



DATA ANALYTICS

NOW & TOMORROW

PRASHANT GURBAXANI

18-FEB-2021

OUTLINE



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

EXECUTIVE SUMMARY



- Data is a part of our everyday lives.
- Data not only describes the present but can also be used to predict the future.
- Data Analytics is changing with development in
 - Programming languages
 - Databases
 - Platforms
 - Webframes
- Requirement for data analysts has increased & is growing.
- Data Analysts need to upskill their portfolio to be ahead of the competition.

INTRODUCTION



- Data Analytics is a vital part of any business.
- Data can show what the business is doing right & where improvements are needed.
- Data can be used to make accurate predictions for growth.
- Visualizing data is easier for the stakeholders to understand
 - Where the business or industry stands currently
 - What is trending in their line of business
- Data analytics is growing & changing at a rapid pace.
- Sourcing the correct people for the job is a daunting task for many employers with the variety in skillsets & software.

METHODOLOGY



- 3 datasets were used for this presentation.
- The data used for this project is a modified subset of the survey results coming from the source:

<https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>

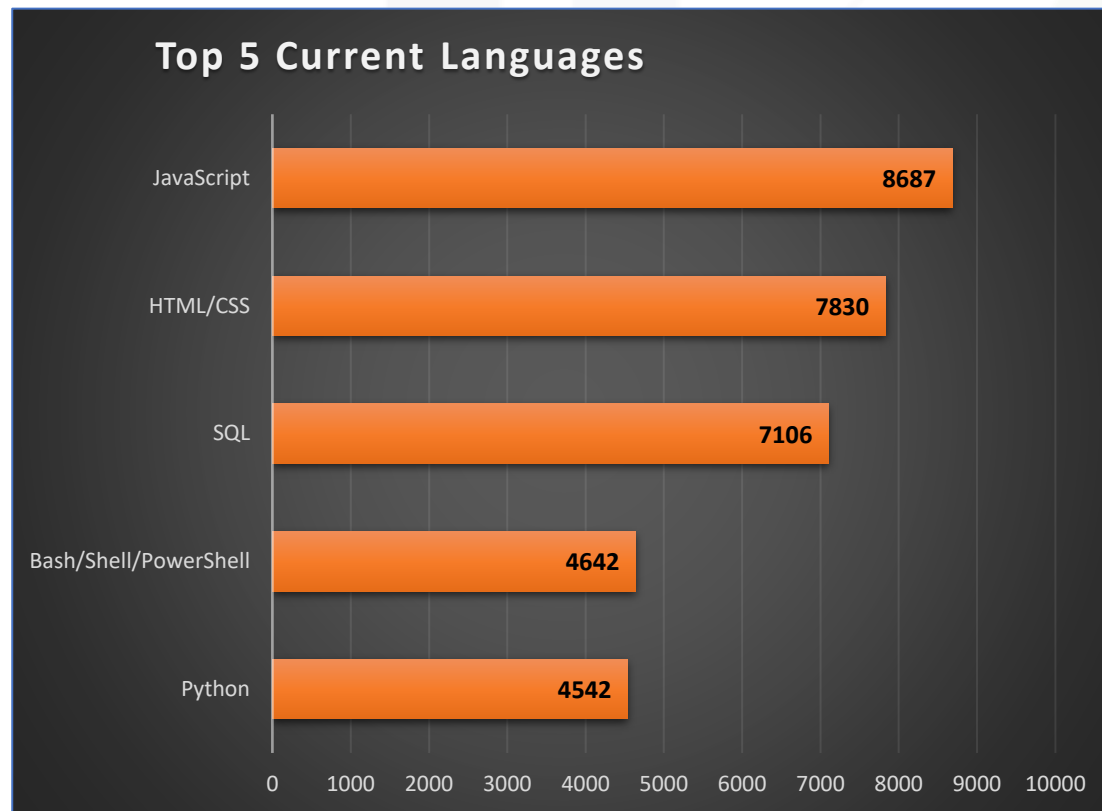
- The parameters used from the survey involve languages, databases, platforms & webframes that developers
 - Work with currently
 - Would like to work with in the future
- A second dataset regarding the different job postings on GitHub was collected using the GitHub API.
- The third dataset was sourced using webscraping tools to get the average annual salary of developers in different technologies from a survey.

