STACK OVERFLOW DEVELOPER SURVEY

By Matt Haysom

25/01/25

IBM Capstone Project

REPORT CONTENTS

- The Story
- Executive Summary
- Background and Context
- Research Approach
- Key Findings
- Data Visualization
- Interactive Dashboard Visualizations
- Conclusions
- Appendices

THE STORY

- Recently, I received an exciting opportunity to join a global IT and business consulting services firm, renowned for its cutting-edge IT solutions and a team of highly skilled consultants. As a new Data Analyst at the company, I knew this was a unique chance to contribute to an organization that values staying ahead of technological changes.
- The firm's mission is to continuously adapt to the evolving technology landscape, and to do this, they rely on data-driven insights to predict future skill requirements. This year, they are focusing on identifying emerging skills that will be in demand in the coming years, ensuring they can offer the right training and talent solutions to their clients.
- My role in this important initiative was clear. I was tasked with collecting data from various sources, including job postings, training portals, and surveys. The goal was to gather the top programming skills that are expected to see a surge in demand. As the technology landscape is constantly changing, the insights I gathered would play a crucial part in helping the company stay competitive and prepare their workforce for the skills of tomorrow.
- Once I had the data, the real work would begin. I would analyze the information to uncover trends and insights that could reveal which skills are emerging and which technologies are gaining traction. This analysis would be the foundation of the firm's report on future skill requirements.
- Each step of this process, from data collection to analysis, is documented in a series of notebooks and supporting files within this project repository. By following this structured approach, I aim to ensure that the findings are clear, actionable, and valuable for our organization’s strategy moving forward.
- This project marks the beginning of a thrilling journey, where data will not only inform decisions but help shape the future of our industry.

EXECUTIVE SUMMARY

- Objective: Data Analysis of Current Technology Usage and Future Direction.
- Overview of Methodology:
  - Data Collection Process
  - Data Examination
  - Visual Representation of Data
- Presentation of Results with Supporting Graphs and Trends.
- Interpretation of Findings and Implications.
- Final Data Analysis Conclusions.

BACKGROUND AND CONTEXT

- The ‘Stack Overflow Developer Survey’ is the most extensive and far-reaching study of the global developer community.
- The findings, however, do not fully capture the entire spectrum of the developer population, with some groups underrepresented.
- The survey gathered insights from nearly 90,000 developers worldwide, offering a vast data pool.
- It highlights emerging trends that can help forecast the future direction of the developer landscape.
- The survey also provides a detailed profile of developers across various regions, showcasing their diverse backgrounds and practices.

RESEARCH APPROACH

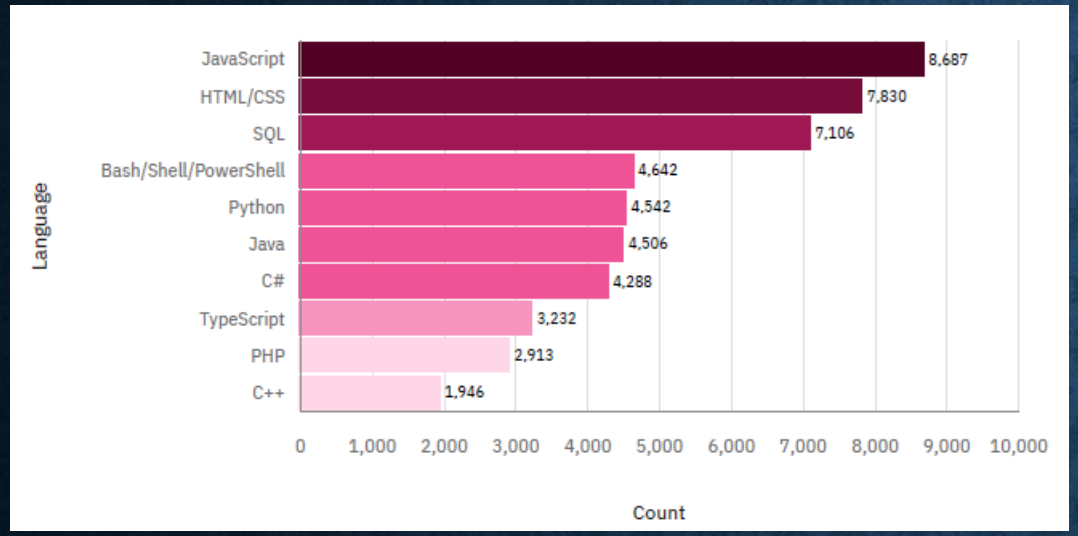
- Explore the Pre-Prepared Survey Data Set Using the Following Methodologies:
  - Familiarization with the dataset was achieved through Data Analysis including:
  - Dataset Exploration
  - Webscraping
  - Data Wrangling
  - Exploratory Data Analysis
  - Data Visualization
  - Building a Dashboard using IBM Cognos Analytics to extract meaningful patterns and trends
- The survey provided valuable insights into global trends and practices pertinent to the topic under review.

