

IBM Data Analyst Capstone Project

Andrea Giada Bruzzano
April 14, 2025



© IBM Corporation. All rights reserved.

OUTLINE



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

EXECUTIVE SUMMARY



- **Core Web Technologies & SQL Remain Foundational:** JavaScript, HTML, CSS are dominant; SQL is essential for data interaction.
- **Python's Ascendancy:** Python is the most desired language.
- **PostgreSQL's Rising Popularity in Databases:** While MySQL is currently top, PostgreSQL shows the strongest future desire. NoSQL and cloud databases are increasingly important.
- **Key Shifts:** Increased adoption of TypeScript, Go, Kotlin; growing preference for PostgreSQL; cloud database influence.
- **Implications:** Continuous learning in modern languages and cloud technologies is crucial. Data skills remain highly valuable.



INTRODUCTION



In this project I take the role of data analyst for a technology company, with the task of collecting data, cleaning them and preparing them for a presentation through the use of dashboards via Google Looker Studio.

The objective is to compare trends in the use of current databases and languages with those of the following year, highlighting the differences of these results.

Target audience:

Executives/Strategic planners;
Engineering Managers/Team Leads;
Recruiters/HR professionals



METHODOLOGY



- Collecting data and exploring its content through:
 - Web scraping
 - APIs
 - Request library
- Data wrangling
 - Finding and removing duplicates
 - Finding and imputing missing values
 - Normalizing data
- Exploring the dataset
 - Analyzing data distribution
 - Handling outliers
 - Finding correlations
- Data visualization libraries
 - Altair
 - Matplotlib
 - Seaborn
- Building dashboards using Google Looker Studio



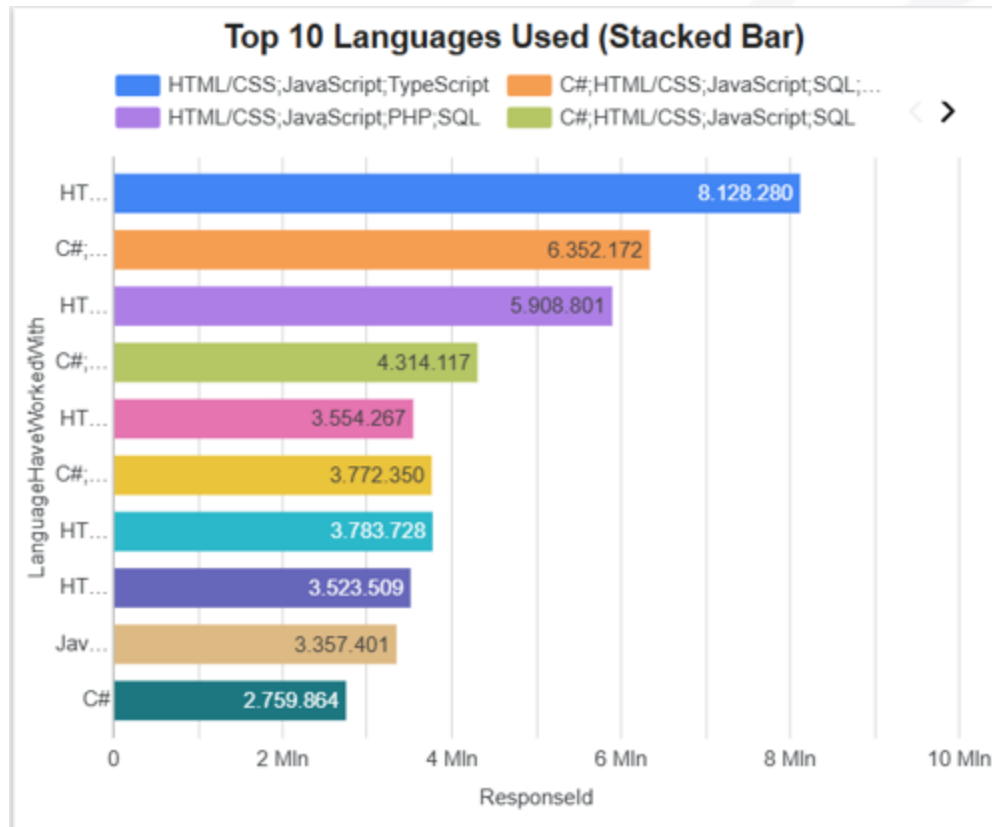
RESULTS

- **Programming Language Trends** (current & future year)
 - Findings and its Implications
- **Database Trends** (current & future year)
 - Findings and its Implications

