

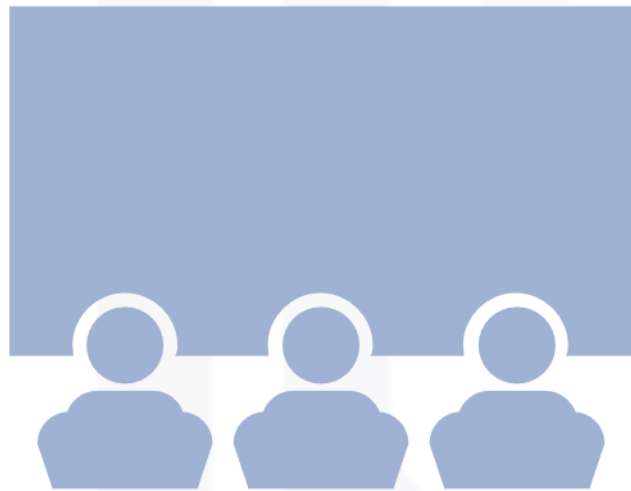


Data Analyst Capstone Project

Cindy Thai

05/31/2023

OUTLINE



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

EXECUTIVE SUMMARY



Stack Overflow's Annual Developer Survey is one of the largest and most comprehensive survey of people in the IT Industry from around the world, this survey covers basic demographic metrics as well as technology preferences in various fields. This survey used to analyze data from comes from the 2019 iteration of this survey.

As a data analyst for a global IT and business consulting firm, I am tasked with analyzing the data to help determine emerging trends and identify insights that answer the following questions:

- What are the top programming languages in demand?
- What are the top database skills in demand?
- What are the most popular IDEs?

By collecting data from various sources, wrangling and visualizing the data, I want to share insights from the data to help my business make data-driven decisions.

INTRODUCTION



About

This capstone project analyzes current trends in programming languages, technologies, and basic demographics based on a 2019 survey on a popular open-forum site, Stack Overflow. In order to keep pace with changing and evolving technologies and remain competitive, our organization must be keen to trends in software and adapt our business practices accordingly.

Purpose

- Collect top programming skills most in demand from various sources (job postings, training portals, surveys)
- Identify key skill requirements for future software engineers
- Make data ready for analysis using data wrangling techniques
- Create a dashboard for easy data manipulation and presentation
- Present insights and findings to key stakeholders

Audience

IT Industry Professionals, Business Executives, Human Resource Officers, and anyone who potentially want to pursue a career in IT

METHODOLOGY



- Data Collection Sources
 - Stack Overflow Developer 2019 Survey
 - Github Job Postings
 - Programming Languages Annual Survey
- Data Exploration / Exploratory Data Analysis
- Data Cleaning and Wrangling
- Data Visualization Using Python
- Data Visualization Using Dashboards
- Presentation

