Data Analysis Capstone Project

Jose Giovanni Guzman Alarcon

08 Sep 2023

OUTLINE

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

EXECUTIVE SUMMARY

The purpose of this presentation is to share the findings of a data analysis project on emerging programming skills. The project involved the following steps:

- Collecting data from various sources such as job postings, training portals, and surveys
- Wrangling the data to make it ready for analysis
- Applying statistical techniques to identify trends and insights
- Creating a dashboard using IBM Cognos Analytics to visualize the results
- Presenting the findings using storytelling skills

INTRODUCTION

- The following presentation contains the results of the data collection and analysis on the top programming skills that are most in demand in the IT industry.

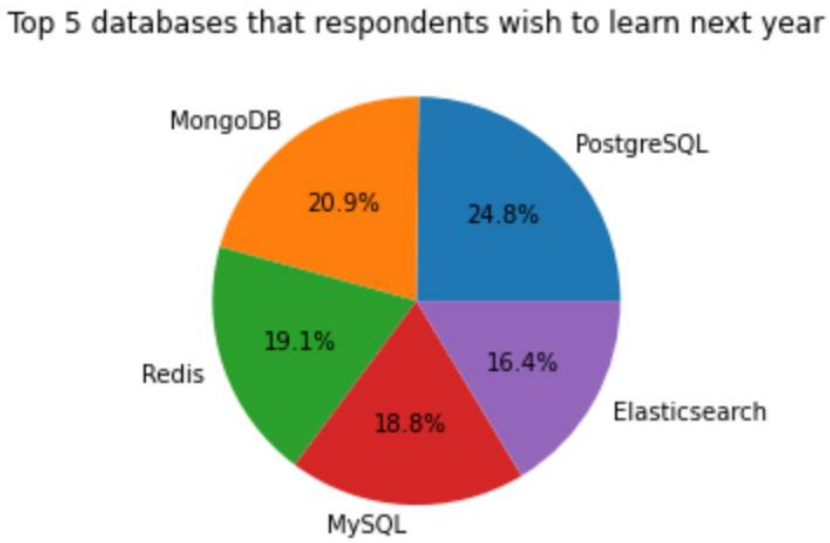
- During this project tools and techniques to collect, wrangle, analyze, and visualize the data were used. The goal is to provide with some insights and trends that can help you plan your future skill development and career growth.

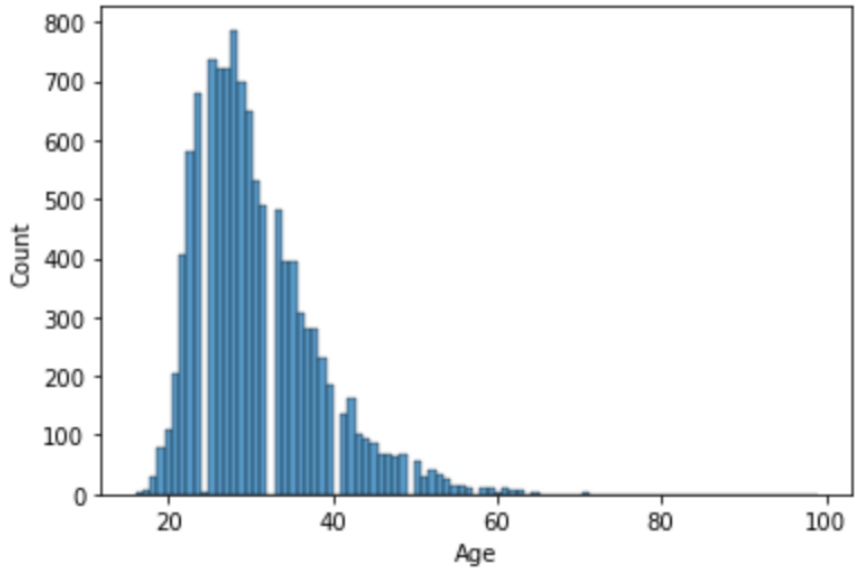
METHODOLOGY

To collect the top programming skills in demand, we used the following sources and methods:

We used job postings, training portals, and surveys to collect data on programming skills. We used Python libraries to wrangle, analyze, and visualize the data. We used IBM Cognos Analytics to create dashboards with key insights.

