



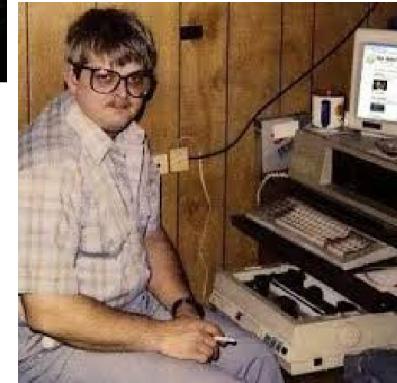
The People Behind The Data

Exploring the world of Data Careers

The Stereotypes around Data Careers

Many people have, in their mind, an idea of what a typical Information Technology or Data worker looks like. Young, white men, hunched over computers, their faces lit from the bright screen.

Today, we are going to look at the results of a few different surveys of coders, and see if that stereotype holds up. In particular, we are going to look at Data Careers, which can include Data Engineers, Data Analysts, and Data Scientists (who typically also work with Machine Learning/ Artificial Intelligence).



Where is our information coming from?

The data we will be using to assess these careers is coming from 3 different surveys.



Stack Overflow is a popular discussion and help site for information technology, and the survey gives us an insight to people who are actively and collaborating in the field.

freeCodeCamp(🔥)

FreeCodeCamp surveyed new coders. This provides us information about those people just entering the field, where they are coming from, and what they are hoping to do.



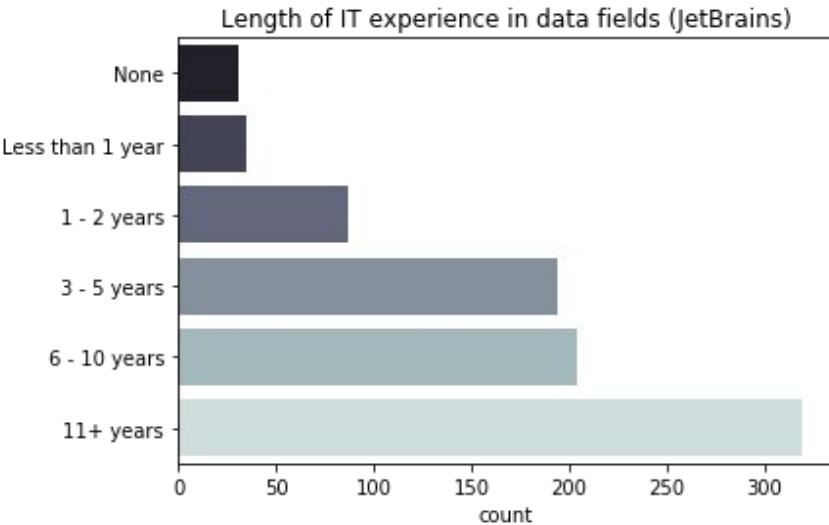
JetBrains is a company that produces software development tools which are used throughout the field. Their users will tend to be people actively working in the field and using these tools.

The Path to Data

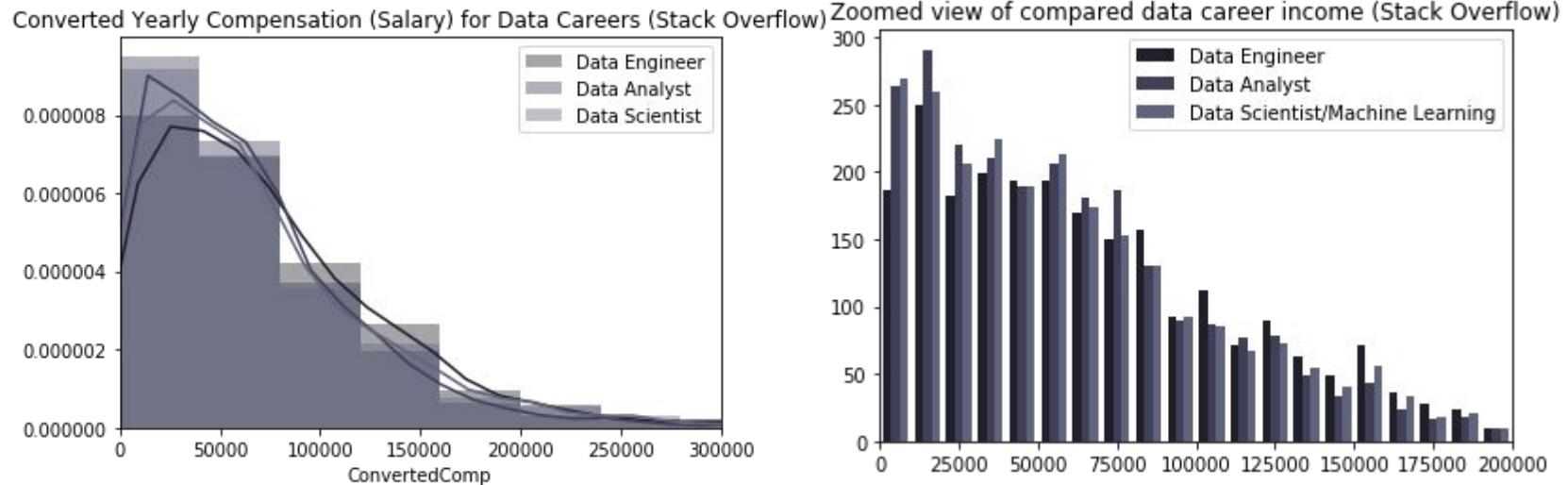
Previous Experience

First, we will look at the career and the paths people take to get to these jobs. What sort of skills, education, and experience do the people working in this career typically have? What sort of self-development do they engage in?

Right off the bat, we can see that there is an interesting range of previous IT experience for those working in data fields already. Some people have little to no experience working in another IT field.



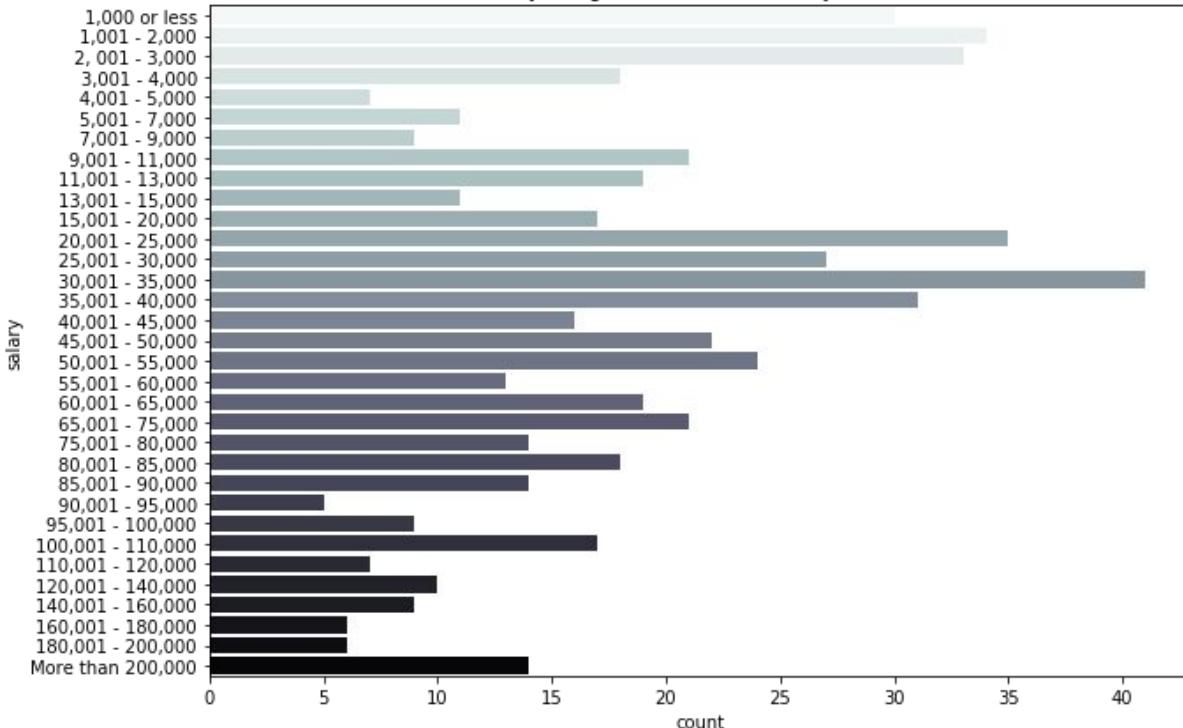
Data Career Salaries



We can see that, according to Stack Overflow, there is little significant difference in income between different data careers. The most common salary tends to be towards \$20-70k.