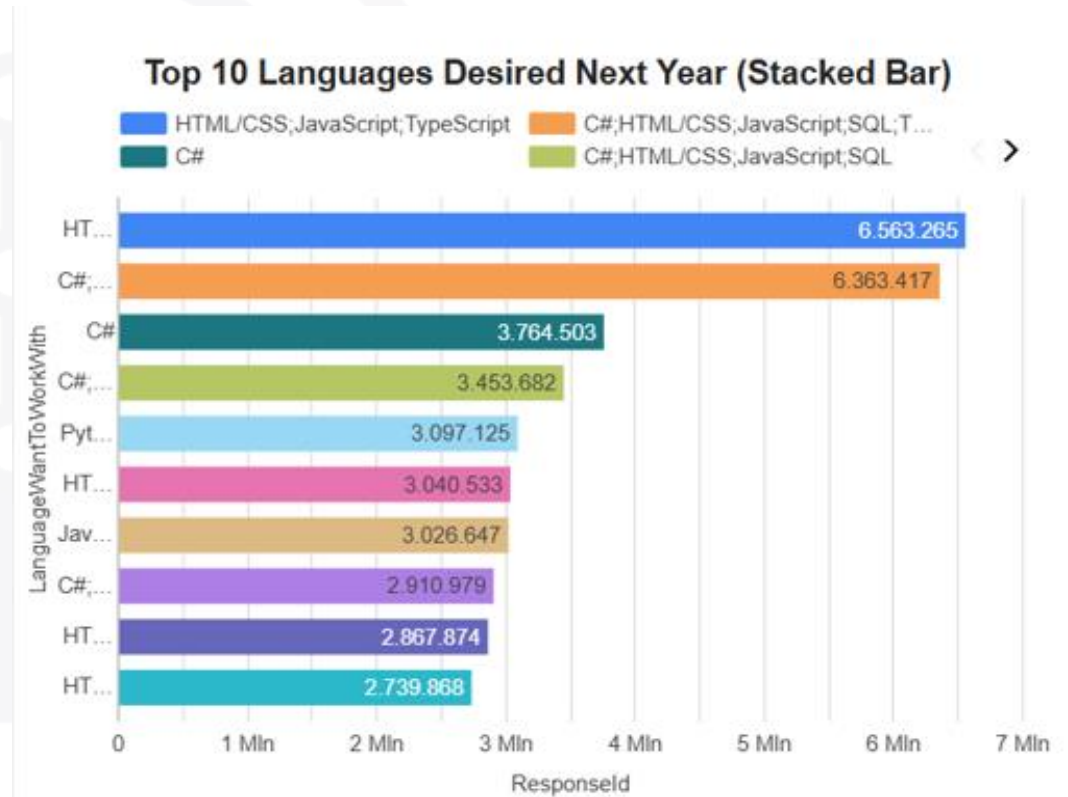


PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- **JavaScript** is overwhelmingly the most used programming language. This highlights its central role in front-end web development and its expanding use in back-end and other areas.
- **SQL** is highly prevalent, indicating the fundamental need to interact with databases in most software applications.
- **HTML/CSS** are also at the top, which is expected as they are foundational for web development alongside JavaScript.
- **Python** is the most desired language for the next year. This suggests a strong future growth and interest in its capabilities.

Implications

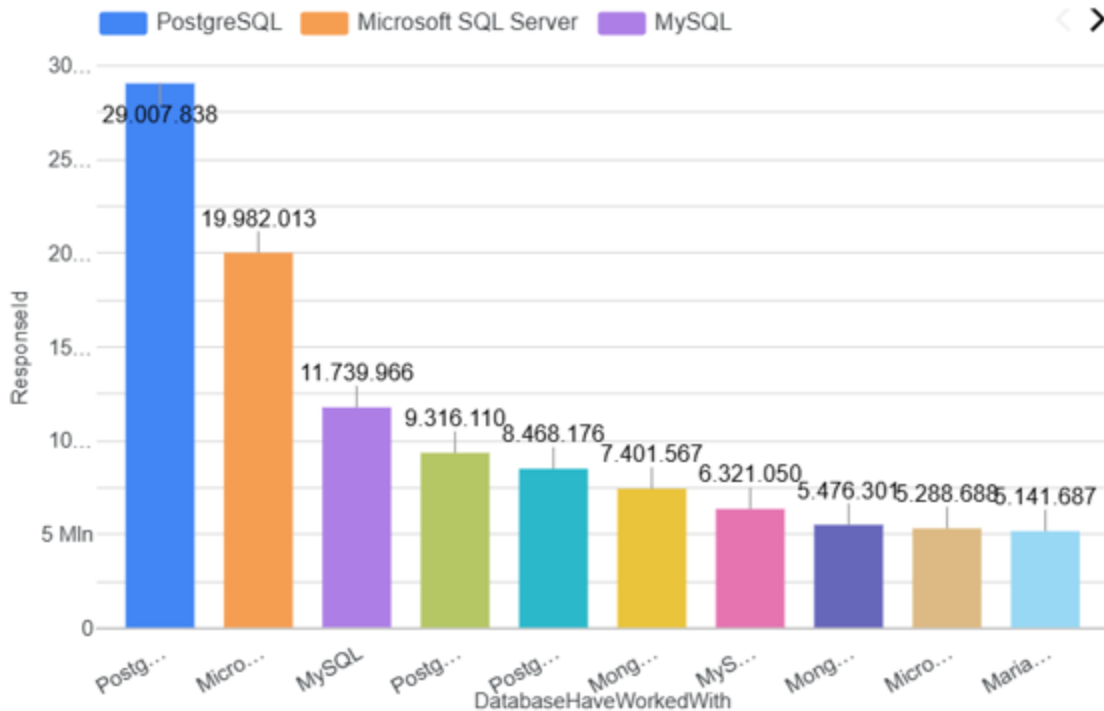
- **JavaScript** remains highly desired, indicating its continued importance in the web development ecosystem.
- **SQL** continues to be a desired skill, reinforcing the ongoing need for database interaction and its fundamental role in software development.
- **HTML/CSS** are still desired by many, as they are fundamental for web development.
- The strong desire for **Python** suggests it will likely continue to grow in popularity and importance in the coming years.