RESULTS

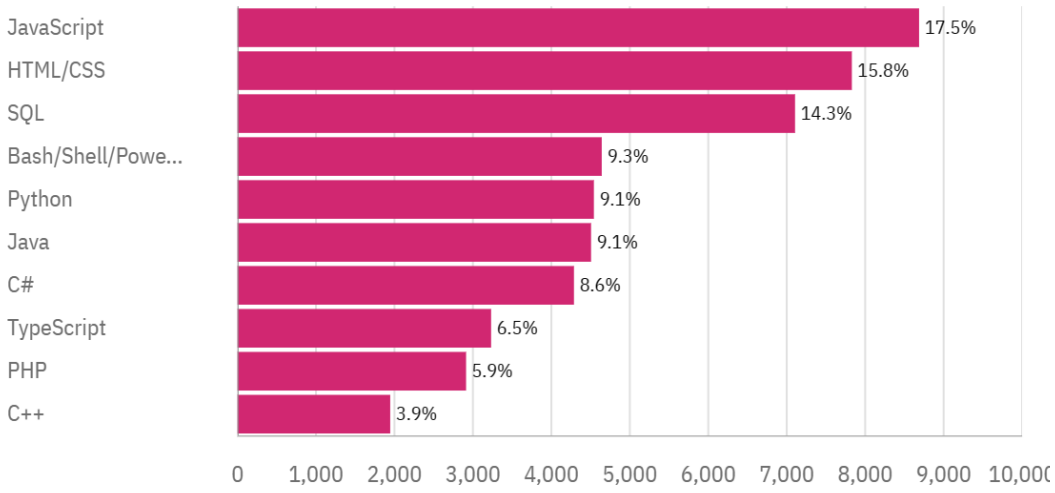




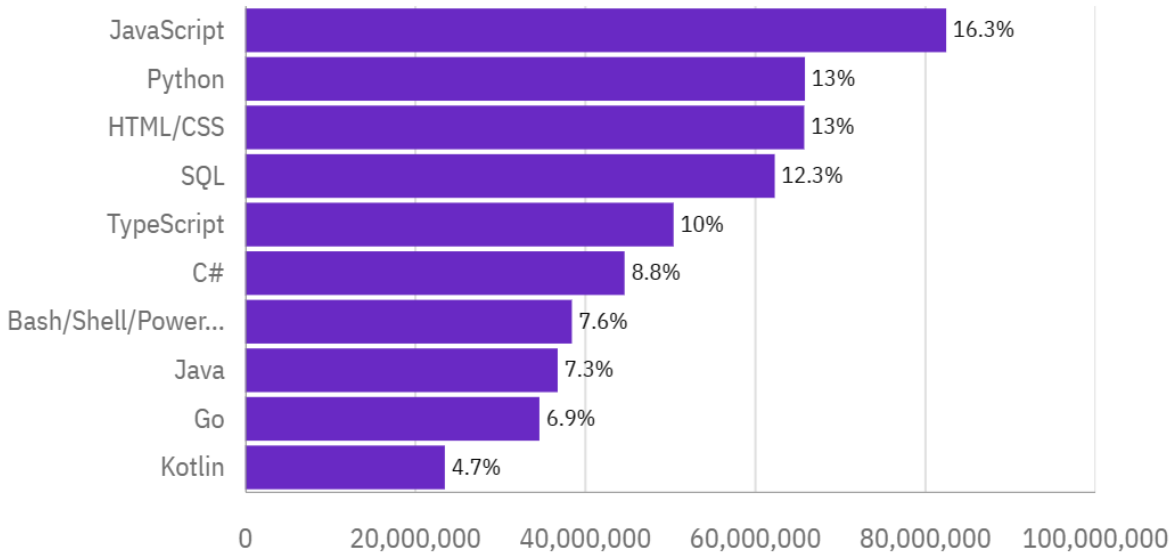
Respondent	12362.21
CompTotal	749993.19
ConvertedComp	131334.00
WorkWeekHrs	42.05
CodeRevHrs	4.76
Age	30.77

