

Is there a Gender Pay Gap in the Florida's Government?

A quick analysis of Florida's State Payroll

Charles McGuinness, charles@mcguinness.us

December 12, 2014

Introduction

Is there is a persistent bias in the wages paid to male vs. female workers in the USA? According to the [Bureau of Labor Statistics](#), women in America make, on average, 81 cents for every dollar a man makes. People have argued if this a result of discrimination against women or are if there benign and innocent causes at play. Many [reasons for the gap](#) have been suggested, some of which are related to the different choice of careers made by men and women, and some reasons which are attributed to outright discrimination. Unfortunately, there are no definitive answers as of yet.

Part of the challenge of understanding the root causes of the pay gap is that it's hard to get good, detailed data on jobs and pay. Researchers often rely on [surveys](#) to discern trends and deduce reasons. But the use of incomplete data leads to interpretation and opens the door to the subjective bias of the researcher.

The state of Florida offers us a rare opportunity to do a more objective analysis across a large and varied employee base. The state provides a complete, downloadable database of its payroll as part of the Governor Scott's [accountability program](#). By looking at this data, we can examine the dimensions of pay differences in the Florida state government, and start to answer the question of whether Florida has a gender pay gap and, if so, how it is manifested in the state's payroll. I write "start", because there can be better answers to the question, but finding those better answers requires far more research and analysis than can be done just by looking at the data. Nonetheless, with the data we can at least see if there is prima facie evidence of bias and start to see the dimensions of the issue and where we can and cannot find signs of a gender pay gap.

Gender Pay Gap Throughout The Entire Workforce

The first way to look at the data is to examine the pay of men and women across all jobs in the state government. To do that, I'll look at the pay for salaried, full time employees. Here is a histogram of the salaries for men and women: it shows, in a glance, what they're getting paid. Across the x-axis (horizontal) is the range of salaries divided up into small ranges. On the y-axis (vertical), it shows the number of people getting paid that amount. Taller means more people. Further to the right means better paid. This gives us a rough idea of how men and women are paid by the State of Florida:



And we can see that female pay rates are lower than male's, in general. Let's do some elementary stats to quantify that:

Male_Mean	Male_Number	Female_Mean	Female_Number
45040	40567	39789	53415

The average woman's pay is \$5252 lower than the average man's. Stated differently, the average woman makes 88 cents for every dollar the average man makes.

This number is fairly consistent with other estimates of the [US gender pay gap](#) (although better). According to [this report](#), the median (not mean) private sector earnings for men in Florida is \$41K compared to women's \$34K. The average pay at the state of Florida has a narrower gap, but generally reflects the same trends as we see in the private sector in Florida as well.

This disparity is not proof of bias, of course. Just the different types of jobs, seniority, and training the different sexes have can account for the differences. The gold standard for detecting bias is whether men and women are paid differently all other relevant things being equal. (What is relevant is, of course, another contentious topic.)

Bias Within Job Classes

The data from Florida also includes job classes, with 3163 different job classes being listed. Example job classifications include:

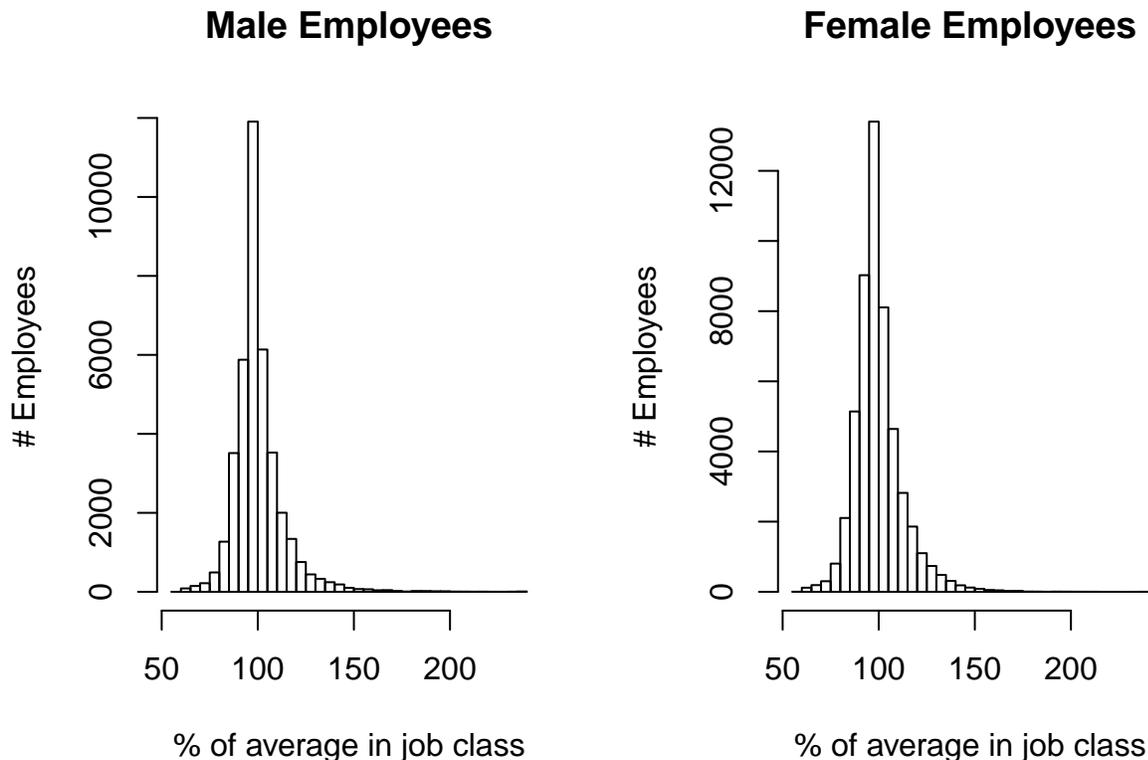
- SAFETY PROGRAM MANAGER - SES
- REGIONAL OPERATIONS MANAGER
- HUMAN RESOURCE TECHNICIAN II
- DIRECTOR OF STAFF & CURR. DEVELOPMENT
- GOVERNOR

(Interestingly, the Governor is the lowest paid full time salaried employee in Florida, getting \$0.12 per year)

The use of titles provides us another, and perhaps better way of looking at the data. Now we can ask whether on average, within each job class, women are paid less than the men.

To do this, I compute the average pay in each job class across men and women. Then I divide each man's pay and each woman's pay by the average pay in that class to compute each person's "normalized" salary. A normalized salary below 100% indicates the person is getting paid less than average for people in the same job. Conversely, a number above indicates the person is getting paid more than average.

Here is a histogram of the normalized salaries for men and women. Like the histogram above, it shows, in a glance, how men and women are getting paid. But in this case it shows what each person's salary is in comparison to what other people in the same job class are paid. 100% means the person is being paid the average wage. 110% means the person is being paid more, while 90% means the person is being paid less. If women are getting much less than 100%, it means that there is a pay gap within individual, specific jobs.



Although the chart looks like there might be some bias, it is hard to tell. To get a more precise answer, we average all the men's normalized salary and all the of the women's normalized salary to get a view of how the average male and female employee of the State of Florida is doing.

The numbers are:

Mens.Mean	Womens.Mean
100.5	99.79

By looking at this data, we can see that, for job classifications that have both male and female employees, the female employees on average receive 99 cents for every \$1 a male makes – virtually indistinguishable.

Clearly, for men and women performing the precise same job (as determined by the specific job class they are in), they are being paid essentially the same. This is good, but it leaves us with no explanation for the overall gender pay gap in the state’s payroll.

Bias Within Similar Job Classes

Even though the pay within job classes was unbiased, there could still be a source of bias if women are not promoted as frequently through the levels as men are. That would result in women being paid less for the same job by being slotted into lower paying levels.

Here is an example of job classes with levels in the payroll database:

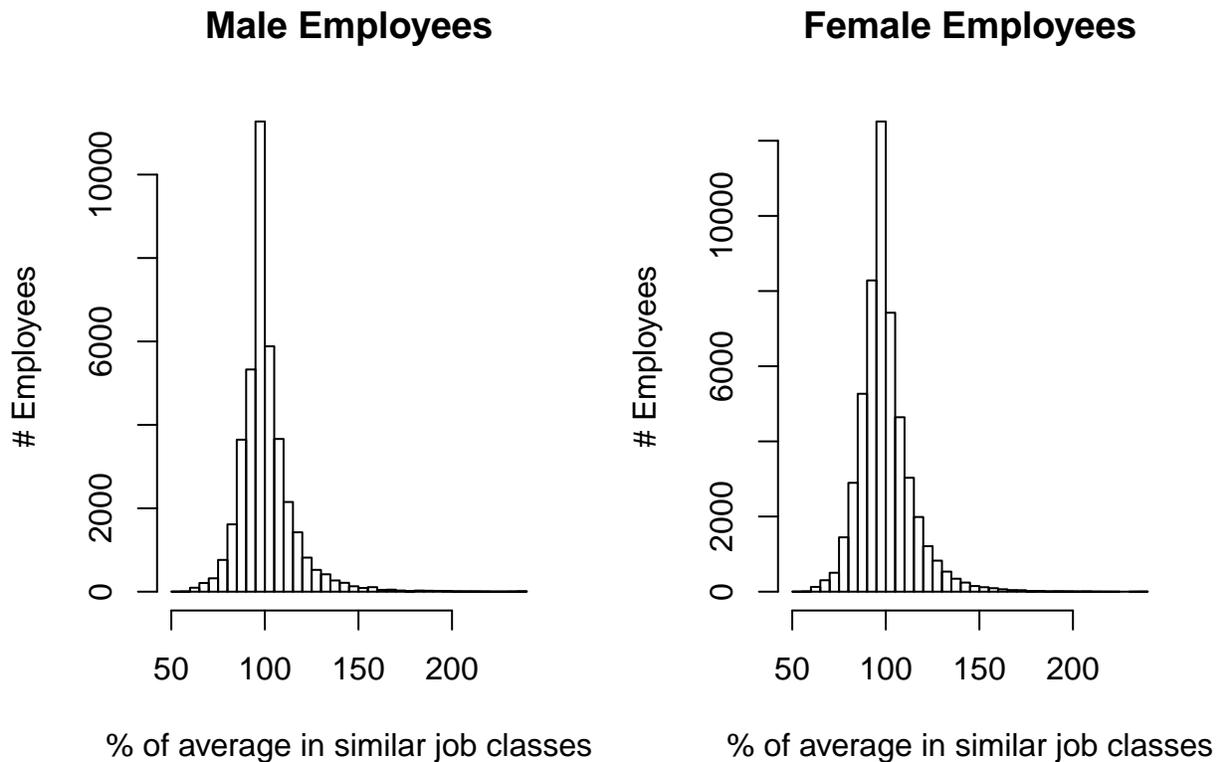
- GOVERNMENT OPERATIONS CONSULTANT I
- GOVERNMENT OPERATIONS CONSULTANT II
- GOVERNMENT OPERATIONS CONSULTANT III

You could imagine that, if there were a pay bias, women would be disproportionately assigned to the level I jobs while men were over-represented in the level III jobs.

An easy way to check is to take the average pay for all levels of a job class (in this example, anyone with a “GOVERNMENT OPERATIONS CONSULTANT” job regardless of level), and see if, within these broader classes, women were being paid less than men.

Let’s run the same analysis as the previous section, but this time consider all I, II, III, etc. jobs to be the same basic job. We want to see if we can now find a gender pay gap because women are not promoted as fast or often.

Here is the distribution of normalized salaries for men and women, when the normalization is done against the average for all levels of a title. As before, 100 means being paid the same as others in the same group of job classes, >100 means more, and <100 means less:



As it turns out, for job classes that have both male and female employees, the female employees on average receive 99 cents for every \$1 a male makes – which make the pay virtually indistinguishable.

The full statistics are:

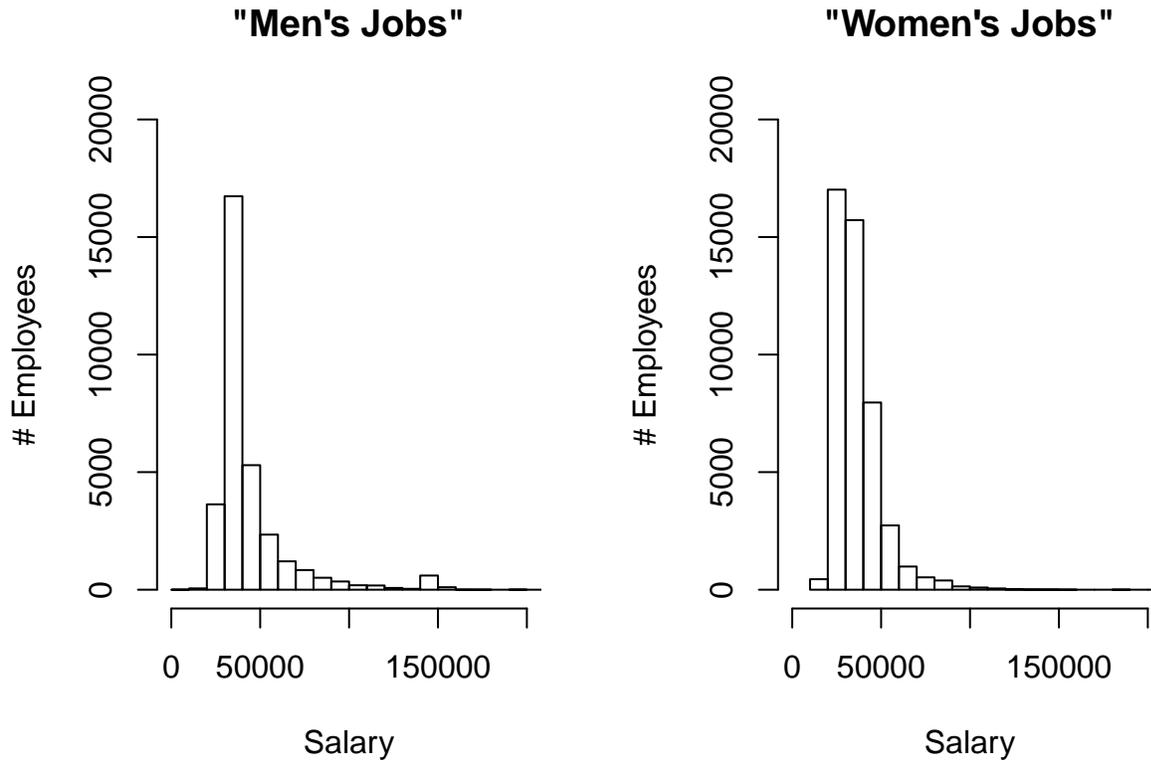
Mens.Average	Womens.Average
100.7	99.67

We still do not see an explanation for gender pay gap; men and women who are broadly doing the same job are being paid essentially the same. The state has done a good job of ensuring equity in pay within job classes and promotion through the job classes for both men and women.

Bias Between Heavily Male and Female Job Categories

We see that men and women are generally paid the same within job classes. And yet there's still an overall pay gap between male and female state employees. So what is causing that? There's a logical suspicion: that job classes that have more female employees tend to be paid worse than those that have more men.

Let's test that by looking at jobs that are strongly female or strongly male. Using a criteria that we only look at job classes that are at least 2/3rds male or at least 2/3rds female, we can see the following:



There's a strong bias in pay between male dominated and female dominated job classes. The average pay for each sex is:

Mens.Average	Womens.Average
43817	36111

So when you compare jobs that are heavily female versus jobs that are heavily male, we see that women earn 82 cents for every dollar a man earns.

It turns out the type of the job you do drives your pay (no surprise there), and that if your job is mostly done by women you will be paid less.

A Closer Look at Male and Female Dominated Job Classes

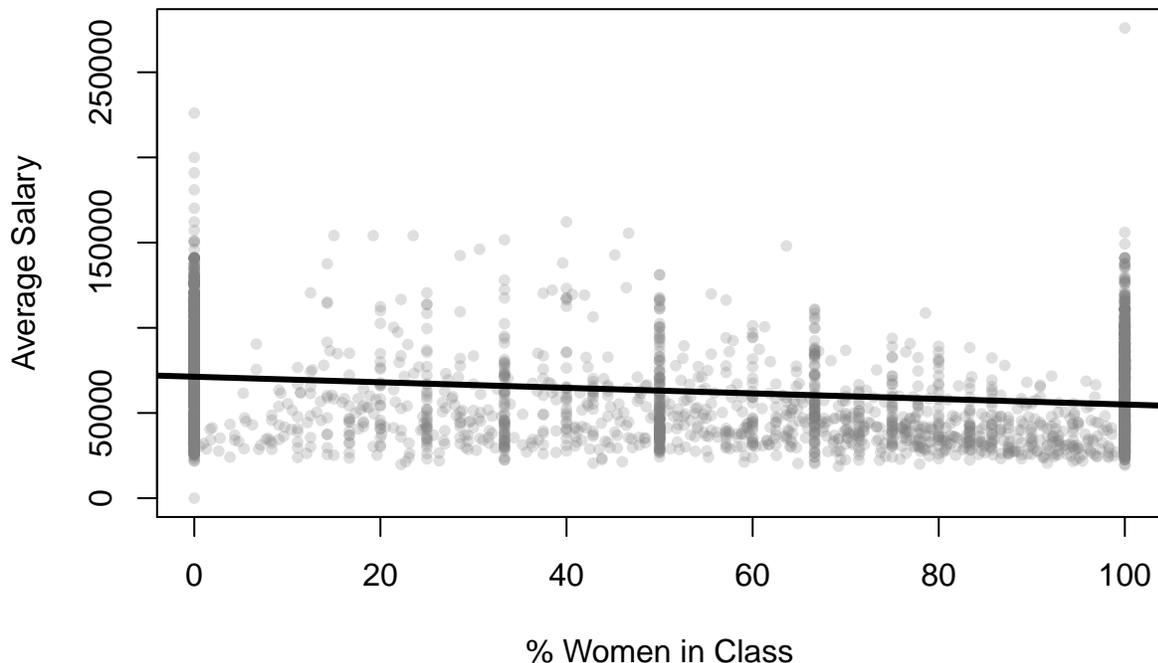
What are these heavily male and heavily female job categories? Could it explain the pay gap in a way that does not imply bias?

