Analysis on Emerging Technology Skills and Trends

A Work For the IBM Data Analyst Capstone Project

Osvaldo Olivares, PMP

August 10th, 2024

IBM Developer

SKILLS NETWORK

OUTLINE

1. Executive Summary
2. Introduction
3. Methodology
4. Results
   a. Visualization – Charts
   b. Dashboard
5. Discussion
   a. Findings & Implications
6. Conclusion
7. Appendix

EXECUTIVE SUMMARY

In the rapidly evolving global IT sector, staying competitive requires continuous adaptation to new technologies. This report leverages data analytics to identify current and future trends in essential skills, focusing on programming languages, databases, and other technological domains. It also examines the demographics of professionals in the technology sector.

Data was meticulously gathered from sources such as the Stack Overflow survey, IBM's site, and GitHub job postings. After thorough cleaning and exploratory analysis, the data was visualized on interactive dashboards.

Key Findings:

- JavaScript Dominance: JavaScript is the most popular programming language and is expected to maintain its lead.
- Database Trends: MySQL currently has the highest usage, but PostgreSQL is projected to see increased demand.
- Demographic Insights: The majority of professionals are male, based in the USA, and have an average age of 28 years.
- This comprehensive analysis provides valuable insights for strategic planning and skill development in the IT sector.

INTRODUCTION

- This presentation leverages data analytics to uncover current and future trends in essential IT skills, focusing on programming languages, databases, platforms, and web frameworks. The research addresses key questions such as:

- Which programming languages are most in demand today?
- What are the most sought-after database skills?
- Which IDEs or web frameworks are popular?

- Targeted at IT professionals, HR managers, and anyone interested in the IT sector, this report provides valuable insights into the top on-demand skills that will remain relevant in the future.

METHODOLOGY

Data was collected in various formats, including the number of job postings for different technologies and locations, using the GitHub Jobs API in Python.

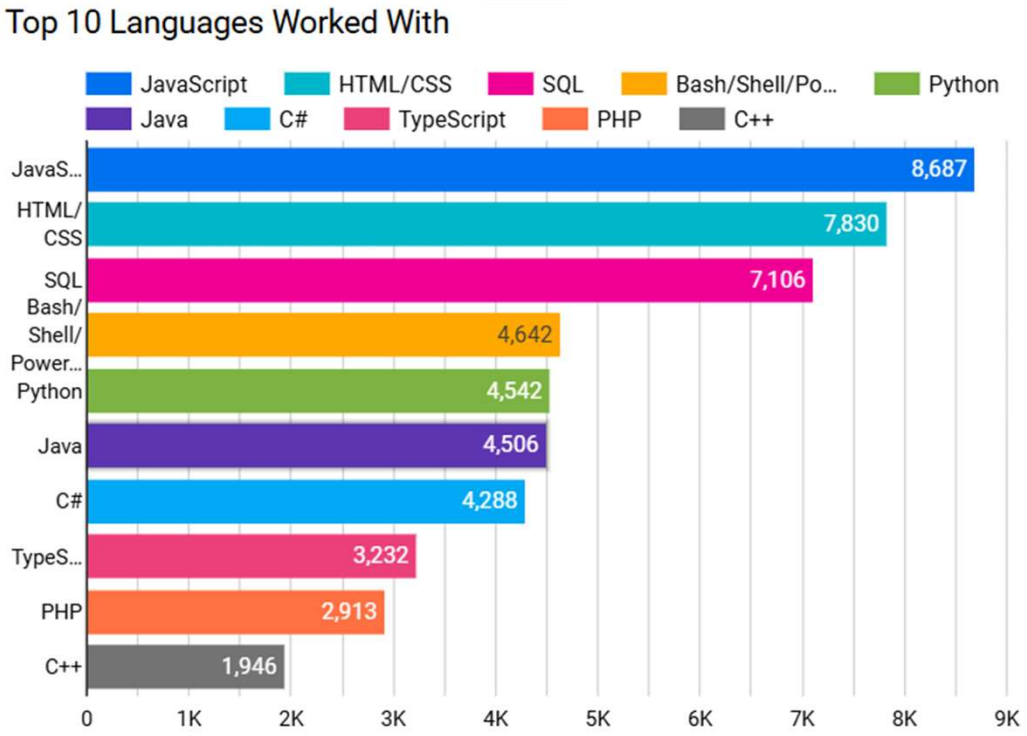
To gather the names of programming languages and their corresponding yearly wages, data was scraped from the IBM website. Additionally, the dataset from the 2019 Stack Overflow Developer Survey was downloaded and saved for further analysis.

