

A STUDY OF GENDER ISSUES  
IN DATA JOURNALISM

---

A Project  
presented to  
the Faculty of the Graduate School  
at the University of Missouri-Columbia

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

---

by  
MADI ALEXANDER  
David Herzog, Chair  
MAY 2016

## ACKNOWLEDGMENTS

I'd like to thank my committee—Amy Simons, David Herzog and Mark Horvit—for their help and guidance during my time at the University of Missouri. I couldn't have asked for better professors, mentors and advisors.

I would like to thank my mom and my grandfather for their unwavering support and continuous encouragement. I'd like to thank my wonderful partner, Anthony Hoehner, for being the most loving, patient and understanding person I know. I would also like to thank the friends, professors and unofficial mentors who encouraged me, listened to me complain, helped me apply for jobs and internships, and continuously told me to not give up. Especially Anna Facci, Farris Willingham, Katy Mersmann, Beatriz Costa-Lima, Liz Lucas and Ryann Grochowski Jones.

Last but not least, I would like to thank the strong, talented journalists who were willing to share their triumphs, struggles and vulnerabilities with me for this project.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
ABSTRACT .....	iii
Chapter	
1. INTRODUCTION .....	1
2. SCHEDULE AND FIELD NOTES .....	2
3. EVALUATION .....	12
4. PHYSICAL EVIDENCE .....	16
5. ANALYSIS .....	22
REFERENCES .....	43
APPENDIX .....	45

## ABSTRACT

Women in data journalism face a myriad of obstacles during their careers. Men with similar experiences and educations are paid more for the same work. Sexual harassment policies do little to protect victims or prevent harassment. Male editors favor other men for promotions and better assignments. Twelve women—all professional data journalists—share stories of inequality and discrimination, both in and out of the newsroom. In general, women in data journalism report that they feel like outsiders in a male-dominated profession. They also agree that companies, editors and individual journalists need to make concerted efforts to establish a friendlier environment for women in data journalism.

## CHAPTER 1: INTRODUCTION

During my two years at the Missouri School of Journalism, I have worked to develop sharp news judgment and specialized data skills. My degree program leaves me with the ability to use multiple languages and tools for a variety of data-driven projects: analysis with SQL, statistics in R, interactive graphics in D3, maps with ArcGIS, scraping with Python, etc. I have also learned how to use all those skills to tell better stories and to focus on the people behind the data. These skills, combined with my previous journalism experience, have enabled me to work, intern and freelance for some of the most highly respected news outlets in this country. I am confident in my ability to produce in-depth, meaningful, high-quality stories using data.

This project—working for ProPublica and researching gender inequality in data journalism—will no doubt be a stepping stone to the rest of my career. My career goal is to be a full-time data journalist, so this project encompasses everything I have prepared for during graduate school. Additionally, it gives me a solid foundation for the future. I am employable and my skills desired in the newsroom. As a woman and as a journalist, it makes sense that I should research this topic. Women deserve to have their stories told. They endure so much in the workplace, including in data journalism, which is why I am focusing my project on the women who work in data journalism.

## CHAPTER 2: SCHEDULE AND WEEKLY FIELD NOTES

### Schedule

**Week 1: January 11** – NFIRS, Alabama, referrals work; compiling list of potential interviewees

**Week 2: January 18** – Alabama, data slice, referrals work; contacting potential interviewees

**Week 3: January 25** – PDMP data, Alabama data; interviewing subjects

**Week 4: February 1** – PDMP data, Dollars for Docs; conducting interviews for research

**Week 5: February 8** – Alabama data, PDMP data; conducting interviews for research

**Week 6: February 15** – Dollars for Docs, PDMP, Alabama data; interviewing subjects

**Week 7: February 22** – PDMP data, Hell and High Water; transcribing interviews

**Week 8: February 29** – PDMP data, Alabama; transcribing more interviews

**Week 9: March 7** – NICAR conference; outlining structure for analysis

**Week 10: March 14** – PDMP, Dollars for Docs, dialysis data; writing analysis

**Week 11: March 21** – Dialysis data, Alabama data; writing analysis

**Week 12: March 28** – Alabama, PDMP data, dialysis; finishing first analysis draft

**Week 13: April 4** – Dollars for Docs, dialysis data; editing second draft of analysis

**Week 14: April 11** – Dialysis, Ohio data collecting; finishing draft of project report

### Weekly Field Notes

#### January 15, 2016

This is the first project memo of the semester. I just finished my first week at ProPublica, which was nothing short of wonderful.

I checked .csvs of chemical endangerment arrest data in Alabama for parsing errors. Chemical endangerment could be a pregnant woman doing drugs or someone cooking meth in a house with a child. The data originally came in as a fixed width text file with upwards of 200 columns, so one of the developers parsed it for us. However, I found 58 errors in the three tables because of an extra double quote at the beginning of the row. That means the rest of the entire row is off by one space. I plan on tackling that next week.

Also, I used data from the National Fire Incident Reporting System to pull out fire incidents with a death or injury in Alabama or Illinois. I did this for 2010-2012. In one of the tables, we got this error: [SQL Server]Unclosed quotation mark after the character string 'KEMP COUNTY ROADS 4071'.

The files are all .dbfs, so I can't import them into Excel on a Mac. Ryann was able to import the file into Excel, then filter to IL and AL. Then I joined tables that had basic incident data, address data and structure information, per Derek's request. He wanted address information, structure information (like property type and property use), injury and death counts for each fire.

The other issue I ran into with the NFIRS data is that in one of the tables, it is possible that all the fields I was joining are null except for the field I'm joining on. In that case, SQL Server skips the nulls and doesn't join on the next table at all.

Ryann and I had to figure out if there was a problem with the unique identifiers I created for each entry or if there was some issue with the data types. Eventually I rearranged the query to have the table with all the nulls be the last table joined in the query so nothing would be skipped. The first time I tried, about 80 of the 900 final rows ended up with blank addresses even though I could clearly see they existed and the unique IDs matched.

ProPublica also has a database of doctors who refer to other doctors, but it's all coded. So in order to get usable information to put into a news app, I joined the table with the codes to the table with the doctor info. That way instead of just the codes, you can see the doctor's name, practice name and other contact information, as well as number of patients referred from one doctor to the other.

Next week I plan on posting on the JAWS and NICAR listservs to ask for volunteers to be interviewed. I also plan on making a list of journalists to email and ask if they'd like to be interviewed (if you have suggestions, I will take them). I also need to fix those errors in the chemical endangerment data.

## **January 22, 2016**

This week I worked on building a database that ranks the top 5 doctors that one doctor refers to, and then the top 5 doctors that refer to one doctor. I'm still in the process of

making the final tables for that because the queries take a few hours to run. There are several million records, plus I am ranking them based on the number of patients referred.

I also cleaned the errors in the Alabama chemical endangerment data. I met with the reporter working on that story to discuss what she wants out of it. Essentially I am going to be looking for discrepancies in prosecution between black and white women, differences in bond amount for pregnant women and people of color, and differences in country prosecution under the chemical endangerment law. I have already spotted a few notable things, but I need to load everything into SQL Server now that I've fixed all the errors.

In order to fix the errors, I manually copied each row from the original fixed width .txt file (one of the reporters had already parsed them into .csvs, but with errors) into a new file. Then I parsed them again using the record layout in Google Refine. I still got errors (commas in the wrong spots), so I checked each column until I found the problem. The record layout says "8 + 2 DECIMAL" for three columns, which worked as a length of 10 for all the other records. But for these 58, for some reason that length was supposed to be 11. No idea why, but at any rate, they're fixed.

I also created two slices of some psychiatric medical provider information for someone buying data from the ProPublica data store.

I have my posts written for the listservs, but won't post until Monday or Tuesday when people are actually checking their email. This weekend I will craft a generic message to send directly to data reporters about my project, asking if they'll participate, telling them the basics of the project, why I'm doing it, etc. Then I will start reaching out to individual people.

## **January 29, 2016**

On the ProPublica front, this week I've been working with data from the DOJ's Prescription Drug Monitoring Program. The data came in two sets because the DOJ changed the form halfway through the data collection period, so the questions are slightly different.

First I went through 5 PDFs from each quarter and checked every single number against the data to make sure it was entered correctly (around 100 questions per PDF, 5 PDFs per quarter, 9 quarters). After doing this, we wanted to narrow down the questions to the data we wanted to use for analysis. However, there's an issue with some of the columns in one dataset being duplicates.

It appears as though a third form may have been used because there are multiple questions for what look like the same thing, but we don't know because some of the PDFs are incomplete. I'm talking to the DOJ about why that is and why the PDF reports are missing so many questions.



Next week I will be working on the Alabama chemical endangerment analysis more, as well as loading all of this data into SQL server for further analysis.

In the meantime, I'm creating a key between the two forms (maybe three) so we know what questions are the same. I suspect the DOJ will tell me the questions are about the same thing, but I don't want to do the analysis without knowing for sure.

On the research front, I interviewed two subjects already. I have three more subjects willing to participate, so I need to schedule times to interview them. I also have five more people to email on Monday.

### **February 5, 2016**

On the ProPublica front, I reached out to Charlie's contact at the DOJ on Monday about getting whatever forms were used for PDMP reporting. Essentially I just requested whatever form the states would have used during that reporting period, so that will help me figure out if some of these fields that I suspect are duplicates are actually the same field.

I have also spent a good part of this week helping spot check the Dollars for Docs data that's going to be updated fairly soon. I checked a list of doctors to make sure that matches from two datasets are indeed the same person. Essentially there are two unique identifiers for doctors in the data, but there's no crosswalk between the two so they had to match names and other identifying information. I just have a list of numbers so when I pull up each number online and in the app, the doctor should be the same person. This took most of my time this week, and will likely take up a good chunk of next week too.

On the research front, I finished three more interviews and I have two more scheduled for next week and one the week after. This puts me at seven interviews completed with three pending.

### **February 12, 2016**

On the ProPublica front, this week I talked to the DOJ about creating a crosswalk between the three different sets of questions. As it turns out, there are actually three editions of the questionnaire states have to fill out, so they're going to crosswalk all three for me.

I also started to go through the Alabama chemical endangerment data. I noticed immediately that multiple women are listed more than once. Essentially, they have both a district court case and a circuit court case for each offense. I have a list of pregnancy cases that had previously been put together by hand for individual women. So, the caveat here is that in the second data set, pregnancies are listed by women, not by case number.

This poses a problem for me because I'm trying to look at individual case numbers for two reasons. First, for many women, the circuit court case is the main case that I need to look at. If a woman has a charge only at the district court level, I need to be able to pull that out. I need to distinguish between four different variables: circuit court pregnancy cases as the main case, district court cases as the main case, district court cases as the secondary case, and non-pregnancy cases. I know, it's confusing.

So, to do this, I created a lookup table of sorts so I could deal with the pregnancy cases. I joined the original pregnancy cases with the new data by case number in order to mark the women with pregnancy cases. Now I'm starting to go through each file by hand to manually mark the cases in the particular way I need them to be marked. Tedious, yes. But not really a better option here. I can sort by last name and then just scroll through the file and see where there are pregnancy cases. If the names are the name and the incident AND offense dates are the same, I mark the case. For some, Nina needs to look at the actual cases to be sure, but I'll be able to do most of this myself.

On the research front, I've done another two interviews and scheduled two more. This leaves me with nine completed interviews and three scheduled interviews. One interview is next week, two are the week after next.

## **February 19, 2016**

On the ProPublica front, I mostly spent this week helping Ryann spot check Dollars for Docs. She sent me a list of drug names, companies, doctors, etc. to check between the app and the CMS website. D4D combines two years of CMS data into one sum, so essentially I ran a bunch of queries on a random list to ensure that all the math added up. I also checked companies that were merged from two separate entries in the CMS data. CMS listed some companies as two institutions, usually as the result of a spelling or punctuation difference. ProPublica combined them, so I just had to make sure things like payment amounts, dates, doctor names, etc. added up properly.

