Stack Overflow Developer Survey dataset Analysis

Iman Ezatfar

2025-07-11

OUTLINE

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

EXECUTIVE SUMMARY

- Top 10 Databases Have/Want work
- Top 10 Languages Have/Want work
- Top 10 Web frames Have/Want work
- Top 10 Platforms Have/Want work
- Geographic Spread of the Respondents
- Number of Respondents by :
  - Employment
  - Age
  - Education Level
  - Coding Activity Category

INTRODUCTION

Purpose:

- Identify trends in technologies developers use and want to use (languages, platforms, databases).
- Analyze developer demographics, education, and employment patterns globally.

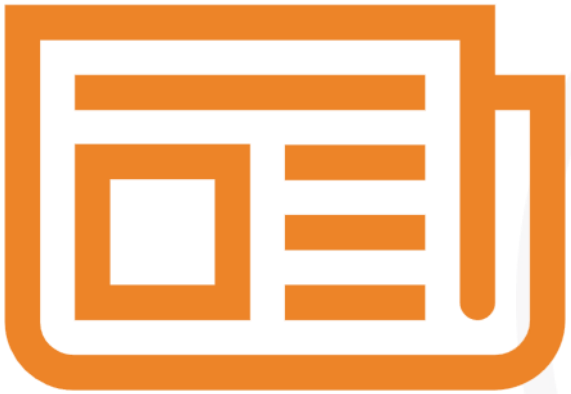
Target:

- Tech employers and recruiters seeking insights on in-demand skills and talent distribution.
- Educators and training providers aiming to align programs with industry needs.

Value:

- Helps stakeholders make data-driven decisions on hiring, training, and product development.
- Empowers developers to benchmark their skills and career direction against global trends.

