

# Data Analyst Capstone Project

**Manpreet Padwal**  
March 21st 2025

# Contents

- Executive Summary
- Introduction
- Methodology
- Results
- Discussion
- Conclusion

# EXECUTIVE SUMMARY

- ▶ To stay competitive in IT, keeping up with evolving technologies is key.
- ▶ This report analyzes skill demand trends in programming languages, databases, and tech demographics using data from Stack Overflow, IBM, and GitHub.
- ▶ Findings show JavaScript as the most popular language, with PostgreSQL expected to surpass MySQL in demand. Most survey respondents are male, based in the USA, and around 28 years old.

# INTRODUCTION

- ▶ This report analyzes trends in programming languages, databases, platforms, and web frameworks using data analytics.
- ▶ It explores: The most in-demand programming languages; Key database skills in demand; Popular IDEs and web frameworks Targeted at IT professionals, HR managers, and anyone interested in future-proof IT skills.

# METHODOLOGY

1. Data on job availability across technologies and locations was collected using the GitHub Jobs API in Python. Exploratory data analysis (EDA)
2. Programming languages and salaries were scraped from IBM, while Stack Overflow's 2019 survey data was downloaded and saved.
3. Python was used for data cleaning and analysis, including exploratory data analysis to check distribution, outliers, and correlations.
4. Charts, graphs, and dashboards were created using Python and Cognos analytics to visualize the data. All the python analyses were carried out on Jupyter notebook through visual studio.