RESULTS

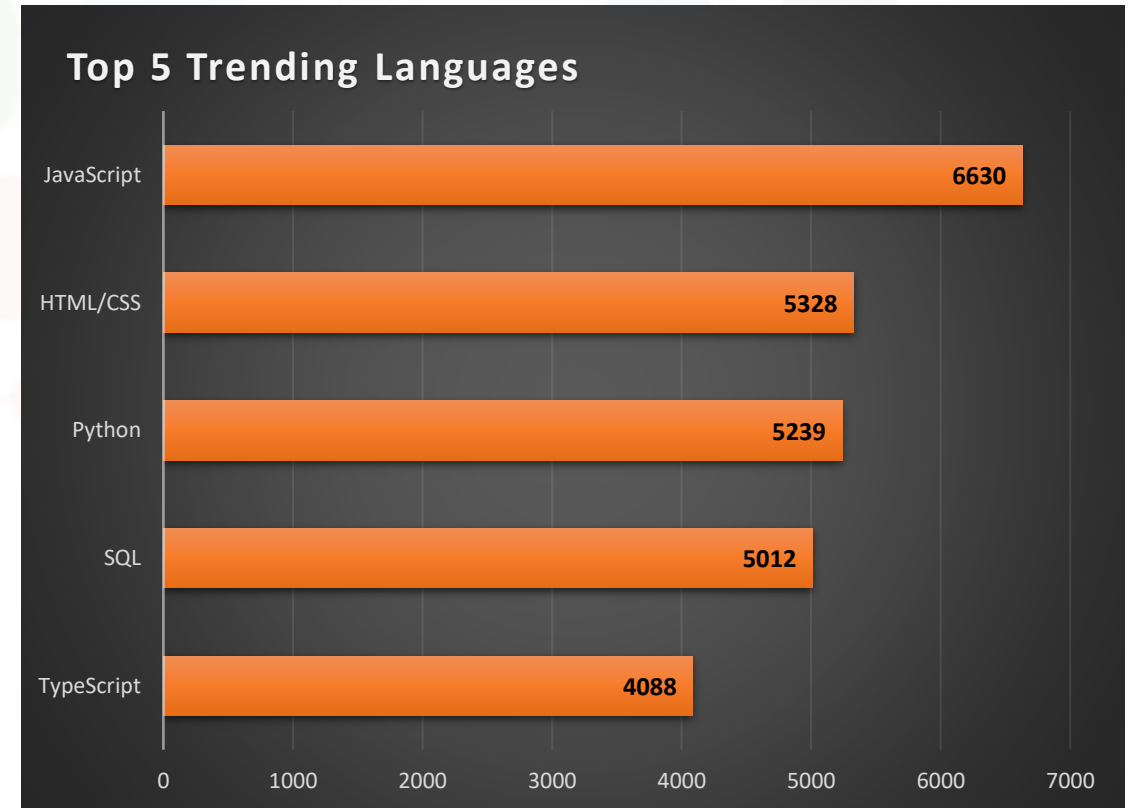
- The subset of the StackOverflow survey has 11399 respondents.
- There are 29 programming languages in the survey.
- The number of unique databases was 15.
- 18 different platforms were identified.
- 14 Webframes were recognized.
- The data obtained from the GitHub API collected job postings related to 12 data technologies
- The data from the webscraping resulted in annual average salaries for the top 10 programming languages.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

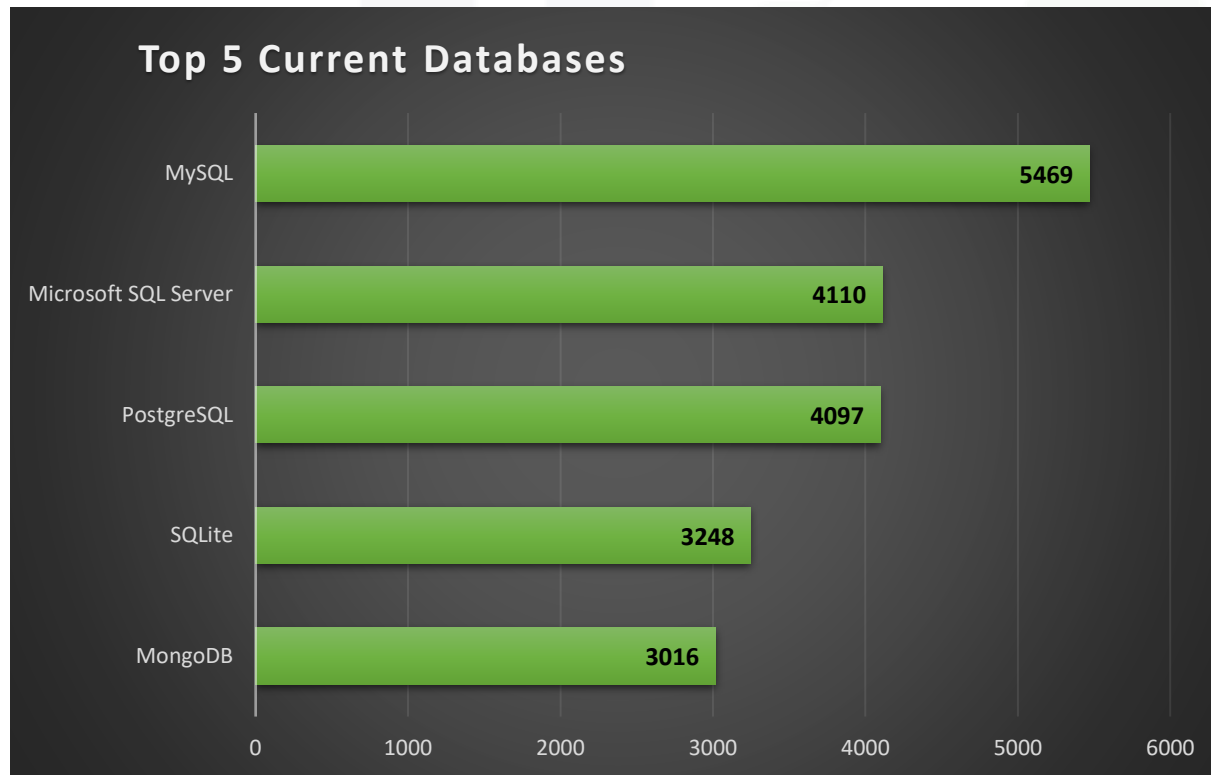
- JavaScript & HTML continue to be the topmost languages.
- Python has risen to number 3 in the desired language table.
- Exit of Powershell from Top 5 trends

