

From the stars back down to Earth

Getting a job as a data scientist

Sarah Jane Schmidt ([@sjs917](#))

German Astronomical Society Meeting 2021

Shape of this talk

- About me
- My decisions
- Data jobs and skills
- The experience of applying
- How things are now
- Advice

About Me

CV

- Data Scientist, everphone (2021-?)
- Leibniz Institute for Astrophysics, Karl-Schwarzschild Fellowship (2015-2020)
- The Ohio State University, Columbus Prize Fellowship (2012-2015)
- University of Washington, PhD in Astronomy (2012)
- Barnard College, BA in Physics and Astronomy (2006)

Identities

White

US citizen

Cis-woman

Bisexual/queer

Polyamorous

Chronic depression

Generalised anxiety

Realising academia is not for me anymore

- I never seriously considered a career outside astronomy until I got fully rejected from faculty positions in my last round of applications. From that point on, I was sure I was done.
- My productivity had plummeted as I felt less comfortable and welcome in parts of the astronomy community, and as I dealt with a divorce and mental health issues.
- I was tired to working independently, completely uninterested in moving to a new town and recreating my community, and exhausted by academic expectations of super-stardom.

**Leaving academia also meant
recreating my identity and
mourning my past self during the
height of pandemic isolation.**

What do I want from work?

How did that lead to data science?

- The stability and consistency of **full time employment.**
- A salary similar to or greater than my postdoc salary.
- To use some of the **same skills** I used in astronomy.
- Enough space to live the rest of my life.
- To stay in the same city, even if I wanted to **change jobs**

Data jobs (incl. analyst, scientist, and engineer) are common in Berlin and in many places with a strong presence in tech and/or startups.

Data jobs in industry

What are they like?

Use data to make better business decisions
Work in SQL, Python, Excel, and/or dashboards
Member of engineering or finance or business intelligence

Data Engineer

Builds database infrastructure
So much SQL
Close to developers

Data Scientist

Models data
SQL and python

Data Analyst

Puts together reports
Excel and dashboards
Close to finance

Skills overlap

- Python (pandas, matplotlib)
- SQL
- Github
- Data cleaning
- Problem solving
- Fitting models to data
- Statistics
- Visualisation
- Communication
- Teamwork/mentoring

I chose to skip major re-training (i.e. bootcamps). I sharpened my SQL and moved from IDL to python while I wrote applications.

Sarah Jane Schmidt (@sjs917)

Reframing skills to match data jobs

- The overlap was never complete - either looked for specific industry experience or specific tools/python packages that I didn't know.
- Knowing 50-75% of the specific skills listed in a job ad is enough to apply. Some of it is not essential, and it's hard to tell what from the job ad.
- People who have completed PhDs are generally better than average at independent work, learning new things quickly, and solving complex problems.
- A team you want to work on will be open to helping you fill in the gaps; if they need very specific experience they are probably a bad fit.

The data science resume (Or at least, mine)

- More action than accomplishment
- More skills than scientific results
- Ask google, ask social media, and look at all examples you can find (here is mine)
- Read and match job ads where possible
- ASK FOR FEEDBACK from anyone willing to give it

Sarah Jane Schmidt

Ex-astronomer looking to create data-driven solutions to problems on Earth.

sarahjaneschmidt@gmail.com
www.sarahjaneschmidt.com

SKILLS

Technical

- **Programming:**
 - Python (proficient)
 - SQL/postgres (proficient)
 - IDL (expert)
- **Tools:**
 - Git/github (familiar)
 - Excel/spreadsheets (proficient)
 - Command line (proficient)
- **Analysis:**
 - Data cleaning (proficient)
 - Data visualization (expert)
 - Statistical reasoning (proficient)
 - Hypothesis testing (familiar)
- **Modelling:**
 - Classification (proficient)
 - Dimensionality reduction (proficient)
 - Regression analysis (expert)
 - Chi-squared minimization (expert)

Communication

- **Written:** Authored more than 60 papers.
- **Presentation:** Delivered over 30 seminars to both specialist and general audiences.
- **Teaching:** Designed and gave lectures and coursework for 3 university courses.
- **Workshops:** Led workshops on data visualization, inclusive mentoring, and vocal technique.

