



Analyzing Programming Realm

Fatih Çağlar
5/22/2024

OUTLINE



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

EXECUTIVE SUMMARY

- We aim to analyze the trends between developers and job postings at this project.
- Some questions:
 - What is the distribution of the compensations of developers?
 - How other features are correlated with age?
 - How is gender affected on compensation?
- The findings can be used to inform educational programs, developer training initiatives, and recruitment strategies.



INTRODUCTION



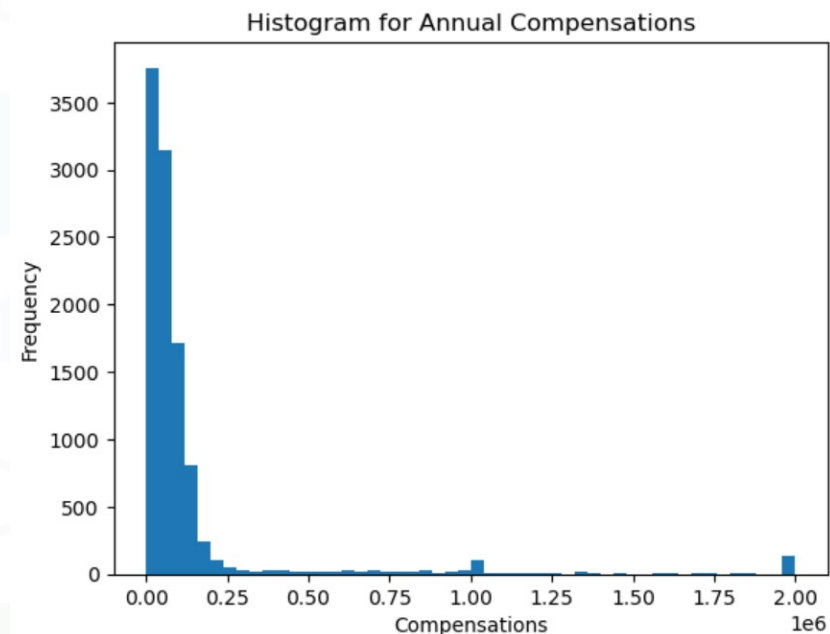
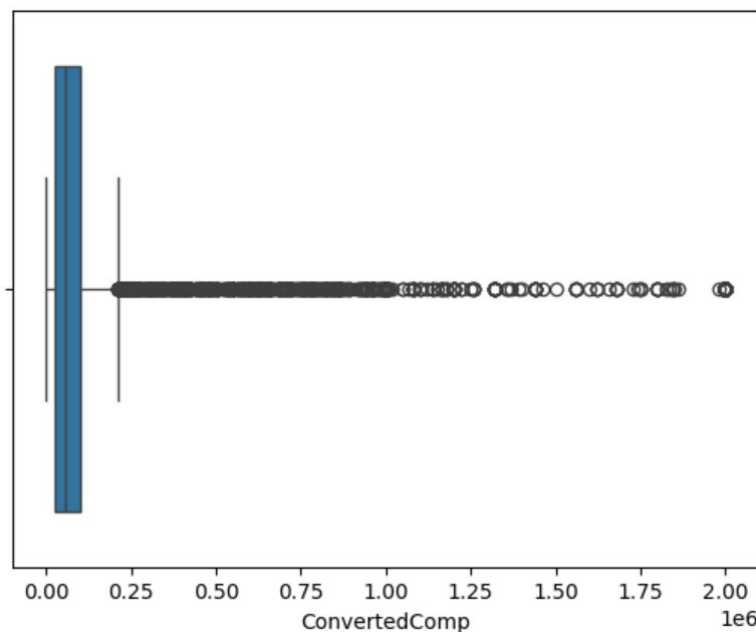
- We mainly focus on Stackoverflow survey data.
- The null value proportions for columns is tolerable.
- There are 11552 observations for 85 different features.
- The mean age for developers is 31.
- The developers are from 135 different countries and most of them are from United States and India.