DATABASE TRENDS

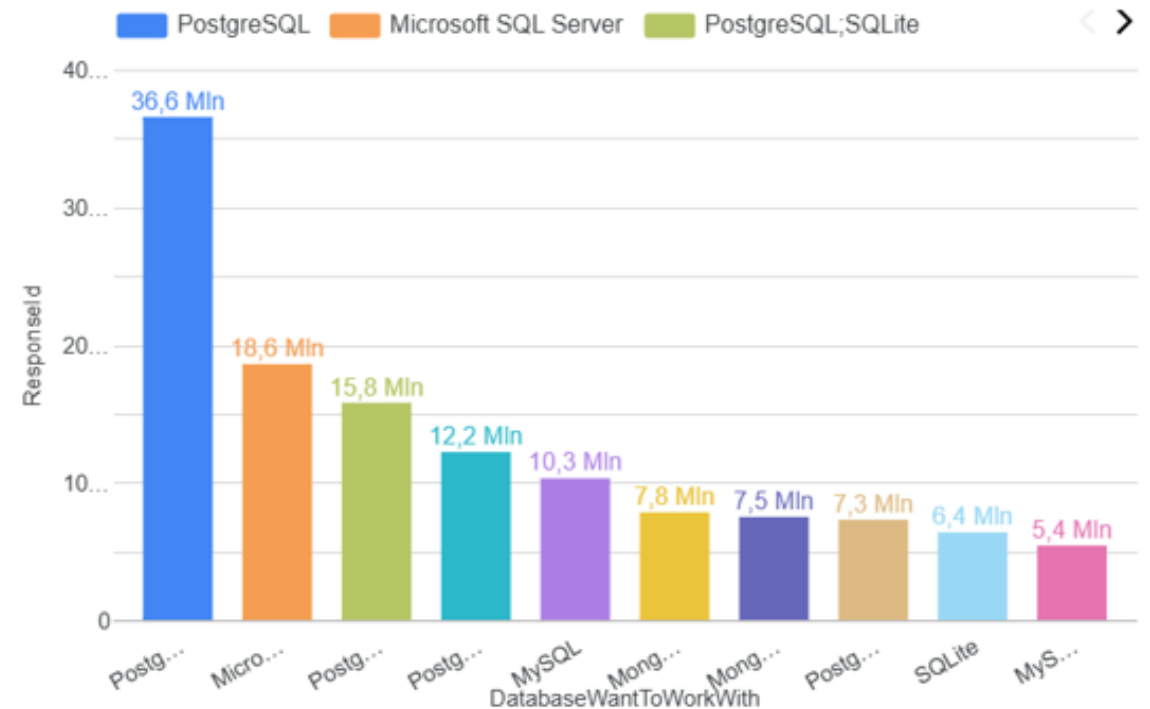
Current Year

Top 10 Databases Used (Stacked Column)



Next Year

Top 10 Databases Desired Next Year (Stacked Column)



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- **MySQL, SQL** and **PostgreSQL** collectively dominate current usage, highlighting the continued importance of structured data management.
- **MongoDB**'s high usage confirms the significant role of NoSQL databases in modern application development.