Implications

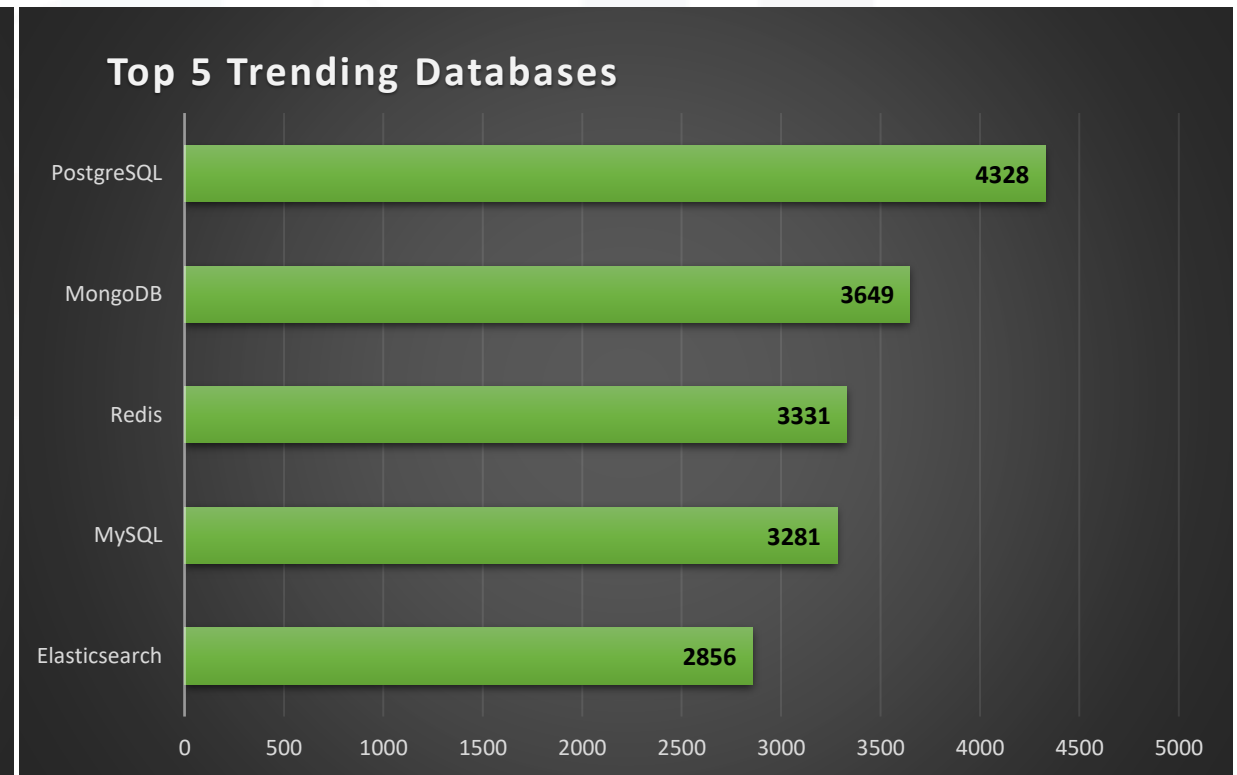
- Though JavaScript is still trending on Top, Python is catching up.
- The ease of using Python is its biggest advantage.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Databases based on SQL on top currently.
- Trending databases are NoSQL based.
- PostgreSQL is trending at the top for next year databases.

Implications

- Growth in NoSQL databases shows the change in database structures.
- PostgreSQL shows the global trend for people to prefer open-sourced technology.

DASHBOARD



[Click here to access](#)

The dashboard is a presentation of the current technologies as well as what the future trend are.