# RESULTS

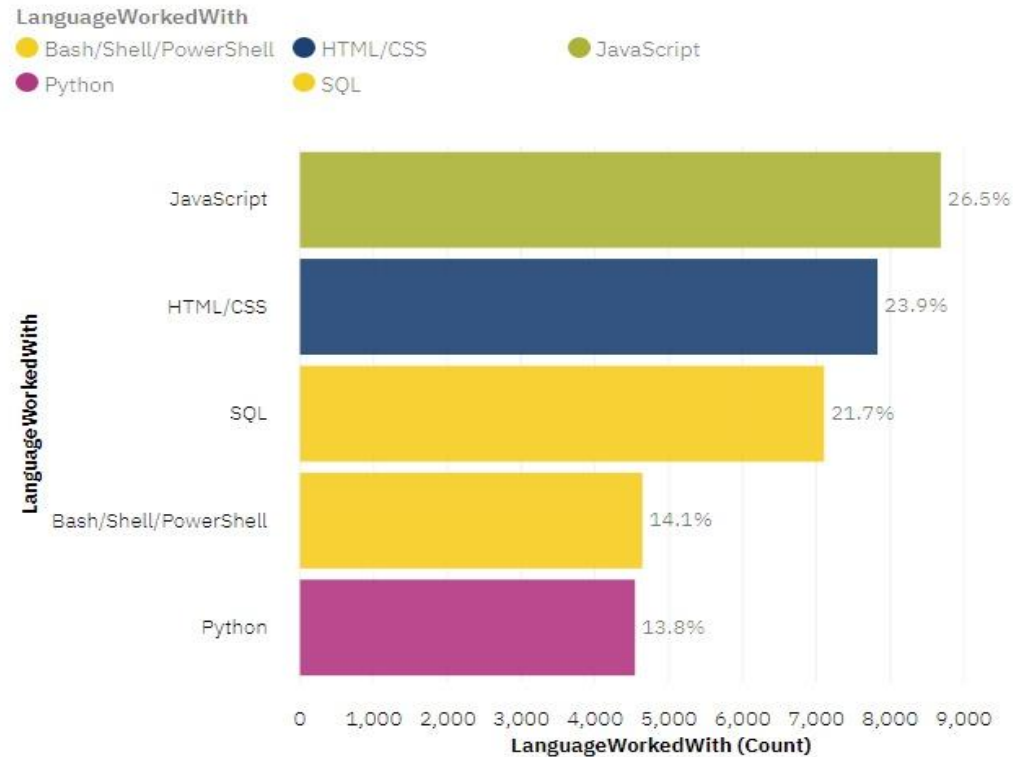


# PROGRAMMING LANGUAGE TRENDS

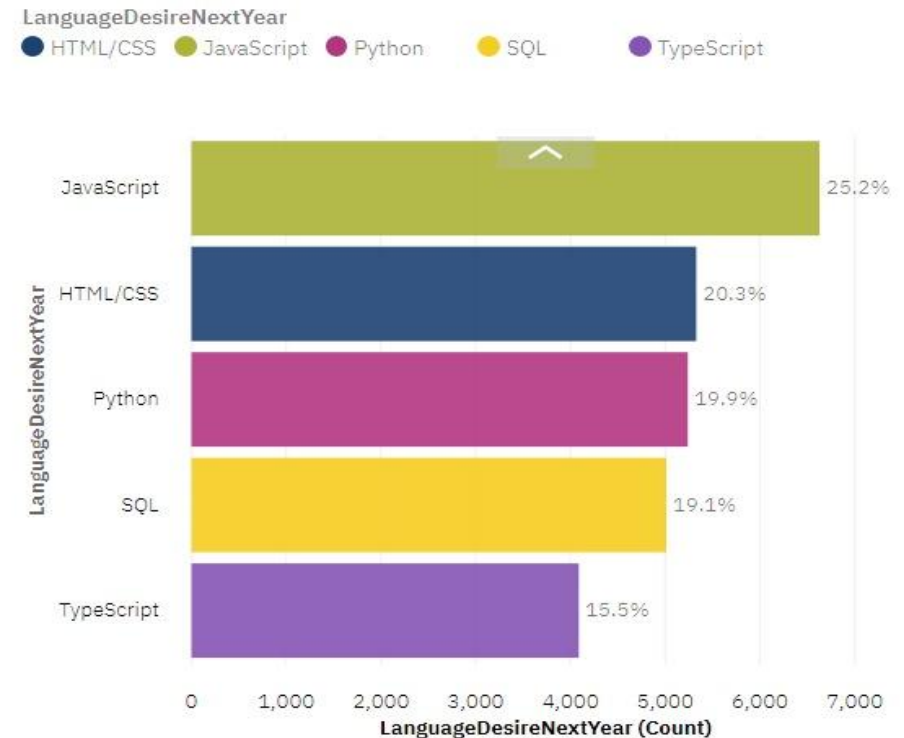
## ► Current year

## ► Next year

Top 5 programming languages for the Current year



Top 5 programming languages for Next year



# PROGRAMMING LANGUAGE

## Observations

- ▶ Currently, the most widely used programming languages are JavaScript, HTML/CSS, SQL, Shell languages, and Python. Looking ahead to next year and beyond, JavaScript, HTML/CSS, Python, SQL, and TypeScript are expected to dominate. Python, in particular, is projected to surpass SQL in demand next year.

## Suggestions

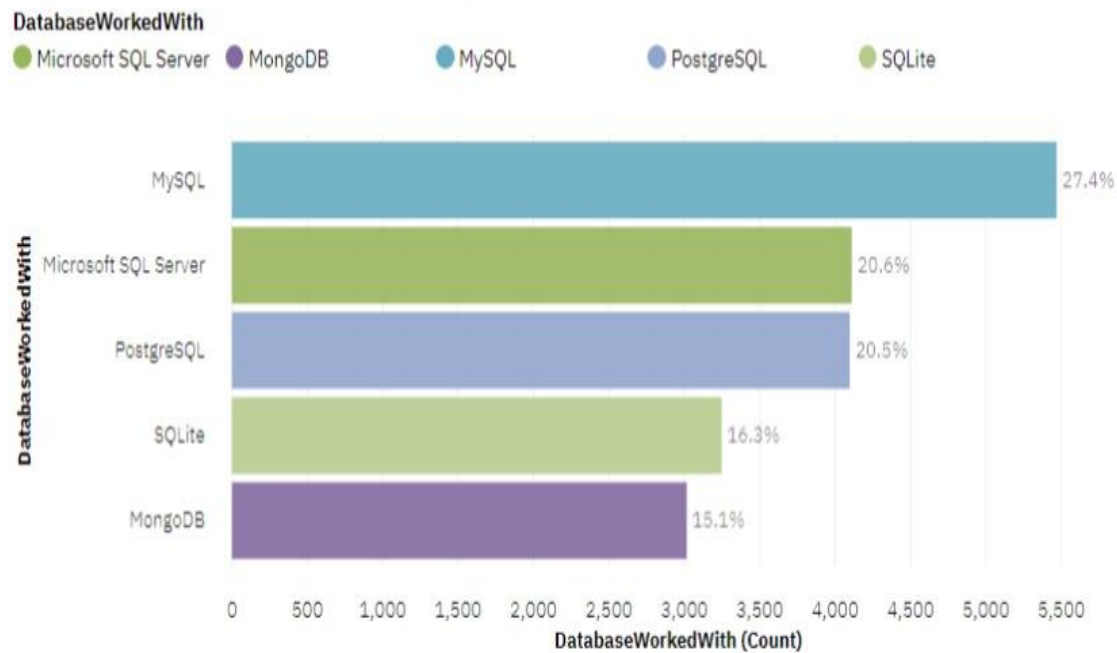
- ▶ JavaScript and HTML are essential for web development, making web development one of the most in-demand tech skills, especially with the rising popularity of TypeScript. Python is increasingly sought after due to the growing demand for AI and ML expertise. SQL remains the most relevant language for data professionals, making it crucial for aspiring data analysts, scientists, and business analysts to master SQL skills.