Implications

- The fact that **PostgreSQL** is the most desired database suggests it may see increased adoption in the future.
- The high ranking of **PostgreSQL, MySQL, MariaDB,** and **MongoDB** reinforces the preference for open-source database technologies among many developers.



DASHBOARD



[IBM Capstone Project - Google Looker dashboards link](#)

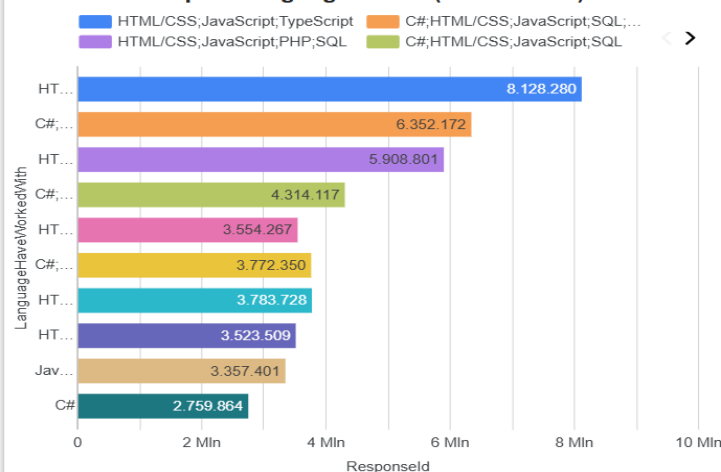
DASHBOARD TAB 1

I. Current Technology Usage

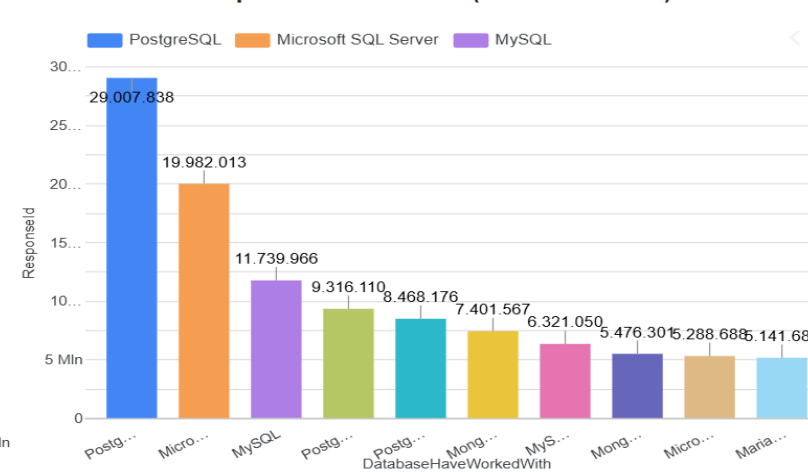
II. Future Technology Trends

III. Demographics

Top 10 Languages Used (Stacked Bar)



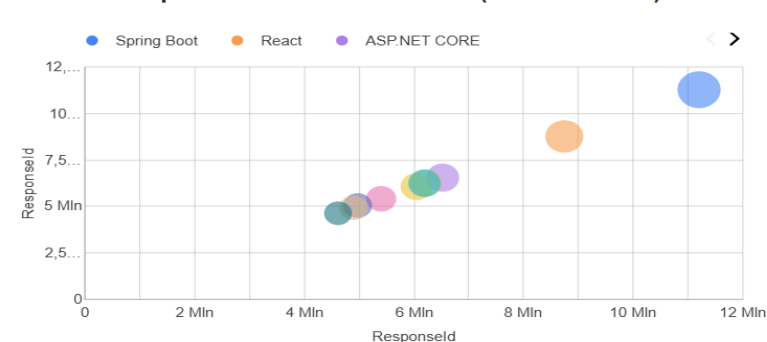
Top 10 DatabasesUsed (Stacked Column)



Top 10 Platforms Used (Word Cloud)



Top 10 Web Frameworks Used (Scatter Bubble)