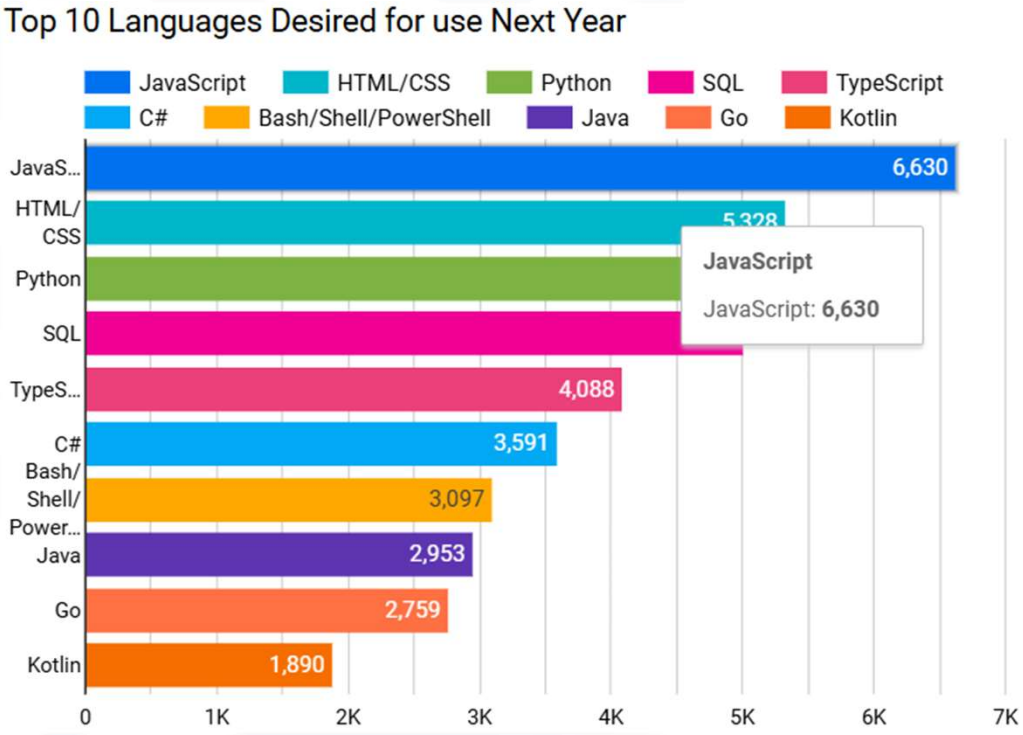
Python was utilized for data cleaning and analysis. An exploratory data analysis (EDA) was conducted to evaluate the data distribution, identify outliers, and examine the correlations between various dataset columns.

Data visualization was achieved through the creation of charts, graphs, and dashboards using Python and Google Looker. All Python analyses were conducted in Jupyter Notebook within Visual Studio.


Results

PROGRAMMING LANGUAGE TRENDS





IBM Developer

SKILLS NETWORK

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

Current Most Used Languages:

- JavaScript
- HTML/CSS
- SQL
- Shell languages
- Python

Projected Most Used Languages:

- JavaScript
- HTML/CSS
- Python
- SQL
- TypeScript

Future Demand:

- Python is expected to surpass SQL in demand next year.

Implications

Web Development Demand:

- JavaScript and HTML are essential for web development, indicating that this tech skill is in high demand.
- TypeScript is rapidly gaining popularity, further boosting the need for web development expertise.