Data Career Salaries

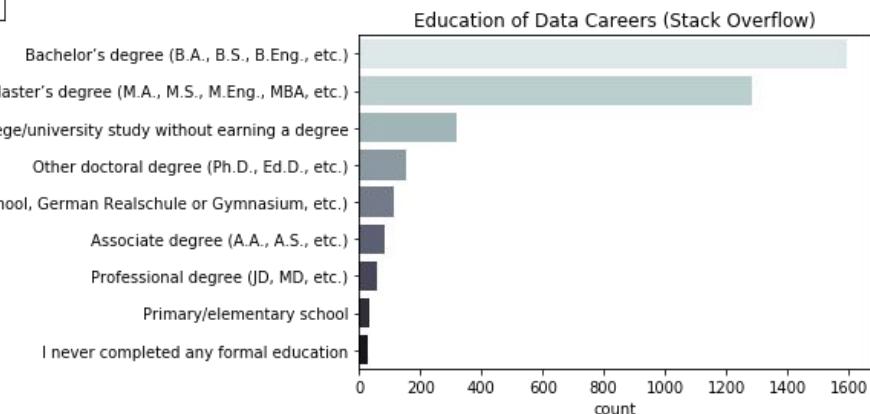
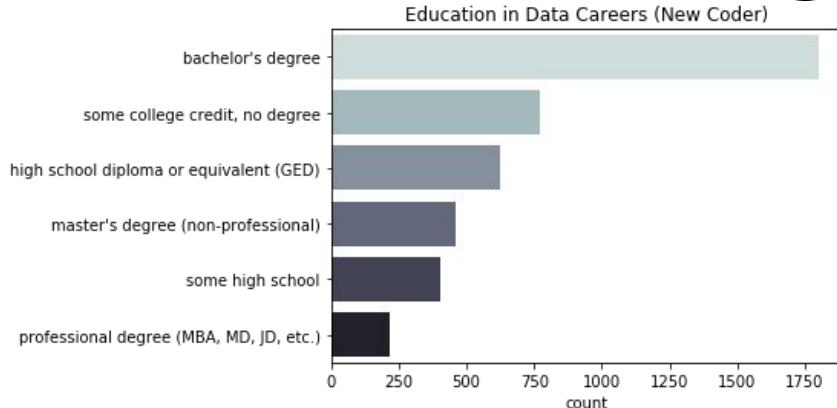
Salary Ranges for Data Careers (JetBrains)



JetBrains shows us a peak salary range between \$20-50k, and does not differentiate between careers.

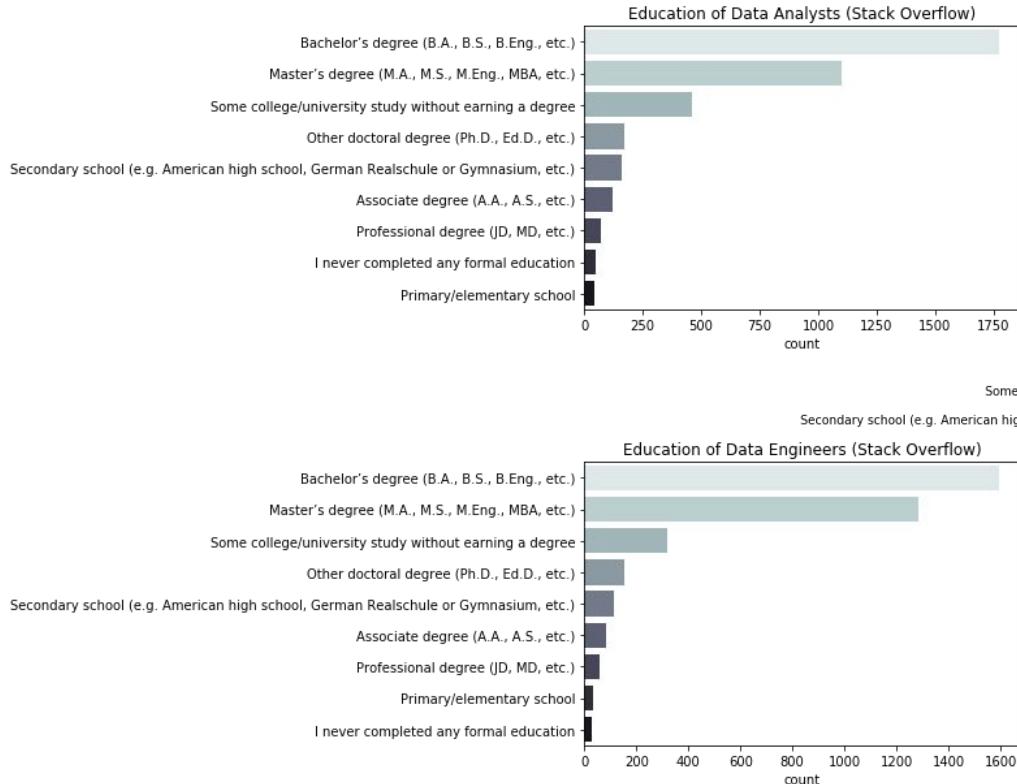
A quick ANOVA test also tells us there is no significant difference in salaries between data careers, with an f-value of .65, which is quite close to 0 (no difference).

Educational Background

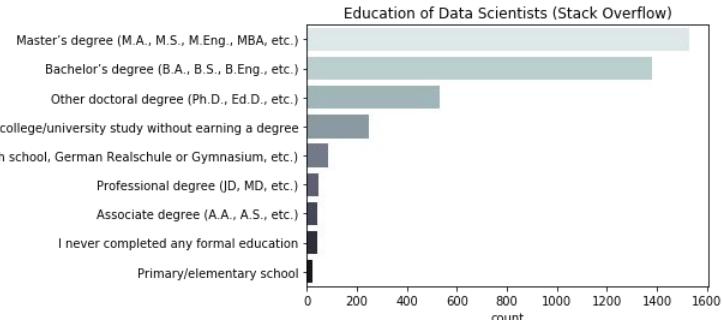


We can see that between these two sources, the most common level of schooling is a Bachelor's degree. There are some minor differences as we look through the rest of the options, notably that the New Coder survey records a much smaller percentage of Master's degrees and other advanced degrees.

Educational Background



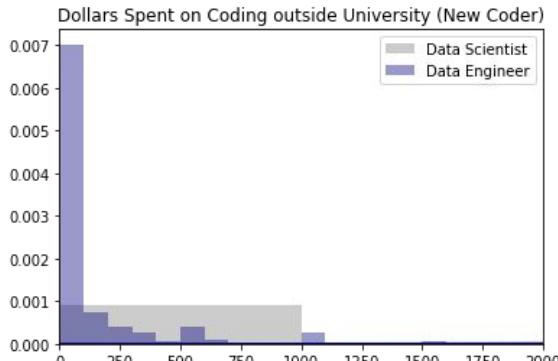
Stack Overflow allows us to look a little closer into the breakdown of education within the data careers. With this, we can see that there is a slight difference in that data scientists tend to skew toward higher education, in that most common degrees are Masters degrees, and other doctoral degrees are also highly present.



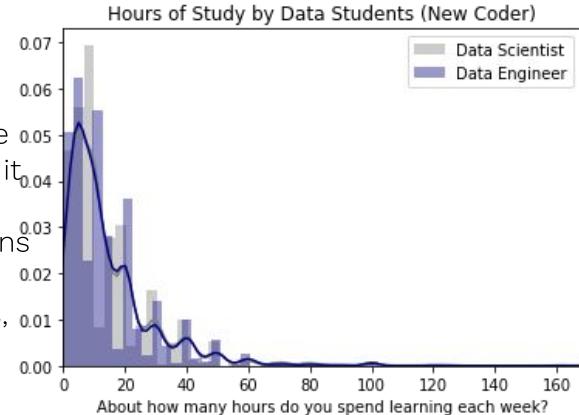
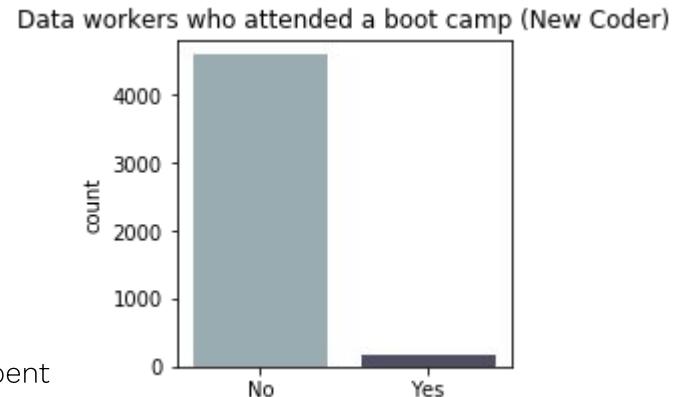
Coding Bootcamps

Interestingly, the New Coder survey tells us that most of the respondents did not utilize a full-time bootcamp to break into the industry. This could imply a few conclusions. New coders may be coming directly from university, or perhaps are taking time to learn on their own without the structure of a bootcamp.