RESULTS

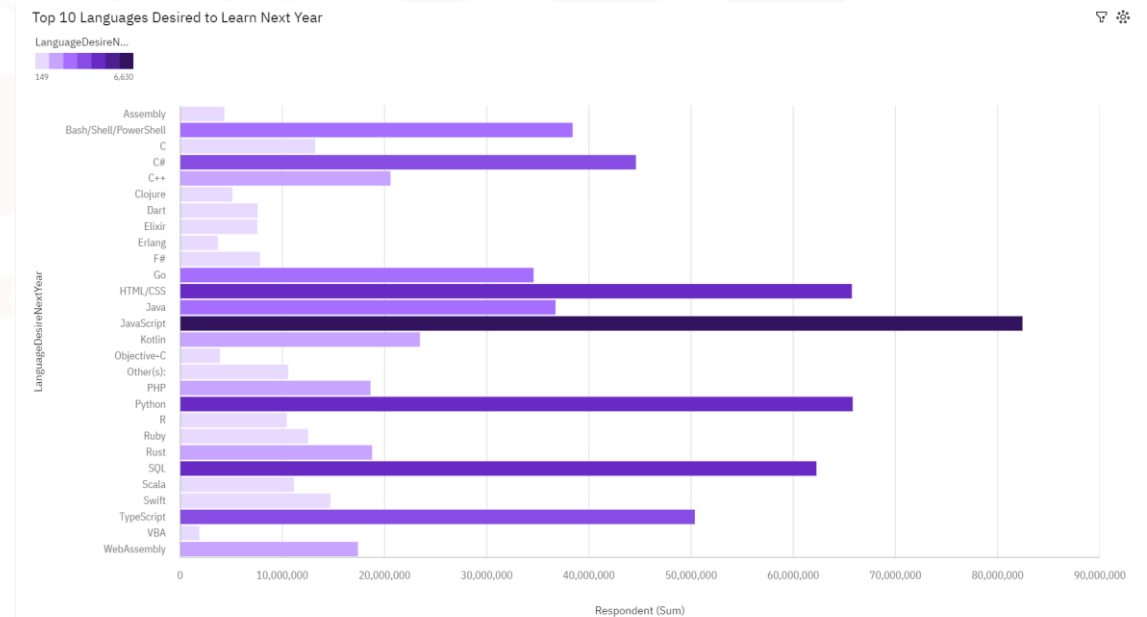
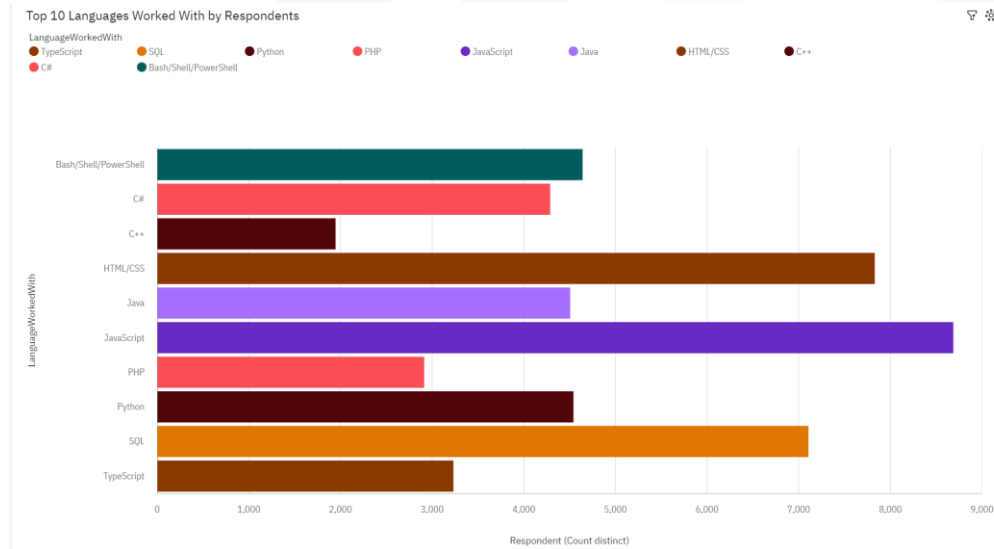
	Respondent...	LanguageWorkedWith	LanguageDesireNextYear	DatabaseWorkedWith	DatabaseDesireNextYear
1.	25142	R			
2.	25142	Python			
3.	25142	C	C		
4.	25142	Go	Python		
5.	25142	C++	C++		
6.	25142	HTML/CSS	R		
7.	25142	PHP			
8.	25142	Bash/Shell/PowerShell	Bash/Shell/PowerShell	PostgreSQL	PostgreSQL
9.	25141	C#	F#		Microsoft SQL Server
10.	25141	JavaScript			
11.	25141	TypeScript			
12.	25141	HTML/CSS			Redis
13.	25141	Java			
14.	25141	Python			
15.	25141	SQL			

1 - 100 / 72277 < >

PROGRAMMING LANGUAGE TRENDS

Current Year

Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Top 5 Programming Languages include: Javascript, HTML/CSS, SQL, Bash/PowerShell and Python
- Java was projected to outrank Python in popularity the next year
- Mobile Development was projected to rise the next year