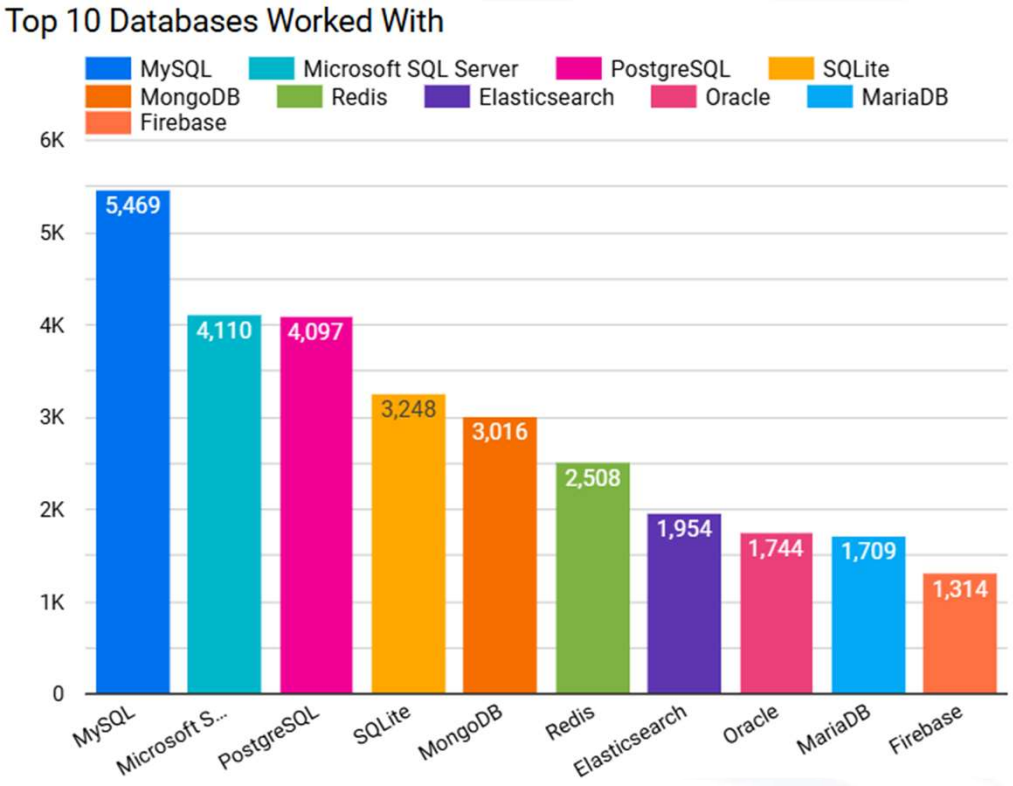
Rising Popularity of Python:

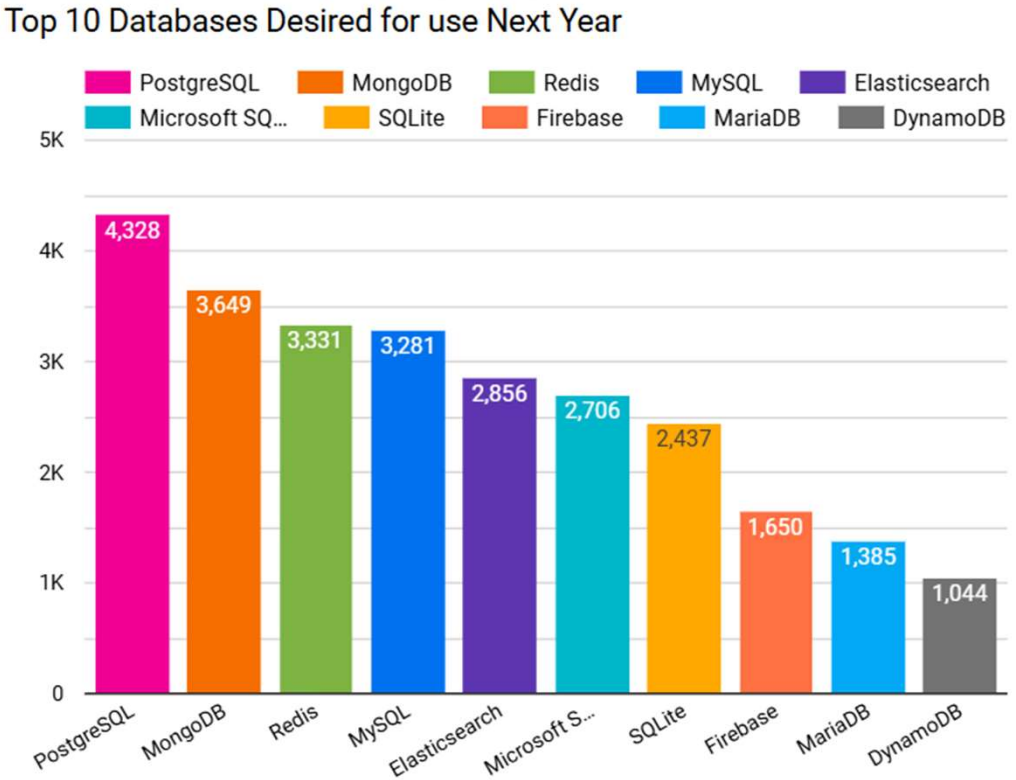
- Python's increasing traction is driven by the growing demand for AI and ML skills.

SQL's Relevance:

- SQL remains the most crucial language for data professionals.
- Aspiring data analysts, scientists, and business analysts must possess SQL skills to stay competitive

DATABASE TRENDS





IBM Developer

SKILLS NETWORK

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

Top 5 Most Used Databases Currently:

- MySQL
- Microsoft SQL Server
- PostgreSQL
- SQLite
- MongoDB
- **Projected Popular Databases:**
- PostgreSQL
- MongoDB
- Redis
- MySQL
- Elasticsense
- **Emerging Tools:**
- Redis and Elasticsense are relatively new and are expected to gain significant traction in the IT space

Implications

- SQL remains a crucial tool for data specialists.
- Companies continue to favor open-source databases.
- Oracle SQL is losing relevance, as it did not rank among the top 5 databases.