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

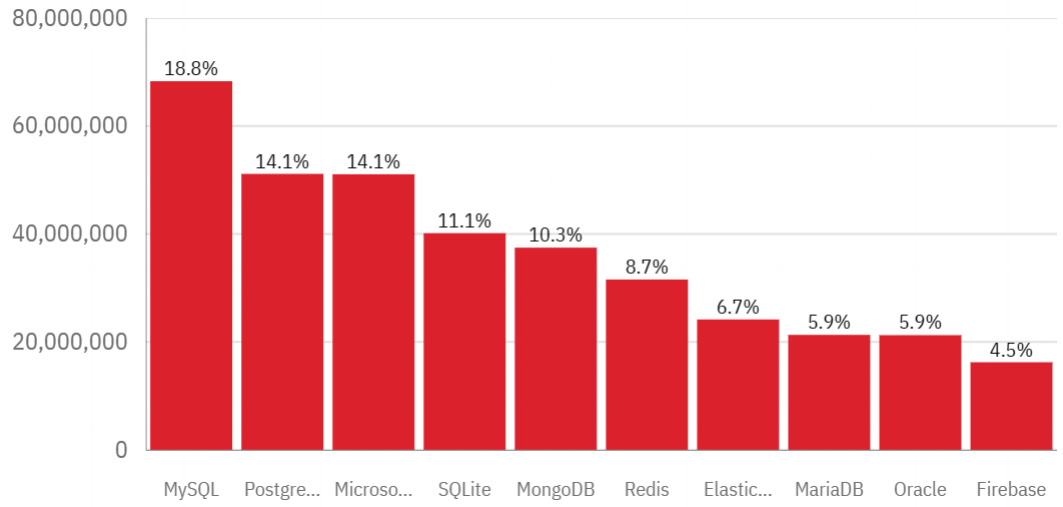
- JavaScript has Respondent of over 35 million
- JavaScript, Python, HTML/CSS, and SQL are the most popular programming languages among developers worldwide.

Implications

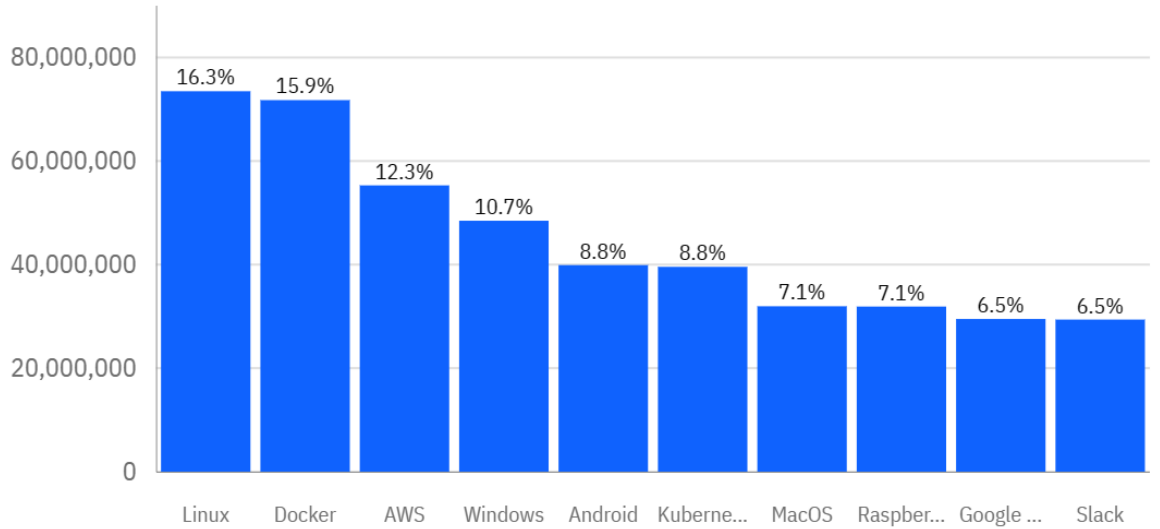
1. The popularity of programming languages reflects the demand and supply of developers in the software industry.
2. The popularity of programming languages influences the choice and design of new languages and frameworks.
3. The popularity of programming languages affects the interoperability and compatibility of different software systems and platforms.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

PostgreSQL is the second most popular DBMS with around 60,000,000 users.

MongoDB is the most popular non-relational DBMS in the world.

Linux is the most popular open-source operating system in the world

Docker is the most popular container platform in the world

Implications

The popularity of database management systems:

1. Reflects the demand and supply of database developers and administrators in the software industry.
2. Influences the choice and design of new database systems and technologies, such as Firebase.
3. Affects the interoperability and compatibility of different software systems and platforms, such as PostgreSQL.