When we look at money spent outside of university, we see that very few engineers spent more than \$100 dollars outside of tuition. Data Scientists were willing to spend more, but no more than \$1000. Looking at the hours per week spent studying, it looks more like studying is a part-time occupation. This means that it is probable that new coders are teaching themselves, outside of a formal school or bootcamp setting.

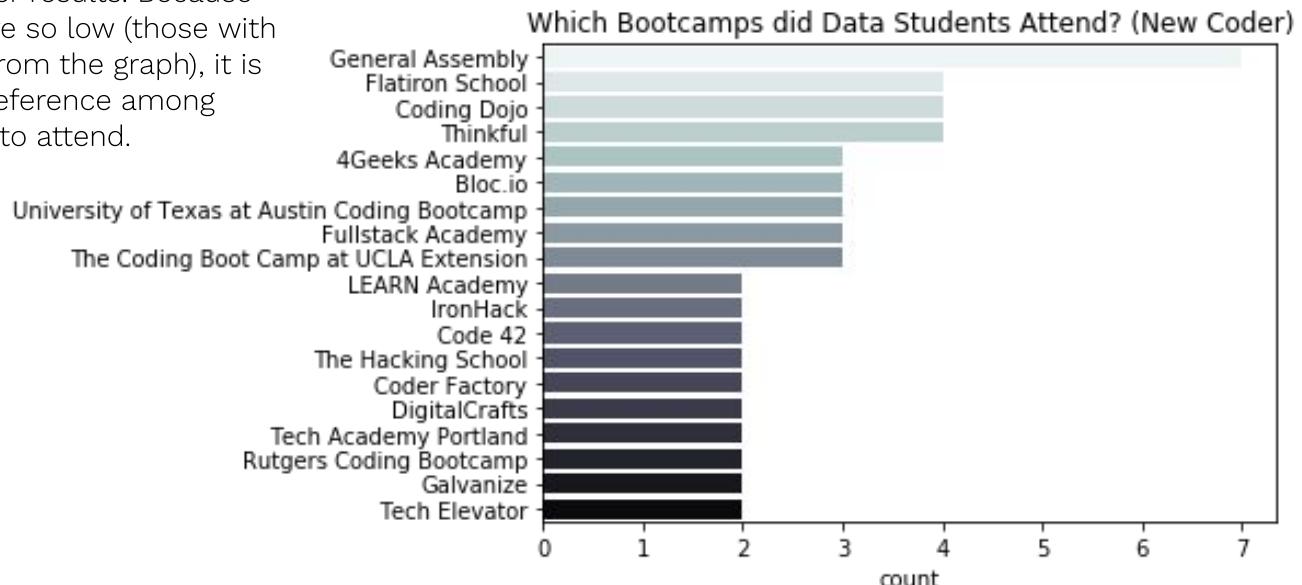


Aside from university tuition, about how much money have you spent on learning to code so far (in US dollars)?

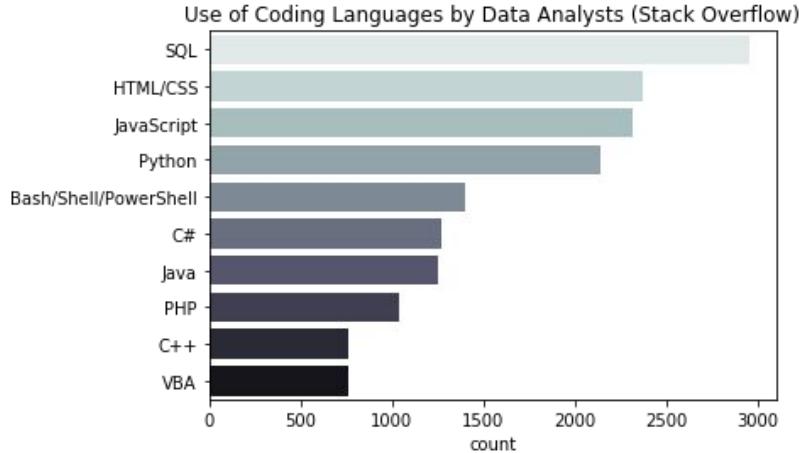


Coding Bootcamps

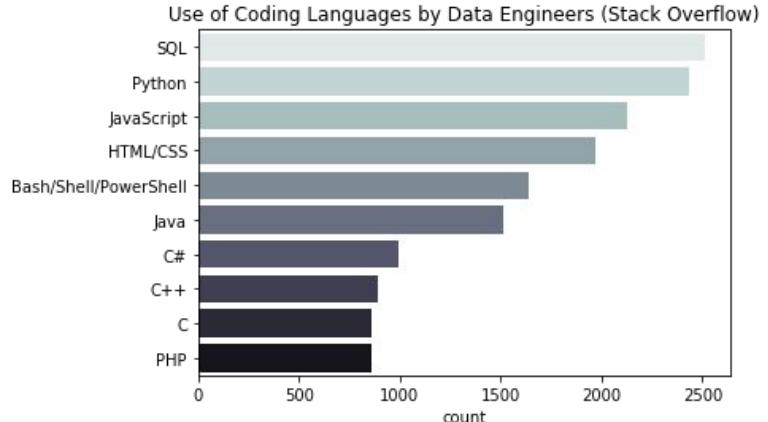
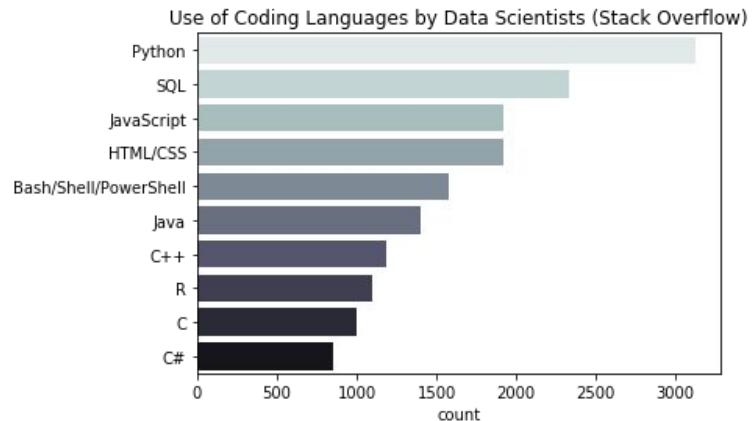
We don't have a very significant sample size of popular bootcamps, but it is interesting to see the data we do have. General Assembly, Thinkful, Flatiron School, and Coding Dojo are the best represented in the New Coder results. Because the counts for each camp are so low (those with only 1 mention are omitted from the graph), it is fair to say there is no real preference among new coders for which camp to attend.



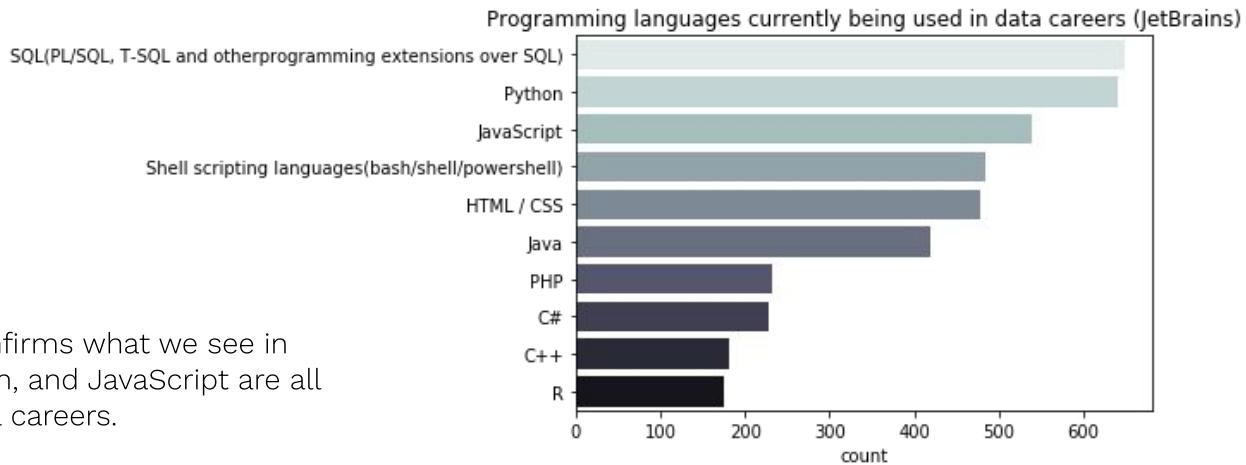
On-the-job Skills



Stack Overflow allows us to look at the breakdown of language use between different careers. Data Scientists and engineers tend to use similar languages, while analysts show a difference in that they use HTML/CSS more often.



