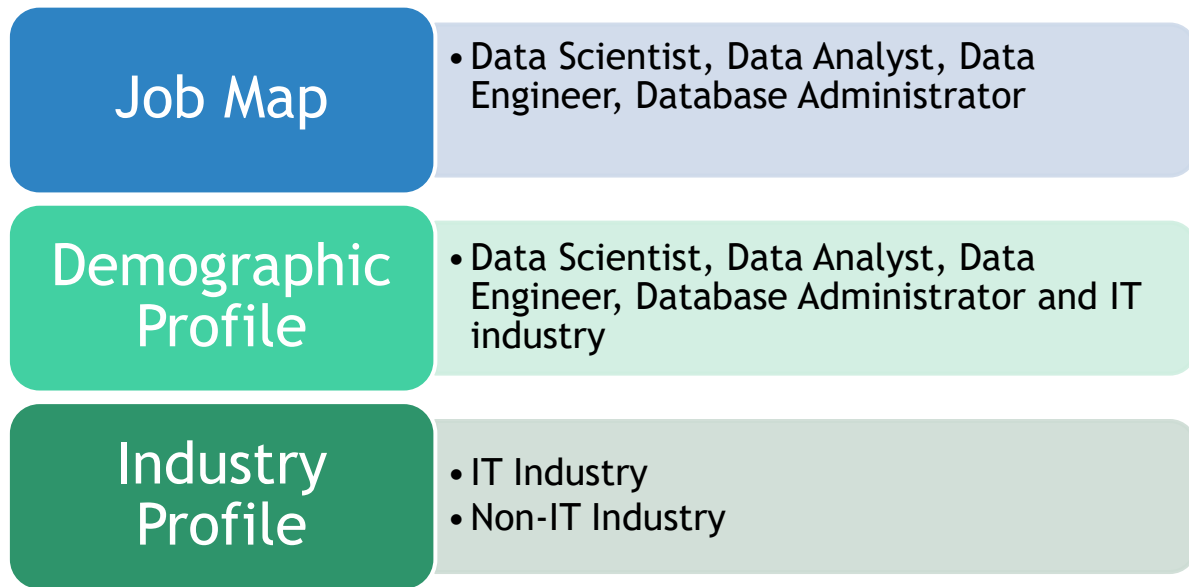


Exploring Public Survey Data - Employment in Data-Related Job Roles

Daniel Garcia Rodriguez
8/21/2020

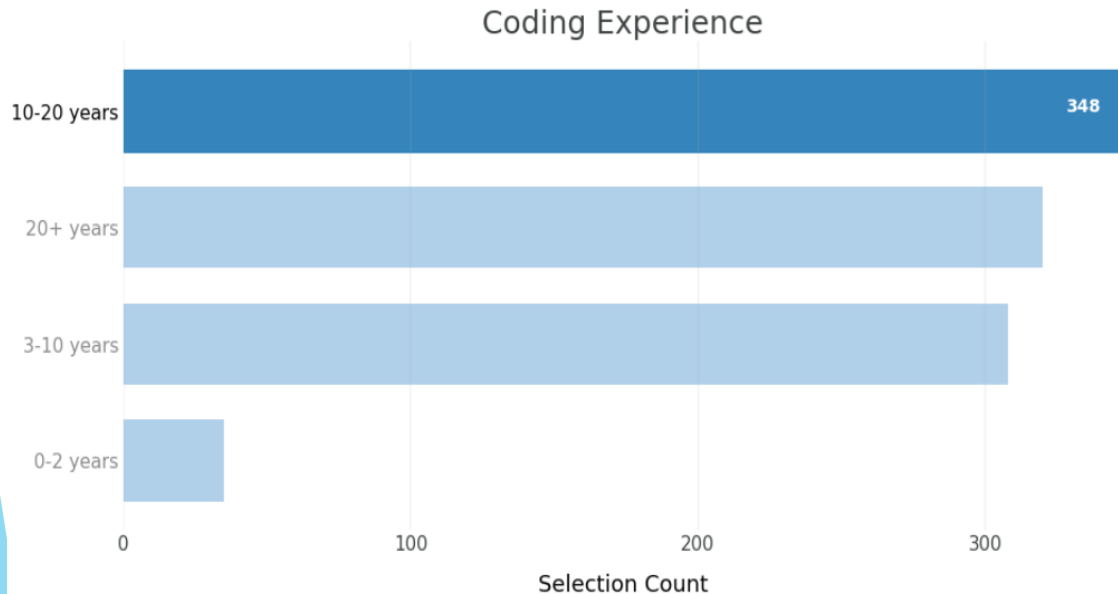
Overview



Data Sources:

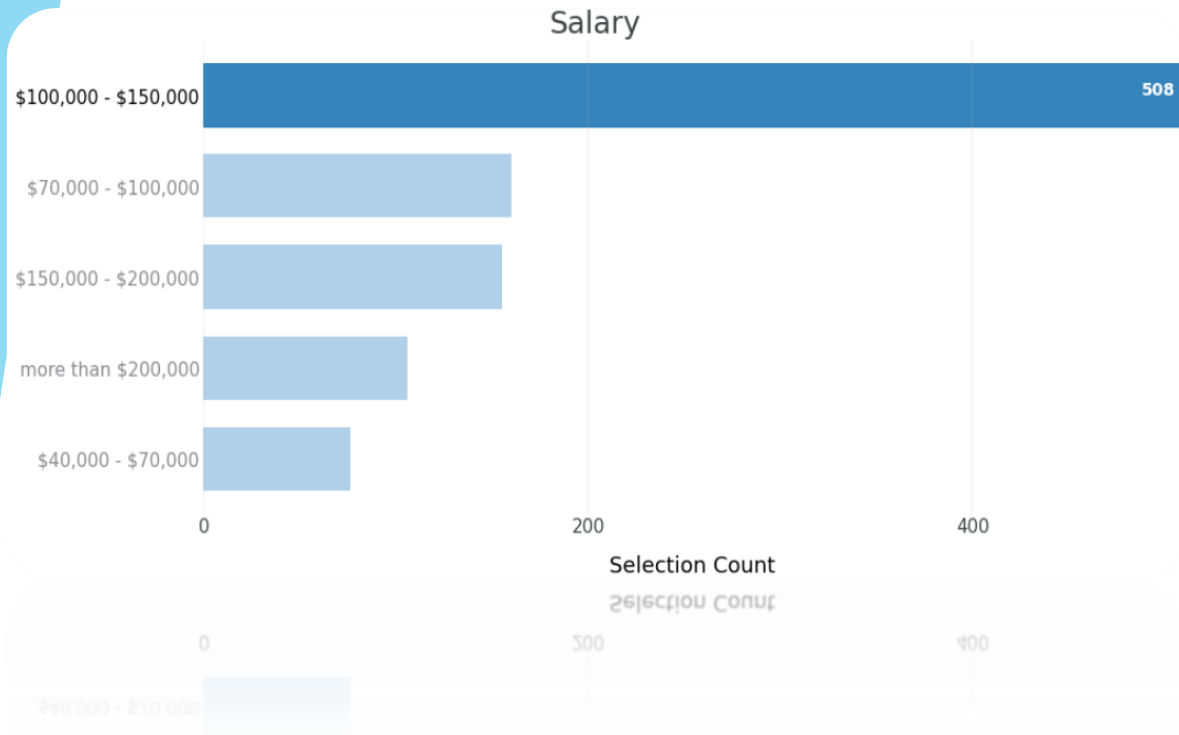
- (1) FreeCode Camp - New Coder Survey (2018)
- (2) Stack Overflow Developer Survey (2019)
- (3) JetBrains - The State of Developer Ecosystem (2019)
- (4) Kaggle - ML & DS Survey (2019)

Data Engineer



No clear most common coding experience. On the aggregate though, coding experience is more likely to be >2 years

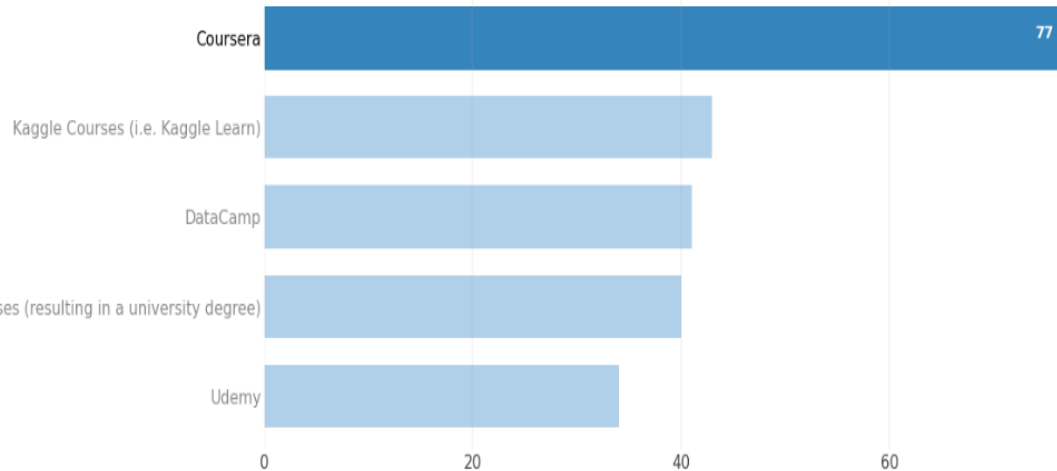
Data Engineer



Data Engineers in the U.S. are reportedly paid well with \$100k to \$500k being the most common salary range

Data Engineer

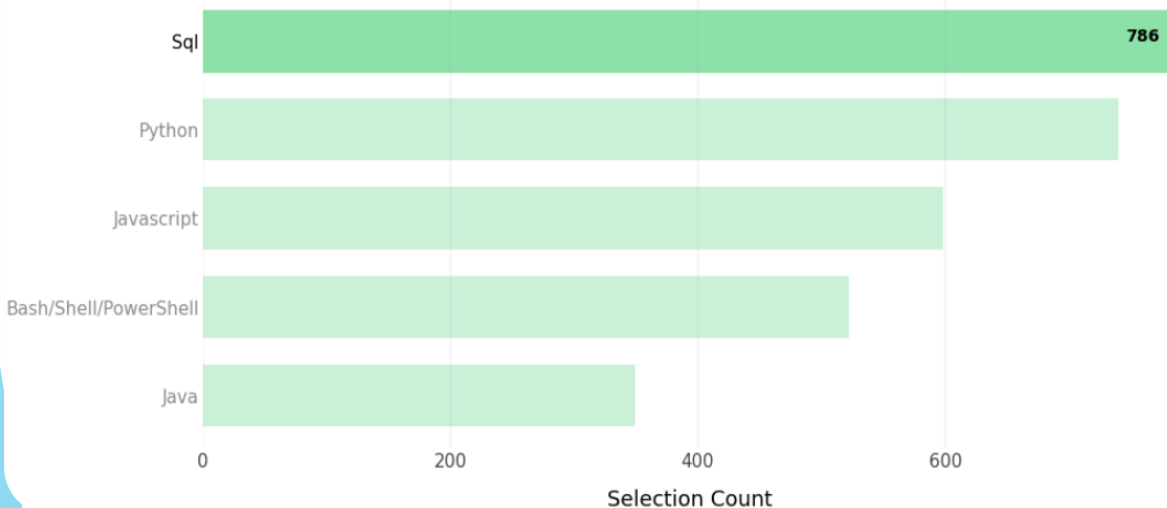
Top 5 Learning Platforms



Coursera is the most common learning platform reportedly used by U.S. Data Engineers

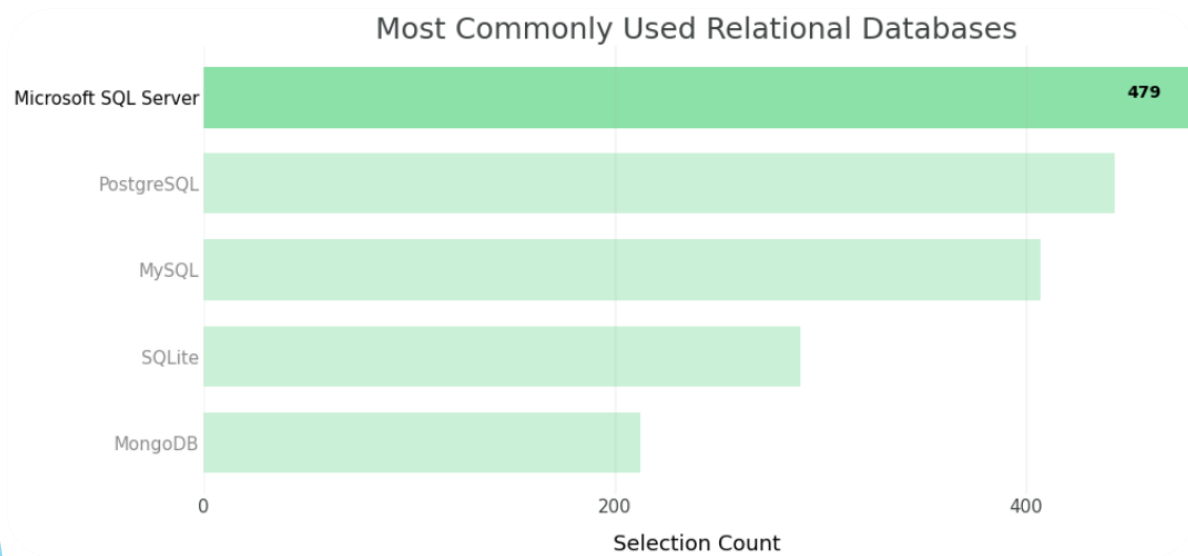
Data Engineer

Most Commonly Used Programming Languages



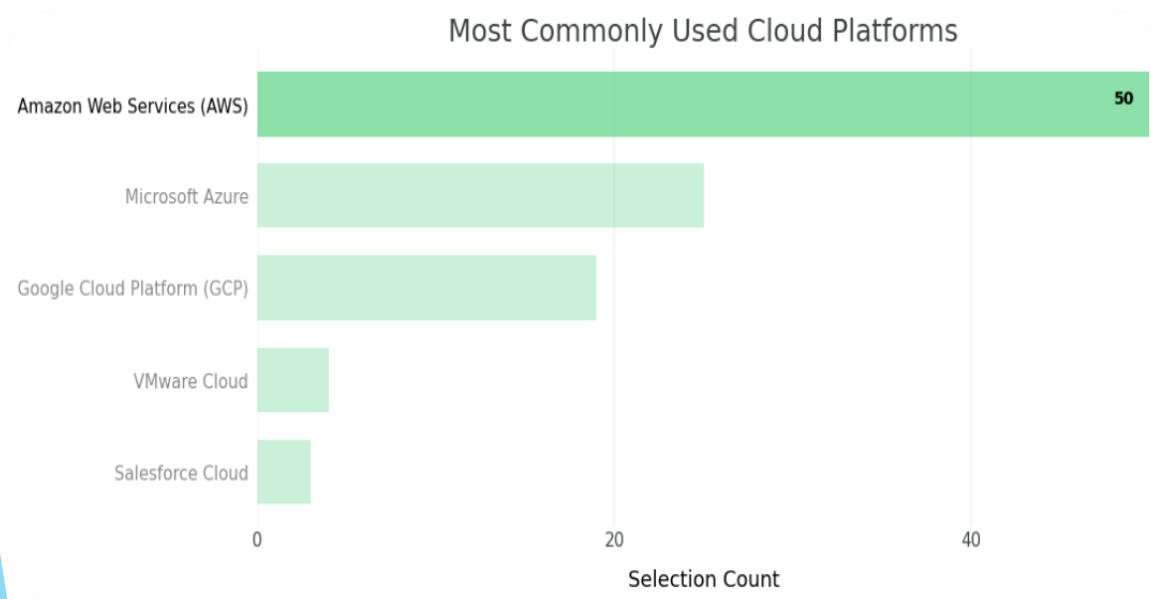
SQL and Python are the most commonly used programming languages for Data Engineers; JavaScript surprisingly falls in third place

