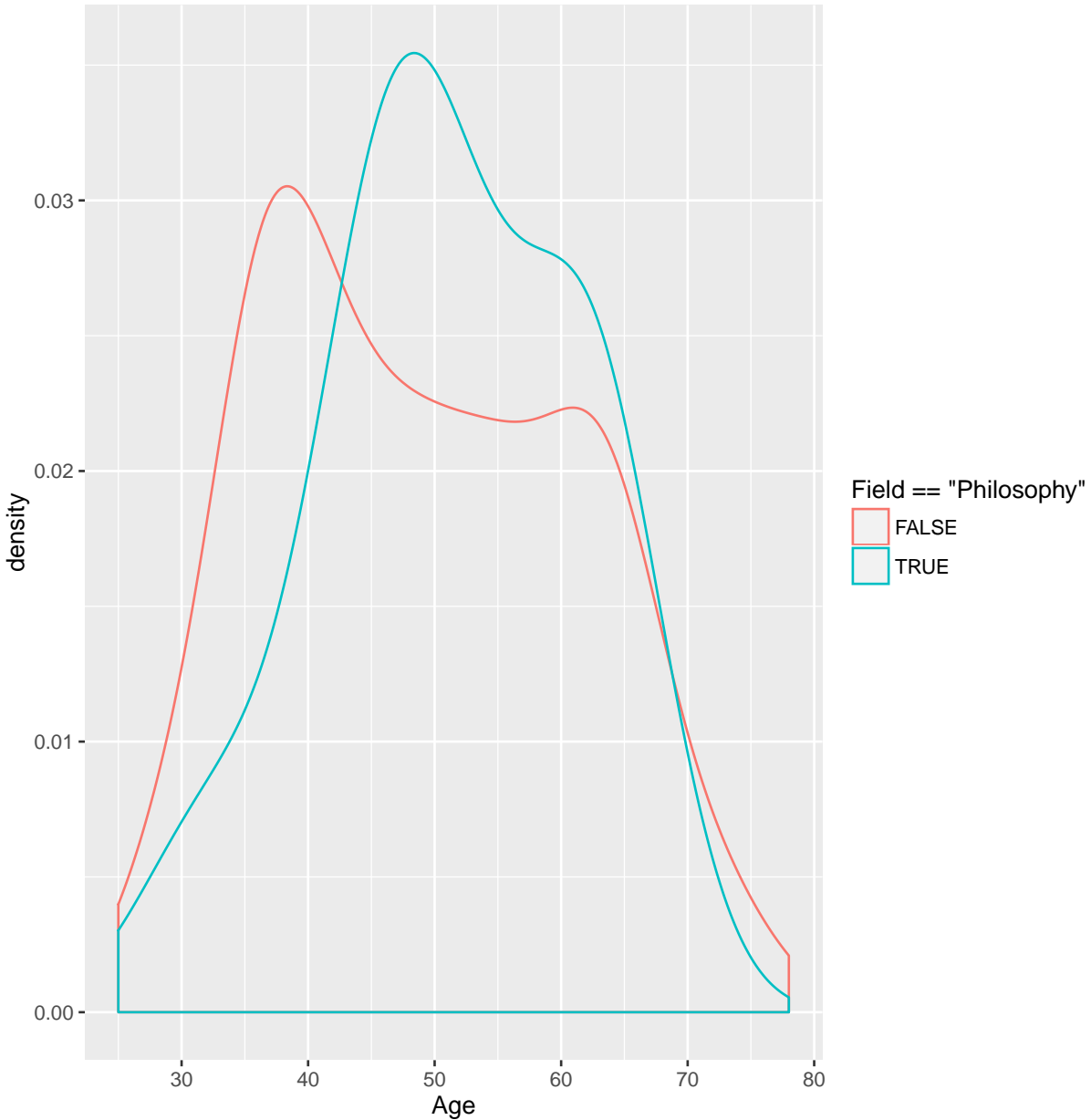


Figure 3: In the figure above, the density smooth plot indicates that the faculty members in the Philosophy field seem to have greater densities in the older age ranges (from 50 to 65) and lesser densities below 45. This indicates that Philosophy faculty members seem to have a greater distribution towards older ages.

If we compare the distribution of ages between the Economics and Philosophy department relative to all of the other departments, then the Economics department could be seen as “younger” than the Philosophy department.