Also, the DOJ sent me the crosswalk we asked for between old PDMP forms and new ones. I am honestly surprised at how fast they did it. I didn't have the time this week to make sure it's exactly what I needed, but in the three minutes I spent looking at it, this is going to work. Essentially now they've documented all the changed measures (i.e. if some questions changed from 3-month to 6-month measures) or deleted measures, so I know which questions from different report years I can combine. Thankfully the ones I needed can be used across all three versions.

The other thing I've been doing this week is going through the Alabama data by hand to flag pregnancy cases. There are about 500 cases in a 5,000-record dataset, so I just have to do it slowly but carefully.

On the research front, I officially have 10 interviews done, with two more pending for next week. Half of the women work for traditional print outlets, the other half work for online outlets.

### **February 26, 2016**

This week I had time to sit down and look at the crosswalk the DOJ sent us. I had to divide one of the datasets into two different files to import them. From there, I was able to import all three datasets into SQL using the crosswalk. Additionally, the data is divided up into reporting periods by quarter. We decided to go ahead and code this 1 to 15 (we have 15 quarters worth of data) because it's easier to look at all the data consecutively this way.

After building one final dataset in SQL, I exported the dataset into Excel so I could create a pivot table and share it with Charlie. It's a lot easier to look at the data by state and reporting period in a pivot table.

In looking at the pivot table, we noticed some things (a lot of things) seem wrong or just strange. For example, some states have zeroes in fields where there should definitely not be zeroes. Some other states have one single outrageously high number in a field that has pretty low numbers for the rest of the reporting periods.

I spent yesterday and most of today looking at each field individually so I could make a list of any states, award numbers and fields that have errors. So far, Wisconsin, Massachusetts and Maryland seem to be totally flawed. I've gone through about half of the fields and looked at each state individually. I'm making a list of errors so we can go back to the DOJ next week and ask about how they audit the data, if certain states are unreliable, etc.

I also helped do some spot-checking of the Hell and High Water project, which is about Houston's lack of preparedness for the next big hurricane. Essentially I just plugged Houston addresses into TAMU's model and into ProPublica's model to make sure the flood levels were relatively close between the two. You can find that project here:

<https://projects.propublica.org/houston/>

On the research front, I am slowly but surely transcribing all my interviews. I suspect this is what I will be doing for most of the next couple weeks. Then the writing begins!

### **March 4, 2016**

On the ProPublica front, this week I completed the list of fields/states/awards from the PDMP data that had errors (or any that looked remotely strange). Each state has its own requirements for reporting the data, so it could very possibly be that some states are unusable and others are fine. I emailed the folks at the DOJ to see what they have to say

about it. I've also just been looking at the states with the least amount of errors to see if I can identify any trends right off the bat.

Today I also finished up compiling a list of case numbers for the Alabama data for women who were pregnant and charged with chemical endangerment. I sent that list to the reporter, Nina Martin, so she could look up the cases and tell me which cases I should include in my analysis. She should get those back to me next week, so from there I can finish creating the dataset of pregnancy cases and can move forward with analysis.

On the research front, nothing too exciting this week. I'm continuing to transcribe my interviews. All of them are about an hour long (some more, some less), so it takes a decent amount of time to transcribe each one. I'm also outlining the final paper so I have a good idea of how it'll be structured.

### **March 11, 2016**

Not too much to report this week since we're all at NICAR right now. On Monday and Tuesday, I spent the days mostly just trying to make heads or tails of the PDMP data. Charlie and I plan to get together next week to go over that. Still waiting on Nina to get me the info back for undetermined pregnancy cases.

I've started outlining the basics of my analysis. I'm going to break it up into subject (pay, harassment, NICAR, etc.) and give anecdotes for three or four women in each section. This way, I can make sure they at least get to comment on one or two things.

### **March 18, 2016**

On the ProPublica front, Dollars for Docs launched: <https://projects.propublica.org/docdollars/>. My first official ProPublica credit line is on there, which is pretty cool. I spent a couple days bulletproofing the app with the rest of the team before it launched. Clicking around, looking doctors up, trying to spot errors or typos, etc.

This week Charlie and I talked to a guy at Brandeis about PDMP data. He's basically the expert on this subject and one of the few people who does analysis with the data. Essentially, he said some states have very thorough methods and some don't. There's been some confusion about some of the measures that involve thresholds or the counting of multiple kinds of drugs.

Some of the questions measure schedule II, III and IV drugs. But some states have taken that to mean II AND III AND IV, while some have taken that to mean II OR III OR IV. The measures differ greatly between those two interpretations, unfortunately. And some states switch between two interpretations halfway through the grant period. This same issue is also present in the field that measure threshold for pharmacies i.e. if a person

filled a prescription for painkillers from 5+ doctors or at 5+ pharmacies. Some states interpret it as 5 AND 5, or 5 OR 5.

I also got a new project today: working with dialysis data. Back when Jen LaFleur worked here, ProPublica launched a dialysis facility app, which you can view here if you want: <http://projects.propublica.org/dialysis/>

So, we have new data for 2014. Jen left a Word doc of some queries outlining how she did the analysis of 2012 data in 2013. Unfortunately, the 2014 data didn't come with a data dictionary, but I've found something similar on the CMS website that should suffice. I'm just starting on this, so I need to read all the documentation and figure out exactly what needs to be done. I know I need to join this data with another dataset from CMS to get address and contact information for all the facilities.

On the research front, I'm starting to write and piece together some interesting (and horrifying) anecdotes. Every woman I talked to had multiple stories to tell about everything from pay inequality to being called obnoxious pet names to being talked over by men, so I'm glad I've chosen to address this topic.

### **March 25, 2016**

On the ProPublica front, I'm focusing a lot on dialysis this week. The agency sent the files in parts, but I've looked at all the files and there seems to be no difference between the parts. Each part has the full table and then subsets of the tables because the full data is a few thousand fields wide. But all the data looks exactly the same cross all six parts. Ryann is asking about this.

Anyway, the 2014 data needs to be joined with an additional dataset. The deal is that both the facility number and facility name can change when a facility changes hands. So, there's a change that once I get this all in, I'm going to have a chunk of facilities that have no matches on name or number. But first, I need to make sure I'm using the right timeframe for facility info so it'll better match with 2014.

For now, I'm just getting this very long list of queries to work in SQL Server (they're written for MySQL). A couple field names are just slightly different in the new data, so I need to go through Jen's notes to make sure I'm using the proper field names in the queries I'm writing.

Nina also sent me the pregnancy cases she checked, so now I need to go through and add them in by hand again. I also pulled and formatted some census data so that I (and mainly Nina) can compare demographics against prosecutions and arrests.

I also talked to Charlie and Ryann yesterday about PDMP stuff. The consensus is that I'm going to spend next week reaching out to a couple states to ask about their programs. If there's anything they don't measure, how they measure it, if the data the BJA gave me is

right, etc. The states on my list are: Kentucky, Wisconsin, Montana, Maryland and Indiana. Once we figure out if the data is usable, we can analyze it. But for now, we just need more info about how it's maintained and collected.

On the research front, I've got about 1,000 words of my draft done. I'm not shooting for a word count here, just trying to keep track of the progress I'm making.

### **April 1, 2016**

This week I spent Monday and Tuesday finishing up the Alabama project regarding chemical endangerment arrests and prosecutions. Ultimately, I ended up compiling some tables of pregnancy and non-pregnancy cases by race and county. I also provided some breakdowns of bail amounts and number of cases so Nina could look more closely at those.

Additionally, I've spent this week contacting a few states about discrepancies in their PDMP data. Specifically, I've contacted Montana, Maryland, Kentucky, Indiana and Wisconsin.

An interesting thing I learned is that states have to build their own programs and software to keep track of this information. So in Montana, their software engineers have not built in the capability to track certain measurements. The BJA does not allow states to report blanks, so instead they have to input zeroes for several measures.

I was definitely right about the threshold measures being interpreted differently. Kentucky switched up their interpretation halfway through the grant, so the earlier part of the reporting period has incredibly high numbers for several measures, while the latter half has much, much lower numbers.

In some states, only a fraction of pharmacies participate in the PDMPs. When more pharmacies are added, the numbers look like they spike dramatically, but that really has to do with the number of entities reporting. Charlie, Ryann and I are going to talk about this on Monday to decide how we're going to proceed with the data, given that some of it is definitely flawed.

On the research front, I've got a decent amount of my draft completed. I'm sitting at about 2,500 words right now.

### **April 8, 2016**

On the ProPublica front, I helped Ryann create a record layout for Dollars for Docs. Most of the fields in the dataset come directly from CMS, so I just need to verify that they're the same and input the field description. This data is going onto the data store, so it needs its own record layout.

I also finish my analyzing and joining of dialysis data. Well, that's really to say I got to a stopping point. The app isn't getting updated, so it's not imperative that I finish this right now. I also spot-checked the dialysis data against the PDF reports. I got all the queries working and ready for the next person. I also joined the dialysis data with the facility information. Sadly, since so many numbers and names changes, I tried a few different ways to match them.

There are 818 facilities that don't match on name + number. Some of those are misplaced commas, dashes, etc. but some are due to the name or number changing when a facility changed hands. So, there are 79 facilities that don't match on number alone. And then there are 70 facilities that don't match on name, number or a combination of name and number. Essentially, someone's going to have to go through those 818 facilities and verify them by hand. It's not an easy feat, especially for facilities that have both changed number AND name. Scott and Ryann are discussing what to do about this.

On the research front, my analysis is done. David is reviewing it and will edit some things. Once I finish the second draft, the project will go to the full committee. So, as of right now, my defense will be Monday, April 25 at 1 p.m. The 14<sup>th</sup> week of my project is next week, so one I finish the memo and whatnot for that, I'll be able to send my full project draft out to the entire committee. I want to at least give everyone the weekend.

## **April 11, 2016**

Here it is, my fourteenth and final week of this project. On the ProPublica front, PDMP is on hold for now since the story may end up being about how PDMPs aren't functioning how they should be. Instead, I'm going to be working on an Ohio project for the next couple weeks.

I'm also writing a how-to of sorts regarding dialysis so the next person can spend much less time than I did getting situated with it.

The goal is this project is to show the changing landscape of voting in Ohio. So, my contribution to this is to track down a couple data sets from the Bureau of Labor Statistics. Specifically, I'm looking for data regarding manufacturing versus service jobs. I'm also looking for the number of companies with more than 1,000 employees, the proportion of retirees with defined pension plans, and union membership. The goal is to get data going back... well, going back as far as I can or as far as it exists.

Anyway, this is my final memo. My project is completed except for this and a conclusion section. It'll go out to all of you so you have a week to look at it before I defend. That's it! I'm almost finished with grad school!

### CHAPTER 3: EVALUATION

Overall, I am very pleased with the work I've done at ProPublica and how I've developed as a professional journalist during the past months. The level of complexity in the data I've worked with in the past 14 weeks has been a challenge, and I am proud of the quality of work that I produced.

One unique thing about this project is that I worked remotely as an intern for a newsroom in New York City. Having never worked remotely before, I felt a little unsure about communication with the newsroom, as well as ensuring that I not get distracted doing other stuff at home. I honestly felt like this wasn't much of an issue at all once I got started. Between Slack, Skype and phone calls, I kept in constant contact with the newsroom and specifically with Ryann.

Furthermore, I felt like this experience was an excellent transition between being a student/intern and being a full-time working professional. My assignments were given to me, which is not how it works in most newsrooms, but it was still my job to track down information, call government agencies, etc.

Not only that, working for ProPublica really helped me figure out how communication works in a position where a data reporter is working as a part of a team and with individual reporters throughout the newsroom. Communicating with my boss, with the reporters assigned to projects and with their editors was a new experience for me. I'm glad I had the opportunity to figure out how a position like this, i.e. doing data analysis for other people's stories, fits into the larger newsroom environment. This job was also a mix of data analysis and some reporting. Though I wasn't interviewing people,



I was calling government agencies and working with people outside of the newsroom to better assist my data analysis. Learning to communicate and negotiate with other humans for data is important in this line of work.