# DATABASE TRENDS

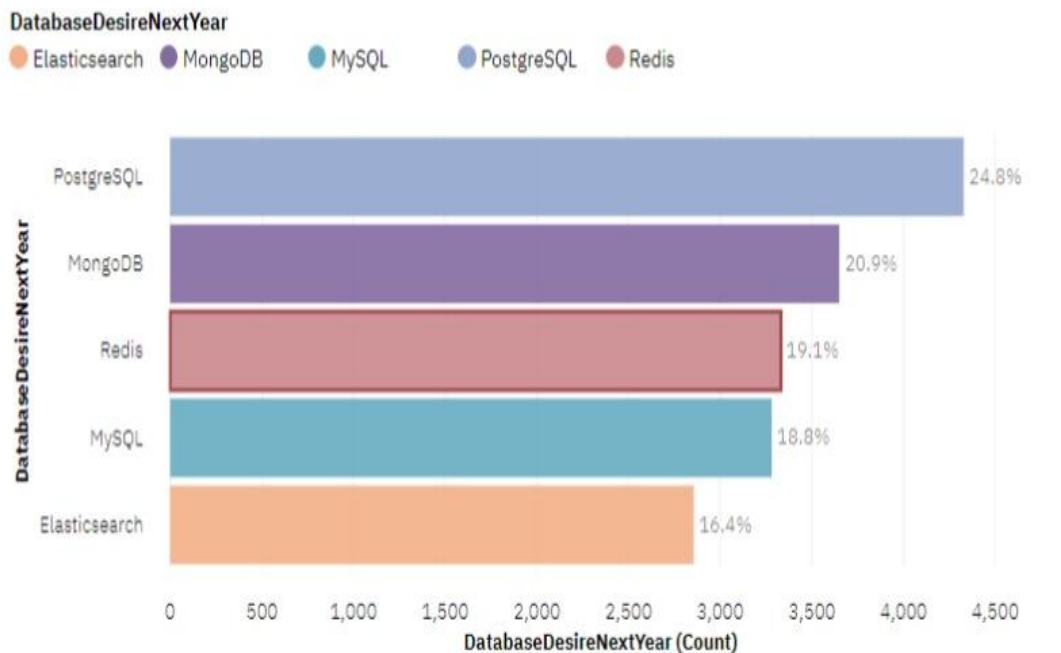
► Current year

Top 5 Databases in the Current year



► Next year

Top 5 Databases for the Next Year



# DATABASE TRENDS

## Observations

Currently, the top five most used databases are MySQL, Microsoft SQL Server, PostgreSQL, SQLite, and MongoDB. However, PostgreSQL, MongoDB, Redis, MySQL, and Elasticsearch are projected to become more popular in the future. Redis and Elasticsearch, being relatively new tools, are set to gain more traction in the IT space.

