Emerging Trends in Data Analytics - Capstone

Emily Sanchez Flores – July 2025

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

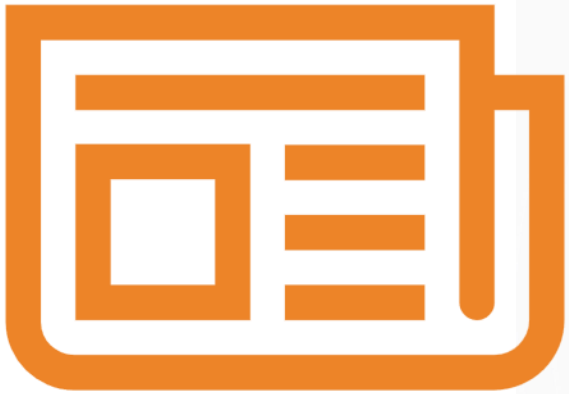
EXECUTIVE SUMMARY

Introduction and Metholology	•Stack Overflow Annual Developer survey background •Methodology used to analyze survey results include data collection, web scraping, data wrangling, exploratory data analysis, data validation, and dashboarding
Programming and Database Trends	•Top 10 current and future programming languages and databases used by survey respondents •Programming and database trends findings and implications
Dashboards	•Dashboards representing programming languages, databases, platforms, and web frames used by survey respondents, as well as user demographics
Overall Findings and Conclusion	•Programming language, database, platform, and web frame usage trends indicate what tools are widely used, what tools will continue to be popular among users, and tools that are falling out of popularity and user opinion

INTRODUCTION

- Stack Overflow is a question-and-answer site for programmers and developers. Their annual Developer Survey is the largest and most comprehensive survey of programmers and developers in the world.
- The survey represents Stack Overflow users, but it is a good indication of language choice trends, database usage, platform usage, and demographics of the users.
- This survey gives readers a chance to understand the characteristics of how people are working, how they want to work, and who is working. It could also assist in deciding what tools to use in the future as their communities grow and usage increases.

