

Data Findings Report



Emerging Skills in the Technology

Outline



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

Executive Summary



AIM OF ANALYSIS:

Find the most in demand programming languages and database skills.

KEY FINDINGS:

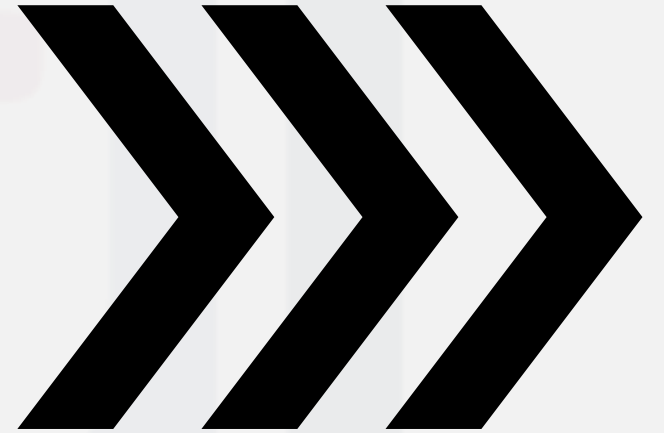
- Top programming languages: Javascript, Python, HTML/CSS, SQL, TypeScript
- Top databases: PostgreSQL, MongoDB, Redis, MySQL, Elasticsearch
- SQL programming is on a downward trend; PostgreSQL is an exception becoming the most in demand database next year.
- NoSQL programming is becoming more popular, particularly in the database skills market.

CONCLUSION:

The industry remains fast-moving, previously popular technology is being overtaken in demand by newer options. Analysis must be integral part of ongoing business plans to keep pace in a competitive market.

Introduction

- The aim of this report is to present all team members with the most up-to-date analysis of our emerging technologies data which will allow us to keep up with the fast-moving industry.
- The report will give you the key analysis needed in order to stay relevant with our competitors by identifying the key trends for future skills requirements.



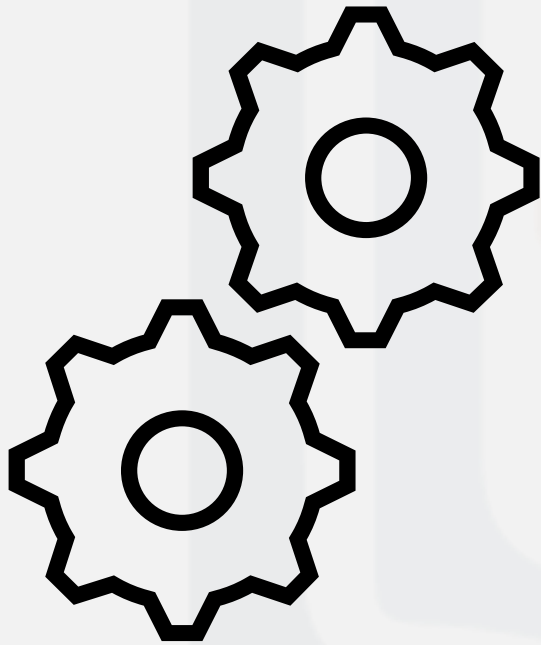
Introduction



The key questions I aim to answer with the data analysis are:

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

Methodology




The original source for the data used in this report was collected by a user survey from Stack Overflow:

<https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>, used under a ODbL: Open Database License.

When analyzing the data I used the CSV file: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv

Methodology

- In order to explore the data the CSV file was stored in a Pandas data frame.
- The wrangling process made sure the data was clean by removing duplicates and missing values and normalizing the data so it was ready for analysis.
- Pandas methods were used to identify data distribution, correlated variables and statistical values (e.g. mean average).

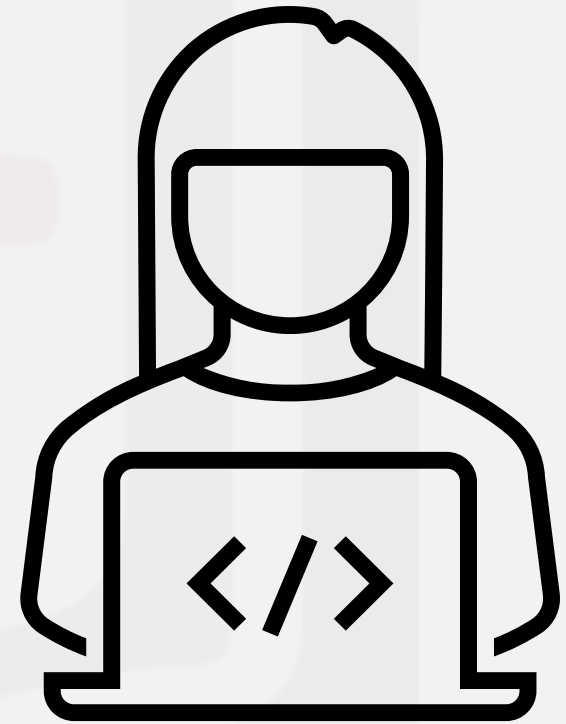


Methodology

The most relevant variables from the dataset for the analysis we are doing are:

- LanguageWorkedWith
- LanguageDesiredNextYear
- DatabaseWorkedWith
- DatabaseDesiredNextYear

By comparing and contrasting these key variables, I was able to gather the results.



Results



- Having completed the analysis of the data, I can now present you with the results in relation to the questions posed earlier.
- By comparing the popularity of programming languages and databases from this year to next, I have identified current technology trends which will help us keep pace with the industry over the coming year.

