



# Software Professionals Survey Analysis

Buka Mulwila

21-March-2024

# OUTLINE

---



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

# EXECUTIVE SUMMARY

---



- An online survey was conducted of software professionals across the world.
  - Mean age of participants was 30.77 years
  - Survey participants came from 135 unique countries
  - Most active software professionals are in the age range of 21-40 years of age
- Statistical analysis of the collected data concluded that:
  - Javascript is the top in-demand programming language
  - MySQL is the most used, with Postgres the top in-demand database skill
  - JQuery is the most used IDE, with React set to be the top in-demand Webframe
- The software industry is dominated by men who have at least a Bachelor's degree or Master's degree

# INTRODUCTION

---



- To keep pace with changing technologies and remain competitive, regular analysis of data is done to help identify future skill requirements.
- Data from various sources was collected about software professionals across the world. The aim was to identify trends on emerging skills.
- Data was analysed, insights and trends were identified which focussed on:
  - top programming languages in demand
  - top database skills in demand
  - popular IDEs
- Statistical techniques were applied to analyse the data. All the information was then brought together via dashboards.

# METHODOLOGY

---



The methodology used to analyze the data involved:

- Data collection - scraping the internet and using web APIs
- Data wrangling - finding duplicates, removing duplicates, finding missing values, inputting missing values and normalizing data
- Data exploration
  - Analyzing data distribution
  - Handling outliers
  - Finding correlation
- Visualization
  - Distribution of data with **histograms** and **box plots**
  - Relationships between data points with **scatter plots** and **bubble plots**
  - Composition of data with **pie charts** and **stacked charts**
  - Comparison of data with **line charts** and **bar charts**
- Bringing it all together with dashboards

# RESULTS

---

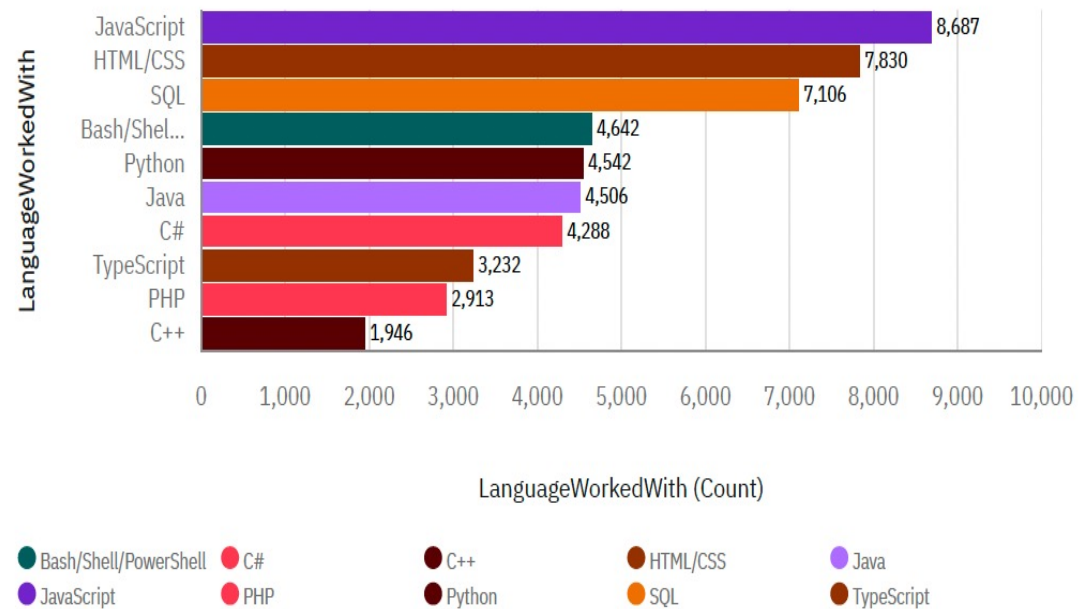
Insights and trends were achieved as follows:

- The dataset was available as a csv file
- Exploratory data analysis was performed by reading from csv using pandas dataframes, matplotlib and seaborn
- Data visualization was performed by extracting data using SQL from data in the form of a RDBMS. Plots were constructed using pandas, numpy and matplotlib
- Dashboards were created using the IBM Cognos Analytics platform.