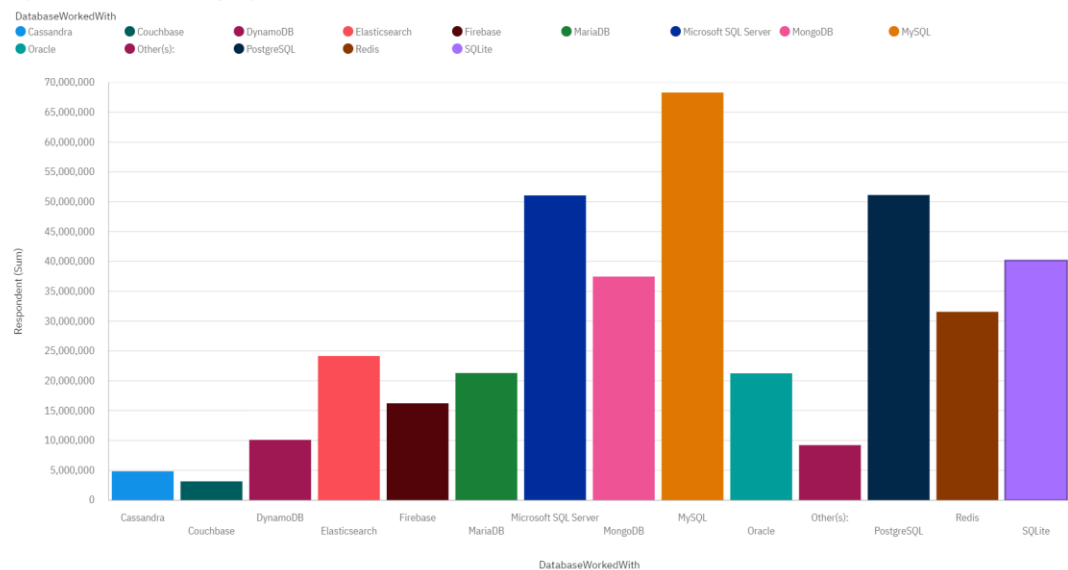
Implications

- Web and Mobile Development is among one of the most desired reasons for programming
- SQL still serves as the backbone for Big Data Programming
- With the rise of AI and Machine Learning, Python is among the most popular language for data science

DATABASE TRENDS

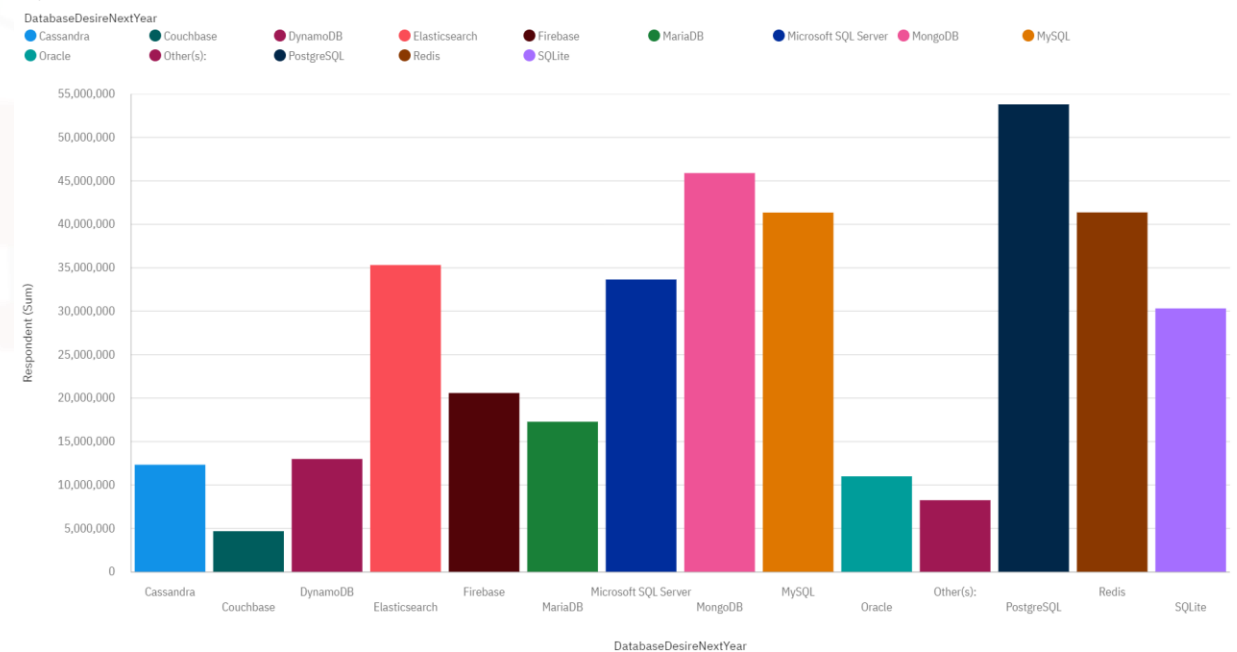
Current Year

Top 10 Databases Worked With by Respondents



Next Year

Top 10 Databases Desired to Learn Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- MySQL was the most popular database management system this year
- Microsoft, PostgreSQL, and SQLite follow close behind
- PostgreSQL was projected to beat MySQL in popularity next year