METHODOLOGY



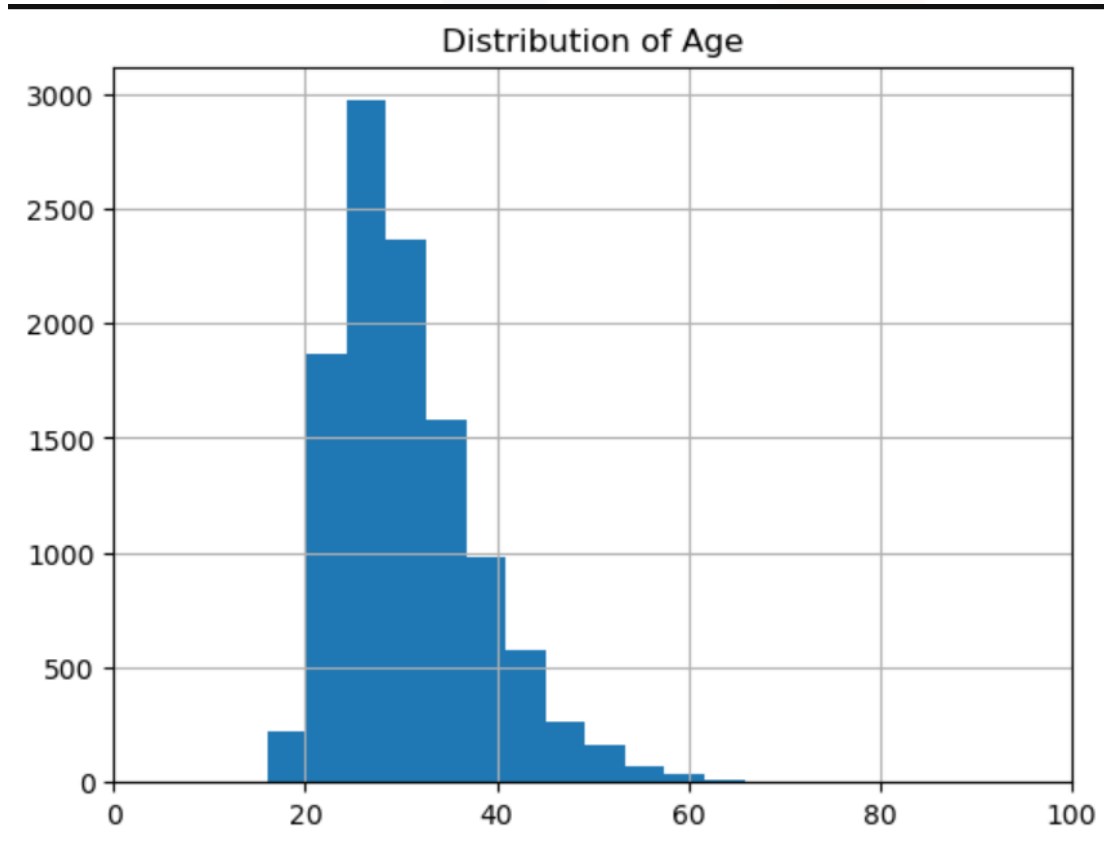
- We acquired the data using Github jobs API, web scraping and Stackoverflow 2019 Developer Survey. We cleaned and transformed the data for analysis.
- Data Wrangling:
 - Dropped duplicates
 - Imputed missing values
 - Normalized data
- Exploratory Data Analysis
 - Comparisons using gender, age & annual Compensation
- Data Visualizations & Dashboards
 - Visualized trends over programming languages, databases, web frameworks, countries and more other features.

RESULTS – Salary Distribution



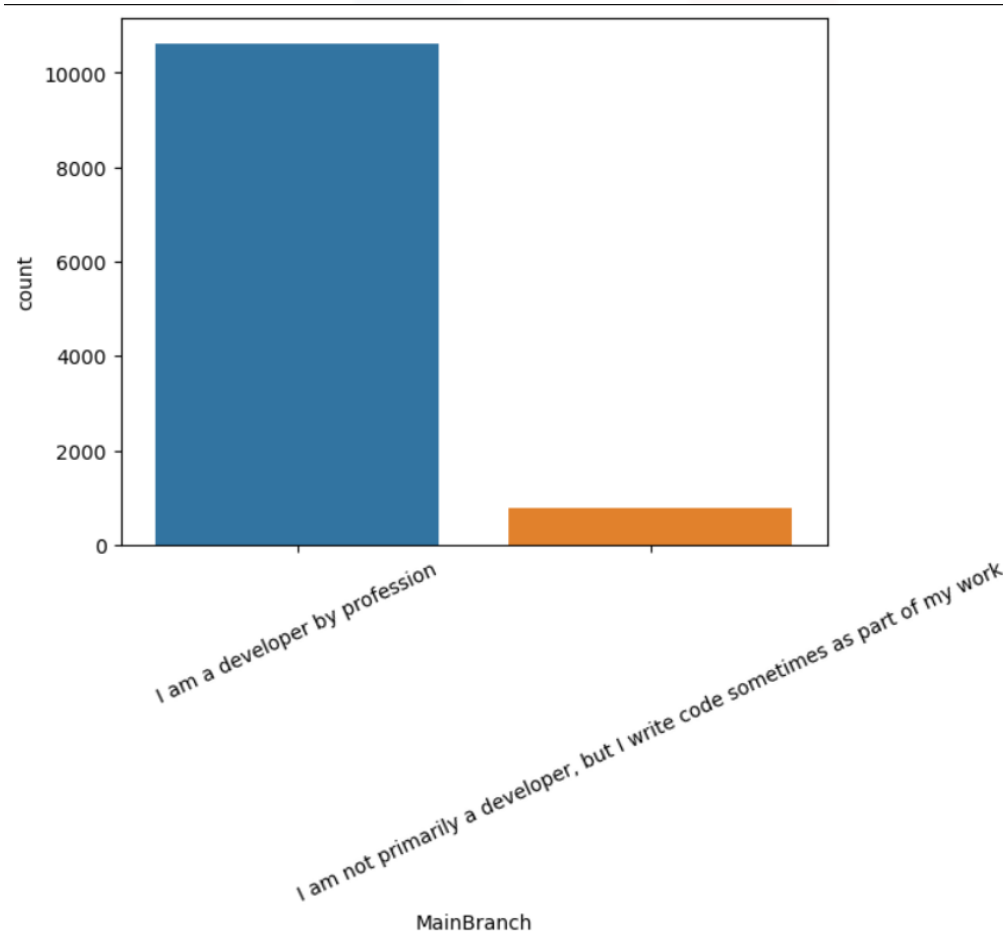
- The distribution of annual compensations is highly right-skewed and there are so many outliers. Median compensation is \$58k and mean compensation \$132k as a result of this skewness.