## Degree

**facultyephs** can do a similar thing with the type of degree a faculty member acquired as it can with the field that a faculty member is in. In this case, we can begin to look at relationships between faculty members with different degrees: BA, BS, BFA, etc. You will also notice that Masters and PhD are included in the degree section. This is due to the fact that sometimes faculty members do not have information of Bachelors or a

Masters/Doctorate was the first degree listed in the catalog. For this example, we will be using the catalog from the 2010-2011 academic year. We could start with an interesting graphic that gives us a distribution with all of the different degrees subsetting by different colors, as shown below.

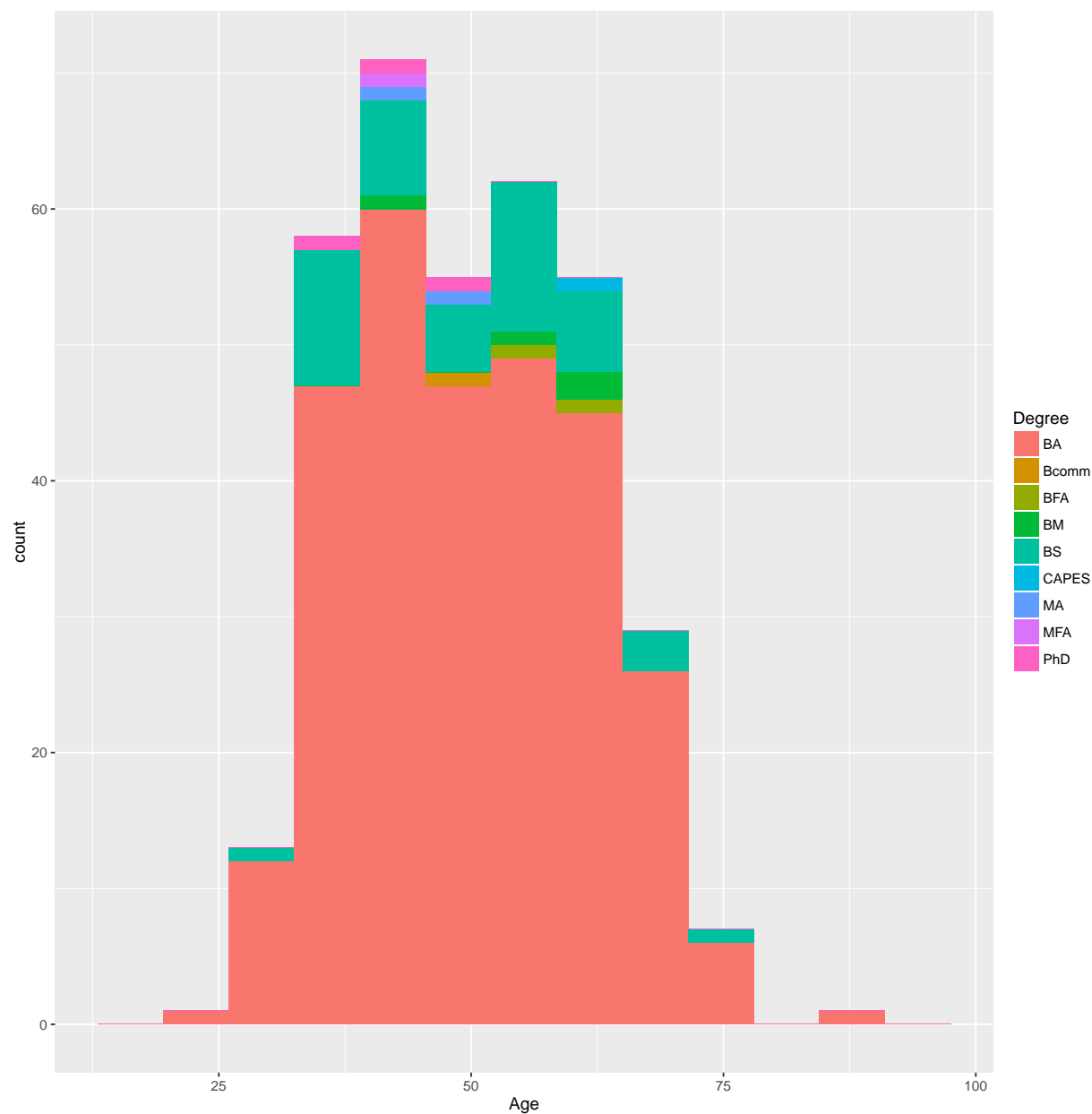


Figure 4: The figure above does not give us much information as it is cluttered and the counts are simply stacked– making it hard to compare the distribution of ages across degrees.