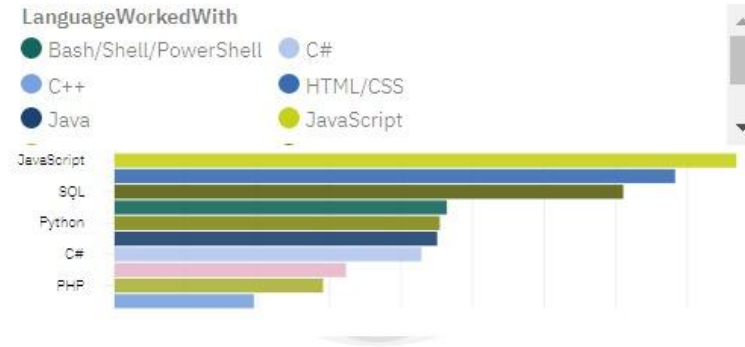
## Consequences

SQL remains a top tool for data specialists. Companies continue to favor open-source databases. Oracle SQL, however, was not among the top five and is gradually losing relevance over time.

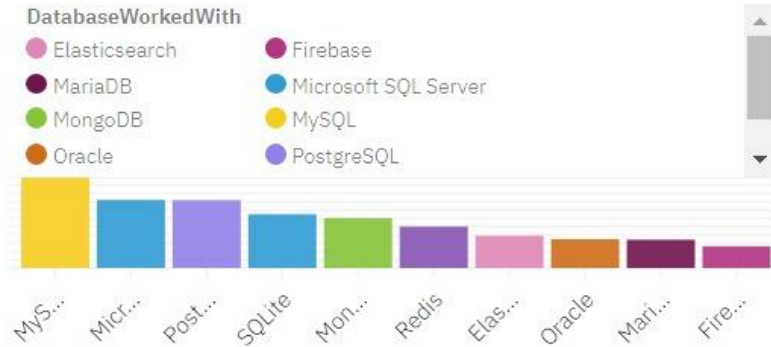
# Dashboard code link in Jupyter Notebook

- ▶ [Data\\_Analyst\\_Course/Dashbord code.ipynb at main · MPS90MP/Data\\_Analyst\\_Course](#)

### Top 10 LanguageWorkedWith



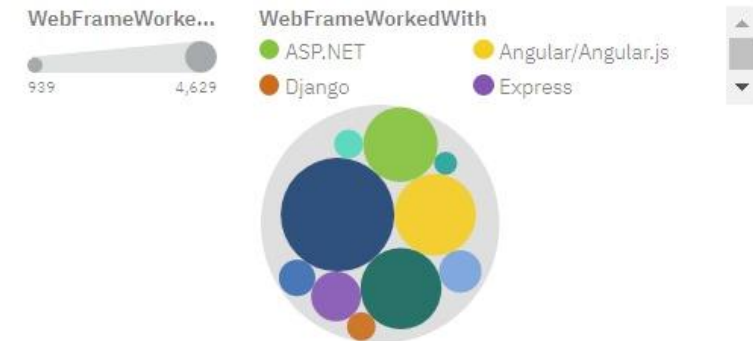
### Top 10 DatabaseWorkedWith



### PlatformWorkedWith



### Top 10 WebframeWorkedWith

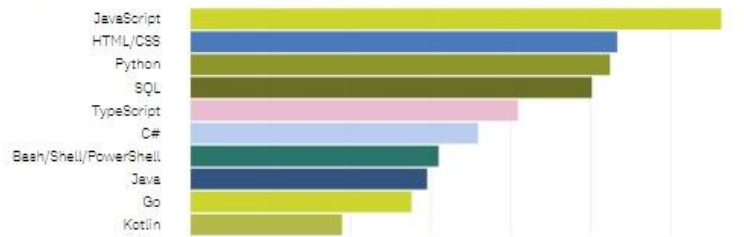


DASHBOARD TAB 1

## Top 10 LanguageDesireNextYear

LanguageDesireNextYear

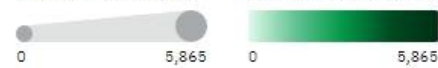
● Bash/Shell/PowerShell ● C#  
● Go ● HTML/CSS



## Most desired Platforms next year

PlatformDesireN...

PlatformDesireN...