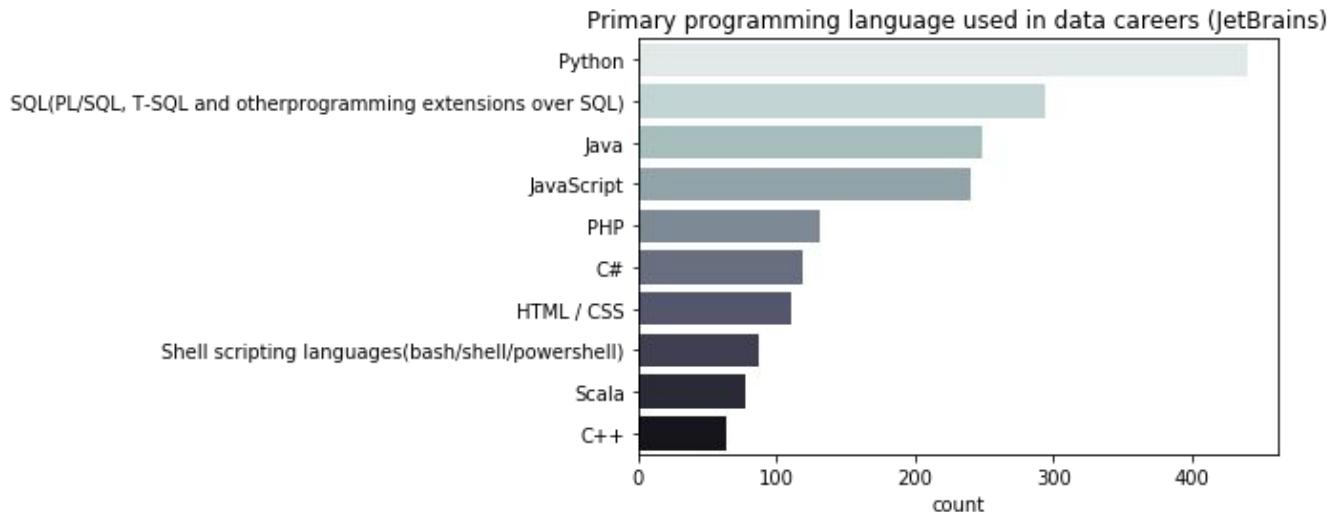
On-the-job Skills



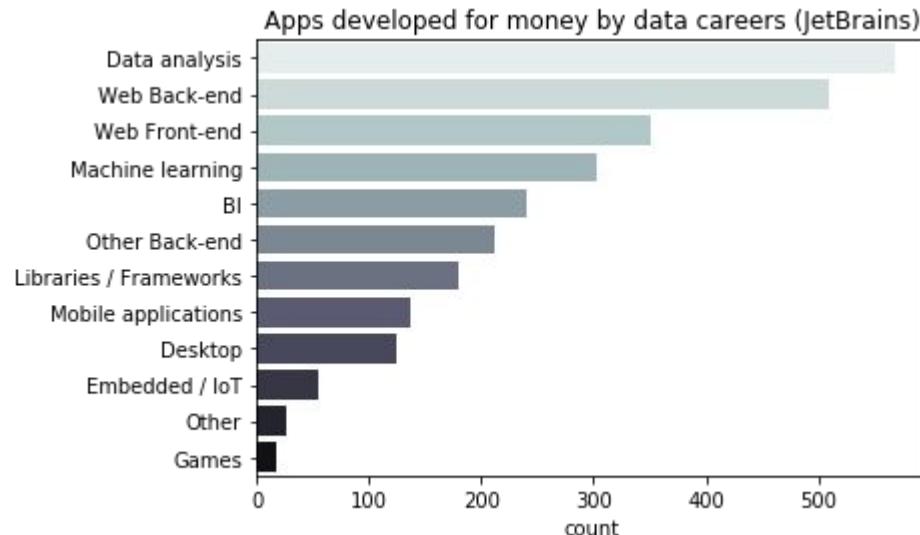
The JetBrains data also confirms what we see in Stack Overflow. SQL, Python, and JavaScript are all very popular in various data careers.

On-the-job Skills

JetBrains gives us a look not just at languages used, but the primary languages used in these careers. Both of these graphs confirm the fact that new data workers should ensure a familiarity with Python, SQL, JavaScript, HTML, and a shell scripting language.



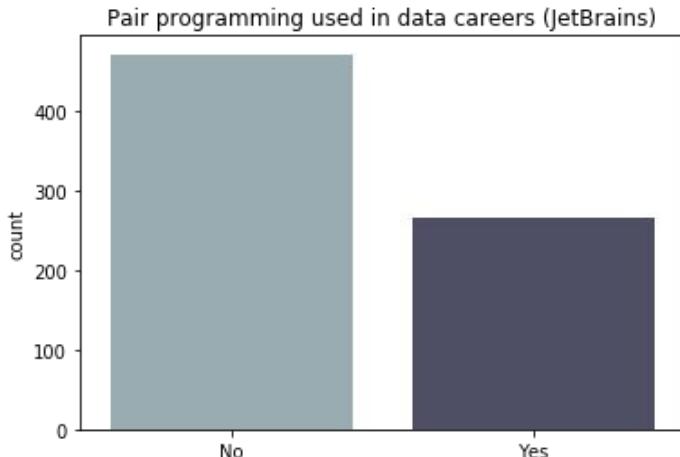
On-the-job Skills



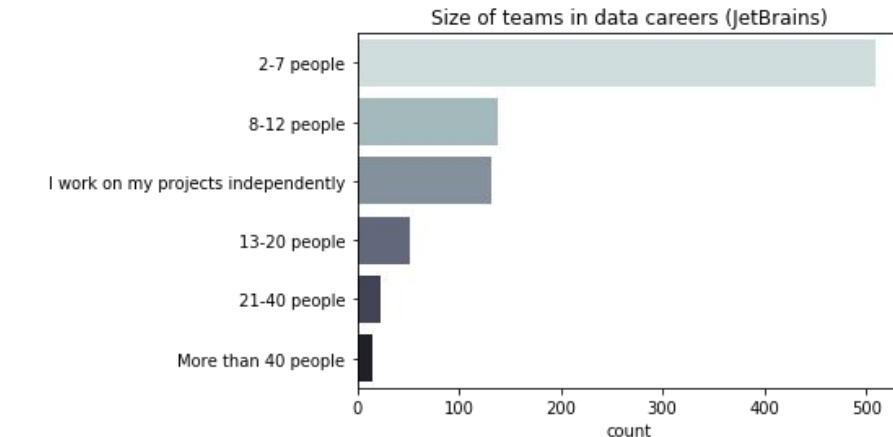
The JetBrains survey allows us to see where a lot of focus is placed in programming by data careers. Unsurprisingly, data analysis is the most popular, but web back-end is also very high, along with web front-end and machine learning. This gives us some insight into what skills a new employer might be expecting an employee to have.

On-the-job Skills

JetBrains tells us that many people working in data careers tend to work in small teams of less than 10 people. Teams of more than 20 people are quite rare. Working independently and working in small groups means that those looking to succeed in the field should be flexible working in different team situations.

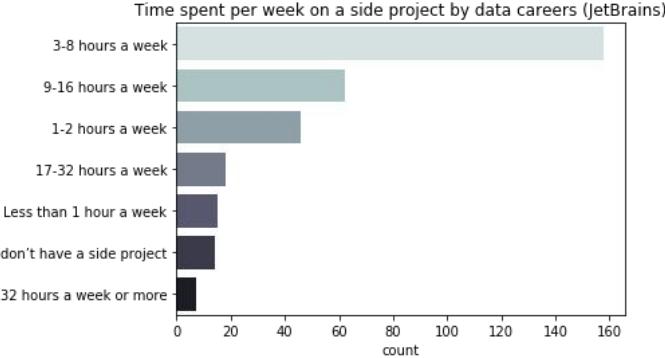


count

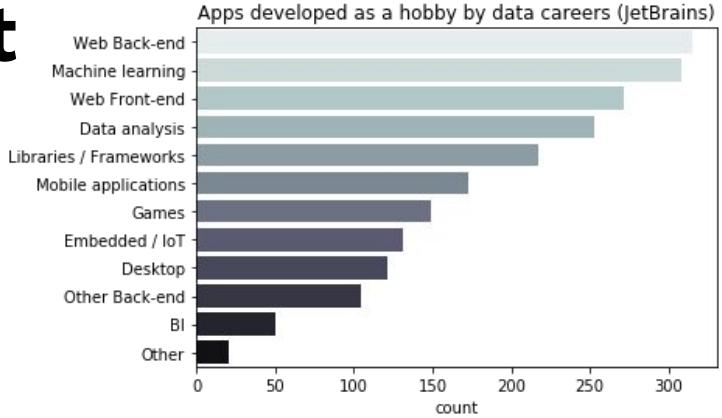


Additionally, we can see that a significant amount of employees engage in pair programming, where you code with a partner. This also implies skills beyond just a programming language or app development, and more towards working with a partner, with effective communication and cooperation.