On a technical level, working at ProPublica was a great way to test my skills, as well as learn new things. Specifically, I had never used SQL Server before. I know the syntax is similar and obviously some of the major concepts are the same, but I had to learn so many new things. This also provided me with the opportunity to learn more advanced SQL than I had previously written in class or for the data library. Doing data work constantly and consistently had taught me a lot about writing code efficiently.

Another aspect that I hadn't previously considered is that I paid so much extra attention to documenting my work. Between saving queries and naming files coherently, I really feel like I've gained an understanding about what it's like to work with a team on projects, rather than just with myself or with one other person. Showing your work is an important aspect of data journalism, especially at ProPublica.

As far as research was concerned, that progressed along nicely and I'm incredibly grateful that I was able to write about such a sensitive topic. One thing I noted is that almost everyone I interviewed asked me again during the interview if these were going to stay confidential. So, while there are a lot of problems out there in data journalism with regards to women, many of the women I interviewed are still hesitant to address those issues with full force for fear of backlash.

Listening to other women talk about their careers also helps me know what I'm likely to encounter on the road ahead. That's really something I take away from this project. And there's also a lack of mentorship in data journalism for women. For me,

having someone five, ten or twenty years older than you explain what it was like for them gave me a lot of hope for the future. Also a lot of hesitation, but I'd rather be prepared than surprised.

This project was a great journey for me, both personally and professionally. These past 14 weeks taught me great lessons about journalism, communication and feminism. I have no doubt that I am fully prepared for the next step: moving to Arlington, Virginia, to work as a data reporter for Bloomberg BNA.

## Supervisor's Evaluation



April 15, 2016

To Whom It May Concern:

In her short time with us, Madi Alexander has proven to be an invaluable member of ProPublica's News Applications team. When we hired Madi, we were looking for a person who could assist with the everyday tasks – both big and small – of data journalism. We needed someone who was smart, self-starting, curious and well-acquainted with cleaning, analyzing and bulletproofing data. Madi more than exceeded those expectations.

Madi was dependable and easy to get in touch with – no small feat considering she worked remotely and in a different time zone. She consistently met her deadlines. Madi needed no hand-holding with her assignments, but she also was not afraid to ask for help when she needed it. She was friendly and helpful when dealing with the reporters and editors in our newsroom, and I know others at ProPublica enjoyed working with her.

I was quite impressed with Madi's journalism and reporting skills. She was not afraid of picking up the phone and making calls when she needed to. Madi has shown to be adept at finding stories, leads and ideas in data and I'm excited to see what ideas she comes up with when she starts her full-time position at Bloomberg.

On Madi's first day in January, we already had a list of things for her to do. I'm happy to say she hit the ground running and never stopped. As a Mizzou graduate and NICAR data library alum, I had no doubt that Madi would have the ability to complete the tasks assigned to her. But Madi really went above and beyond what was expected, and I'll be sad to see her move on to her new position in a few weeks. Madi is an excellent data journalist and has a bright future ahead of her.

Sincerely,

Ryann Grochowski Jones  
Data Reporter, ProPublica  
155 Avenue of the Americas, 13<sup>th</sup> Floor  
New York, NY 10013  
917-512-0249  
[ryann.jones@propublica.org](mailto:ryann.jones@propublica.org)

## CHAPTER 4: PHYSICAL EVIDENCE

### Dollars for Docs

<https://projects.propublica.org/docdollars/>

## Dollars for Docs

### How Industry Dollars Reach Your Doctors

*By Charles Ornstein, Lena Groeger, Mike Tigas, and Ryann Grochowski Jones, ProPublica. Updated March 17, 2016*

Pharmaceutical and medical device companies are now required by law to release details of their payments to a variety of doctors and U.S. teaching hospitals for promotional talks, research and consulting, among other categories. Use this tool to search for general payments (excluding research and ownership interests) made from August 2013 to December 2014. | [Related Story: Now There's Proof: Docs Who Get Company Cash Tend to Prescribe More Brand-Name Meds](#) »

**Has Your Doctor Received Drug Company Money?**

All States

For example: [Andrew Jones, Boston, 10013](#)

 **\$3.49B** in disclosed payments  **684,915** doctors  **1,135** teaching hospitals  **1,565** companies

#### Top 50 Companies

Click on a company to see how its payments break down by drug, device or doctor. Or, [see all companies](#) »

COMPANY	PAYMENTS
<a href="#">Genentech, Inc.</a>	\$388M
<a href="#">DePuy Synthes Products LLC</a>	\$94.7M
<a href="#">Topera, Inc.</a>	\$93.1M
<a href="#">Stryker Corporation</a>	\$90.8M
<a href="#">AstraZeneca Pharmaceuticals LP</a>	\$90.7M
<a href="#">Medtronic Sofamor Danek USA, Inc.</a>	\$85M
<a href="#">Pfizer Inc.</a>	\$82.5M

#### Highest-Earning Doctors

NAME	PAYMENTS
<a href="#">SUJATA NARAYAN</a> Family Medicine	\$43.9M
<a href="#">KAREN UNDERWOOD</a> Pediatric Critical Care Medicine	\$28.5M
<a href="#">STEPHEN BURKHART</a> Orthopaedic Surgery	\$24M
<a href="#">SANJAY YADAV</a> Cardiovascular Disease	\$23.1M
<a href="#">KEVIN FOLEY</a> Neurological Surgery	\$22M

*Additional reporting, design & development by Madi Alexander, Al Shaw, Annie Waldman, Tobin Asher, Eric Sagara, Jeremy B. Merrill, Dan Nguyen, and Sisi Wei. Pharmacy, building, hospital and first aid icons from The Noun Project.*

#### About the Dollars for Docs Data

Details behind our drug company money database.

#### Download the Data

The entire data set is available for purchase in the [ProPublica Data Store](#).

#### Source

The Centers for Medicare and Medicaid Services [Open Payments](#) data.

#### Archive

Search for payments made by 17 drug companies between 2009 and 2013.

#### Background Stories



### Alabama Chemical Endangerment

```
SELECT [criminal].[COUNTY],
[criminal].[COUNTY_NAME],
[criminal].[CASE],
[criminal].[NAME],
[criminal].[SSN],
[criminal].[SEX_CODE],
[criminal].[RACE_CODE],
```

```
[criminal].[BIRTH_DATE],
[criminal].[INCIDENT_DATE],
[criminal].[OFFENSE_DATE]
INTO offender_info
FROM [criminal];
```

```
ALTER TABLE [offender_info]
ADD pregnancy varchar(255);
```

```
SELECT *
FROM [chem_pregnancies]
INNER JOIN [offender_info] ON [offender_info].[CASE] =
[chem_pregnancies].[Casenum] AND [offender_info].[COUNTY_NAME] =
[chem_pregnancies].[County];
```

```
UPDATE [offender_info]
SET [offender_info].[PREGNANCY] = [chem_pregnancies].[pregnancy]
FROM [chem_pregnancies]
INNER JOIN [offender_info] ON [offender_info].[CASE] =
[chem_pregnancies].[Casenum] AND [offender_info].[COUNTY_NAME] =
[chem_pregnancies].[County];
```

```
UPDATE [dbo].[offender_info]
SET [dbo].[offender_info].[PREGNANCY] = '4'
WHERE [dbo].[offender_info].[PREGNANCY] = '0';
```

```
ALTER TABLE [dbo].[criminal]
ADD pregnancy varchar(255);
```

```
UPDATE [criminal]
SET [criminal].[pregnancy] = [offender_info].[pregnancy]
FROM [offender_info]
INNER JOIN [criminal] ON [dbo].[offender_info].[CASE] = [dbo].[criminal].[CASE]
AND [dbo].[criminal].[COUNTY_NAME] = [dbo].[offender_info].[COUNTY_NAME];
```

```
SELECT [dbo].[criminal].[COUNTY_NAME],
[dbo].[criminal].[RACE_CODE],
sum(cast([dbo].[criminal].[BOND_AMOUNT] as numeric)) AS pregnancy_bond_amt,
count([dbo].[criminal].[RACE_CODE]) AS pregnancy_race_count
FROM [dbo].[criminal]
WHERE [dbo].[criminal].[SEX_CODE] LIKE 'F' AND ([dbo].[criminal].[pregnancy] IN
('1','3'))
GROUP BY [dbo].[criminal].[COUNTY_NAME], [dbo].[criminal].[RACE_CODE]
ORDER BY [dbo].[criminal].[COUNTY_NAME] ASC;
```

```
SELECT [dbo].[criminal].[COUNTY_NAME],
```

```

[dbo].[criminal].[RACE_CODE],
sum(cast([dbo].[criminal].[BOND_AMOUNT] as numeric)) AS
nonpregnancy_bond_amt,
count([dbo].[criminal].[RACE_CODE]) AS nonpregnancy_race_count
FROM [dbo].[criminal]
WHERE [dbo].[criminal].[SEX_CODE] LIKE 'F' AND ([dbo].[criminal].[pregnancy] IN
('4'))
GROUP BY [dbo].[criminal].[COUNTY_NAME], [dbo].[criminal].[RACE_CODE]
ORDER BY [dbo].[criminal].[COUNTY_NAME] ASC;

```

```

SELECT [dbo].[criminal].[RACE_CODE],
sum(cast([dbo].[criminal].[BOND_AMOUNT] as float)) AS pregnancy_bond_amt,
count([criminal].[pregnancy]) AS pregnancy_case_count
FROM [dbo].[criminal]
WHERE [dbo].[criminal].[SEX_CODE] LIKE 'F' AND ([dbo].[criminal].[pregnancy] IN
('1','3'))
GROUP BY [dbo].[criminal].[RACE_CODE];

```

```

SELECT [dbo].[criminal].[RACE_CODE],
sum(cast([dbo].[criminal].[BOND_AMOUNT] as float)) AS nonpregnancy_bond_amt,
count([criminal].[pregnancy]) AS nonpregnancy_case_count
FROM [dbo].[criminal]
WHERE [dbo].[criminal].[SEX_CODE] LIKE 'F' AND ([dbo].[criminal].[pregnancy] IN
('4'))
GROUP BY [dbo].[criminal].[RACE_CODE];

```

### **Patient Referral Data**

```

SELECT [npidata_20150712_subset].[entity_code] AS referrer_entity_code,
[npidata_20150712_subset].[first_name] AS referrer_first_name,
[npidata_20150712_subset].[last_name] AS referrer_last_name,
[npidata_20150712_subset].[org_name] AS referrer_org_name,
[npidata_20150712_subset].[practice_address_1] AS referrer_practice_address_1,
[npidata_20150712_subset].[practice_address_2] AS referrer_practice_address_2,
[npidata_20150712_subset].[practice_city] AS referrer_practice_city,
[npidata_20150712_subset].[practice_state] AS referrer_practice_state,
[npidata_20150712_subset].[practice_zip] AS referrer_practice_zip,
[physician-shared-patient-patterns-2013-days30].[bene_count] AS bene_count,
[physician-shared-patient-patterns-2013-days30].[referring_npi] AS referrer_NPI,
[physician-shared-patient-patterns-2013-days30].[referred_npi] AS referee_NPI INTO
"shared-patient-patterns-2013-referrer-info"
FROM [physician-shared-patient-patterns-2013-days30]
LEFT JOIN [npidata_20150712_subset] ON [physician-shared-patient-patterns-2013-
days30].[referring_npi] = [npidata_20150712_subset].[NPI]
ORDER BY [npidata_20150712_subset].[last_name] DESC;