RESULTS – Age Distribution



- Median age is 25 and distribution of age is slightly right-skewed.

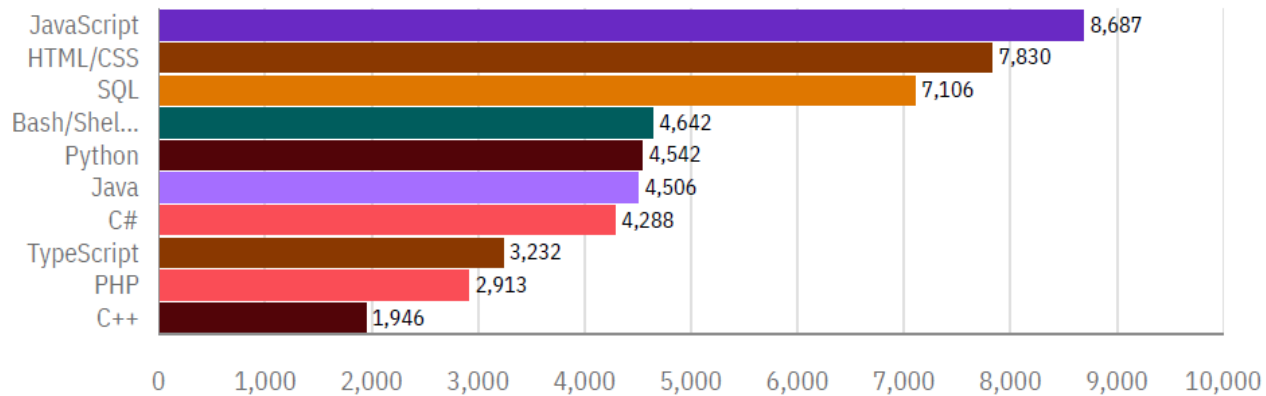
RESULTS – Main Branches



- The number of developers using Stackoverflow are nearly ten times greater than users from other branches

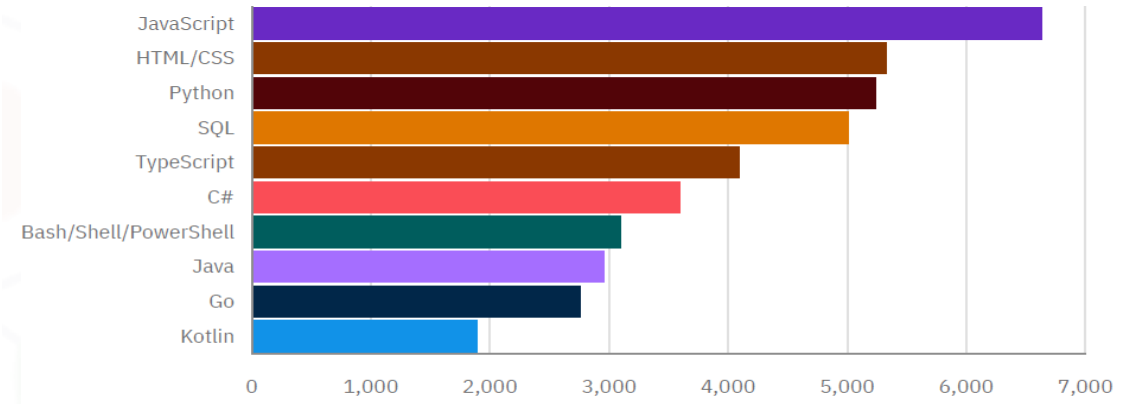
PROGRAMMING LANGUAGE TRENDS

Current Year



Respondent (Count)

Next Year



Respondent (Count)