Personal Development



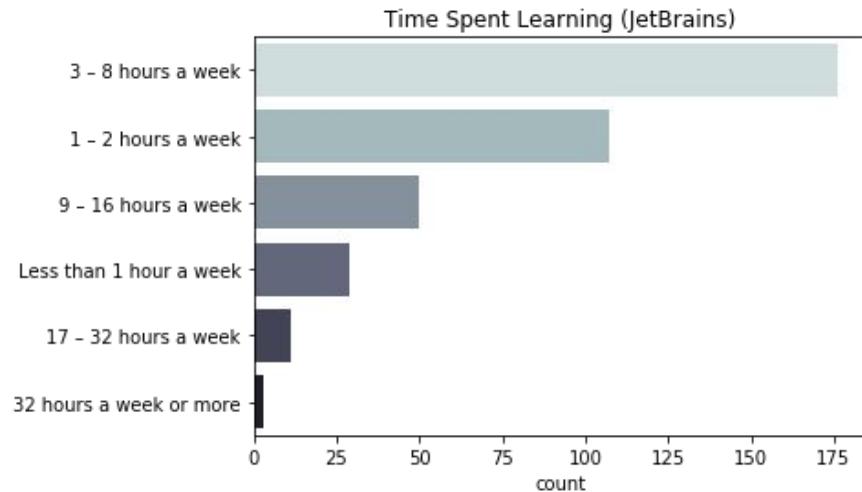
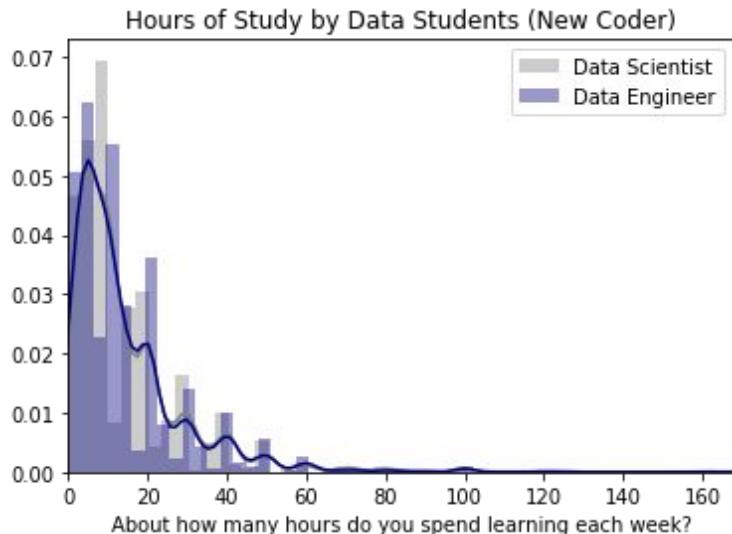
JetBrains tells us that it's incredibly common for those in the data field to spend 3-8 hours per week working on a hobby side project.



Unsurprisingly, the types of apps worked on as a hobby mirrors what is used in employment. The count of Web back-end apps is higher in hobbies, though, which may mean that developers are trying to hone their skills independently to support their paid work experience.

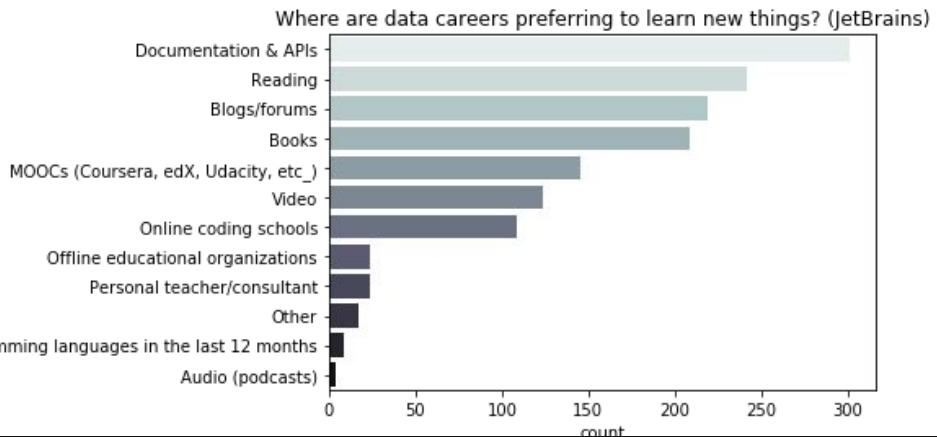
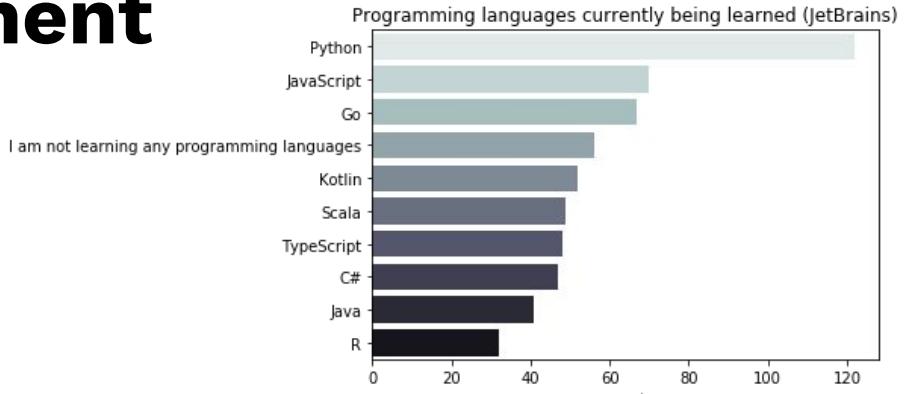
Personal Development

We can also see that time spent learning independently tends to be around 3-8 hours a week as well, which implies that most personal learning is done through side projects. Independent projects are a huge part of data careers!



Personal Development

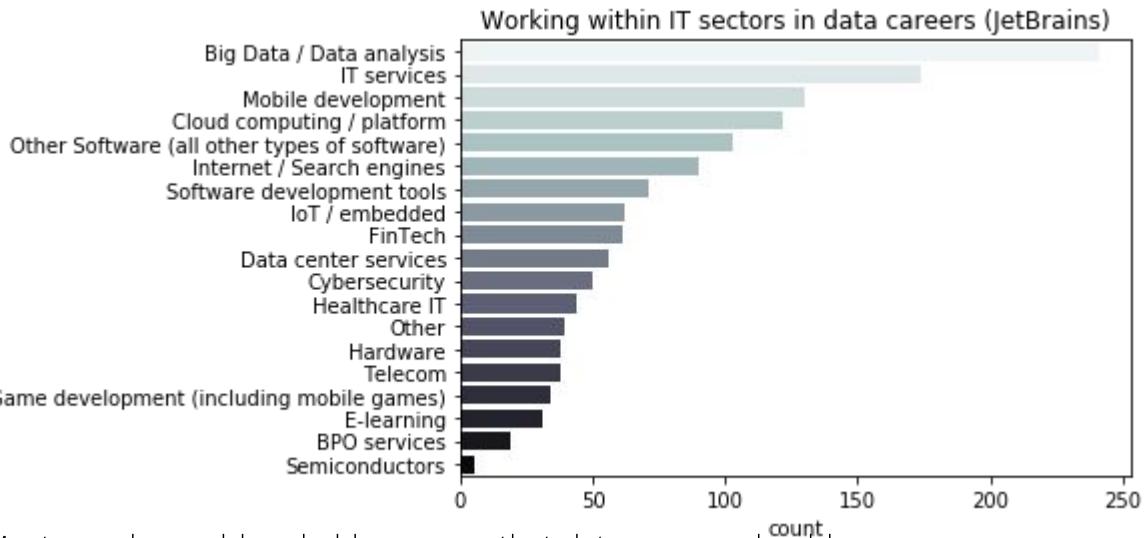
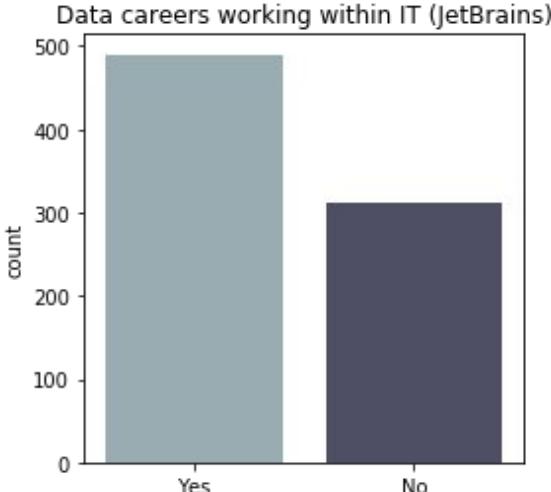
What are they learning and from where? JetBrains again gives us some insight to personal development. Once again, Python and Javascript are considered important languages to be learning. Surprisingly, Go shows up in this response, which previously has not shown up in the top rankings. Many people prefer to learn as they explore documentation on these languages, but reading, such as on blogs and in books, is also very popular.