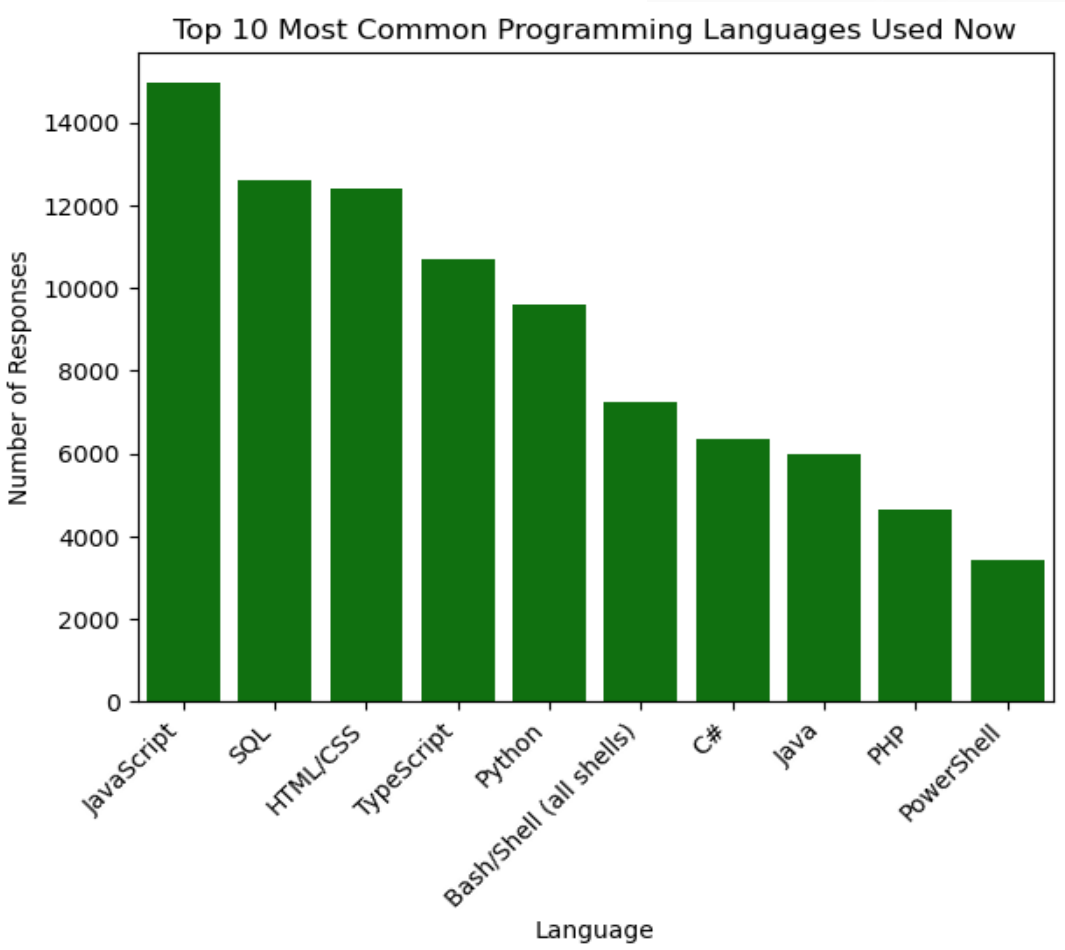
METHODOLOGY



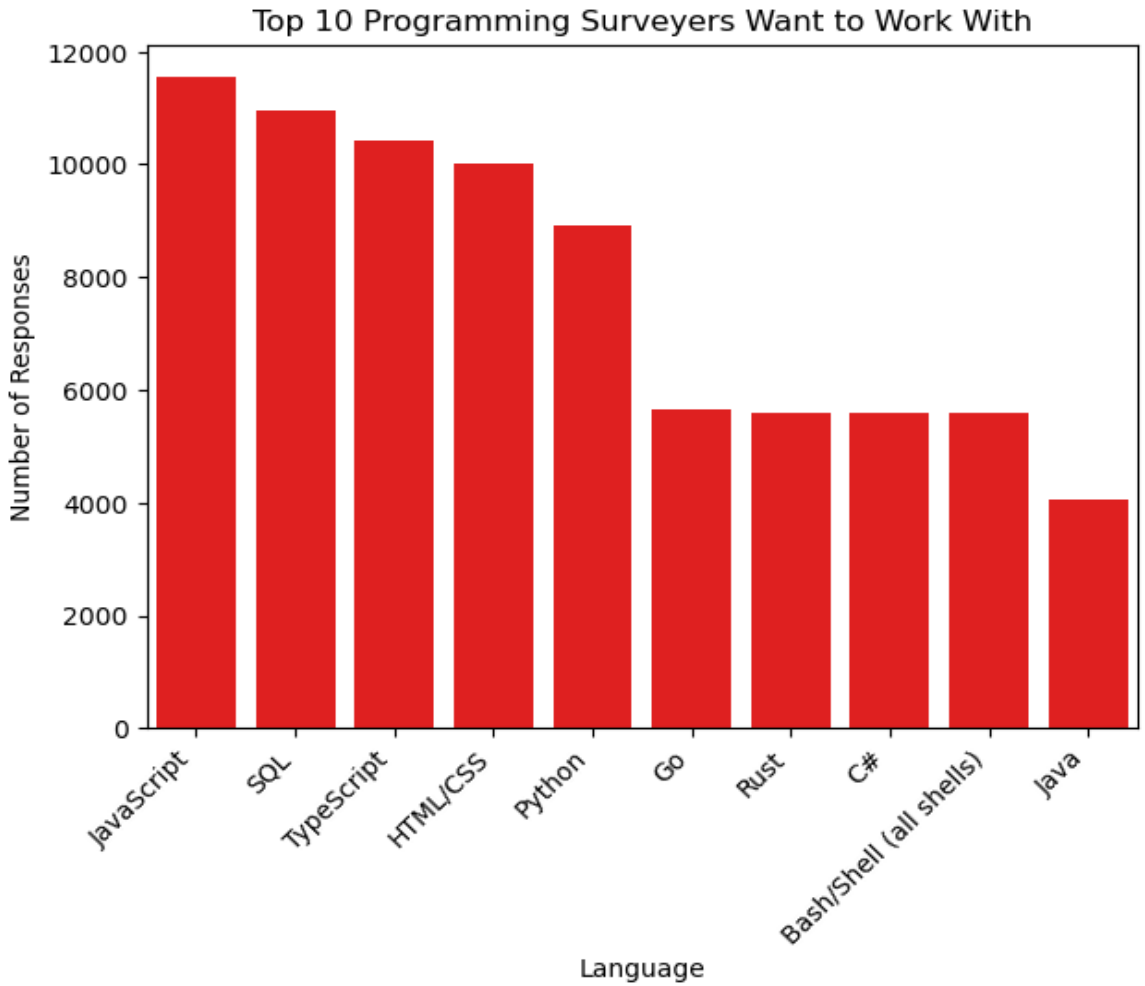
- Collect survey data and explore the content
  - Collected through web scraping
  - APIs
  - Requests (library)
- Data Wrangling
  - Discovery, cleaning, removing duplicates, validating
- Exploratory Data Analysis
  - Analyzing distribution of data collected
  - Identifying and processing outliers, skewness.
  - Calculating and identifying correlations
- Data Visualization
  - Creating various charts and plots to display the story the data is telling.
- Creating Dashboards to compile data and key information into an easy to read and update format

PROGRAMMING LANGUAGE TRENDS - 2025

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

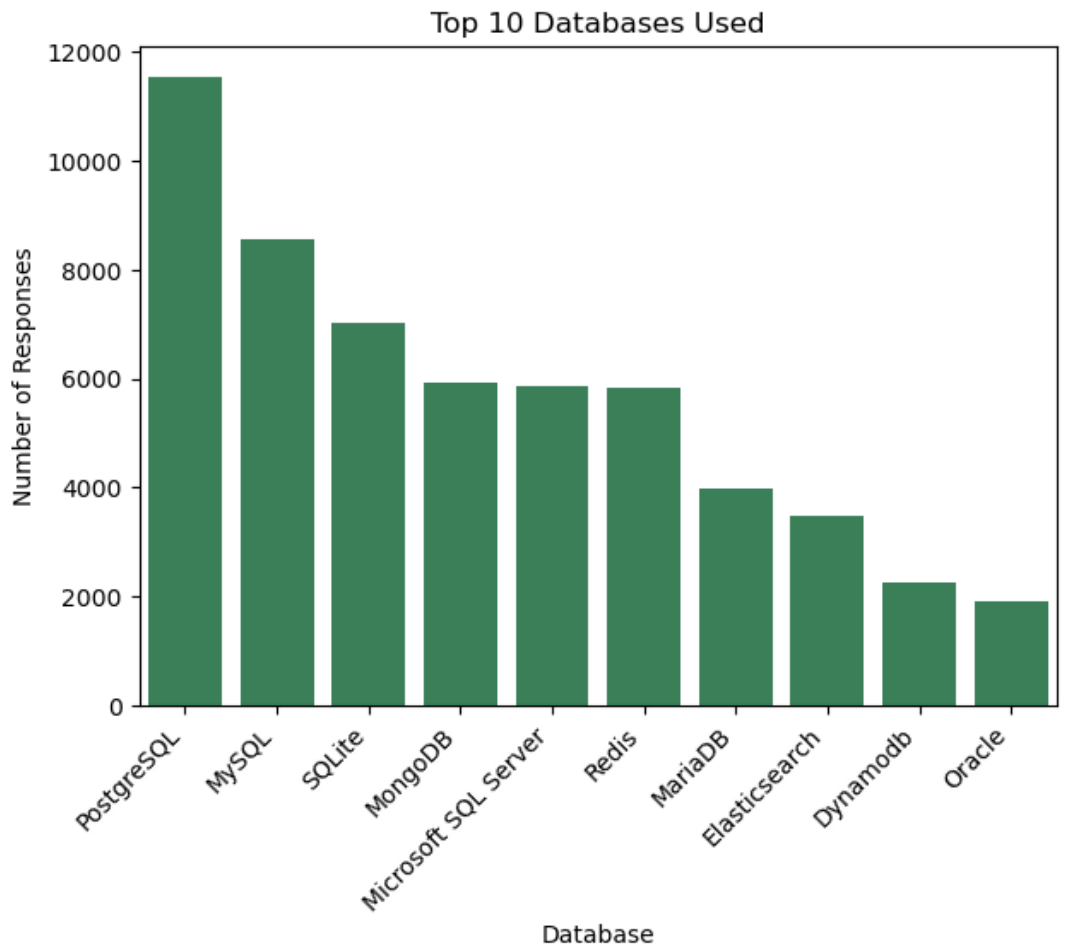
- JavaScript is currently the most popular programming language worked with, according to the Stack Overflow survey. It is also a language that survey respondents *want* to work with the most.
- SQL, HTML/CSS, Typescript, and Python make up the next 4 languages worked with. Typescript goes alongside JavaScript, aiding its popularity. SQL and Python are two common options for analytics solutions. They also work well together.
- The top 10 popular languages are mixed with tools for data analytics, app development, web development, and command-line interfaces