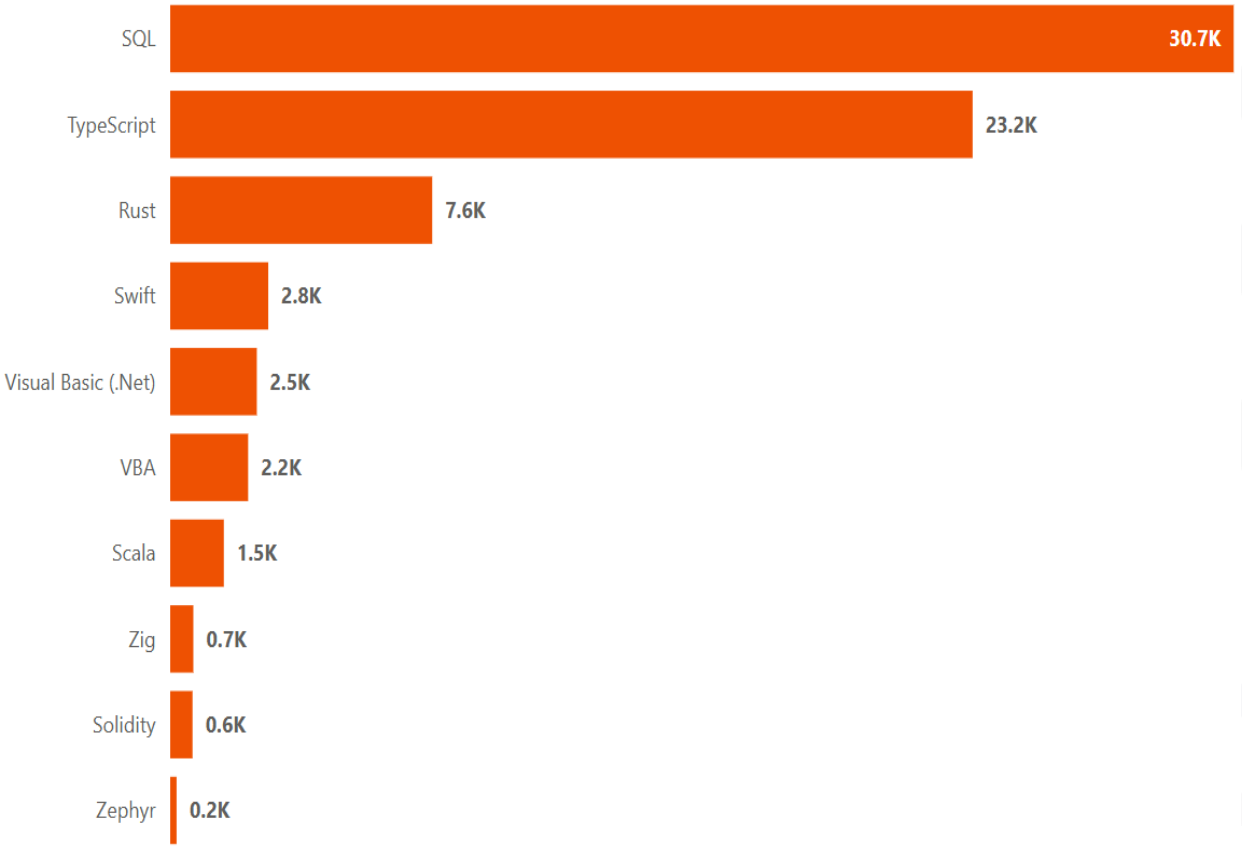
METHODOLOGY



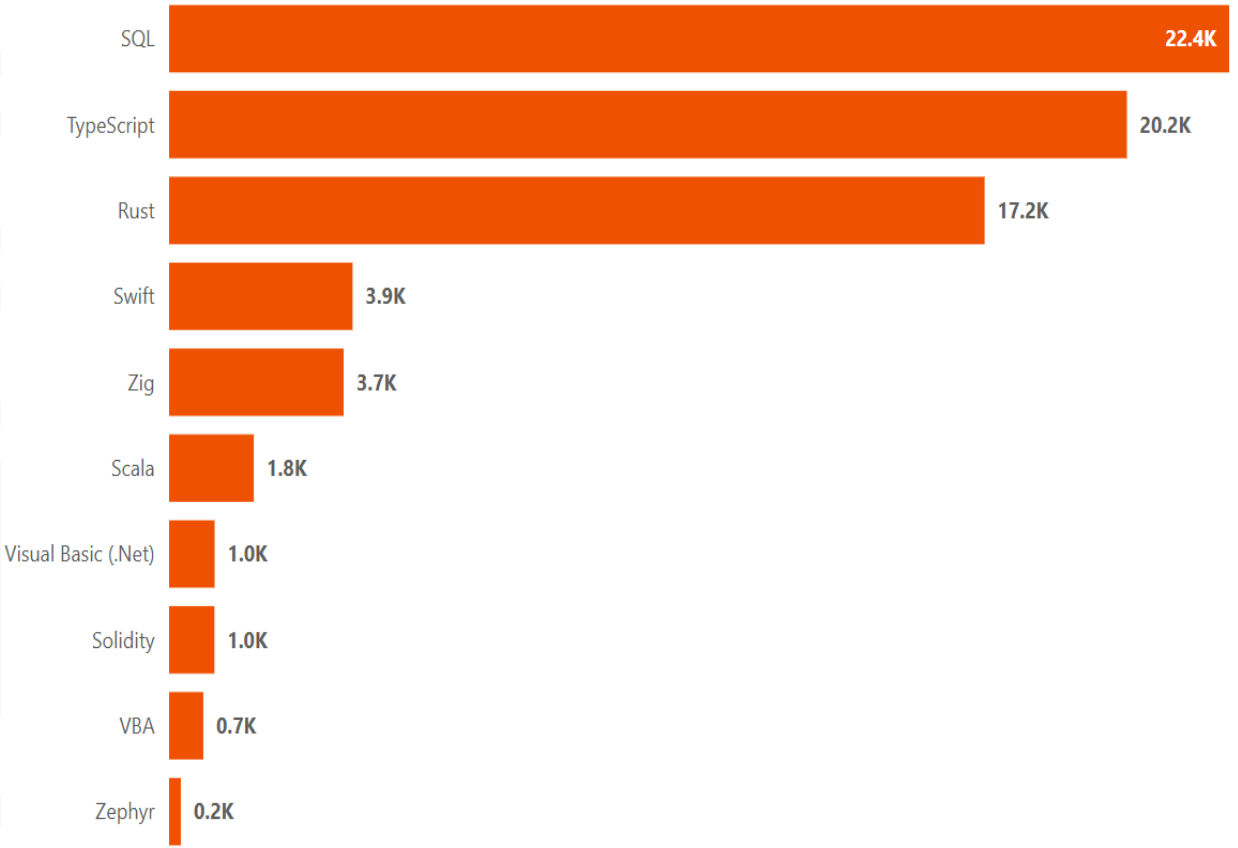
- Stack Overflow Developer Survey dataset Open Database License (ODbL)
- Contains responses from thousands of developers globally in CSV format
- Survey conducted by Stack Overflow via voluntary online participation.
- Questions covered a wide range of developer-related topics (skills, tools, work status).
- Responses collected anonymously and randomized using ResponseId.
- Collected via Stack Overflow's internal survey tool and shared publicly.
- Wrangling
  - Cleaning, Converting, Normalizing, Dealing with Outliers
  - Feature Engineering, Mapping, Binning, Splitting, Exporting
  - Visualizing