## Top 10 Most Desired Databases

DatabaseDesireNextYear

● DynamoDB ● Elasticsearch  
● Firebase ● MariaDB



## Most Desired Web frames Next Year

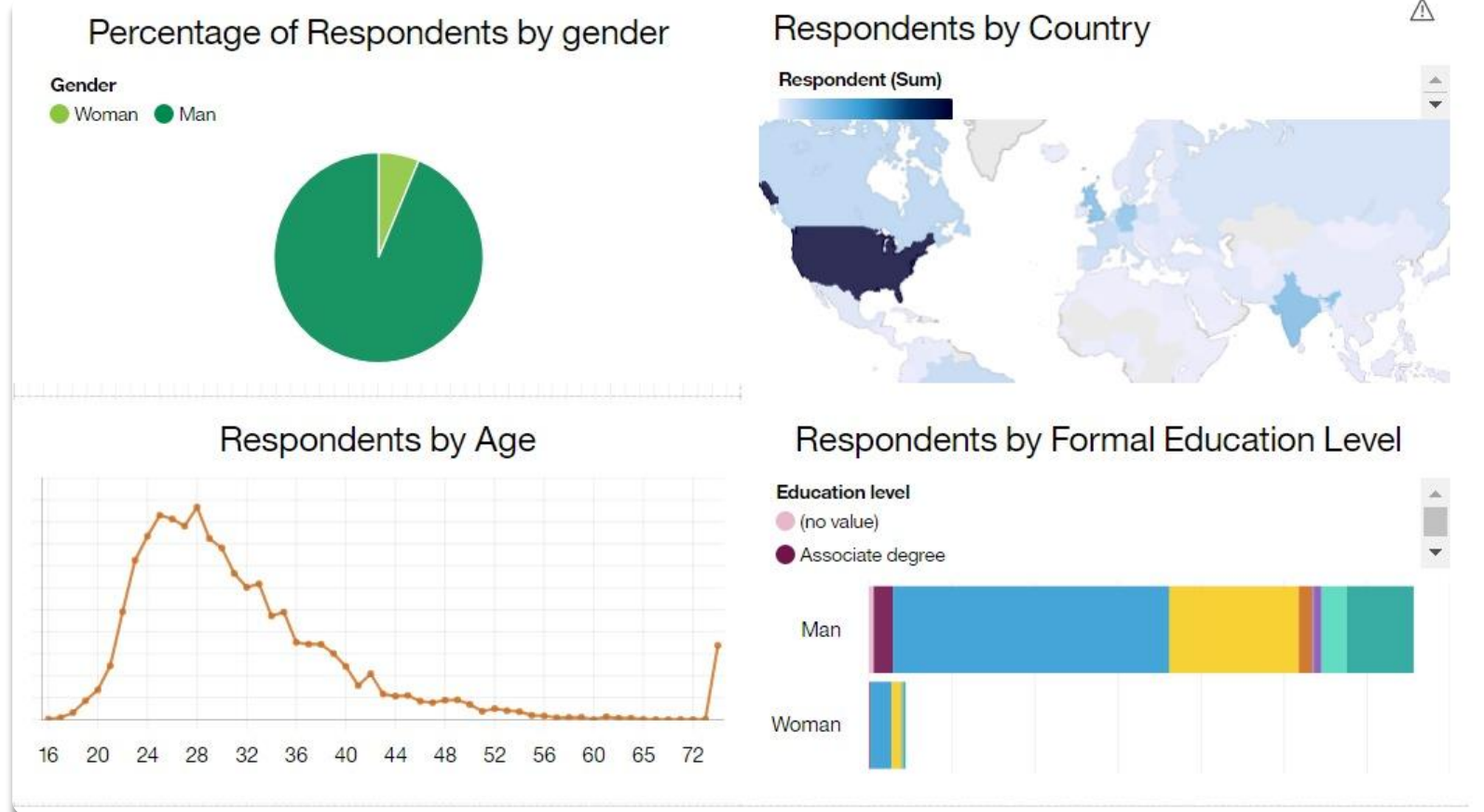
WebFrameDesire...

WebFrameDesireNextYear



DASHBOARD TAB 2

# DASHBOARD TAB 3



# DISCUSSION

- Upskilling in the tech sector.
- Closing the gender gap in technology.
- Is a master's or doctorate degree necessary?
- Rising demand for mobile development with Kotlin.
- Expanding tech education and access in less developed regions (South East Asia, South America, Africa, parts of Europe).
- Future relevance of Oracle SQL.