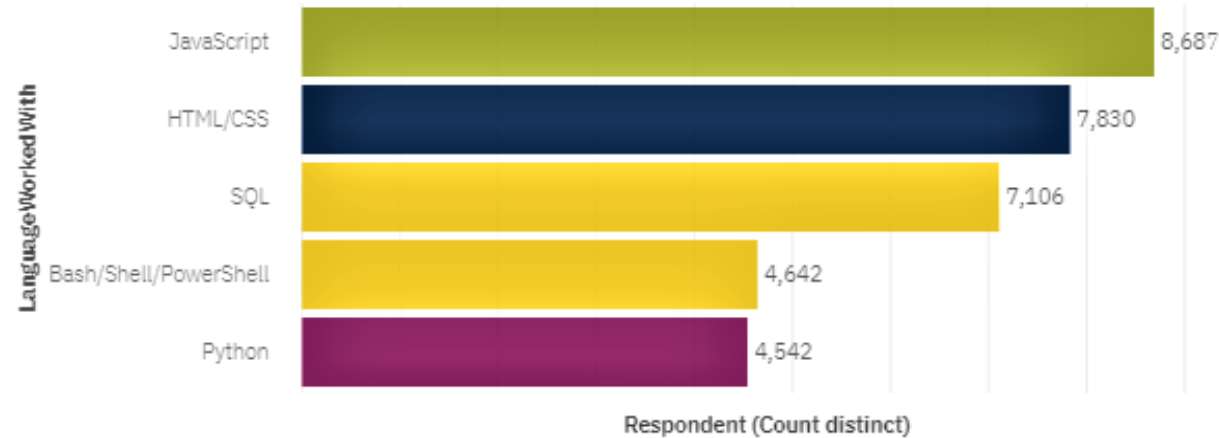
Trends: Programming Languages

Current Year

Top 5 Languages Worked With

LanguageWorkedWith

● Bash/Shell/PowerShell ● HTML/CSS ● JavaScript ● Python ● SQL

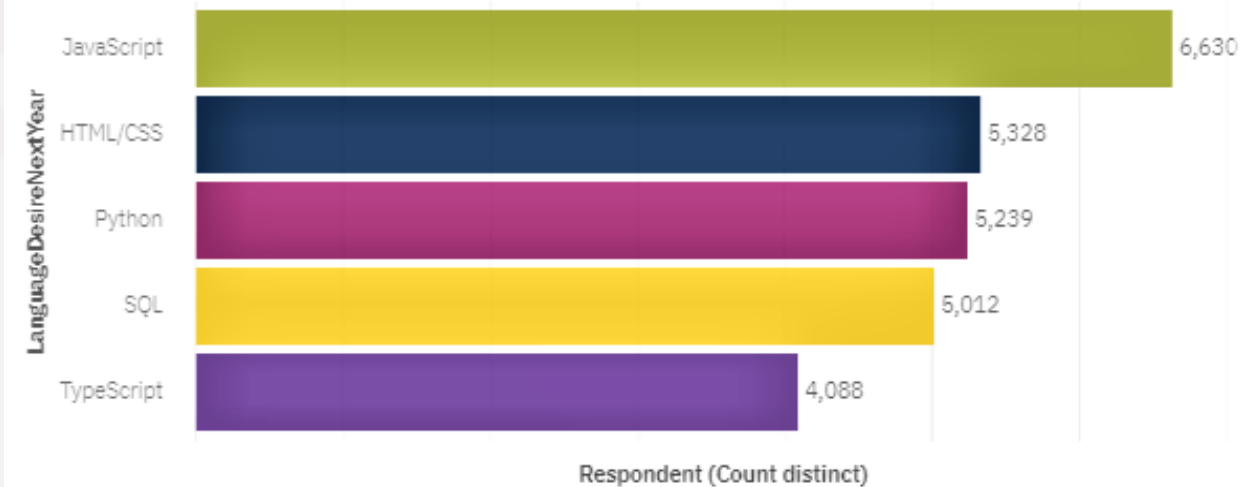


Next Year

Top 5 Languages Desired Next Year

LanguageDesireNextYear

● HTML/CSS ● JavaScript ● Python ● SQL ● TypeScript



Findings & Implications: Programming Languages

Findings

- JavaScript will remain the most popular language to work with into next year, but it will be slightly less in demand than this year.
- HTML/CSS will remain the second most popular programming language.

Implications

- JavaScript and HTML/CSS are still high priority programming languages for the coming year
- We must continue to diversify our interests in other popular languages which may take over in the next few years.

Findings & Implications: Programming Languages



JavaScript stays #1
- 24% Next Year



HTML/CSS stays #2
- 32% Next Year

Findings & Implications: Programming Languages

Finding

Python has overtaken SQL to become next year's third most popular programming language.

Implication

Python is a language that is growing in popularity and if that trend continues it may become the top language in the next few years.

Findings & Implications: Programming Languages



Python up to #3
+ 15% Next Year

SQL down to #4
- 30% Next Year

Findings & Implications: Programming Languages

Finding

TypeScript is growing in popularity, knocking Bash/Shell/Powershell out of the Top 5 to become the fifth most popular language next year.

Implication

TypeScript must take more of a focus in the next year as it grows whereas less time must be spent on Bash/Shell/Powershell as it is becoming less popular.