Implications

- Open Source Databases are still preferred amongst companies
- NoSQL will be the biggest source for storing non relational databases
- Databases will continue to be used by all programmers alike

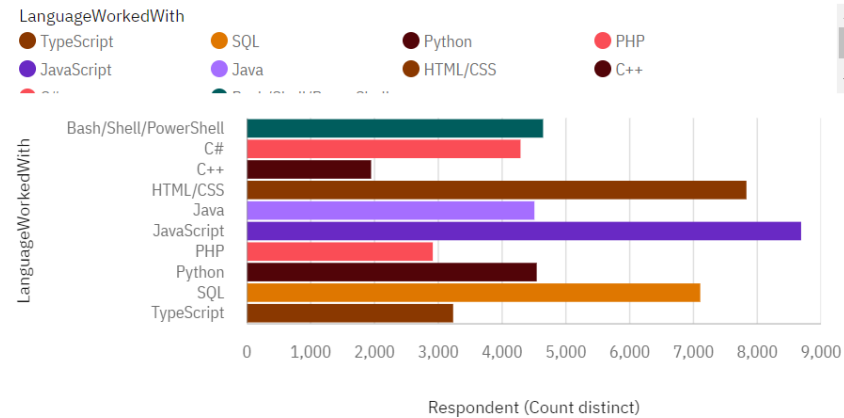
DASHBOARD



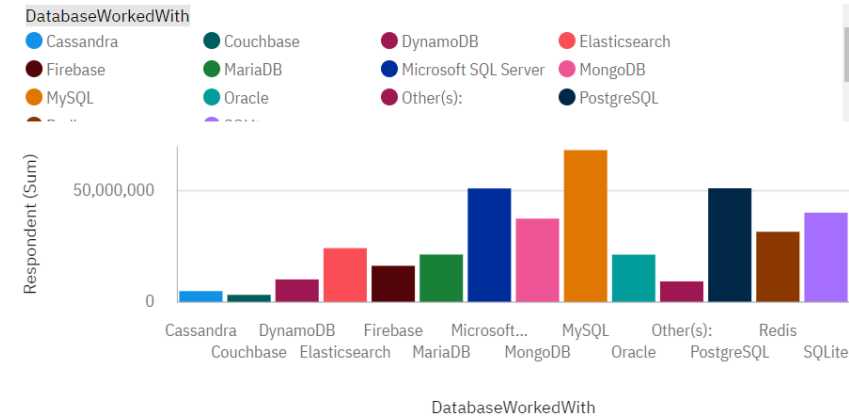
https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FDashboard-Assignment&action=view&mode=dashboard&subView=model000001887347075a_00000000

DASHBOARD TAB 1

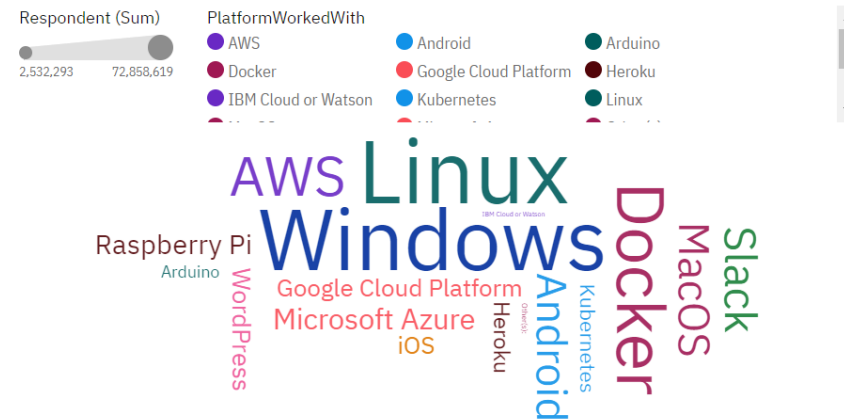
Top 10 Languages Worked With by Respondents



