



Analysis of Stack Overflow Developer Survey Data: Implications for Business Strategy and Skills Development.

Svetlana SEDYKH

March 1, 2023

OUTLINE



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

EXECUTIVE SUMMARY



- ▶ This report highlights some important findings regarding programming languages, databases, platforms and web frameworks.
- ▶ In terms of programming languages, TypeScript and Python are becoming more and more popular among developers.
- ▶ When it comes to databases, PostgreSQL, MongoDB, Redis, and Elasticsearch are of more interest to developers than other databases.
- ▶ Linux, Docker and AWS remain popular platforms with developers, while Windows popularity has declined.
- ▶ Interest in jQuery has declined in favor of React.js and Vue.js web frameworks.
- ▶ These findings have important implications for the firm. To remain competitive in the market, the firm must invest in developing consultant skills in TypeScript, Python, PostgreSQL, MongoDB, Redis, Elasticsearch, Linux, Docker, and AWS.
- ▶ Additionally, the firm should prioritize delivering solutions using popular web frameworks such as React.js and Vue.js over outdated frameworks such as jQuery.
- ▶ The declining popularity of Windows among developers suggests that the firm should rethink its approach to building solutions for the Windows platform and shift its focus to other platforms, such as Linux or AWS.
- ▶ These steps will help the firm stay on top of current trends and provide high-quality services to the clients.

INTRODUCTION



- ▶ The purpose of this report is to analyze current trends and identify future trends for emerging technologies.
- ▶ The report will focus on in-demand technology skills, data will be collected from various sources, including job postings, blog posts, and surveys.
- ▶ The collected data will be prepared for analysis using data processing techniques, and statistical methods will be applied to identify insights and trends.
- ▶ Specifically, the report will look into which programming languages and database skills that are most in demand, the popularity of integrated development environments (IDEs), and demographic data like gender and age distribution of developers.
- ▶ To aid in the analysis, IBM Cognos Analytics will be used to create interactive dashboard for dynamic data presentation. The findings of this analysis will be presented in the last part of the report.

METHODOLOGY



- ▶ A subset² of the Stack Overflow Developer Survey data¹ in the form of CSV file was used for analysis in this report. The dataset was loaded into a Jupyter Notebook and explored using the Pandas Python Library.
- ▶ Upon analyzing the data and pinpointing problems related to both data quality and completeness, the dataset was wrangled to eliminate duplicate values, impute or remove missing values, and normalize data in the dataset. A new column called 'NormalizedAnnualCompensation' was added to the dataset to make comparison of developers' salaries easier.
- ▶ Exploratory data analysis was performed on a cleaned dataset³ to find out the distribution of data, presence of outliers and to determine the correlation between different columns in the dataset.
- ▶ Different visualizations were created to highlight the distribution of data, relationships between data, the composition of data, and the comparison of data. The dataset⁴ was presented in the form of a RDBMS and SQL queries were used to extract the data.
- ▶ Then a dashboard was created using IBM Cognos Analytics. Two CSV files: *m5_survey_data_demographics.csv*⁵ and *m5_survey_data_technologies_normalised.csv*⁶ were uploaded as data assets into IBM Cognos Analytics. Current Technology Usage, Future Technology Trends and Demographics slides were created with Cognos Dashboard Embedded (CDE).
- ▶ Appendix
- ▶ To determine the number of jobs currently open for various technologies, the GitHub Jobs API⁷ was used. The API was called for all the given technologies and the results were written in an Excel spreadsheet⁸.
- ▶ To get the names of popular programming languages and average annual salaries, web scraping was used to extract information from the website⁹ and the result was saved in CSV file¹⁰.