# PROGRAMMING LANGUAGE TRENDS

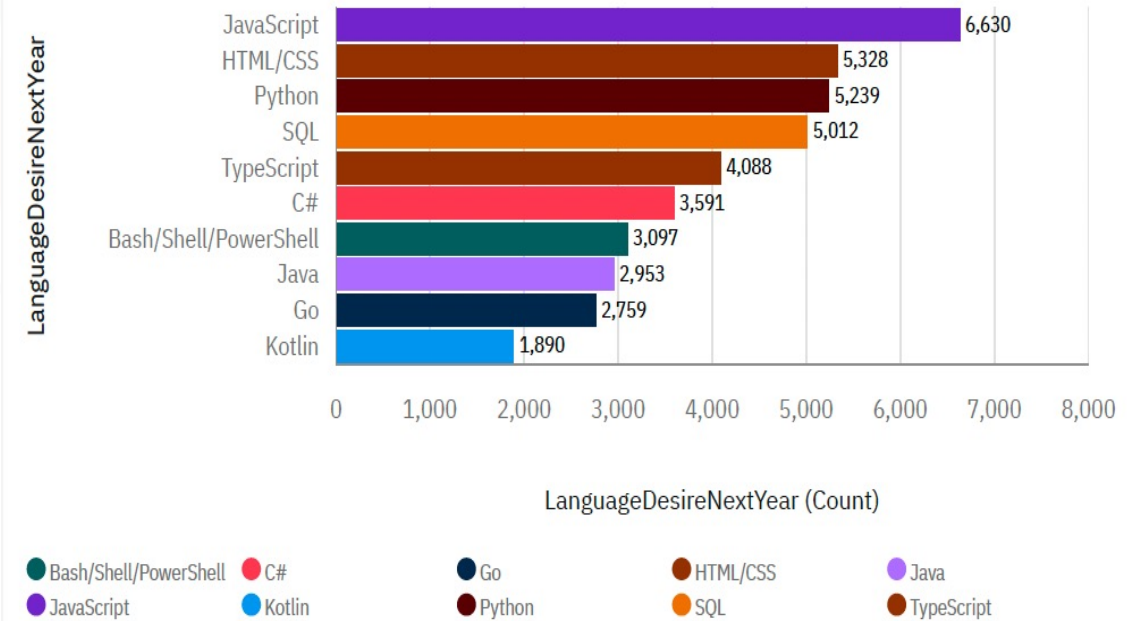
## Current Year

Top 10 Languages Worked With



## Next Year

Top 10 Language Desire Next Year





# PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

---

## Findings

- Kotlin and Go are not in top 10 languages currently used, but desired for next year
- C++ and PHP dropped out of top 10 indicating their declining popularity
- Javascript will remain the most desired language next year with Python being the fastest growing

