Capstone Project(Data Analysis) IT Business

Name: Hadeel Salman Dawood Alsaadi

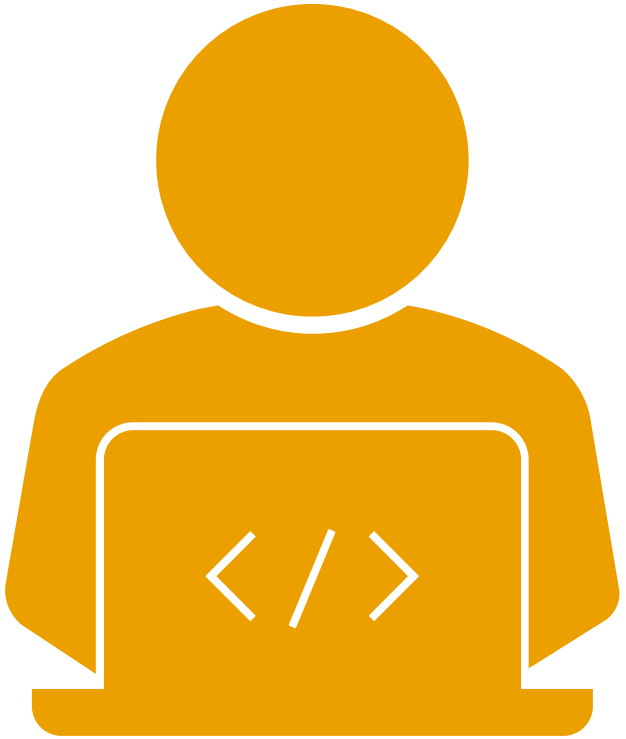
Date: 27.02.2025

Outlines

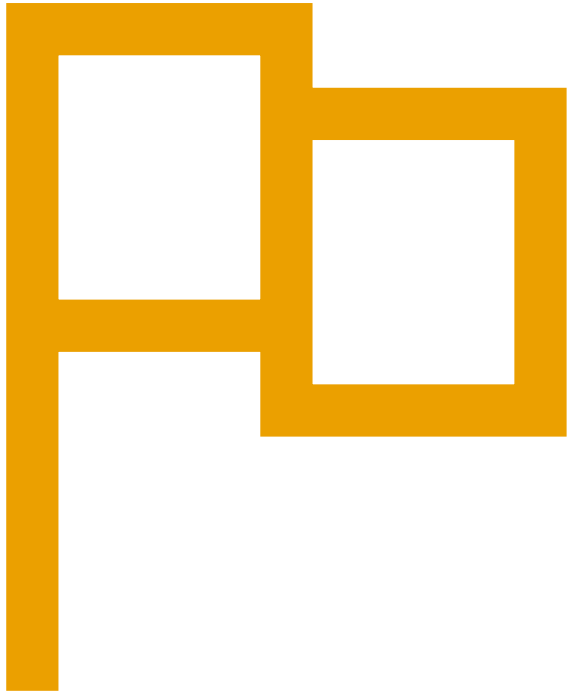
- Executive summary
- Introduction
- Methodology
- Results
  - Programming Languages Trends
  - Database Trends
  - Dashboards
    - Current Technology
    - Future Technology
    - Demographics
- Conclusion

Executive Summary

- This analysis aims to identify the trends and future skill requirements in IT Business an that will include all the programming languages, databases, web Frameworks and platforms have worked and want to work with.
- Data sources
  - Collecting data from Job platforms and web pages
  - Developer surveys
- Statistical analysis using python and SQL queries
- Key Findings
  - Top 10 technologies in use and those expected to be in demand in the near future.
  - To create a complete picture of the IT domain, we will compare current and anticipated technologies across key areas.
- Demographic distribution and education level of respondents.



Introduction



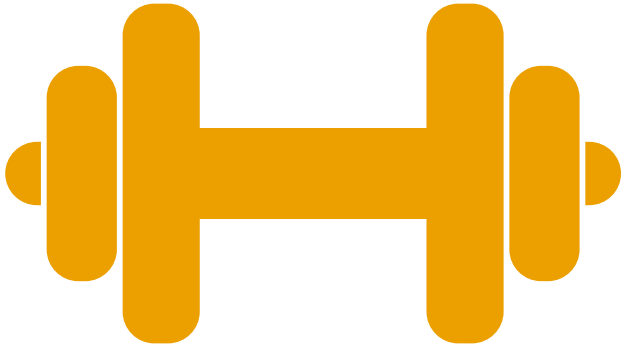
Technology is rapidly evolving, and we must maintain a fast pace to stay competitive in this landscape.

The purpose of this analysis is to answer the following questions:

•What are the top programming languages, databases, web frameworks and platforms in demand today?

•Which skills are needed in the job market?

Methodology



Data sources

- Job postings
- Training portals
- Developer surveys

Wrangling the data set(python pandas)

1. Data Collection

Gather data from different sources (Web scraping, APIs, Spreed sheets, etc.).

Ensure the data format is compatible for further processing.

2. Data Exploration

Load and inspect the dataset to understand its structure and quality.

Identify missing values, duplicate records, and inconsistencies.