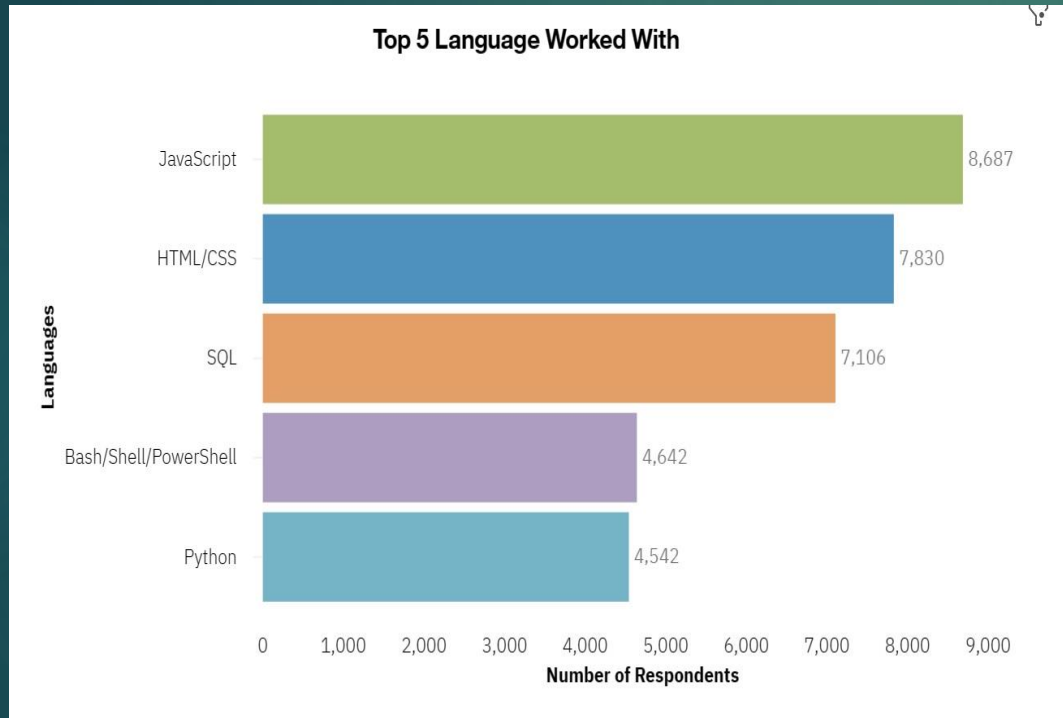
Results



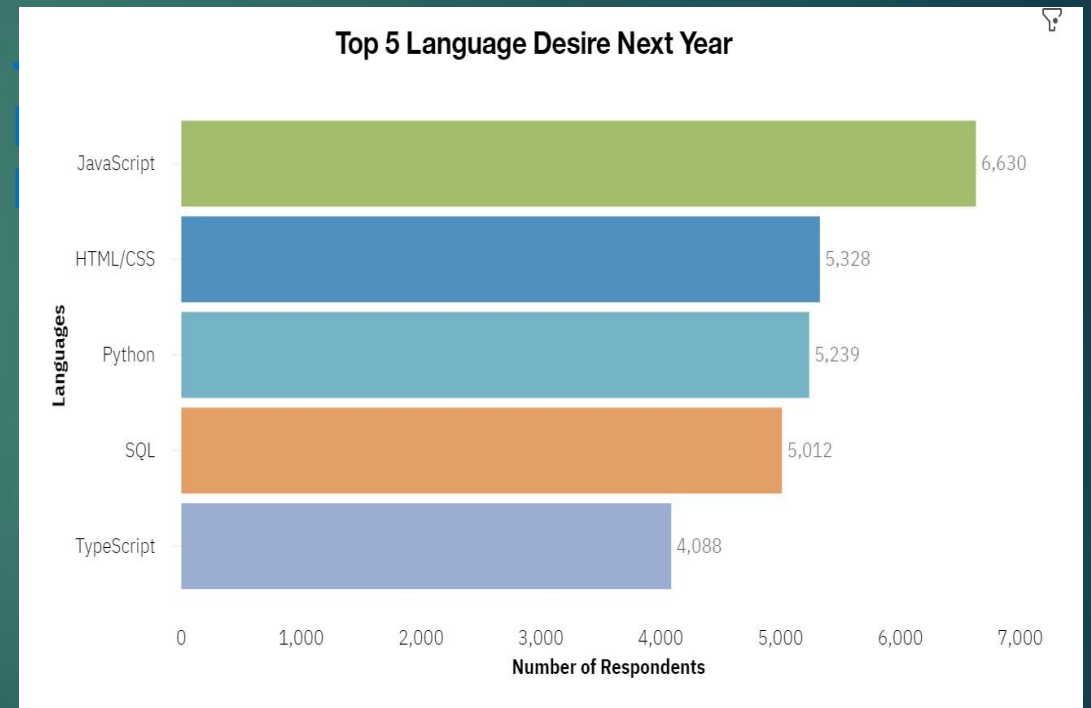
In the following slides you will find the results of this analysis.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

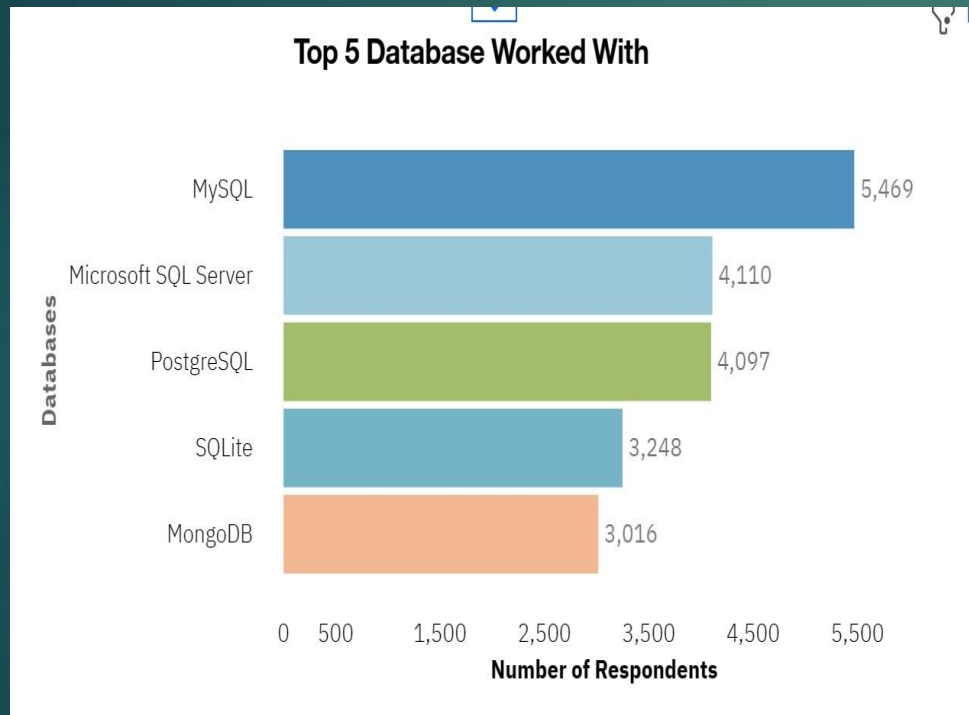
- ▶ Our analysis showed that among the programming languages respondents gave the highest preference to JavaScript, followed by HTML/CSS. This trend is expected to continue in the coming years.
- ▶ In addition, we have seen a growing interest among developers in Python and TypeScript.
- ▶ However, we found that SQL and Bash/Shell/PowerShell have seen a decrease in interest among developers.