Data Engineer



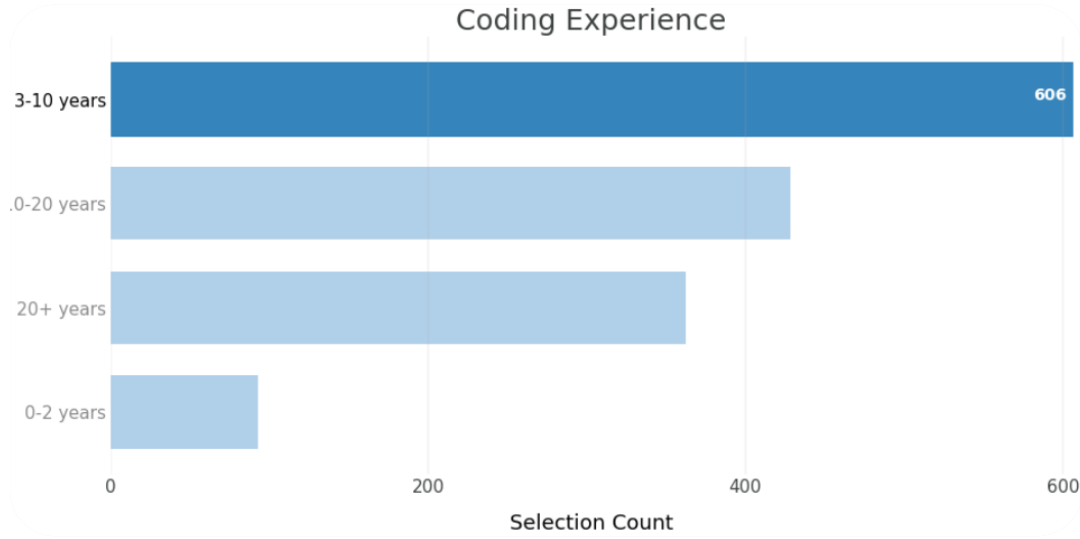
Microsoft SQL Server, PostgreSQL, and MySQL show large use for Data Engineers

Data Engineer



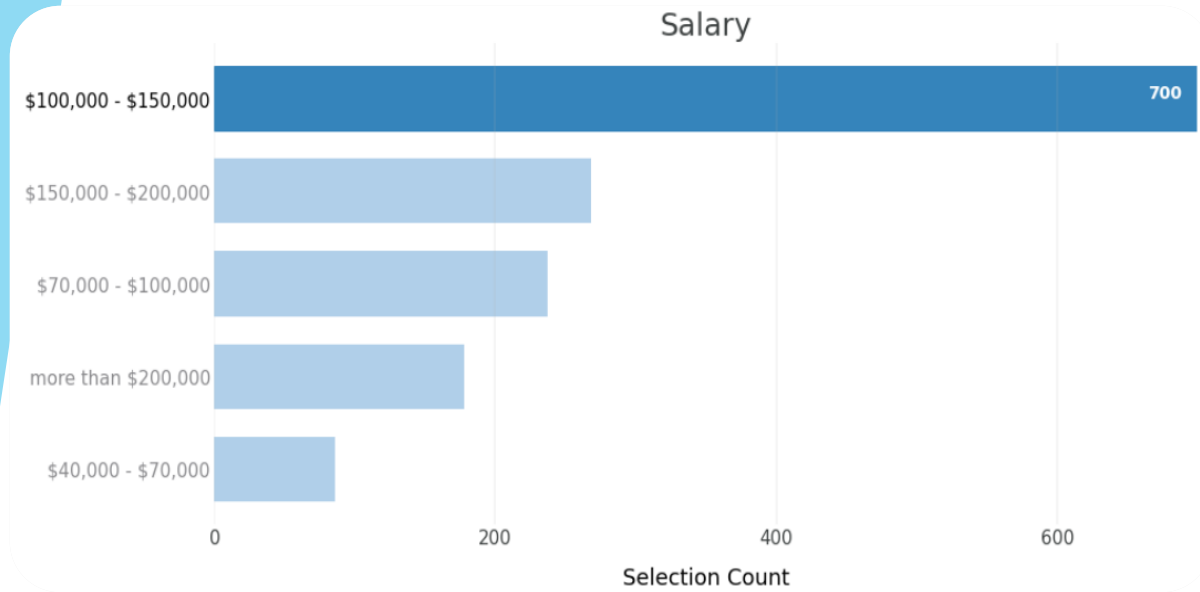
Amazon Web Service dominates in terms of usage for Data Engineers

Data Scientist



3 to 10 years of coding experience is common for U.S. data science professionals

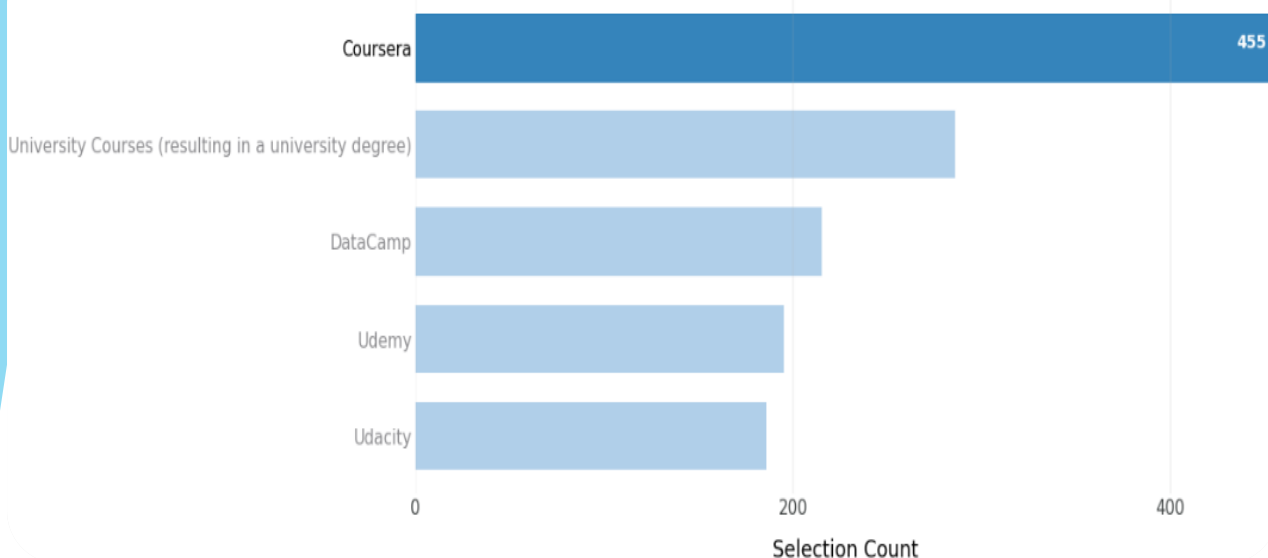
Data Scientist



Data Scientist in the U.S. are reportedly paid well with \$100k to \$500k being the most common salary range

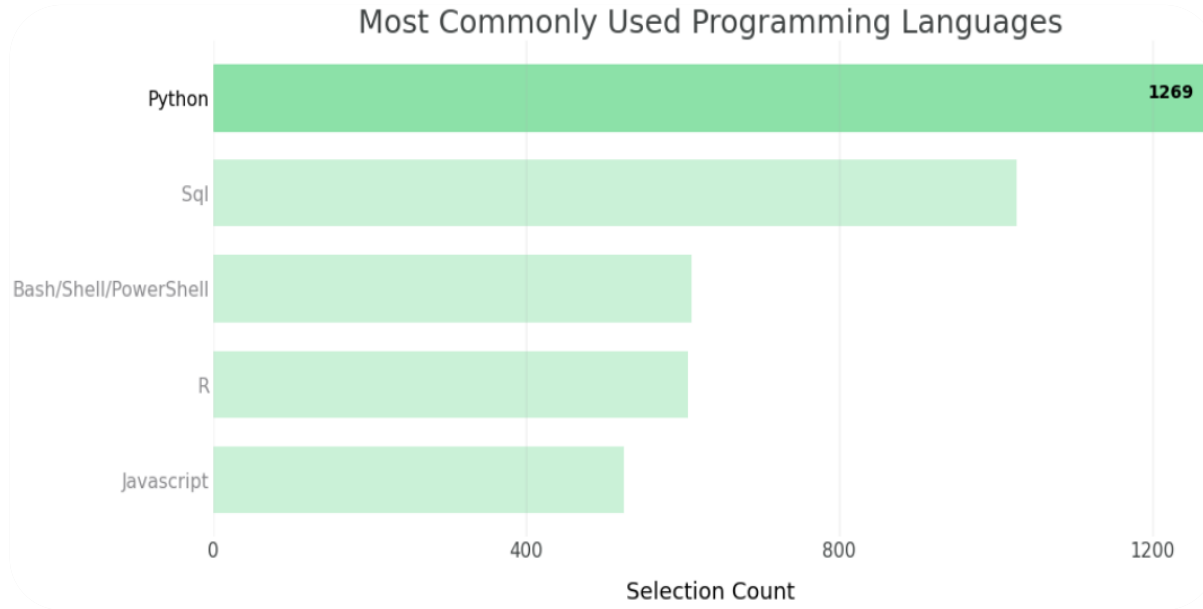
Data Scientist

Top 5 Learning Platforms



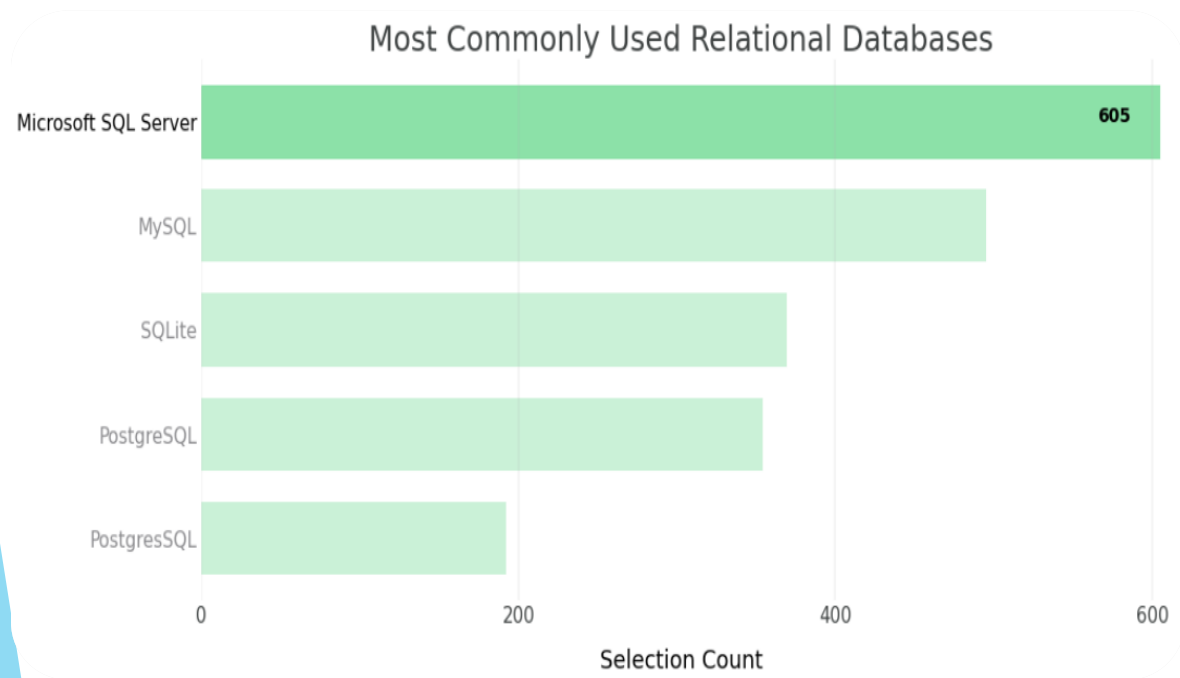
Coursera is the most common learning platform reportedly used by U.S. Data Scientist

Data Scientist



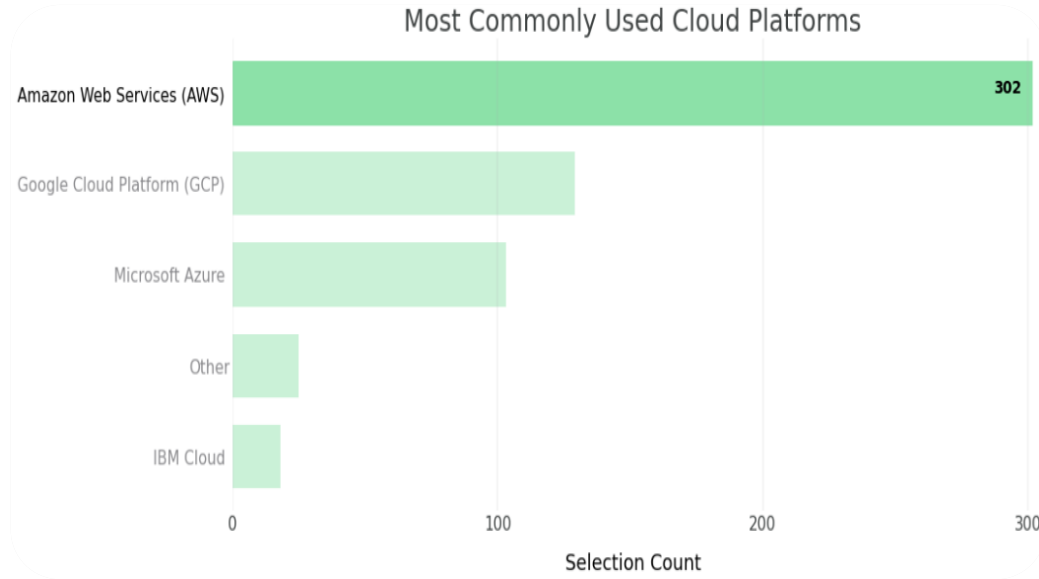
SQL and Python are the most commonly used programming languages for Data Scientist