Implications

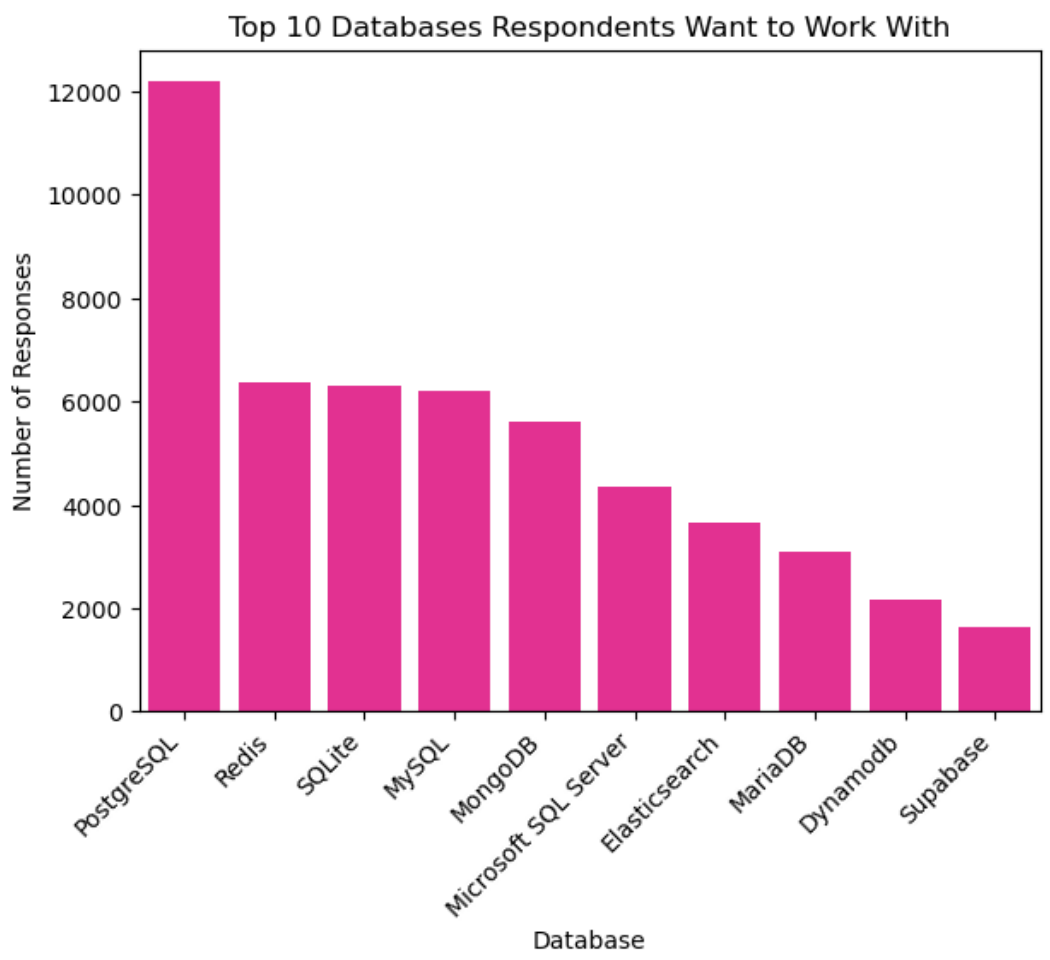
- The top 5 programming languages survey respondents work with are also the top 5 languages they want to work with (though ranked differently). Not only do respondents work with these 5 languages often but they also enjoy working with them.
- Go and Rust are 2 languages in the 'Top-10 admired' list, giving them potential to grow in industry use due to positive user opinions. About half as many people who use admire Javascript want to work with Go and Rust.
- SQL, Python, and HTML/CSS have a large percentage of respondents admiring them as tools, implying they will remain highly popular languages used in the future.