## Implications

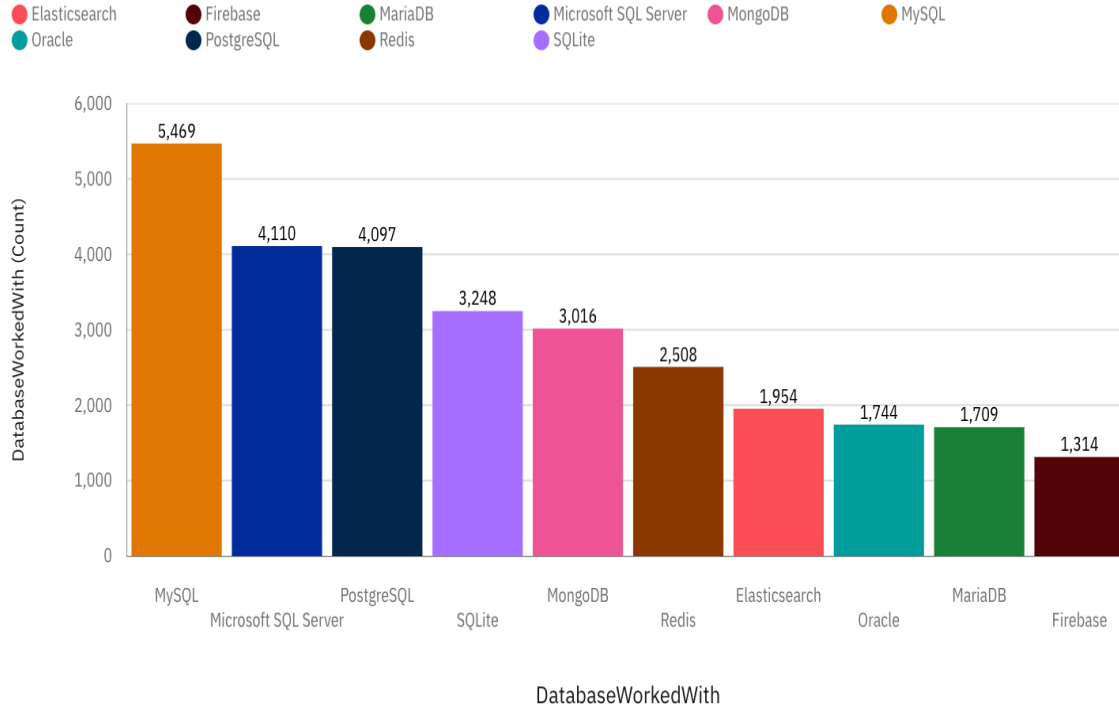
- There will be a requirement to upskill in Kotlin and Go programming languages
- A reduction in investment in C++ and PHP is expected
- Despite 24% decline in usage next year, Javascript will remain popular and python uptake will increase indicating a need for continued investment