Data Scientist



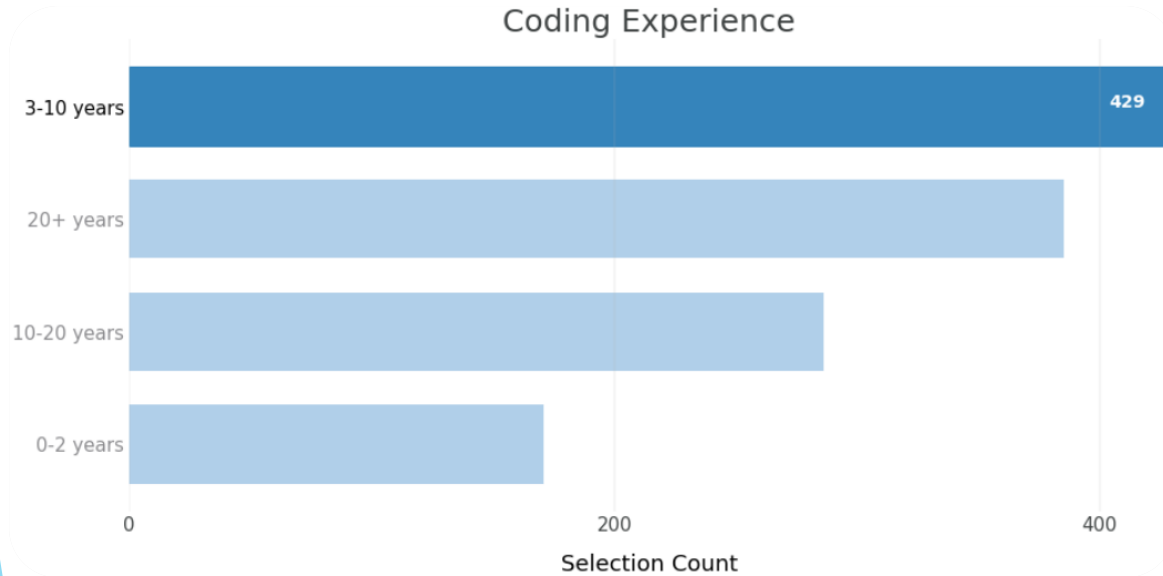
Microsoft SQL Server and MySQL show large use for Data Scientist

Data Scientist



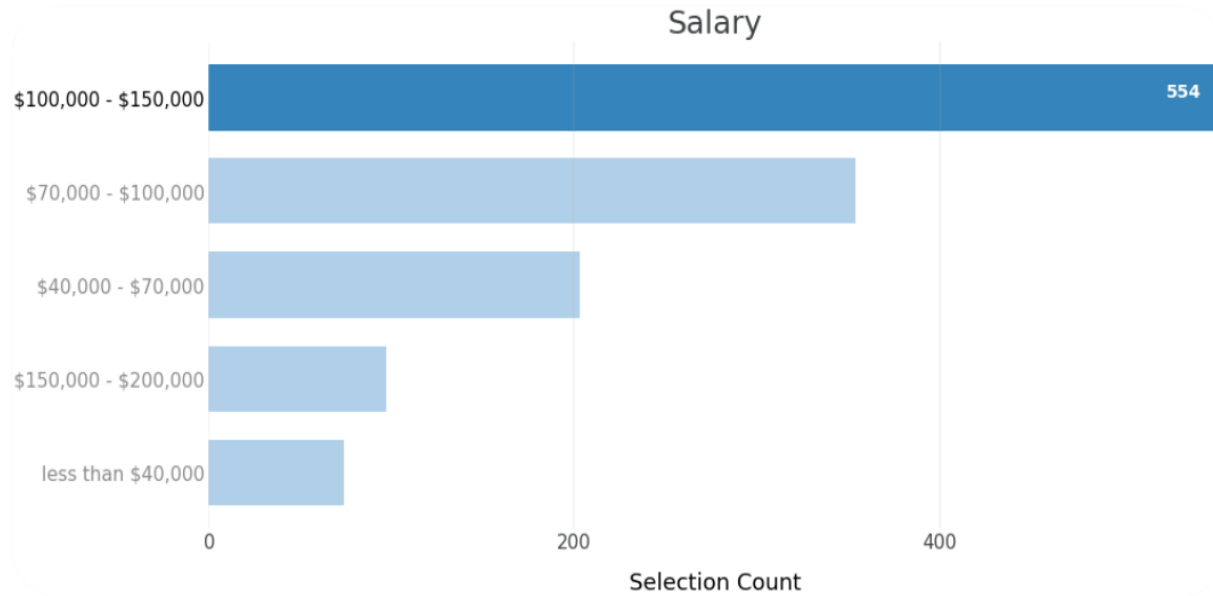
Amazon Web Service
dominates in terms
of usage for Data
Scientist

Data Analyst



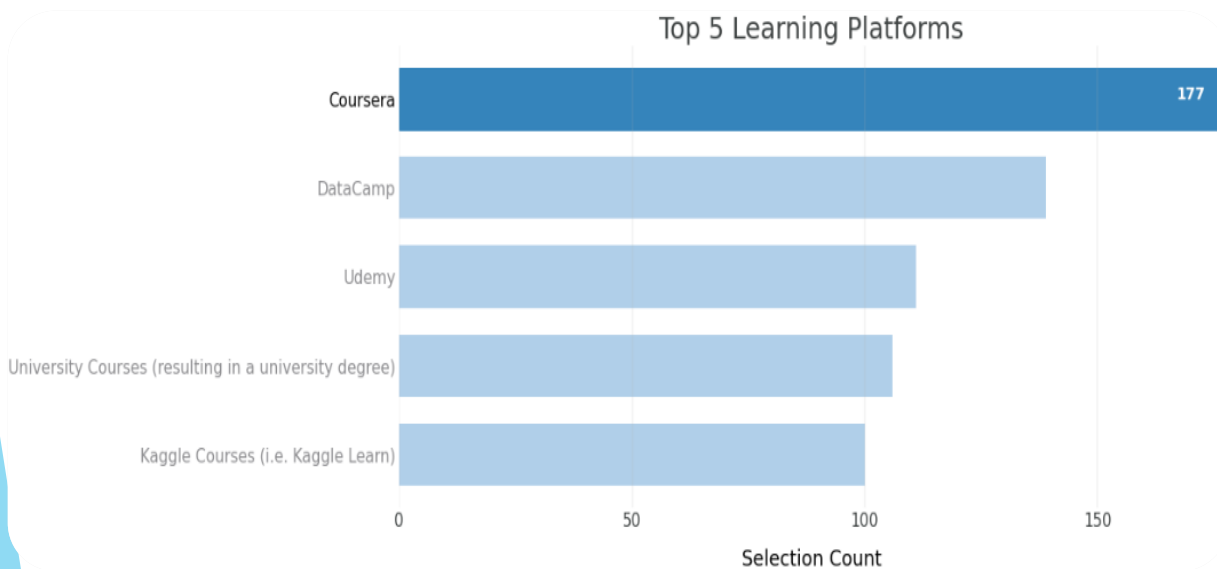
3 to 10 years of coding experience is common for U.S. data science professionals. On the aggregate though, coding experience is more likely to be >2 years

Data Analyst



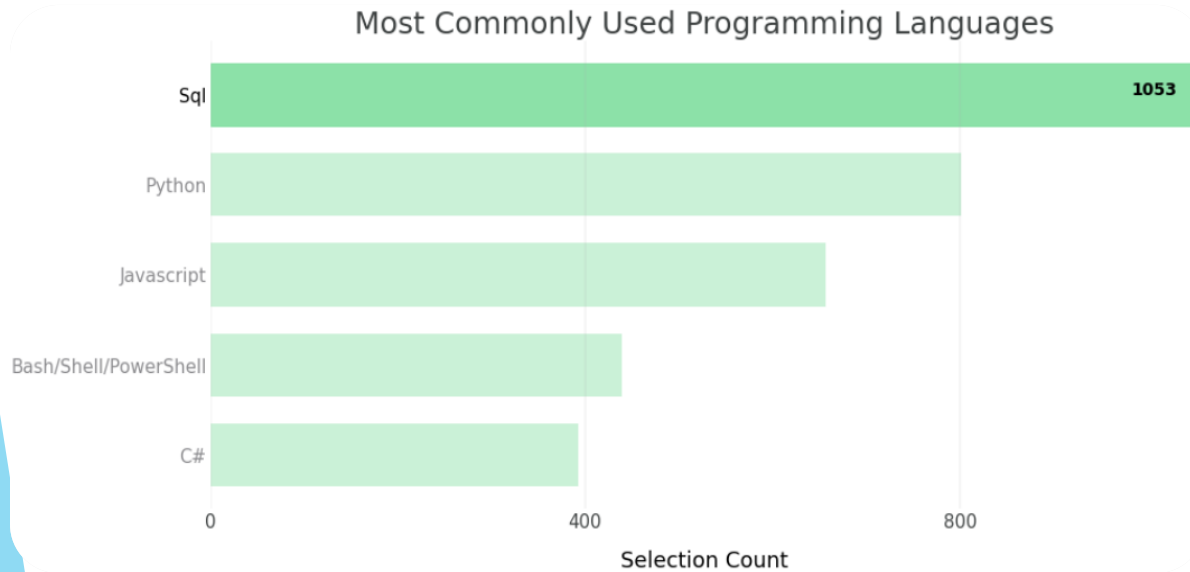
Data Analyst in the U.S. are reportedly paid well with \$100k to \$500k being the most common salary range

Data Analyst



Coursera is the most common learning platform reportedly used by Data Analysts; Data Camp follows closely behind

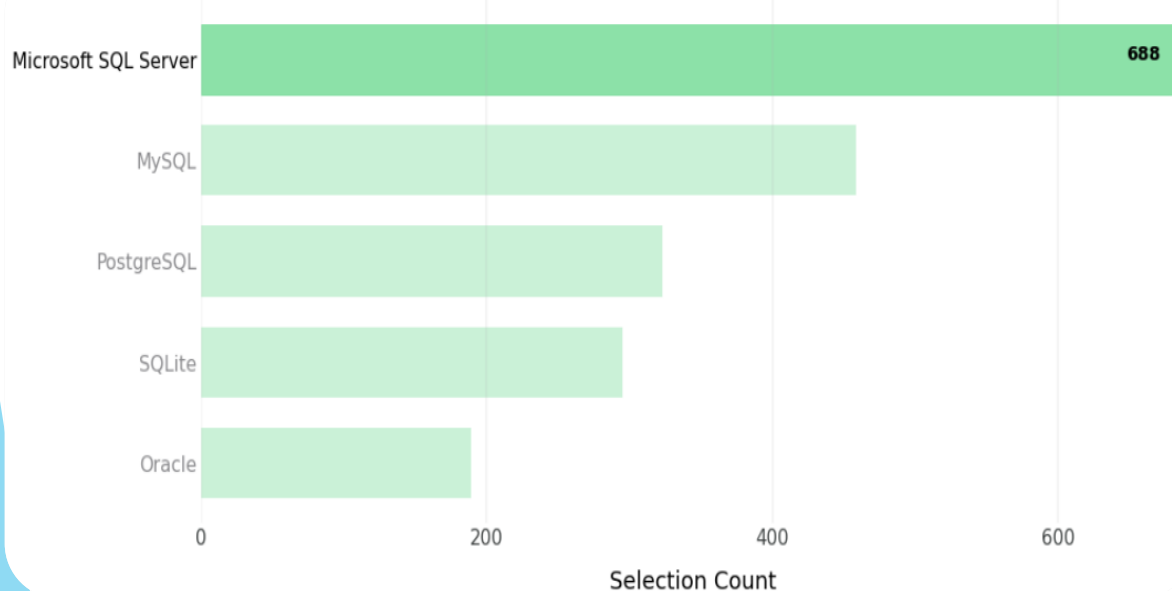
Data Analyst



SQL and Python
are the most
commonly used
programming
languages for Data
Scientist

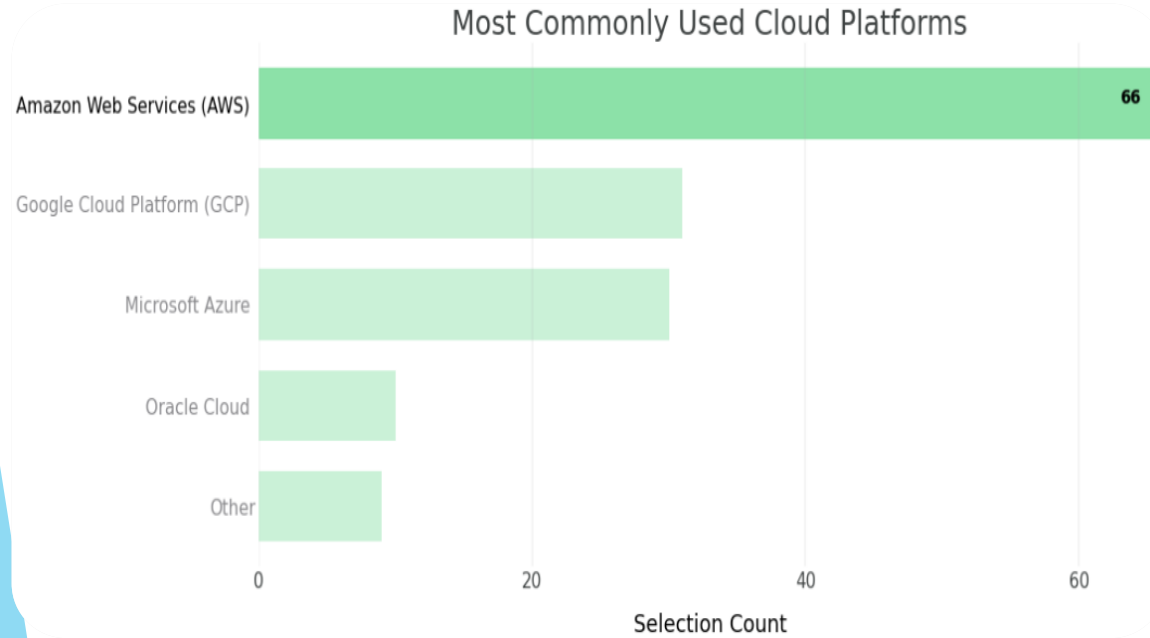
Data Analyst

Most Commonly Used Relational Databases



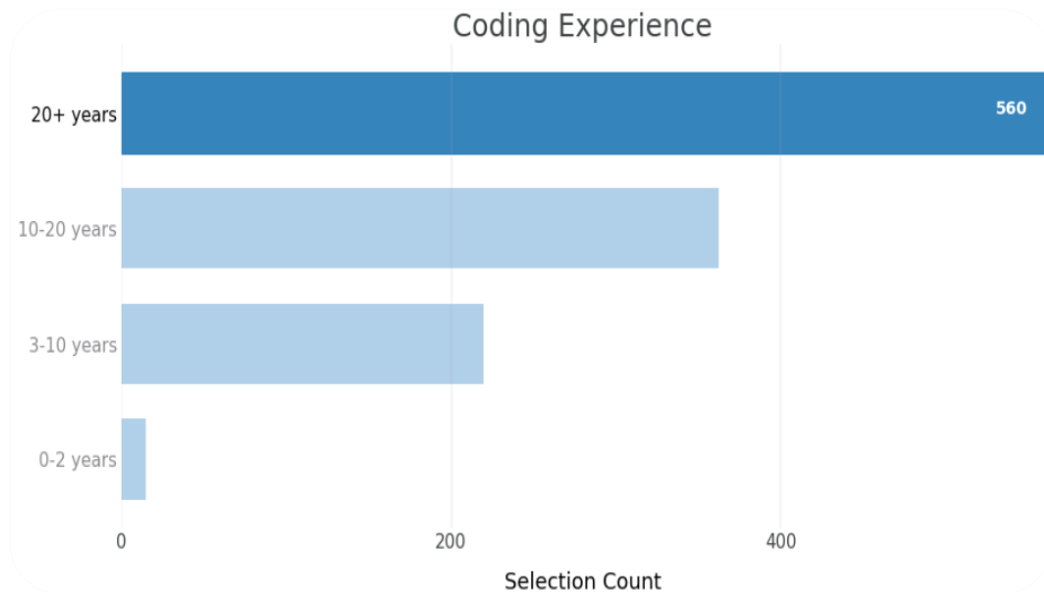
Microsoft SQL Server show large use for Data Analyst