Current Technology Usage	Future Technology Trend	Demographics
--------------------------	-------------------------	--------------

Demographics

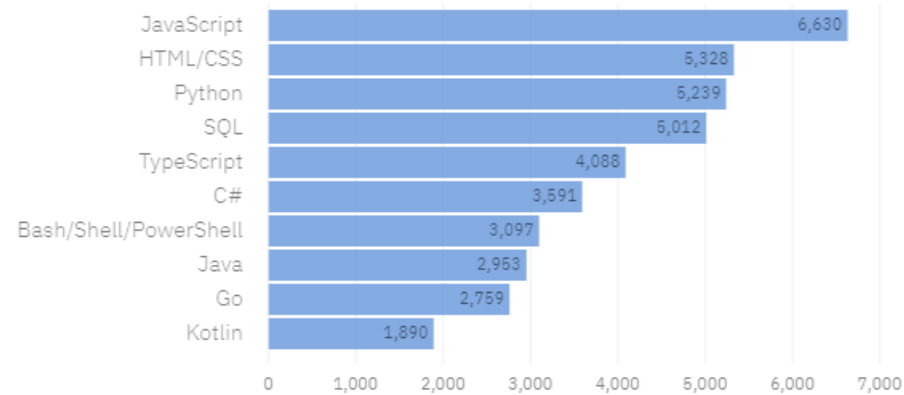
DASHBOARD TAB 2 – Future Technologies

Current Technology Usage

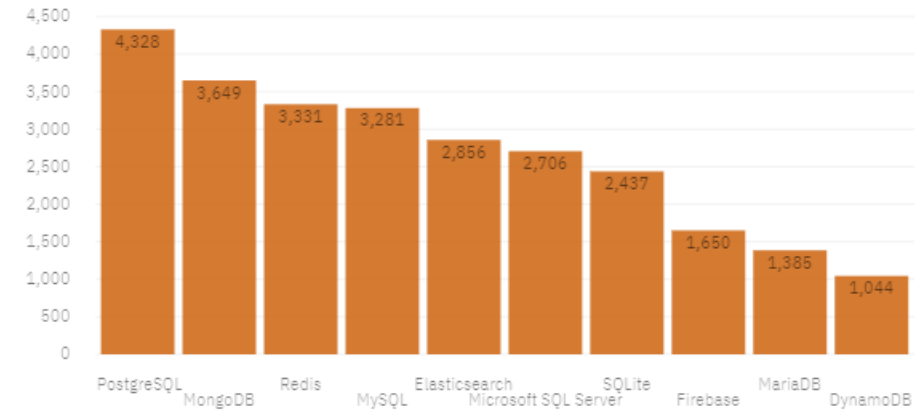
Future Technology Trend

Demographics

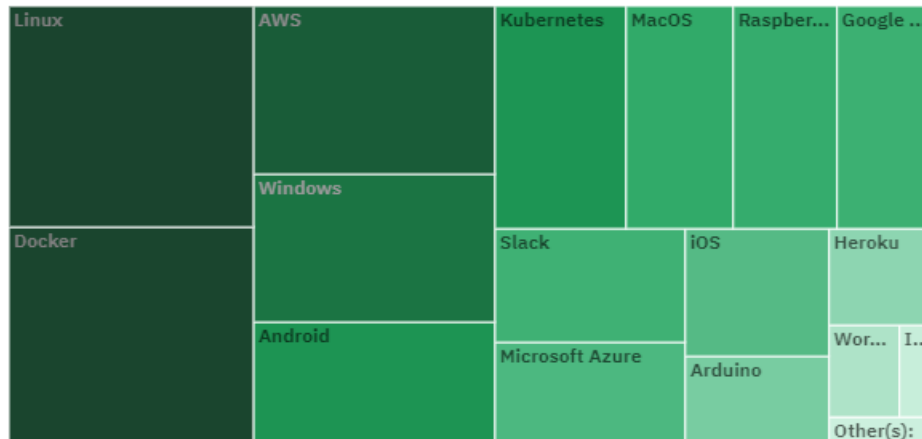
Top 10 Trending Languages



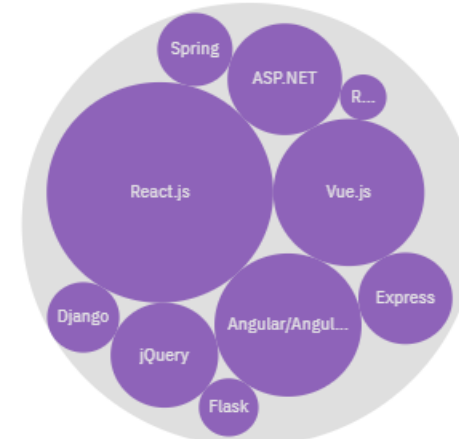
Top 10 Trending Databases



Trending Platforms



Top 10 Trending Webframes



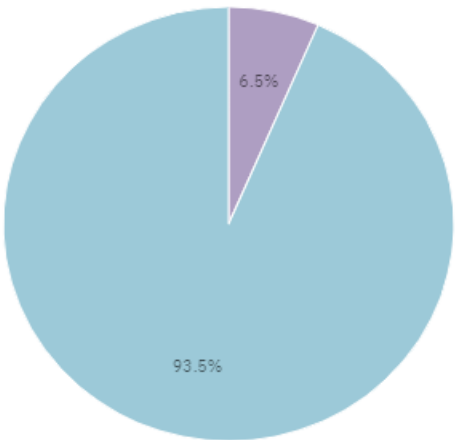
DASHBOARD TAB 3 – Demographics for survey