Findings & Implications: Programming Languages



TypeScript up to #5
+ 26% Next Year

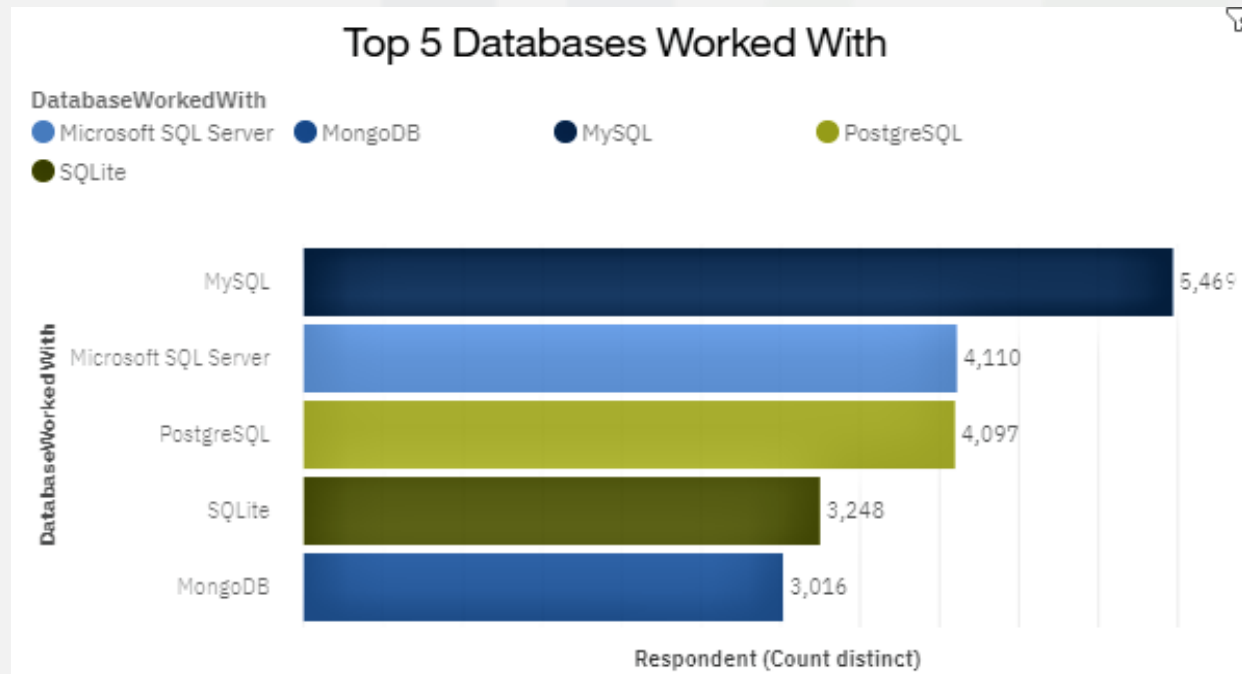


BASH
THE BOURNE-AGAIN SHELL

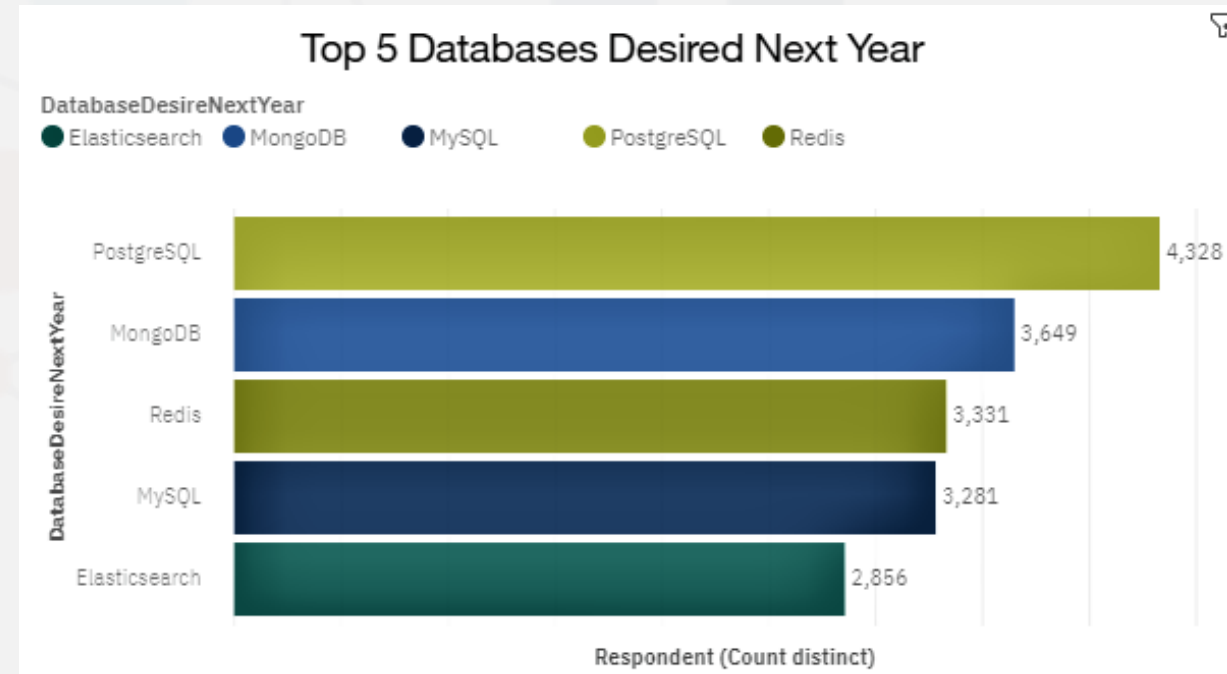
Bash/Shell/Powershell out
of Top 5
- 34% Next Year

Trends: Databases

Current Year



Next Year



Findings & Implications: Databases

Finding

PostgreSQL and MongoDB will be the two most popular databases to work with next year.

Implication

Both databases are rising in popularity and must be a focus for the emerging technology industry going forward.

Findings & Implications: Databases



PostgreSQL

PostgreSQL now #1
+ 6% Next Year



mongoDB®

MongoDB now #2
+ 18% Next Year

Findings & Implications: Databases

Finding

MySQL and Microsoft SQL Server are both significantly less in demand next year, out of proportion with the growth of PostgreSQL.

Implication

Despite PostgreSQL being the most popular database next year, in general SQL databases less in demand overall.