Leadership

- **Mentoring:** Supervised ten students through data-focused research projects.
- **Team Lead:** Weekly research team meetings.
- **Committee Lead:** Chaired the executive level inclusion committee for SDSS, a 2,000 member international collaboration.
- **Collective:** Practiced non-hierarchical leadership within choir and feminist activist spaces.

EXPERIENCE

Research Fellow

Leibniz Institute for Astrophysics, Potsdam
September 2015 - September 2020

- Led a four person team that used machine-learning techniques (e.g., k-nearest neighbor, gaussian processes) on large datasets to predict the ages of stars.
- Designed and analyzed a survey to quantify gender-bias in how audiences question conference presenters and used the results to train moderators in best practices.

Research Fellow

The Ohio State University, Columbus
September 2012 - August 2015

- Pioneered a new method to model sparse time-series data of stellar magnetic outbursts using chi-squared minimization.
- Verified the results of a large, collaborative data analysis pipeline for ten thousand stars.
- Used regression analysis to calibrate trends between the chemical composition of a star and its observed color.

Research Assistant

University of Washington, Seattle
September 2006 - August 2012

- Extracted and cross-matched data for twelve thousand stars from five different databases to identify trends in the evolution of stellar magnetic fields.
- Wrote custom IDL code to process thousands of multi-dimensional data files to measure the temperature of magnetic stellar outbursts.
- Analyzed a dataset of grades for 100,000 students to measure the mean increase in performance caused by a mentoring program targeted at a diverse group of students.

EDUCATION

PhD in Astronomy

University of Washington
August 2012, Seattle

BA in Astronomy & Physics

Barnard College
June 2006, New York City

LANGUAGES

- English (native)
- German (A2)

WORKSHOPS

Data Scientist in Python Path

Dataquest, February 2021 - present
Interactive online course with a broad, data-focused curriculum.

Machine Learning for Astronomers

September 2020
Introduction to supervised and unsupervised learning techniques.