# DATABASE TRENDS

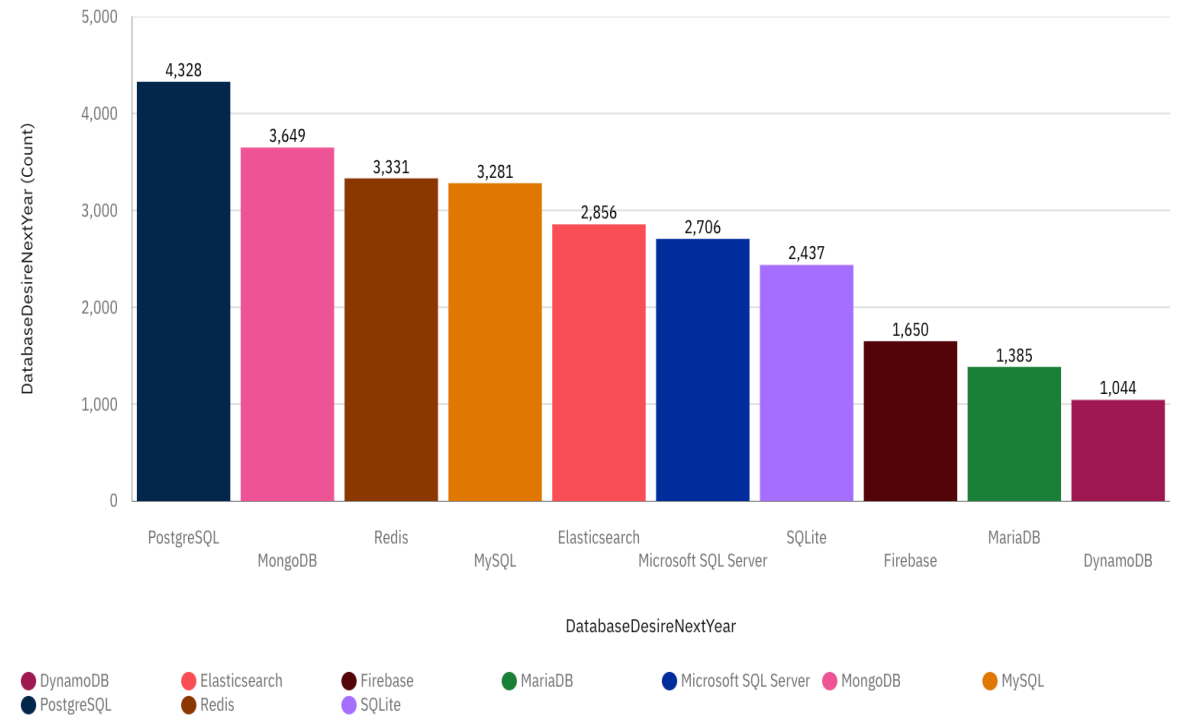
## Current Year

Top 10 Database Worked With



## Next Year

Top 10 Database Desire Next Year



# DATABASE TRENDS - FINDINGS & IMPLICATIONS

---

## Findings

- Oracle is expected to lose popularity and drop out of the top 10 next year
- Dynamo DB is expected to increase in popularity next year
- PostgreSQL will replace MySQL in popularity next year

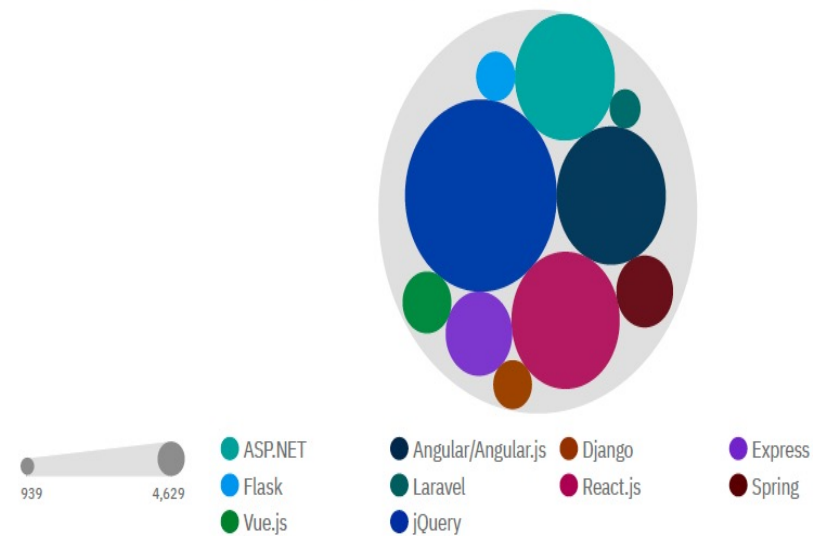
## Implications

- A reduction in investment in the oracle database is expected next year
- There will be a requirement to upskill in Dynamo DB
- There will be a requirement to upskill in PostgreSQL