I did not learn any new tools / technologies / programming languages in the last 12 months

**Where to find a
data career?**

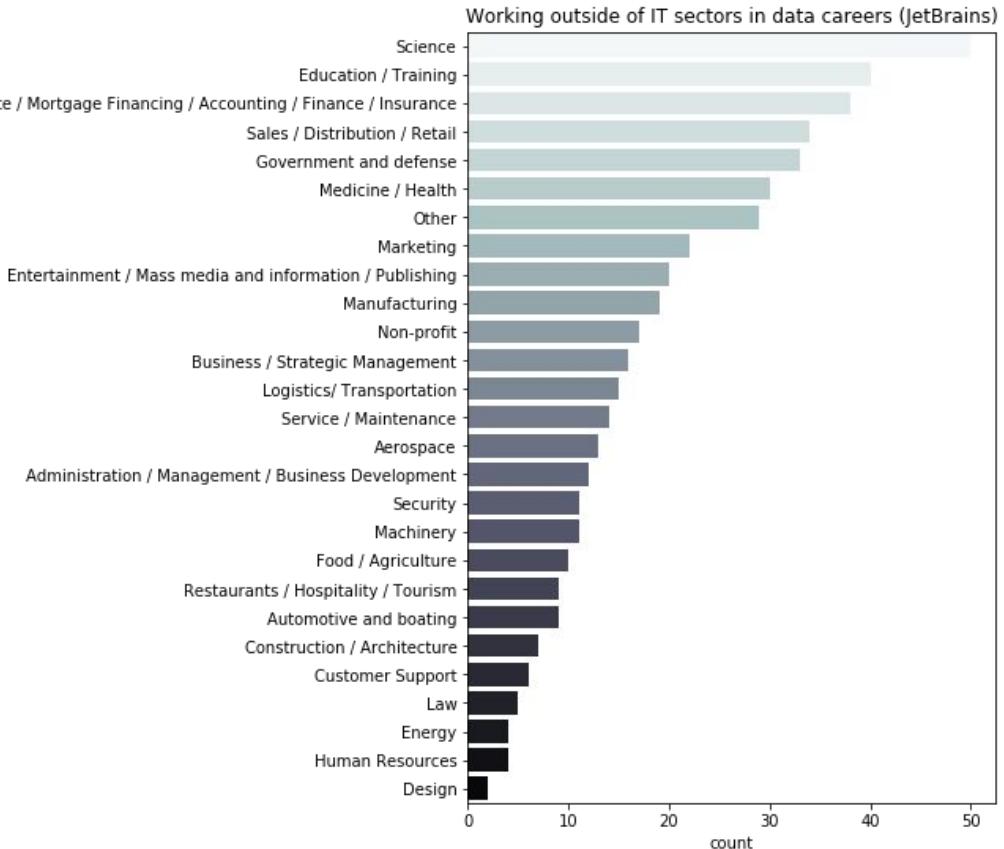
Industries



Most people would probably assume that data careers should focus on the IT field to find a new position. Data from JetBrains shows us that this isn't necessarily so. While a majority of respondents work in the typical IT fields, there is a good amount of those who don't. The graph above shows the range of different sectors where data workers find employment. Unsurprisingly, the most popular sector is in Big Data/ Data analysis. Many different sectors employ data workers, though, implying that this career has broad applications.

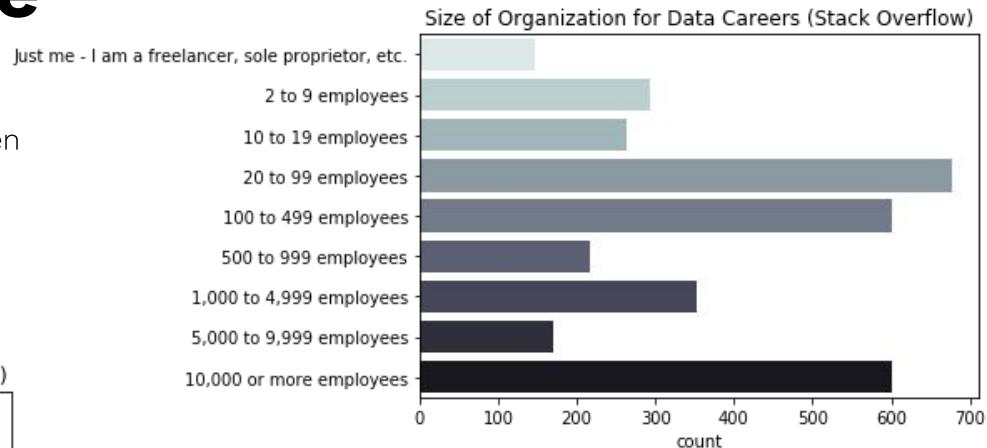
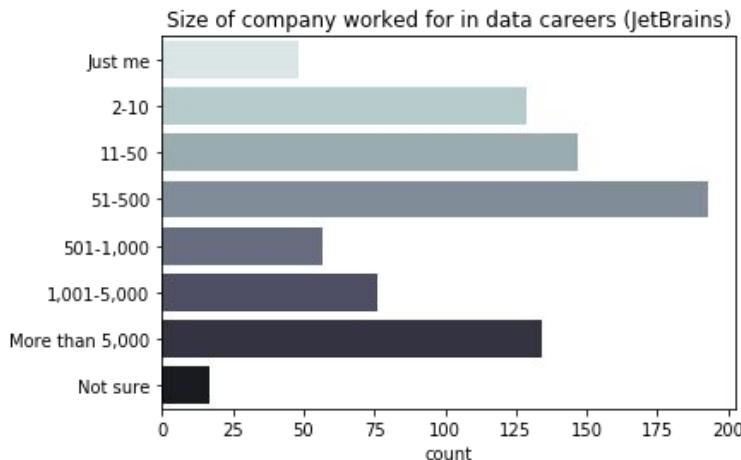
Industries

Very exciting is this graph from JetBrains showing the adaptability of data careers outside of the IT industry. Manipulating and interpreting data is becoming a huge aspect of many different lines of business. Sharpening your skills in this career can lead to a wide variety of potential careers. Based on this survey, we can see that Science, Education, Government, Banking, and Sales tend to utilize data careers, and may be a good place to look for people entering the field.

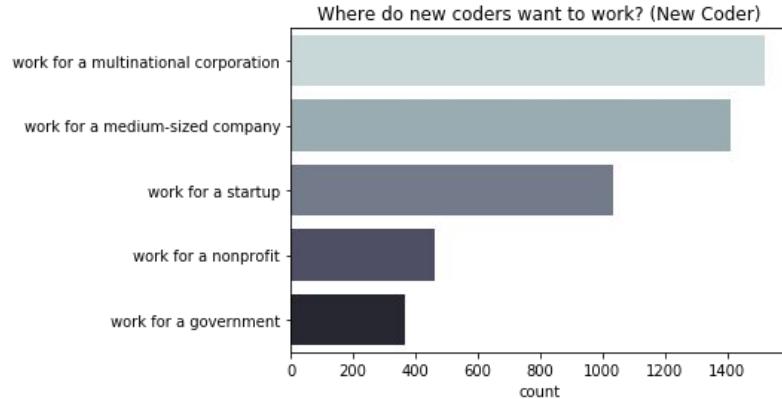


Organization Size

Looking at employment organization size, it appears most respondents land somewhere between 50-500 employees. There are those working anywhere between self-employed, working alone, all the way up to more than 5,000 or 10,000 employees. Organizations of all sizes have a use for the work of data careers.

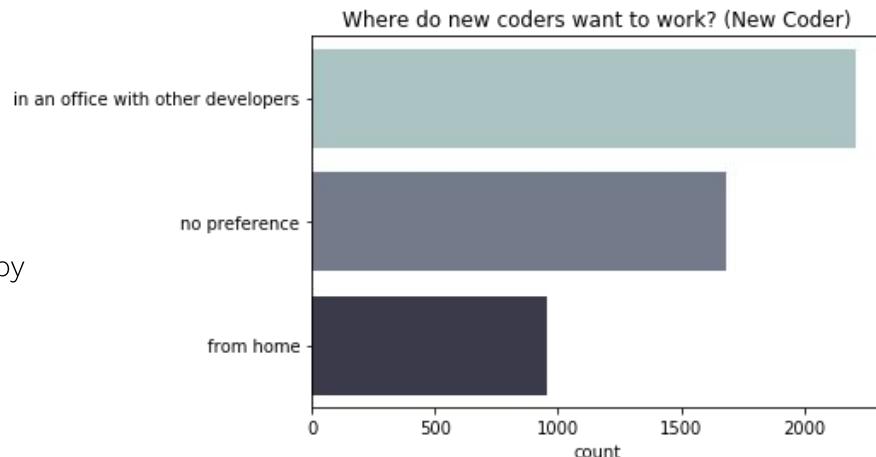


Where do new coders want to work?