DATABASE TRENDS - 2025

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

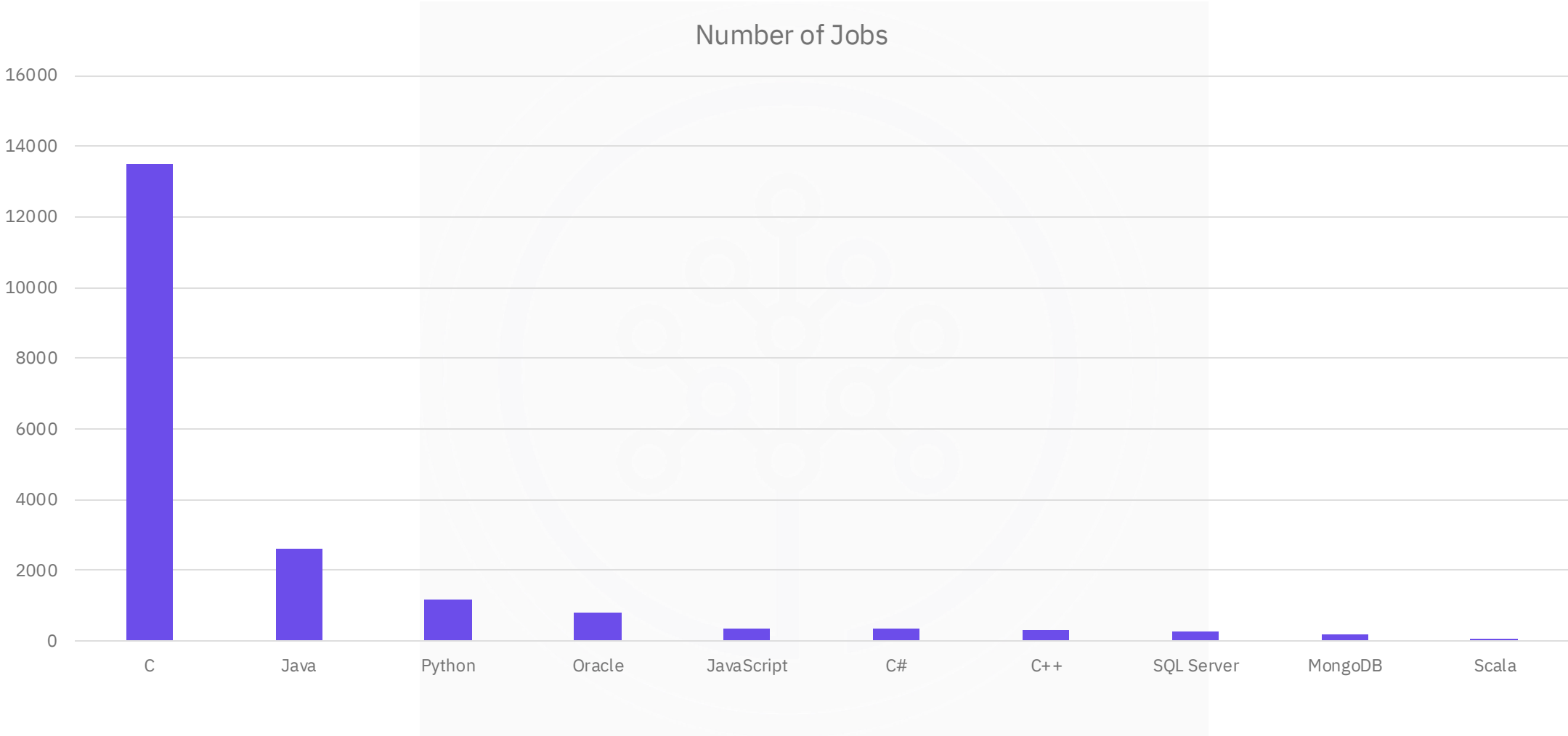
Findings

- PostgreSQL, MySQL, SQLite, MongoDB, Microsoft SQL Server are the top 5 databases used by people surveyed. 4 of the 5 are based on SQL, a highly popular and used programming language. MongoDB is NoSQL.
- Among the top 10 are databases used are ones with AI capabilities, key-value search, relational databases, and joining capabilities.
- PostgreSQL is by far the database respondents want to work with the most. The next 4 are Redis, SQLite, MySQL, and MongoDB.

Implications

- More respondents said they want to work with PostgreSQL than actually use it, implying that it has potential for growth in the workplace.
- Redis is the 2nd most wanted to work with database and 6th most used by respondents. This database is highly likely to mode up in the usage ranks.
- Lightweight and non-lightweight variants of SQL remain both well used and are similar enough that users can switch back and forth. This implies that the usage is likely to stay strong, and given the amount of people that want to use these databases, they will remain popular.