Data Analyst



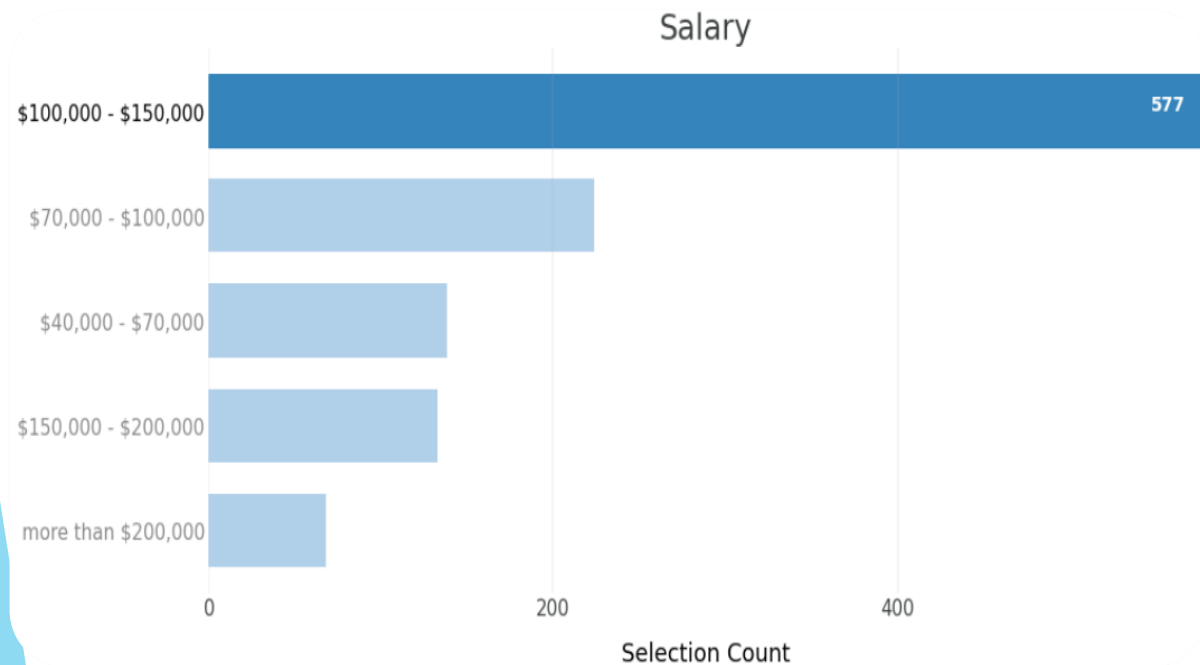
Amazon Web Service dominates in terms of usage for Data Analyst

Database Administrator (DBA)



It's not common to see 20+ years of coding experience for U.S. DBAs survey respondents

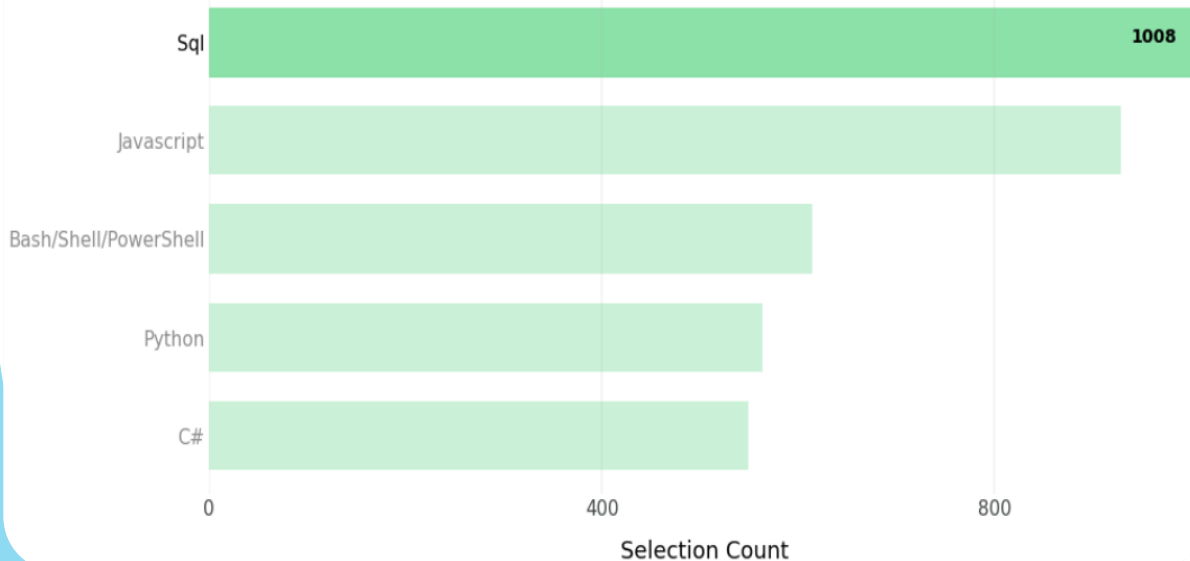
Database Administrator (DBA)



DBAs in the U.S. are reportedly paid well with \$100k to \$500k being the most common salary range

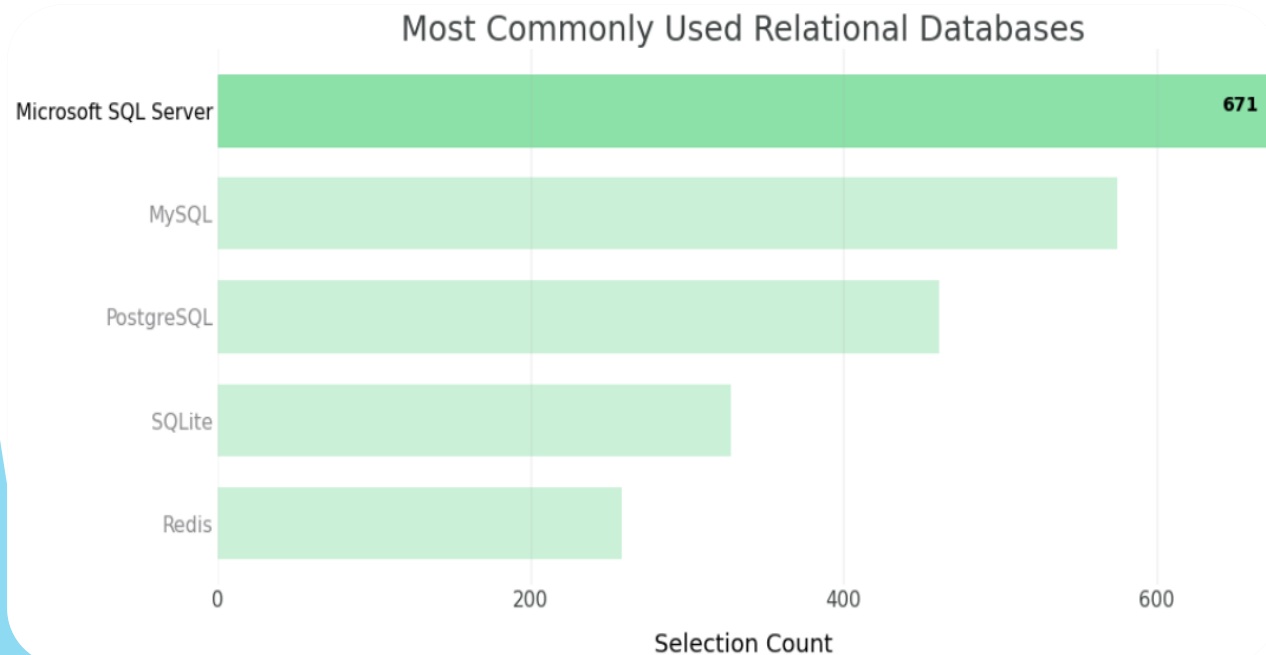
Database Administrator -Skills

Most Commonly Used Programming Languages



SQL and Python are the most commonly used programming languages for Data Scientist

Database Administrator -Skills



Microsoft SQL Server and MySQL show large use for DBAs

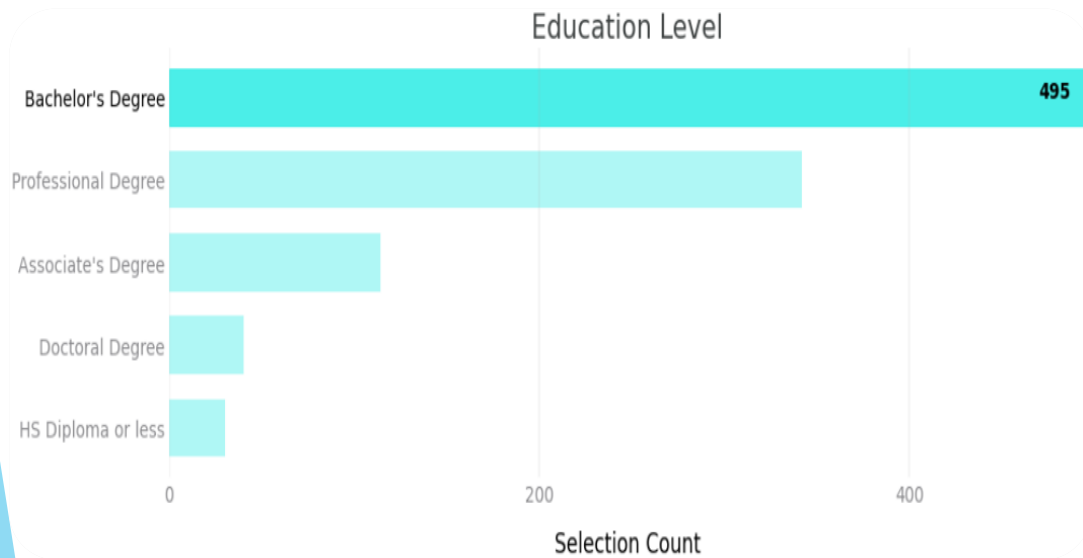
JOB MAP - Overall Conclusions

- Overall data-related jobs earn on average \$100,000 - \$150,000, most commonly have 11+ years of professional experience. Their most commonly used programming languages are Python and SQL. Moreover, their most commonly used relational database language and cloud platform were indicated to be Microsoft SQL Server and Amazon Web Service (AWS), respectively.
- Individuals surveyed seem to be well-established in their fields as indicated by their most common work experience.

JOB MAP - Overall Recommendations

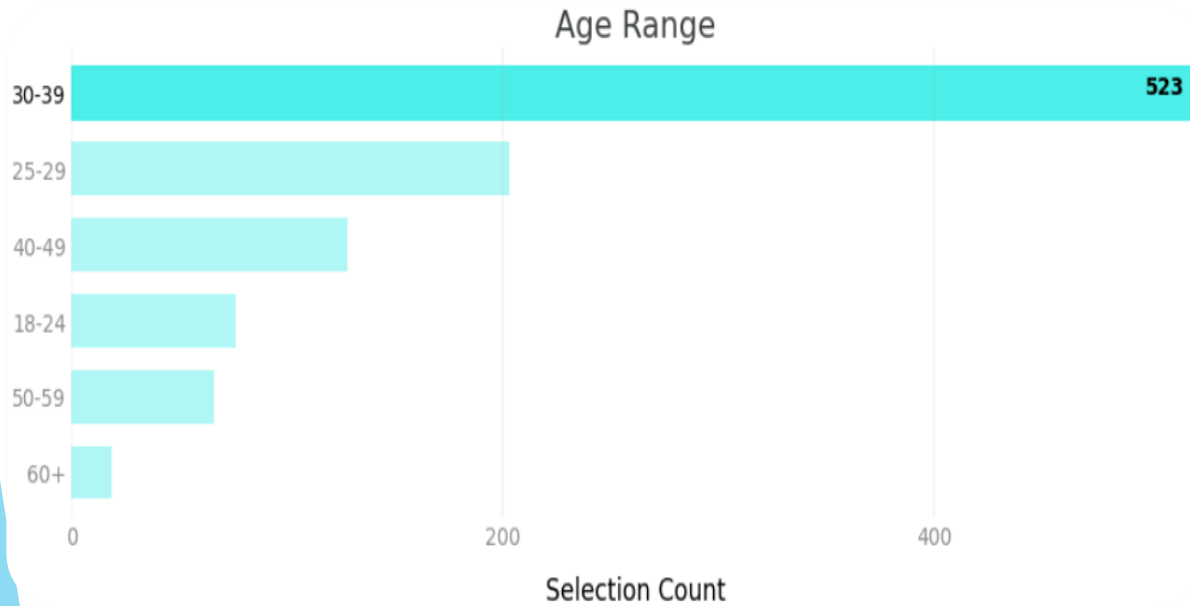
- ↩ An individual aiming to obtain a job as a data scientist, data analyst, database administrator, or data engineer, should look into learning SQL (more specifically, Microsoft SQL Server), cloud-based platforms (more specifically Amazon Web Service) and Python as they are the main tools used in a data-related role.