Here are the top 5 job classes that are “most male” and “most female”:

Top_Womens	Top_Mens
SENIOR CLERK	CORRECTIONAL OFFICER
ECONOMIC SELF-SUFFICIENCY SPECIALIST I	CORRECTIONAL OFFICER SERGEANT
REVENUE SPECIALIST II	LAW ENFORCEMENT OFFICER
STAFF ASSISTANT	FOREST RANGER
REGISTERED NURSE SPECIALIST	VOCATIONAL INSTRUCTOR III - F/C

It seems that much of the discrepancy can be attributed to the dominant male presence in law enforcement jobs. (Note that “VOCATIONAL INSTRUCTOR III - F/C” involves “supervising an inmate work crew” – still a sort of law enforcement job) This pay gap appears to be an example of the [danger wage premium](#) concept (even if that concept is still contentious).

When you look at the job classes over all, sorted by the imbalance of women in the role, we can see a lowering of the salaries as we move from left (job classes with all men) to right (job classes with all women). This chart plots the average salary paid in all the various job classes as compared to the % of women in the job. Dots on the far left represent all-male job classes, while dots on the left represent all-female job classes:



(As an aside, the one outlier in the upper right hand corner is Pamela L. Steward, Commissioner of Education, whose salary is \$276,000 per year. And since she’s the only one in that spot, it’s 100% female.)

The line through the data is the linear regression. It shows the overall trend the data. Knowing the regression means that, if you're going into a job class that is 100% male, you can expect to be paid \$71302 on average, while if you're going into a job class that's 100% female you can expect to be paid \$54986.

The statistics show that job classes pay better, on average, the more male-dominated they are.

Conclusions

As with most other studies of the gender pay gap, looking at the state of Florida's payroll leaves us with many questions even if also leaves us with a better understanding of the situation.

- We have seen that across all state employees, women receive 88 cents to every \$1 male employees make.
- But when we look within various job classes, however, we see that there is broad parity in the pay men and women receive for doing the same job. Even if we look across all levels of a job class, the pay between men and women is bias-free.
- But when we sort job classes into those that are predominantly female vs. predominantly male, we see the gender pay gap open its widest: "men's work" pays substantially better than "women's work", with women only earning 82 cents for every dollar the men make.

There's a common saying that "correlation is not causation". So it is rash to draw any conclusions apart from there being an obvious gender pay gap. The gap between male and female job classes could be a result of completely legitimate reasons (the danger wage premium mentioned above), or it could be for reasons unrelated to gender per se (e.g., law enforcement workers are better able to bargain/lobby for higher pay), or it could be the result of unfairly valuing "women's work" less. Or a combination of any of these.

Certainly, the State of Florida appears to be doing a good job of ensuring that men and women doing the same work are getting paid the same. That part is commendable. But the data also suggest that the state should look carefully at pay differences between traditionally male and female jobs to see if and where women are being unfairly underpaid even though they bring comparable levels of skill, training, and experience to the job. Similarly, the state should make greater efforts to recruit women into traditionally male jobs (and men into traditionally female jobs as well.) Perhaps by pushing (hiring more women into non-traditional roles) and pulling (raising salaries in traditionally female roles) the state can start close the gap.