```

```

SELECT [npidata_20150712_subset].[entity_code] AS referee_entity_code,
[npidata_20150712_subset].[first_name] AS referee_first_name,
[npidata_20150712_subset].[last_name] AS referee_last_name,
[npidata_20150712_subset].[org_name] AS referee_org_name,
[npidata_20150712_subset].[practice_address_1] AS referee_practice_address_1,
[npidata_20150712_subset].[practice_address_2] AS referee_practice_address_2,
[npidata_20150712_subset].[practice_city] AS referee_practice_city,
[npidata_20150712_subset].[practice_state] AS referee_practice_state,
[npidata_20150712_subset].[practice_zip] AS referee_practice_zip,
[shared-patient-patterns-2013-referrer-info].[bene_count],
[shared-patient-patterns-2013-referrer-info].[referee_NPI],
[shared-patient-patterns-2013-referrer-info].[referrer_entity_code],
[shared-patient-patterns-2013-referrer-info].[referrer_first_name],
[shared-patient-patterns-2013-referrer-info].[referrer_last_name],
[shared-patient-patterns-2013-referrer-info].[referrer_NPI],
[shared-patient-patterns-2013-referrer-info].[referrer_org_name],
[shared-patient-patterns-2013-referrer-info].[referrer_practice_address_1],
[shared-patient-patterns-2013-referrer-info].[referrer_practice_address_2],
[shared-patient-patterns-2013-referrer-info].[referrer_practice_city],
[shared-patient-patterns-2013-referrer-info].[referrer_practice_state],
[shared-patient-patterns-2013-referrer-info].[referrer_practice_zip]
INTO "shared-patient-patterns-2013-NPI-info"
FROM [shared-patient-patterns-2013-referrer-info]
LEFT JOIN [npidata_20150712_subset] ON [shared-patient-patterns-2013-referrer-
info].[referee_NPI] = [npidata_20150712_subset].[NPI]
ORDER BY [npidata_20150712_subset].[last_name] DESC;

```

```

SELECT *,
(RANK() OVER (PARTITION BY referrer_npi ORDER BY bene_count DESC)) AS
referrer_rank,
(RANK() OVER (PARTITION BY referee_npi ORDER BY bene_count DESC)) AS
referee_rank
INTO "physician-referrals-2013-ranks"
FROM [shared-patient-patterns-2013-NPI-info];

```

```

SELECT *
INTO "physician-referrals-2013-top5referrers"
FROM [physician-referrals-2013-ranks]
WHERE [physician-referrals-2013-ranks].[referrer_rank] <= 5;

```

```

SELECT *
INTO "physician-referrals-2013-top5referees"
FROM [physician-referrals-2013-ranks]
WHERE [physician-referrals-2013-ranks].[referee_rank] <= 5;

```

```

SELECT
  * INTO "physician-referrals-2013-referee-individuals"
FROM
  [shared-patient-patterns-2013-NPI-info]
WHERE
  [shared-patient-patterns-2013-NPI-info].[referrer_entity_code] = '1';

```

```

SELECT
  * INTO "physician-referrals-2013-referrer-individuals"
FROM
  [shared-patient-patterns-2013-NPI-info]
WHERE
  [shared-patient-patterns-2013-NPI-info].[referee_entity_code] = '1';

```

```

SELECT
  * INTO "physician-referrals-2013-referee-organizations"
FROM
  [shared-patient-patterns-2013-NPI-info]
WHERE
  [shared-patient-patterns-2013-NPI-info].[referrer_entity_code] = '2';

```

```

SELECT
  * INTO "physician-referrals-2013-referrer-organizations"
FROM
  [shared-patient-patterns-2013-NPI-info]
WHERE
  [shared-patient-patterns-2013-NPI-info].[referee_entity_code] = '2';

```

```

SELECT *, RANK() OVER(
  PARTITION BY referrer_npi ORDER BY bene_count DESC
) AS referrer_rank INTO "physician-referrals-2013-referrer-individuals-ranks"
FROM [dbo].[physician-referrals-2013-referrer-individuals];

```

```

SELECT *, RANK() OVER(
  PARTITION BY referrer_npi ORDER BY bene_count DESC
) AS referrer_rank INTO "physician-referrals-2013-referrer-organizations-ranks"
FROM [dbo].[physician-referrals-2013-referrer-organizations];

```

```

SELECT *, RANK() OVER(
  PARTITION BY referee_npi ORDER BY bene_count DESC
) AS referee_rank INTO "physician-referrals-2013-referee-individuals-ranks"
FROM [dbo].[physician-referrals-2013-referee-individuals];

```

```

SELECT *, RANK() OVER(
  PARTITION BY referee_npi ORDER BY bene_count DESC
) AS referee_rank INTO "physician-referrals-2013-referee-organizations-ranks"

```



```
FROM [dbo].[physician-referrals-2013-referee-organizations];
```

```
SELECT *  
INTO "physician-referrals-2013-referrer-individuals-top5"  
FROM [physician-referrals-2013-referrer-individuals-ranks]  
WHERE referrer_rank <= 5;
```

```
SELECT *  
INTO "physician-referrals-2013-referrer-organizations-top5"  
FROM [physician-referrals-2013-referrer-organizations-ranks]  
WHERE referrer_rank <= 5;
```

```
SELECT *  
INTO "physician-referrals-2013-referee-individuals-top5"  
FROM [physician-referrals-2013-referee-individuals-ranks]  
WHERE referee_rank <= 5;
```

```
SELECT *  
INTO "physician-referrals-2013-referee-organizations-top5"  
FROM [physician-referrals-2013-referee-organizations-ranks]  
WHERE referee_rank <= 5;
```

## CHAPTER 5: ANALYSIS

When I started out on this project—interviewing women in data journalism about what it’s like to be women in data journalism—I wondered if the wage gap would be present in data journalism. In the United States and in journalism, women are generally paid less for the same work than men who have similar educations and work experiences. Yes, female data journalists reported that they were paid less than their male colleagues. But the inequality didn’t stop there.

Women said they felt demeaned by male colleagues who called them “kid,” “the intern,” “sweetie,” and a variety of other pet names. Female data journalists reported that they didn’t feel supported by editors or companies when deciding to have children. Women data journalists said they felt like men excluded, talked over and ignored them at conferences and other professional gatherings.

From the wage gap to sexual harassment, women face a number of obstacles in the workplace. Gender inequality isn’t unique to data journalism or to journalism in general. As data journalism grows as a community, so should the fight for making the newsroom an equal place for women.

\*\*\*

In the spring of 2016, I interviewed female data journalists about their professional experiences as women, both in and out of the newsroom. Twelve journalists of different ages, ethnicities and work experiences, have detailed what it’s like to be women working at the intersection of technology and journalism.

Data journalism has grown more and more popular in the last several years.

Attendance at the National Institute of Computer Assisted Reporting (NICAR) conference is at its all-time high. Journalism schools are pushing students to develop data and programming skills. Programmers not native to journalism are crossing over from the technology industry. News organizations are establishing entire teams to build apps, develop graphics and do in-depth analyses. And while data journalism grows in popularity, the number of women going to journalism school is also increasing. As more women enter the field, it is imperative that the data journalism community address complaints of sexism, harassment and wage discrimination sooner rather than later.

My research was not designed to determine if gender inequality is any different, better or worse in data journalism than it might be elsewhere. Every career field, every news medium and every beat might have similar issues, or they might not. Instead, my goal was to document how gender inequality manifests itself in data journalism and what women who work as data journalists experience on a daily basis.

Data journalism is a unique field in that no one seems to be able to precisely define it. People who call themselves data journalists range from graphic artists to reporters to news apps developers. All work with data, and all do totally different things on a daily basis. Some data journalists work in teams alongside investigative reporters in large, national newsrooms. Others work alone, serving as the go-to person in small newsrooms for all things data related. All of these jobs, skills and duties fall under the larger umbrella of data journalism.

### **Participants**

I chose to interview women whose full-time job it is to work with data, regardless of their formal job titles. Their ages range from the mid-20s to the mid-50s. They work

on both the East and West coasts, as well as several places in between. Their jobs span an equally wide range. They are designers, news apps developers and reporters. They work at some of the nation’s largest news outlets, as well as local and regional publications.

I picked my subjects on a first-come, first-serve basis with a goal of interviewing 12 women total—six from traditional print news outlets and six from web-only news outlets. I reached out via email and social media to a number of data journalists who I knew, whose names I had heard or who were suggested to me by other members of the data journalism community. During this process, the NICAR email list also endured some controversy when someone posted a shirt design featuring a bare-skinned woman, which resulted in more than 20 women reaching out to me to discuss my research. In the end, nearly fifty female data journalists volunteered to be interviewed.

In order to protect these women from potential backlash and encourage an environment of honesty, I decided ahead of time to keep names and news organizations confidential. I have given each woman a pseudonym in order to preserve their individual experiences.

Name*	Age Range	Type of Organization	Size of Organization
Becky	26-30	Online	Regional
Andi	26-30	Online	Regional
Rachael	26-30	Online	Local
Ann	26-30	Print	Medium
Melissa	26-30	Print	Large
Ellen	31-35	Print	Large
Alyssa	31-35	Print	Large
Kimberly	36-40	Online	Regional
Jenny	36-40	Online	Regional
Emily	41-45	Print	Medium
Leslie	46-50	Print	Large
Erica	56-60	Online	Regional

\*Names have been changed

### **Women in Journalism**

Research has long shown that women have different experiences in the newsroom because of gender. Women are, and have historically been, the minority in the newsroom. Dating back to the 1980s, women have typically made up about one third of the journalism workforce—a number that has not grown significantly in more than 30 years, according to a 2013 Indiana University survey on the state of American journalism. From television to magazines to newspapers, women in the newsroom have consistently faced a number of challenges in the workplace because of their gender.

In 2012, women working in journalism made 83 cents for every dollar earned by men, according to the 2013 Indiana University survey. In 1991, women earned 81 cents for every dollar made by a man. In three decades, the wage gap shrunk by only two percentage points.

Gender theories often suggest that workplaces, and some fields in particular, heavily favor and benefit males rather than females. Back in 2001, mass media and gender researcher Karen Ross wrote about the methods women use to deal with the imbalance of gender in the newsroom: incorporation, feminist, or retreat.

The incorporation approach suggests that women try to act like one of the boys. By adopting masculine traits and behaviors, female journalists attempt to fit in with and gain respect from male colleagues by acting like them.

Feminist approaches are used when female journalists feel the need to develop styles and behaviors that are different from those of male colleagues. Specifically, feminist behaviors attempt to provide an alternate voice, such as writing from a different frame or including otherwise overlooked sources. Women actively attempt to distinguish themselves from men.

Retreat happens when women choose not to involve themselves in the traditional workplace. This could mean working as freelancers, quitting a job or abandoning the field of journalism altogether.

### **Money, Money, Money**

In general, it is clear that women are uncomfortable with the idea of asking for more money, even though they know that women are often paid less than male colleagues. And in an era of tight budgets and falling revenues, it is not uncommon for news organizations to give out fewer and fewer raises.

Data journalism is also often perceived as a field that pays more than traditional reporting positions. Job applicants are required to have highly-specialized programming skills, so in comparison to other positions, data journalists might have higher salaries. But inequality among men and women is still present, even for high-paying data positions.

Jenny said her editor at a previous news organization acknowledged multiple times that she was being underpaid for the work Jenny was doing, but did nothing about it.

“It was really frustrating to repeatedly be told I was being underpaid and have no one do anything about it,” she said.

Once she realized her skills would be worth more at another news organization, Jenny immediately left the job where she was being underpaid.

As the wage gap increasingly garners attention, some companies are taking it upon themselves to try and close that gap internally. In March, the Independent Association of Publishers' Employees, a guild that represents Dow Jones employees, released a report detailing pay inequality among positions, races and genders. Dow Jones

is the parent company of *The Wall Street Journal*.

The average weekly salary of male reporters at Dow Jones is \$1,999, while female reporters made an average of \$1,800 per week. Out of all Dow Jones employees, white men made an average of \$1,773 per week, while black women made just \$1,141 per week on average. After the study was released, Dow Jones' CEO, William Lewis, and *The Wall Street Journal* editor-in-chief, Gerard Baker, affirmed their commitment to eliminating the wage gap.