3. Data Cleaning

Handle missing data (e.g., filling missing values, dropping rows/columns).

Remove duplicates and correct inconsistent formatting.

Fix structural errors (e.g., incorrect data types, typos).

4. Data Transformation

Normalize or standardize numerical data.

Encode categorical variables (e.g., one-hot encoding, label encoding).

Create new features if needed (feature engineering).

5. Data Integration

Combine multiple datasets if necessary (merging, joining).

Ensure consistency in column names and formats.

6. Data Validation

Check for outliers and inconsistencies.

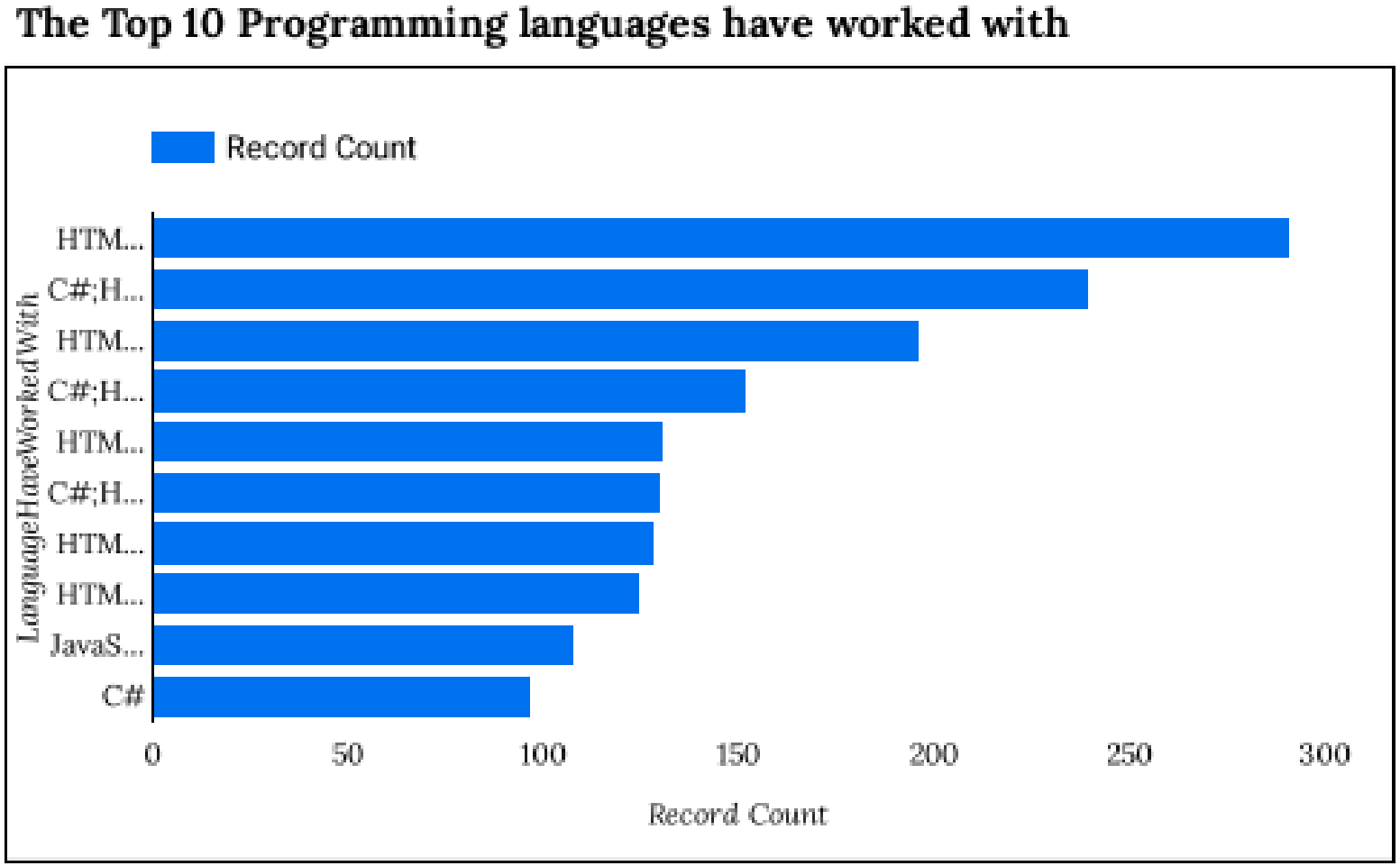
Verify the accuracy and completeness of the processed data.