Implications

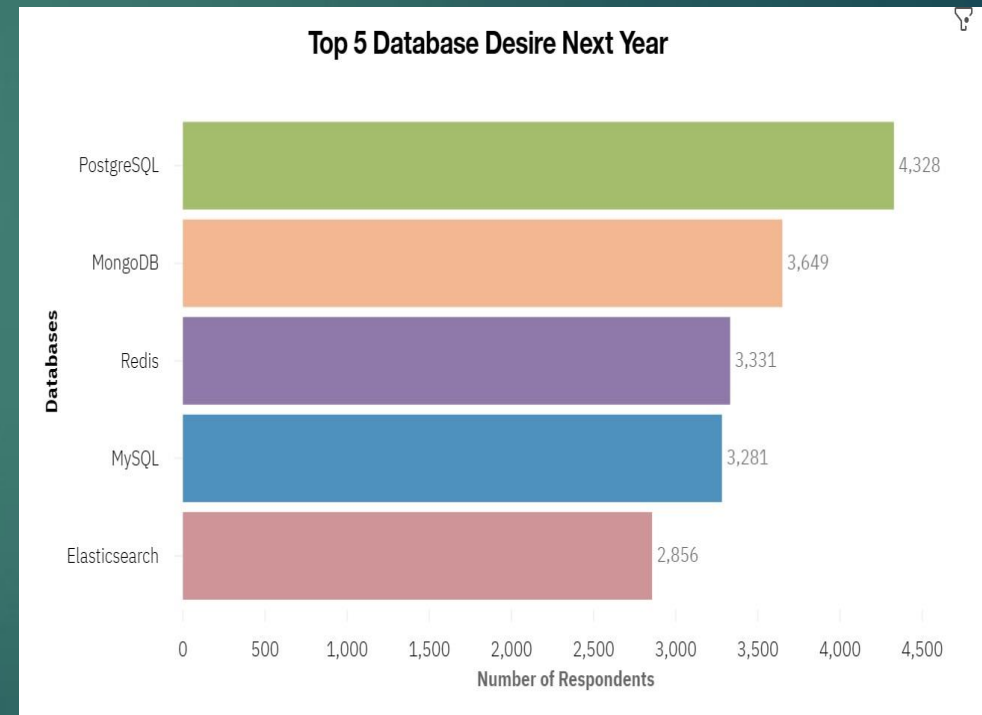
- ▶ The firm should focus on hiring and training developers with extensive JavaScript and HTML/CSS experience, investing in new trends in these languages, and focusing on these languages in the offered services.
- ▶ The firm should consider expanding its expertise in Python and TypeScript, that could include hiring developers with extensive experience in these languages and investing in training existing developers.
- ▶ The firm should consider reallocating resources away from SQL and Bash/Shell/PowerShell and focus on new technologies that are in high demand by developers. This could include reducing investments or downsizing teams focused on these technologies.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- ▶ Our analysis indicates that MySQL is currently the most commonly used database, followed by Microsoft SQL Server and PostgreSQL. However, we have seen a decline in the popularity of MySQL, Microsoft SQL Server and SQLite.
- ▶ On the other hand, we found an increase in interest for PostgreSQL, MongoDB, Redis, and Elasticsearch.

Implications

- ▶ The firm may have to shift its focus from MySQL, Microsoft SQL Server and SQLite to other databases that are in high demand by respondents.
- ▶ As PostgreSQL grows in popularity, the firm must be prepared to provide clients with solutions that use this database technology. Growing interest in MongoDB, Redis and Elasticsearch will require investment in training and development for the consultants to master these new database technologies.

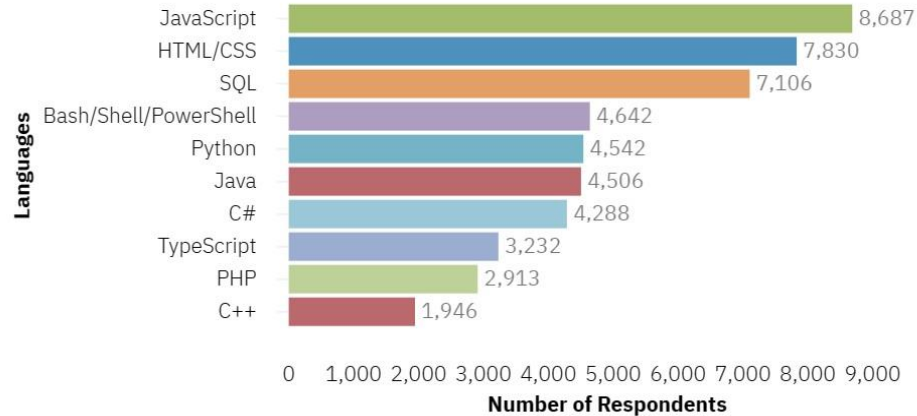
DASHBOARD



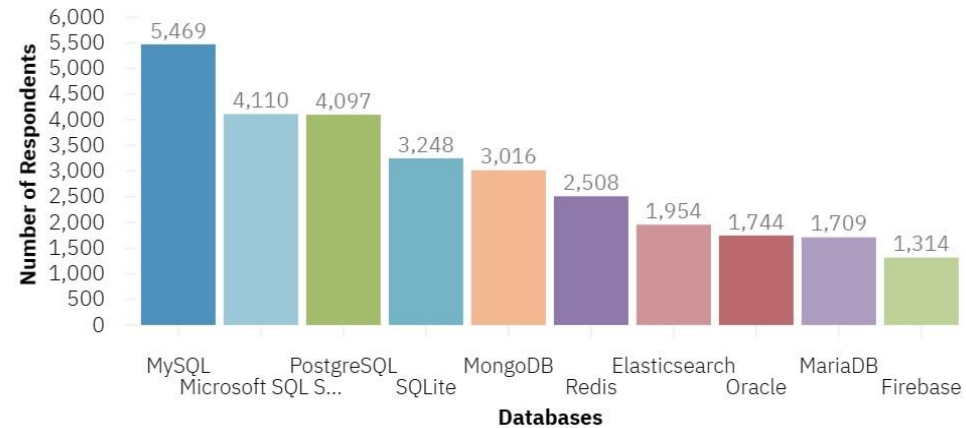
To view IBM Cognos Dashboard click [here](#)