Dow Jones is not the only news organization to attempt to fix an internal pay gap. But some companies have been less open about their process and results. Becky said the CEO of the news outlet she most recently worked for (she has since left her job) did a gender equity study of salaries in the newsroom. And while she appreciated the effort, she was skeptical of that evaluation for a few reasons.

First, the CEO was ultimately in charge of approving salaries, which is a fairly common practice when hiring new employees.

“If you were underpaying us, it was because you chose to,” she said. “I don’t really trust that he’s an objective person to do that and it’s not in his interest to ensure that women are paid more because he and the organization would have less money.”

Second, the results of the study were never released to the staff. Becky never saw any figures or statistics regarding pay equity in the organization.

“I had no idea if they actually found some huge wage gap or if everything was fine,” she said. “No one ever came to me and said I was getting a raise, so I guess I wasn’t being underpaid. But I really have no idea.”

Becky felt comfortable enough to ask some of her male and female coworkers

about their salaries. After those conversations, Becky said she realized her salary was on par with everyone else's pay. But she would have liked some transparency from the brass about whether other women in the organization had been underpaid.

Kimberly took a different approach. When accepting her current job, she asked if she was on the higher end of the salary spectrum. At a previous job, Kimberly made more than her coworkers and she was uncomfortable with that because she didn't enjoy feeling "an unnecessary sense of superiority."

Kimberly's situation is unusual among the women I interviewed. She said she would have taken less money if it meant her salary would have been on par with those of her coworkers, both male and female. That feeling of superiority is not something she wanted in her new job.

Kimberly said the only time she ever negotiated for a raise was when she was headhunted for another job that paid significantly more.

"[My employer] ended up counter offering" she said. "They didn't quite offer me what the other job was offering, but they bumped up my salary quite a bit."

### **Work and Parenting**

The phrase "work-life balance" earned some scoffing laughter from some of the women I interviewed. We've all heard it—raising children while working full-time as a journalist can be incredibly hard, if not impossible. A 2007 study by professors at the University of North Texas and Southern Methodist University showed that women are likely to leave journalism when they become dissatisfied with the long hours, low pay and lack of benefits. This dissatisfaction is often brought on when female journalists have children, according to the same UNT/SMU study. While not all of the women I



interviewed had children, work/life balance plays a major role in deciding when or if to start a family.

Ellen recently started a new job and shortly before starting, she was approved to be an adoptive parent. That means that at any time, she could be notified that a child is waiting for her.

“Someday, somebody’s going to call me and say ‘Congratulations, we have a kid for you,’” she said. “That could be three months from now, that could be tomorrow. This whole process could happen this week or a year from now.”

But as she found out, adopting a child comes with different legal protections than physically giving birth. The company she works for does not offer parental leave until an employee has been working there for one year. Parental leave is not the same as maternity leave.

Since maternity leave is considered short-term disability leave after physically giving birth, maternity leave does not apply to adopted children. Imposing the same time restriction on maternity leave would be discrimination, since women are considered disabled after giving birth. If someone called Ellen tomorrow with a child ready to be adopted, she would not qualify for parental leave.

Ellen said she was satisfied with the salary offered, so she tried to negotiate the leave policy. The organization remained firm— company policy states that they do not negotiate benefits. Instead, her supervisor agreed that she would be able to take an unpaid leave of absence in the event that she adopted a child before the parental leave policy kicked in. But Ellen also said that the news organization has an adoption benefit that applies on the first day, so they do provide financial support to cover some of the costs of

adoption.

Overall, Ellen said she feels relatively supported by the organization she works for in terms of parenting. While most of the people on her team don't have children, the people Ellen works most closely with all have children. She said that helps ease some of the fears she has about balancing parenting with work life.

Jenny said that having an even partnership with her husband has helped to balance her work and professional lives. She can go to conferences and travel for work without feeling like her family would suffer without her.

When I asked Jenny what it was like to have a child while working, she answered with a profanity-ridden statement about how her editors and the company were not at all supportive or helpful during or after her pregnancy.

“I had actually gone to the hospital and I think I was actually hooked up to all the machines at this point, but I'm laying there and I'm calling into my editor trying to finish this project we'd been working on because there was no preparation or anything,” she said. “It was a total disaster.”

Not only that, the news organization stopped paying Jenny and cut off her benefits by mistake while she was on maternity leave. Jenny said that happened because of the organization's unclear policies and procedures regarding how maternity leave is handled. Eventually her pay and benefits were reinstated, but the difficulties continued once she returned to work. Jenny said the company was very unclear about policies regarding new mothers.

“I was in a pumping room that like, didn't have a phone or windows or Internet, so I couldn't do any work for thirty minutes,” she said. “You know, that's thirty minutes

twice a day.”

Some companies establish lactation rooms where women can go for privacy. While federal law generally considers time spent pumping an unpaid break, Jenny said she at least wanted the option to continue working while pumping.

### **Sexual Harassment**

Think back to when you were first hired at your current job. Did you have to do an anti-sexual harassment training? Did you have to sign a form? Perhaps you watched a video or were given a pamphlet. Can you remember?

Most news organizations have official policies and procedures regarding how to handle sexual harassment in the workplace. The problem is not that those policies don't exist. The problem is that no one pays much attention to them.

“I know I had to do a bunch of stuff at orientation,” Melissa said. “I couldn't give you the details of it because I didn't pay that much attention.”

Almost everyone I talked to responded with a similar answer. Yes, they had done something. But what it was or what was said, they cannot recall.

On the whole, Melissa agreed that having a sexual harassment policy is better than nothing. But she thought those policies seemed like more of a formality rather than an attempt to prevent harassment in the workplace.

Even if news organizations require training, there is no guarantee that employees will take the training seriously. Erica said she did a mandated anti-sexual harassment training where a group of reporters had to watch a pre-recorded corporate videos.

“People were practically laughing at the video and repeating lines afterward because it was so just outlandish,” Erica said. “The basic message [of the video] was

‘suck it up.’”

On the other end of the spectrum, Andi said she had no training or information about sexual harassment given to her when she was hired at her current job.

“I think [policies] at least open the conversation,” she said. “I don’t think they’re going to stop [sexual harassment] or fix it, but I think it’s part of the package that helps. I’d like to see it, frankly.”

Leslie said that often times, sexual harassment involves a subtle pattern over time. Catching and reporting it, she said, takes a lot of vigilance on the part of other people in the office.

“It isn’t always obvious,” Leslie said. “People just mostly want to live their lives and go on.”

I asked the women I interviewed if they had ever experienced sexual harassment in the newsroom. All of them said no, but more than half said they had witnessed a coworker be sexually harassed by someone else in the workplace. And perhaps the most telling part of this question was the “but…” that typically followed the “no.”

\*\*\*

I had all the women I interviewed define what sexual harassment means to them. Everyone immediately stated that unwanted touching or overt sexual comments were undoubtedly sexual harassment. But then, most women went on to describe something less concrete, such as commenting on someone’s appearance, asking a coworker on a date or making remarks about someone’s personal life. Is that sexual harassment? Or is it just inappropriate? And most importantly, is it worth saying something?

“Blatant sexual harassment is like inappropriate touching and explicit sexual

comments,” Melissa said. “I think there’s also subtle sexual harassment like comments that make women feel uncomfortable or demeaned in some way.”

This notion of a grey area, so to speak, in sexual harassment policies is something almost every woman I interviewed brought up. Some women said they brushed off particular behaviors because it wasn’t severe enough to warrant a report to their supervisor or HR.

Ann noted one instance in which she wore an “appropriate, knee-length skirt” to the office and was greeted with an unwelcome comment from a coworker.

“One of the guy reporters yelled ‘legs’ when I walked into the office one day,” Ann said. “I just kind of laughed about it at the time, but looking back, I don’t think that was really appropriate.”

This isn’t the only situation in which Ann said she felt uncomfortable, but not enough to report someone’s behavior. And she certainly didn’t think that the incident qualified as sexual harassment. During a fellowship, Ann said she had some problematic encounters with her male supervisor.

“I did, at another job I had, have a boss who frequently asked me about my boyfriend a lot and that did make me pretty uncomfortable,” Ann said. “But I was so reliant on him extending my contract with that fellowship that I wouldn’t have done anything.”

Ann said in that instance, she didn’t feel like the man’s behavior fit the bill of what workplaces typically consider sexual harassment, but that it fell into the grey area between the obvious and the subtle.

Trainings, brochures and videos attempt to define sexual harassment. But for

some women, the line between overt sexual harassment and plain inappropriate behavior is often blurry. Women hesitate to confront coworkers or talk to management about the problem because they do not want to seem dramatic or sensitive. And even more so, women are afraid to report sexual harassment for fear of retaliation.

Alyssa had a coworker who repeatedly asked her to go out for dinner and drinks after work, despite her repeated refusal.

“I think it almost became a game to him,” she said. “This probably happened two times a week for a few months before I finally told my boss about it. He had some words for the guy, but I know we never told anyone in HR about it.”

Alyssa said that she felt like her work life would have been hurt had she told the human resources department about this man earlier since he was in a senior position in the newsroom. She had just started at the news outlet and did not want to be seen as a tattletale.

Instead, she waited and hoped that the man would give up. After several months, Alyssa caved and told her supervisor because she eventually “just couldn’t stand being in the same room” with the man. Afterward, Alyssa said she thought that incident damaged her reputation with her boss.

“You’re supposed to be tough. I’m a reporter. I have thick skin,” she said. “But I don’t think my boss ever really thought I could be that person anymore. I think our working relationship suffered because I essentially had to ask him to come rescue me.”

### **NICAR’s Gender Problem**

For many of the women I interviewed, the NICAR community epitomizes gender inequality in data journalism. NICAR staff has worked to diversify conference attendees

and speakers, but the conference still represents a space where women can see the inequality play out on a smaller scale. In panels, in social groups and at the bar, the division between men and women becomes apparent. The NICAR listserv, the email list, also serves as a point of contention for some female data journalists. And even though more women are attending and speaking at the NICAR Conference, an increase in numbers has not created an entirely welcoming environment for women.

“The NICAR bros,” as Rachael said, make it hard for women to get and stay involved in the larger data journalism community. Groups of young, white men tend to push women out or exclude them in the first place, she said.

“I’ve honestly considered not going back just because it’s a bunch of men hanging out and there’s no space for women to even try to be a part of that group,” Rachael said.

Rachael has no doubt that this camaraderie at conferences trickles into the workplace, where men are more likely to hire and promote friends or people they know well in professional circles.

“It’s not even just about the conference and the listserv,” Rachael said. “It’s that the bro attitude there carries over into how we work together, who gets hired and who gets promoted. I see it. Any woman you ask at NICAR probably sees it, too.”

That “bro attitude” Rachael described is essentially the sense of camaraderie and family between younger, predominately white men who works in journalism and technology. While bonding with coworkers is not out of the ordinary. Rachael said it is particular strong among this particular demographic of data journalists.

Rachael said she stays involved in NICAR because it provides excellent

professional development and networking opportunities. The training and career prospects that NICAR provides are essential to advancing in the data journalism world, she said.

Leslie also feels like the heavy male presence at NICAR has been detrimental to the overall organization. She said she feels like young men have somewhat “hijacked” the original purpose and intent of NICAR by turning it into a place to show off, rather than a place to support and help other journalists. Leslie notices that members are much younger than they have been in the past. But, she said, even though there are more women involved nowadays, the men are still much louder when it comes to who’s actually speaking on panels and who’s asking the questions.

“I’ve been a member of IRE and NICAR for 20-plus years now,” Leslie said. “There are times when I feel like it’s not really my group anymore... but I feel like it’s still a very useful organization.”