Data Engineer - Demographics



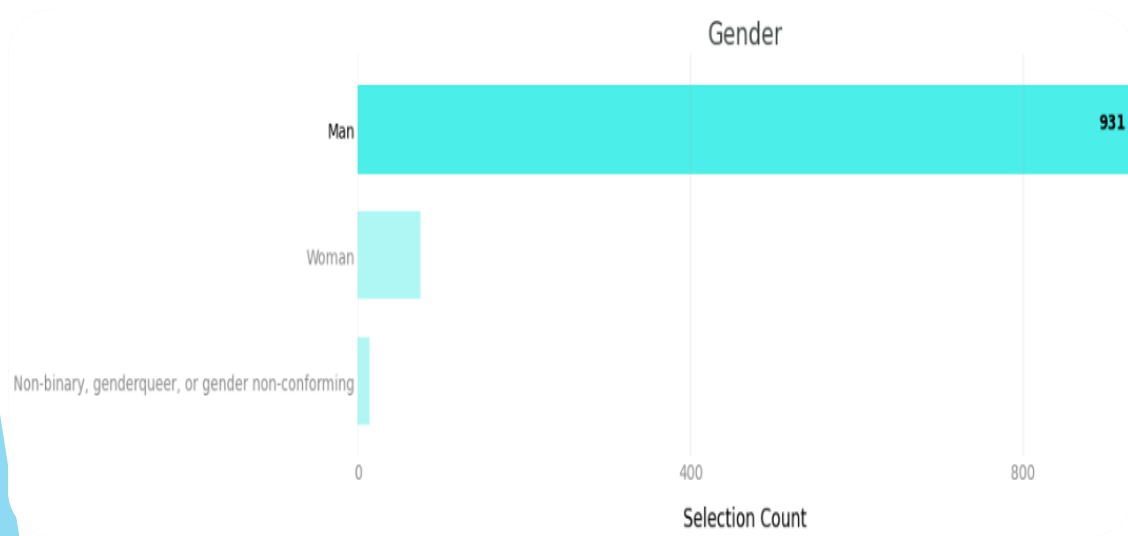
Data Engineers are mostly educated with a Bachelors degree and Professional degree being the most common education level

Data Engineer - Demographics



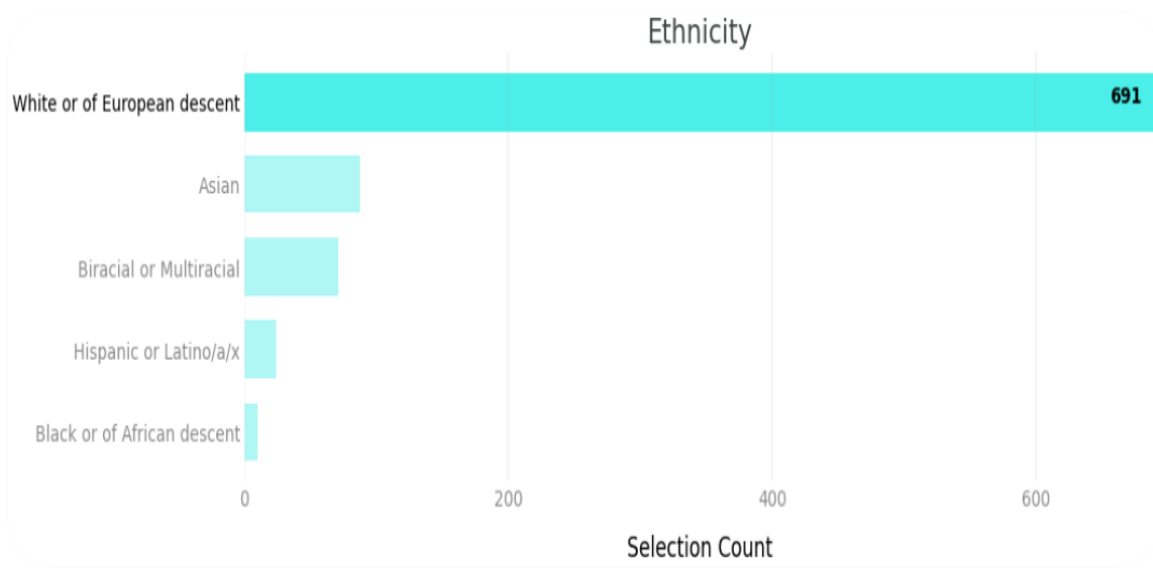
Data Engineers
are most
commonly in
their 30s to 39
year old

Data Engineer - Demographics



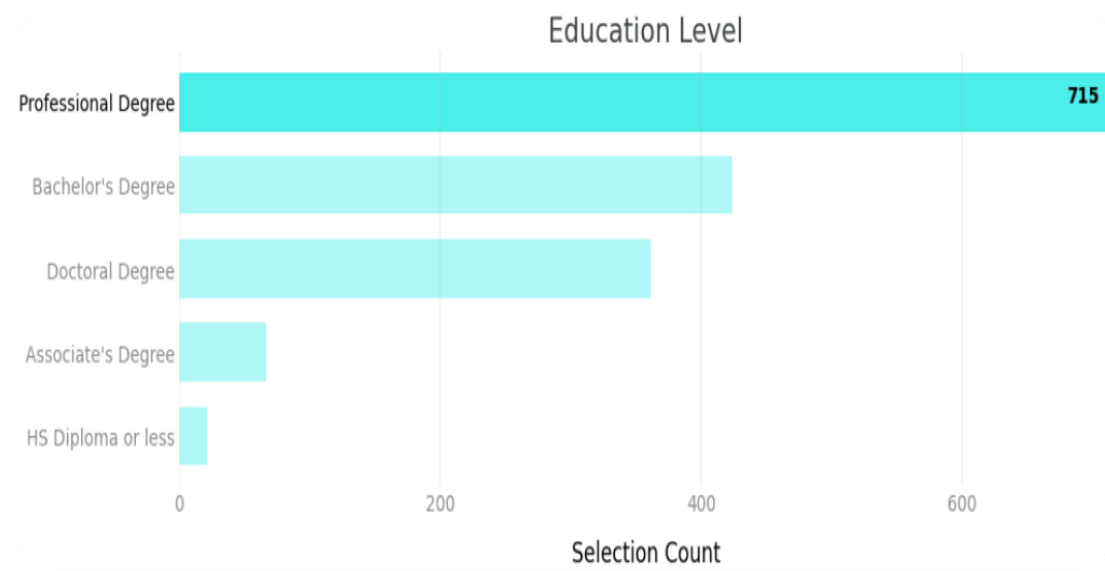
Data Engineers
are
overwhelmingly
males with
women being
heavily
underrepresented

Data Engineer - Demographics



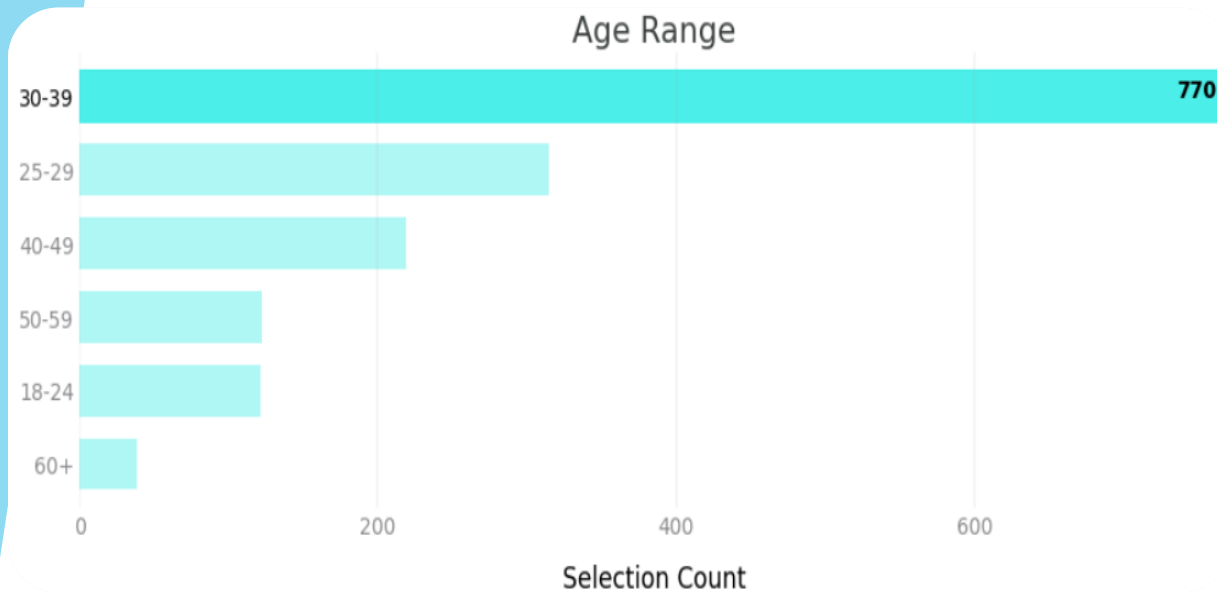
Data Engineers are most commonly White or of European descent with other races not being as represented

Data Scientist - Demographics



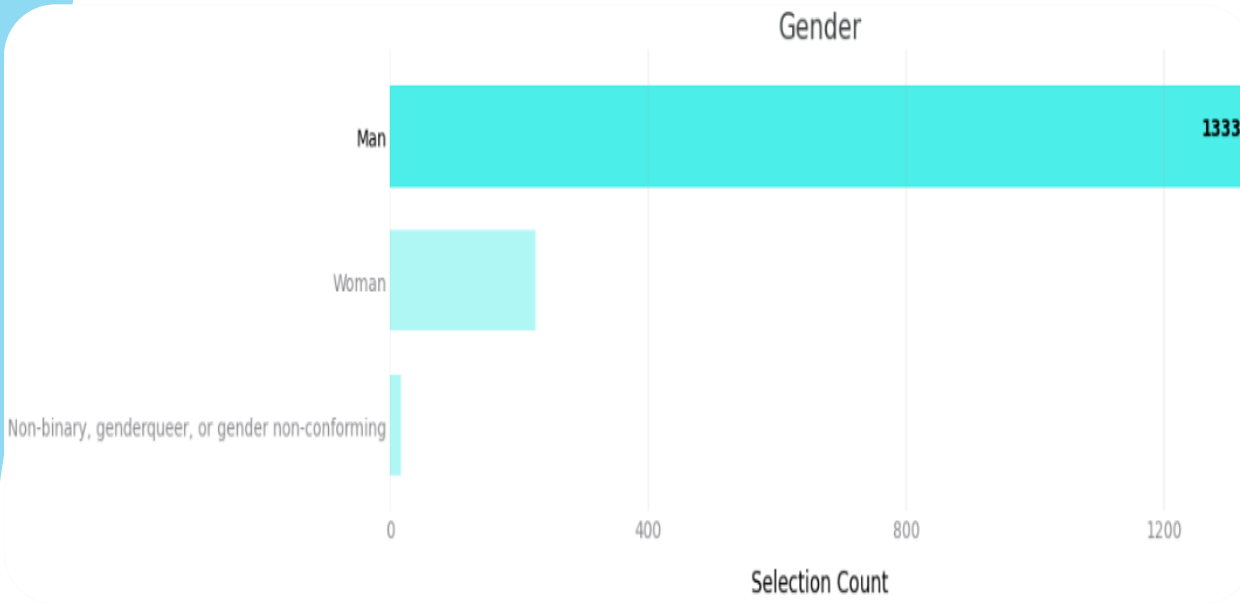
Data Scientist are mostly educated with a Professional degree and Bachelors degree being the most common education level

Data Scientist-Demographics



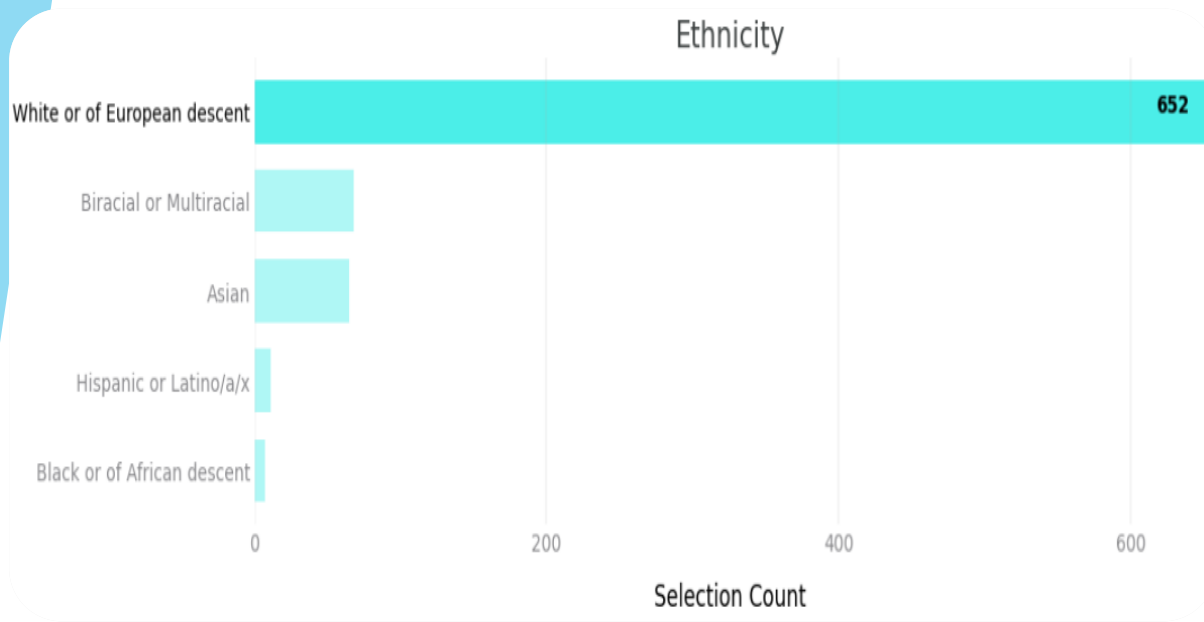
Data Scientist
are most
commonly in
their 30s to 39
year old

Data Scientist - Demographics



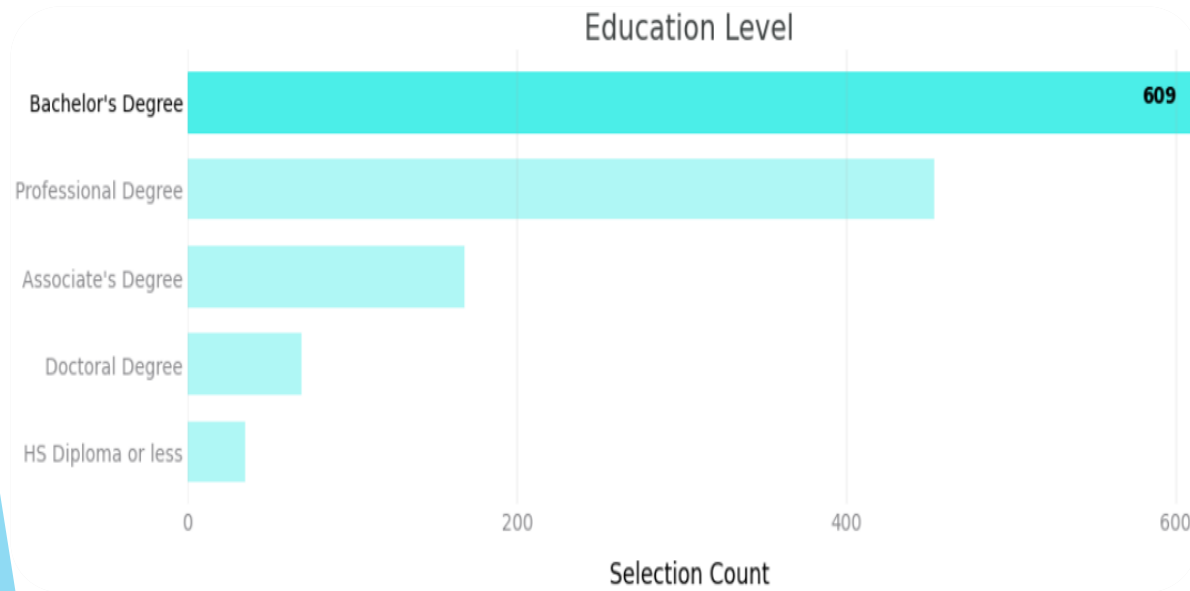
Data Scientist are overwhelmingly males with women being heavily underrepresented