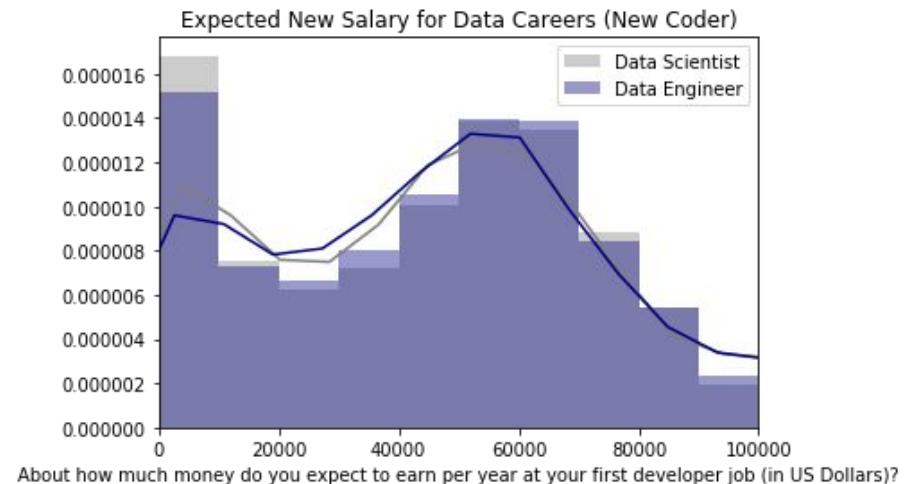
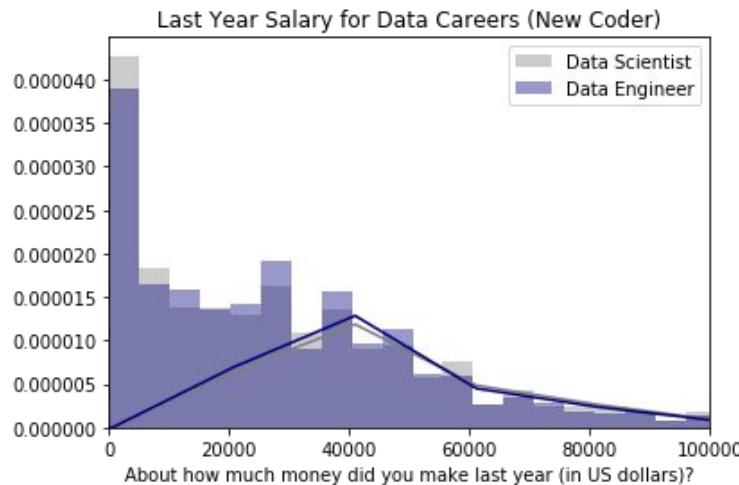
Perhaps most relevant in the times we live in today, do new coders have a preference for where exactly they would prefer to work? While most would prefer an office environment, many have no preference or would be happy to work at home. Seeing as working from home is probably part of our 'new normal', this is encouraging for the current climate.

The survey of new coders showed us that new coders are interested in working for a variety of different types of companies, from startups to multinational corporations. Less popular are the options of government and nonprofits, but the difference is not highly significant.



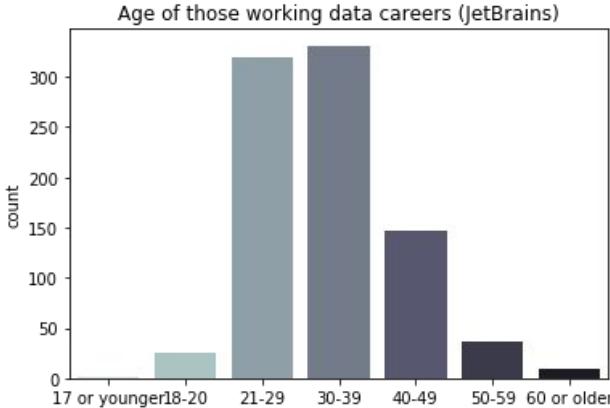
Salary Transition

As an aside, the New Coder survey allows us to view the change in salary from those newly coming into the IT/data field. The graph on the left shows us that most of the salaries tended to be in the \$10-50k range, ostensibly from before they began pursuing their new careers. If we assume they are estimating their future salary from offers and reliable research, we can see that this job transition should ensure a generally higher income, usually between \$40-80k.

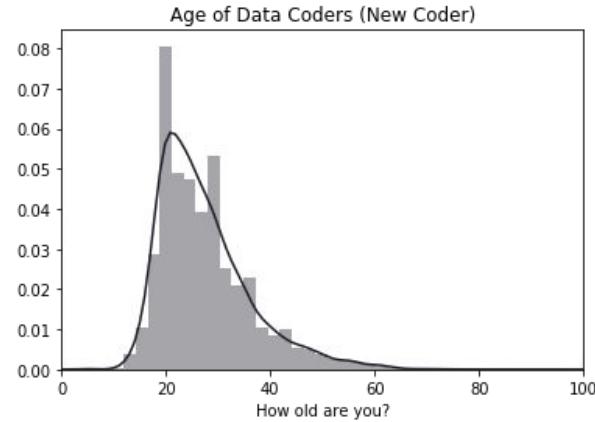


Data Demographics

Age Ranges

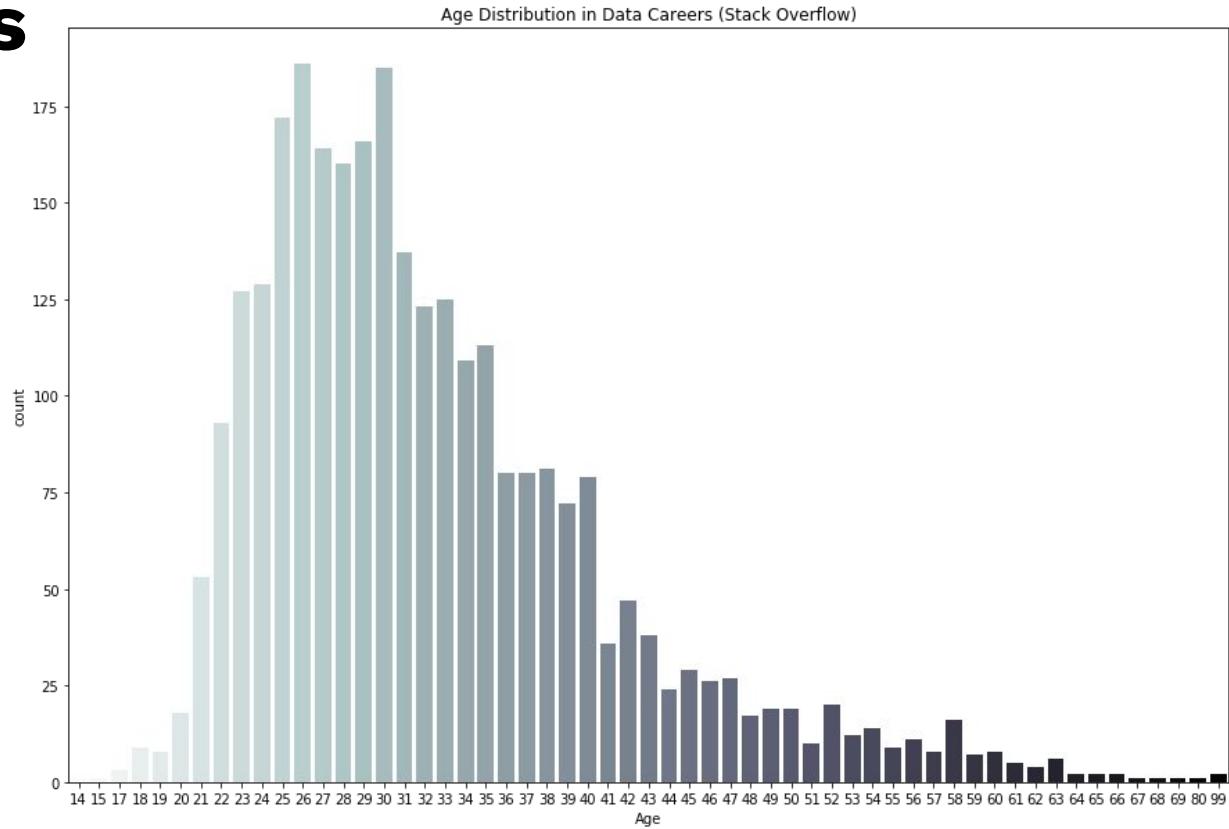


Across all the surveys, there is a definite trend that most coders land within the age range of 20s-30s, but we do see quite a few coders even all the way up into their 60s.
There is no age limit on code!
(Stack Overflow data shown on next slide).