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

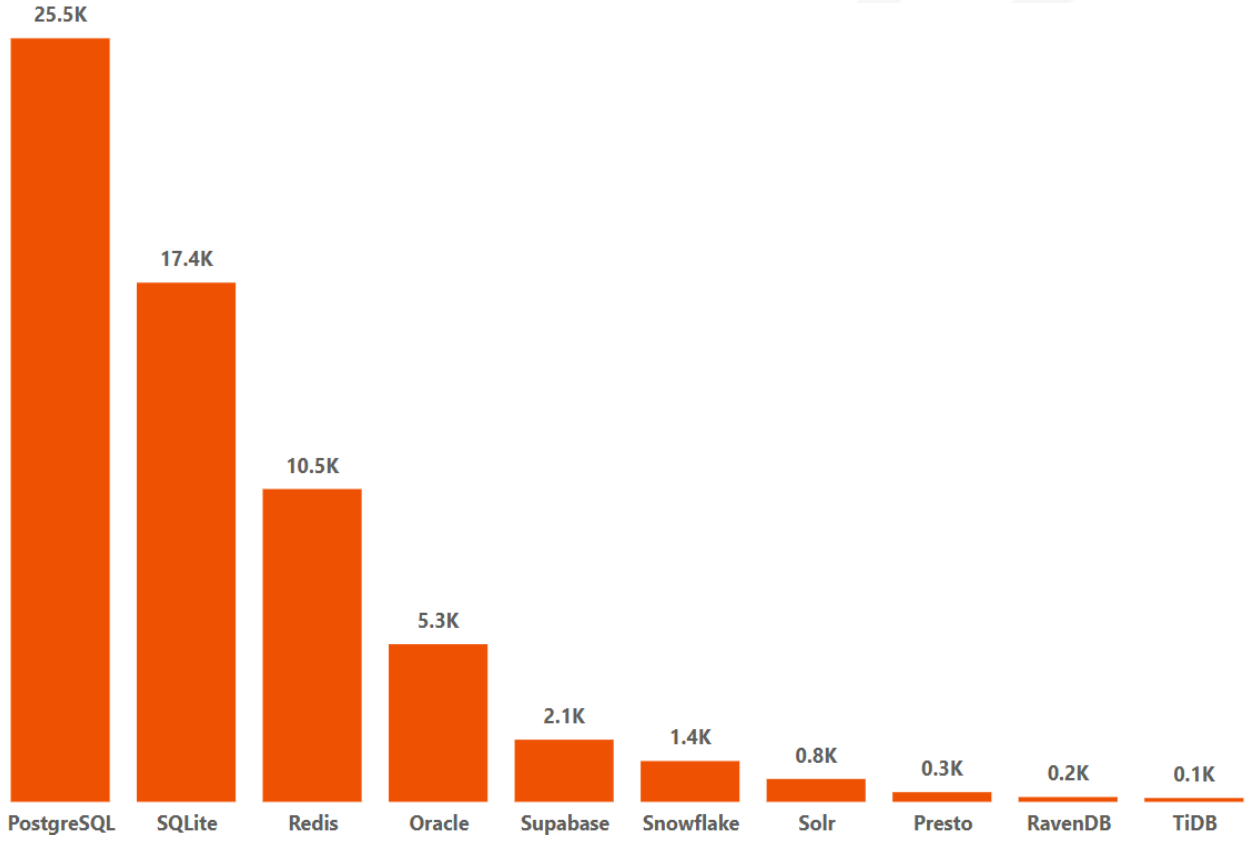
- **Significant Decline in SQL Mentions:** SQL usage dropped from 30.7K to 22.4K, indicating a 27% decrease, though it remains the most used language.
- **Rust on the Rise:** Rust saw a huge jump from 7.6K to 17.2K, climbing from 3rd to 3rd place but with a much narrower gap behind TypeScript.
- **Emergence of Zig:** Zig went from 0.7K to 3.7K, overtaking several languages including Visual Basic (.Net) and VBA, indicating growing interest in low-level programming languages.