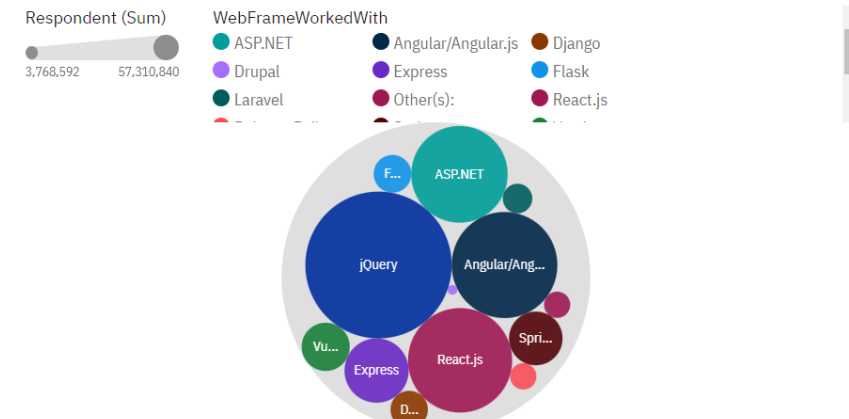
Top 10 Databases Worked With by Respondents



Top 10 Platforms Worked With by Respondents

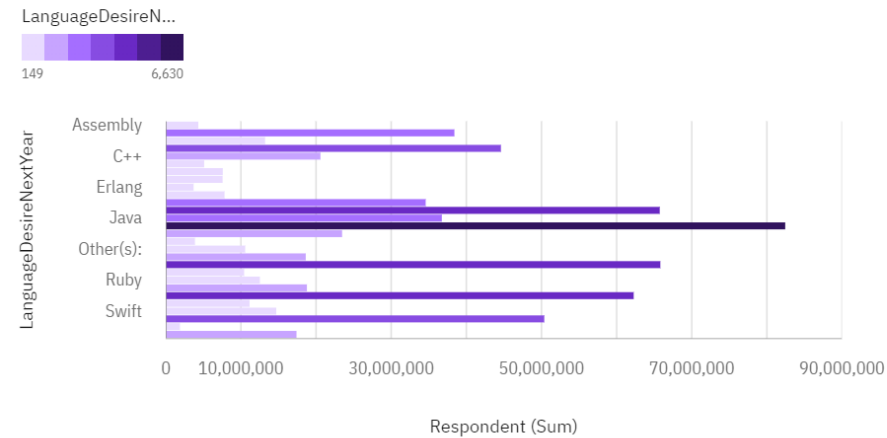


Top 10 Web Frames Worked With by Respondents

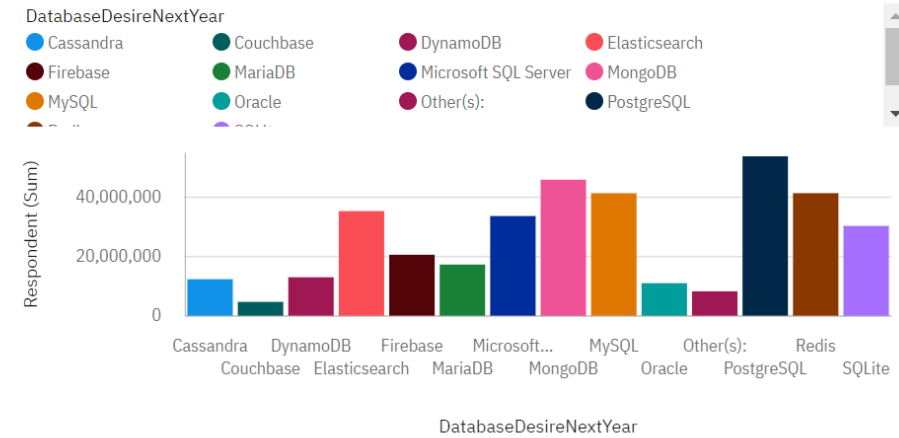


DASHBOARD TAB 2

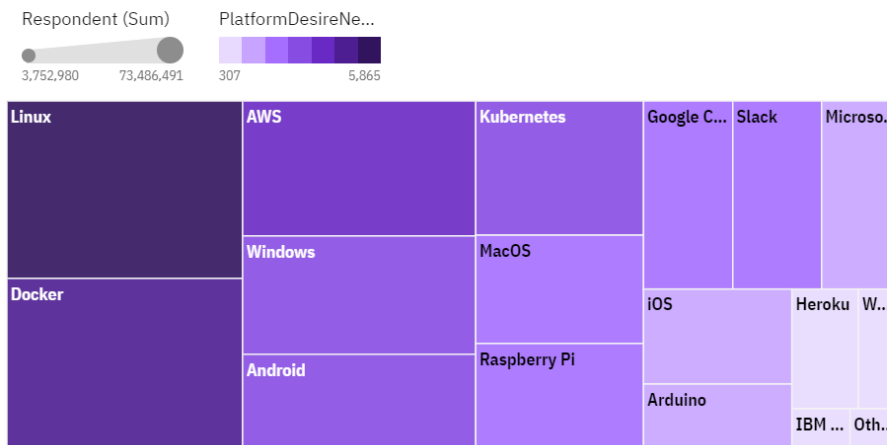
Top 10 Languages Desired to Learn Next Year