Current Technology Usage

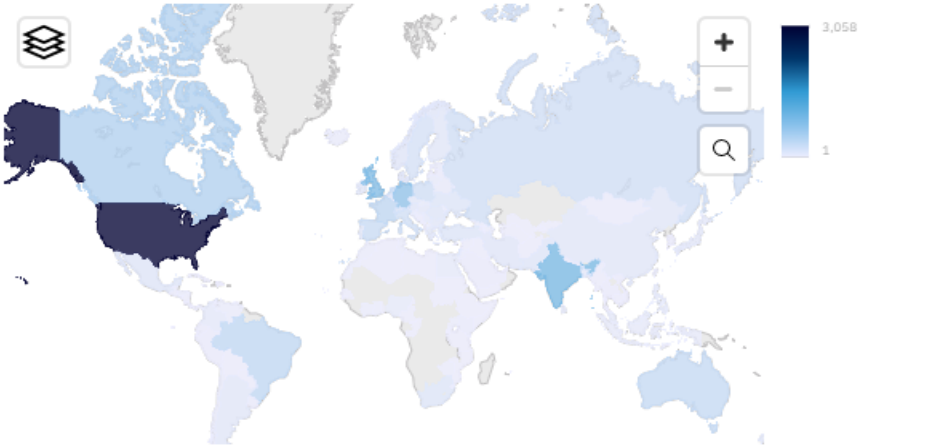
Future Technology Trend

Demographics

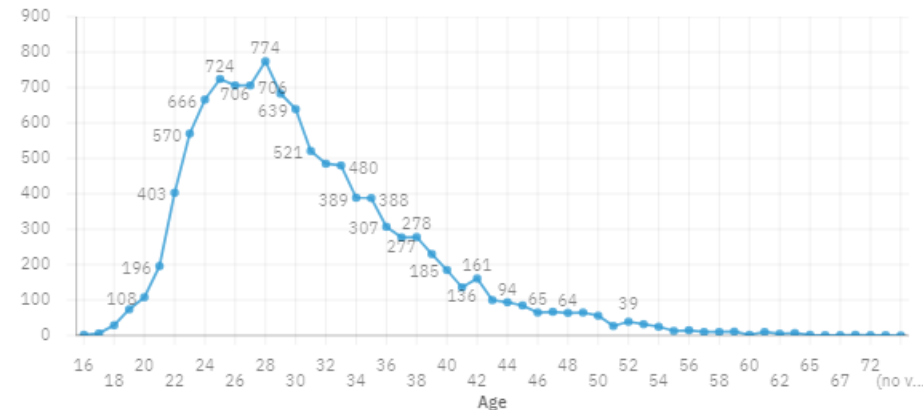
Respondents by Gender



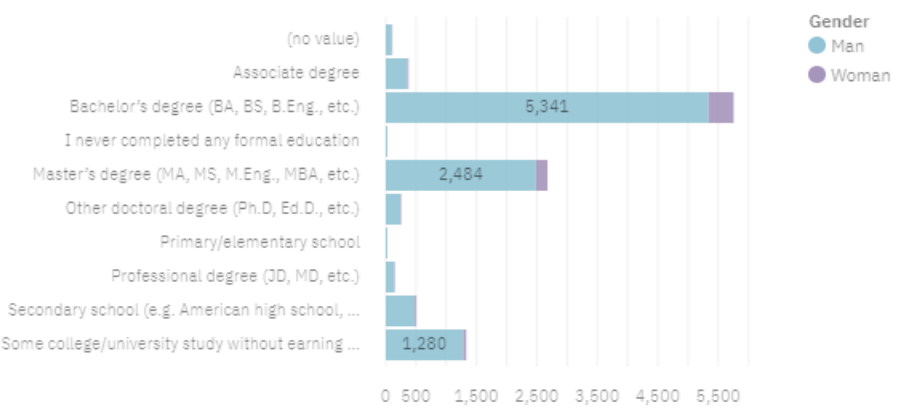
Distribution of Respondents



Age Distribution



Formal Education Level



DISCUSSION



- Data Analytics is an important part of any business nowadays.
- Data Analysts need to be able to transfer numbers into pictures to make data analysis more impactful.
- Open source technologies are becoming more popular, especially when related to data.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Growth of open source technologies.
- Python looks to becoming the main programming language.
- NoSQL databases on the rise.

Implications

- Data Analytics becoming more versatile.
- Python can be used for other types of programming and software development besides web development.
- Data Analysts need to evolve and upgrade skillsets.

CONCLUSION



- Digital transformation is being led by open source technologies.
- A Data Analyst has to be a mathematician as well an artist.
- Understanding data will help enterprises grow.
- Data Analytics can help in securing a brighter future for all.