KEY FINDINGS

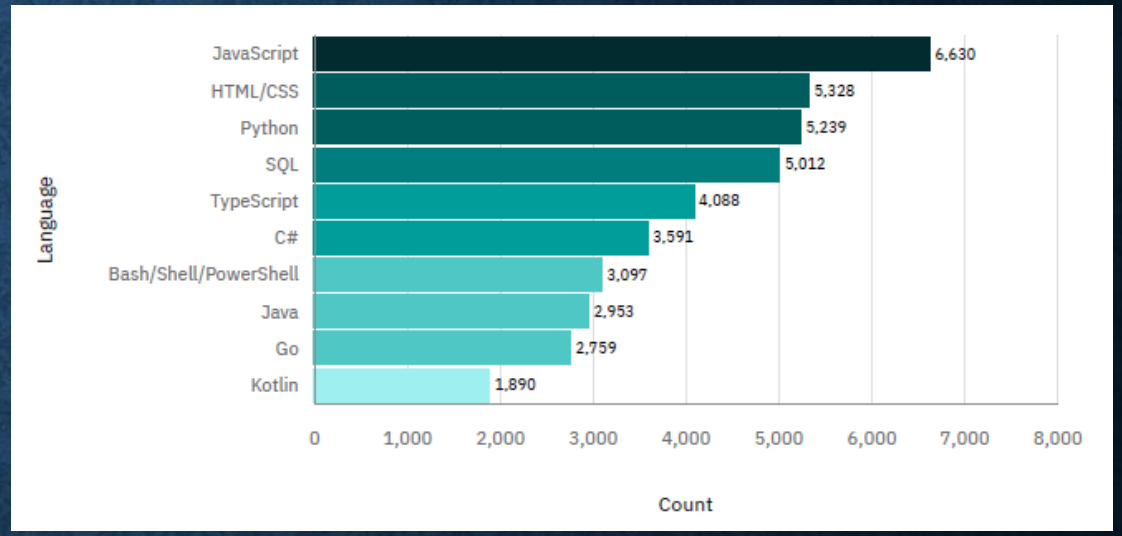
- Approximately 90% of survey respondents are male, with higher female representation among students than in professional roles, particularly in regions like the US, India, and the UK.
- JavaScript continues to be the most widely used programming language globally.
- Python has surpassed Java to become the 5th most preferred programming language, showing significant growth as the fastest-growing major language.
- jQuery remains the most popular web framework, while React.js has overtaken Angular in developer usage this year.
- Around 75% of professional developers worldwide hold at least a bachelor's degree, consistent with previous trends, and 75% of respondents in professional developer roles are under 35 years of age.