Appendix

Methodology

The State of Florida provide weekly salary updates on its [official site](#). There, you can download the complete payroll (“Export All Salaries”). This report starts with this data and then enriches it to prepare for the analysis.

The enrichment takes two steps:

1. Some simple clean up of the data (trivial)
2. Estimating the gender of each employee (non-trivial)

With the enriched data, this report looks at gender differences across the entire state, as well as differences within each job classification. For the purposes of this analysis, only employees who are full-time and salaried were looked at (a vast majority of state employees), to order to avoid distortions caused by part-time work or the variable pay earned by hourly workers (because of differences in hours worked).

For this analysis, the important items from the data are:

- First Name, from which we deduce the gender of each employee (see below)
- Class Title, which is the job classification of each employee. We use this to group employees into the same or similar jobs.
- Salary, the yearly pay
- Full Part Time, a flag that indicates whether the employee is a full or part timer. We only look at full time employees to avoid distortions in pay caused by variable hours
- Employee Type, which indicates whether the worker is paid a set salary or is paid hourly. We look only at salaried workers to remove distortions caused by variable hours and/or overtime.

Estimating Gender

Since the question at hand is the pay differences for men and women, a key piece of information is knowing which employees are male and which are female. That information, however, is not present in the data downloaded from the state. Thus, the first step of the analysis is to estimate the gender of each employee from their first name.

The actual mapping of first names to gender is a fairly straight-forward process, driven by a database of names. The source of this database is a Social Security Administration published list of the 1000 most popular baby names from each year. By downloading this data and aggregating it across the years, one can build up with a very comprehensive list of names and the sexes of those names.

For this analysis, names from 1945 to 2013 were used, giving a database with the following # of unique names:

Total_Male_Names	Total_Female_Names
36031	60154

The list of names requires a bit of processing, as it turns out that many first names have been given to both male and female babies. Usually, there is such an overwhelming trend in the use that the name is unambiguously male or female. For example, for Charles, my name, the counts of male vs. female babies:

Male_Charles	Female_Charles
1291047	6054

That’s pretty definitive. But for other names, like “Pat”, not so much:

Male_Pat	Female_Pat
14954	15589

For this analysis, I have used a rule that 90% or more of the babies with a given name have to be male (or female) to decide the name is unambiguously male or female. There’s a little room for error (at worst, a 10% chance of mis-identification), but that’s going to affect only a very small number of names on the edge. But if a name falls below 90% male or female, I declare it to be androgynous (like “Pat” is). If the database doesn’t have the name, I try a few simple transformations to find it. For example, sometimes the first name field in the employee data has both the first and middle name in it; so when I find a compound name I test each component of it. “Mary Grace” may not be in the database, but “Mary” is.

There are some names that remain where we just don’t know. Lots of initials are in the database. Other names are very rare or have spelling variations not recognized. Fortunately, the vast majority of employees can be assigned a gender. Here are the final counts:

Males	Females	Androgenous	Unknown
44472	59367	4800	5174

We could loosen the 90% restriction (perhaps to 80%) to move a few more androgynous names into definitive ones, but I do not think the trade-off is beneficial

Analysis

I have attempted to document clearly the statistics used in this paper. However, English prose is not a good substitute for precise formula or algorithms. If there is any question or ambiguity in the document, it can be resolved by reading through the code which is publicly available (see the next section).

Author's Note

This analysis was motivated by a desire to see what could be learned from the open data provided by the state of Florida. I appreciate that given the results I have generated, some will attempt to attribute bias or agenda.

Since nothing I can say will convince a skeptic, I have one thing to say to people who disagree with me. . .

Fork me!

This project has been done with 100% transparency. I have followed the principles of reproducible research, which means all of the data and code is open source and available to you. The analysis is written in the statistical programming language R, including the verbiage of the report (using the knitr package). Thus, when you rerun the program, it generates not only the data and statistic, but the actual text of the report.

All of the data comes from public sources. The code is available online. Thus, you can download the complete analysis and run it yourself. And if, when reading through my code, you spot any errors in my logic you can fix it. Or you can try a different analysis and argue why its results are more relevant. Whatever you like.

You can find the complete sources for this analysis on [Github](https://github.com/cm McGuinness/Florida-State-Gender-Pay-Gap) at <https://github.com/cm McGuinness/Florida-State-Gender-Pay-Gap> .

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), so there are no restrictions on what you do with this (apart from the attribution requirements). Build your own using this code. Or look at a different state's payroll.