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

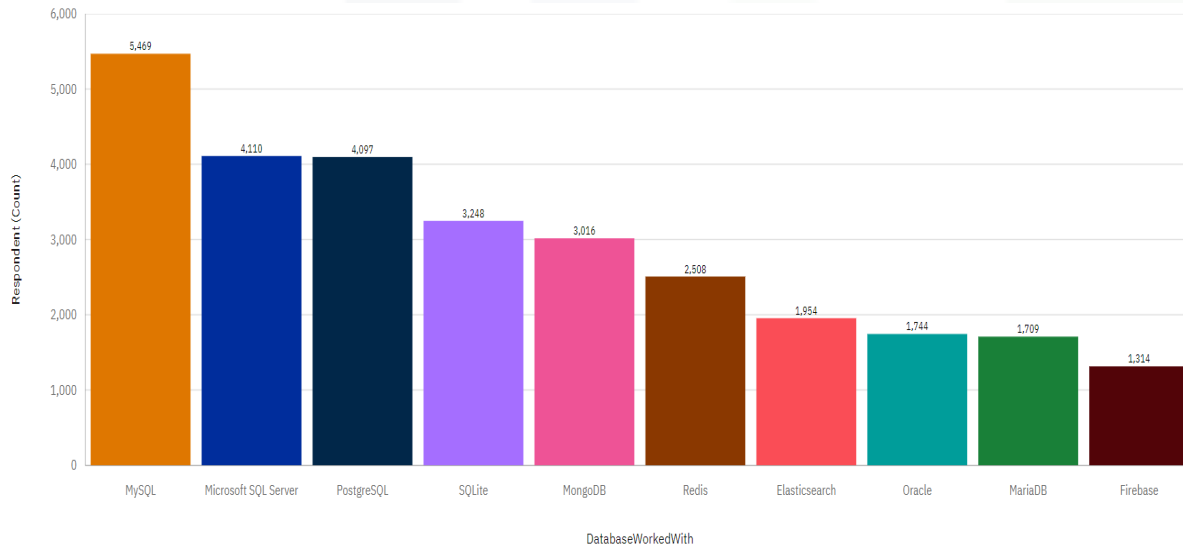
- Javascript, HTML/CSS and SQL are the top 2 languages used in 2019 and desired to use in 2020. Python & SQL comes right after.
- C++ & PHP are least used languages out of top 10 and do not exist in top 10 desired languages for 2020.



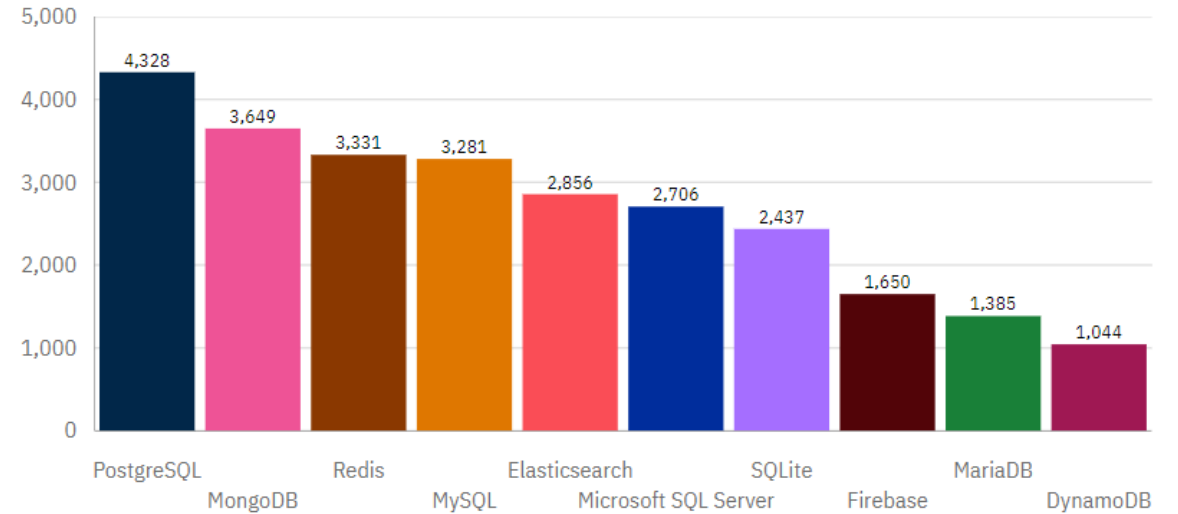
- The data suggests a move towards web-based technologies and away from traditional desktop applications.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

- MySQL is most used database in 2019 but losing its popularity for next year.
- MongoDB and Redis are gaining popularity while SQL servers stay behind except PostgreSQL.



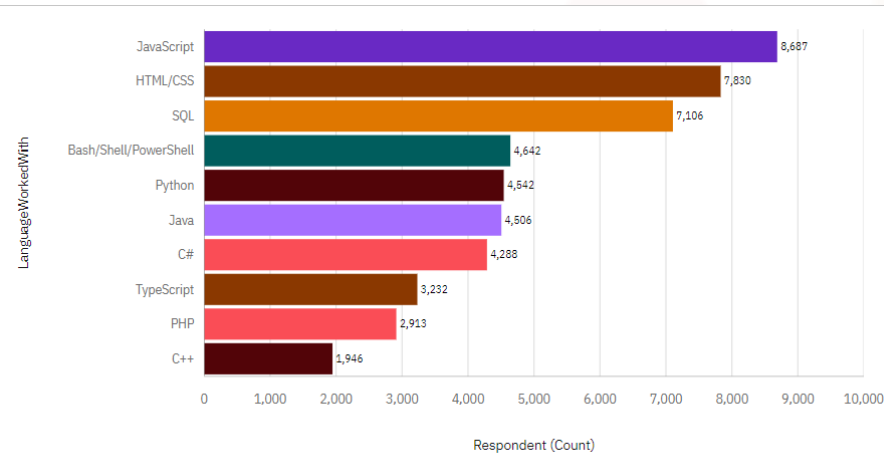
- PostgreSQL's leading position suggests a potential shift towards more scalable and feature-rich database solutions.
- NoSQL databases are desired to use by developers.