7. Data Exporting

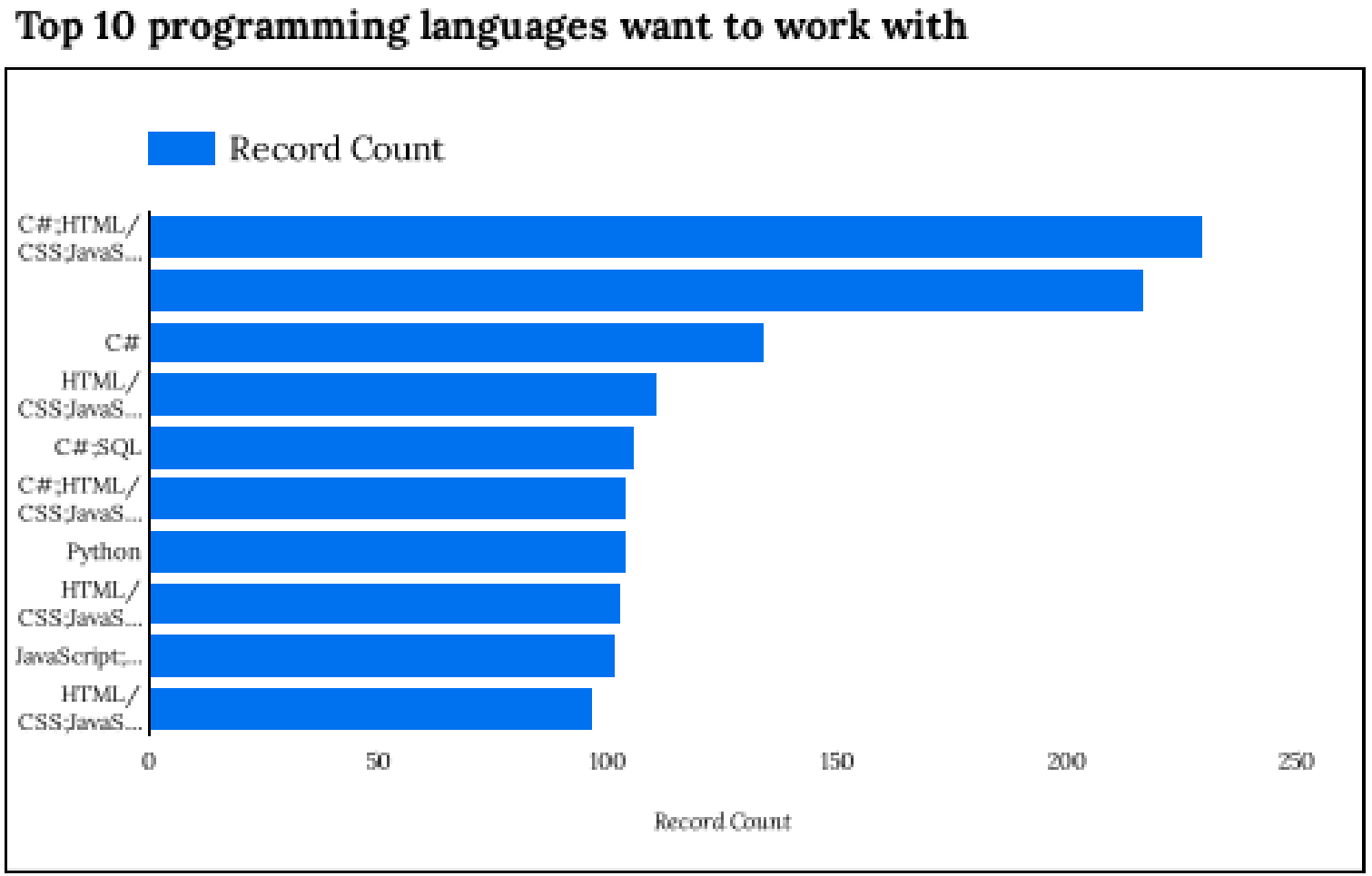
Save the cleaned and structured data in a suitable format (CSV, SQL database, etc.).