Findings & Implications: Databases



MySQL down to #4
- 40% Next Year



MSFT SQL Server out of Top 5
- 34% Next Year

Findings & Implications: Databases

Findings

- Redis and Elasticsearch have both grown in demand for next year
- Three of the top five databases are now NoSQL

Implications

- NoSQL databases are becoming more popular than SQL databases
- The industry looks likely to continue trending in this direction in the following year

Findings & Implications: Databases



redis

Redis up to #3
+ 33% Next Year



elasticsearch

Elasticsearch up to #5
+ 46% Next Year

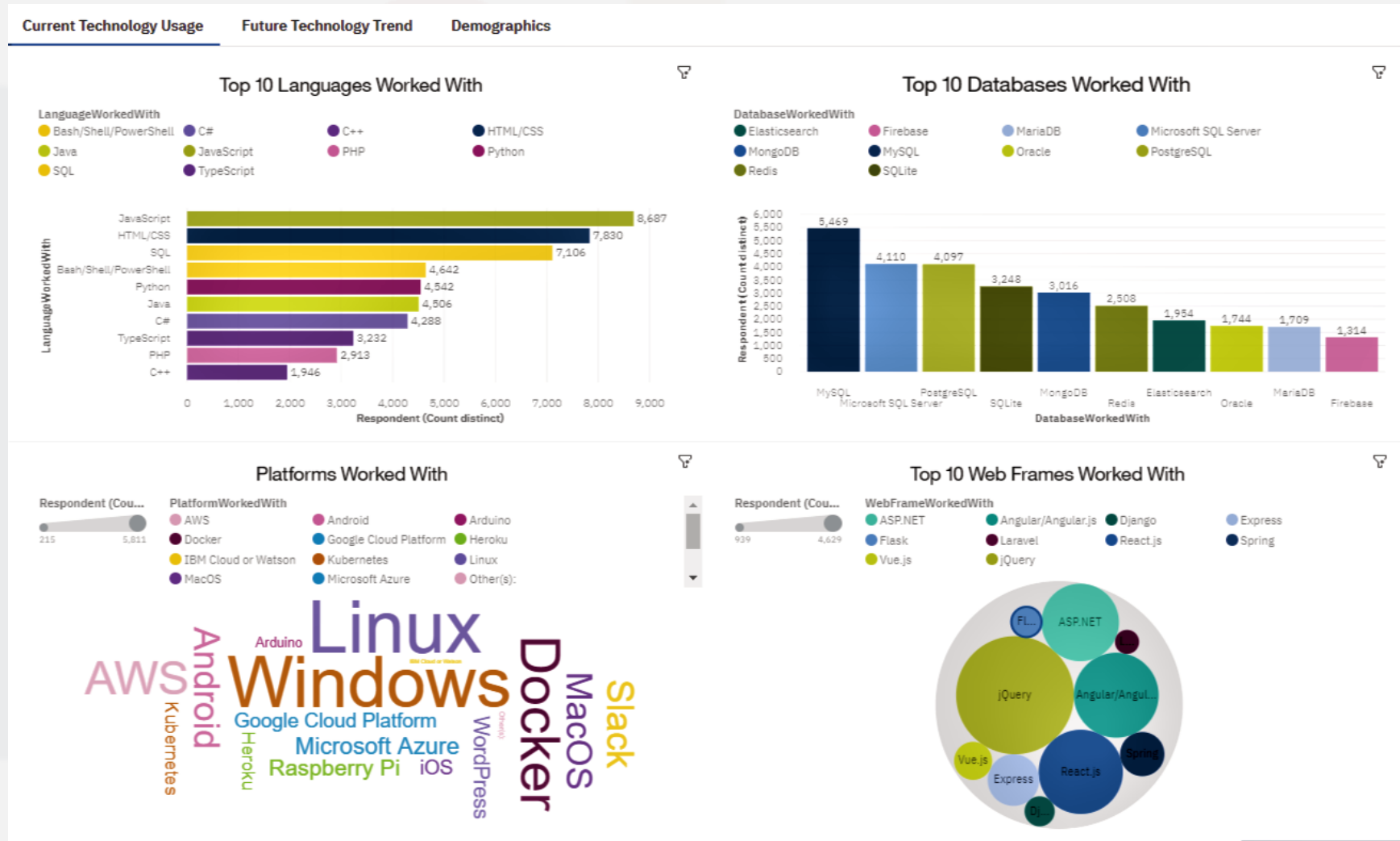
Dashboard



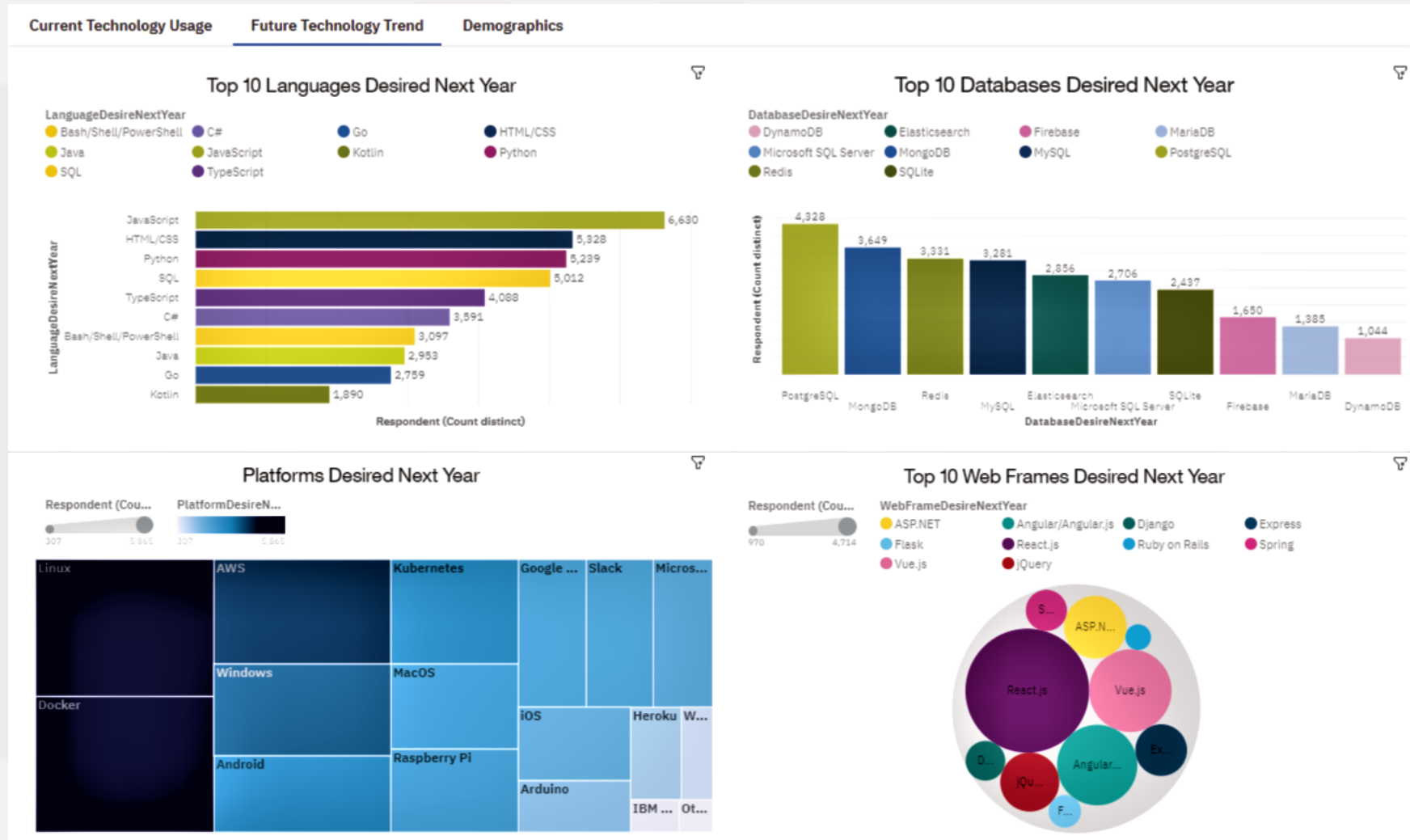
My emerging technology data analysis dashboard:

[https://eu-gb.dataplatform.cloud.ibm.com/
dashboards/9b591ba4-8cc2-49cb-a7ba-b89de352bd6a/
view/063edc2566ef3fea09cfd4e4079b7a027b647358b4bb80078
7d17b490e647897a83f1591c82d4b5a8f145461a1ea43589c](https://eu-gb.dataplatform.cloud.ibm.com/dashboards/9b591ba4-8cc2-49cb-a7ba-b89de352bd6a/view/063edc2566ef3fea09cfd4e4079b7a027b647358b4bb800787d17b490e647897a83f1591c82d4b5a8f145461a1ea43589c)