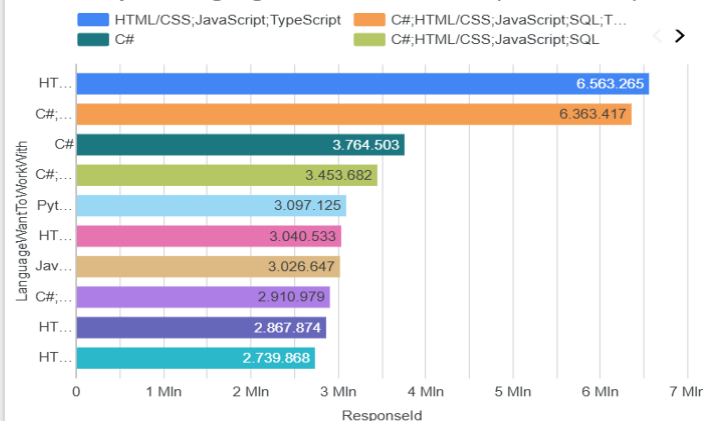
DASHBOARD TAB 2

I. Current Technology Usage

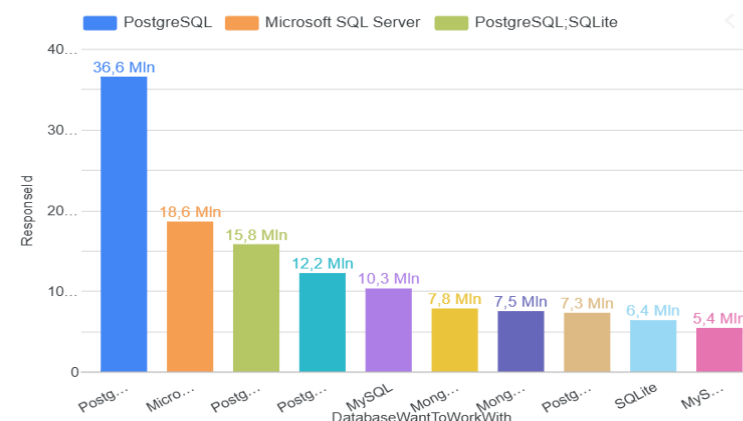
II. Future Technology Trends

III. Demographics

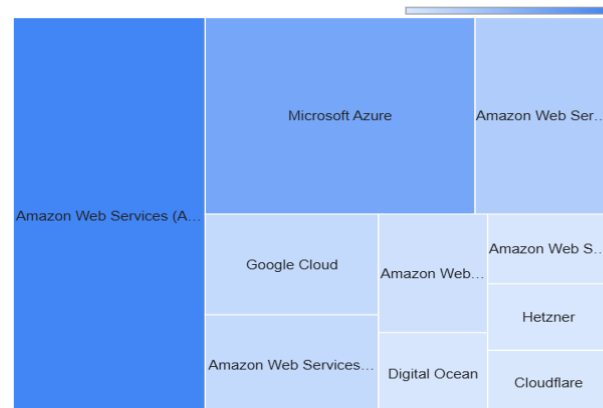
Top 10 Languages Desired Next Year (Stacked Bar)



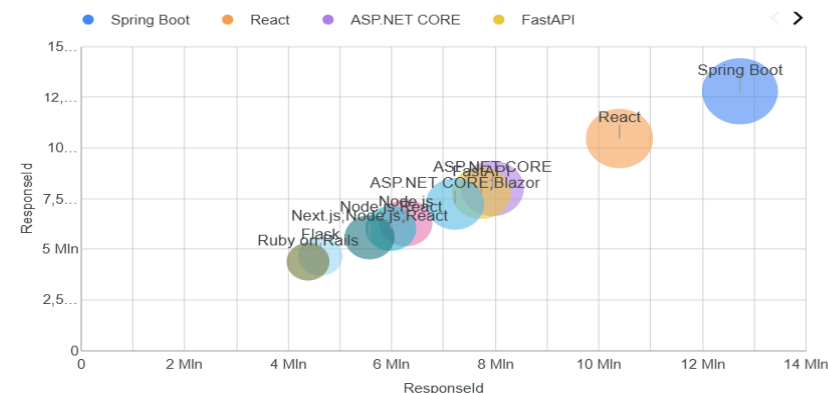
Top 10 Databases Desired Next Year (Stacked Column)



Top 10 Desired Platforms (Tree Map)



Top 10 Desired Web Frameworks (Scatter Bubble)



Ultimo aggiornamento dati: 10/04/2025 21:48:05 | [Norme sulla privacy](#)



