

2021 Data Skills Survey Report

Spring 2022



Table of Contents

<u>Background</u>

<u>Methodology</u>

Demographics

Education & Opportunity

Job Roles & Responsibilities

Time Spent Working

The Data Position Title is Important

Final Remarks and Recommendations

<u>References</u>



Background

The DataedX Data Skills survey conducted between September and October 2021 helps us understand the data skills landscape for Black professionals. Publications like the <u>Harnham Diversity Report</u> [1] break down the demographics of race, age, gender, and more for the state of data professionals. The goal of this survey however is to empower job seekers, career changers, along with new and seasoned data workers alike with information about the most sought after skills and tools in the increasingly lucrative job market. 140 responses were analyzed to tell the story of mainly Black people with data skills, job roles, and proficiencies in 2021.

Data skills have consistently increased in marketability, as well as the compensation. Not only are companies ramping up their data strategies, but there is an emergence of more data companies whose focus is making data tools to meet demand. In 2020, Snowflake, a cloud data warehousing platform, had the biggest software technology company IPO ever [2]. This pattern likely won't slow down. The need to inform and empower Black professionals pivoting into the data industry so they can maximize their earning potential given their skills.

Methodology

The data was collected as part of a survey between the months of September and October 2021, and shared on several online platforms.



Some respondents heard about the survey through multiple platforms

Data practitioners were asked about their location, age, income, and education levels. They also responded to questions about their jobs and roles, comfort levels with different data skills and tools, years of experience, and time spent on different tasks during the work week (expected and unexpected). **140 people participated** after learning about the survey through social media (Twitter and Instagram), LinkedIn, email, and in online community spaces.



The data was processed and cleaned using Excel and python in Jupyter notebooks. Most of the visualizations were created using Google Data Studio and slide images.

Demographics

For this online survey, most of the respondents were based in the United States. The U.S. also has some of the highest pay rates for professionals in data-related fields [3].



Across the top responsive race groups and ethnicities, the majority identify as women. Black women in particular are the most prominent respondents. They are spread out across different age groups with over 50% being professionals under the age of 40. The table below displays the race and gender identities shared by the respondents.

Race	Female	Male	Non-Binary	Prefer not to answer
Black	39.3%	8.6%	0.7%	-
White	29.3%	10%	0.7%	-
Asian	4.3%	2.9%	-	-
Multiracial	-	1.4%	-	0.7%
Hispanic	-	0.7%	0.7%	-
Pacific Islander	-	0.7%	-	-



Education & Opportunity

Another informative demographic we explored is education where respondents indicated their highest degree level. Companies still weigh degrees in high regard making them a useful indicator of job titles and income. It is interesting to observe that close to three quarters of participants hold at least a Master's degree, with others having terminal and professional degrees or a combination of both.



Paired with the revelation that the largest segment of our respondents were Black women and in the U.S., we see another factor impacting data literacy, equity and - Opportunity. The American Association of University Women (AAUW) <u>reports that women are more likely than men to earn degrees</u>. In the U.S., "Black women get 64.1% of bachelor's degrees, 71.5% of master's degrees and 65.9% of doctoral, medical, and dental degrees."



Below are the demographic compositions for each income level. For each level, we observe the percentage of respondents whose earnings match the income category; along with the breakdown of race and ethnicities of individuals who reported said income. The income buckets are not widely distributed with a majority of respondents (26%) reporting \$75,000 - \$100,000 annual earnings, while the least amount of people (12%) in the survey make less than \$50,000 a year.

Though most respondents earn over \$50,000, white data professionals make up nearly **18%** of earners below this amount with Black data workers making up **over 50%** of the lowest income bracket.





Job Roles & Responsibilities

Next, we explored the job positions and their corresponding duties with the survey responders. We have a healthy ratio of team members to team leaders in our survey. In speaking to opportunity, there are different avenues for data literacy enthusiasts to enter and stay in the field. For historically excluded groups, the decisions that affect our communities happen in rooms that most of us aren't allowed to enter.



Implementing equity in data operations has a lot to do with effective leadership participation as much as knowing how the code is designed. Increasing leadership participation among Black and Brown data professionals is an important next step. This phase of data empowerment involves getting into those rooms and having more absent names mentioned and championed.

Time Spent Working

We explore how individuals in different job roles divide their time between five data task categories.

1. Data Acquisition and Cleanup focuses on data sourcing, collection and (re)formatting.

2. Data Storage and Management considers how to organize data collected for effective handling of complex requests.

3. Data Analysis leverages algorithms and processes to transition raw data into actionable insights or information.



4. Data Visualization is centered on presenting findings in a visual form to help data practitioners determine if their analyses make sense.