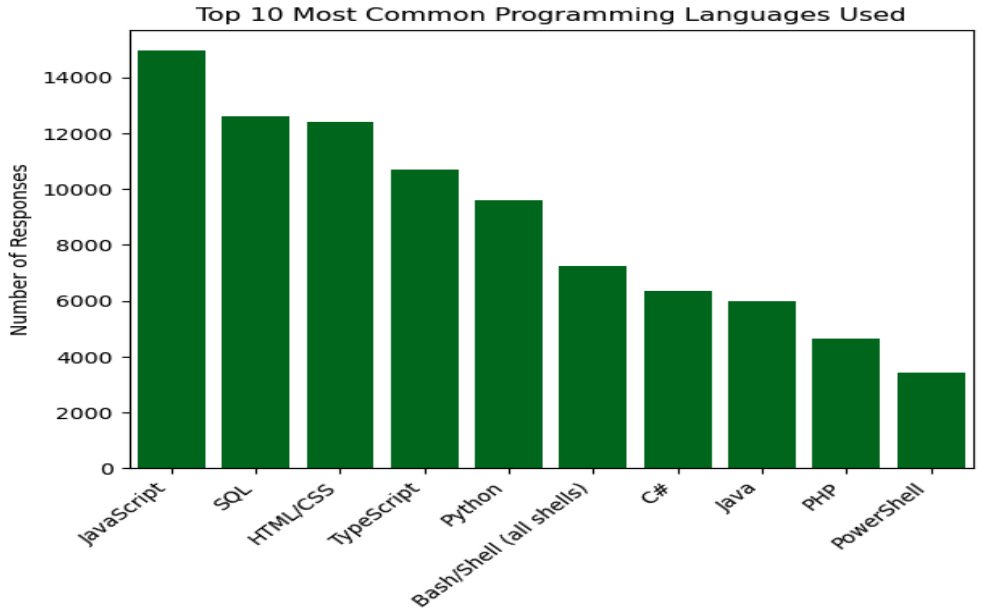
JOB POSTINGS

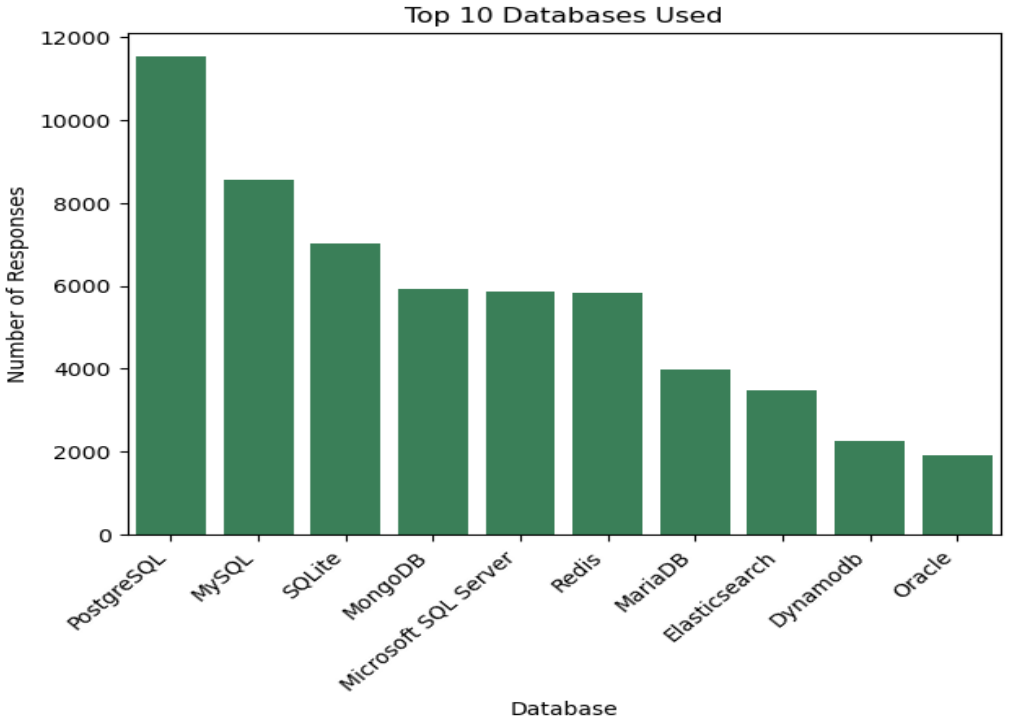


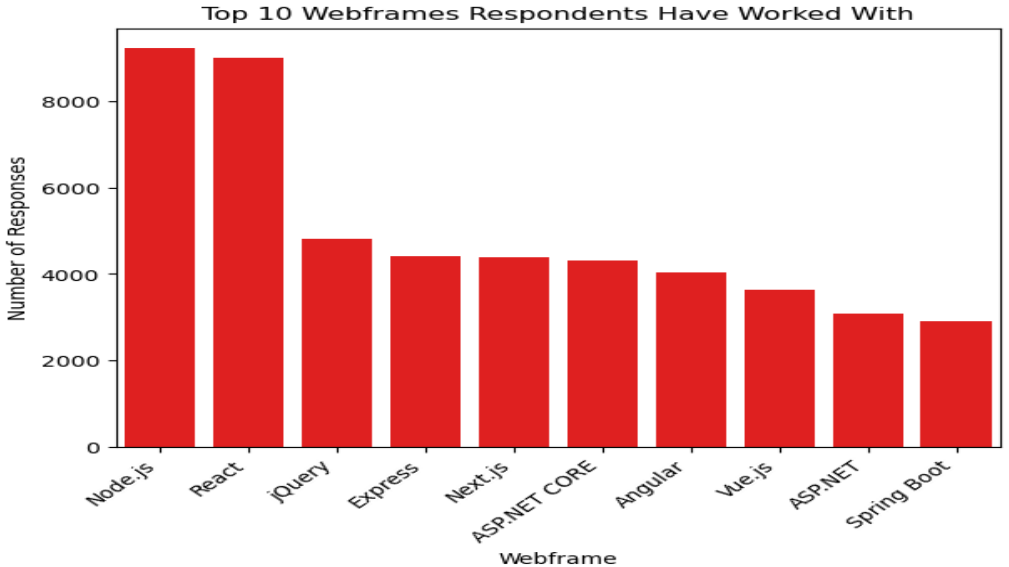
DASHBOARD

The following 3 slides contain dashboards diving further into the data

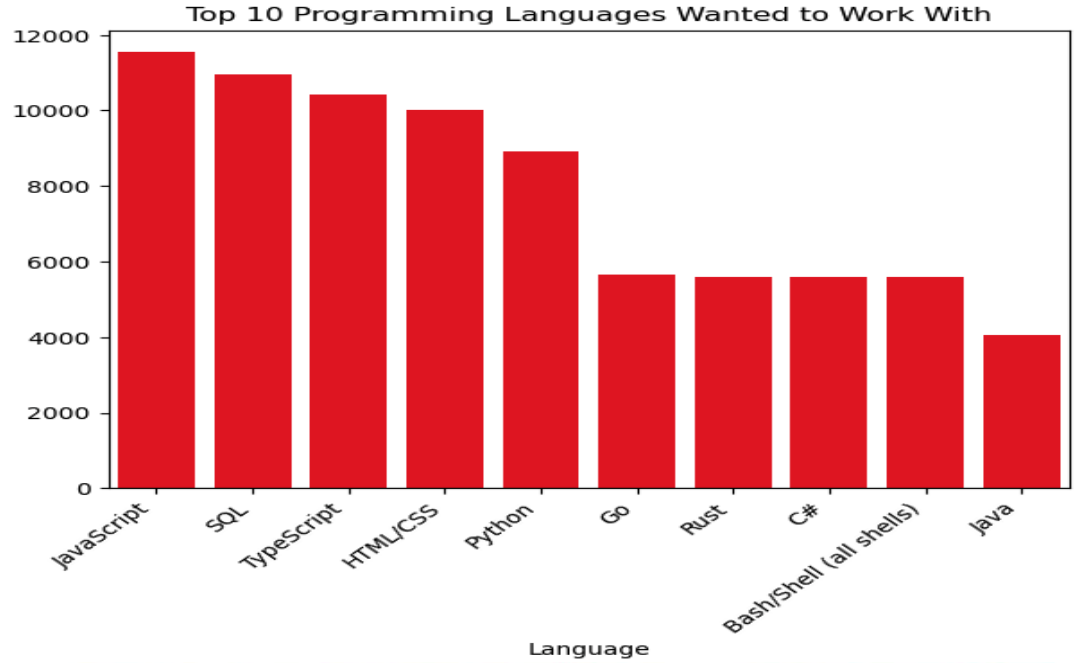
DASHBOARD TAB 1

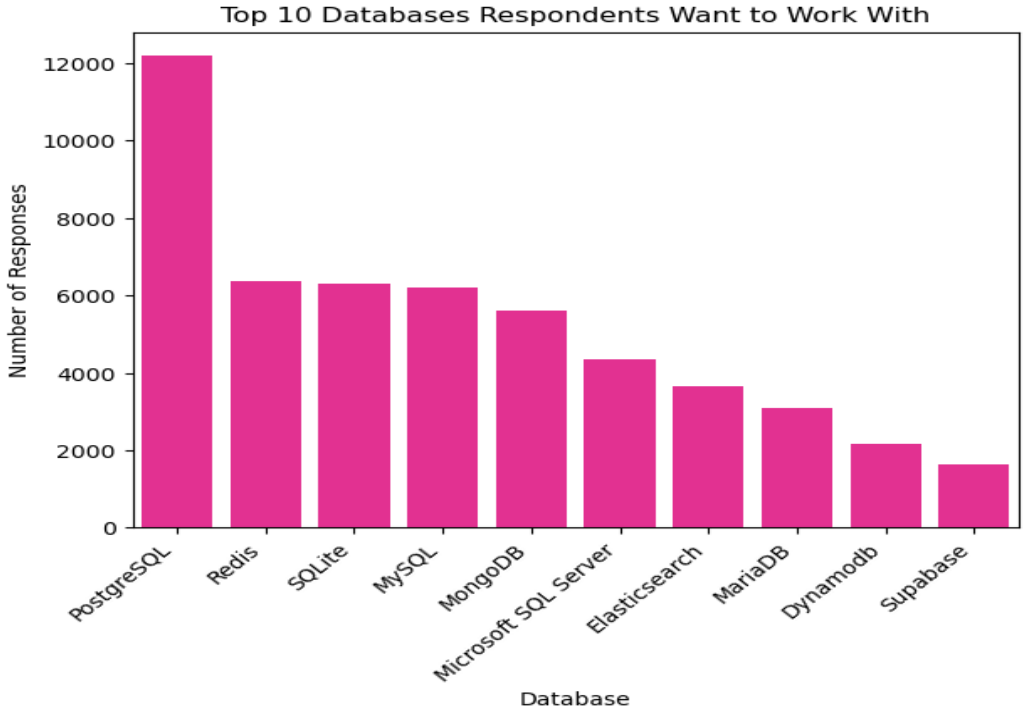


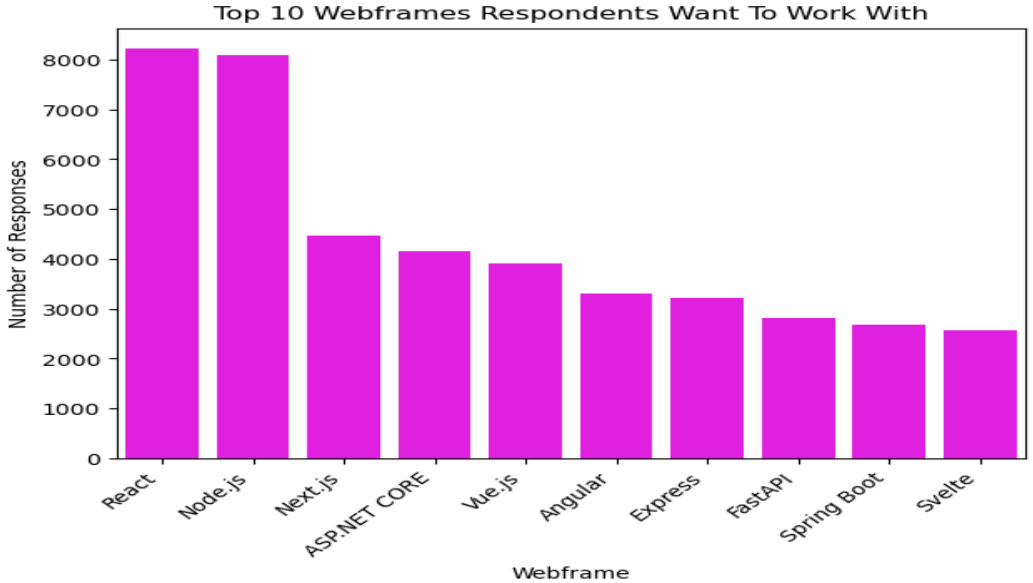




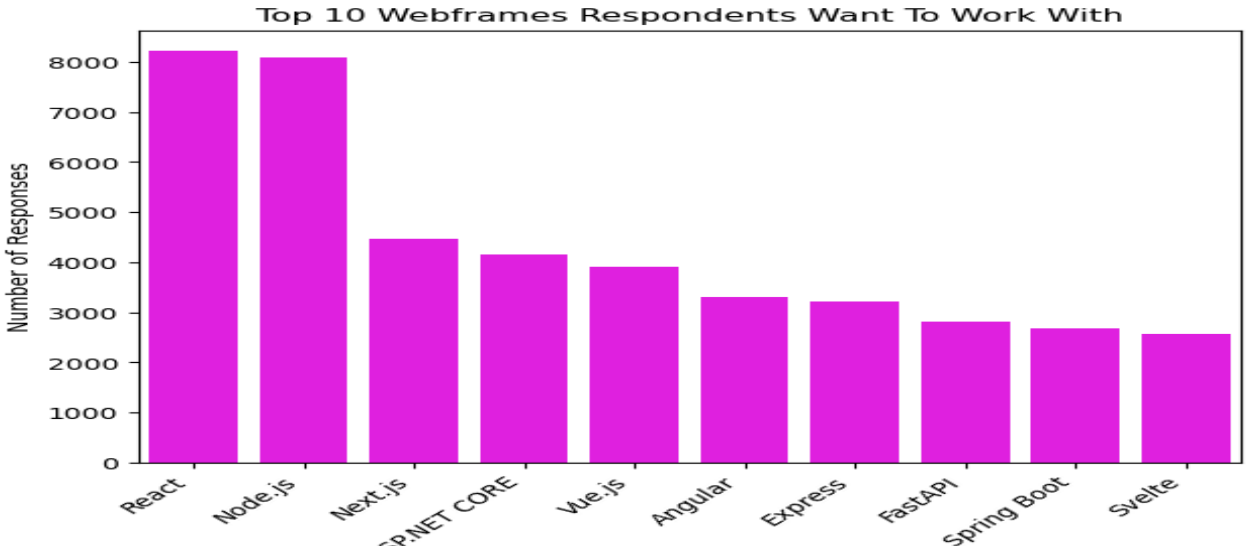
DASHBOARD TAB 2

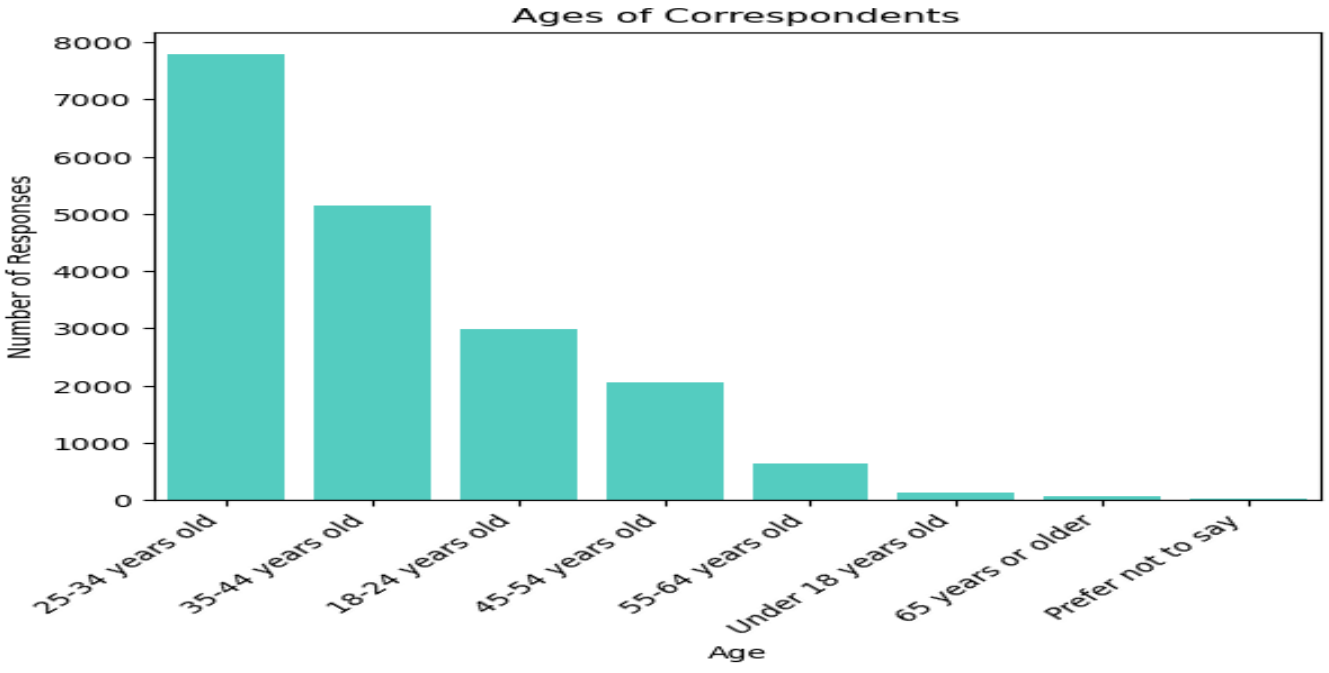


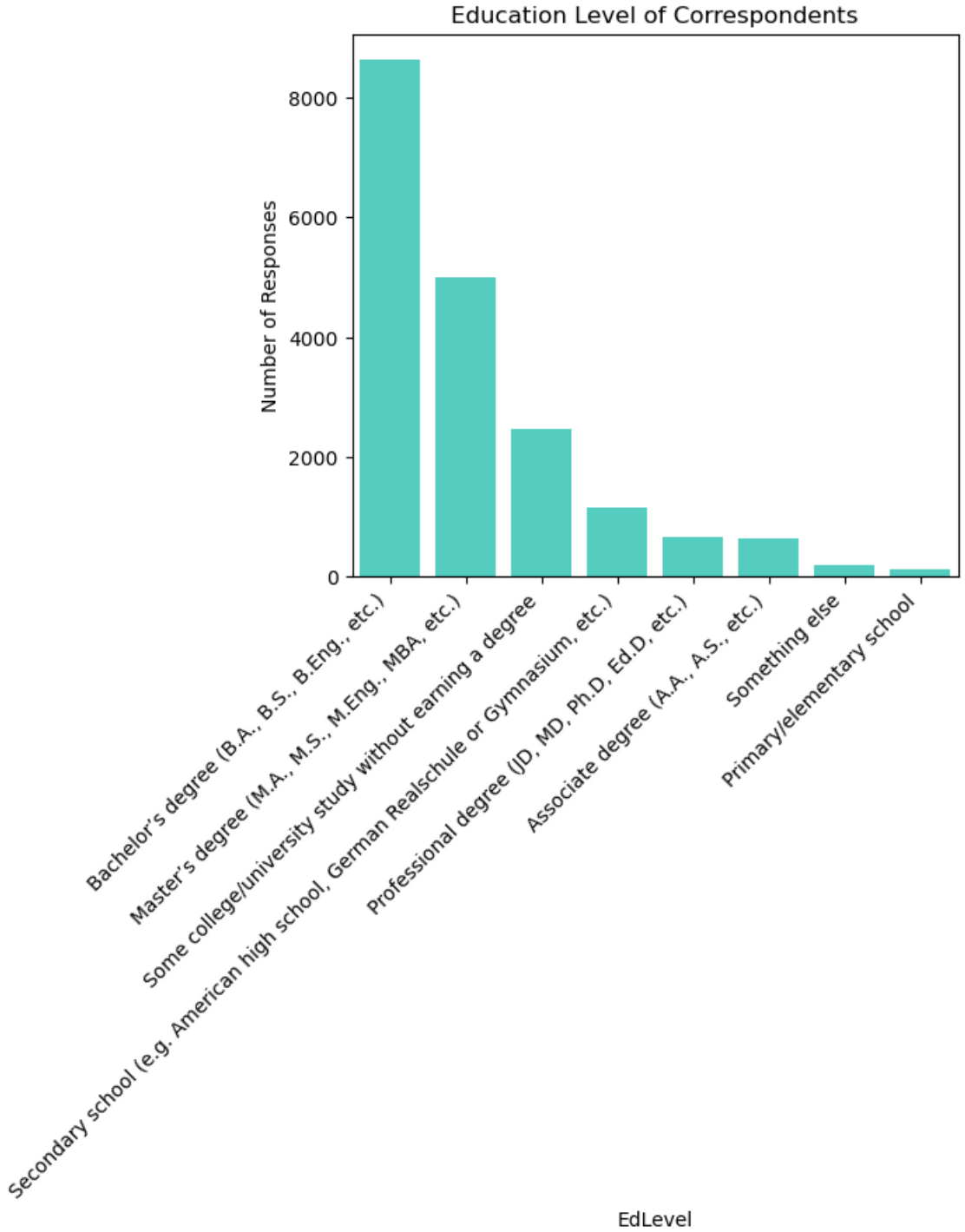




DASHBOARD TAB 3







DISCUSSION

The following slides contain a discussion on overall findings, implications, and colclusions

OVERALL FINDINGS & IMPLICATIONS

Findings