Data Scientist- Demographics



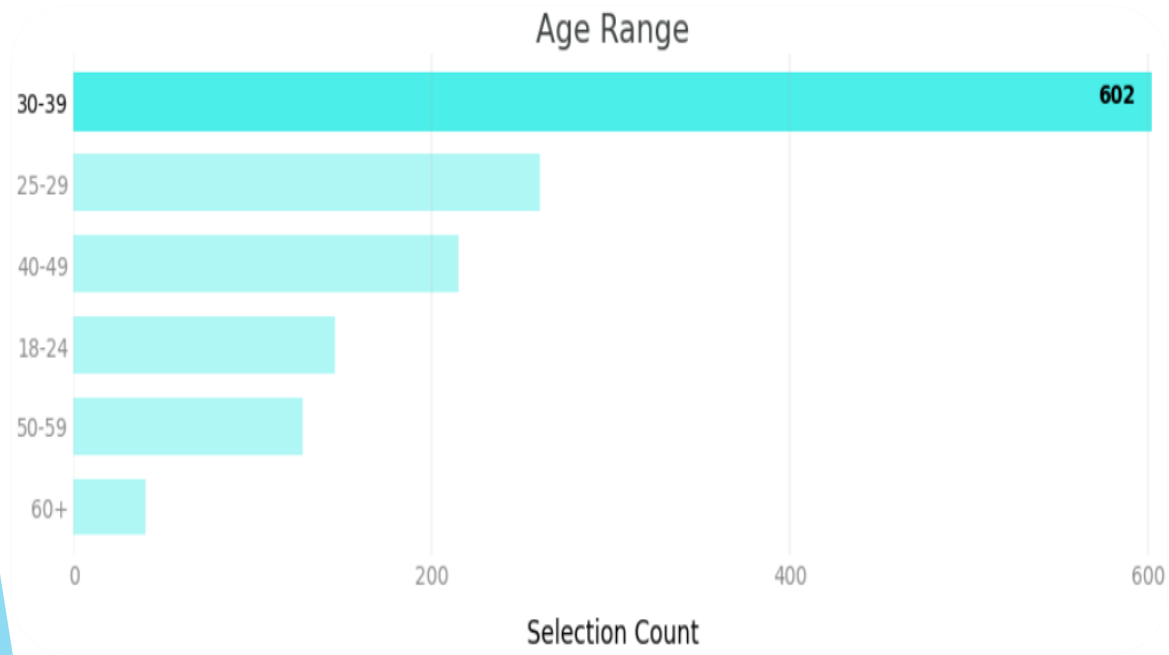
Data Scientist are most commonly White or of European descent with other races not being as represented

Data Analyst - Demographics



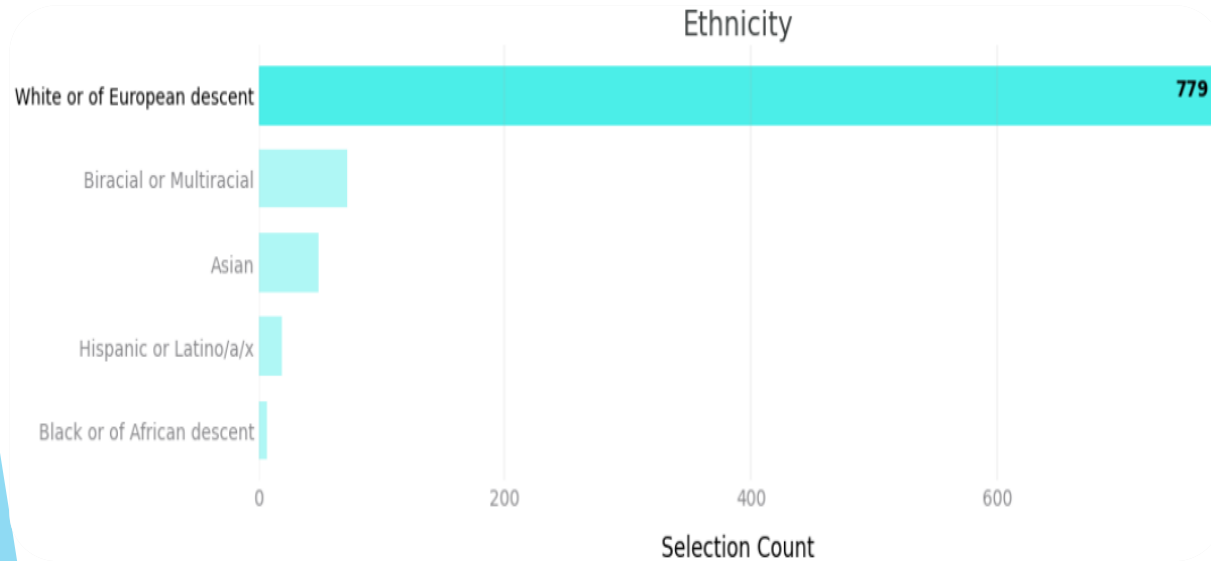
Data Analysts are mostly educated with a Professional degree and Bachelors degree being the most common education level

Data Analyst -Demographics



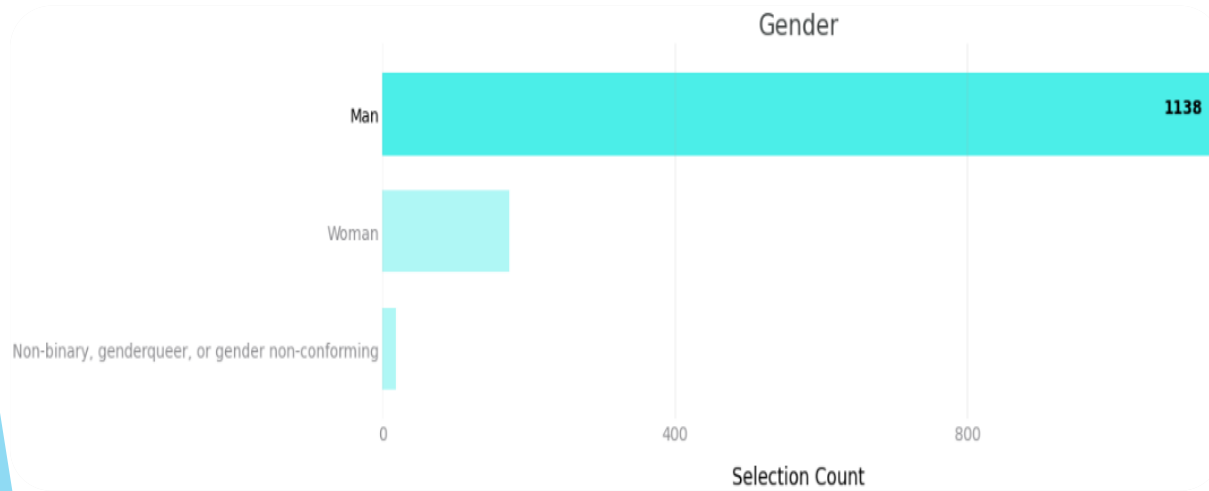
Data Analysts
are most
commonly in
their 30s to 39
year old

Data Analyst - Demographics



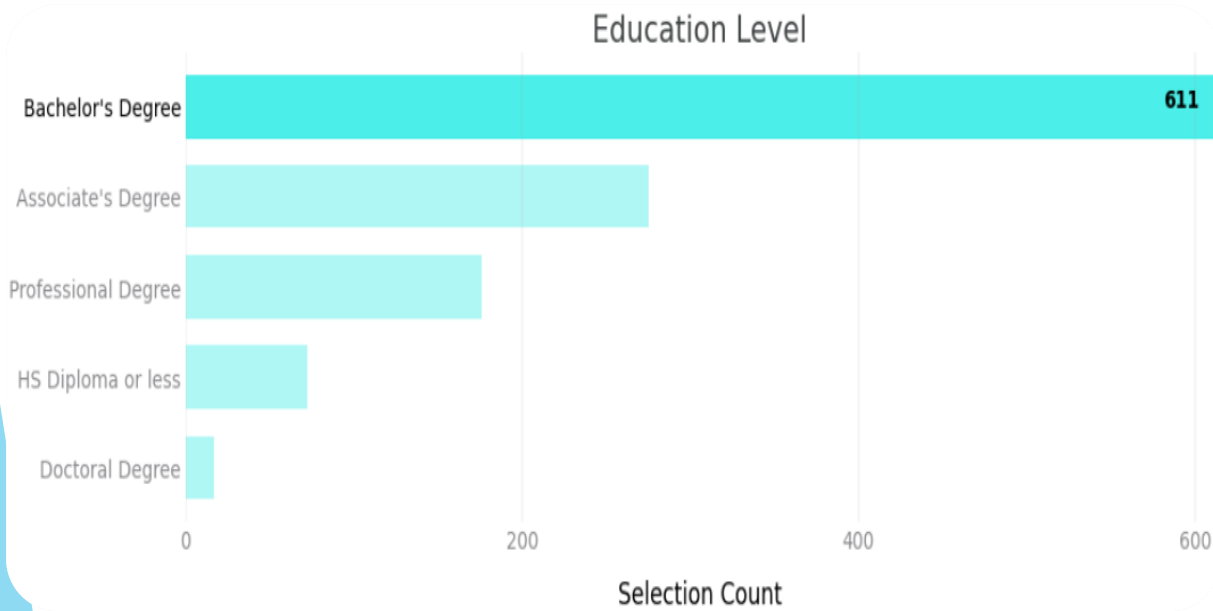
Data Analysts are most commonly White or of European descent with other races not being as represented

Data Analyst - Demographics



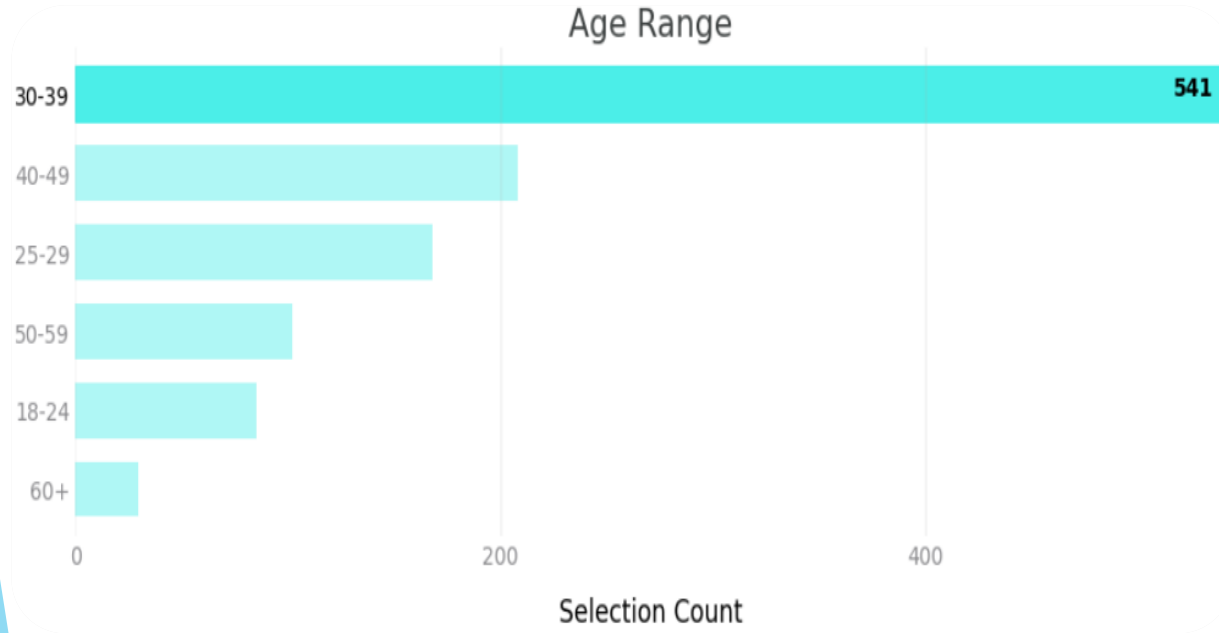
Data Analysts are overwhelmingly males with women being heavily underrepresented

DBA - Demographics



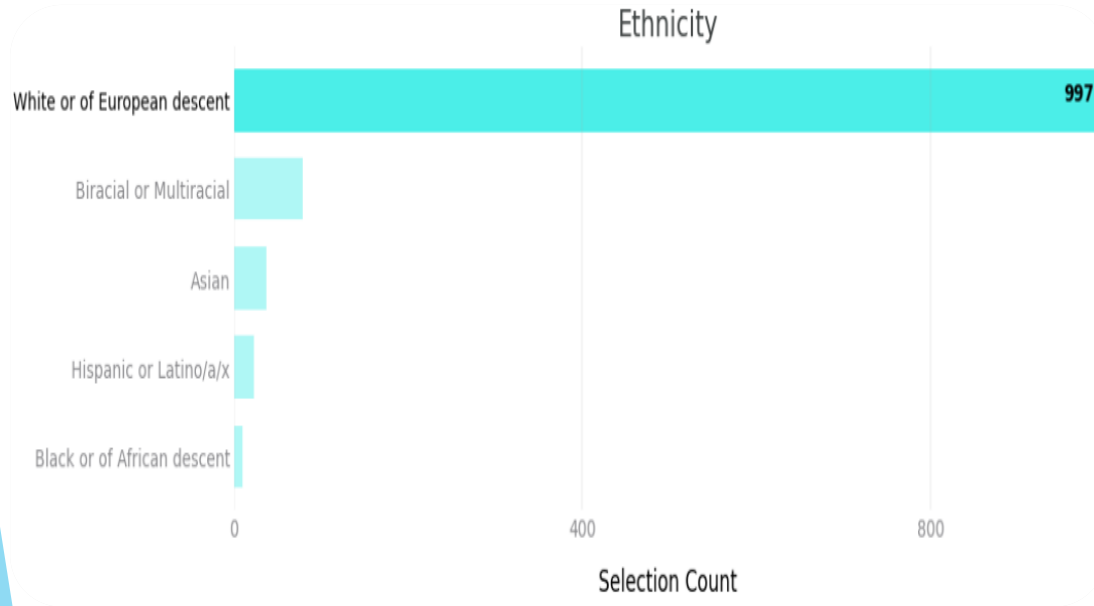
DBAs are mostly educated with a Bachelors degree being the most common education level