Let us look at the distribution of ages across different degrees in a more orderly fashion by using the figure below:

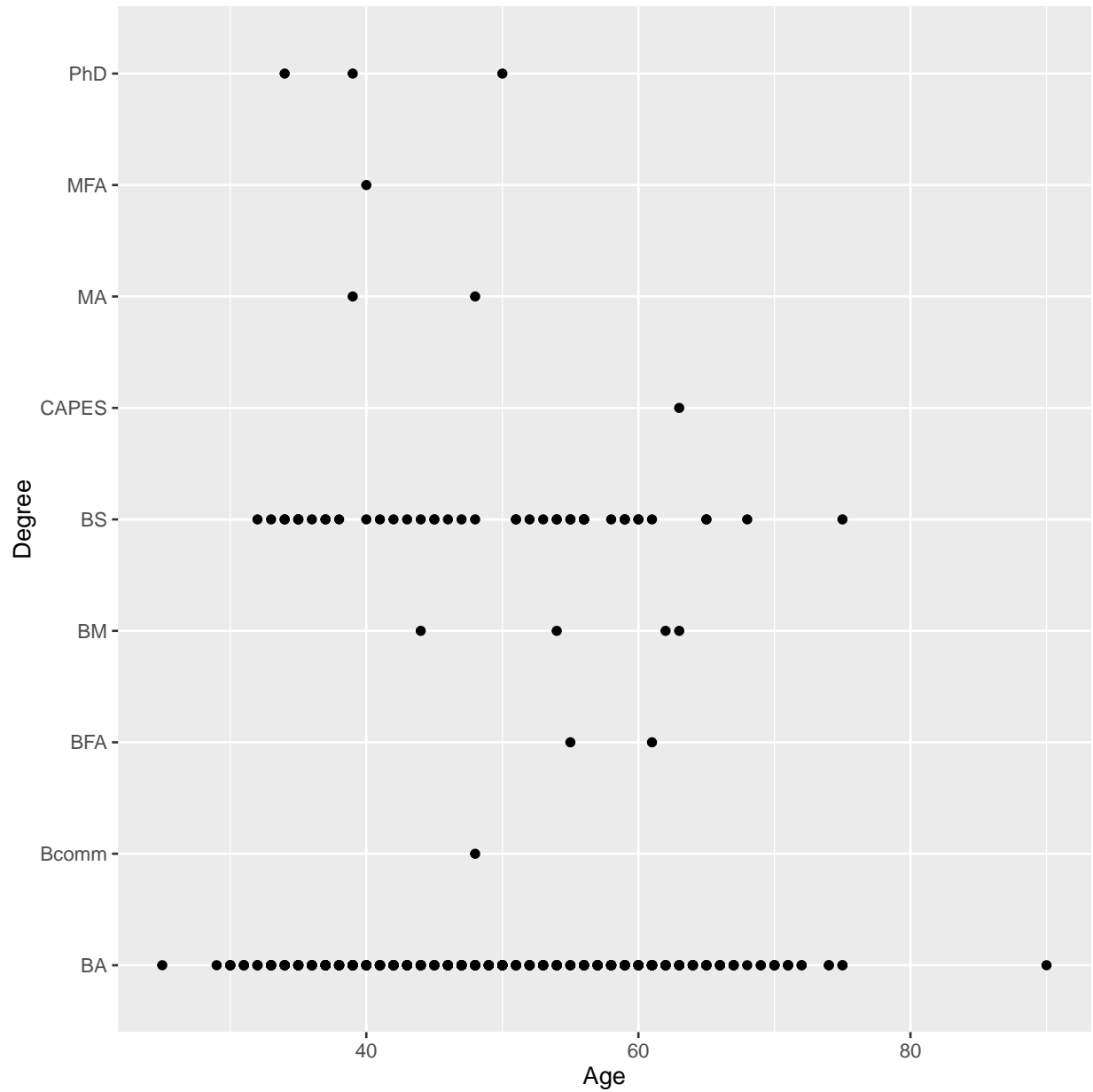


Figure 5: Here you can see that faculty members with a BA degree have a wider range of ages than those with a BS degree. We are also able to see that BM, BFA, Bcomm are very few and not as widely distributed.