The popularity of software platforms:

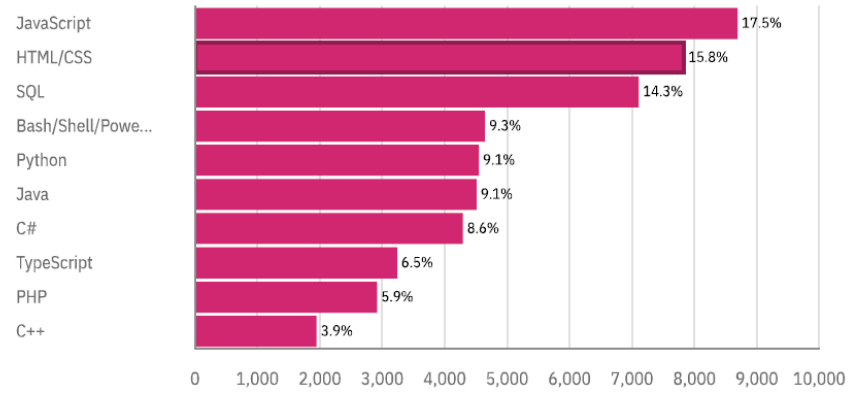
1. Reflects the demand and supply of software developers and engineers in the software industry.
2. Influences the choice and design of new software systems and technologies, such as Docker.
3. Affects the interoperability and compatibility of different software systems and platforms, such as Linux.

DASHBOARD

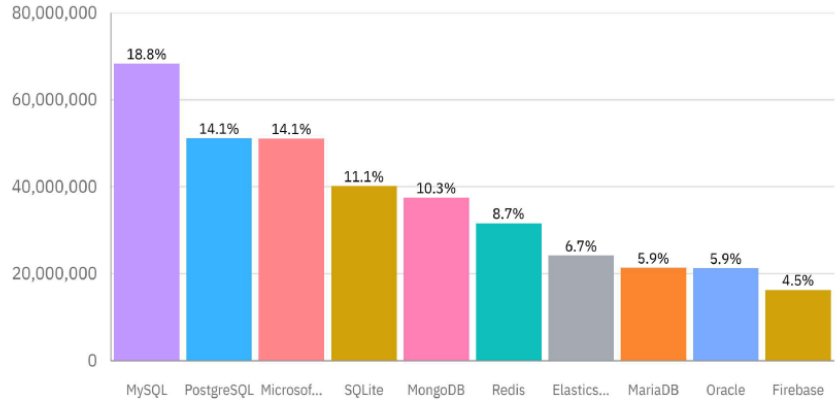
https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCapstone%2BProject%2B1&action=view&mode=dashboard&subView=model0000018a738c1839_00000000