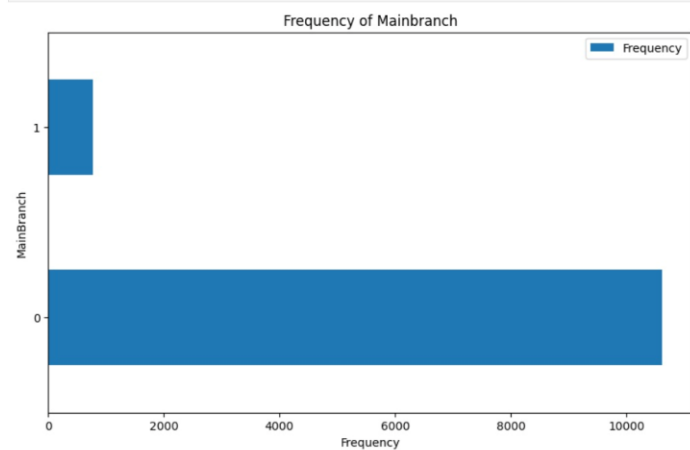
# CONCLUSION

- Most IT professionals have a bachelor's degree.
- Web development languages are currently the most popular and in-demand tools in IT.
- The tech sector is predominantly composed of individuals under 40.
- Many respondents aim to learn PostgreSQL and React JS next year.
- Data professionals should master NoSQL alongside SQL databases.
- Web development remains a lucrative skill.
- Less developed countries need greater access to tech training and education.



	MainBranch	Frequency
0	I am a developer by profession	10618
1	I am not primarily a developer, but I write co...	780

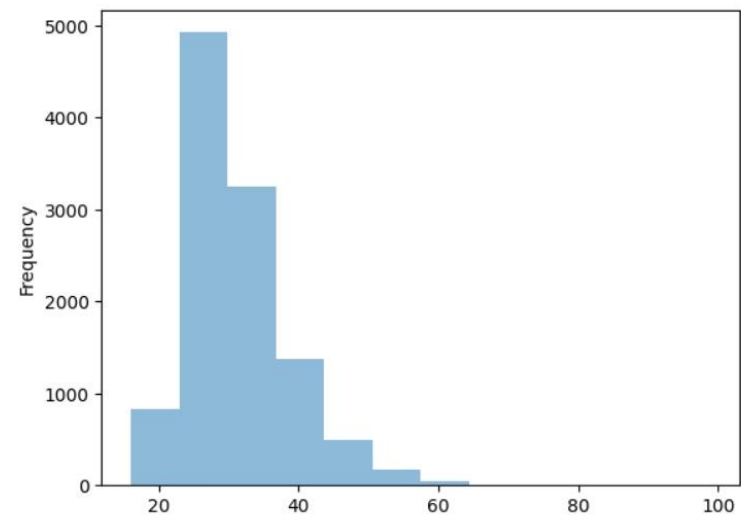
```
bar_df.plot(kind = 'barh', figsize=(10,6))
plt.xlabel('Frequency')
plt.ylabel('MainBranch')
plt.title('Frequency of Mainbranch')
plt.show()
```



Plot a histogram of the column `Age`.

```
df['Age'].plot.hist(bins=12, alpha=0.5) #
```