Implications

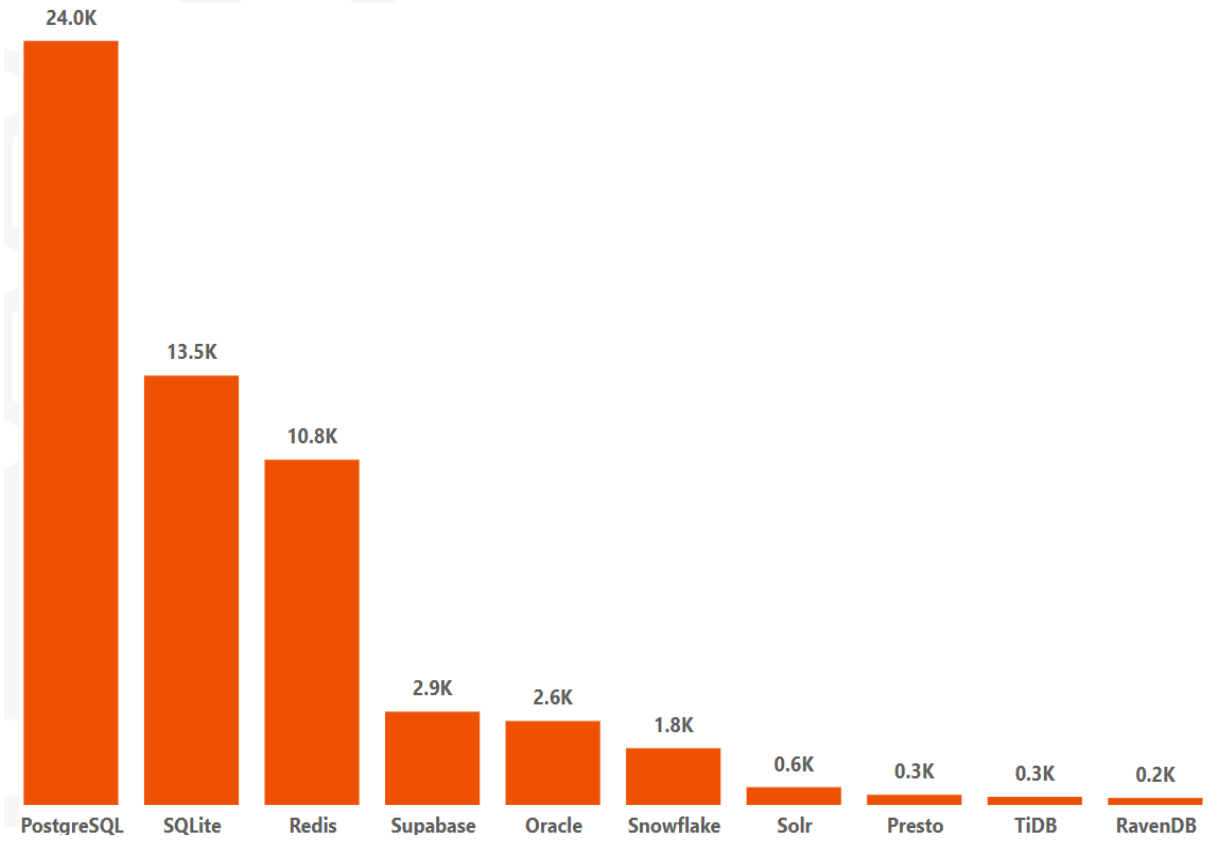
- **Shift Toward Modern Systems Languages:** The rise in Rust and Zig suggests increased industry focus on performance, safety, and systems-level development.
- **Potential Need to Diversify SQL Skills:** Despite being top-ranked, SQL's decline hints at growing integration with cloud-native, typed, or procedural tools, prompting professionals to pair SQL with other languages like Python or Rust.
- **Legacy Technologies Losing Ground:** Older languages like Visual Basic (.Net) and VBA are seeing major drops, suggesting reduced demand and prompting companies to migrate from legacy tech stacks.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- **PostgreSQL and SQLite Slightly Decline:** PostgreSQL dropped from 25.5K → 24.0KSQLite dropped from 17.4K → 13.5KThese still remain the top two databases but are seeing reduced interest.
- **Database Surges in Popularity:** Database usage jumped from 2.1K to 2.9K, surpassing Oracle in the next year.
- **Redis Holds Steady:** Redis slightly increased from 10.5K to 10.8K, showing consistent and growing usage as an in-memory data store.

Implications

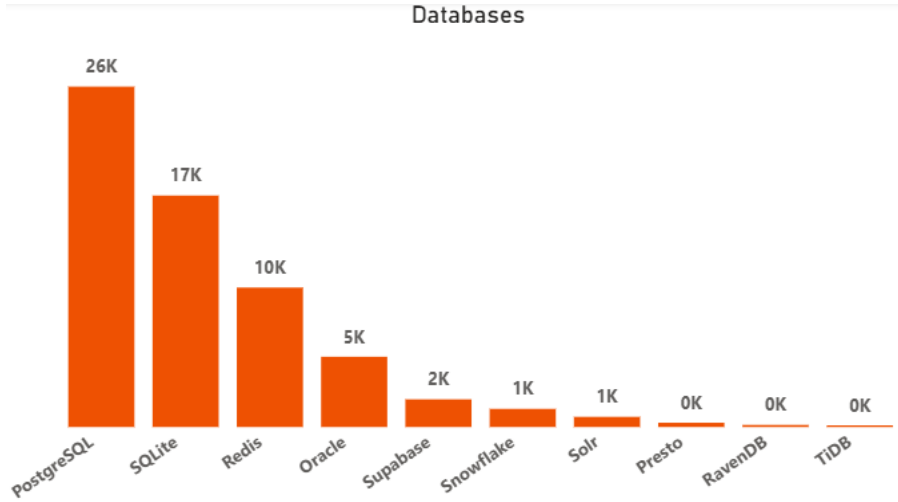
- **Developers Leaning Toward Modern, Open-Source Stacks:** The growth of Database reflects increased adoption of full-stack open-source alternatives to Firebase.
- **Traditional Databases Losing Some Ground:** Oracle's decline from 5.3K to 2.6K may indicate a shift away from legacy enterprise solutions to more developer-friendly databases.
- **No Clear Leader Among Emerging Databases:** Technologies like TiDB, Presto, RavenDB remain niche with minimal traction, suggesting they haven’t broken into mainstream adoption yet.