In the past few years, the need for a women-only space has grown into a Women in CAR Dinner at the NICAR conference, as well as a private Slack channel dedicated to women working in data journalism and development. This exemplifies the feminist approach to gender structures in the workplace. Women carve out spaces for themselves where they can support and mentor each other. Even those efforts, though, have been met with some hostility.

“We hear some guys complain about why they’re not invited to [the dinner] and how we’re being sexist by excluding them,” Erica said. “You know, I think they’re joking sometimes, but sometimes I really think they’re being serious about it.”

Establishing environments exclusively for women is a feminist approach to



dealing with patriarchal organizations, like what we see in newsrooms and in the data journalism community. Women create their own groups in order to differentiate themselves from the norms established by men.

\*\*\*

Earlier this year, the NICAR email list experienced some controversy regarding a NICAR Conference t-shirt suggestion featuring a seemingly nude woman and the phrase “clean data... dirty minds.” The shirt was a play on the Ashley Madison logo, a reference that some listserv subscribers did not understand at first glance.

This incident led to multiple lengthy threads discussing sexism, codes of conduct and other issues surrounding how women are treated in the NICAR community. Jenny, who has been a member of IRE since the 1990s, said she was outraged at how few people were willing to stand up and call out the blatant sexism.

“The fact that that wasn’t universally acknowledged as totally inappropriate is completely demoralizing,” Jenny said.

To Jenny, and to other women who subscribe to NICAR’s listserv, this incident was emblematic of a larger problem. For them, it was a defining moment in a NICAR’s history of data journalism’s gender inequality problem.

“When we can’t even get on the same page about overt sexism, how are we supposed to address the issue of things like excluding women or shouting over women or crowding them out?” Jenny said.

Leslie said she was dismayed at the casual attitudes of many NICAR listserv contributors more than the shirt suggestion itself.

“I think it was the lack of professionalism that bother me more than anything,”

she said. “Clearly there was a need to talk about it, but at some level, people were forgetting that that’s a professional listserv. And not only that, but your future boss was probably on that listserv or maybe they’re not going to be your future boss now.”

Leslie said that she and her female colleagues often don’t get involved when topics like this arise on the listserv because they see it as a professional organization where journalists go for job-related help.

“That listserv is supposed to be our professional help, not our therapist,” she said.

### **Women with Imposter Syndrome**

Imposter syndrome is a hot topic in journalism right now. Imposter syndrome is essentially feeling like you’re unqualified or undeserving of something, like an award, job or professional achievement. High-achieving individuals are unable to see that they have earned certain recognitions, which makes them feel like they just got lucky or are frauds.

According to a 2013 Agnes Scott College study, women are significantly more likely than men to report imposter feelings than men. The women I interviewed overwhelmingly said men rarely seemed to let imposter syndrome show or effect their daily work lives. Piled on top of all the other inequalities women experience in the workplace, it is understandable how women see their imposter syndrome as being worse than that of their male peers.

Unsurprisingly, age was cited as a major factor in imposter syndrome among both the younger and older women I interviewed. Despite Melissa having an impressive resume, she said being in her 20s is probably the biggest contributor to her imposter syndrome.

“All the time I feel constantly aware that I’m the youngest person,” she said.

Ann, who is also in her late 20s, echoed this sentiment, saying she was acutely aware of her imposter syndrome at her first permanent full-time job because she had been doing temporary contract work for various other news outlets before that.

“For the first four months that I was here, I walked in pretty much every day thinking today’s the day they’re going to realize that I’m just not qualified,” Ann said.

“Today’s the day they’re going to realize they made a total mistake in hiring me.”

Likewise, Andi works at a large national news organization with several men who have more job experience than she does. That alone makes her feel insecure about her abilities. Like Ann and Melissa, Andi is also between 26 and 30 years old.

“In the role I’m in, I’m definitely one of the younger people on the team, so there is an experience level differential,” Andi said. “It’s something that I’ve definitely wrestled with.”

For Emily, not having a serious coding background makes her more self-conscious about her skills and abilities. She said most of her background is in storytelling and reporting, rather than in programming. Compared to recent graduates, Emily feels like her skills don’t quite measure up, even though she’s in her 40s.

“The last couple people who filled this job either had no previous journalism experience or only a couple years of previous journalism experience,” she said. “I’ve been working mostly as a reporter and editor for the last 20 years, so I don’t have a lot of the programming background that a lot of young folks coming out of school nowadays have.”

Another telling example of imposter syndrome lies in the way women describe

themselves. They describe themselves as designers, journalists and data analysts, but they rarely say “programmers.”

“Somebody on my team told me once told me that I could call myself a programmer and I felt really uncomfortable saying that, even though I use programming for my job,” Becky said. “I don’t feel like I know enough programming to call myself a programmer.”

Even though many of the women I interviewed spend a lot of time writing code in a multitude of languages, few were comfortable describing themselves as a programmer.

“Even being able to own the word ‘programmer’ was really complicated for me,” Andi said.

Achieving “programmer” status is typically reserved for individuals who are experts in a particular language. Even though these women spend hours every day writing in Python, JavaScript and SQL, hardly anyone wanted to call themselves an expert. Additionally, Andi pointed out that the men on her team are more likely to agree to do a project, even though they may not currently possess the skills to do it.

“If you ask them later, they’ll say that they have no idea how to do it,” she said. “They’re just so confident that they can figure it out.”

Andi tries to not do that because she doesn’t like making promises she can’t keep. Being unable to figure something out or deliver on a project is scary, she said. Sometimes that makes her feel inadequate because her male colleagues constantly appear capable of doing everything the job requires.

### **“It’s the little things”**

Time and time again, the women I interviewed said the blatant, overt

discrimination is not always the biggest problem because that is often the easiest to address. Instead, it is the day-to-day interactions and comments that slowly eat away at them.

Jenny said she could tell the difference between how she was treated versus how her male colleagues were treated. Most differences were subtle, but stacked next to each other, she said those little things started to add up.

“Often times, questions, especially technical questions, would go to a male colleague, even if you were the person in charge of working on the project or were the one most familiar with the data,” Jenny said. “One reporter repeatedly called me ‘kid.’ I’m almost 40. I’m not a kid. That’s not okay. Don’t call me ‘kid.’”

Minimizing and demeaning comments from male coworkers is a problem faced by women. Alyssa said she had a coworker refer to her as “the intern” for several months after she was hired as a full-time staff member.

Others said that examples of subtle discrimination are hard to find and even harder to document because those interactions have become so nuanced. After a while, women feel defeated and start to ignore problems, rather than trying to confront issues. Melissa said she often feels like she’s being talked down to by male coworkers, even though she knows just as much or more than they do about a particular topic.

“I experience subtle sexism,” Melissa said. “I experience feeling like men aren’t respecting me like they would respect another man in my position. I don’t really know how to quantify that, though. It’s just such a normal thing that happens now.”

Instead of trying to fight the system, Ann tries to manipulate it in her interest. Her approach to this problem stands out the most among the women I interviewed. This is a

quintessential example of how women incorporate themselves into patriarchal environments in order to succeed.

“I also think that [men] have a tendency to underestimate young female reporters, and so there’s a huge opportunity to exceed those expectations,” she said. “This is the reality of the situation and of work life. I’m trying to come to terms with it and use it a little bit to my advantage.”

## REFERENCES

- Anderson, M. (2014). *As Jill Abramson exits the NY Times, a look at how women are faring in newsrooms*. Pew Research Center. Retrieved 18 November 2015, from <http://www.pewresearch.org/fact-tank/2014/05/16/as-jill-abramson-exits-the-ny-times-a-look-at-how-women-are-faring-in-newsrooms/>
- Bendl, R. (2000). Gendering Organization Studies: A Guide To Reading Gender Subtexts In Organizational Theories. *The Finnish Journal of Business Economics*, 373–393.
- Cassidy, W. (2008). Traditional in Different Degrees: The Professional Role Conceptions of Male and Female Newspaper Journalists. *Atlantic Journal Of Communication*, 16(2), 105-117.
- Cusack, C., Hughes, J., & Nuhu, N. (2013). Connecting Gender and Mental Health to Imposter Phenomenon Feelings. *Psi Chi Journal Of Psychological Research*, 18(2), 74-81.
- Djerf-Pierre, M. (2007). The Gender of Journalism. *Nordicom Review*, (Jubilee Issue), 81-104.
- Elmore, C. (2009). Turning Points and Turnover among Female Journalists: Communicating Resistance and Repression. *Women's Studies In Communication*, 32(2), 232-254.
- Elmore, C. (2007). Recollections in Hindsight from Women Who Left: The Gendered Newsroom Culture. *Women and Language*, 30(2), 18-27.
- Everbach, T. & Flournoy, C. (2016). Women Leave Journalism For Better Pay, Work Conditions. *Newspaper Research Journal*, 28(3), 52-64.
- Engstrom, E., & Ferri, A. (1998). From Barriers to Challenges: Career Perceptions of Women TV News Anchors. *Journalism & Mass Communication Quarterly*, 75(4), 789-802.
- Fink, K., & Anderson, C. (2014). Data Journalism in the United States. *Journalism Studies*, 16(4), 467-481.
- Flatow, G. (1994). Sexual harassment in Indiana daily newspapers. *Newspaper Research Journal*, 15(3), 32-45.
- Hardin, M., & Whiteside, E. (2009). Token responses to gendered newsrooms: Factors in the career-related decisions of female newspaper sports journalists. *Journalism*, 10(5), 627-646.

- Hearn, J. & Parkin, W. (1993). "Organizations, Multiple Oppressions and Postmodernism".
- Hassard, J. & M. Parker (Eds.). *Postmodernism and Organizations*. London-Newbury Park-New Delhi: Sage Publications, 148–162.
- McLaughlin, H., Uggem, C., & Blackstone, A. (2012). Sexual Harassment, Workplace Authority, and the Paradox of Power. *American Sociological Review*, 77(4), 625-647.
- Payne, L. (2003). 'Gendered Jobs' and 'Gendered Workers': Barriers to Gender Equality in Gendered Organizations. *University Of Guelph Centre For Families, Work & Well-Being*.
- Reinardy, S. (2009). Female Journalists More Likely To Leave Newspapers. *Newspaper Research Journal*, 30(3), 42-57.
- Reskin, B. (1993). Sex Segregation in the Workplace. *Annual Review Of Sociology*, 19(1), 241-270.
- Ross, K. (2001). Women at Work: journalism as en-gendered practice. *Journalism Studies*, 2(4), 531-544.
- Seymour, K. (2009). Women, Gendered Work and Gendered Violence: So Much More than a Job. *Gender, Work & Organization*, 16(2), 238-265.
- Steiner, L. (2012). Failed Theories: Explaining Gender Difference in Journalism. *Review Of Communication*, 12(3), 201-223.
- Tsui, C., & Lee, F. (2012). Trajectories of women journalists' careers in hong kong. *Journalism Studies*, 13(3), 370-385.
- Willnat, L. & Weaver, D. (2013). *The American Journalist in the Digital Age* (1st ed.). Bloomington: School of Journalism, Indiana University.



**APPENDIX**

**PROPOSAL**

**Professional Project**

Madi Alexander

Missouri School of Journalism  
Spring 2016

---

**Committee:**

David Herzog, Chair

Mark Horvit

Amy Simons

## INTRODUCTION

When I started at the University of Missouri in August 2014, my goal was to become a journalist who could use data to tell deeper investigative stories about society, government and community. Data gives us solid numbers, most of the time. Along with documents and public records, data adds a layer of validity and context to stories that anecdotes cannot. The courses I have taken and the reporting I have done has helped shape my ability to enter the professional world as a functional and successful data journalist. Having these skills not only makes me an all-around better journalist, but also makes me a more employable investigative journalist.