DASHBOARD

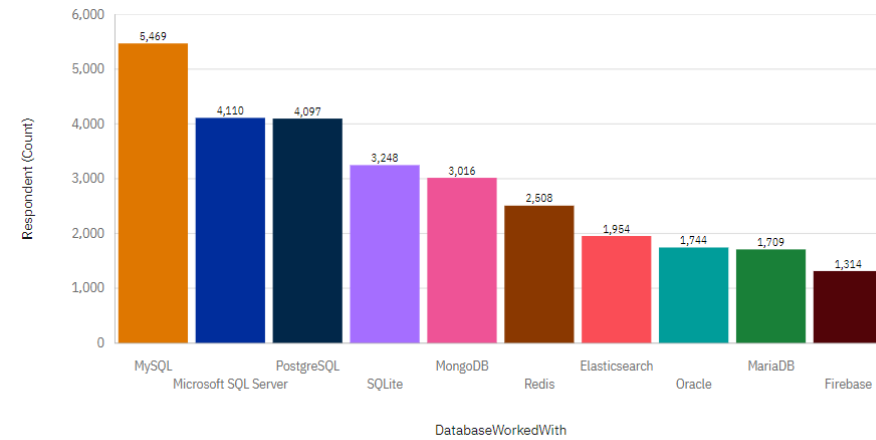


[https://github.com/mf-caglar/ibm-data-analyst-capstone/blob/main/Survey Dashboard.pdf](https://github.com/mf-caglar/ibm-data-analyst-capstone/blob/main/Survey%20Dashboard.pdf)

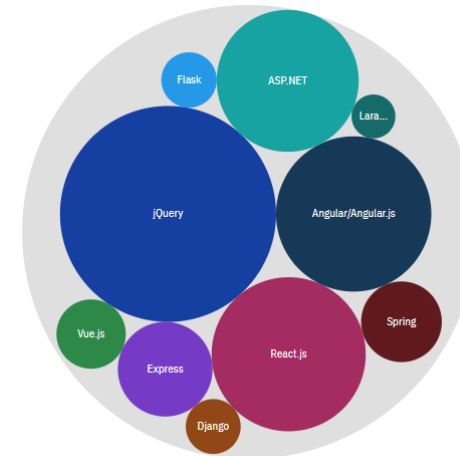
DASHBOARD TAB 1



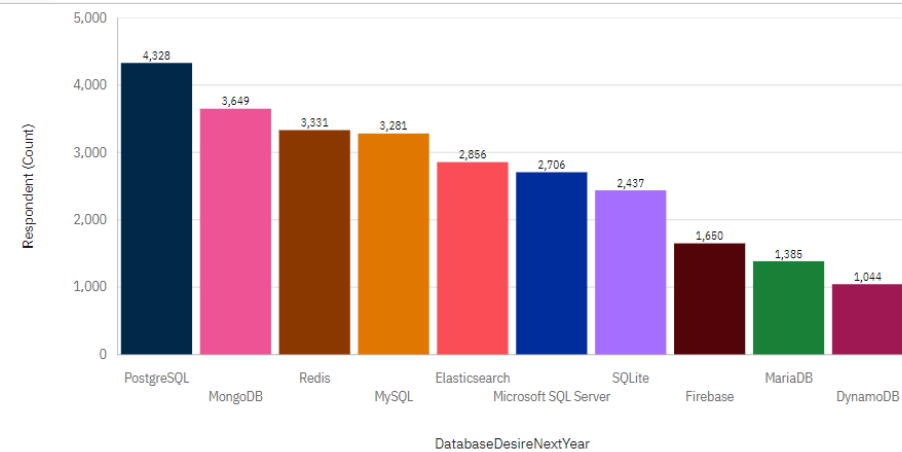
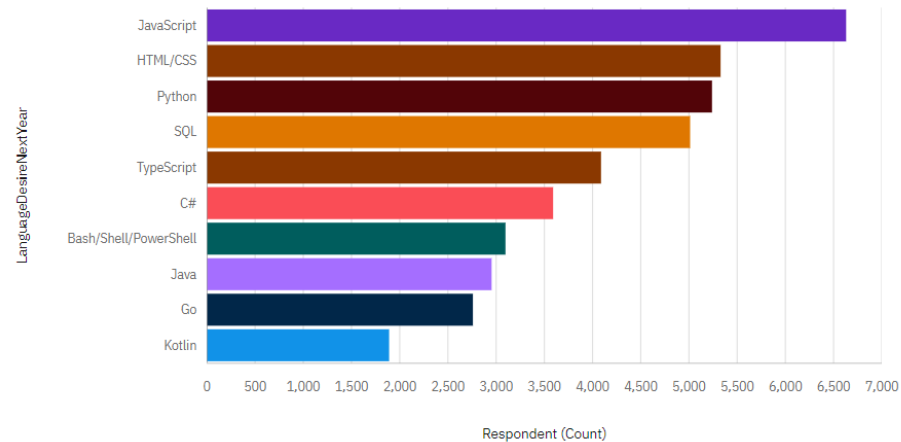
Top 10 Platforms Developers Worked With