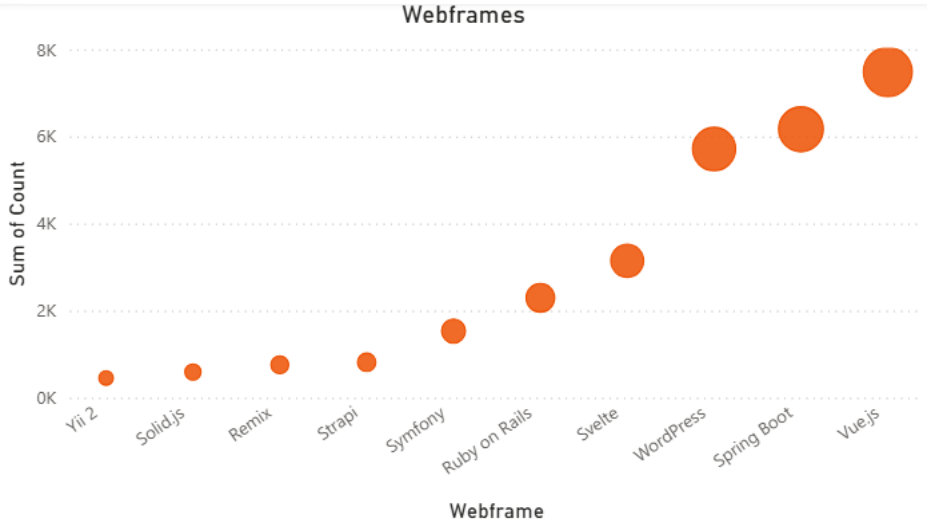
DASHBOARD

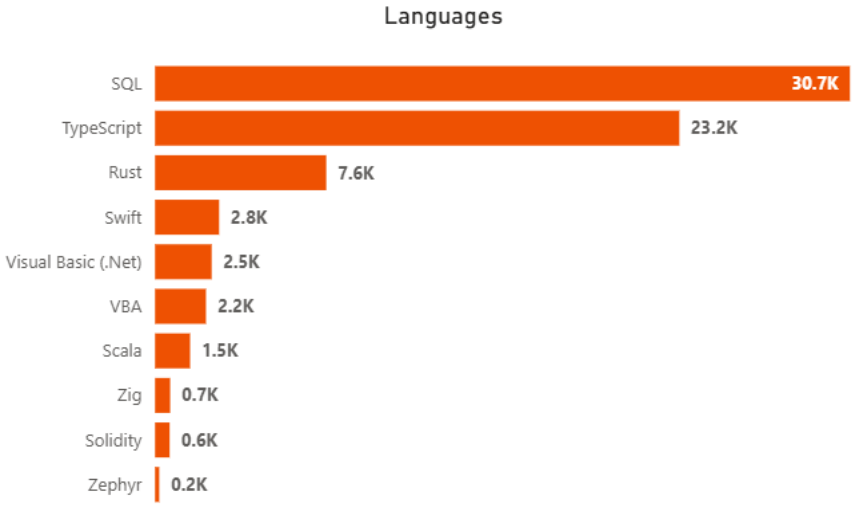
Power BI:

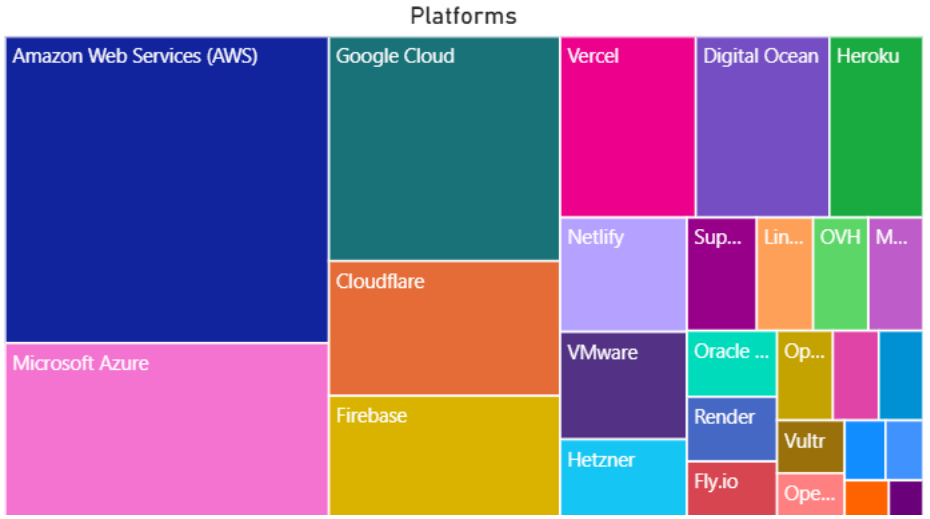
- Dashboard: Current Technology Usage
- Dashboard: Future Technology Trends
- Dashboard: Demographics