DATA VISUALIZATION 1: PROGRAMMING LANGUAGE TRENDS

Top Ten Languages Worked with This Year



Anticipated Language Trends Next Year

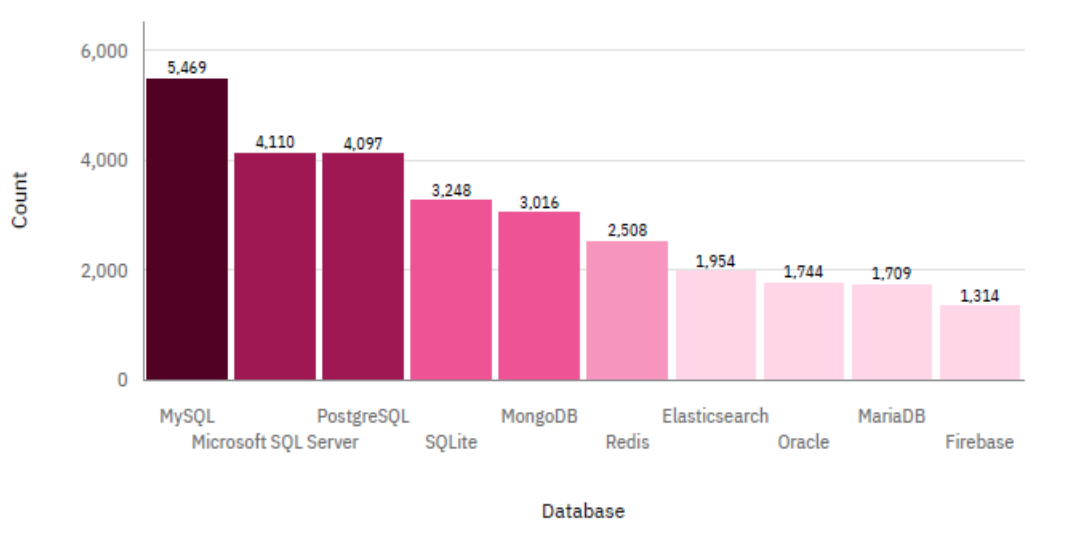


Implications:

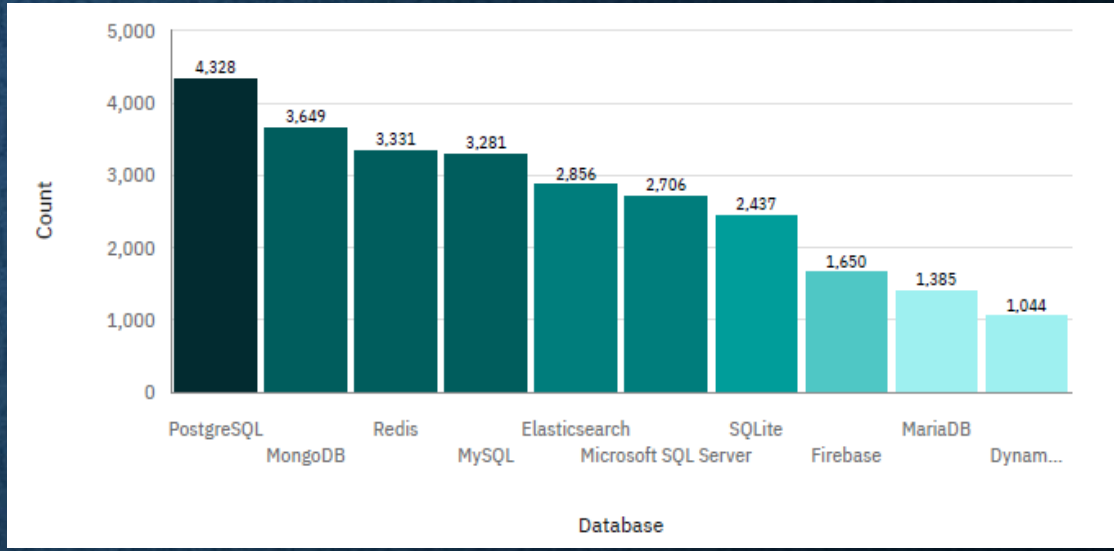
- Though Java Script is still anticipated to be the preeminent working programming language, python is seen to be becoming more popular.
- Data shows that the rise of Python is likely at the expense of aging languages like PowerShell and C++.

