

Wage Gap in Salaries by Predicted Gender at UNC-Chapel Hill

Caleb Sigmon

Abstract—The motivation for this project is to answer the question of whether there is evidence of a wage gap at UNC-Chapel Hill in general between faculty of different genders, and if present, then specifically, which departments are the most dramatic examples of this. This study used R and JavaScript charts to examine the topic and found there is indeed a general trend of male faculty making higher salaries.

1 INTRODUCTION

UNC-Chapel Hill is a public university so its salaries are guaranteed access to the public under the Freedom of Information Act. The UNC System maintains a website for the public to access their database of salaries, making inquiries simpler and faster for both parties - the inquirer and the data provider - than the typical process of submitting a FOIA request. Looking directly at the salaries in the database is a luxury not every state provides for their own universities, but still it lacks details such as race and gender that would aid in investigating inequity in the system.

2 PRELIMINARY RESEARCH

In the general workforce in the US, “Women earn an average of 16% less than men.”[1]

Academia is no better, with professors that are women making 82 cents for every dollar earned by their male counterparts.[2]

In North Carolina, the average salary for full-time faculty is 104.9k; for professors that are tenured or on the tenure track it’s 112.6k. A last significant statistic: the national average for male salaries is 130.8k, but for female salaries it’s 109.9k.[3]

One explanation may be the difference in ranks held by men and women. UNC’s Office of Institutional Research and Assessment maintains a dashboard of statistics across genders for professor ranks and tenure status.[4] While initially, this may seem to solve the problem and enable easy investigation into inequity - and to some degree it does - the dashboard lacks the correlating salaries and a proper degree of granularity. It only lists the

statistics across rank and gender, but not by department. This makes it much more difficult to answer the question of which departments experience the worst wage gap.

Before arriving on the methods used to make estimates, the HR office in the department of Business and the Office of Research were contacted to investigate if it is possible to access the information on how much men and women make in a particular department. Quite simply, the answer was reported to be no.

OIRA, on its dashboard named “Permanent Full-Time Faculty by Tenure Status and Gender or Race/Ethnicity,” does show some key statistics, like that overall there are more female full-time faculty than male.

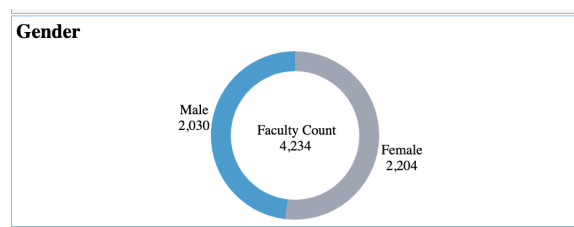


Fig 2.1 Faculty gender ratio - all full-time faculty

In contrast, there is a greater difference in the opposite direction for tenured professors, with male professors outnumbering female professors by 276.

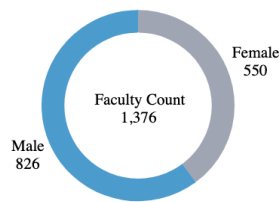


Fig 2.2 Faculty gender ratio - tenured faculty

On the tenure track, however, female professors outnumber male ones again.

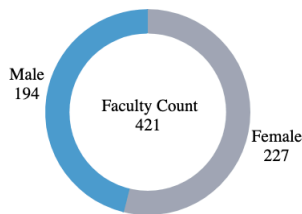


Fig 2.3 Faculty gender ratios - tenure track

The fixed-term professors have the most dramatic ratio of all, also led by female professors.

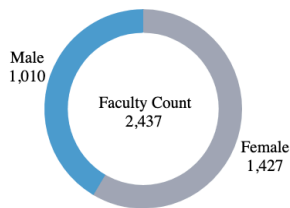


Fig 2.4 Faculty gender ratios - fixed-term

Since tenure status corresponds to higher pay, these figures may provide some optimism that as the tenure track professors continue on their track it will lead to a higher equity in pay between male and female professors.

What's interesting about all this is that it goes against expectations from national statistics. "Women make up the majority of nontenure-track lecturers and instructors across institutions, but only 44% of tenure-track faculty and 36% of full professors." [5]

At UNC-Chapel Hill, according to the Trustee Policy definitions, [6] faculty ranks primarily consist of four main positions: Professor, Associate

Professor, Assistant Professor, and Instructor. These ranks encompass both tenure-track/tenured positions and fixed-term appointments, with the latter often distinguished by modifiers such as "research" or "clinical."