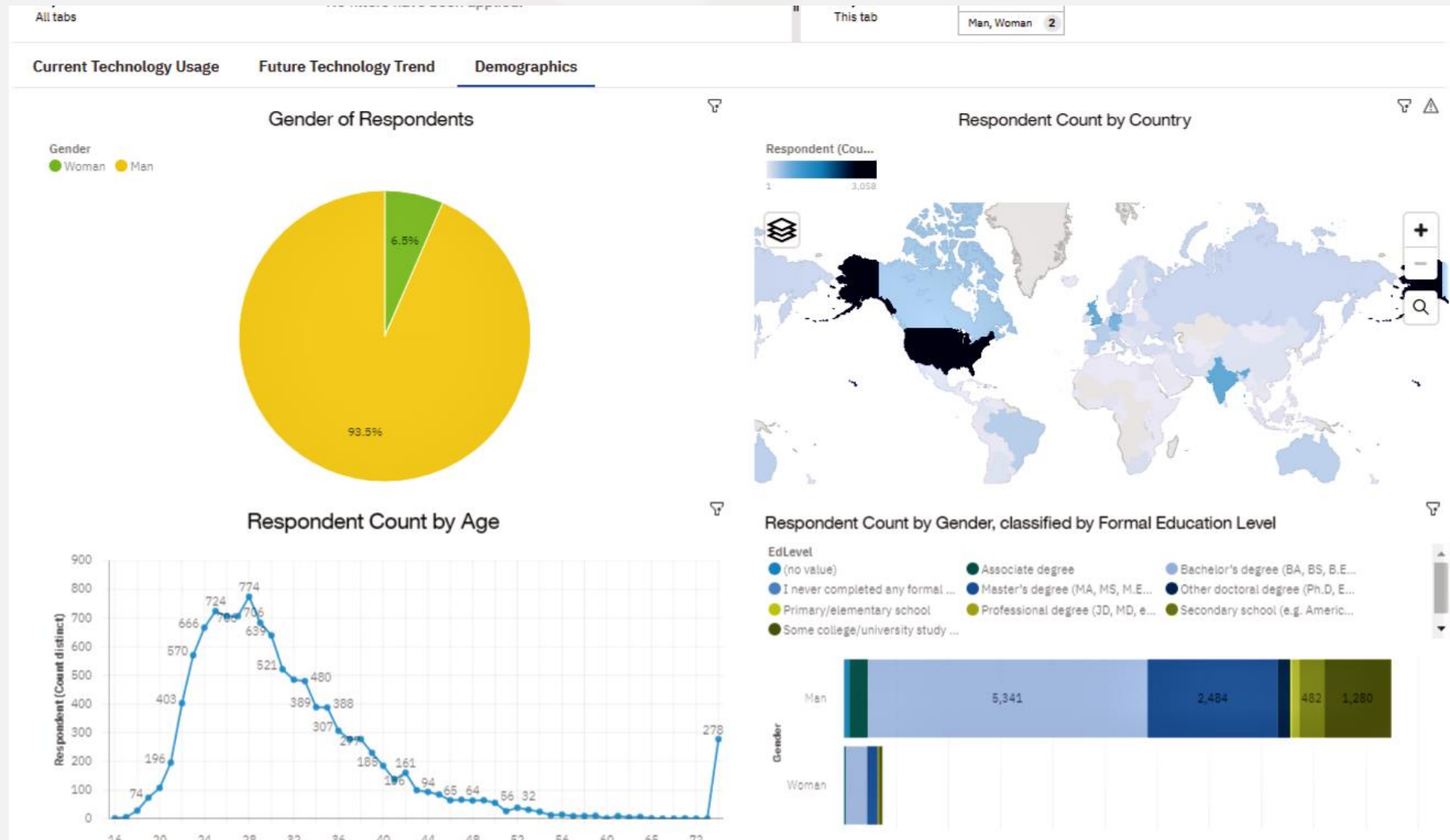
Dashboard Tab 1



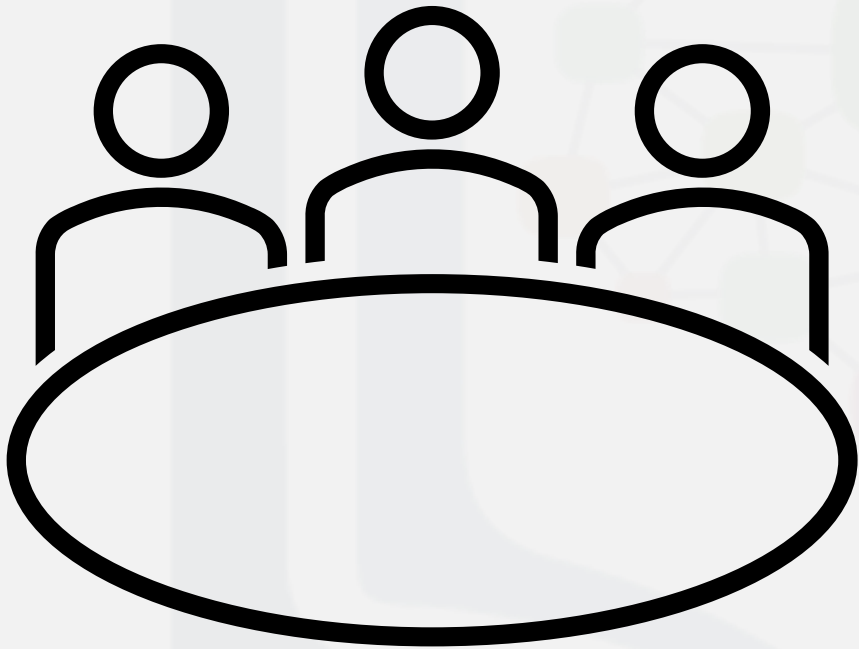
Dashboard Tab 2



Dashboard Tab 3



Discussion



- Surprises in the data?
- What findings are the most relevant?
- Ideas for action to be taken?
- Other sources of data to analyze?

Overall Findings and Implications

Findings

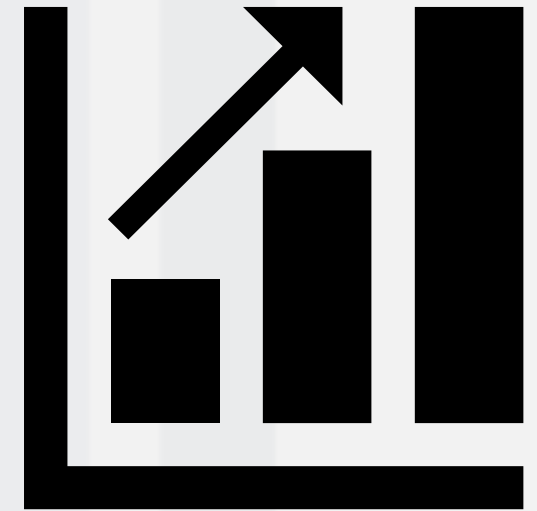
- Most in demand languages: Javascript, Python and TypeScript
- Most in demand databases: PostgreSQL, MongoDB
- SQL & SQL Databases are generally on a downward trend, demand is dropping despite PostgreSQL's popularity

Implications

- Javascript must remain a priority whilst shifting focus to upcoming languages
- New databases have overtaken demand from previous market leaders
- Forward planning must work on the expectation that demand for NoSQL will continue growing

Conclusion

- Most in demand languages: Javascript, Python and TypeScript
- Most in demand in demand databases: PostgreSQL, MongoDB
- The industry remains fast-moving, previously popular technology is being overtaken in demand by newer options
- Analysis must be integral part of ongoing business plans to keep pace in a competitive market



Appendix: Data Sources



Primary sources:

- <https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>, used under a ODbL: Open Database License.
- https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv

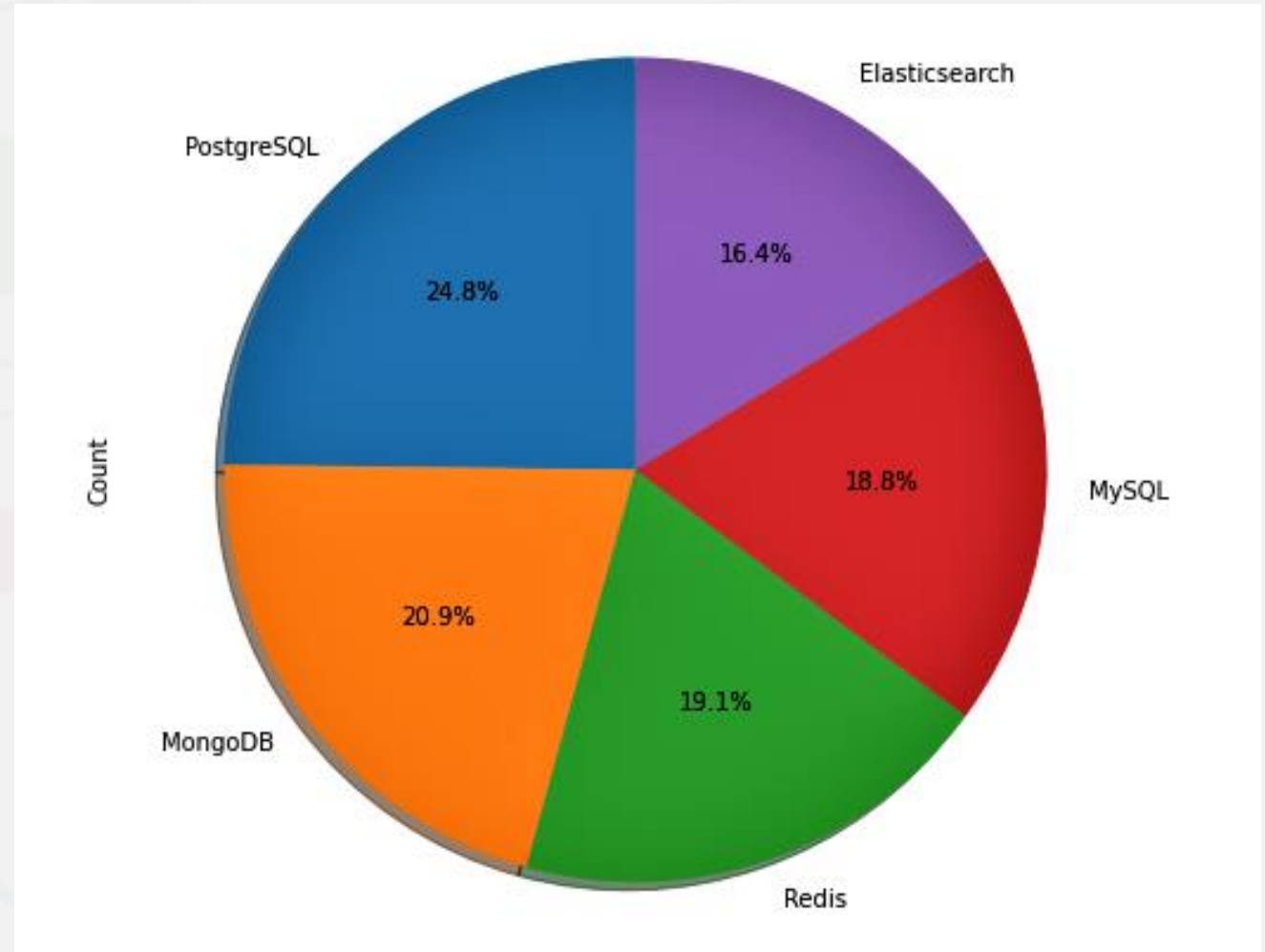
Additional sources:

- <https://jobs.github.com/api>
- https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/labs/datasets/Programming_Languages.html