DBA - Demographics



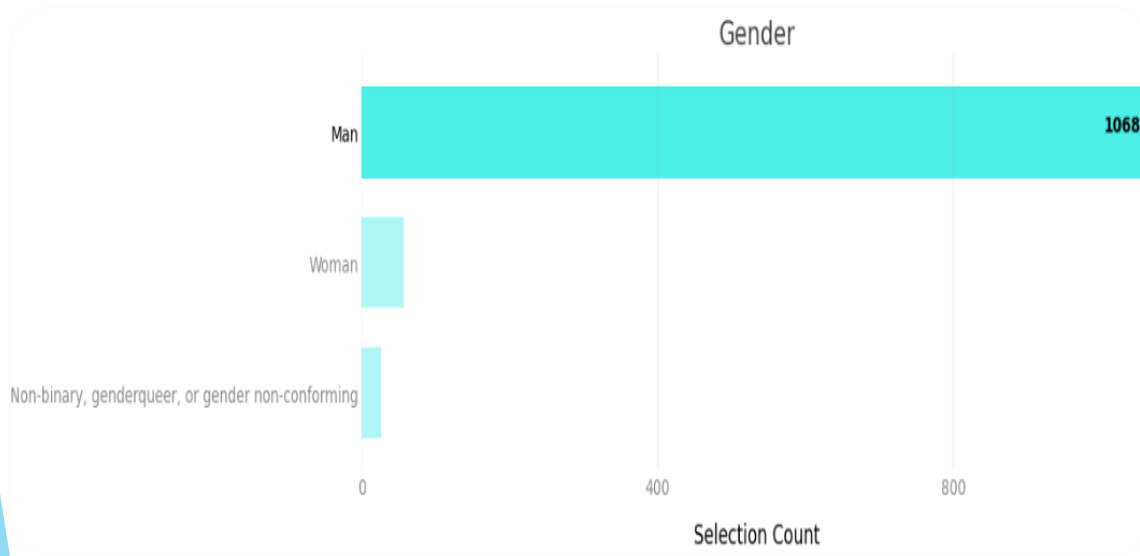
DBAs are most commonly in their 30s to 39 year old

DBA - Demographics



DBAs are most commonly White or of European descent with other races not being as represented

DBA - Demographics



DBAs are overwhelmingly males with women being heavily underrepresented

DEMOGRAPHIC MAP - CONCLUSION

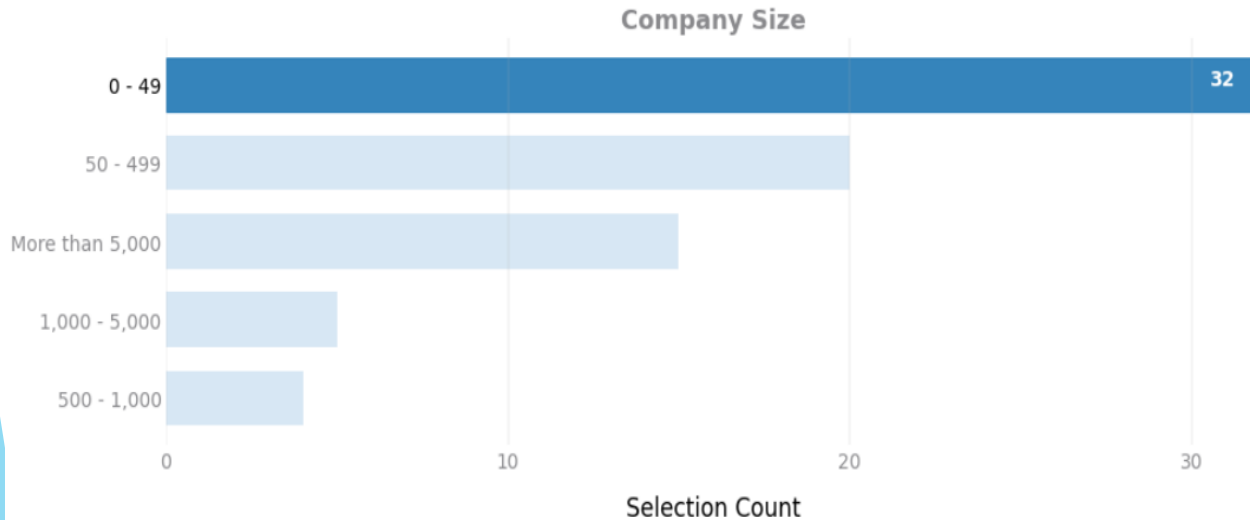
- ⌞ Across all the data-related jobs explored as well as in the tech industry, the typical demographic profile is the following: Straight/heterosexual White or European descendant Male whose age range is between 30 - 39.
- ⌞ For data-related jobs in the tech industry, the common salary, oddly enough, was indicated to being \$40,000-\$70,000 with education attained being most commonly in the form of a Bachelor's degree or Associate's Degree (or some college). Moreover, very few of respondents identified as having a disability status.

DEMOGRAPHIC MAP

- RECOMMENDATIONS

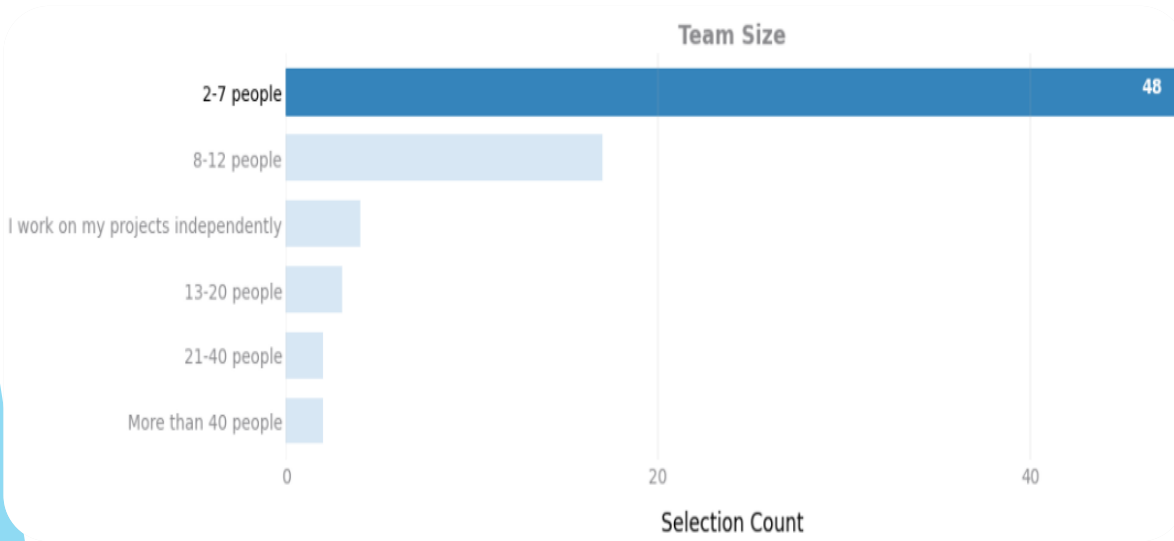
- ↳ There is certainly a lot of room for improvement in terms of diversity in data-related roles. Nothing to discourage but something worth considering for those not represented in the field. More initiatives or programs of the like should be launched to address this gap in diversity which is clearly prevalent in the tech field and data-related roles.

IT Industry



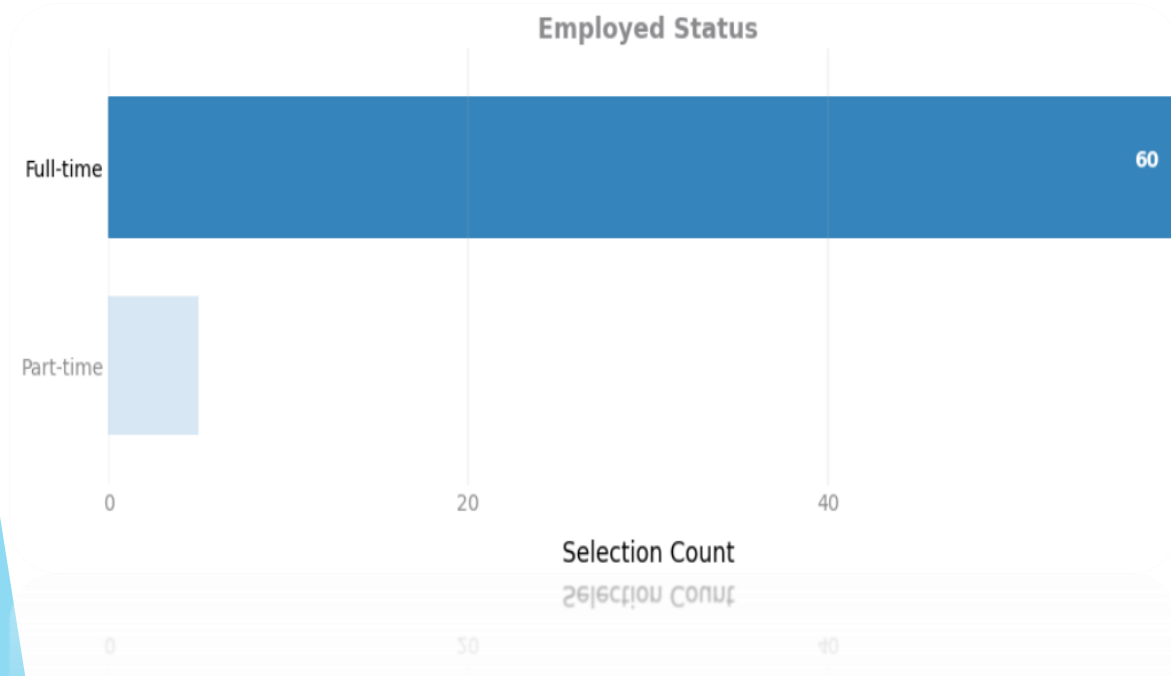
Companies in the IT industry are small-to-mid size (0 to 499 people)

IT Industry



Companies in the IT industry most commonly have a small team size of between 2 to 7 people

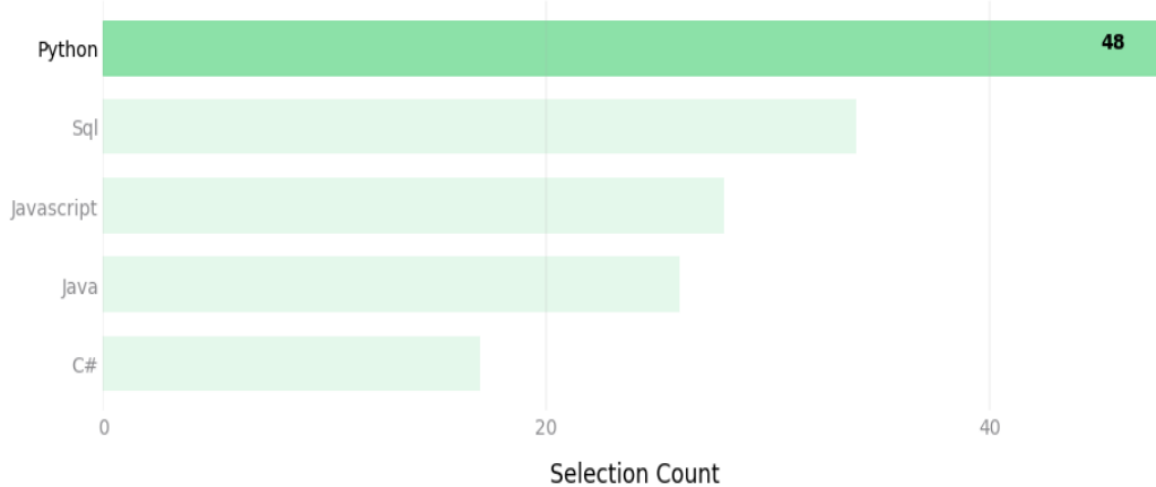
IT Industry



Most companies in the IT industry hire people full-time with employees working part-time

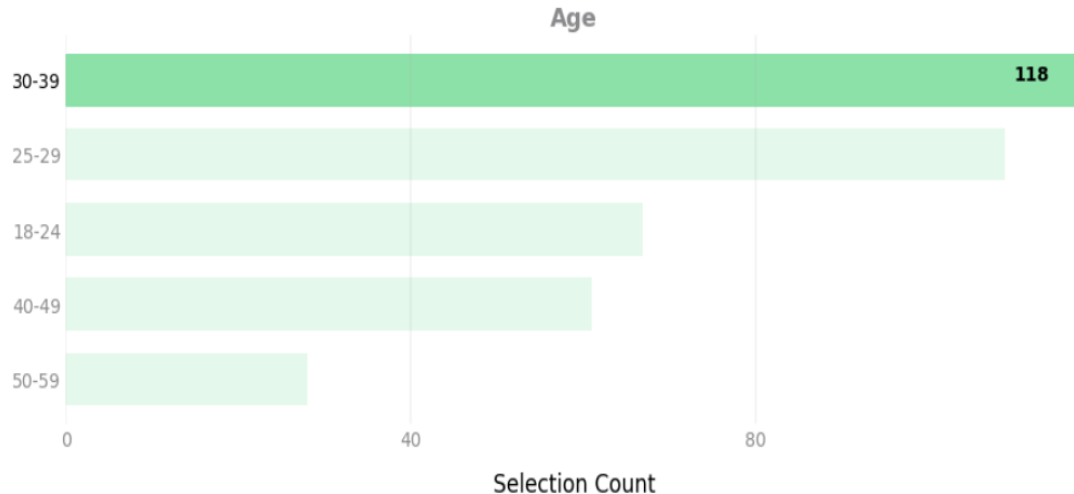
IT Industry

Most Commonly Used Programming Languages



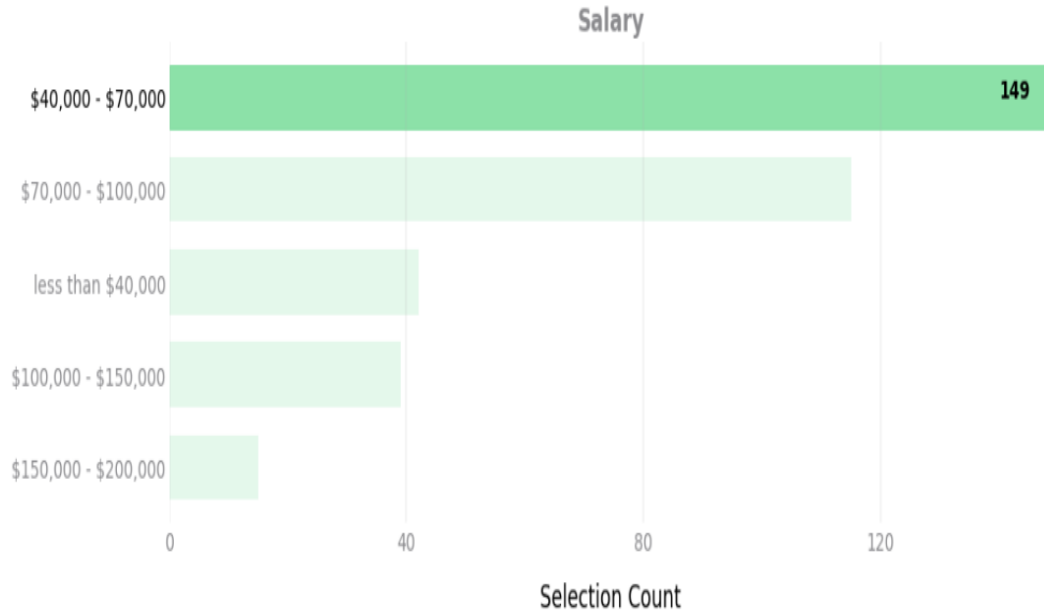
For companies in the IT industry Python and SQL are the most commonly used programming languages