Current Technology Usage

Top 10 Language Worked With



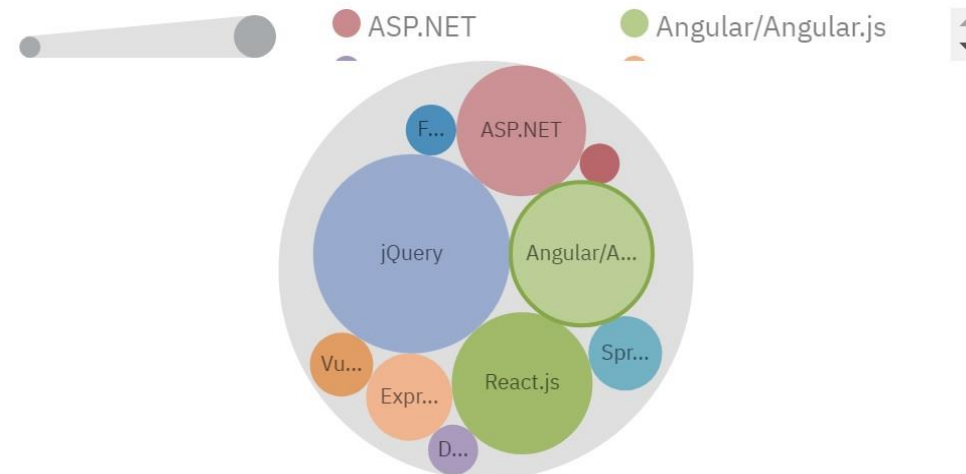
Top 10 Database Worked With



Platform Worked With

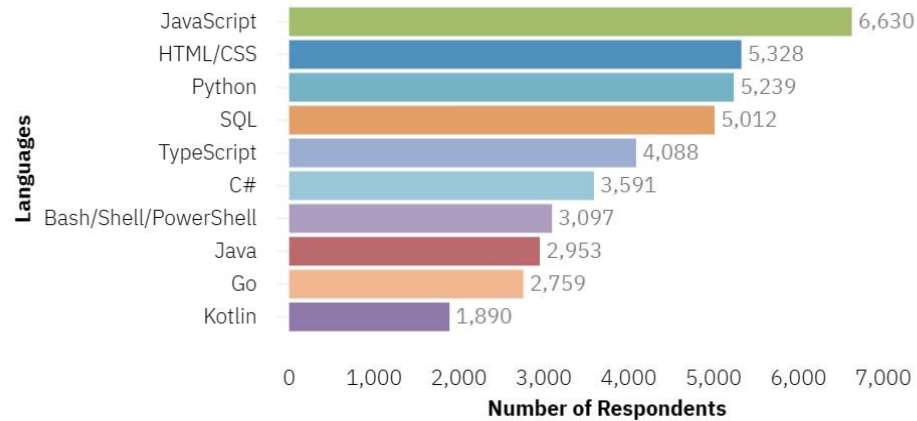


Top 10 Web Frame Worked With

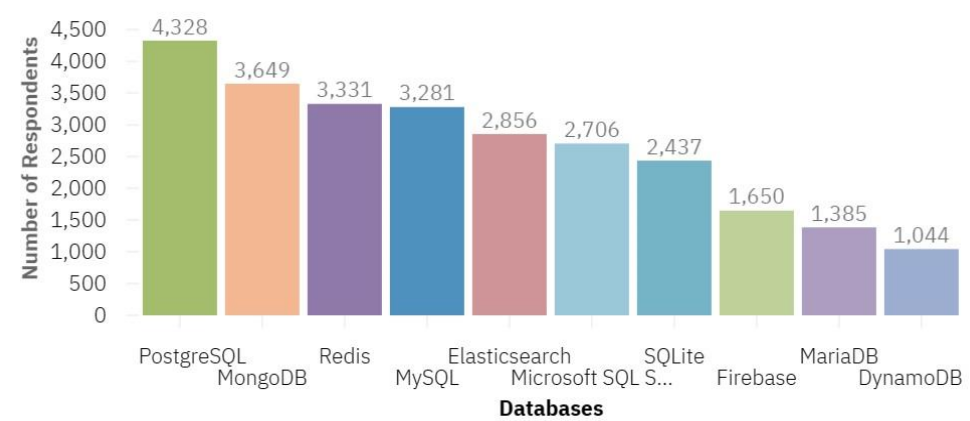


Future Technology Trends

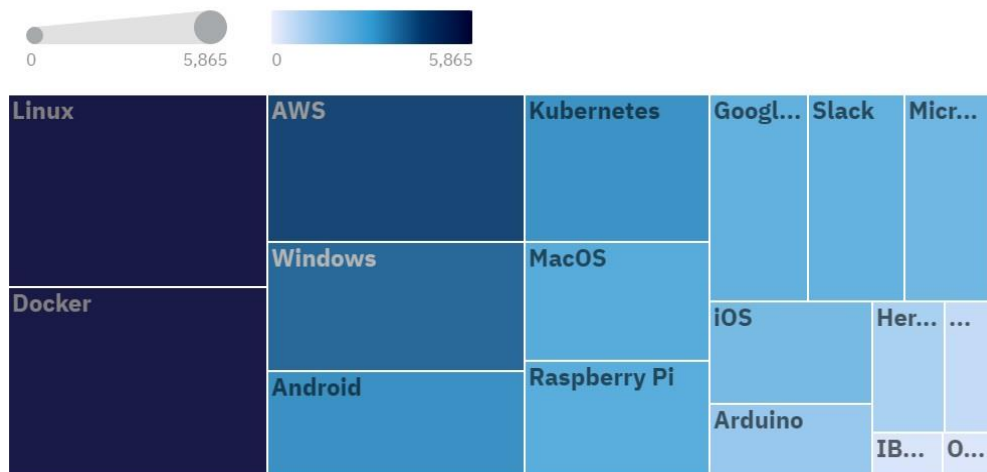
Top 10 Language Desire Next Year



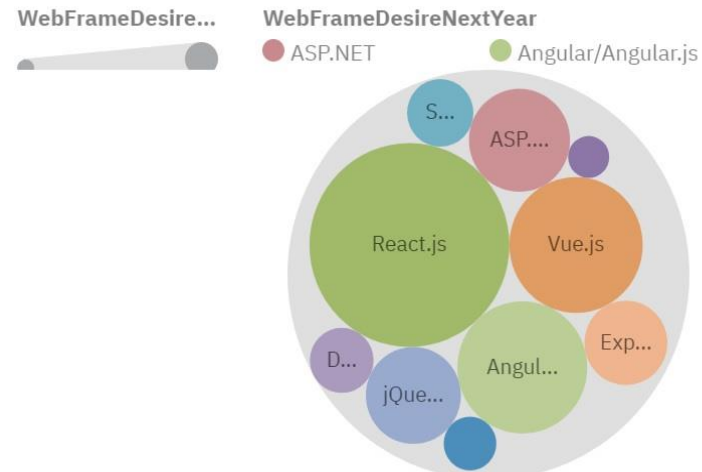
Top 10 Database Desire Next Year



Platform Desire Next Year



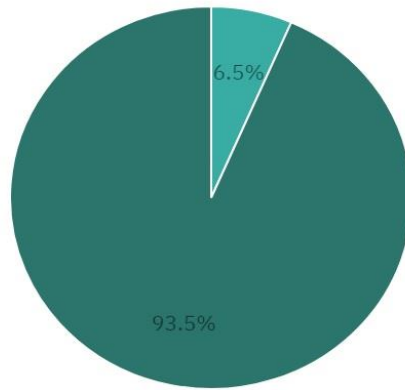
Top 10 Web Frame Desire Next Year



Demographics

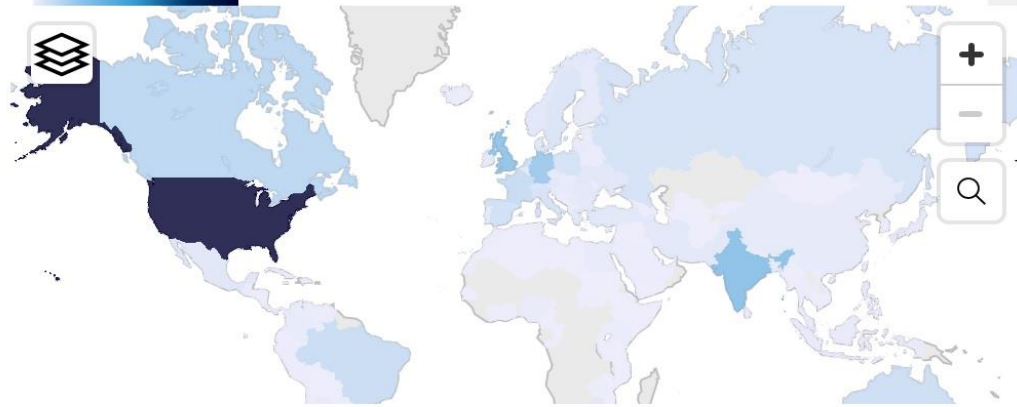
Respondent classified by Gender

Woman Man

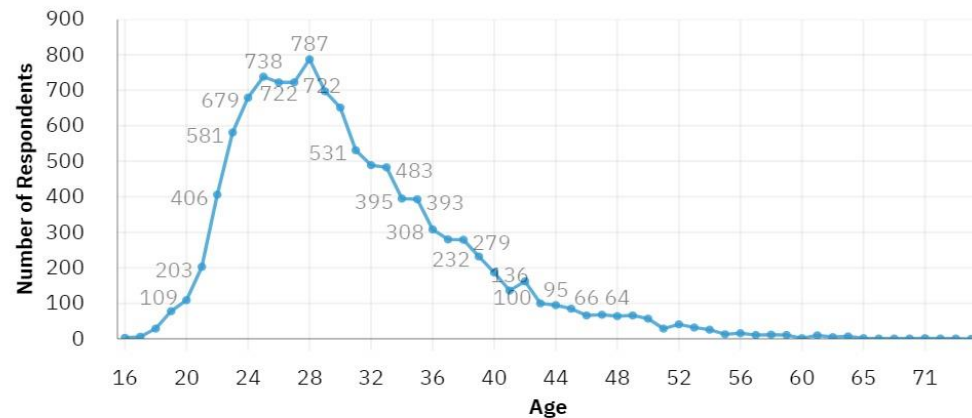


Respondent Count for Countries

Country (Count)



Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level

EdLevel

Associate degree Bachelor's degree (BA, BS, B.E...
I never completed any formal ... Master's degree (MA, MS, M.E...
Other doctoral degree (Ph.D, E... Primary/elementary school



DISCUSSION



- ▶ In the following slides you will find the overall finding of this analysis and the implications for the firm.

OVERALL FINDINGS & IMPLICATIONS

Findings

- ▶ Our analysis revealed some key findings in the areas of programming languages, databases, platforms, and web frameworks.
- ▶ Firstly, we observed a rise in popularity for TypeScript and Python among developers.
- ▶ Secondly, there was an increase in interest for databases such as PostgreSQL, MongoDB, Redis, and Elasticsearch.
- ▶ Thirdly, the interest in jQuery has significantly decreased, whereas React.js and Vue.js have gained popularity among web developers.
- ▶ Finally, we found that Linux, Docker, and AWS remained popular while the popularity of Windows has decreased among developers.