DATA VISUALIZATION 2: DATABASE TRENDS

Top Ten Databases Worked with This Year



Anticipated Future Database Demand

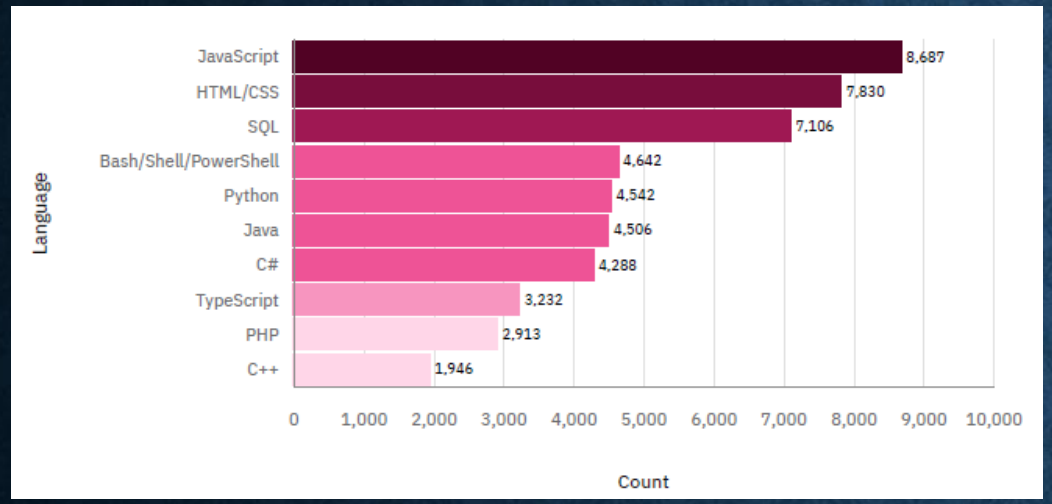


Findings:

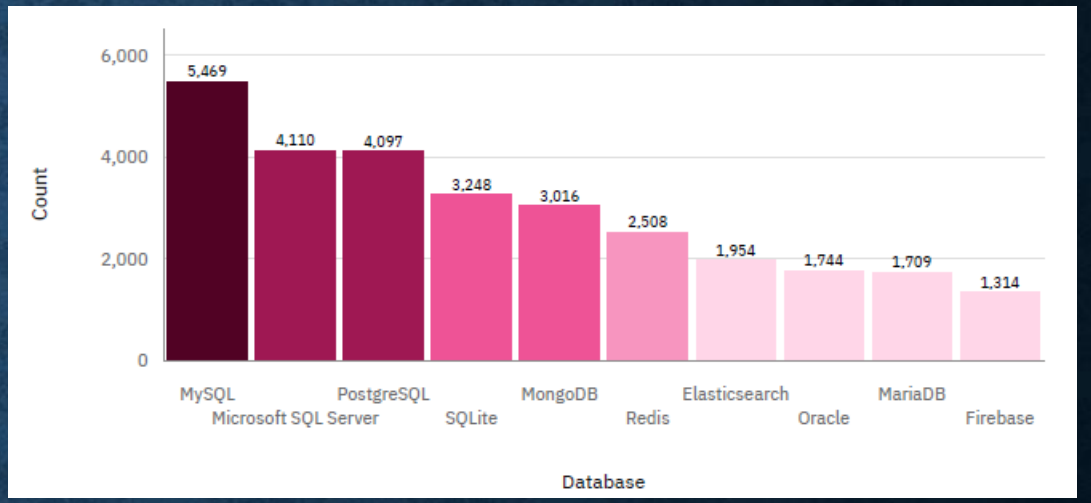
- PostgreSQL seems well favoured to take a more dominant role going forward.
- Data also shows that MongoDB and Redis will also feature prominently ostensibly at the expense of MySQL, MSSQL Server and others.