Ensure the data is ready for analysis

Programming Languages
Trends
Top 10 in use

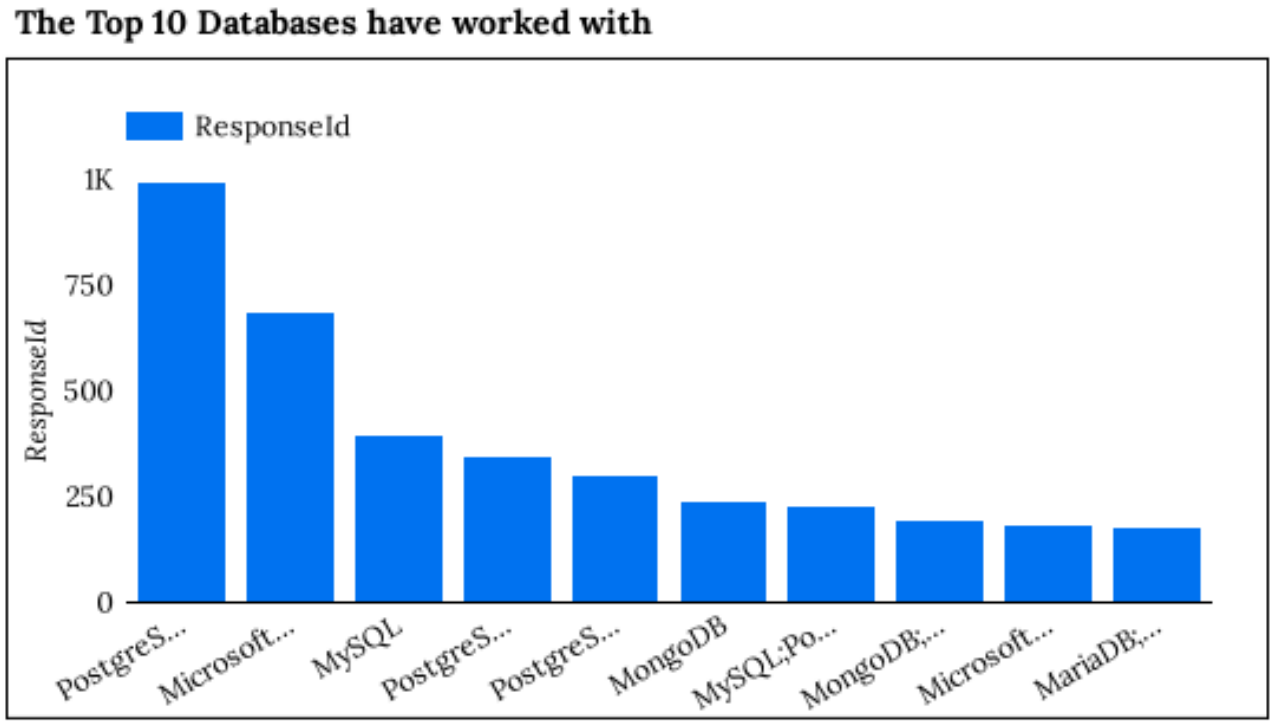


Programming Languages Trends
Top 10 want to work with



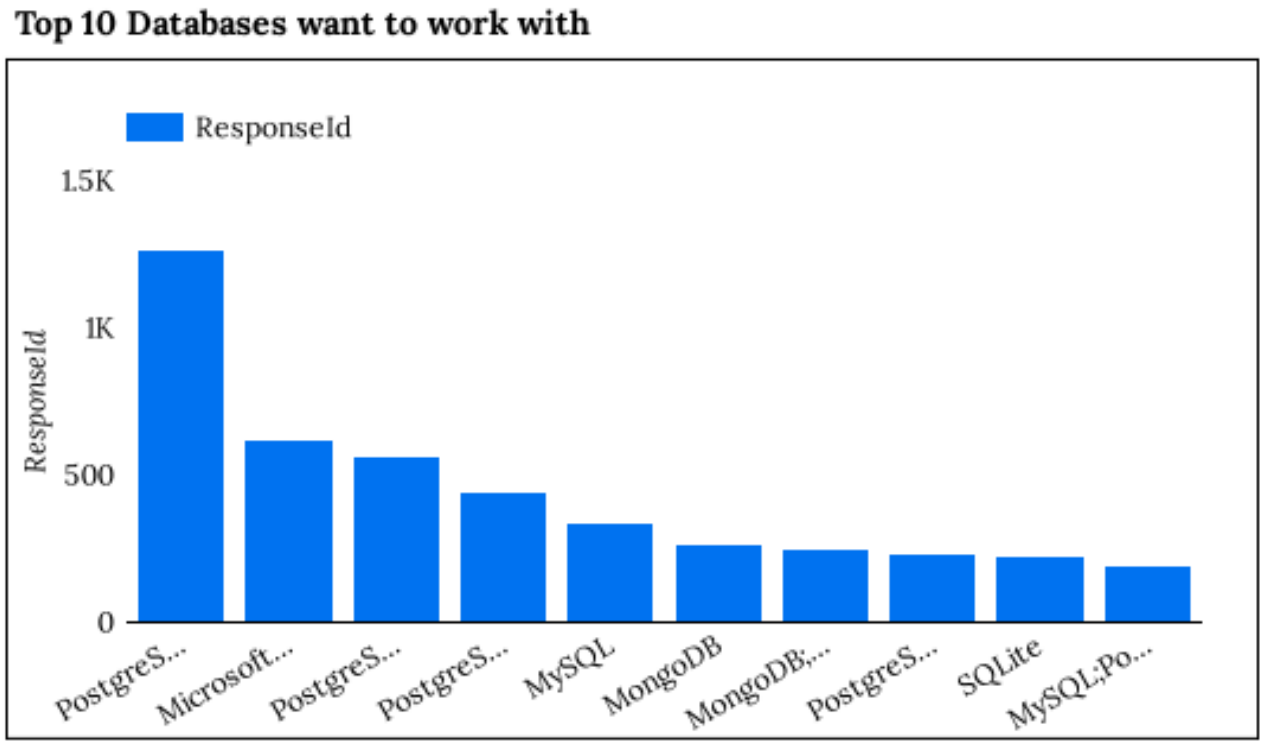
Databases in use