DASHBOARD TAB 1

Respondent by programing Language

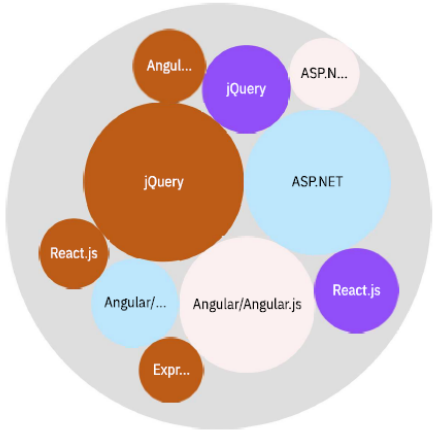


Respondent by Database



Most used platforms

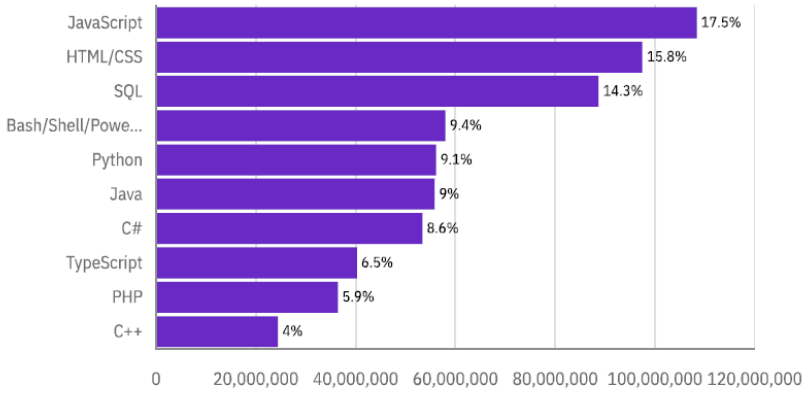
Respondent by Webframe



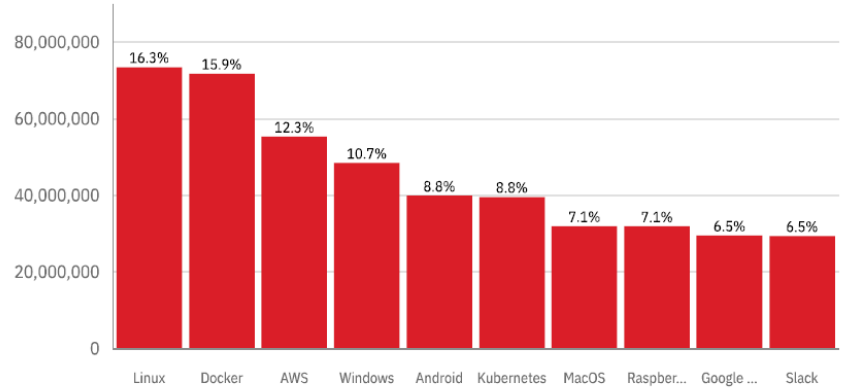
DASHBOARD TAB 2

Future Technology Trend

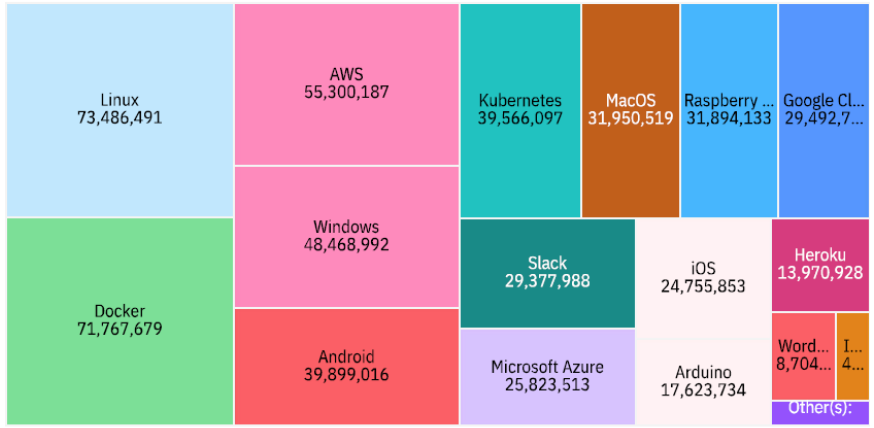
Respondent by Language



Respondent by Platform Desire in the Next Year



Platform Desire in the Next Year hierarchy



Most Web Frame Worked With

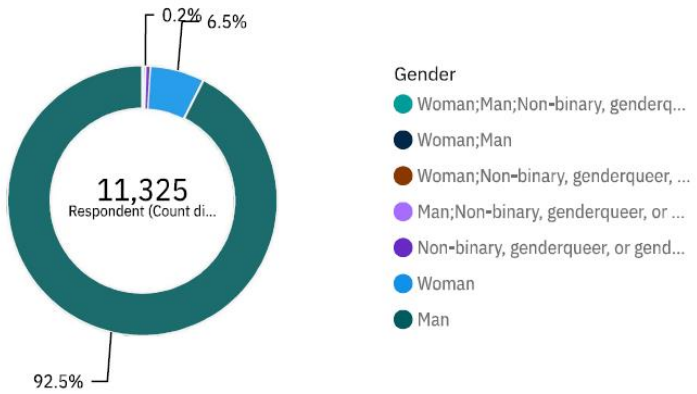
IBM Developer

SKILLS NETWORK

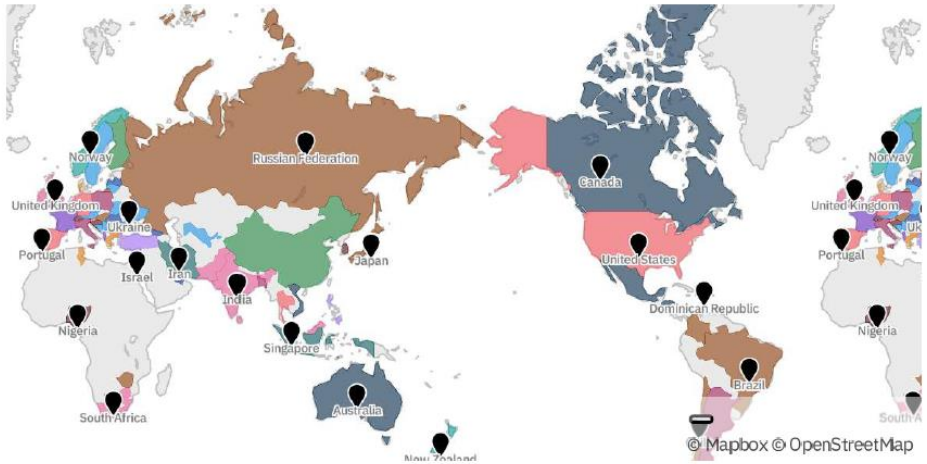
DASHBOARD TAB 3

Demographics

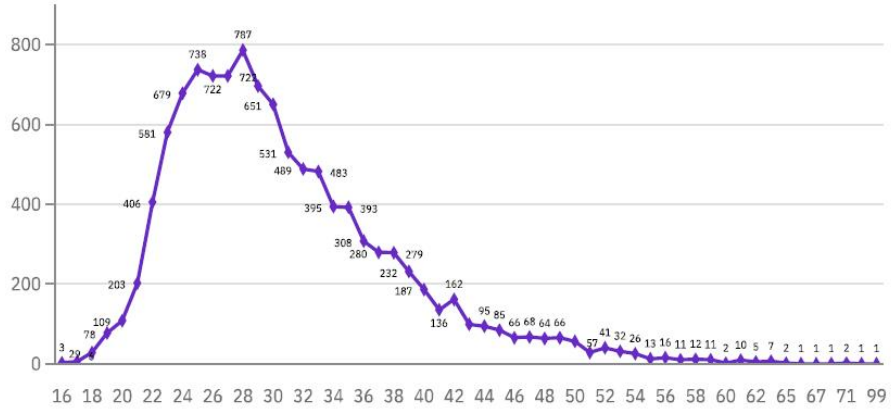
Gender



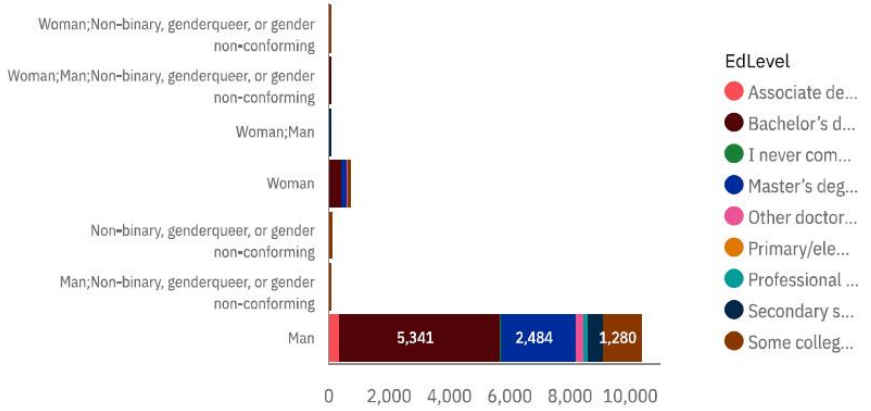
Respondent by Country



Respondent by Age



Respondent Count by Gender, classified by Formal Education Level



DISCUSSION

- How the popularity of database management systems and software platforms reflects the trends and needs of the software industry and the digital economy.
- How the choice and design of new database systems and software platforms are influenced by the features and benefits of the existing ones, such as Firebase, Docker, and Linux.
- How the interoperability and compatibility of different software systems and platforms are facilitated by the use of multiple data models, such as PostgreSQL, or real-time data synchronization, such as Firebase.
- What are the challenges and opportunities for software developers and engineers who work with different database management systems and software platforms, such as learning new skills, finding jobs, or creating innovative applications.