Implications

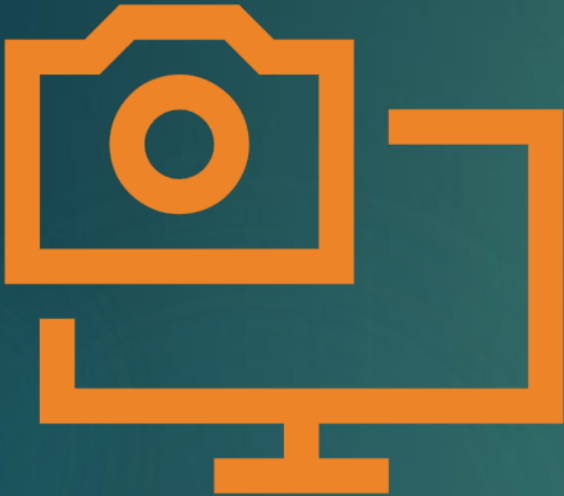
- ▶ These findings have important implications for the firm. To remain competitive in the marketplace, the firm should invest in developing consultant skills in TypeScript, Python, PostgreSQL, MongoDB, Redis, Elasticsearch, Linux, Docker, and AWS.
- ▶ The firm should also focus on providing solutions using popular web frameworks such as React.js and Vue.js rather than outdated frameworks such as jQuery.
- ▶ Moreover, the decline in the popularity of Windows among developers suggests that the company may have to rethink its approach to creating solutions for the Windows platform. It may have to shift its focus to other platforms, such as Linux or cloud solutions like AWS.

CONCLUSION



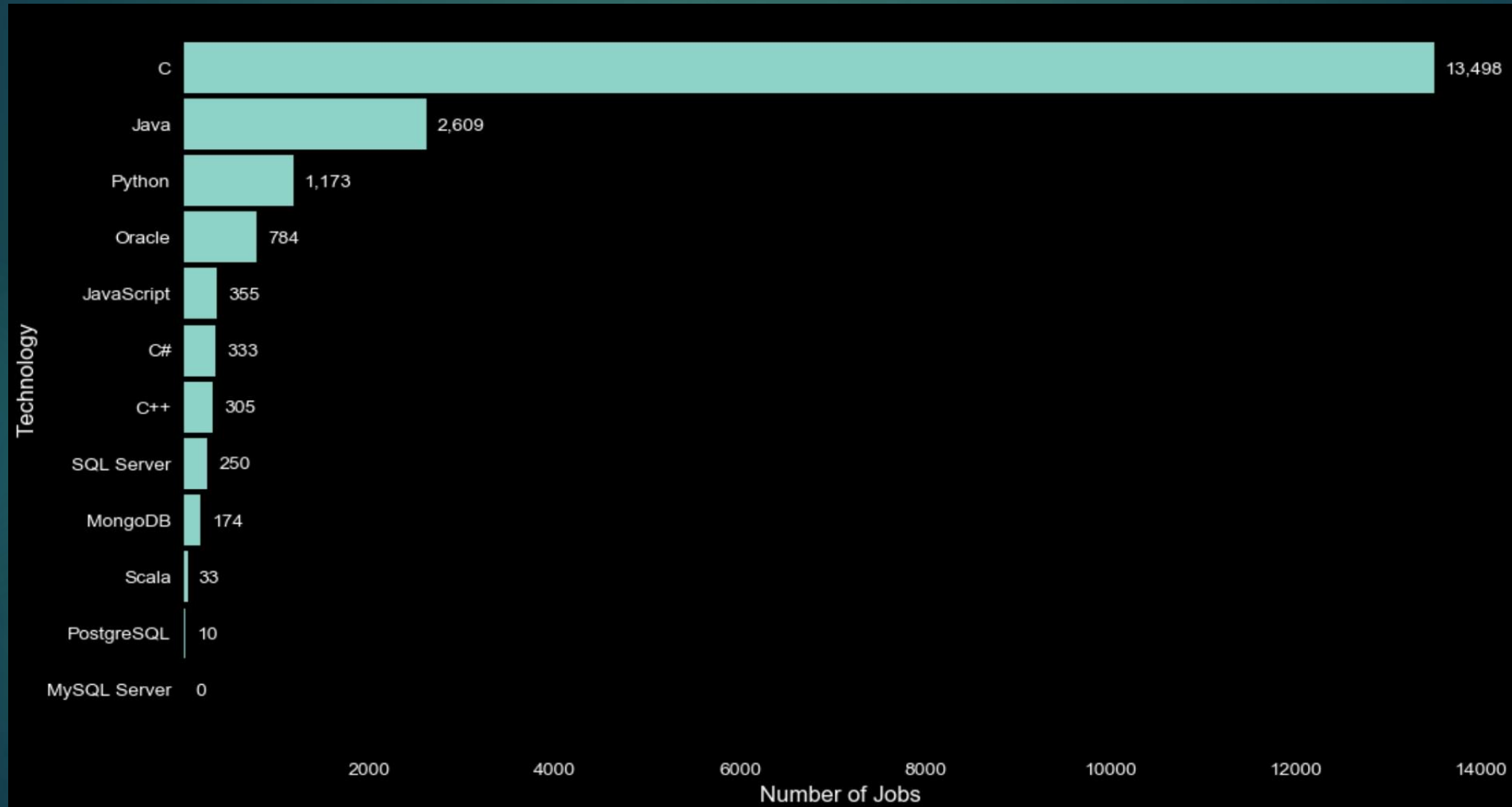
- ▶ Our analysis reveals current trends and developer preferences in programming languages, databases, platforms and web frameworks.
- ▶ The popularity of TypeScript and Python, interest in databases such as PostgreSQL, MongoDB, Redis and Elasticsearch, and the continuing popularity of Linux, Docker and AWS have important implications for the firm. To remain competitive, the firm should invest in developing consultant skills in these areas and focus on popular Web frameworks such as React.js and Vue.js.
- ▶ The firm may need to shift its approach for building solutions on Windows platform to other platforms.
- ▶ Overall, staying updated and investing in skills development is crucial for the firm to remain competitive and drive growth.