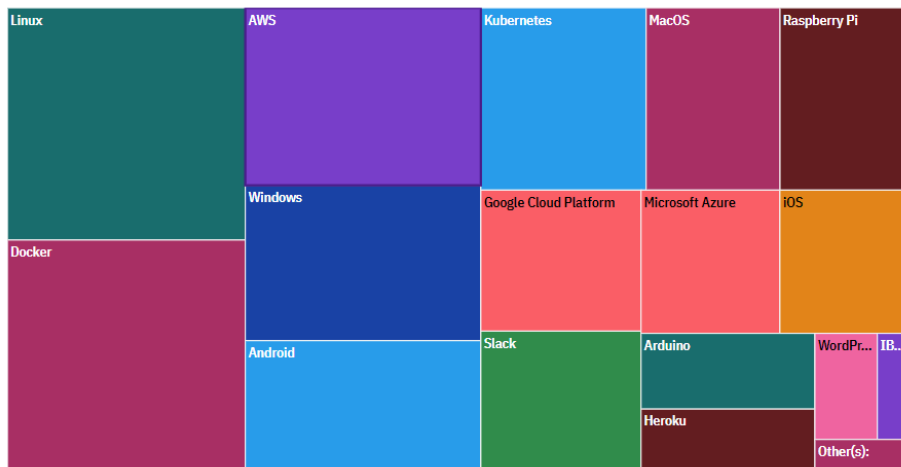
Top 10 Webframes Developers Worked With



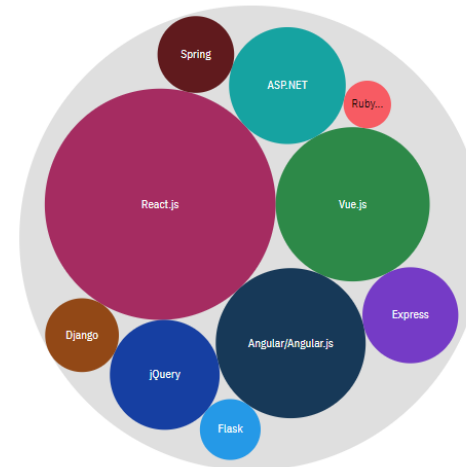
DASHBOARD TAB 2



Platforms Desired Next Year



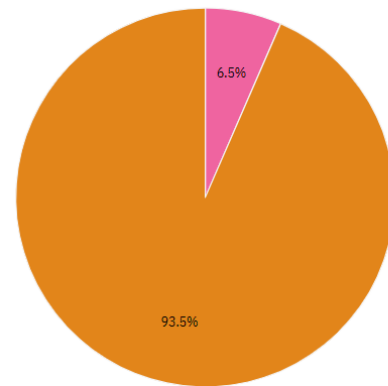
Top 10 Webframes Desired Next Year



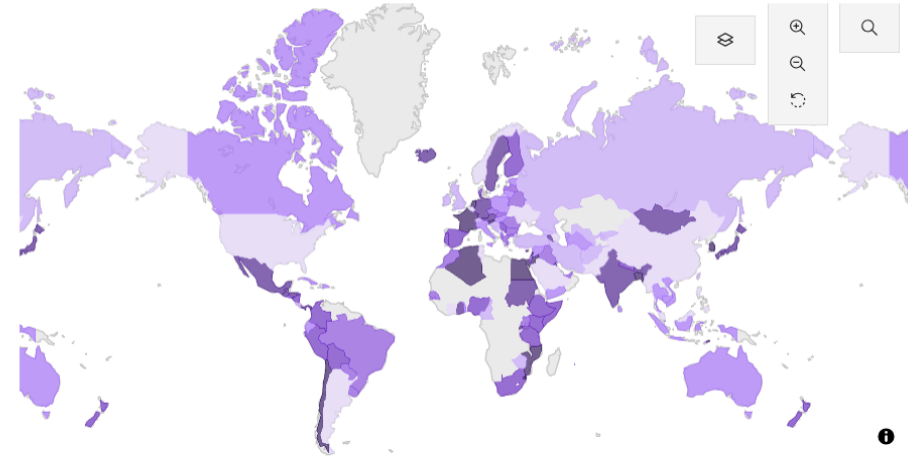
DASHBOARD TAB 3

Respondent Count by Gender

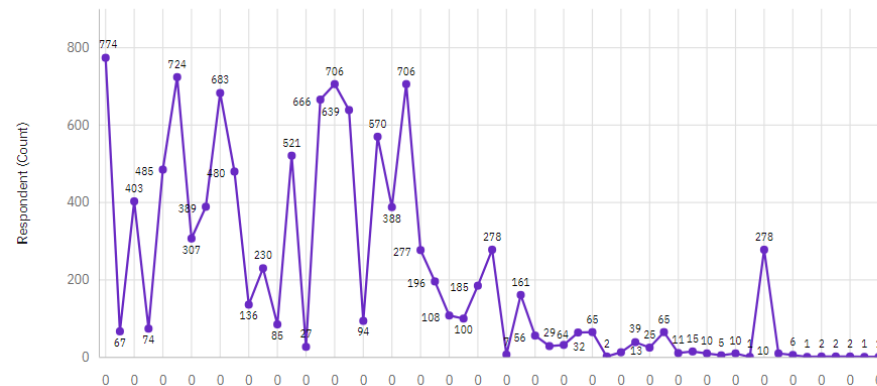
Gender
● Woman ● Man



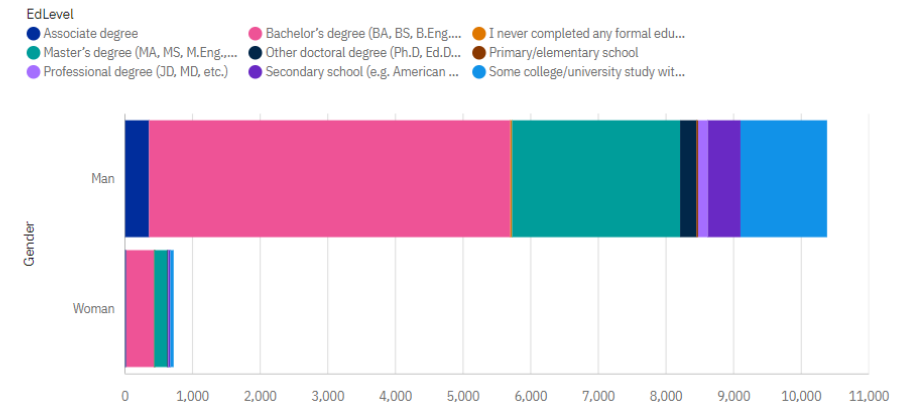
Respondent Counts by Countries



Respondent Count by Age



Respondent Count by Gender Classified by Education Level