5. Data Storytelling (communication) addresses the data practitioner's journaling of these phases alongside the dissemination of insights in an accessible manner for non-data professionals.

Of those who responded to the question, a majority of the primary job responsibilities for data professionals are data analysis tasks - including coding and transitioning data from its raw form to information and actionable insights. This is followed in popularity by data storytelling - the all important skill of using data to express a narrative to stakeholders and non-technical audiences. Data storytelling has to do with the ability to create persuasive documentation and communicate effectively using data. This skill is highly transferable for new data professionals who previously worked in customer centered domains.



Though data acquisition and cleaning - the core of data engineering - are not identified as primary data tasks for most professionals, they feature prominently and in different forms as **unexpected data responsibilities** such as database querying, programming, engineering, storage, and even visualization as data transformation and cleaning is an integral part of the data visualizing process. A lot of data practitioners do not think of acquisition, cleaning, and transformation as its own step, rather a preparation for analysis.



The Data Position Title is Important

Job titles in any industry are an external indicator of a professional's status in the workplace. Data is everywhere. In our survey, we observed that 18.5% of data professionals identified as researchers in an academia setting. The same skills when applied in industry would have different job titles i.e. data analyst, AI researcher as opposed to associate professor. The rate at which promotions occur in industry versus academia also changes how workers are paid. In academia, tenure and professorship can take decades with some only experiencing raises for cost of living increases.



In the U.S. for instance, changing academic locations also doesn't mean the job title can change as the system is more standardized and enforced countrywide. Meanwhile, a person who starts out as a data analyst can move companies twice in three years and double their salary along with a promotion to senior roles.

We therefore cannot weigh the salary, titles, and responsibilities the same when informing data professionals because they are not uniformly assigned based on skill and experience levels.

With the level of uniqueness in job titles, 69% of professors and researchers in our survey reported making between \$50,000 and \$100,000 a year. 68% of our analyst respondents also fell in this range. Those with "scientist" in their job title had a majority (9 of 12) making over \$75,000 a year in 2021. On the other hand, most respondents who identified as engineers (7 of 10) made over \$125,000. Scientists

1000 unique job titles out of 140 respondents



and engineers in the data industry garnered salaries above the US national average of about \$50,000. Unfortunately, those in academia or analyst professionals weren't commanding over \$50,000 a year salaries. These varying observations illustrate some instances of how job titles in different industries inform pay bands more than education and skills do.

Final Remarks and Recommendations

Women of color, and Black women in particular have always needed to provide more proof of competence than their white male counterparts. With increased access to online resources and interactions, there is opportunity to inform while eradicating the gatekeeping tactics. Additionally, income transparency helps more people enter and sustain careers in the data field regardless of their educational background, geographic location or socio-economic status.

We encourage you to share this report with your teammates and leadership group in your organization as a professional development planning resource for data teams. As an individual, now that you know the five data skills categories, pay closer attention to which data tasks you do regularly. Consider enhancing them, and check out datacareerpaths.com for resources about data degree programs in your area. DataedX also continues to have workshops and educational experiences. Join us at our signature event, the <u>Black Women in Data Summit</u>, to be held in Atlanta on October 1-2 2022. For more insights from the survey, check out our <u>interactive dashboard</u>*.

*To view the report, click the three ellipses on the top right side, and select "Present"



References

- [1] Harnham. (2021). USA Diversity in Data Report 2021-22 [White paper]. Retrieved March 18, 2022, from https://www.harnham.com/harnham-data-analytics-diversity-report
- [2] La Monica, P.R. (2020, September 17). *Snowflake shares more than double. It's the biggest software IPO ever.* CNN Business. <u>https://www.cnn.com/2020/09/16/investing/snowflake-ipo/index.html</u>
- [3] Kaggle. (2021, October 14). *State of Data Science and Machine Learning 2021 (slide 23)*. Retrieved March 18, 2022, from https://www.kaggle.com/kaggle-survey-2021

The Survey Team

- **Brandeis Marshall, PhD:** Brandeis is leading DataedX Group since 2018. Trained as a computer scientist and an educator-scholar by profession, she teaches, speaks and writes about the impact of data practices on technology and society. She focuses on educating and producing scholarship that amplifies Black women thriving in data and tech careers.
- **Sia Seko:** Sia is a data generalist with a specialty in data analytics engineering who consults on healthcare and ed tech industry projects. She performed the data cleaning, analysis, and storytelling tasks for this survey report.
- **Susan Geier, PhD:** Susan is independent research and program evaluation consultant specializing in STEM related educational outreach and career development projects. She led this survey's design and development.