APPENDIX

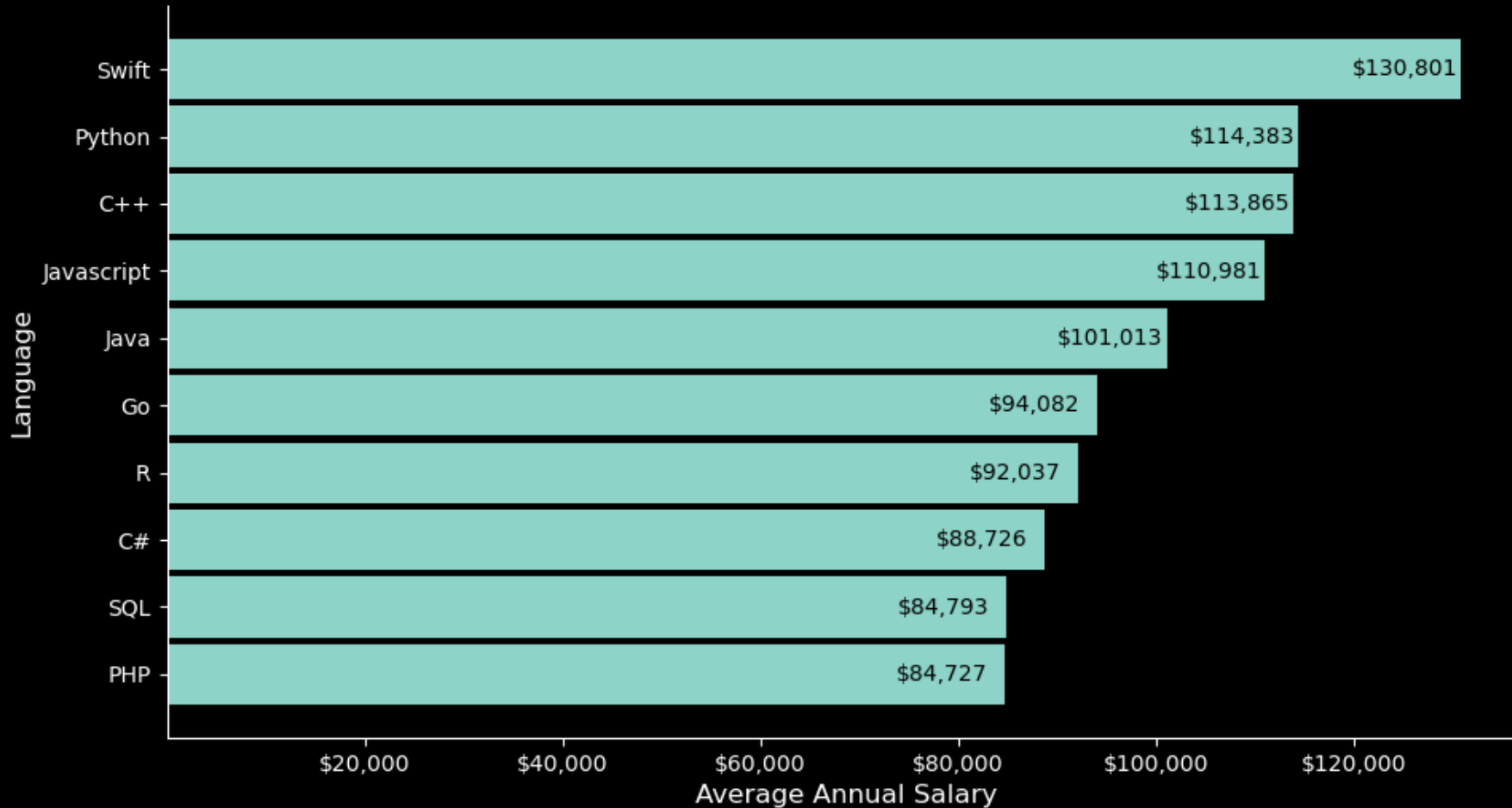


- ▶ GitHub Job Postings Chart
- ▶ Popular Programming Languages Chart
- ▶ Additional Information (Age Distribution and Converted Compensation Charts)
- ▶ Resources