- The majority of survey participants have some level of post-secondary education (most respondents have at least a Bachelor's degree) and fall mostly between the ages of 25-44.
- Javascript, SQL and similar variants, HTML/CSS, and Python are all very common languages used and pair well with the most popular databases used (PostgreSQL, MySQL, SQLite, MongoDB)
- Amazon Web Services and Microsoft Azure are the most common platforms, likely due to their accessibility and widespread availability.

Implications

- SQL and its variants are still widely preferred by users and growing in users. It will likely stay within the most popular languages and database methods.
- Users want compatibility with their programming languages, database choice, webframe, and platform choices. Tools that have high integrability secure future growth, popularity, and positive user opinion.
- Languages, databases, webframes, and platforms that are less robust and lack modern features (such as Ai integration in IDEs or database) or lack community support/documentation are falling out of use.

CONCLUSION

As younger data analysts enter the workforce, a variety of platforms, languages, databases, and webframes are being introduced into the workforce, all with varying pros and cons. Tools with high satisfaction reports tend to be ones with high documentation, are updated frequently with new libraries and AI features, have high integrability with other tools, and are proven to be reliable business solutions.

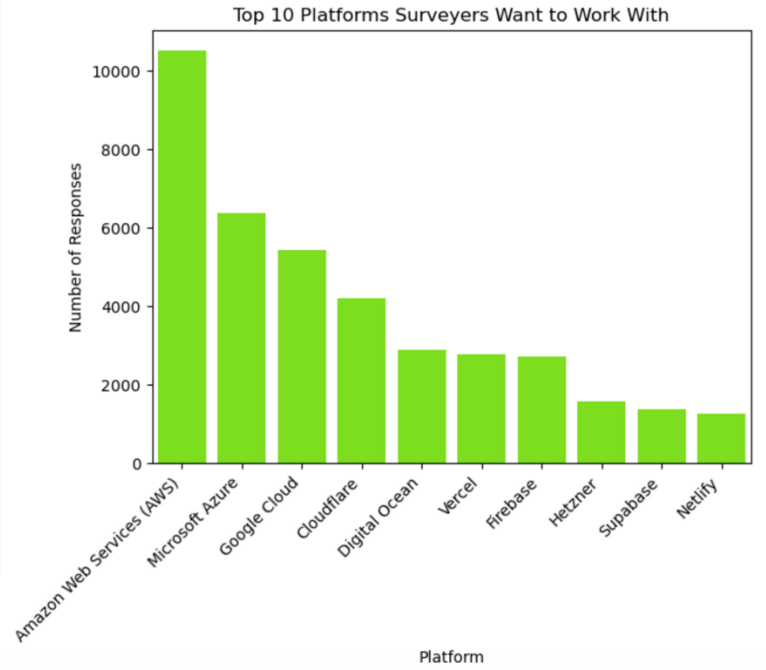