INTERACTIVE DASHBOARD 1: CURRENT TECHNOLOGY USAGE

Top Ten Languages Worked with This Year

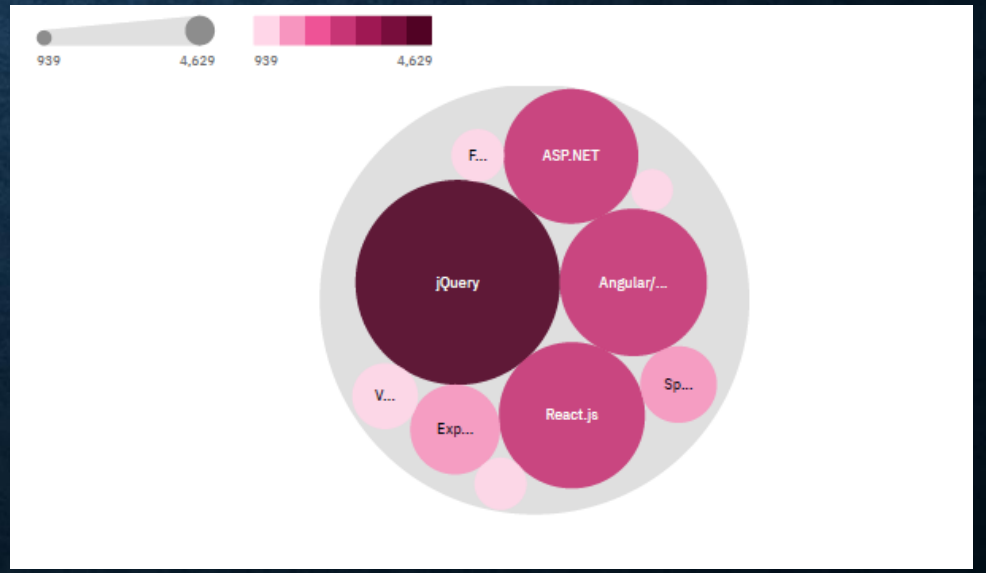


Top Ten Databases Worked With



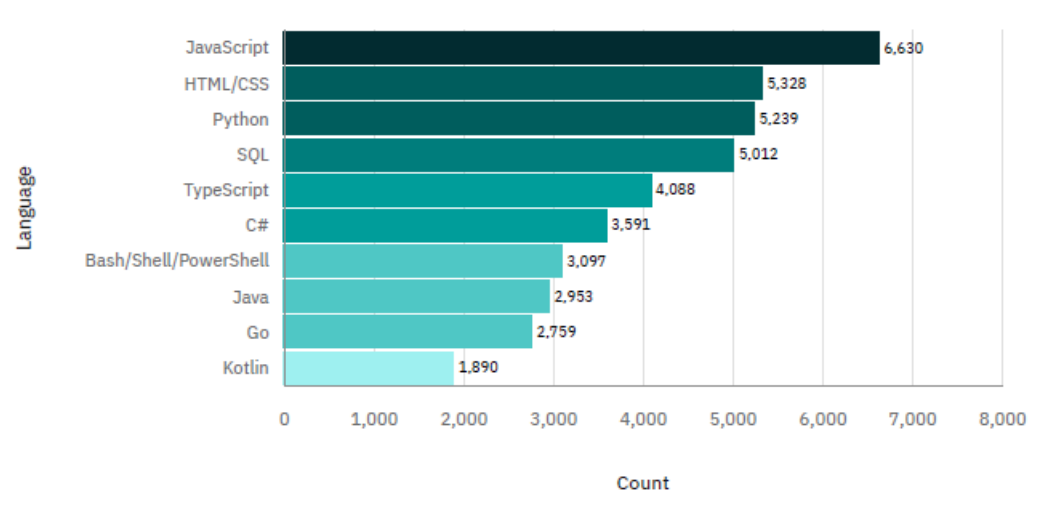
Top Platforms Worked With

Top Ten Web Frameworks Worked With

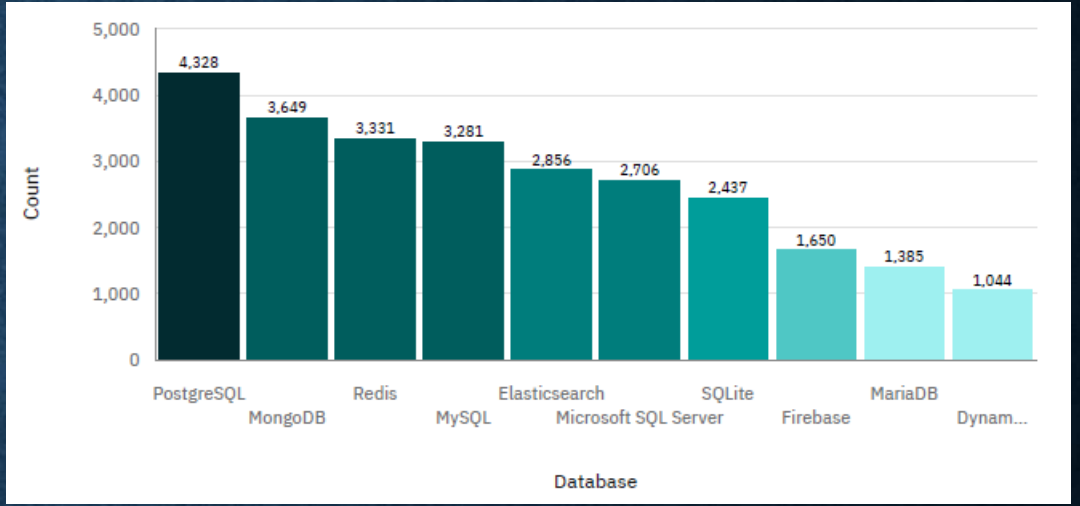


INTERACTIVE DASHBOARD 2: FUTURE TECHNOLOGY TRENDS

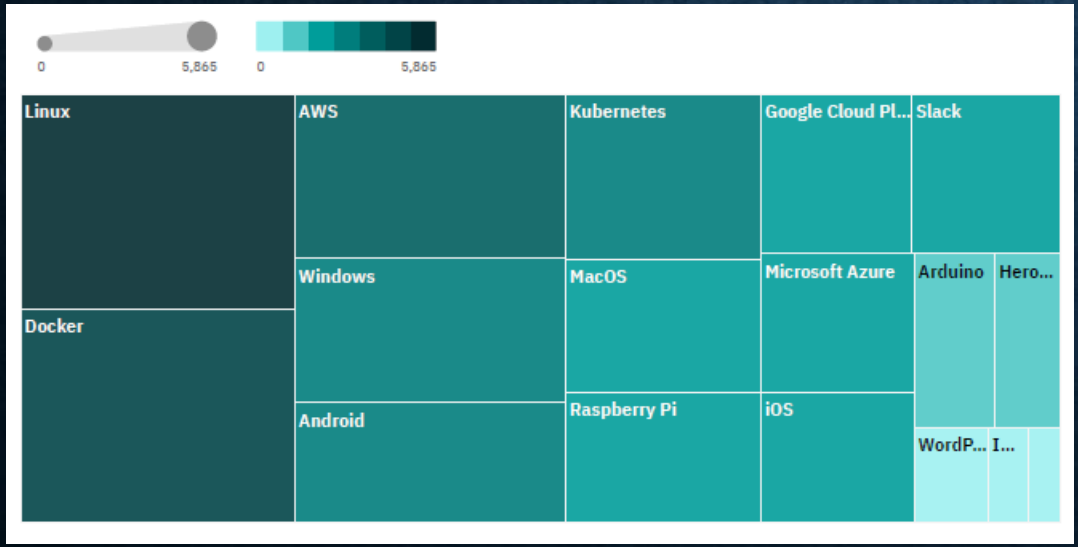
Top Ten Languages Desired In the Future



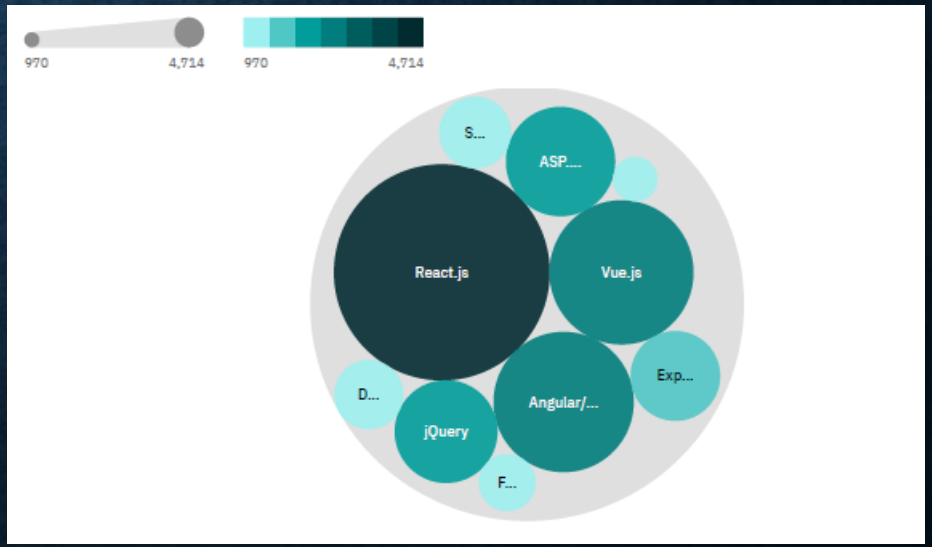
Top Ten Databases Desired in the Future



Top Platforms Desired Next Year