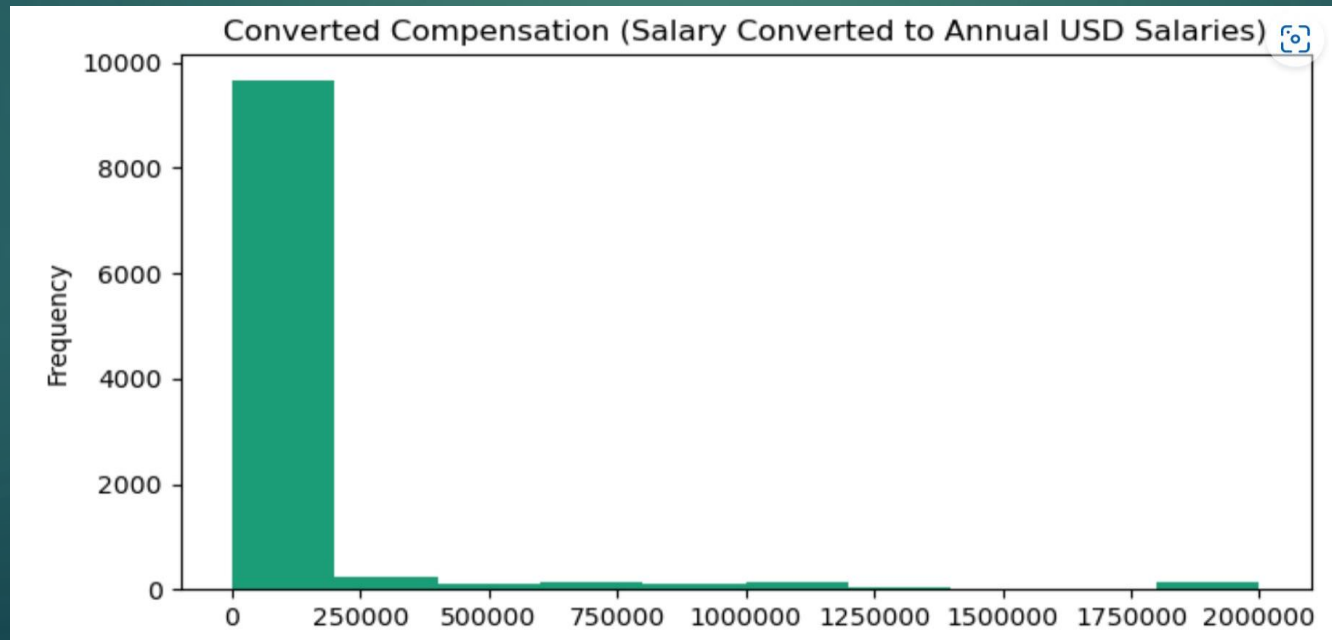
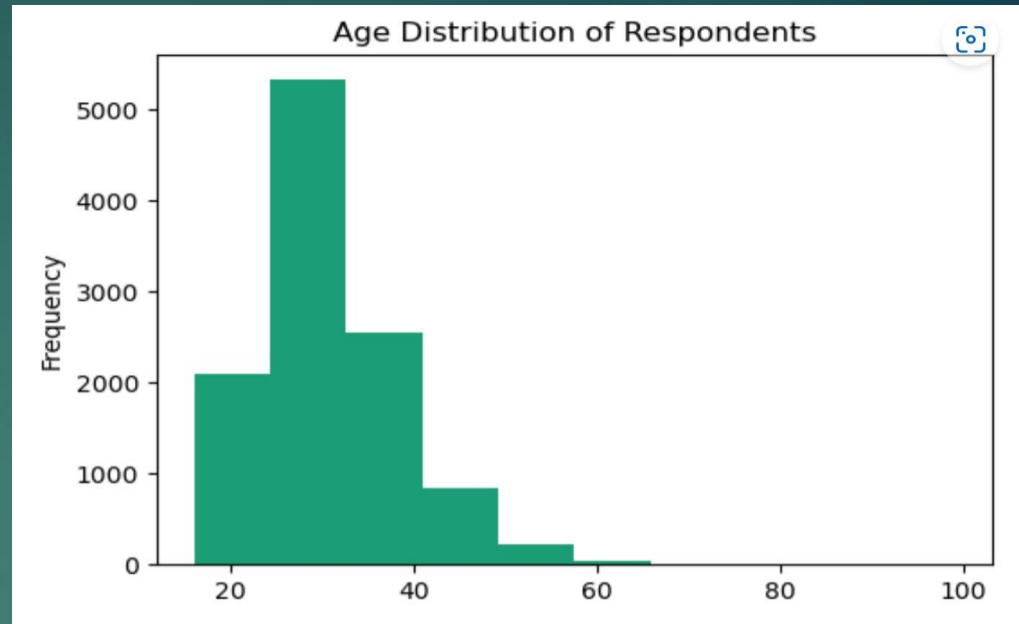
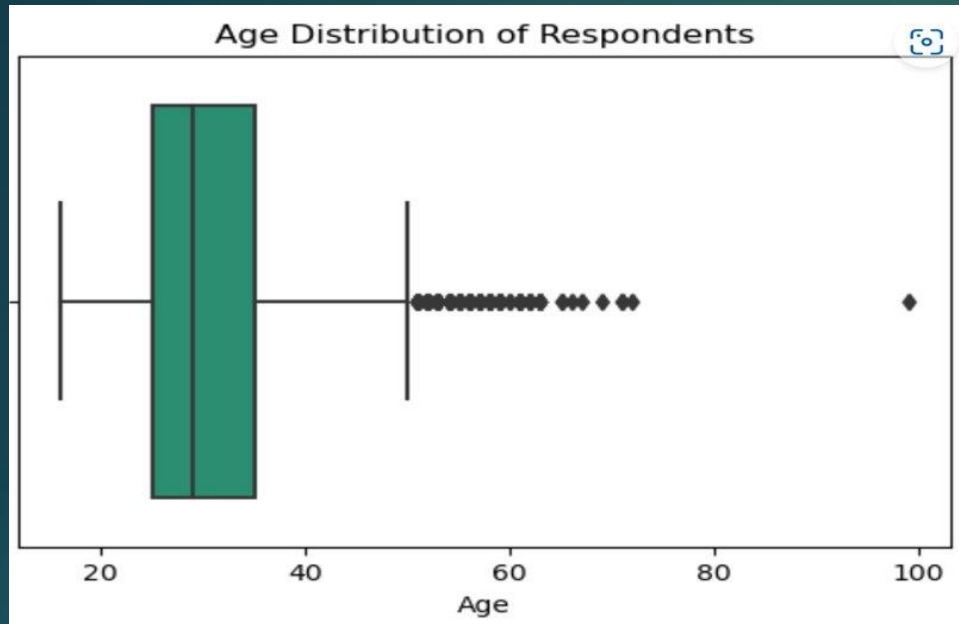
GITHUB JOB POSTINGS



POPULAR PROGRAMMING LANGUAGES



Additional Information



Resources



1. [Stack Overflow Developer Survey Results](#)
2. [m1_survey_data.csv](#)
3. [m2_survey_data.csv](#)
4. [m4_survey_data.sqlite](#)
5. [m5_survey_data_demographics.csv](#)
6. [m5_survey_data_technologies_normalised.csv](#)
7. [GitHub Job Postings](#)
8. [job-postings_tech.xlsx](#)
9. [Popular Programming Languages](#)
10. [popular-languages.csv](#)