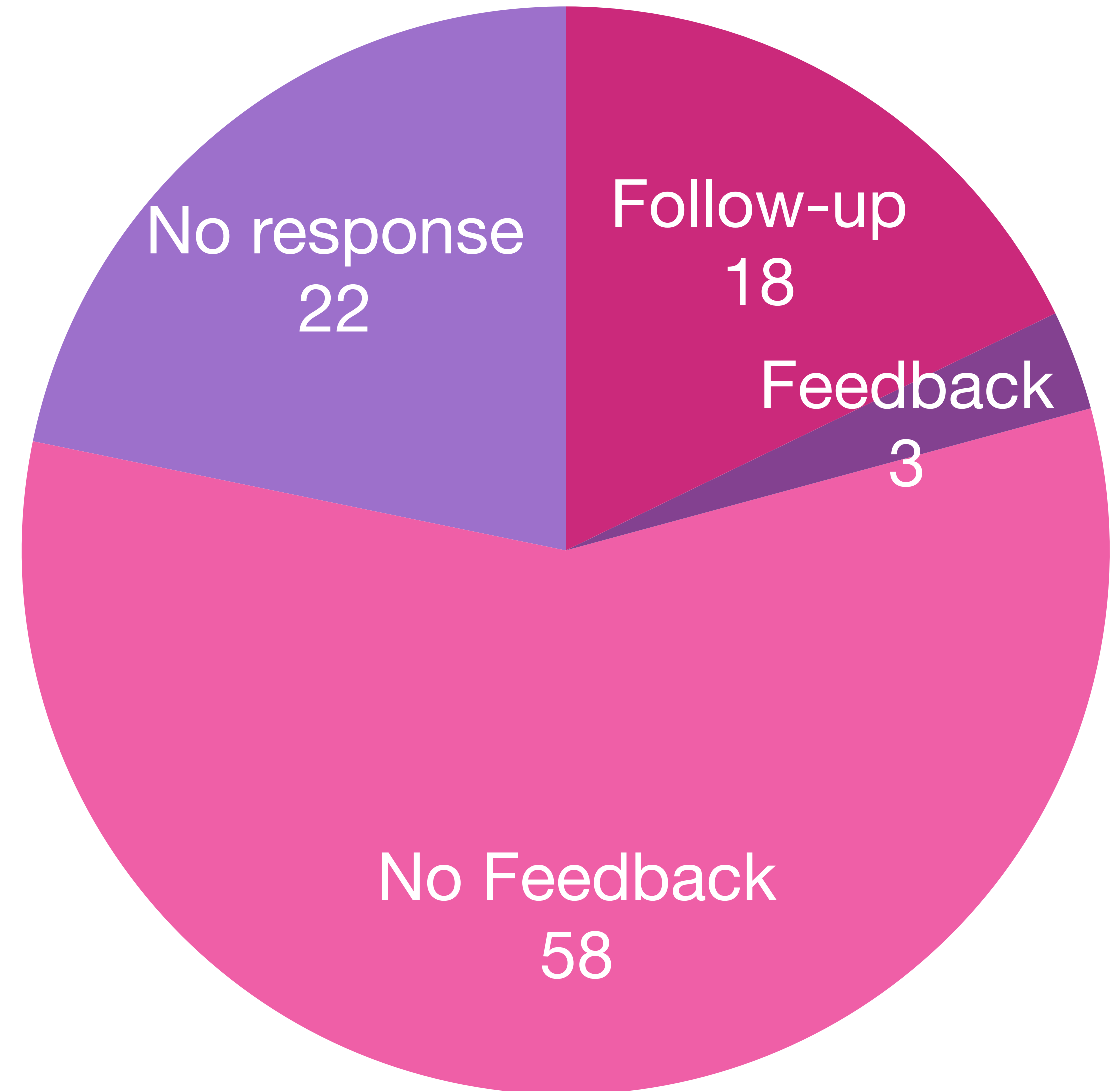
Astro Hack Week

August 2019
Included tutorials on pytorch, keras, tensorflow, and visualization.

Sarah Jane Schmidt (@sjs917)

Applications by the numbers

- Three months of serious effort
- 101 applications
- 22 no response
- 58 rejections with no feedback
- 3 rejections with feedback
- 18 companies requested follow-up



Job titles:

55 data scientist, 36 data analyst, 7 data engineer, 3 other

Sarah Jane Schmidt (@sjs917)

What did follow-up look like?

In general, HR interview -> tech test -> team lead interview -> senior interview

- For 18 companies, I did a total of 26 interviews and 8 tech take-home tests.
- Only one of these interviews was in person!
- Ten rejected me after the first step (feedback was helpful though).
- Two decided to interview me for different positions than I applied for.
- Went through the entire interview process with 3 companies
- Got two verbal offers (in the same day), accepted one.
-

On interviewing and tech screenings

- Some companies have hiring practices that deeply disrespect their candidates, like asking for 8+ hours of work on a tech test before meeting you. It can be worth having firm boundaries around what you will do.
- Every interview and tech screening gave me the opportunity to learn something about these jobs, myself, or the company.
- My application materials improved A TON after I started to see what questions interviewers would ask me. I learned to frame my experience in ways that they could interpret and understand it.
- The best interviews were easy - it was a chat with someone I wanted to work with about whether or not my skills fit the needs of the team.

It was difficult.



Image credit: [KC Green](#)

... I was not fine.

Sarah Jane Schmidt (@sjs917)

Four months into my data science job

How does it compare to what I thought?

- At first, it was overwhelming how much there was to learn, and how unqualified I felt to make much progress.
- Despite my own feelings of slowness, my manager says I adapted quickly.
- Company processes for working and collaborating - I very much want to work on teams, but my default is to be independent.
- My main growth area has been going from coding to software engineering - writing code that is understandable and usable by others.
- Manager and team have made me feel welcome and helped me grow!

Is it really better for my mental health?

- Yes.
- I can leave work at work, and take genuine vacations.
- I am not constantly worried that I haven't achieved enough.
- But.
- I am still trying to settle in to my post-academia identity.
- Plagued by some persistent visa issues that should be fine, but not settled.

Advice

- Strengthen your support network.
- Know your own worth.
- Ask yourself what you want and need.
- Gather as much info as you can.
- Solicit and accept feedback on everything.

Resources

Me

[Website](#)
[@sjs917](#)

Job wisdom
[Ask a manager](#)

Job ads

[linkedin](#)
[Indeed](#)

Tech skills

[Dataquest](#)
[SQL zoo](#)