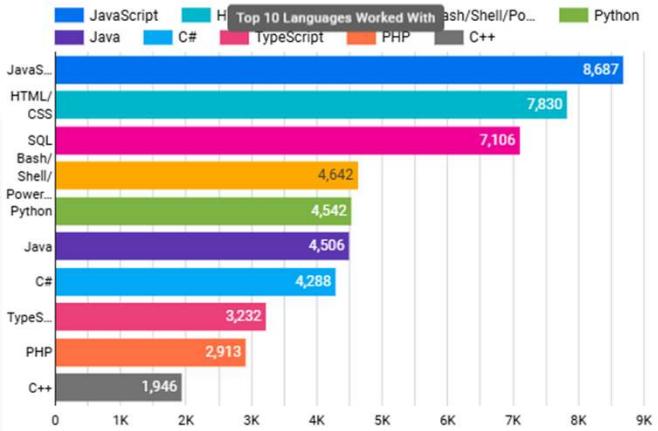
DASHBOARD

Permanent link to the Google Looker Studio Dashboard

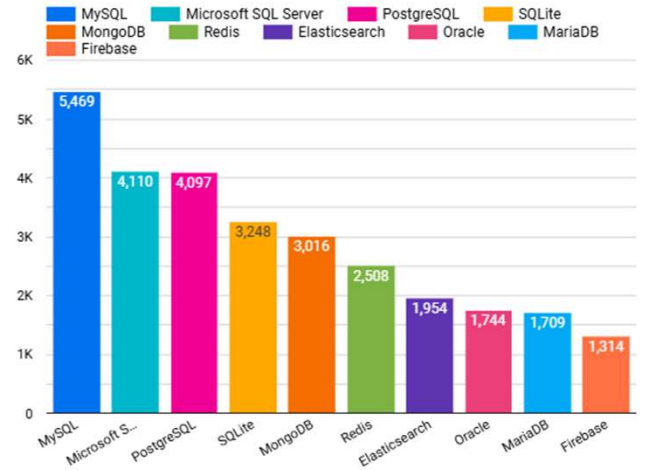
IBM-Data-Analyst-Capstone-Project/Google_Looker_Dashboard_Link_at main · OsvaldoOlivaresR/IBM-Data-Analyst-Capstone-Project (github.com)