Skills Network



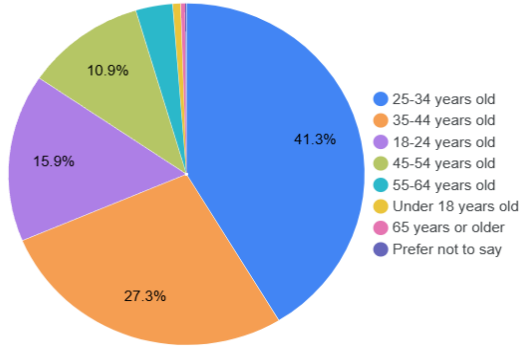
DASHBOARD TAB 3

I. Current Technology Usage

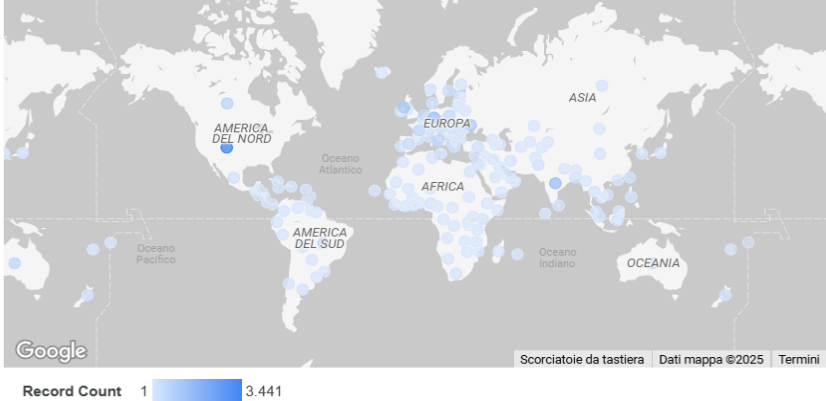
II. Future Technology Trends

III. Demographics

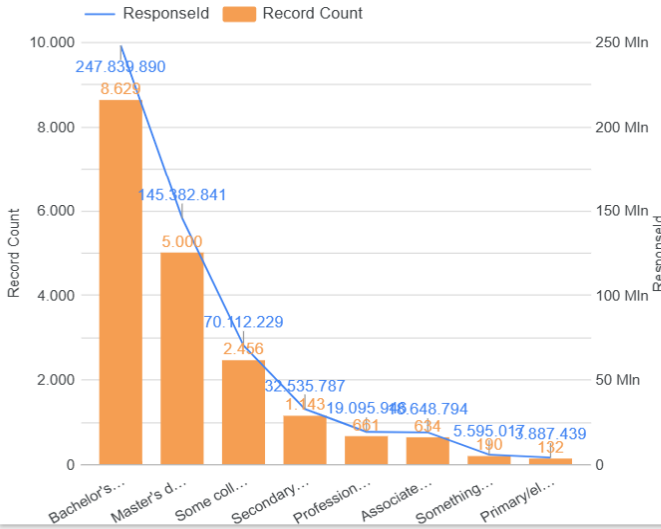
Respondents by Age (Pie Chart)



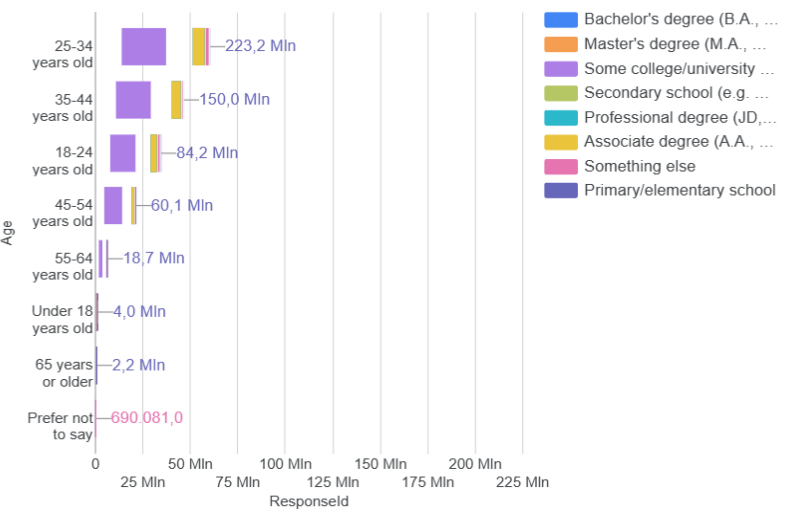
Respondent Count by Country (Map Chart)



Respondent Distribution by Education Level (Line Bar Chart)



Respondent Count by Age, Classified by Education Level (Stacked bar ...)



DISCUSSION



Important Considerations:

- **Intersectionality:** The age, country and education level factors often interact. For example, a young developer in a developed country might have a very different technology profile than an older developer in a developing country.
- **Rapid Change:** The technology landscape changes quickly, so these correlations are not static.
- **Data Availability:** The strength of these correlations depends on the richness and completeness of the demographic data.

OVERALL FINDINGS & IMPLICATIONS

Findings

- **Web Development Remains Crucial:** JavaScript, HTML, and CSS consistently rank high in both current usage and future desires. This confirms that web development is a dominant force in the software industry.
- **Python's Growing Influence:** Python shows strong adoption and even stronger future desire, indicating its increasing importance across various domains.
- **SQL's Enduring Relevance:** SQL remains a fundamental skill, as it's essential for interacting with databases, regardless of the specific application or technology stack.
- **PostgreSQL's Ascendancy:** PostgreSQL is becoming increasingly popular in the database space, potentially challenging MySQL's dominance.
- **Cloud's Impact is Growing:** The interest in cloud-specific technologies like DynamoDB highlights the increasing influence of cloud computing on database choices.
- **Open Source is Favored:** Open-source technologies like Python, JavaScript, PostgreSQL, MySQL, and MariaDB are consistently popular, reflecting a strong community and a preference for flexibility and cost-effectiveness.