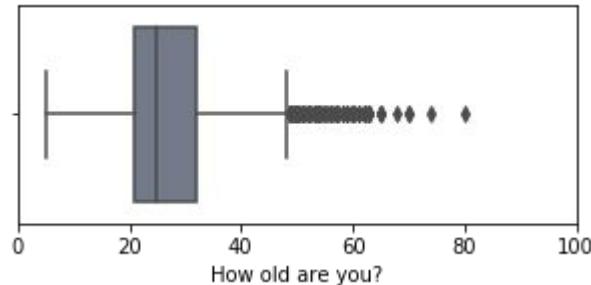
Age Ranges

A similar distribution to the previous slides can be seen in the Stack Overflow data.

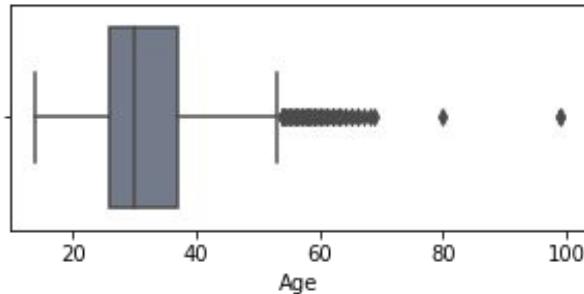


Age Ranges

Age of Data Coders (New Coder)



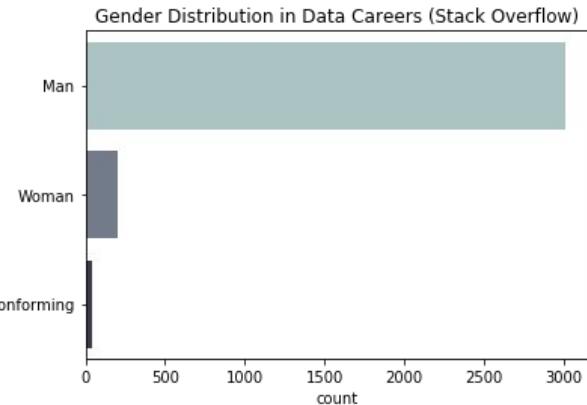
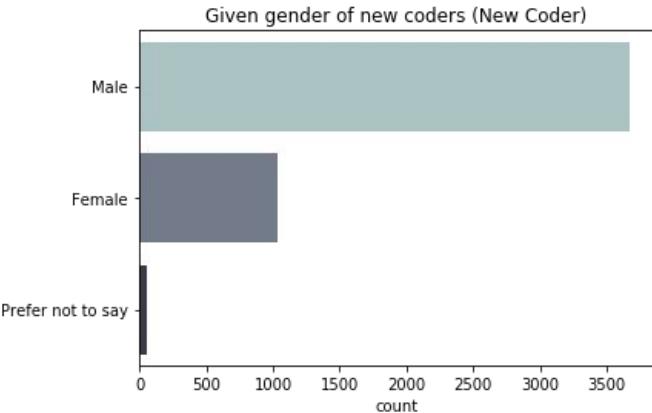
Age Distribution in Data Careers (Stack Overflow)



Stack Overflow and New Coder allow us to see age distributions a bit better. We can see that SO has a median of 30 and a mean age of 32. New Coder shows a median of 25 and mean age of 27. This makes sense as the more established coders would be a little older.

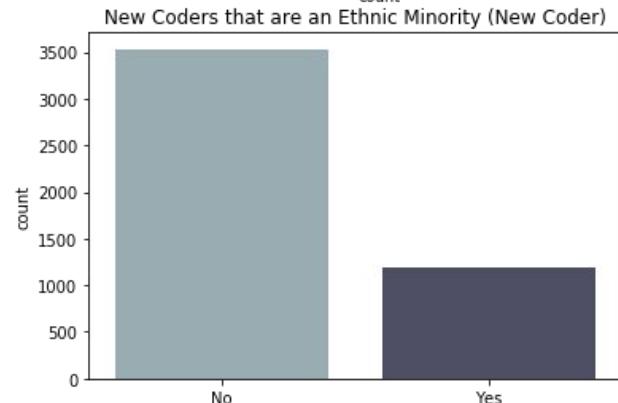
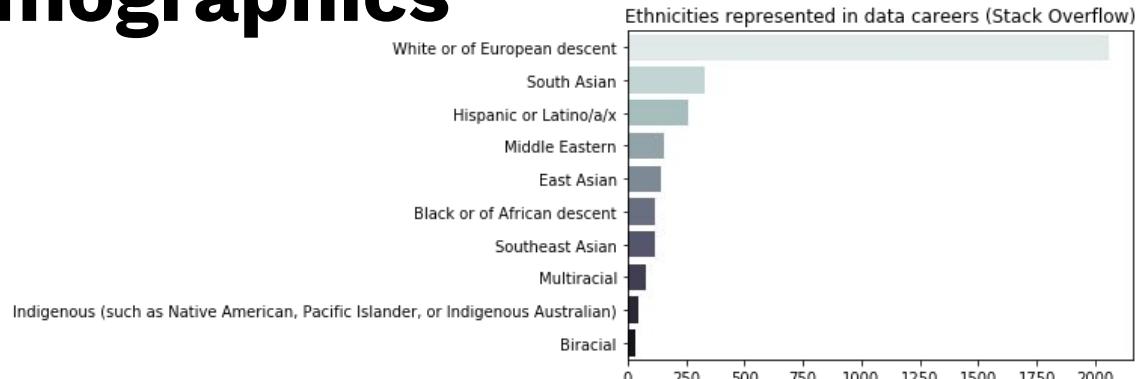
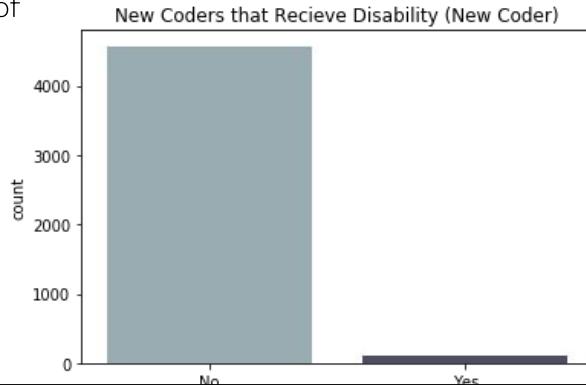
Gender Distribution

Our gender data does show a large disparity between men and those who identify differently. What is encouraging is seeing that the new coders seem to show a larger percentage of non-male coders entering the workforce. It is not unsurprising to see a small portion of people identify outside the gender binary, but having it as a posted option in the survey also implies progress.



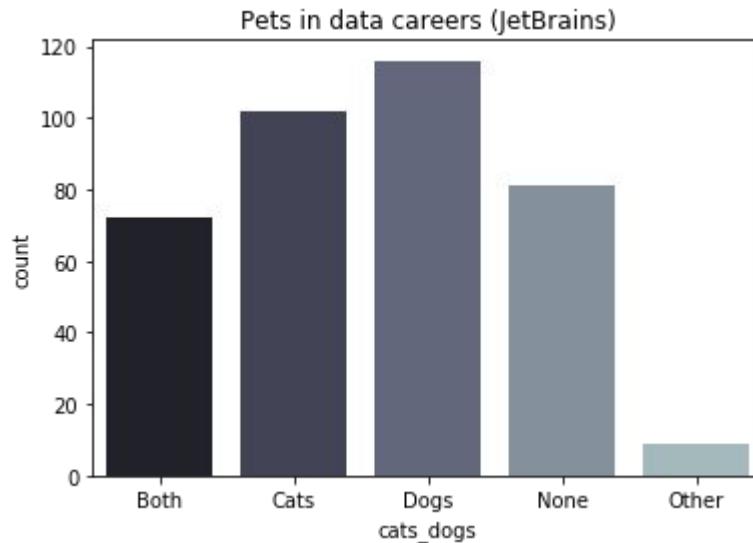
Assorted Demographics

We are also able to look at the distribution of ethnicities and even an insight into disability. While an overwhelming majority of those in the data field are white, there is a good spread of coders in various minority ethnicities. We can also see that while it is a small percentage, there are some coders who identify as disabled, showing that this career is open to lots of different people.



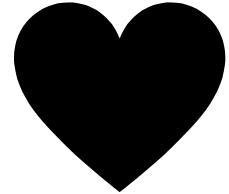
Final Demographics

Finally, the most important and telling demographic we can glean from these various surveys. JetBrains gives us an insight into what kinds of pets those in the data field take care of. Dogs are the winner, with cats a close second. Many cannot choose and have both. A small percentage of data careers have a different pet, possibly a bird or a snake, while an astonishing 80+ respondents said they had no pets. Further research into this aspect of data careers is desperately needed.



Final Conclusions

We started out with a preconceived notion of what someone in the data fields looks like, a young, white, man. While some of our data supports that, we can see that these careers are incredibly diverse, covering a wide range of companies, ethnicities, business sector, and ages. We can also see that new coders are entering the industry from many different backgrounds. The future of data careers is definitely something that is growing and changing!



Credits

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by [SlidesCarnival](#)