My goal is to make a career out of data journalism, not necessarily just crunching numbers, but using data to pinpoint larger societal problems that impact human lives. In pursuing that goal, I have taken several courses that have given me the necessary data skills: Computer-Assisted Reporting, Advanced Data Journalism, Advanced Data Visualization and Introduction to Geographic Information Systems. Through these courses, I have learned how to acquire data, analyze it, look for flaws and, ultimately, make data journalism about people. I learned to use MySQL in Computer-Assisted Reporting, which is how I landed an assistantship at the National Institute for Computer Assisted Reporting. I now regularly use SQL in a professional capacity to analyze, clean and check data for inaccuracies. In Advanced Data Journalism, I learned to use Python to clean data and scrape websites, which is useful when databases are not readily and freely accessible.

In Advanced Data Visualization, I learned how to use D3.js to build powerful interactive graphics, as well as the fundamentals of making sure graphics are truthfully

representing the data. Introduction to GIS taught me how to manage data, conduct spatial analysis and ensure that maps are truthful and accurate.

While data is obviously important to data journalism, the human element cannot be forgotten. Investigative Reporting and News Reporting taught me to look deeper and push further than what is easily accessible on the surface. Investigative reporting didn't just teach me how to be a good reporter, it taught me how to think critically and analyze information before simply repeating it in my stories.

My professional project will allow me to take all of the concepts and apply them in a professional setting for ProPublica. ProPublica is one of the leading news organizations for data journalism. Their innovation has not only earned them numerous awards, but also the respect of the journalism community and the communities they cover. Working for ProPublica will challenge me because this is the first time I will be working as a data journalist full-time. I will be operating as a professional journalist, helping other reporters acquire, clean and analyze data. Since I want to do this kind of work professionally, it is a perfect fit for my professional project.

Moreover, as I want to enter the field of data journalism, my professional analysis will focus on that. However, as a woman, I am also interested in the interpersonal impact of my gender in the workplace, which is why I have chosen to interview female data journalists for my professional analysis. Previous research shows that women in journalism, as a whole, are subjected to lower wages, fewer raises, fewer promotions, sexual harassment, and a number of other gender-specific disadvantages. I want to document the experiences of professional data journalists in order to understand the gender dynamics in the newsroom and hopefully shed enough light on the problems that

newsrooms make an effort to fix them. In doing this research, I am also preparing myself for what to expect when I do leave school and enter the professional world of data journalism. The Columbia Journalism Review, the IRE Journal and American Press Institute are target publications for this research.

## **PROFESSIONAL SKILLS COMPONENT**

### **Area of emphasis**

My project focuses on data journalism as a component of public-interest investigative journalism. I will be a data intern for ProPublica for the duration of my project. ProPublica is a non-profit news outlet in New York that focuses on high-impact investigative journalism.

### **Qualifications**

During my time at the University of Missouri, I have focused my coursework specifically on the field of data journalism. The following courses are particularly relevant to this professional project: Computer-Assisted Reporting, Investigative Reporting, Advanced Data Journalism, Advanced Data Visualization, and Intro to GIS. Additionally, I have worked at the National Institute for Computer Assisted Reporting since January 2015. My main duties at NICAR are to acquire, clean and analyze large datasets for resale in the NICAR database library, as well as outside news organization that contract NICAR for data work. I have also interned with the following publications in various capacities:

- The New York Times
- The Oklahoman
- The Journal Record (Oklahoma City, Oklahoma)
- The Shawnee News-Star (Shawnee, Oklahoma)

I also produced freelance work for the following news outlets:

- Reuters
- The Washington Post

- The Frontier (Tulsa, Oklahoma)
- KHOU (Houston, Texas)
- Religion News Service
- CNN

### **Dates and schedule**

I will begin my professional project on January 11, 2016 and end on April 15, 2016 for a total of 14 weeks. I will be working 30 hours per week, with the workday starting at 9 a.m. I will be working remotely from Columbia, Missouri. I will stay in regular contact with the newsroom in New York using Skype, Slack and email. I will also visit the ProPublica office in New York City for one week in January and potentially one week later in the semester.

### **Type of work**

My main duties will include cleaning databases, formatting databases for easier use by reporters, and analyzing databases. I will also be responsible for bulletproofing data analyses done by other reporters to ensure accuracy. I will also have the opportunity to do some writing and reporting, but the main focus of the work will be cleaning and analyzing data for use in stories by other journalists.

### **Evidence of work**

Since this professional project focuses more on data components of journalism, the evidence of my work will be proven through a combination of SQL queries for a particular dataset, bylines or contributing lines in stories, and in blog posts on the ProPublica website detailing how analyses for certain stories were done.

### **Supervision**

My direct supervisor will be Ryann Grochowski Jones, the data reporter at ProPublica. I will also report to Scott Klein, the assistant managing editor.

## THEORETICAL FRAMEWORK

### Introduction

For the professional analysis component of my project, the question I will be answering is: How do female data journalists perceive the impact of their gender on their daily newsroom experiences?

A relatively new field, data journalism presents itself at the intersection of the technology industry and journalism. But what is the definition of data journalism? Fink and Anderson (2015) studied data journalists who defined themselves as such and reported several of the common skills self-identified data journalists possessed. Fink and Anderson (2015) stated that organizational relationships with the University of Missouri, Investigative Reporters and Editors and the National Institute for Computer Assisted Reporting were “[one] of our most important findings,” (p. 470). Some data journalists interviewed by Fink and Anderson (2015) focused more on data visualizations, like charts and maps, while others focused more on interactive news applications. Still, other self-identified data journalists reported using databases and public records as a part of a beat, like the city hall or education beat, rather than working on a designated data, computer-assisted reporting or graphics team (Fink and Anderson, 2015). In the end, Fink and Anderson (2015) chose to allow data journalist to self-identify for their research because skills, job titles, organizational affiliations and work products varied greatly among journalists who considered themselves data journalists (p. 472).

This research is timely and relevant for several reasons. First, it appears to be the first of its kind. There is essentially no research on the role of female data journalists in newsrooms. Previous research has covered women as a whole in newsrooms, female



supervisors in newsrooms, and female sports journalists. As data journalism programs are becoming increasingly more popular at universities, this research will be especially relevant in the coming years. Second, this research will provide newsroom leaders and supervisors with perspective. By understanding how women view their gender in the newsroom, supervisors can make adjustments and changes to how women are treated. Third, my research comes during a time that is rife with feminist movements. From pay equality and stronger anti-harassment policies, women are beginning to take a stand both in and out of the newsroom.

This topic relates to my professional component because I am a female data journalist and will be working as such during the spring semester. Understanding how my gender influences my life in the workplace is incredibly important not only to me, but also to my future coworkers and other data journalists.

### **Gender organization theory**

Gender organization theory broadly suggests that organizations and institutions promote normative gender expectations that favor men and disadvantage women (Payne, 2003). Fundamentally, gender organization theory states that workplaces inherently discriminate between men and women, whether or not that discrimination is apparent or intentional.

Traditionally, organization theory has been presented as non-gendered, yet has typically been “written through a male perspective, culture and discourse,” (Hearn & Parkin, 1993, p. 149). Gender organization theory takes a feminist approach in that it does not inherently condone masculine concepts like rationality and hierarchy (Bendl,

2000). Rather, gender organization theory acknowledges the implicit advantage given to men in the workplace and recognizes that typical workplace structures either intentionally or unintentionally disadvantage women.

Gender organization theory also gives way to two important notions that are also important in journalism: the glass ceiling phenomenon and the glass elevator phenomenon (Bendl, 2000). The glass ceiling is the unbreakable barrier that prevents women from rising to the top in an organization. The glass elevator suggests that men in female-dominated fields rise to positions of power faster than women in the field.

Traditional measures of gender as a variable include evaluating job satisfaction, leadership, use of power, organizational commitment, sexual harassment and work/family balance (Bendl, 2000). In the workplace, these indicators can reveal that an organization may seemingly be gender-neutral and unbiased, while unknowingly perpetuating a patriarchal hierarchy of management and interpersonal interaction.

Previous research, such as Elmore's study on sexual harassment in newsrooms, has also shown that a majority of women in journalism have experienced some form of sexual harassment, abuse or assault on the job. This is just one of the few measures of how a particular field may not outright endorse a discriminatory behavior, but in effect does relatively little to correct or even acknowledge a problem. Using all of these various indicators, it is possible to evaluate how gender organization theory is exemplified in newsrooms.

### **Gender organization theory in newsrooms**

According to the Pew Research Center, the number of women in newsrooms has largely remained unchanged in the last 15 years. In 1998, women made up 36.9 percent of newsroom employees (Anderson, 2014). In 2012, women made up 36.3 percent of newsroom employees (Anderson, 2014). The number of women in newsroom supervisor positions increased to 34.6 percent in 2012 from 33.8 percent in 1998 (Anderson, 2014). Women are still the minority in newsrooms and news organizations do little to make space for women. While the newsroom may be touted as a gender-neutral environment, the organization is in fact organized “around a man-as-norm and woman-as-interloper structure,” (Ross, 2001, p. 535). In order to fight the patriarchal structure of the newsroom, women often feel the need to adopt more masculine traits as a way to succeed. Female journalists embrace male norms in “an effort to break stereotypes,” (Elmore, 2007, p. 18).

Essentially, women feel the need to be more like men in order to be successful journalists. This is also true of women in supervisory positions in newsrooms. Previous research has shown that “women in positions of leadership who take on more masculine traits often are perceived as more competent within organizations than are women leaders with more feminine traits,” (Elmore, 2007, p. 25). Women in journalism have also recounted instances where male coworkers felt the need to rescue to save them (Elmore, 2007).

Another indicator of the biased workplace is the refusal of some women to acknowledge the extent of the bias for fear of seeming like a victim. Previous research has shown that “many women in journalism have denied or downplayed the patriarchal newsroom culture, and that many women in general do not recognize or want to

recognize the larger social issues of discrimination against women,” (Elmore, 2007, p. 26). That lends itself to another issue: How do you fix a problem when the people it impacts aren’t willing or able to admit that the problem exists because that might make them look less like equals?

## **Conclusion**

Even though there is a significant lack of research in this area, the gender aspect of data journalism must be acknowledged. As a developing and rapidly growing field, women who work in this particular area of journalism face the same battles as women in other areas. It is crucial to evaluate all the ways in which women feel they are disadvantaged, targeted and discriminated against simply because of their gender. In order for more women to succeed in this field, newsrooms can no longer ignore gender roles and the perpetuation of inequality within their organizations. That is why my research will center on this question: How do female data journalists perceive the impact of their gender on their daily newsroom experiences?

## **METHOD**

### **Interviews**

To investigate this research question, this study will use interviews to qualitatively assess how female data journalists view their gender in the workplace. Interviews will be a useful tool for this research because they will allow me, the researcher, to go more in depth with each subject than a survey would allow. Semi-structured interviews are common among mass media researchers, especially when studying women in the media, because they allow for more detailed qualitative data to be gathered. For example, Elmore (2009) used open-ended interviews to research women who left journalism because interviews “allowed participants to raise issues on their minds that were not imposed upon them, related to their newspaper experiences, turnover decisions, or subsequent roles,” (p. 236).

Additionally, a semi-structured interview will allow for follow-up questions in order to dig deeper into gender issues in the workplace. Tsui and Lee used in-depth interviews to research female journalists in Hong Kong because interviews allowed them “to tackle the complexities involved in how various forces interact to shape individuals’ careers,” (p. 371). Using in-depth semi-structured interviews is appropriate for this research because it allows for deeper qualitative investigation and provides a more complete picture than other methods, while also covering similar topics with each interviewee.

### **Subjects**

At least 10 female data journalists will be interviewed, half coming from traditional newspapers and half coming from online-only news outlets. For this research,

I am seeking to interview female data journalists. “Data journalist” will include journalists whose full-time job it is to acquire, clean, analyze, visualize and report with data, or some combination of the aforementioned duties. These journalists could be working as the sole data journalists in a newsroom, as part of a team (i.e. if a data journalist were assigned to the business beat), or as a part of a larger data team. I do not intend to include beat reporters who use data in the course of their reporting i.e. an education reporter who sometimes uses data, but is not primarily tasked with doing so.