- PostgreSQL is the most used database in the IT world

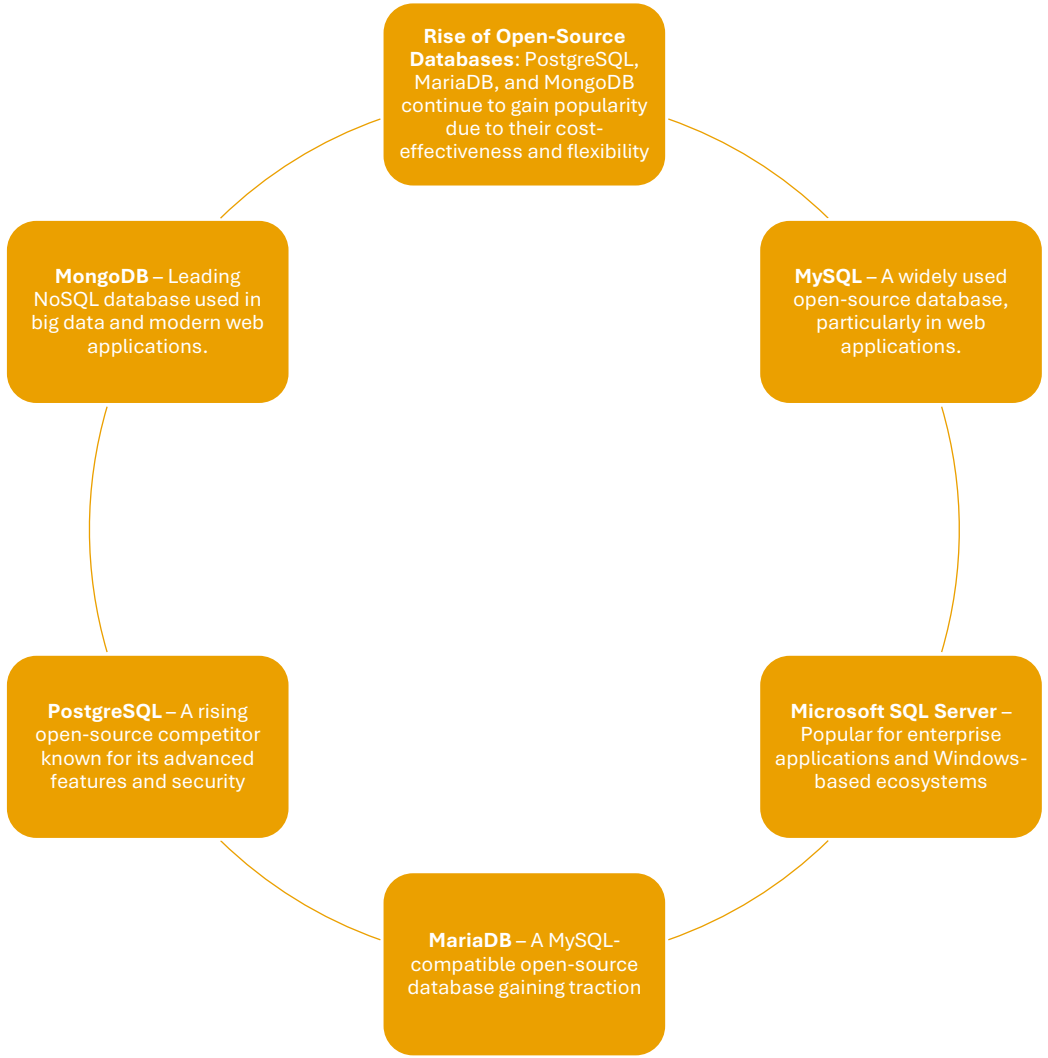


Database want to work with

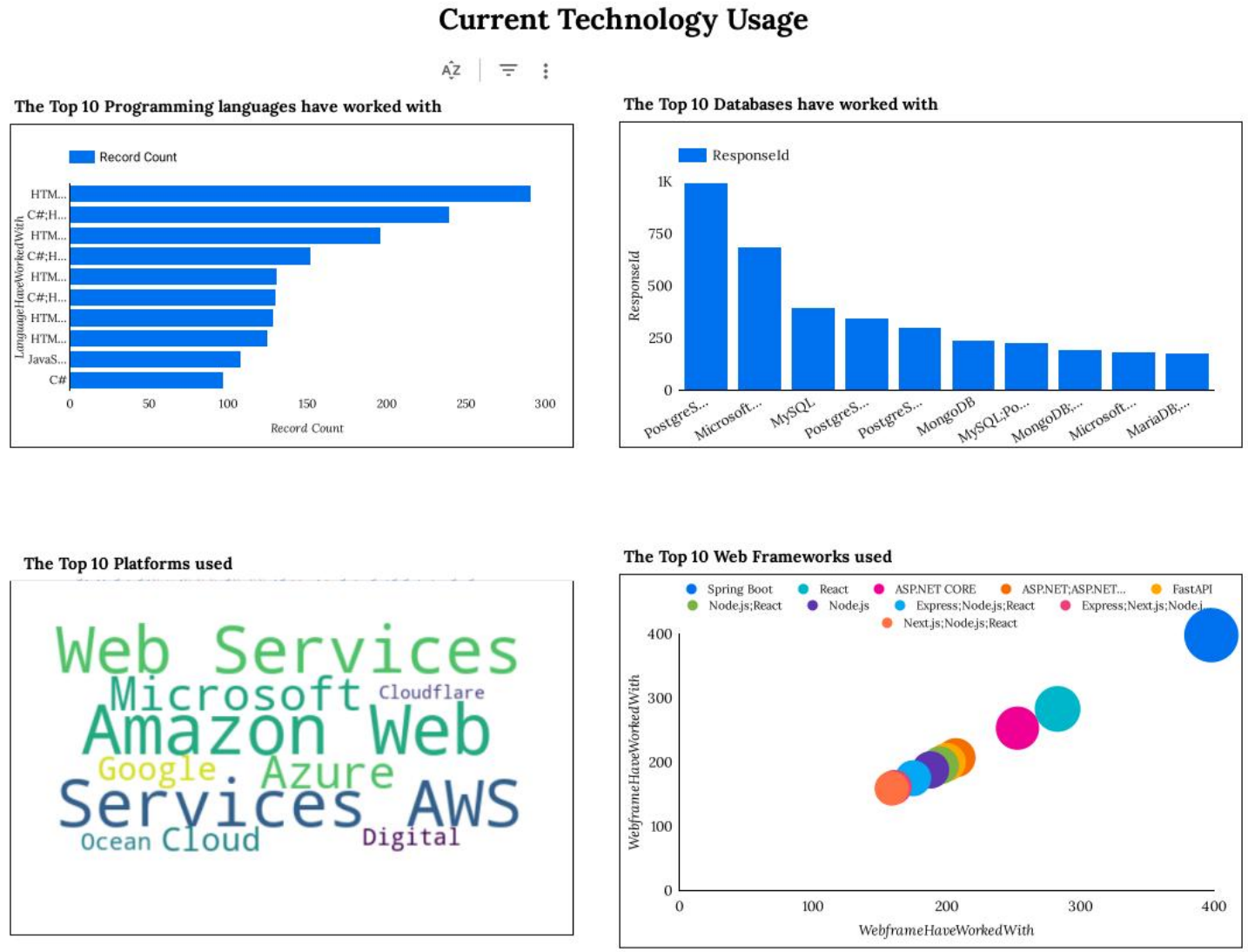
- PostgreSQL dominates the database world and remains in the top ten databases