DASHBOARD TAB 1

Top 10 Languages Worked With

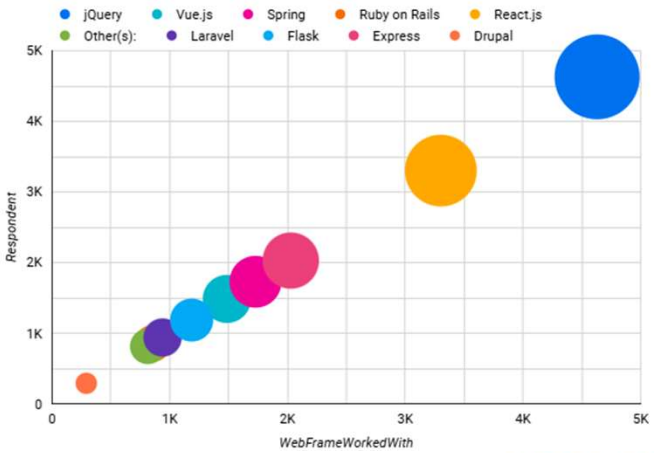


Top 10 Databases Worked With



Platform Worked With Word Cloud

Top 10 Web Frames Worked With

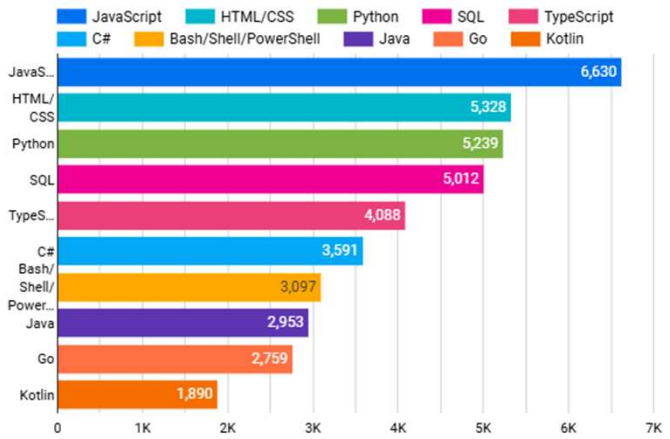


IBM Developer

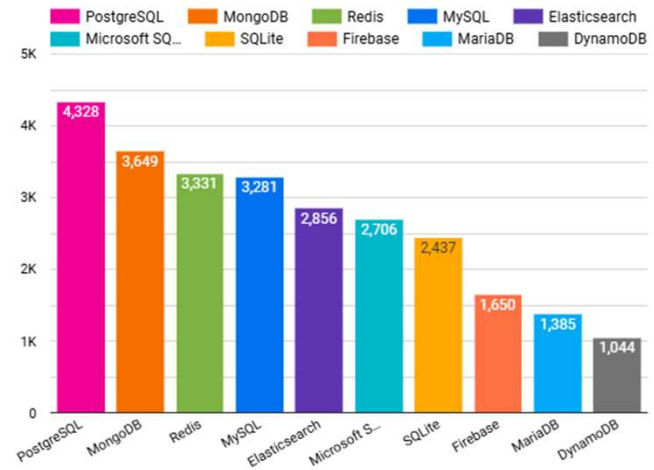
SKILLS NETWORK

DASHBOARD TAB 2

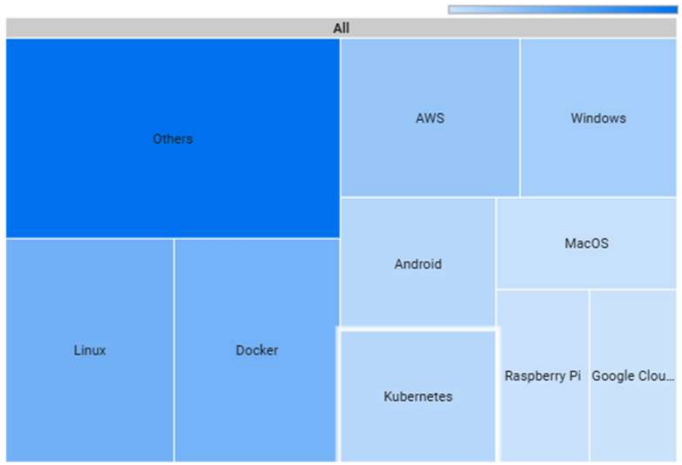
Top 10 Languages Desired for use Next Year



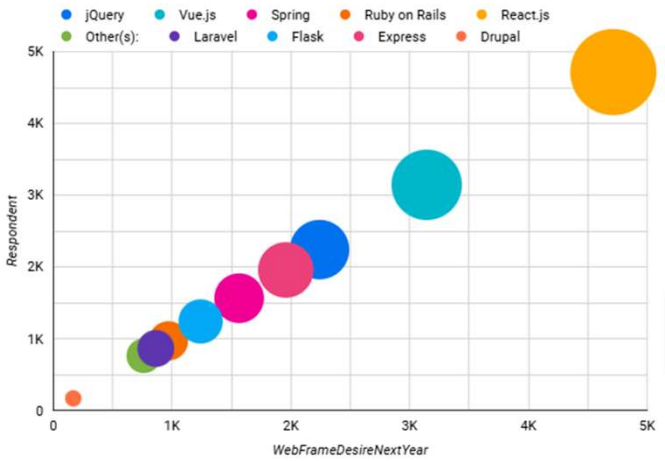
Top 10 Databases Desired for use Next Year