DASHBOARD TAB 1

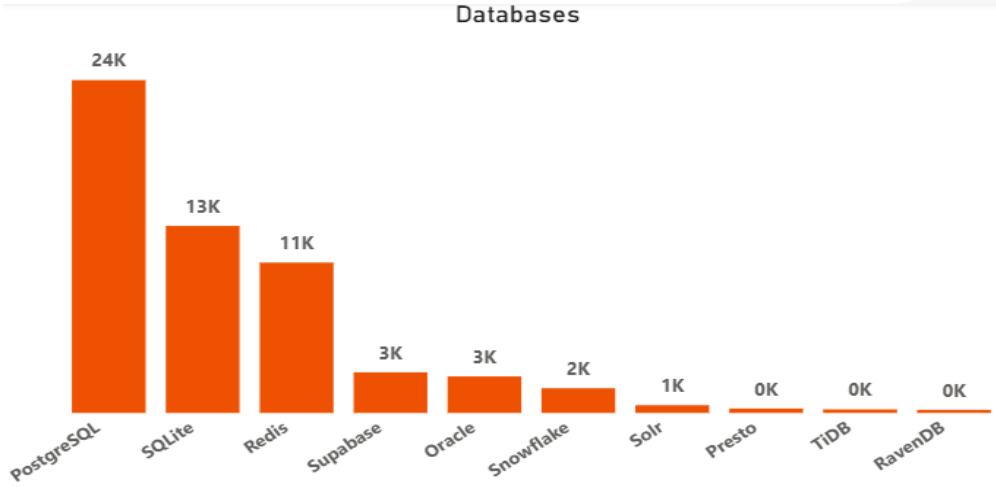


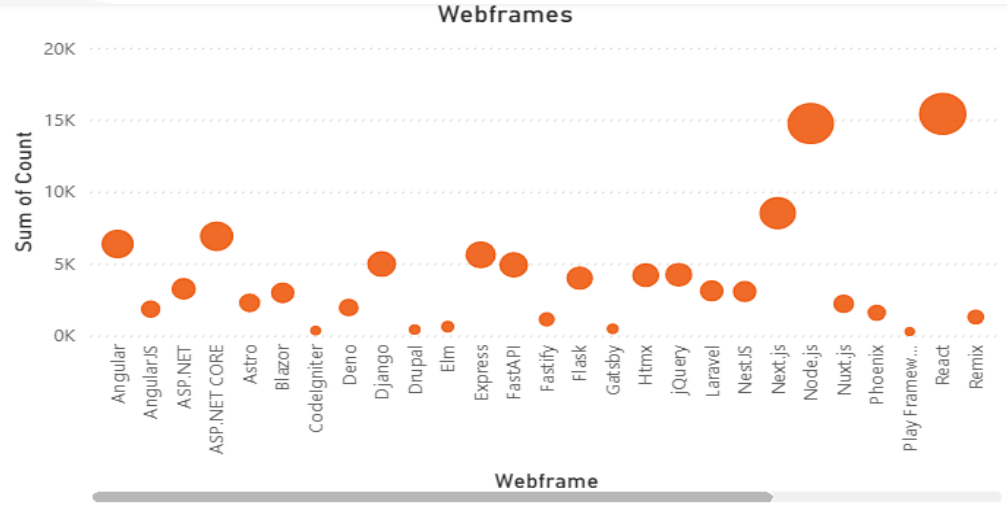


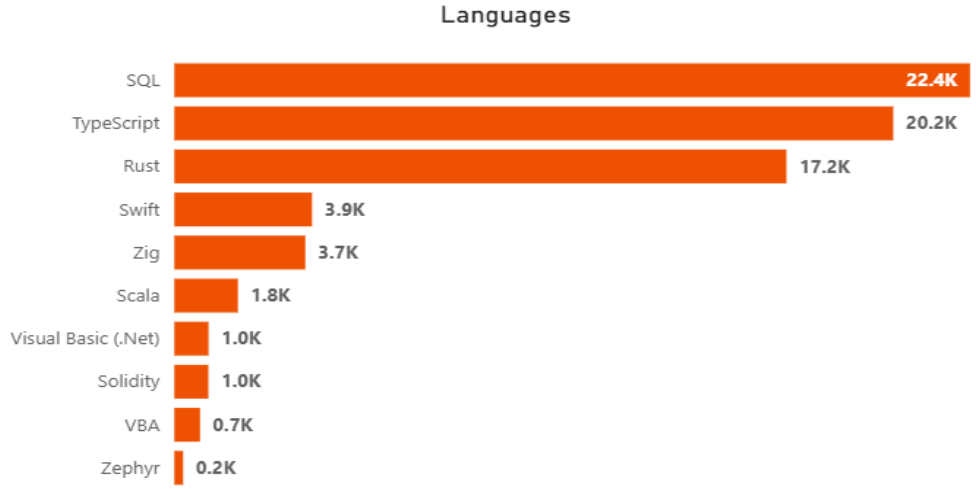


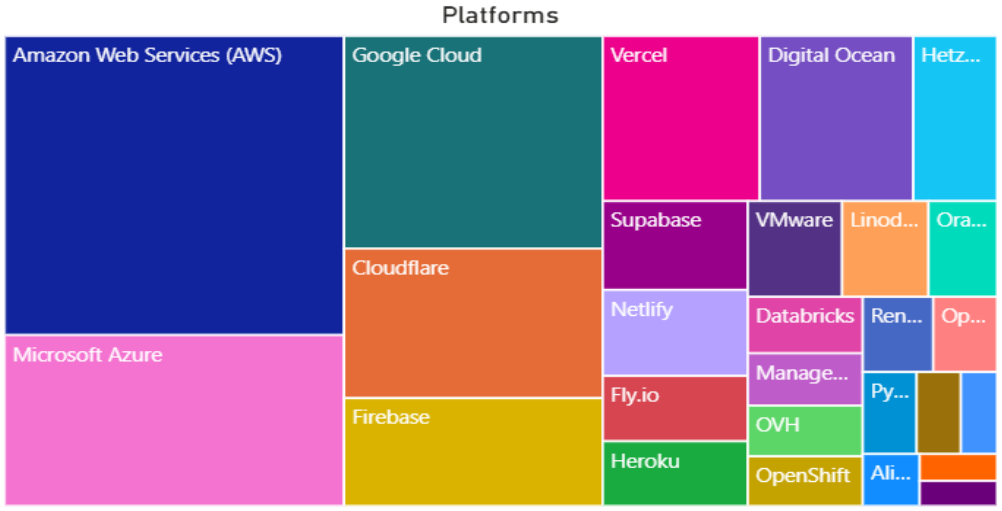


DASHBOARD TAB 2



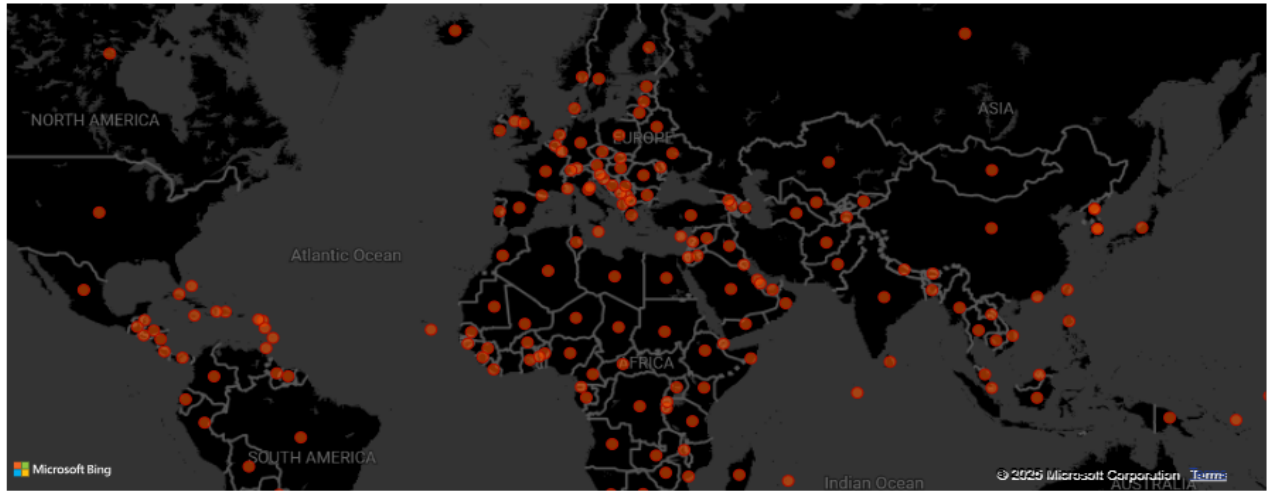




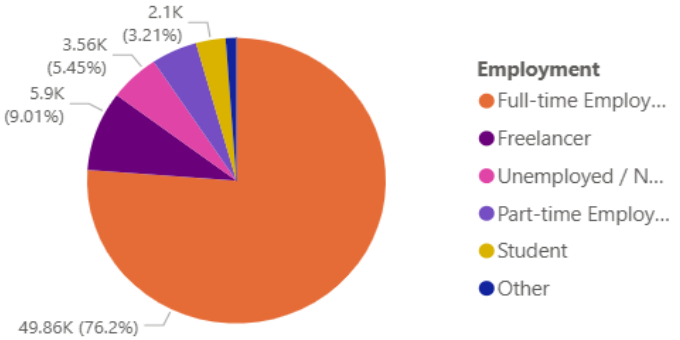


DASHBOARD TAB 3

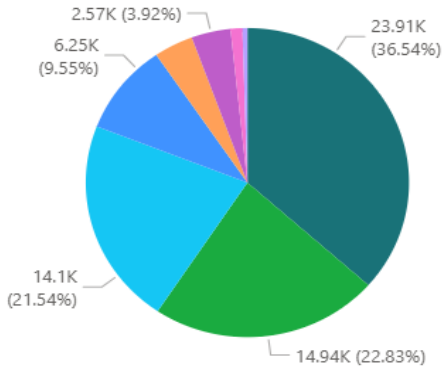
Country

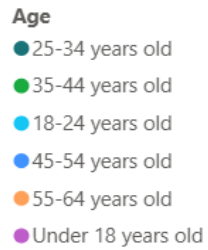


Employment

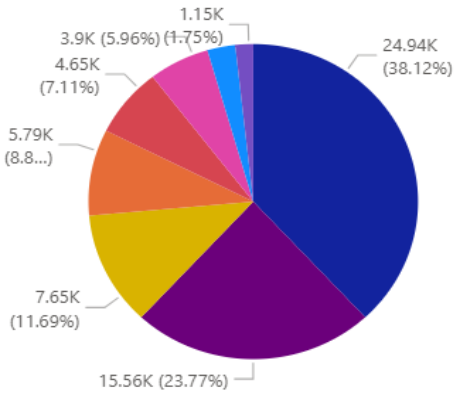


Responseld by Age

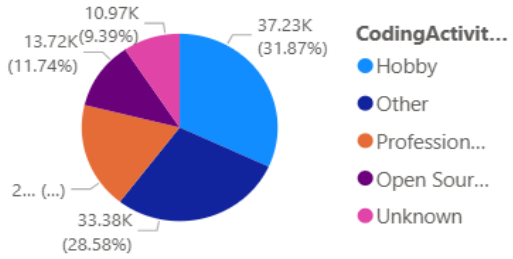




Education Level



Coding Activity Category



DISCUSSION

- Most of the respondents live in Europe and Africa
- More than half of the respondents are between 25-44 years old
- Around 40% of them are Bachelors holders