DISCUSSION



OVERALL FINDINGS & IMPLICATIONS

- The number of male developers are 15 times greater than the number of female developers.
- On average, male developers earn more than females. But their median are very close.
- 90% of developers have a college degree or at a higher education level.



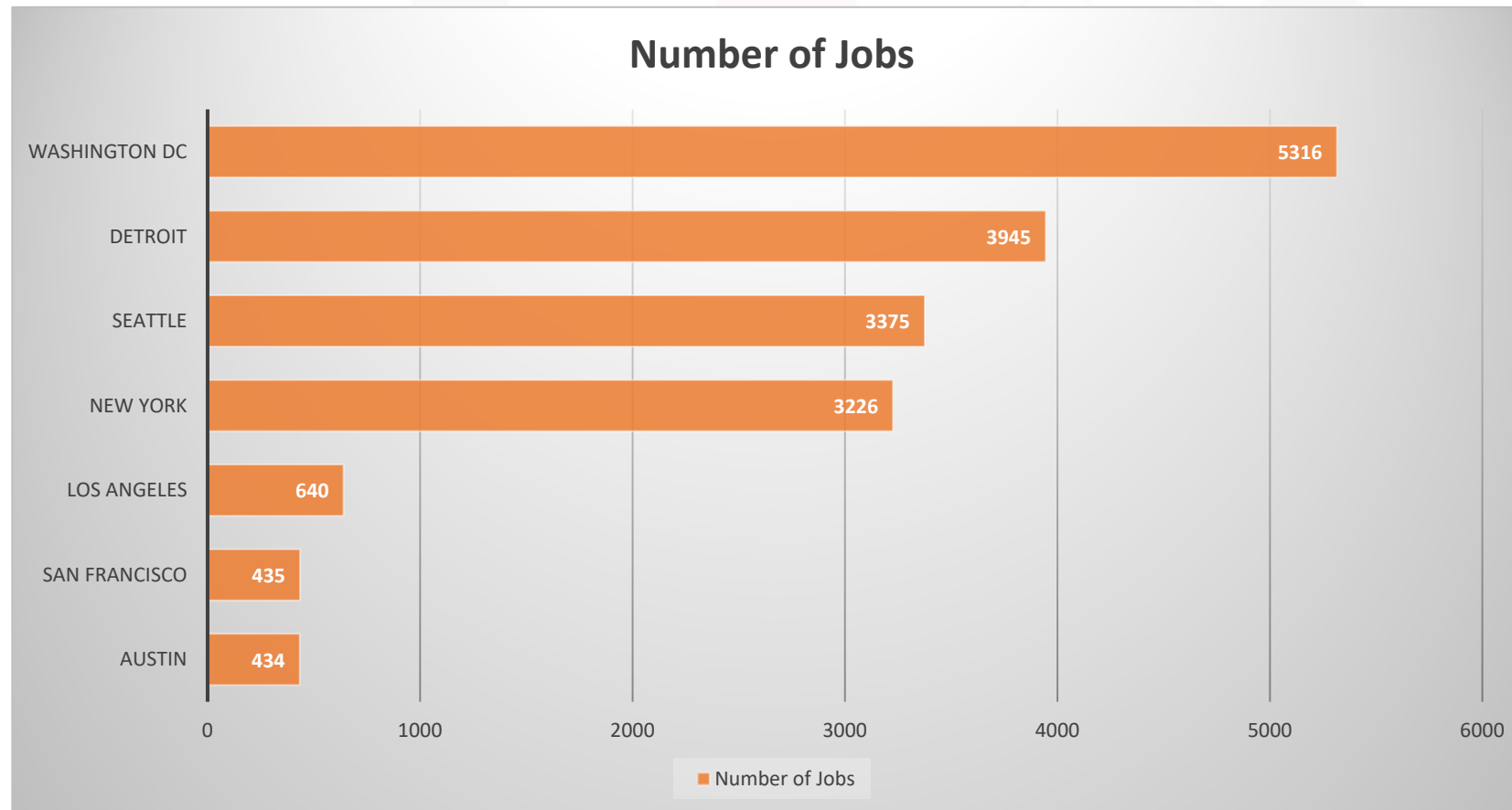
- There is a gender gap and gender pay gap in tech industry. But this is because of the outliers in the compensations of male developers.
- Getting a degree is a standard between developers.