Implications

- **Continuous Learning is Essential:** Developers need to stay updated with the latest trends, especially in web technologies, modern languages, and database systems.
- **Data Skills are in High Demand:** With the increasing importance of data-driven applications, skills in SQL and NoSQL databases, as well as data analysis and machine learning (often using Python), are becoming crucial.
- **Cloud Computing is Transforming Development:** Developers should be familiar with cloud platforms and cloud-native technologies, including cloud-based databases and serverless computing.
- **Focus on Performance and Scalability:** There's a growing need for applications that can handle large amounts of data and users, so developers should prioritize languages and tools that offer performance and scalability.
- **The Technology Landscape is Dynamic:** The software development field is constantly evolving, so adaptability and a willingness to learn are essential for success.



CONCLUSION



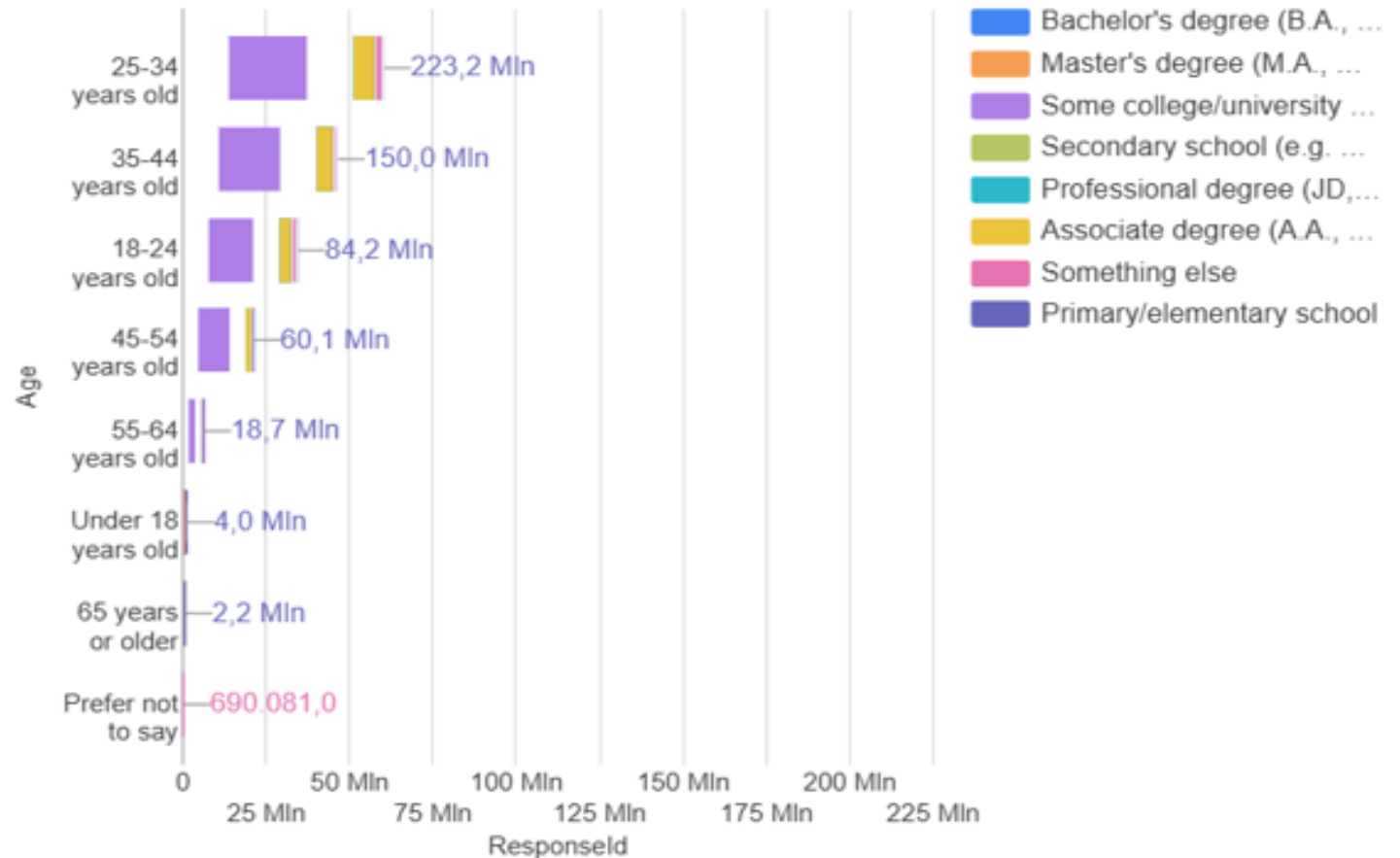
The data paints a picture of a dynamic industry where web development remains central, Python is becoming a dominant force, and modern, scalable technologies are gaining prominence. Developers who embrace these trends and continuously expand their skills will be well-positioned for the future.