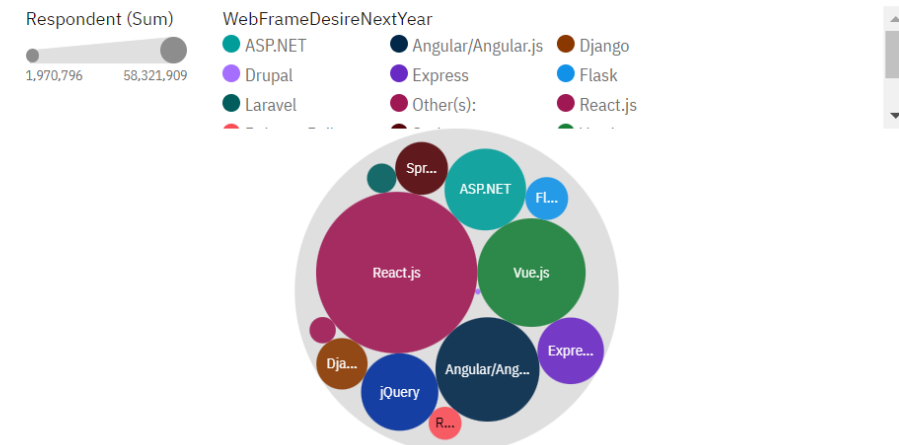
Top 10 Databases Desired to Learn Next Year



Platforms Desired to Learn Next Year



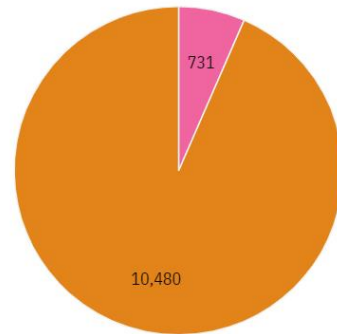
Top 10 Web Frames Desired to Learn Next Year