I will attempt to interview data journalists with diverse educations, ethnicities, ages, sexual orientations, and numbers of years working in journalism in order to get a broad, intersectional perspective of what it’s like to be a woman in the newsroom. I will document demographic data about the interviewees, as well as ask for information about their current employers.

I will allow interviewees to remain confidential unless otherwise agreed upon with the interviewee. Since I will be asking about sensitive topics like sexual harassment, I do not want to reveal the identity of someone who has been sexually harassed, abused or assaulted without their explicit permission. Given how frequently this happens to female journalists, it is not unlikely that more than one of my interviewees will have experienced sexual violence or harassment in the newsroom. I also do not want interviewees to face backlash or punishment from current employers if they reveal that their current employer does not foster a safe and inclusive environment in the newsroom. Additionally, I will have an interview guide in order to ask the same questions of each subject, while also leaving room for follow-up questions.

## **Interview guide**

- How old are you?
- What is your race or ethnicity?
- What is your sexual orientation?
- How long have you worked in journalism?
- What is your current job title?
- Who is your current employer?
- Are you in a supervisory position?
- Tell me about your role in the newsroom. What is a typical day like for you?
- What is the gender breakdown of the people you work most closely with?
- How confident are you in your ability to negotiate for a higher salary or ask for a raise?
- Do you have feelings of “imposter syndrome” in your professional life? Why do you think that is?
- How do you think your employer promotes or inhibits gender equality?
- What do you think your employer could do to better promote gender equality?
- How would you define sexual harassment?
- What kinds information or training has your employer provided to you on sexual harassment? How effective do you think that is?
- If you reported sexual harassment in your current workplace (happening to you or someone else), what kind of impact do you think that would have on your career?
- Can you tell me about any instances of sexual harassment you have seen or experienced in the workplace, either at your current or past employers?

- As a whole, how would you describe your coworkers' attitudes toward family?
- Have you ever been asked about family, marriage or children in a job interview?
- In terms of maternity leave, health benefits and other support, how would you characterize your employer's attitude toward motherhood?
- What kind of impact do you think a child would have (or had) on your career?
- What are some barriers you have experienced during your journalism career?
- How do you think gender inequality has personally impacted you?
- What advice would you give to young women who want to be data journalists?



## LITERATURE REVIEW

### Introduction

Women have long been the minority in the newsroom. Dating back to the 1980s, women have typically made up about one third of the journalism workforce—a number that has not grown significantly in 30 years (Anderson, 2014). From television to magazines to newspapers, women in the newsroom have consistently faced a number of challenges in the workplace simply due to their gender.

At its core, gender influences human interaction and ultimately who is given a preconceived idea of power in society and in the workplace. With gender comes certain powers, or lack of powers, in the workplace (McLaughlin et al., 2002). These power structures create an environment typically centered on male-dominated norms and patriarchal values. Men are undoubtedly the ones with power. For women in journalism, these patriarchal power structures manifest themselves in a number of ways: sexual harassment, wage gaps, gender-based favoritism in promotions and more (Elmore, 2009).

This research attempts to answer this question: How do female data journalists and interactive news developers perceive the effect of their gender in the newsroom? Even though previous research usually does not focus on a specific field or beat, but rather journalism as a whole, it is important to understand the general role of gender and power in newsrooms and how the male-dominated structure of the newsroom is already alienating and disadvantaging women. By examining the role of gender in the workplace, discussing how female journalists perceive themselves as a part of the newsroom, and evaluating research on gender-based discrimination, this literature review creates a

roadmap for determining how newsrooms could do a better job of ensuring that women are not outcasts in a male-dominated workplace.

### **Gender in the workplace**

Gender is central to how everyone interacts in the workplace, at home and in public. At the very forefront of human interaction, gender identities are “are fundamental to the everyday relations and conduct of normal life,” (Seymour, 2009, p. 240). Seymour (2009) states that this relationship means power and control are fundamentally connected to gender identity. Gender is “a complex and powerfully effective domain of social practice” that has the proven potential to control vast parts of society and human interaction (Seymour, 2009, p. 240). Gender is not simply an identity, but rather a defining characteristic that is inherently associated with power and dominance in society (Seymour, 2009). The relationships between gender, sex, power and the workplace are inseparable, since all four of those aspects commingle to create specific environments (Seymour, 2009). Reskin (1993) suggests that explanations for sex-segregated workplaces either focus on occupational choices of males and female or on the actions of the employer.

The first explanation focuses on how males and females prefer different structures and characteristics in professions (Reskin, 1993). Namely, women tend to prefer workplaces with more stable hours, shorter hours and more reliability (Reskin, 1993). The second explanation deals with how employers treat the different sexes, mainly in terms of sexism and discrimination. Sex-segregated workplaces are perpetuated when women are discouraged from applying for certain positions because they believe their sex

puts them at a disadvantage (Reskin, 1993). Men are also more likely to hire men, thus perpetuating male-dominated workplaces (Reskin, 1993). This again ties into the perpetuation of a patriarchal culture in male-dominated professions like journalism.

### **Women in the newsroom**

Female reporters are no exception to fulfilling organizational roles imposed upon them by the male-dominated field. According to Steiner (2012), female reporters “enjoyed substantially diminished journalistic authority” in the workplace (p. 209). Ross (2001) states that what is perceived as a neutral environment in journalism is actually inherently skewed to favor men:

Newsroom culture that masquerades as a neutral “professional journalism ethos” is, for all practical (and ideological) purposes, actually organised around a man-as-norm and woman-as-interloper structure. The consequences for women who choose to work in this male-ordered domain, then, are to either beat the boys at their own game by becoming more assertive and more macho or else to develop alternative ways of being a journalist, often by working in concert with other professionals who are on the outside of the inner circle by dint of the same or different reasons for exclusion, for example, journalists of colour, self-identified gay and lesbian journalists, and so on. (p. 535)

Women cope with this disenfranchisement in one of three ways: incorporation, feminist, or retreat (Ross, 2001). The incorporation approach suggests that women act like “one of the boys,” imitating masculine style, values and reporting technique (Ross, 2001, p. 535). Feminist approaches are employed when journalists “make a conscious decision to

provide an alternative voice, for example, writing on health in order to expose child abuse and rape,” (Ross, 2001, p. 535). The retreat from disenfranchisement happens when women actively chose to not participate in the workforce and therefore operate as freelancers outside of a traditional newsroom (Ross, 2001). With all three of these approaches, women have to make conscious evaluations of where they fit into journalism and how their genders play a part in their reporting.

Cassidy (2008) determined through a survey of men and women journalists that women reporters have carved out their own interpretations of how their gender determines their role in the newsroom. Women and men were asked to gauge the importance of certain professional roles, such as “interpretive/investigative” and “disseminator,” (Cassidy, 2008, p. 111). The disseminator role is interpreted to mean the quick relay of information to the public and the avoidance of “unverifiable facts,” (Cassidy, 2008, p. 111).

In the study, women placed more emphasis on the traditional disseminator role than did men (Cassidy, 2008). Based on the results, Cassidy (2008) states “female newspaper journalists are even more aligned with the majority culture of their profession than the men,” (p. 114). Cassidy (2008) suggests that this might be because female journalists feel pressured to adopt the more traditional norms of the profession in order to fit into the male-dominated culture of the newsroom. Female journalists, according to Cassidy (2008), actually end up being more closely aligned to the values of the profession simply because of the organizational pressure to behave like male journalists.

In addition to perceiving their roles differently, women and men also have different views on how gender plays a role in the newsroom. Djerf-Pierre (2007)

surveyed Swedish reporters about what role gender plays in reporting. More than twice as many women agree, “[male] reporters are favored in the assignment of stories,” (Djerf-Pierre, 2007, p. 95). Fifty-one percent of women in the study reported that they think men receive that favoritism, while only 21 percent of men agreed with the same statement (Djerf-Pierre, 2007). Additionally, 28 percent of men agreed that editors assigned stories based on gender, compared to 48 percent of women agreeing with the same statement (Djerf-Pierre, 2007).

Gender norms also work their way into journalism through certain beats being considered more feminine or masculine (Tsui and Lee, 2012). Arts and cultural beats are considered feminine, while business and sports are interpreted as the masculine fields (Tsui and Lee, 2012). In this case, both the newsroom and the subject matter being covered in the newsroom are interpreted as being gendered. Additionally, some female reporters face other barriers that are compounded into the already difficult requirements of being a journalist. For broadcast journalists, one of those challenges is appearance. In Engstrom and Ferri’s (1998) study, female broadcast journalists reported that an overemphasis on physical appearance, on top of long hours and low pay, was one of the biggest challenges in their careers:

One, for example, commented, "Men can grow old, have facial hair and be bald, where women must be young, pretty and perky." Another wrote, "A male anchor can age, gray, gain weight, wear glasses, and he's considered seasoned, a woman is considered old." Especially telling were comments that showed how appearance could detract from an anchor's credibility, such as, "A bad hair day negates what I'm saying on the news. I get more comments on my clothing and make-up than on stories." (p. 794) Even though

men who appear on television news might also feel some of the same pressures, research shows “women bear the burden of ‘looking good’ more so than men do,” (Engstrom and Ferri, 1998, p. 798).

### **Why women leave journalism**

In addition to working in an already challenging field, women also face a number of other barriers in the workplace, ranging from criticisms of their appearance to long work hours that conflict with their ability to have families. Elmore (2009) discusses the systematic disadvantage given to women in the workplace and how that creates dissatisfaction and ultimately motivates women to leave the field of journalism. Female newspaper journalists “report higher levels of exhaustion and lower levels of professional efficacy than do men,” (Reinardy, 2009, p. 42). In addition to cataloging exhaustion, Reinardy’s (2009) study reveals that the overarching male-dominated attitude and culture of the newsroom discourages women from continuing their careers in the field:

So while women are feeling overburdened at work, which may be leading to exhaustion and reduced feelings of professional efficacy, they are not receiving the organizational support they seek. Faced with family and childcare issues, sexism, discrimination and the glass ceiling, it appears women are not receiving the organizational encouragement to contend with these issues or perhaps even to remain in journalism. (p. 53)

Men are more likely to accumulate an advantage based on their efforts and be rewarded for their professional performance (Elmore, 2009). Thus, women recognize the gender discrepancy and seek out other avenues and careers for validation (Elmore, 2009). Djerf-

Pierre (2007) also asserts that women in journalism feel frustrated because they are “systematically disadvantaged” in the workplace because of their gender (p. 96).

Female sports journalists also expressed this same dissatisfaction with patriarchal newsrooms, which impacted their decisions to remain in or leave the field (Hardin and Whiteside, 2009). Women who left their positions as sports reporters did so because they “felt no sense of accomplishment and saw themselves as lacking power to achieve professional goals,” (Hardin and Whiteside, 2009, p. 640). The institutionalized glass ceiling for female journalists and “patriarchal culture” of newsrooms were key factors in decisions to quit the field altogether or move on to different outlets (Elmore, 2009, p. 248). Flatow (1994) also asserts that women “are second-class citizens” in the newsroom and that their gender is used as a weapon to maintain that status (p. 42). This patriarchal attitude alienates female journalists, creates a hostile environment for women and ultimately results in women leaving the newsroom (Flatow, 1994).

## **Conclusion**

Women in journalism face a myriad of obstacles in their careers. Men with similar experience and education are paid more than women. Sexual harassment policies do little to protect victims. Long working hours deter women from having children. Male editors tend to favor men for better and more prestigious assignments. Women feel more pressure to maintain a certain look, style and weight. This has become the norm for most women in journalism and the pressure, combined with the lack of basic dignity, is causing them to leave the field. In an age where more women than men are graduating from journalism school, newsrooms will eventually have to come to terms with the fact

that their male-dominated structures are going to lead to high turnovers and dissatisfaction if they do not change. If news organizations truly want to make the workplace an environment that is welcoming to diversity, the way female journalists are treated needs to change.