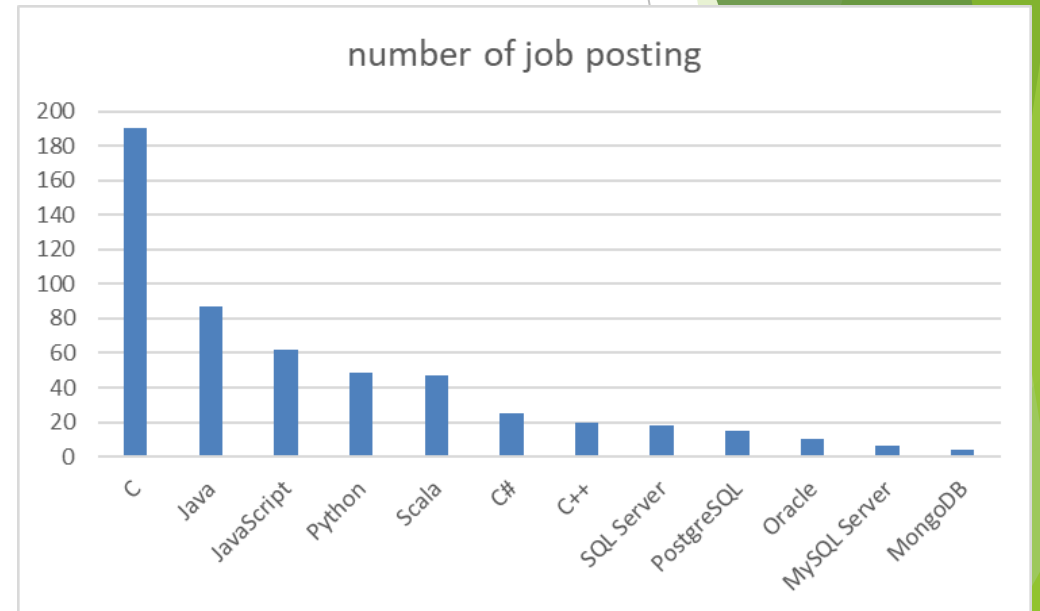
<AxesSubplot: ylabel='Frequency'>



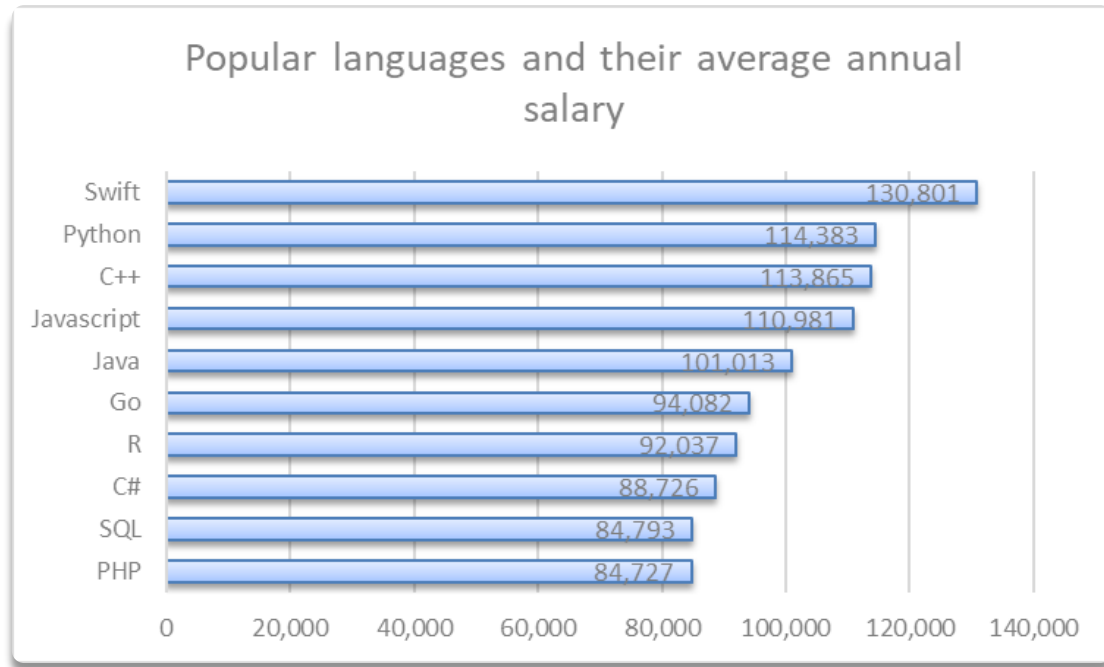
```
Respondent      0.004041
CompTotal       0.006970
ConvertedComp    0.105386
WorkWeekHrs     0.036518
CodeRevHrs      -0.020469
Age             1.000000
Name: Age, dtype: float64
```

# APPENDIX

# JOB POSTINGS



# POPULAR LANGUAGES



Bar chart displaying popular languages and their average annual salary. The data was collected through web scraping the Github jobs data and saved in a csv file.