CONCLUSION



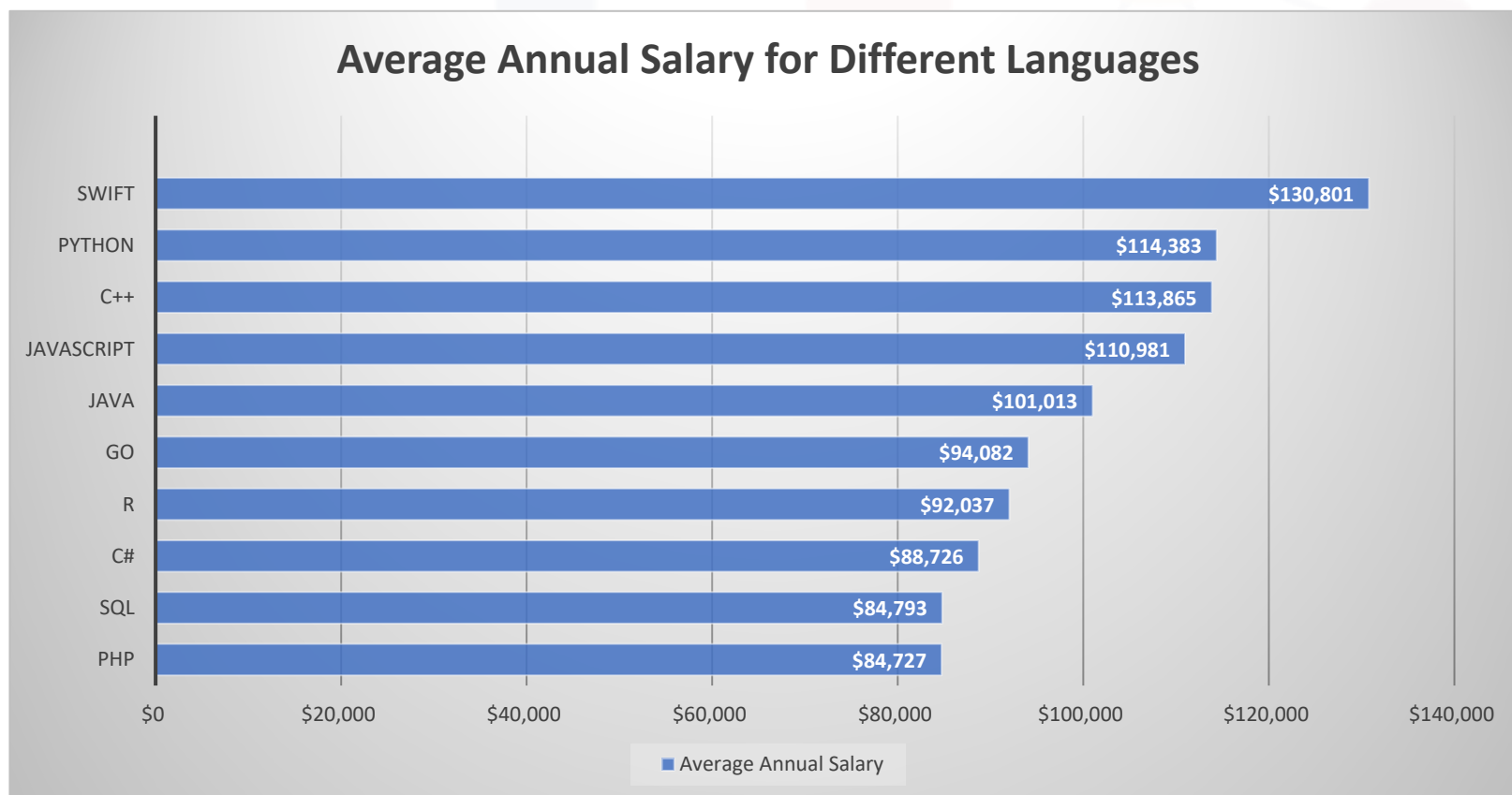
- Imbalanced compensation distribution
- Young developers are more than middle-aged ones.
- Gender pay gap on average
- JavaScript, HTML/CSS, and SQL are most popular languages. Python's popularity is rapidly increasing.

JOB POSTINGS



- The fewest jobs are in Austin, while the most jobs are in Washington DC and Detroit.

POPULAR LANGUAGES



- Swift and Python are the languages with the highest pay, while PHP and SQL are with the lowest.

THANK YOU