- More than two-third are full-time employees
- Around 1/3 of them also code as a hubby

OVERALL FINDINGS & IMPLICATIONS

Findings

- Main characteristics of the respondents
- The Platforms, Databases, Languages, and Web frames
- Know the next year trend

Implications

- Deciding the next year target market regarding age, location, DB, and so on
- Improving infrastructures
- Preparing learning materials

CONCLUSION

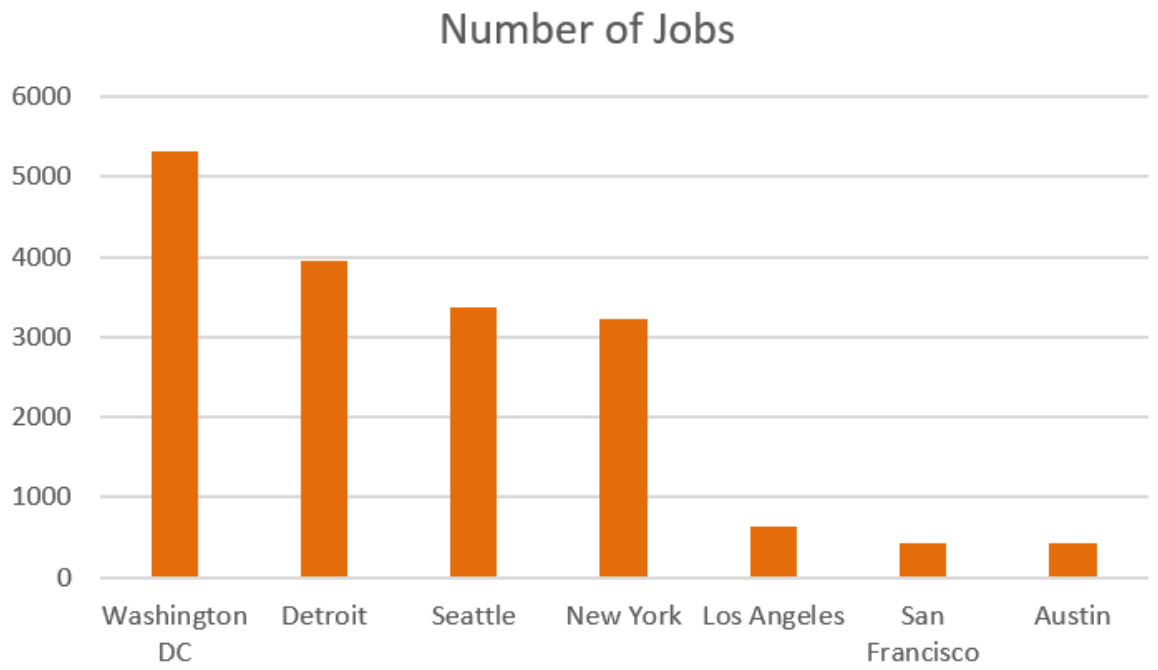
- Discovered that SQL, and Java are consistently the most used languages, indicating their global dominance and relevance in tech careers
- Found that remote databases and platforms preference as well as the web frames
- Identified that younger developers (under 45) containing most of the coders
- These insights can guide policy makers, educators, and employers to tailor training programs, recruitment strategies, and remote work policies aligned with current tech workforce trends.

APPENDIX

- Include any relevant additional charts, or tables that you may have created during the analysis phase.

JOB POSTINGS

In Module 1 you have collected the job posting data using Job API in a file named “job-postings.xlsx”. Present that data using a bar chart here. Order the bar chart in the descending order of the number of job postings.



POPULAR LANGUAGES

In Module 1 you have collected the job postings data using web scraping in a file named “popular-languages.csv”. Present that data using a bar chart here. Order the bar chart in the descending order of salary.