IT Industry



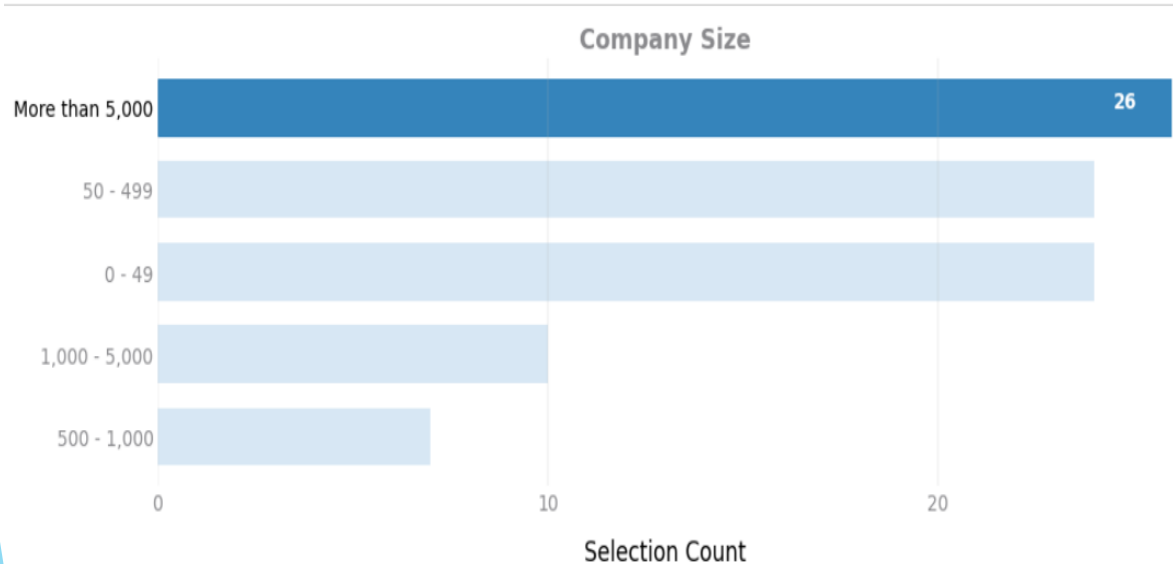
Employees working in an IT industry are young to early mid-age and are between the ages of 25 to 39

IT Industry



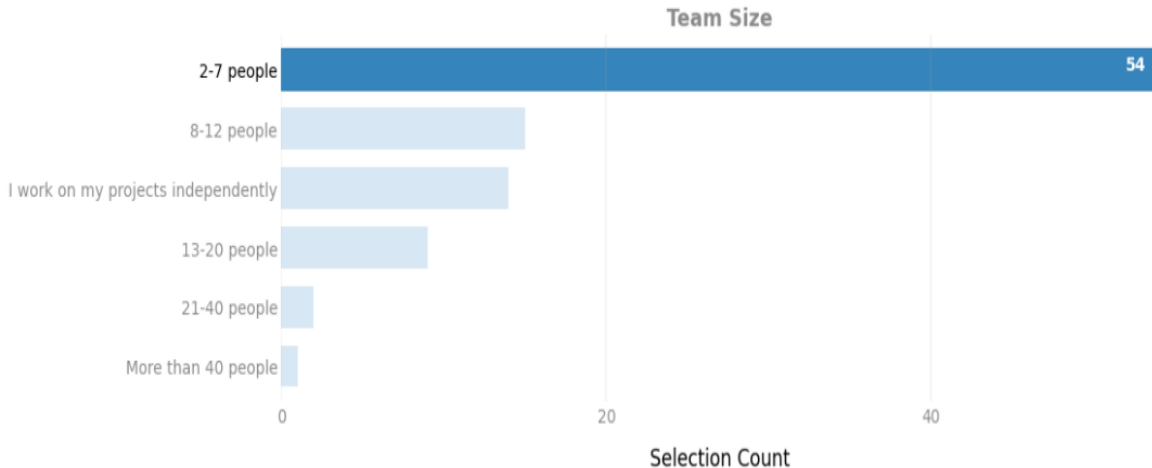
Employees in the IT industry are most commonly paid a salary between the ranges of \$40k to \$100k

Non-IT Industry



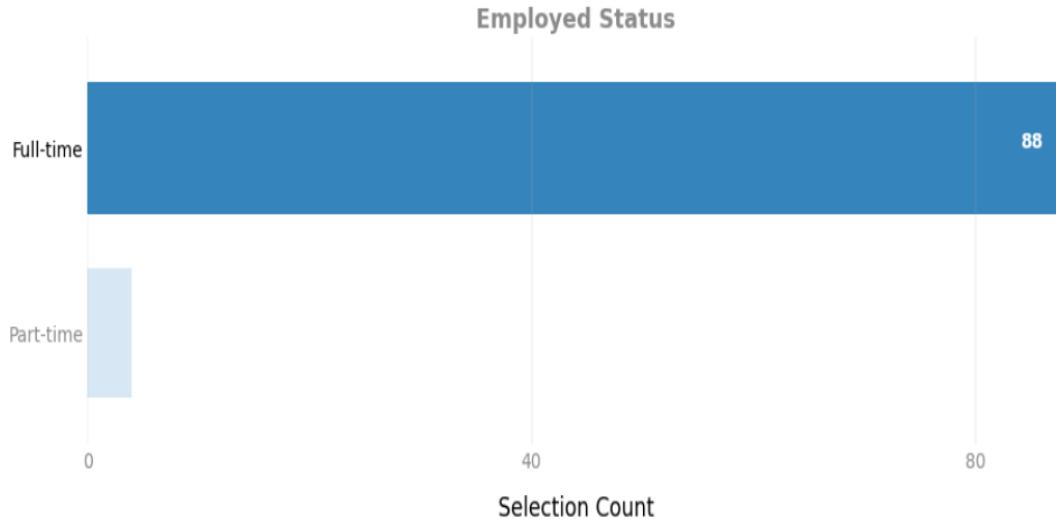
Companies in the non-IT industry are relatively between 0 to 500 people or more than 5,000 people

Non-IT Industry -General



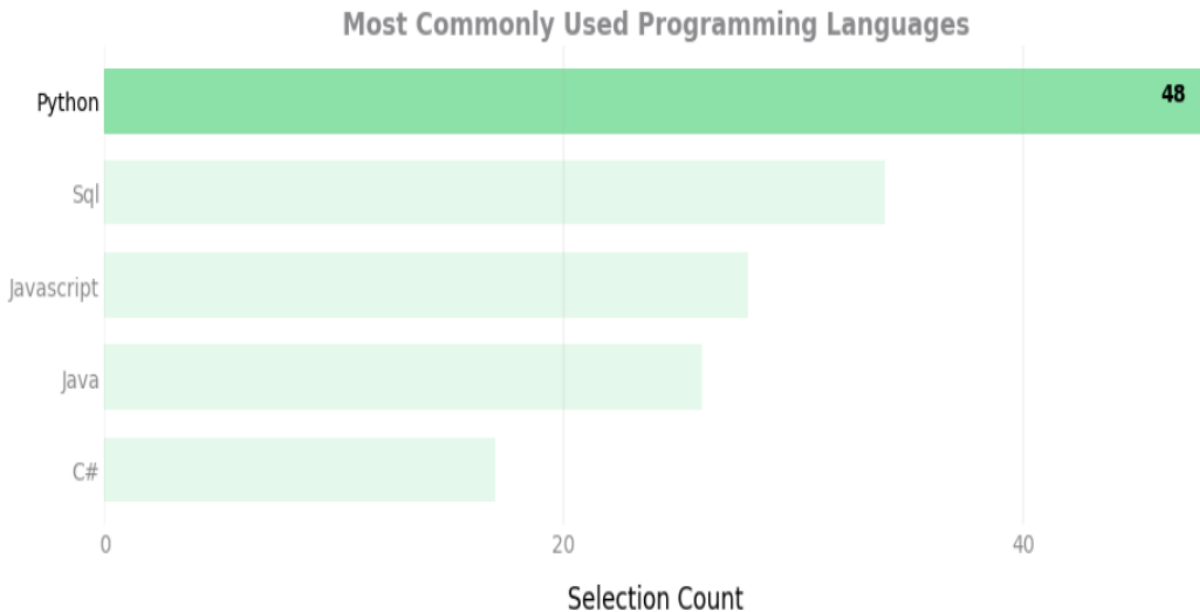
Companies in the non-IT industry most commonly have a small team size of between 2 to 7 people

Non-IT Industry -General



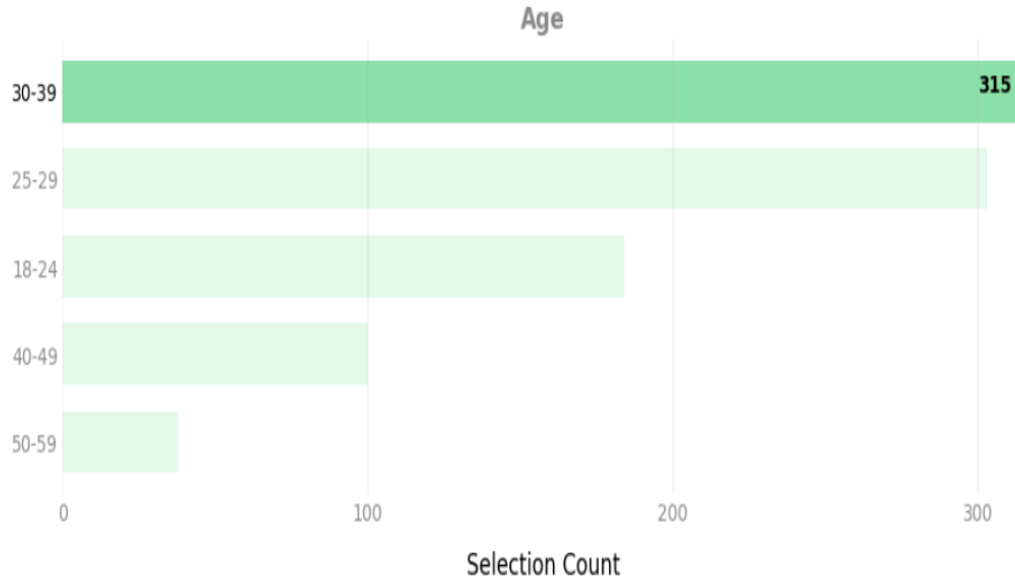
Most companies in the non-IT industry hire people full-time with employees working part-time

Non-IT Industry



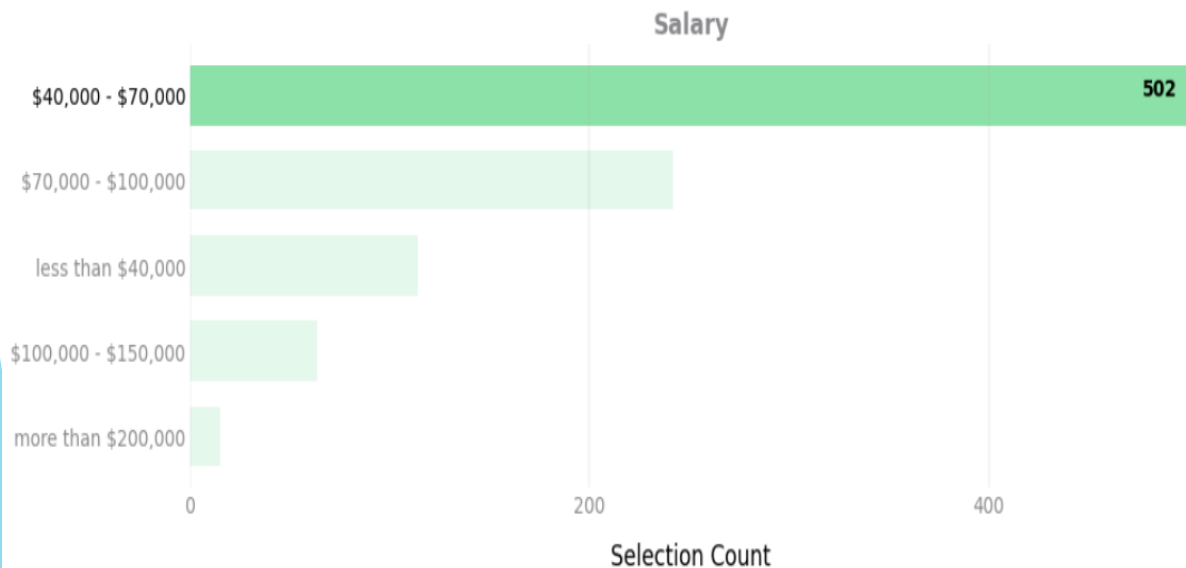
For companies in the non-IT industry Python and SQL are the most commonly used programming languages

Non-IT Industry



Employees working in a non-IT industry are young to early mid-age and are between the ages of 25 to 39

Non-IT Industry



Employees in the non-IT industry are most commonly paid a salary between the ranges of \$40k to \$70k

EMPLOYER PROFILE- CONCLUSION

Across the industries, the most commonly used programming languages are Python and SQL and in terms of team size, 2-7 people was the most common reported range. It was interesting to see that the most common degrees were in the form of a Bachelor's or Associate's Degree (or at least some college) hinting at advanced degree becoming less and less common for data-related roles. Of the respondents, most identified as being in the age ranges of 25-39, male and not part of a minority group.

For non-IT companies, \$40,000-\$70,000 was the most common salary range whereas for IT companies \$40,000-\$100,000 was more common. Lastly, what stood out was the company size of which a large number (> 500 people) was most commonly reported for companies in non-IT industries in comparison with the relatively small company size of 0 to 500 people reported for companies in the IT industry.

EMPLOYER PROFILE

- RECOMMENDATIONS

If we are to consider both industries together, Python and SQL are great programming languages to learn as they are most commonly used in a professional setting. In addition, learning how to utilize cloud platforms such as Amazon Web Service and Microsoft Azure might be a good idea if one is considering to join a team that works with "big data." Moreover, working in small team size of about 2 - 7 people is most probable and thus suggests the importance of soft skills playing a role as a data professional.