Platform Desire Next Year Tree map



Top 10 Webframes Desired for use Next Year

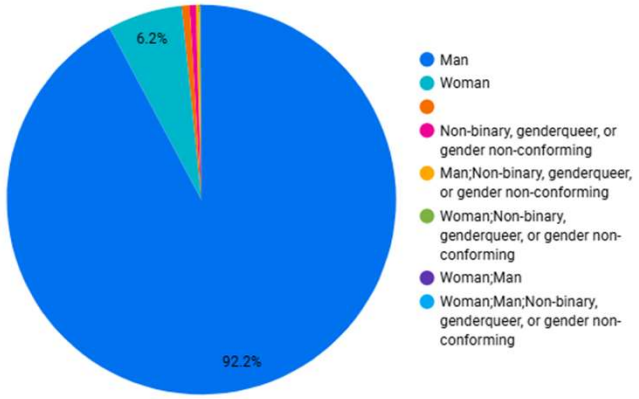


IBM Developer

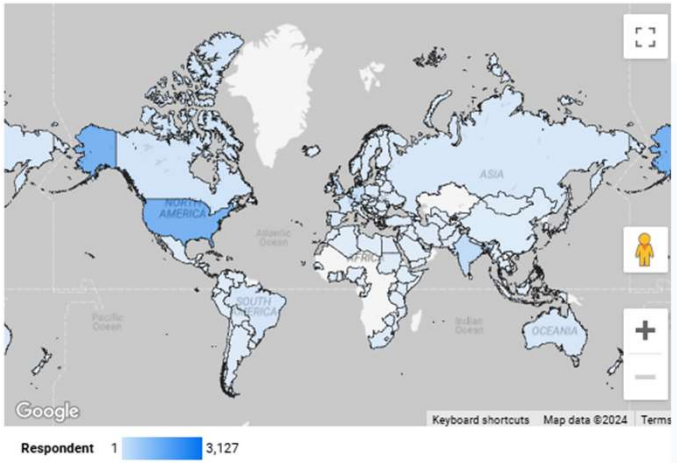
SKILLS NETWORK

DASHBOARD TAB 3

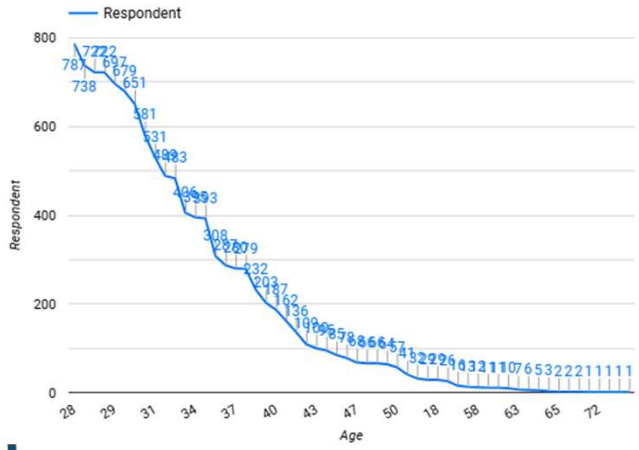
Respondents by Gender



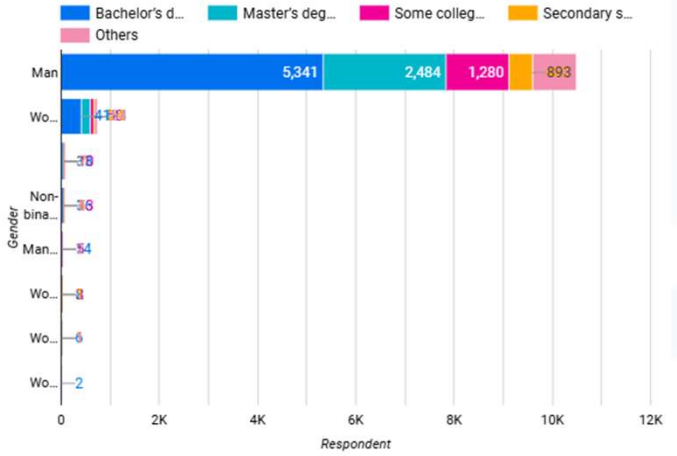
Respondent Count Map



Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level



IBM Developer

SKILLS NETWORK

DISCUSSION

Upskilling in the Technology Sector	• Strategies for continuous learning and professional development.
Closing the Gender Gap in Technology	• Initiatives and policies to promote gender diversity and inclusion.
Educational Requirements	• Evaluating the necessity of advanced degrees (Masters or Doctorate) in the tech industry.
Rising Demand for Mobile Development	• The growing popularity of Kotlin and its implications for mobile app development.
Expanding Tech Education and Access	• Enhancing technology education and development in underdeveloped regions such as Southeast Asia, South America, Africa, and parts of Europe.
Future Relevance of Oracle SQL	• Assessing the long-term viability and demand for Oracle SQL in the evolving tech landscape.

OVERALL FINDINGS & IMPLICATIONS

Findings

- **Educational Background:**

Most IT professionals hold a Bachelor's degree.

- **Popularity of Web Development Languages:**

Web development languages are currently the most popular and in-demand tools in the IT field.

- **Demographics:**

The tech sector is predominantly composed of young professionals under the age of 40.

- **Future Learning Trends:**

Most respondents are interested in learning PostgreSQL and React JS in the coming year.

Implications

- **Proficiency in NoSQL:**

Data professionals should develop skills in NoSQL databases alongside SQL to stay competitive.

- **Lucrative Web Development:**

Web development remains a highly lucrative skill in the IT sector.

- **Tech Training Access:**

There is a critical need for increased access to tech training and education in less developed countries.