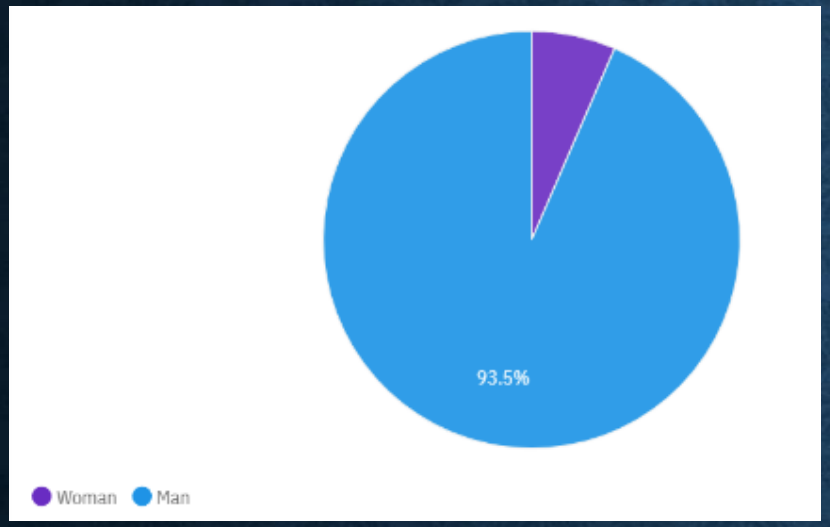


Top Ten Web Frameworks Worked With



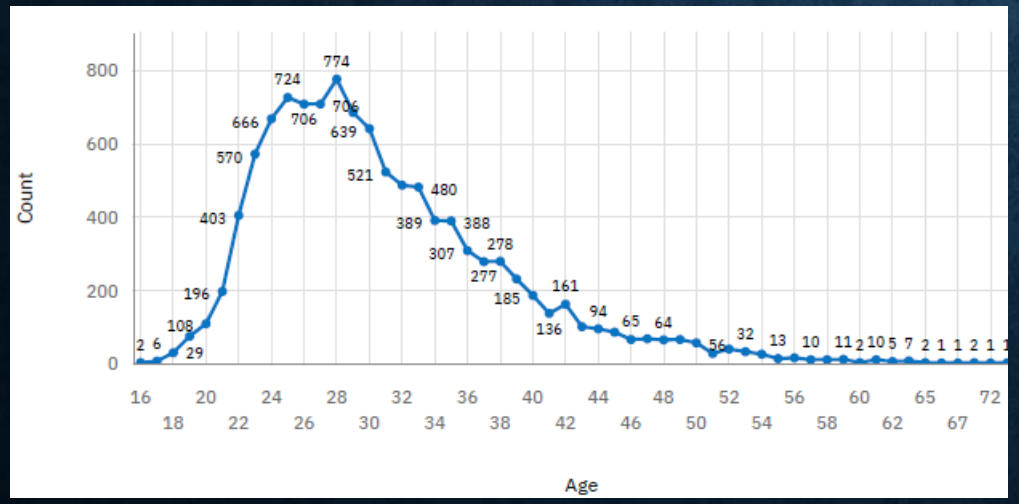
INTERACTIVE DASHBOARD 3: DEMOGRAPHICS

Respondent Count by Gender



Respondent Count by Count

Respondent Count by Age



Respondent Education Level by Gender

CONCLUSIONS

- Technology Usage:
  - JavaScript will continue to be a popular programming language
  - Python is growing rapidly in popularity and some say has already overtaken JavaScript already as the leading language (1)
- Database Preference:
  - MySQL is the leading Database
  - Microsoft SQL and SQLite are losing market share
  - Others such as MongoDB are becoming more popular
- Demographics
  - Predominantly a male-oriented industry, the data seems to bear out the long held belief that Women are facing huge challenges when trying to break into tech and data analytics-related roles(2)

APPENDICES

1. https://dataleum.com/why-python-overtook-javascript-as-the-most-popular-language-on-github/

2. https://www.linkedin.com/pulse/from-glass-ceiling-cliffs-decade-long-progress-women-data-cisneros-ylz5c

3. https://github.com/Mattph1963/GitHub-Repository-IBM-Data-Analyst-Capstone-Project