APPENDIX

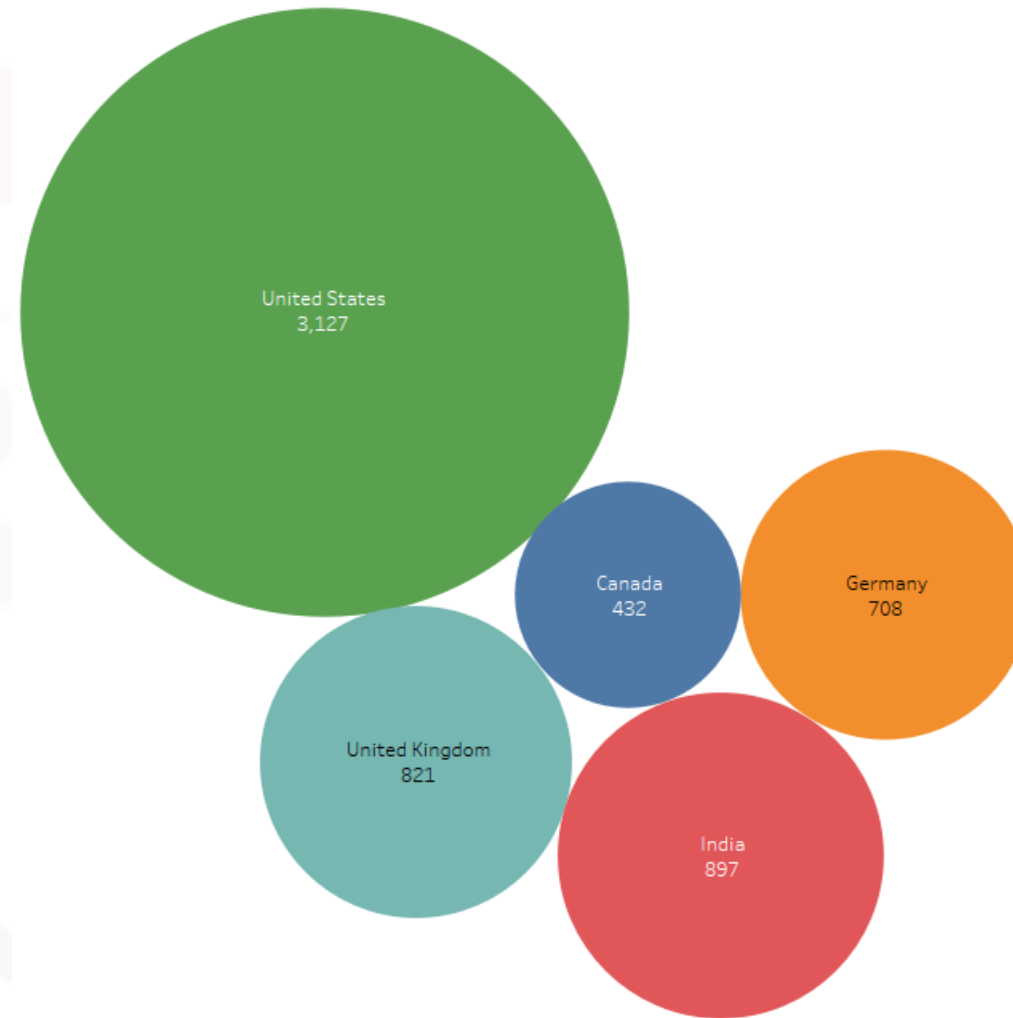


Other visualisations:

1. Top 5 Countries by Respondents
2. Gender Disparity
3. Salary Vs Age

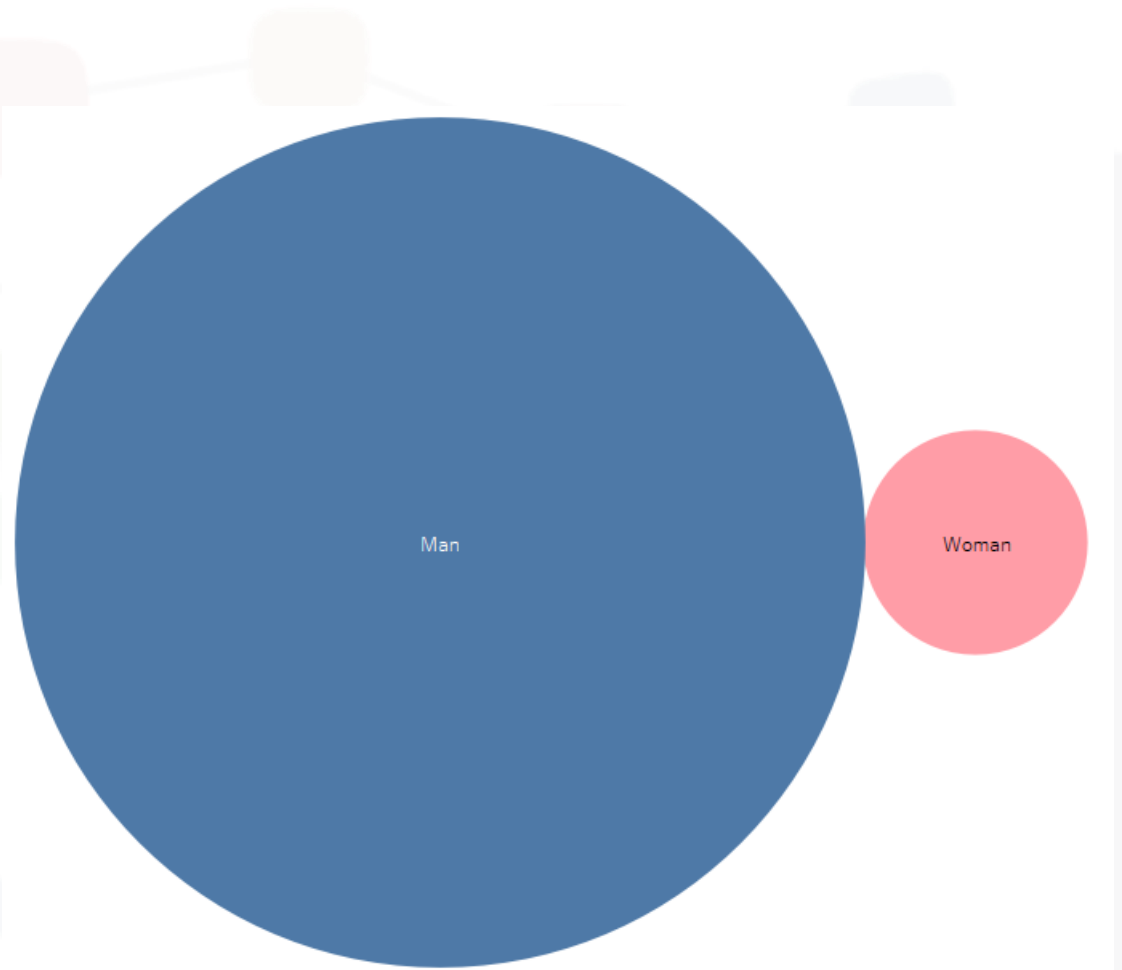
Top 5 Countries by Respondent

- Most of the survey respondents come from United States.
- The emergence of Indian programmers in technology development.