Findings and implications about top 10 databases in use and want to work with



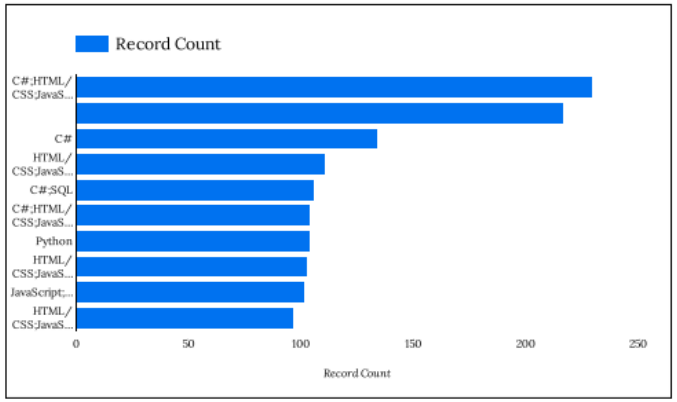
Current Technology Usage



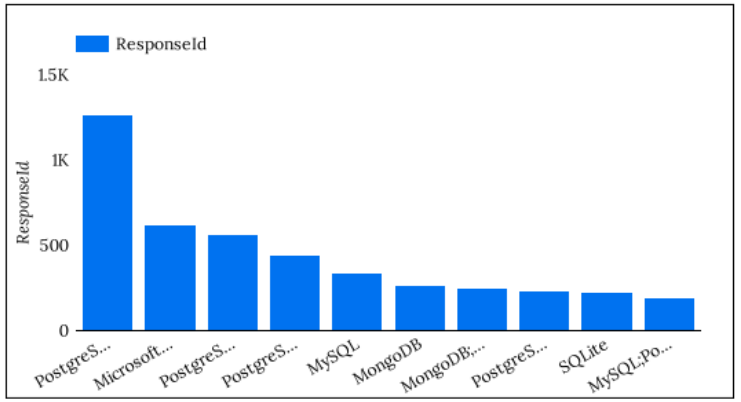
Future Technology Trends

Future Technology Trends

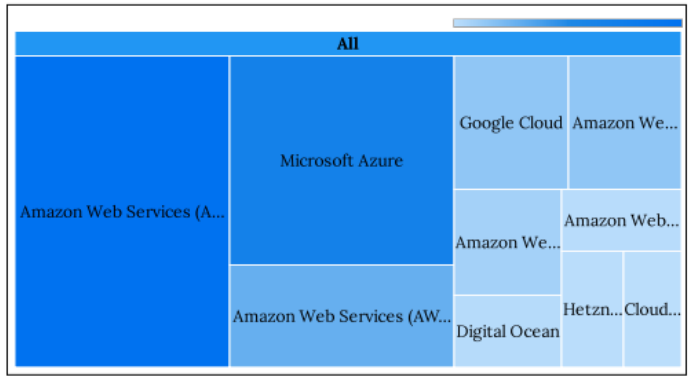
Top 10 programming languages want to work with