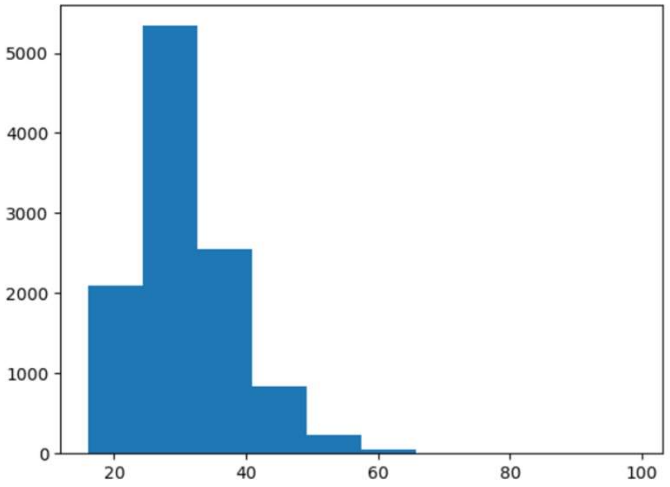
CONCLUSION

- **Embrace Change**: Continuous skill upgrades are essential for developers to stay relevant.
- **Diversifying Market**: The tech market is becoming increasingly diverse.
- **Lucrative Web Development**: Web development remains a highly profitable field for the future.
- **Educational Insights**: A Master's degree is not essential, as most developers hold a Bachelor's degree.

APPENDIX

In [12]:
```
# your code goes here
plt.hist(df["Age"])
```

Out[12]:
```
(array([2.094e+03, 5.337e+03, 2.557e+03, 8.420e+02, 2.250e+02, 4.900e+01,
        6.000e+00, 0.000e+00, 0.000e+00, 1.000e+00]),
 array([16. , 24.3, 32.6, 40.9, 49.2, 57.5, 65.8, 74.1, 82.4, 90.7, 99. ]),
 <BarContainer object of 10 artists>)
```



Finding correlation

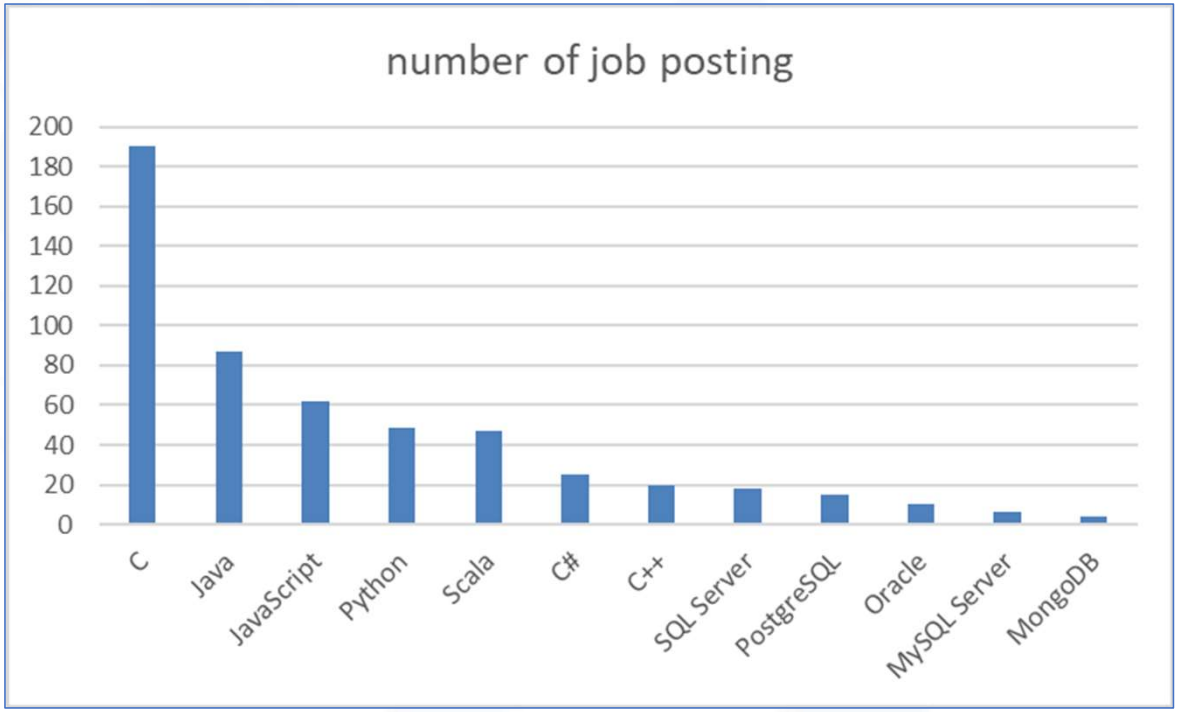
Find the correlation between `Age` and all other numerical columns.