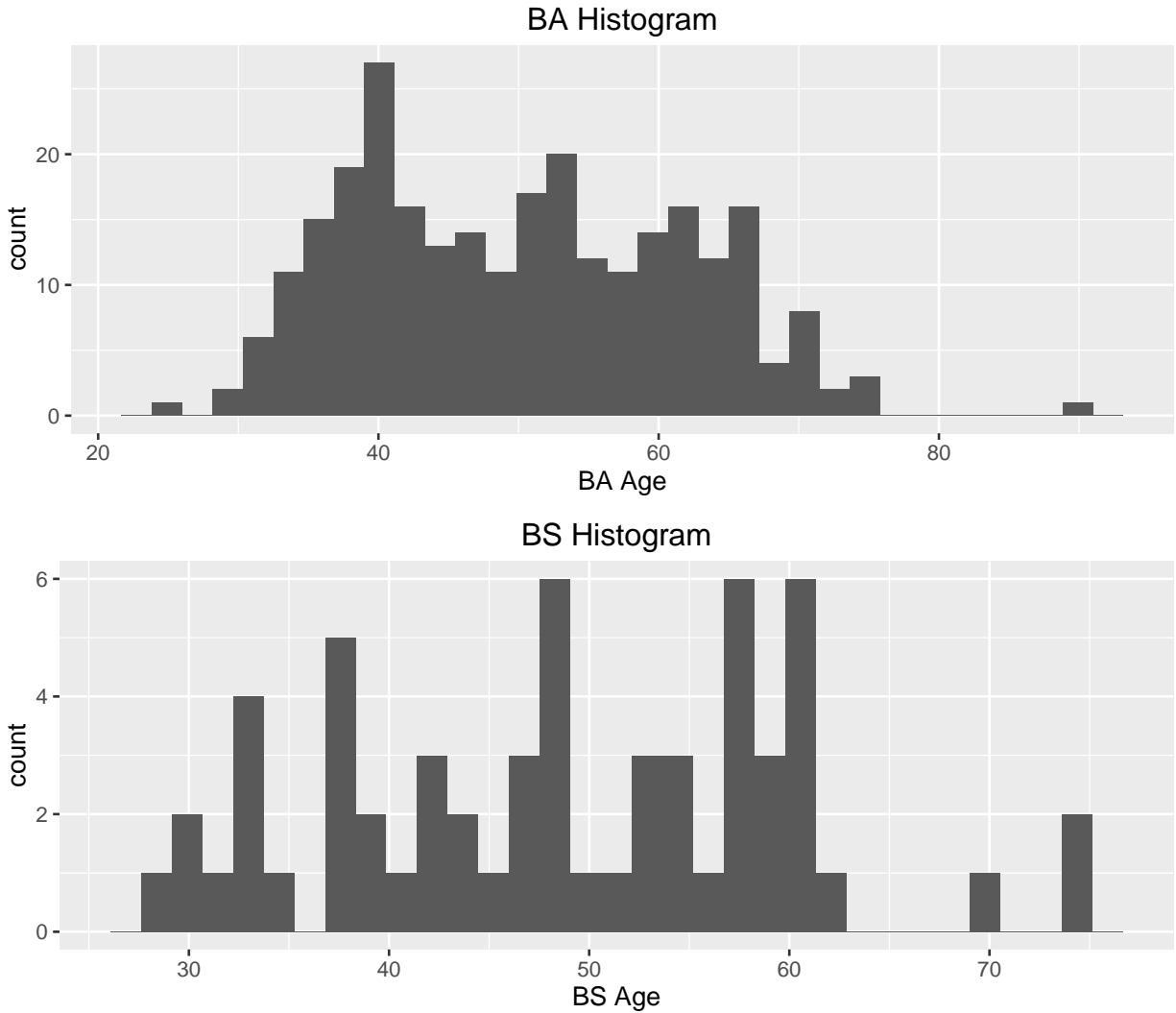


Figure 6: The figure shows a histogram of ages for faculty members who either received a BA or BS.

Ultimately, Williams faculty members have a lot more BA degrees than BS degrees, which makes sense as they are teaching at a liberal arts college that only offers BA degrees. In terms of the age distribution, the BS degree histogram shows a more uniform and unimodal distribution of ages as opposed to the BA degree histogram which suggests that most faculty members with BAs lie are under 50 years old.

We can also look at a intuitive hypothesis, such as the assumption that those who acquired a PhD are older.