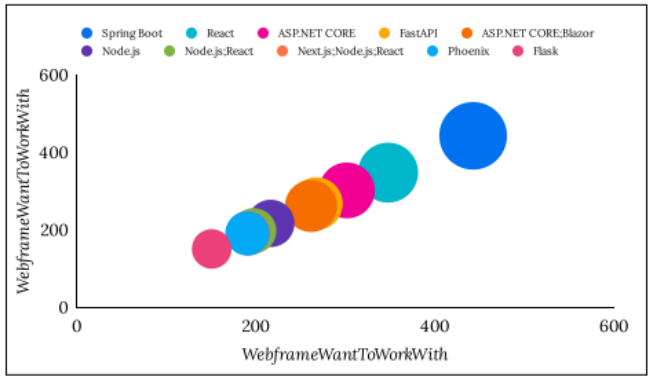
Top 10 Databases want to work with



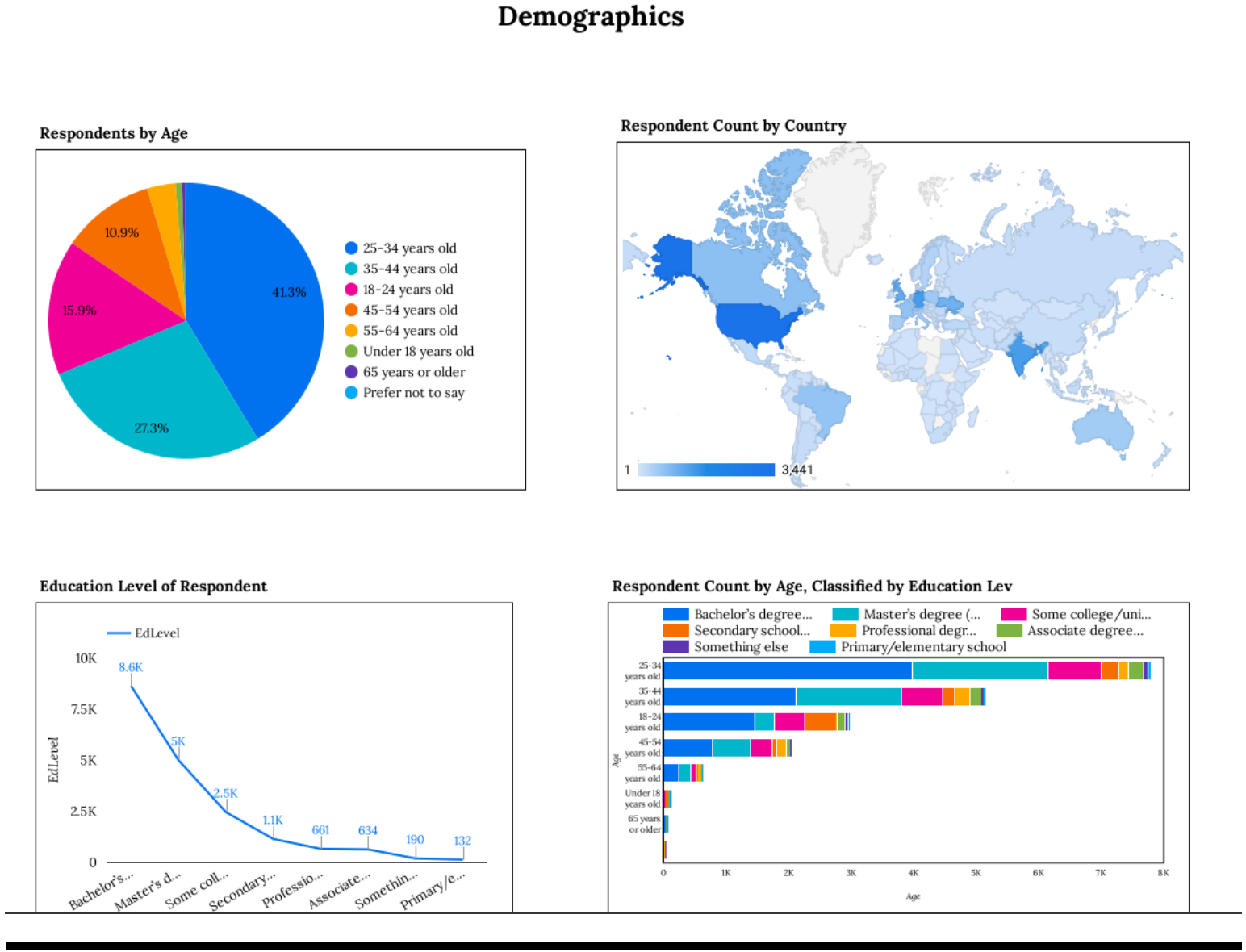
Top 10 Platforms want to work with



Distribution of the top 10 Web Frameworks want to work with



Demographics

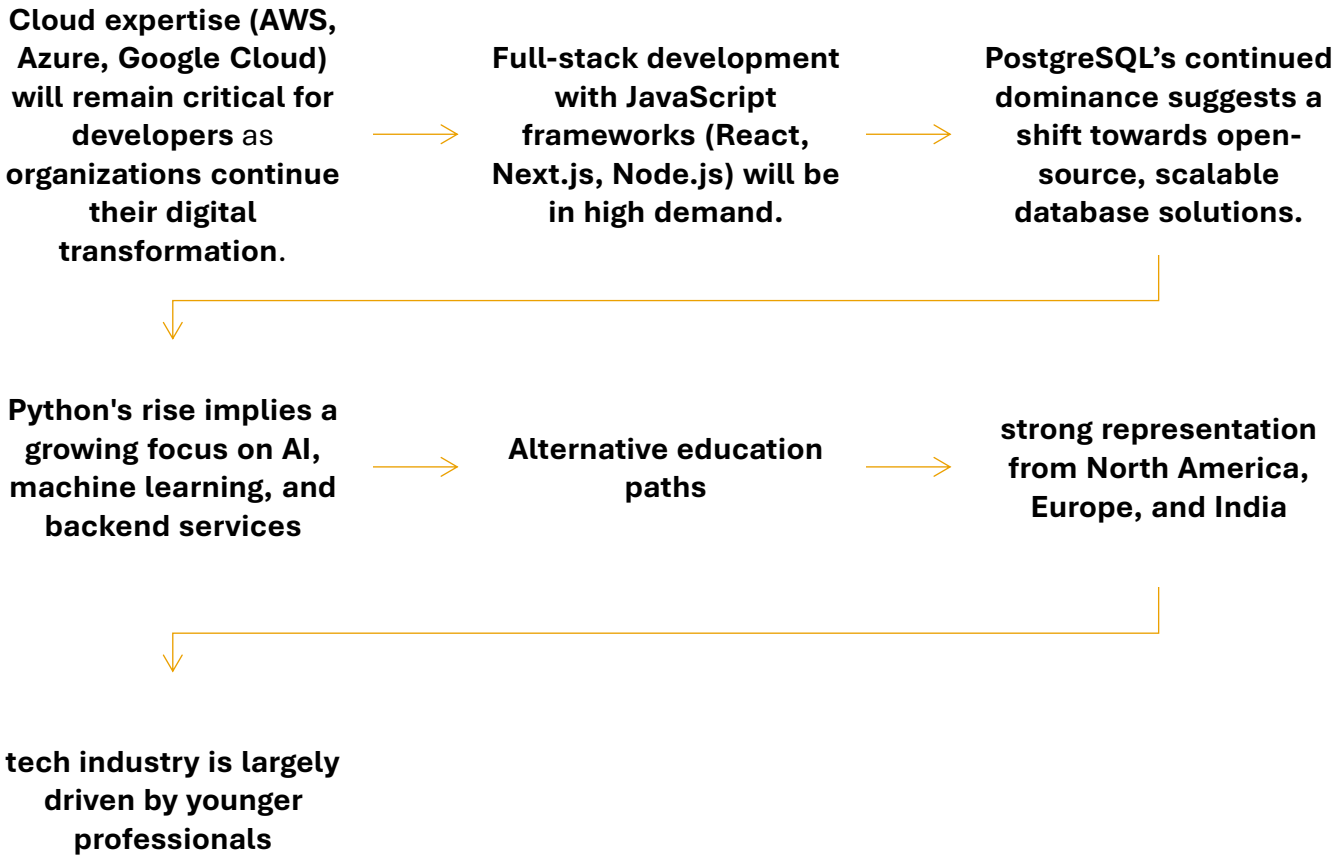


Insights from Dashboards

Findings:

- Programming Languages:
  - HTML and C# appear to be the most widely used languages in this dataset.
  - JavaScript also ranks among the top, reflecting its importance in web development.
  - Python is gaining traction, indicating its growing demand in fields like AI, data science, and backend development.
- Databases:
  - PostgreSQL leads the database rankings, indicating strong adoption.
  - Microsoft SQL Server and MySQL are also widely used.
- Platforms:
  - Cloud services dominate, with AWS, Microsoft Azure, and Google Cloud being the most referenced.
  - Web services are a key focus for developers.
- Web Frameworks:
  - Spring Boot, React, Node.js, and ASP.NET Core are among the most popular frameworks.
  - There is significant usage of full-stack combinations like Express.js with Node.js and React.
- largest group (41.3%) is aged 25-34 years35-44 years (27.3%)

Overall Findings and Implications



Conclusion

The analysis of current and future technology trends, database usage, programming languages, platforms, web frameworks, and demographics reveals key insights into the evolving tech landscape.

1. Current vs. Future Technology Trends

•Popular Programming Languages: Current trends show heavy usage of HTML, C#, and JavaScript, while future preferences lean towards C#, SQL, Python, and JavaScript. This suggests a growing demand for backend and full-stack development skills.

•Database Preferences: PostgreSQL, Microsoft SQL Server, and MySQL are the most commonly used databases, and this trend continues into the future, with PostgreSQL leading the way in anticipated demand.

•Platforms: Cloud services like AWS, Microsoft Azure, and Google Cloud dominate both current and future usage, indicating a continued reliance on cloud-based solutions.

•Web Frameworks: React, Node.js, and ASP.NET Core are widely used today, with a growing future demand for frameworks like FastAPI and Next.js, suggesting a shift toward efficient, scalable web solutions.

2. Demographics & Industry Impact

•The majority of respondents are between 25-34 years old, showing that tech remains a young, dynamic field.

•A high percentage hold Bachelor’s and Master’s degrees, reinforcing the importance of formal education, though non-traditional learning paths (bootcamps, certifications) are gaining traction.

•The largest representation comes from North America, Europe, and India, confirming these regions as global tech hubs.