

EDA - Assignment 8

Aaron Niskin

September 20, 2016

Florida is a “right to know state.” While it may be unclear what this means, there are public sites which host information concerning who gets paid what. One such site can be found at

http://floridahasarighttoknow.myflorida.com/search_state_payroll

Visit the site and download all available salary data.

Prepare an R Markdown file which documents the steps you take in answering the following questions:

```
emp_data <- read.csv("emp.csv")
```

1. How many people in the state of Florida work for the State University System?

```
dim(emp_data)
```

```
## [1] 84315    12
```

But,

```
dim(unique(emp_data[,c("Position.Number", "Last.Name", "First.Name")]))
```

```
## [1] 62324     3
```

So about 84,315 people are listed, but a bunch are going to be duplicates (people who are paid for two

2. How many people in the State University System are full professors?

```
dim(unique(emp_data[emp_data$Class.Title == "PROFESSOR",  
                    c("Position.Number", "Last.Name", "First.Name")]))
```

```
## [1] 4476     3
```

About 4476, give or take a few errors here and there.

3. What is the median salary of a full professor in the State University System?

```
median(emp_data[emp_data$Class.Title == "PROFESSOR", "Annual.Salary"], na.rm = TRUE)
```

```
## [1] 80504.5
```

4. How many full professors are female?
5. Give some measure of confidence involving your results.

To address 4., it might be useful to construct or otherwise locate a list of female first names and think about join.