- **Professor Rank:** This represents the highest level of academic achievement and job security. Both promotion to Professor and initial appointment mean getting permanent tenure.
- **Associate Professor Rank:** Promotion to Associate Professor also grants permanent tenure. Initial appointments are typically probationary for five years, but permanent tenure can be granted upon appointment with department approval.
- **Assistant Professor Rank:** Initial appointment as an Assistant Professor is usually probationary for four years. Reappointment decisions are made periodically, and successful completion of the probationary period leads to permanent tenure.
- **Instructor Rank:** Instructors are appointed with the expectation of progression to higher ranks. Initial appointments are for one year and can be renewed for up to four terms. Promotion to Assistant Professor may occur during this period.

Because of how it is defined in the policies in UNC I am placing instructors under the broader category of "tenure track" since they are expected to be a rank that lasts one year and after that year would ordinarily rise to the rank of assistant professor.

3 METHOD

3.1 Cleaning UNC System Data

The project uses open-source government data from the UNC System salary database. [7] The data was last updated on March 31, 2024, and for this was exported in a Microsoft Excel format. Each row lists such attributes as first and last name, salary, title, and department.

Since this project wants to ask questions of the data from different dimensions - meaning it needs to be grouped by different attributes - it was deemed

expedient to solve some preliminary grouping in the language R.

The R “dplyr” package was installed to perform several operations on the data. First, it was filtered from all the salaries in the UNC System to just those corresponding to UNC-CH. Next it was filtered to just include people whose titles include “Professor”, “Prof”, or “Instructor”. Since these are the terms defined to encompass the faculty types given in the UNC Chapel Hill Tenure Policies and Procedures. As a brief aside, the policies also mention a few other titles for fixed-term faculty that were not found in the UNC System salaries database under the primary working title columns (These were “Lecturer”, “Senior Lecturer”, “Artist in Residence”, and “Writer and Residence”. The term “Lecturer” did appear in a different column, “Job Category”, but every case of a lecturer in that column corresponded to a value of “Professor of the Practice” in the primary working title column, and vice-versa.

After being filtered down to the full-time faculty professors that are either tenured, tenure track or fixed term, two data frames were made to be converted into JSON objects for the web dashboard. For both, the data was cleaned to combine department groups that have very similar names into their respective joint departments. For example, there are over 25 “departments” that begin with “Psychiatry” then some specific subfield. It was deemed sensible to combine these kinds of departments to prevent skew from extremely small departments. See below for a full listing of the departments beginning with or containing the same broader term that were combined. Basic regular expression terminology to note is “^” means beginning with and “*” means any amount of any characters. So “grepl(*Health Sciences*)” means any title that contains the substring Health Sciences.

```
72 # Combine like departments, which begin with one of the following words.
73 # Since many of these contain only 1-3 people and are subdepartments.
74 custom_group <- function(text) {
75   case_when(
76     grepl("^ASOD", text) ~ "ASOD",
77     grepl("^Medicine", text) ~ "Medicine",
78     grepl("^Neurology", text) ~ "Neurology",
79     grepl("^Anesth", text) ~ "Anesthesiology",
80     grepl("^Biomedical Engineering", text) ~ "Biomedical Engineering",
81     grepl("^Biostatistics", text) ~ "Biostatistics",
82     grepl("^Dermatology", text) ~ "Dermatology",
83     grepl("^Emergency Medicine", text) ~ "Emergency Medicine",
84     grepl("^Psychiatry", text) ~ "Psychiatry",
85     grepl("^ENT", text) ~ "ENT",
86     grepl("^*Health Sciences*", text) ~ "Health Sciences",
87     grepl("^Med", text) ~ "Medicine",
88     grepl("^OBGYN", text) ~ "OBGYN",
89     grepl("^Ophthalmology", text) ~ "Ophthalmology",
90     grepl("^Orthopaedics", text) ~ "Orthopaedics",
91     grepl("^Pathology Lab Med", text) ~ "Pathology Lab Med",
92     grepl("^Peds|^Pediatric", text) ~ "Pediatrics",
93     grepl("^Radiation Oncology", text) ~ "Radiation Oncology",
94     grepl("^Radiology", text) ~ "Radiology",
95     grepl("^Surgery", text) ~ "Surgery",
96     grepl("^SOP", text) ~ "SOP",
97     grepl("^Urology", text) ~ "Urology",
98     grepl("^*Nursing*", text) ~ "School of Nursing",
99   )
100 }
101 TRUE ~ text
102 }
103 }
```

Fig 3.1 R code for combined departments

The first data frame to JSON exported was simply 2 columns, “Department,” and “n”, with n corresponding to the number of faculty in that department. This was exported in the project files as “depts_by_n.json”. The second export is just every row of the unc salaries data frame but mutated to now include the predicted gender.