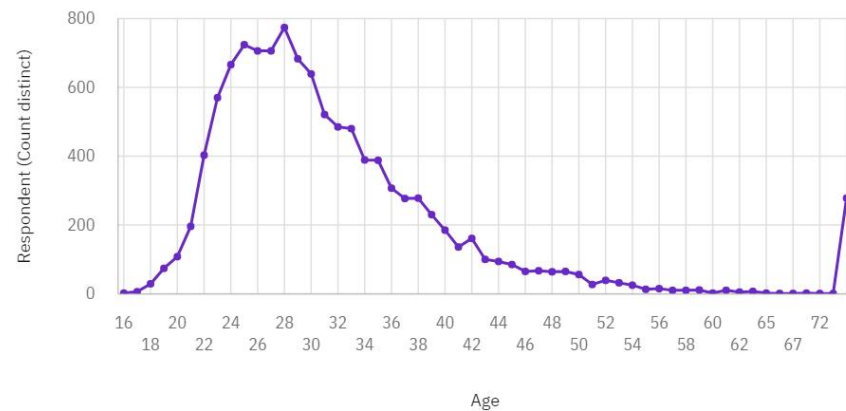
DASHBOARD TAB 3

Respondent Classified by Gender

Gender
● Woman ● Man

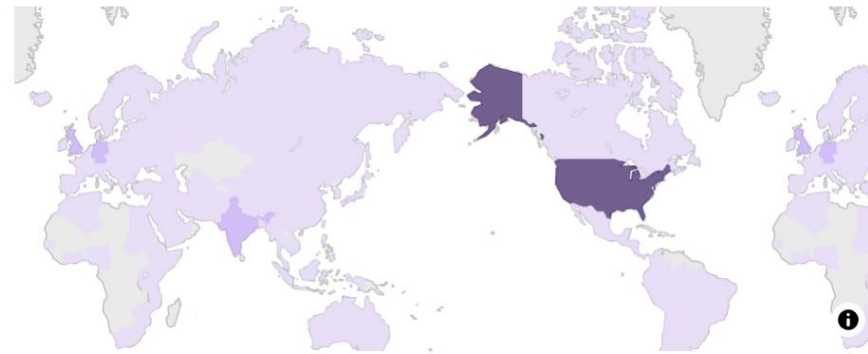


Respondent Count by Age



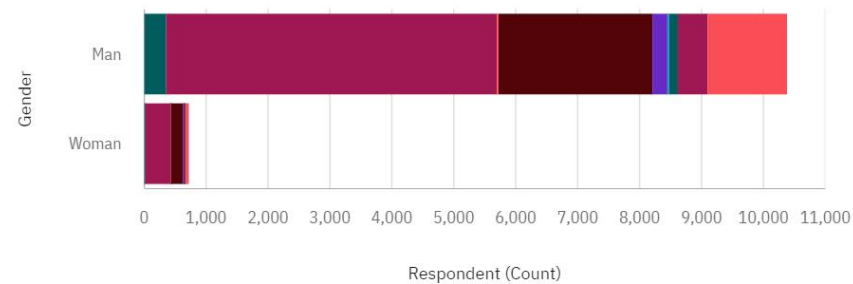
Respondent County for Countries

Respondent (Count)
1 3,058



Respondent Count by Gender, Classified by Formal Education Level

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



DISCUSSION



- Although it was likely easier to focus on one language and one platform previously, in this day, we need to diversify our portfolio
- We need to bridge the gap between developed and developing countries
- We need to eliminate gender discrimination in the industry

OVERALL FINDINGS & IMPLICATIONS

Findings

- Technology Changes at an alarming rate year over year
- We can focus our efforts on expanding into more countries such as India and US
- There are less women working in the field overall
- Cloud Computing has been a hot topic

Implications

- Companies need to adapt to these changes
- We need to expand our efforts into new countries that are adapting these technologies
- We develop at faster rates year over year and we will likely need a mix of multiple skills