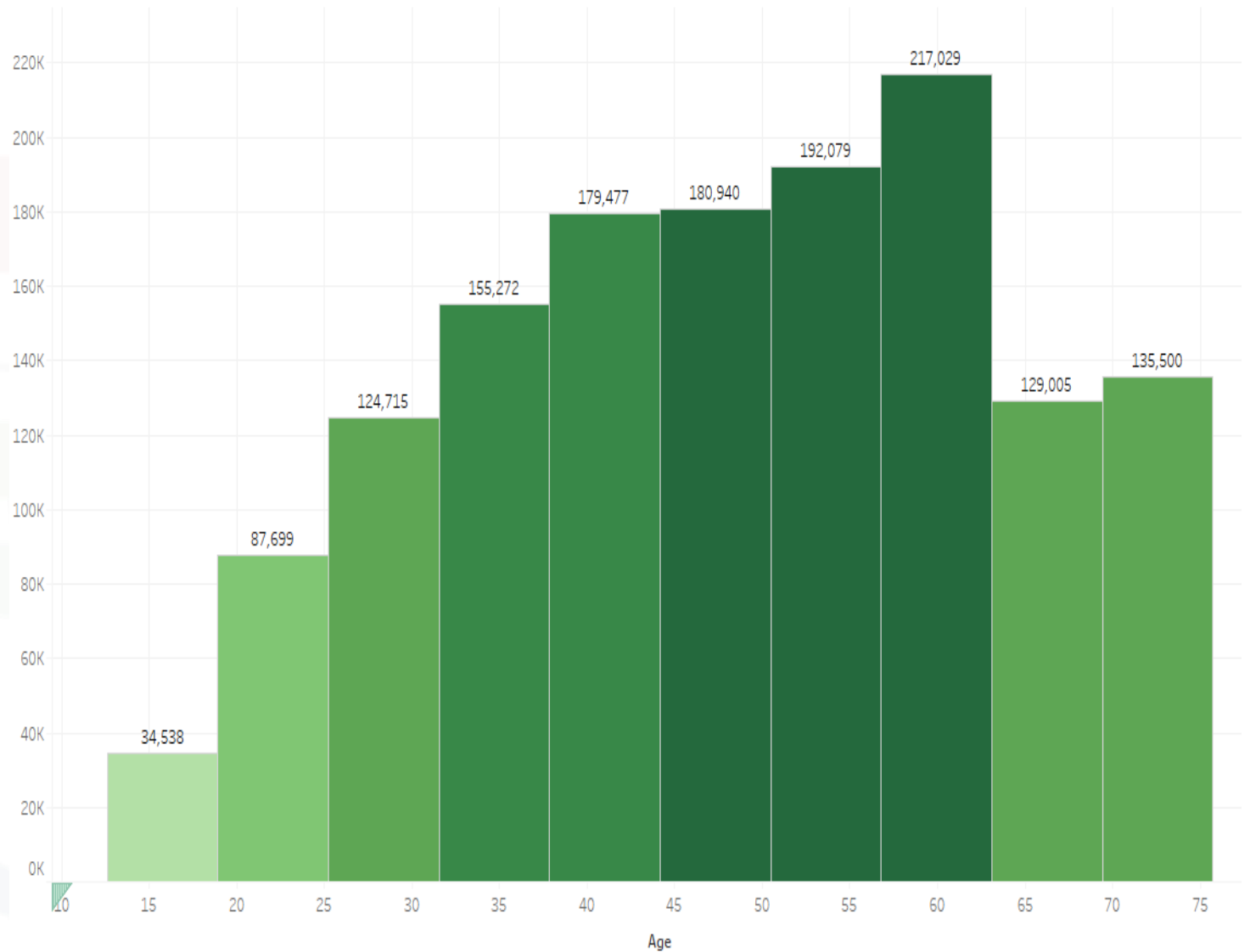
Gender Disparity

- Men outnumber women in the industry by a ratio of 14:1
- Possible causes of industry not attracting women could be long work hours.

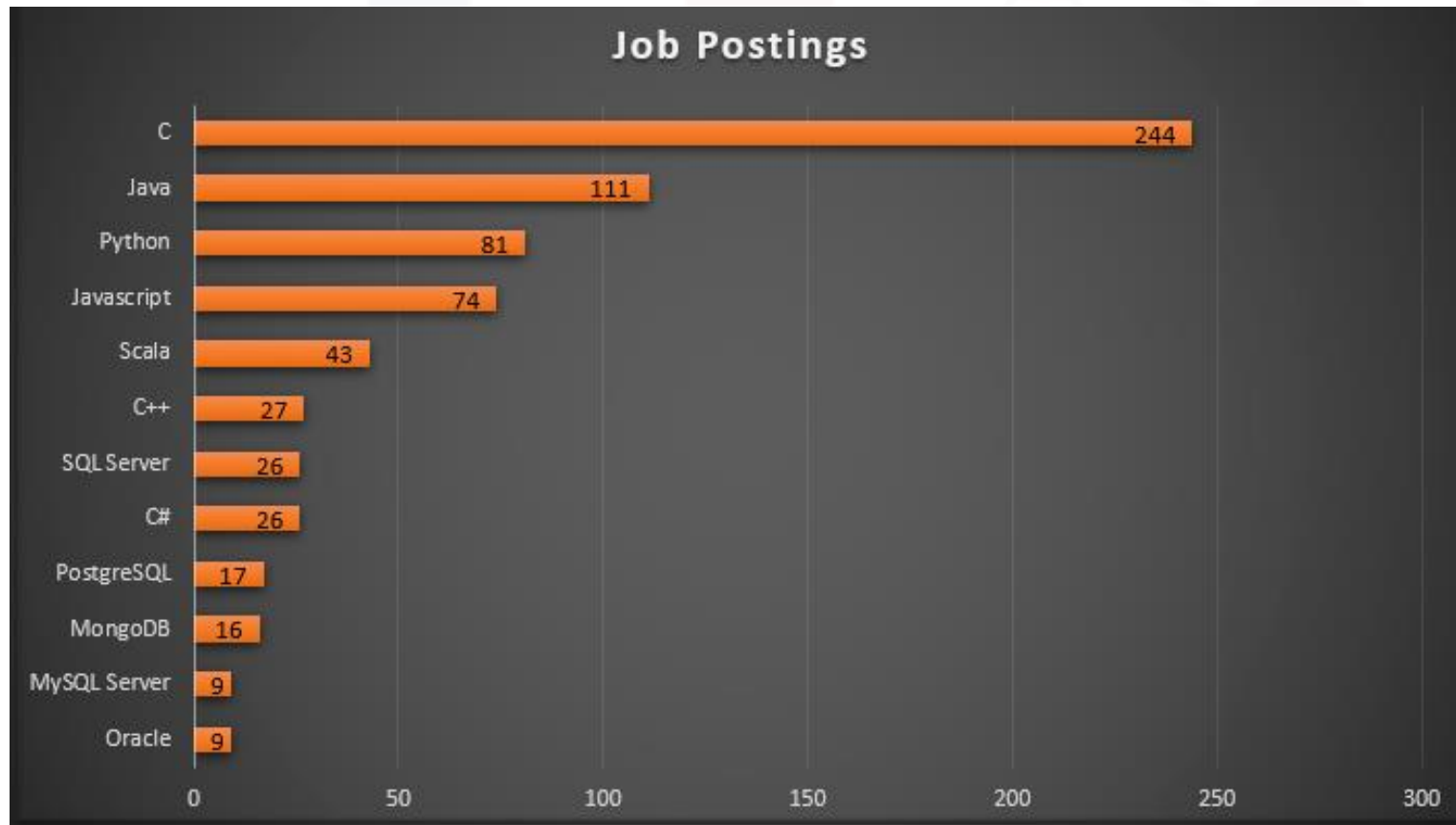


Salary & Age

- Professionals between 40 & 65 earn the most.
- Salaries are proportionate to experience.
- No specific retirement age.



GITHUB JOB POSTINGS



- C language seems to have the maximum requirement in Github job postings.

POPULAR LANGUAGES SALARIES



Python developers are highly paid compared to most other developers.