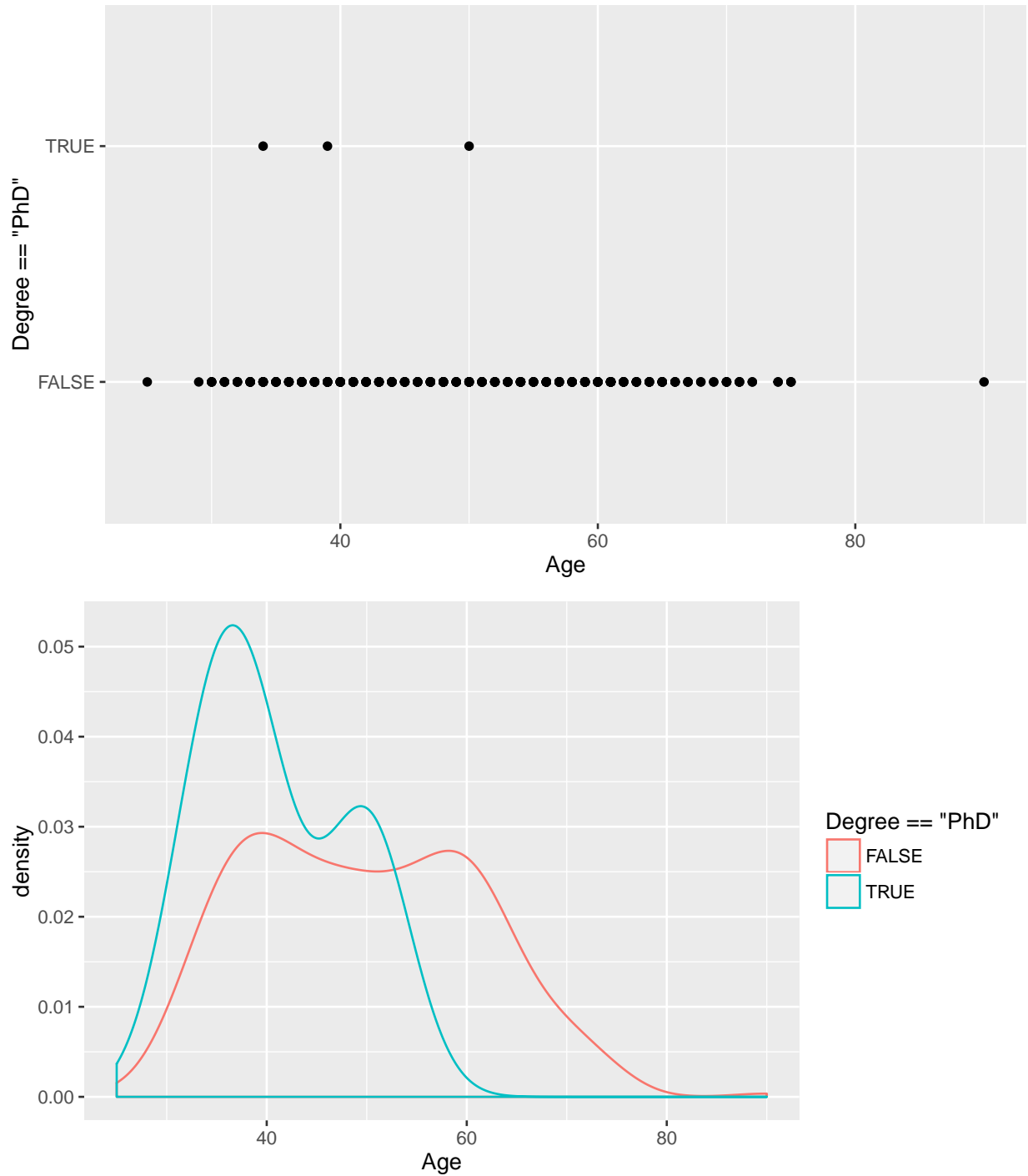


Figure 7: The figure above shows the distribution of faculty members with a PhD compared to all of the faculty members who do not. The top part of the figure gives us a general idea while the density plot in the bottom of the figure allows us to see in more detail.

Ultimately, it seems that faculty members with PhD age's are not as distributed towards high age values as we would expect.

## Conclusion

**facultyephs** is still in its preliminary stages and can have many more capabilities to address more complex questions regarding Williams faculty. **facultyephs** currently provides only a basic skeleton for more work to be carried out. This project could be cleaned and modified and become more fully developed to answer questions such as “How many faculty members are williams alumni?”, or perhaps “How many faculty members were on leave for at least one semester?”. However, if we look at the distribution of ages, whether it is according to field of study or degree, there is a lot to be learned about the Williams Faculty.