In [24]:
```
# your code goes here
num_df = df.select_dtypes(include = ["number"])
corr_matrix = num_df.corr()
print (corr_matrix)
```

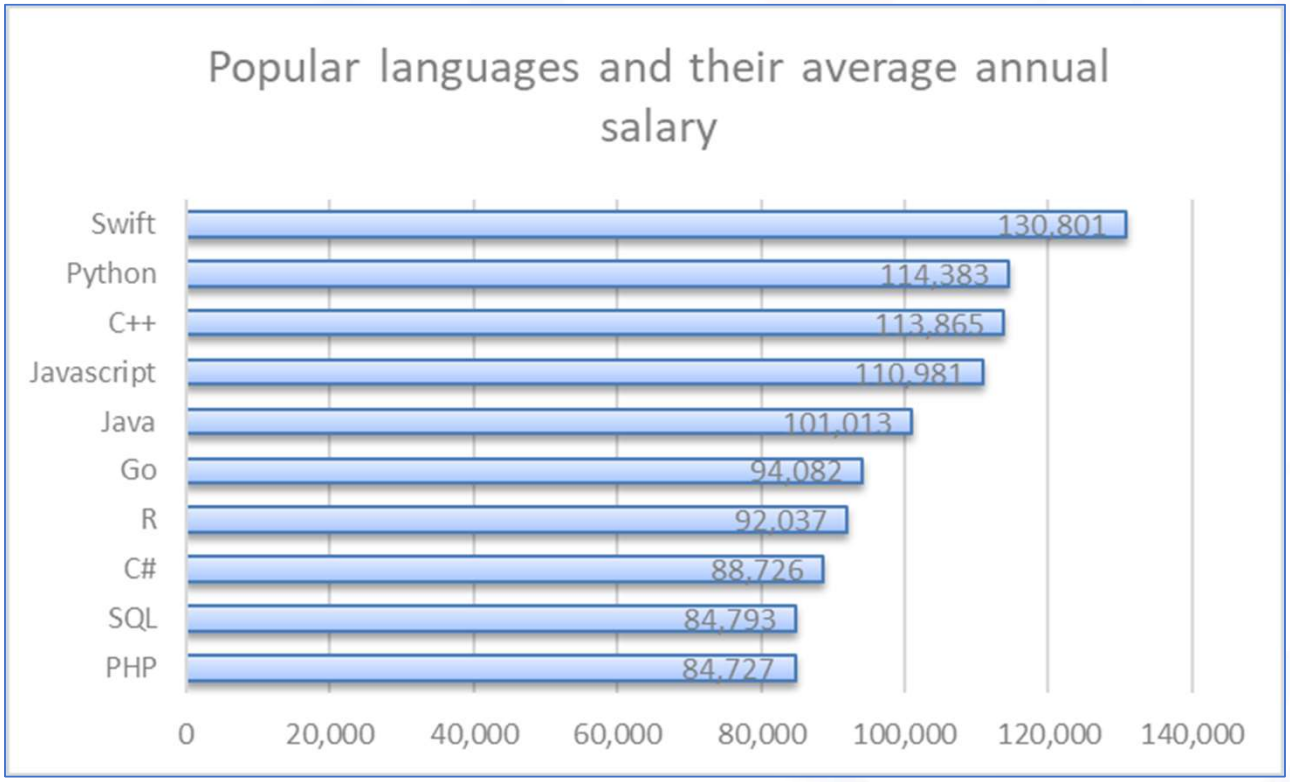
```
                Respondent  CompTotal  ConvertedComp  WorkWeekHrs  CodeRevHrs  \
Respondent        1.000000  -0.013490       0.002181    -0.015314    0.004621
CompTotal        -0.013490   1.000000       0.001037     0.003510    0.007063
ConvertedComp     0.002181   0.001037       1.000000     0.021143   -0.033865
WorkWeekHrs      -0.015314   0.003510       0.021143     1.000000    0.026517
CodeRevHrs        0.004621   0.007063      -0.033865     0.026517    1.000000
Age               0.004041   0.006970       0.105386     0.036518   -0.020469

                     Age
Respondent      0.004041
CompTotal       0.006970
ConvertedComp   0.105386
WorkWeekHrs     0.036518
CodeRevHrs     -0.020469
Age             1.000000
```

JOB POSTINGS

Bar chart illustrating job posting data retrieved from the GitHub Jobs API.



POPULAR LANGUAGES



- Bar chart showcasing popular programming languages and their average annual salaries, based on data scraped from GitHub job listings and stored in a CSV file.