CONCLUSION

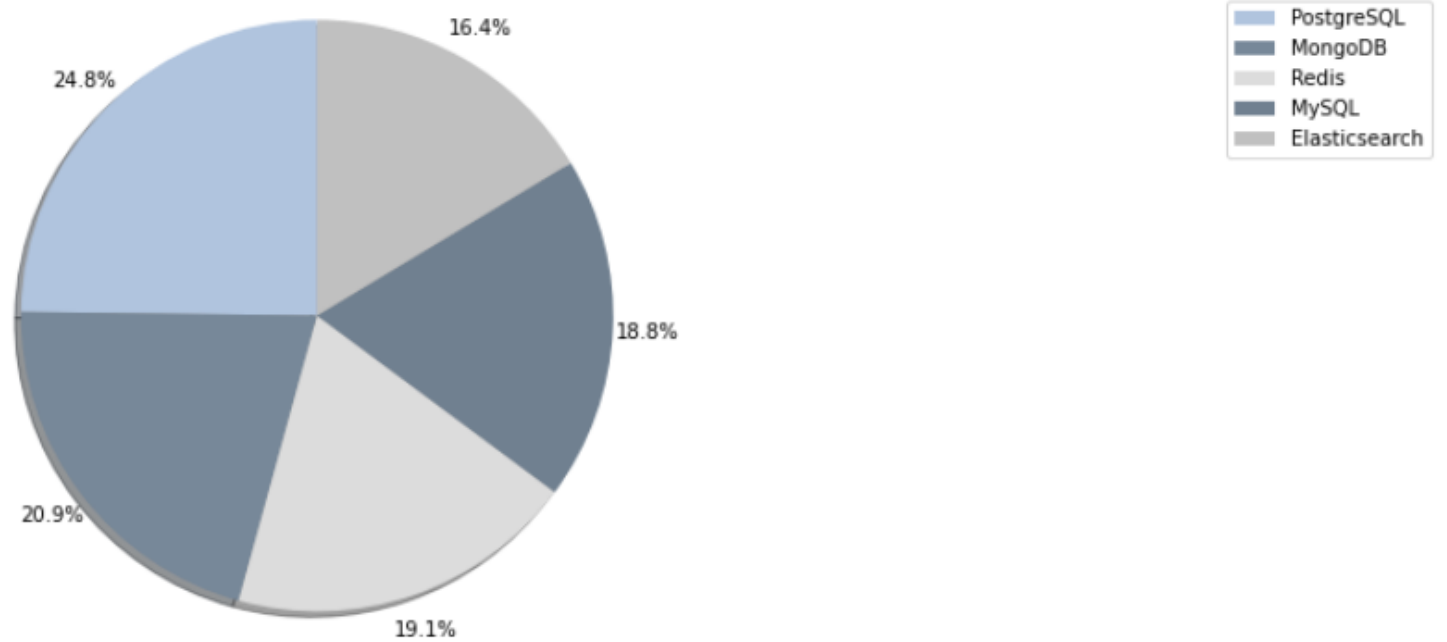


- We analyze survey data to analyze technology trends for this current year and the year after
- We discovered patterns with demographics data
- We found the most popular programming languages, databases, platforms, and IDEs
- In the future, we can incorporate machine learning algorithms to conduct predictive analysis

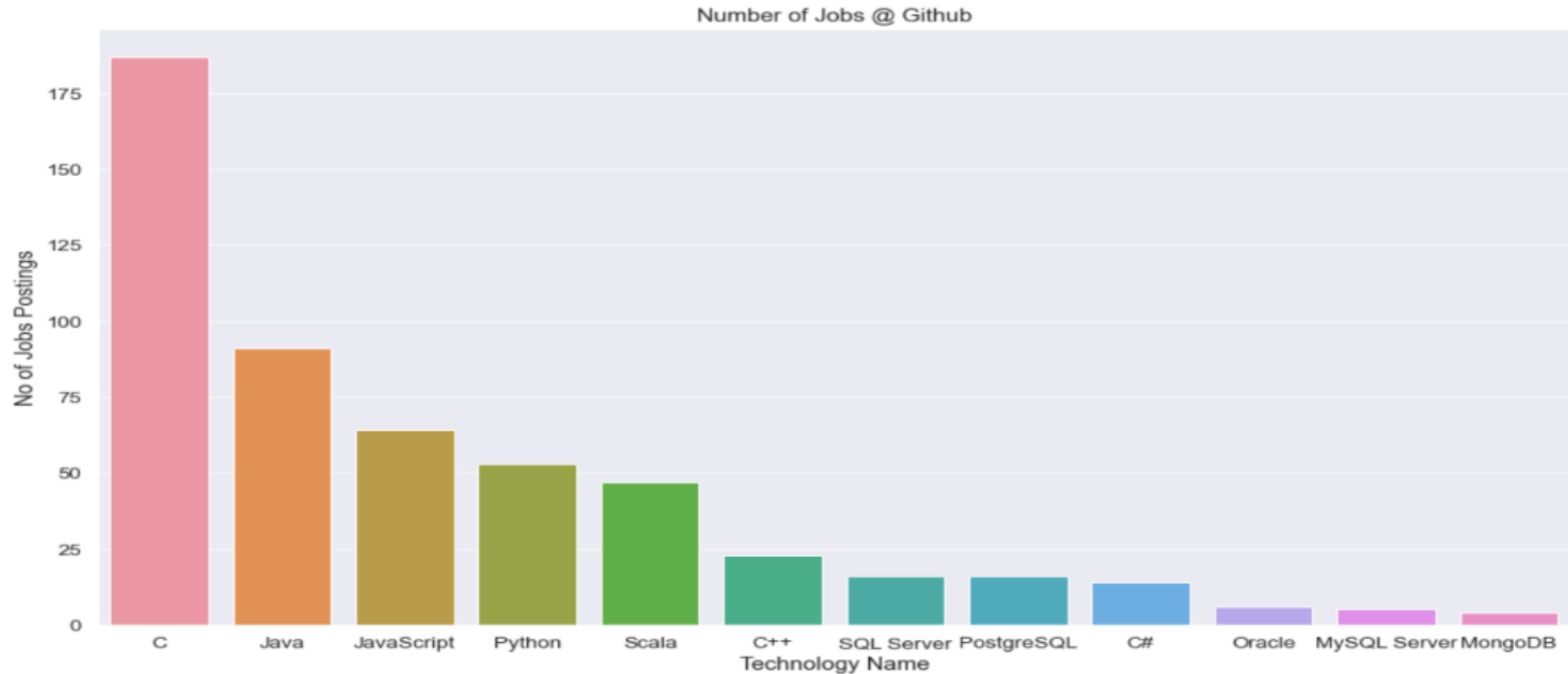
APPENDIX



Top 5 Databases Respondents Wish To Learn



JOB POSTINGS



POPULAR LANGUAGES