OVERALL FINDINGS & IMPLICATIONS

Findings

- JavaScript, Python, HTML/CSS, and SQL are the most popular programming languages among developers worldwide, with JavaScript having over 35 million respondents.
- PostgreSQL is the second most popular DBMS with around 60 million users, and MongoDB is the most popular non-relational DBMS in the world.
- Linux is the most popular open-source operating system in the world, and Docker is the most popular container platform in the world.

Implications

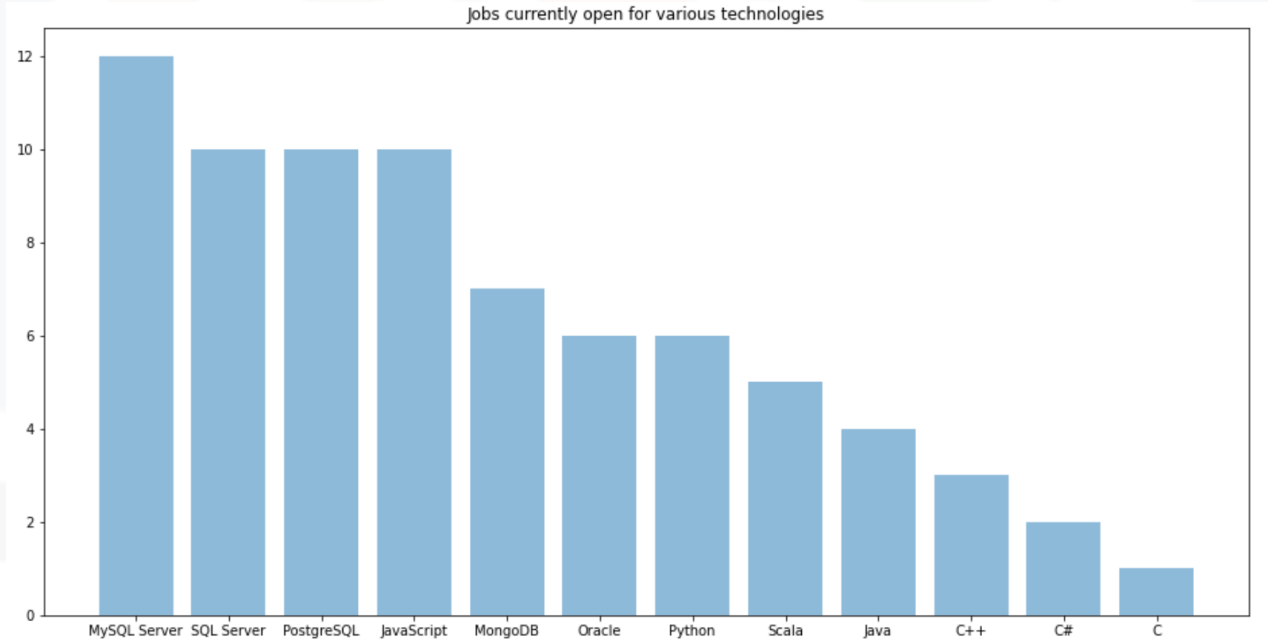
- The popularity of programming languages reflects the demand and supply of developers in the software industry, and influences the choice and design of new languages and frameworks.
- The popularity of database management systems reflects the demand and supply of database developers and administrators in the software industry, and influences the choice and design of new database systems and technologies, such as Firebase.
- The popularity of software platforms reflects the demand and supply of software developers and engineers in the software industry, and influences the choice and design of new software systems and technologies, such as Docker.
- The popularity of programming languages, database management systems, and software platforms affects the interoperability and compatibility of different software systems and platforms, such as PostgreSQL, Linux, etc.

CONCLUSION

- The popularity of programming languages, database management systems, and software platforms reflects the trends and needs of the software industry and the digital economy.
- The choice and design of new languages and frameworks, such as Firebase, Docker, and Linux, are influenced by the features and benefits of the existing ones, such as JavaScript, PostgreSQL, and AWS.
- The interoperability and compatibility of different software systems and platforms are facilitated by the use of multiple data models, such as PostgreSQL, or real-time data synchronization, such as Firebase.
- The skills and knowledge of programming languages, database management systems, and software platforms are essential for software professionals who want to create innovative and effective applications.

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.