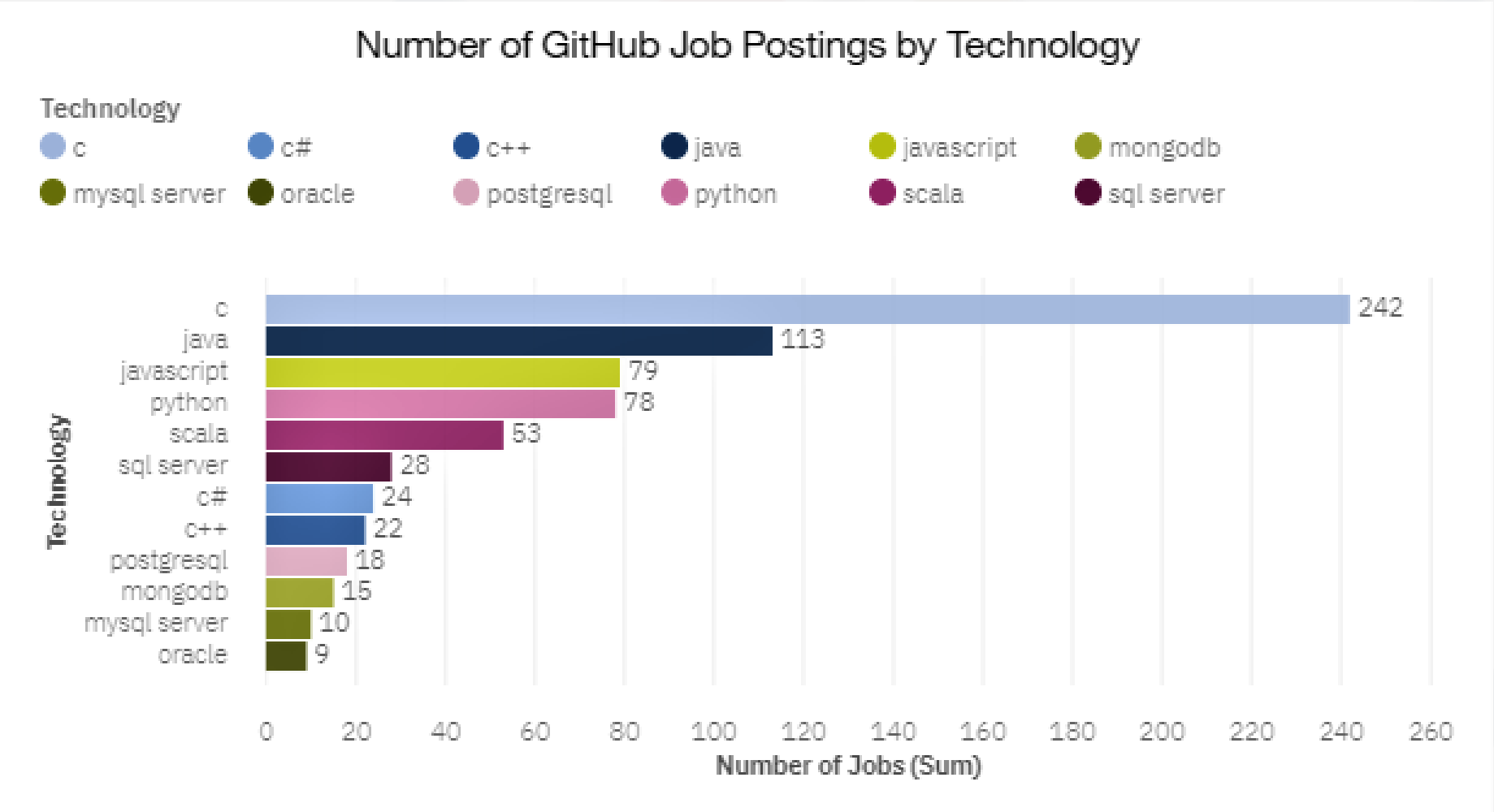
Appendix: Additional Visualizations

Top 5 Databases: Pie Chart

This chart was created during the exploration phase of the analysis. It clearly shows that the demand for the top 5 databases is fairly even, however it does not display the hierarchy of the values very well so was not used in the main results.

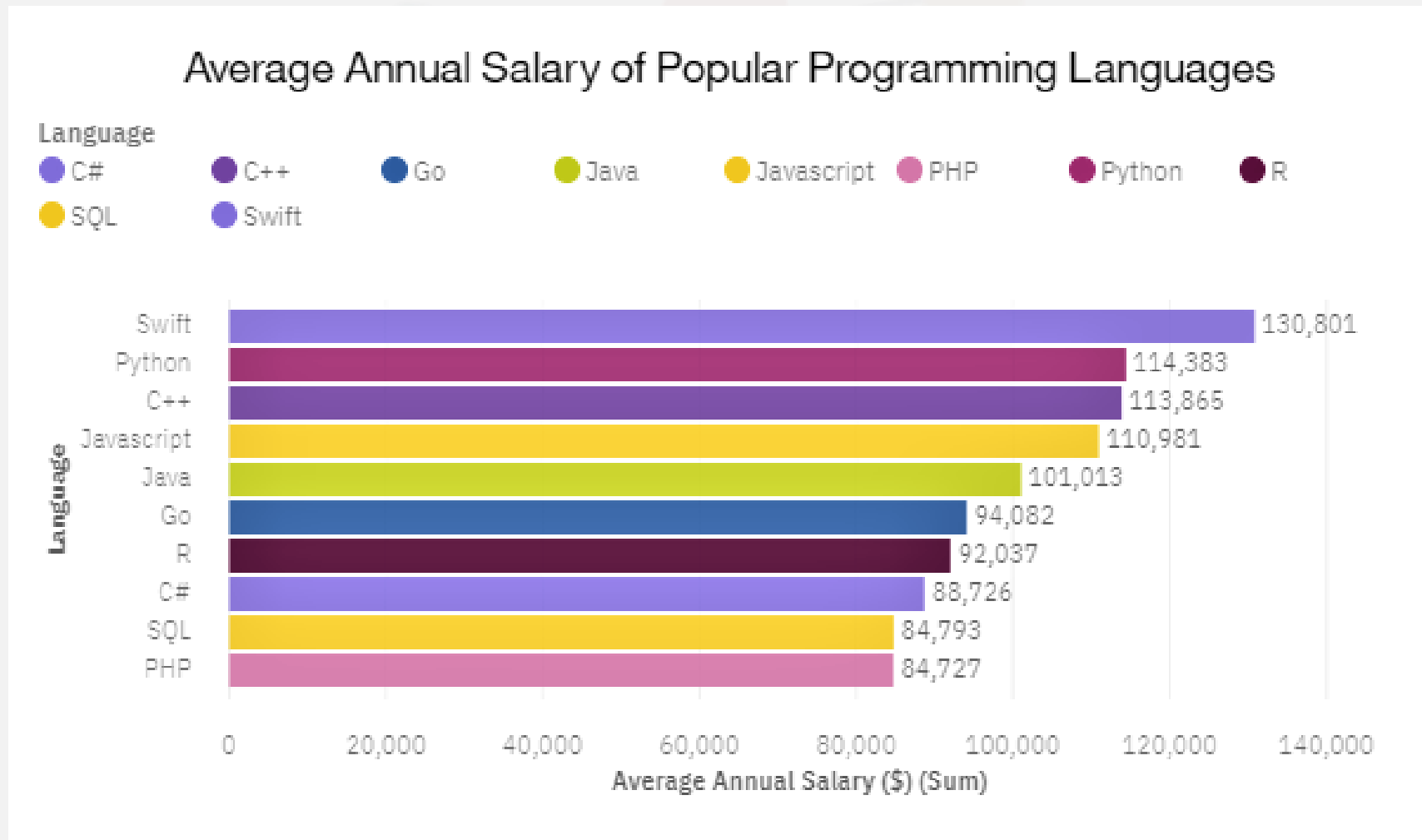


Appendix: GitHub Job Postings Bar Chart



This chart visualising the volume of GitHub job postings per programming languages shows that there is a discrepancy between the Stack Overflow survey desired languages and the current market demand for those skills.

Appendix: Popular Languages Bar Chart



This chart visualising the average salary of popular programming languages shows that many of the most in demand languages are among the most well-paid, with the exception of the best paid language, Swift, which was not high in demand in the Stack Overflow user survey.

Thank You