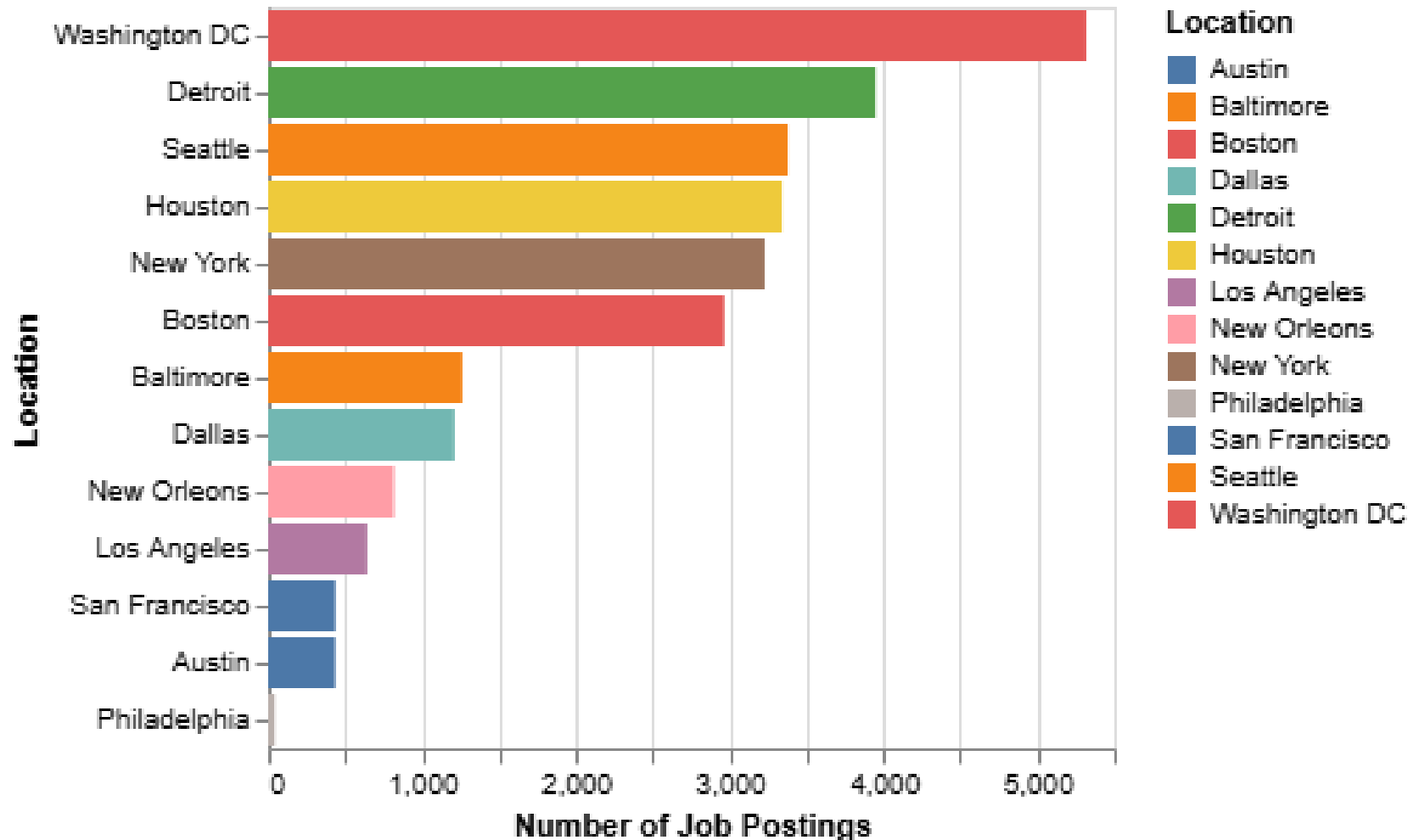
APPENDIX



Respondent Count by Age, Classified by Education Level (Stacked bar ...



JOB POSTINGS (generated with Altair)



POPULAR LANGUAGES (generated with matplotlib)