Data Analysts in 2025 are continuing to use different versions of SQL, Python, HTML/CSS, Typescript, SQL databases, Microsoft and AWS platforms, and React and Node.js. They also want to continue using these tools. Any data analyst entering the industry can be assured that these tools will remain popular. There are other options to stand out, as shown in the dashboards, that could also be great options and integrate well with many other tools.

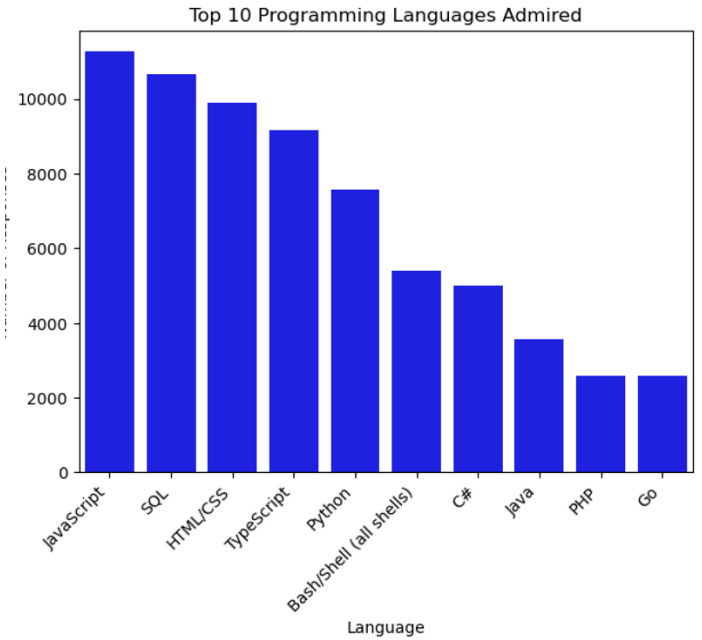
APPENDIX

Information about respondents:

```
Mean years coding: 15.266521230179846
Mean years coding professionally: 10.719962366484031
Median years coding: 13.0
Median years coding professionally: 9.0

Average yearly compensation: 84589.6967539267
Median yearly compensation: 65857.5
```





POPULAR LANGUAGES