3.2 Predicting Gender

To predict the gender where none was included in the data, a simple R package “predict_race”[8] was installed and used. The package includes a “predict_gender” function that simply takes in a person’s first name as input and matches it to one of four options.

1. Male
2. Female
3. Female, male
4. NA

The package makes its predictions by matching the name to US Social Security Administration records and finding if there are more men or women listed with that name. If none are found it returns NA, and if an equal number it returns “female, male”.

3.2.1 Accuracy

An important caveat when considering this research as evidence of inequity is that these are gender estimates. That being said, the gender ratios from the OIRA charts mentioned previously lend support to

the accuracy of the gender ratios predicted with the R package.

After exporting the necessary objects, several chunks were devoted to checking the accuracy of the ratios. For instance, OIRA lists the counts of all female faculty at 2,204 and all male faculty at 2,030. This ratio of women to men comes out to 1.085714. In the predicted gender data, there are 2,052 professors predicted as women, 1,880 predicted as men, (and 329 professor names that were unable to render a gender prediction) making the female-to-male ratio 1.091489, just about 0.01 off from the actual ratio.

The same work was done for tenure track and fixed-term professors, comparing the estimated ratio to the actual one from OIRA. The results are shown below.

estimated ratio for tenure and tenure track: 0.7832817
 actual ratio for tenure and tenure track: 0.7617647
 estimated for fixed-term (non tenure track) 1.419319
 actual for fixed-term (non tenure track) 1.412871

Fig 3.2.1 Output for gender ratio by ranks - predicted compared to actual

3.3 Visualizations

With these JSON exported visualizing is simple. The grouped departments object is used first, to render a 13-page table, each with a button to render the visualizations on the rest of the page based on that department's professors. On the top right is a bar graph that utilizes all the rows in the other JSON object with professors whose departments match the selected one, and displays the average salary in that department for women, for men, and for names the R package was not able to predict, along with the overall average salary for the department.

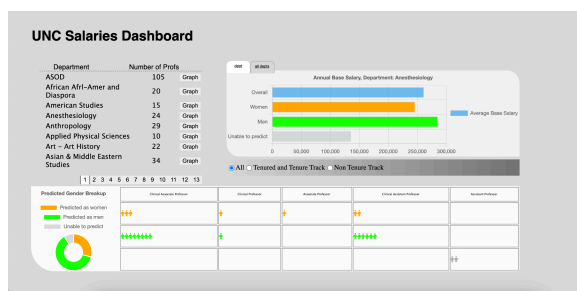


Fig 3.3.1 Salaries dashboard with specific department graph shown

Clicking a graph button on the table also renders two more visualizations - a simple donut chart that shows the divide of the gender prediction counts (by sheer number of professors matching that prediction in the department, not by salary proportion), as well as the specific amounts of that predicted group when the mouse hovers over a donut sector.

The table to the right of the donut chart breaks things down even further, showing where the individual professors line up across gender and rank. Across the board, orange is used to represent professors predicted as women and green is used for professors predicted as men. The simple reason for this is that using blue or a red-adjacent color might unnecessarily give a viewer a reason to insinuate there is a message on gender norms by color (since blue is stereotypically a masculine color and pink a feminine one). So on the chart below, you can see that in the Anesthesiology department, predicted male professors make more on average[9] than predicted female professors, the department has a larger proportion of predicted male professors, and most of the faculty have a fixed-term modifier in their titles such as "Clinical".

There are also a few buttons to filter the data, labeled "All", "Tenured and Tenure Track" and "Non Tenure Track" which perform the obvious operation on the data.

Lastly, there is a visualization to enable easy spotting of the department with the worst wage gap. By clicking the "all depts" tab at the top of the bar chart section, a new horizontally scrollable bar chart displays the differences in average male and female salaries for every department. The difference is just ((average male salary) - (average female salary)) so whatever bars have higher amounts have the highest wage gap in men's favor.

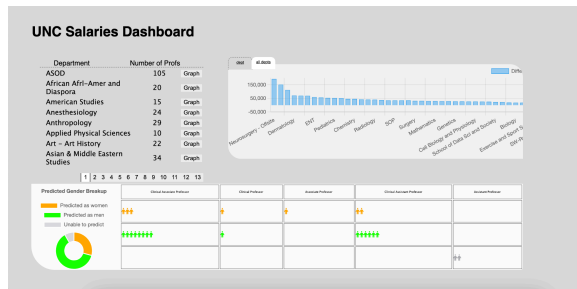


Fig 3.3.1 Salaries dashboard with all departments wage gap graph shown

4 FINDINGS

While one could go department by department to answer the research question of which department has the worst wage gap, we can now simply glance at the chart and see the answer is Neurosurgery. Hovering through we can see all the top four are related to medical fields, and the number five worst disparity belongs to the Kenan-Flagler Business School.

Another finding in general while engaging with the tool is that not only do the departments with the worst wage gaps tend to have more men represented in the higher ranks (as expected) but they also have more men employed in general. Of course, these are still merely estimates on gender so it is entirely possible for some names predicted wrongly to have set some things off.

By looking at the chart for disparity across all departments it is easy to see there are more bars in the positive, favoring men, than the other way around. This supports the notion that male professors at UNC as a whole generally make more. To support the finding, the actual wage gap across the board for all professors at UNC was calculated in the R project, seen below (with male and female salary still based on a best effort to predict professor genders).

avg overall salary: 171941.4
avg male salary: 192156.4
avg female salary: 155268.6

Fig 4.1 Average salaries for all professors

One last finding from the dashboard is that departments tend to not have more than a handful of unpredicted gender professors, and the 329 professors

whose gender were unable to be reasonably predicted were spread pretty evenly across the different departments, which adds some confidence in the accuracy of the findings in the dashboard as a whole.

Bibliography

- [1] K. Haan, "52 Gender Pay Gap Statistics In 2023 – Forbes Advisor," *www.forbes.com*, Feb. 27, 2023.
<https://www.forbes.com/advisor/business/gender-pay-gap-statistics/#:~:text=What%20is%20the%20wage%20gap>
- [2] G. Carlton, "Professor Gender Pay Gap: Colleges Haven't Fixed It. Will Courts? | BestColleges," *www.bestcolleges.com*, Sep. 15, 2023.
<https://www.bestcolleges.com/news/analysis/professor-gender-pay-gap-colleges-havent-fixed-it-will-courts/#:~:text=The%20Gender%20Pay%20Gap%20in> (accessed May 02, 2024).
- [3] L. Allen, "Full-Time Faculty Avg. Salary, by Academic Rank," *AAUP Data*, Jul. 08, 2022.
<https://data.aaup.org/fcs-ft-faculty-salaries/> (accessed May 02, 2024).
- [4] Office of Information & Assessment, "Analytic Reports | OIRA," *Unc.edu*, 2018.
<https://oira.unc.edu/reports/>
- [5] K. Palmer, "Workplace climate pushing women out of the professoriate," *Inside Higher Ed | Higher Education News, Events and Jobs*, 2023.
<https://www.insidehighered.com/news/faculty-issues/diversity-equity/2023/10/24/workplace-climate-pushing-women-out-professoriate#:~:text=Women%20make%20up%2044%20percent> (accessed May 02, 2024).
- [6] Board of Trustees of The University of North Carolina at Chapel Hill, "Trustee Policies and Regulations Governing Academic Tenure in the University of North Carolina at Chapel Hill Section 1. Academic Freedom," Sep. 2020. Accessed: May 02, 2024. [Online]. Available:

<https://facultyaffairs.unc.edu/wp-content/uploads/sites/1353/2023/07/UNC-Chapel-Hill-Tenure-Policies-and-Procedures.pdf>

[7] UNC System, “UNC Salary Information Database,” *Northcarolina.edu*, Mar. 31, 2024.
<https://uncdm.northcarolina.edu/salaries/index.php>

[8] J. Kaplan, “Package ‘predictrace’ Title Predict the Race and Gender of a Given Name Using Census and Social Security Administration Data,” Jul. 2023, doi: <https://doi.org/10.1038/sdata.2018.25>].

[9] The College and University Professional Association for Human Resources, “Tenured/Tenure-Track Faculty Salaries,” *Higheredjobs.com*, 2024.
<https://www.higheredjobs.com/salary/salaryDisplay.cfm?SurveyID=24>