# WEBFRAME TRENDS

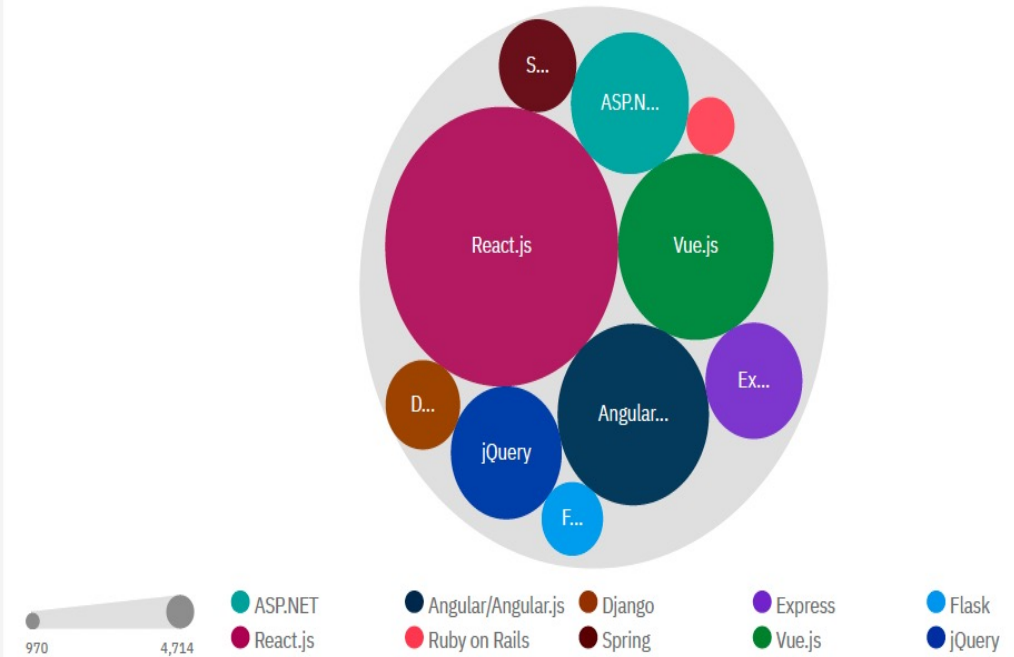
## Current Year

Top 10 WebFrame Worked With



## Next Year

Top 10 Web Frame Desire Next Year



# WEBFRAME TRENDS - FINDINGS & IMPLICATIONS

---

## Findings

- JQuery is most used, but is expected to be replaced by React next year as most popular
- Vue, Angular and ASP.Net also set to be most desired next year, behind React
- Ruby is set to replace Laravel in the top 10 next year

## Implications

- There will be a requirement to upskill in JQuery IDE in preparation for next year
- There will be a requirement to upskill in Vue, Angular and ASP.Net IDEs in preparation for next year
- A reduction in investment in the Laravel is expected next year

# DASHBOARD

---

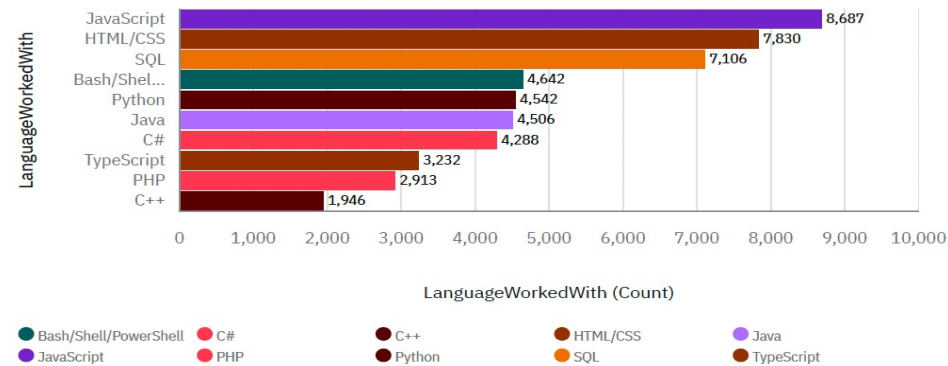


<https://github.com/mulwib/dataAnalysisCapstone/blob/main/Survey%20Data%20Dashboard%20-%20Capstone.pdf>

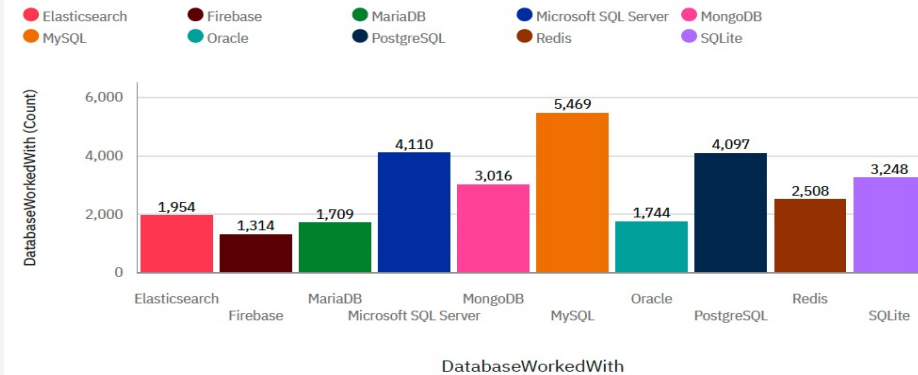
# DASHBOARD TAB 1

## A - Current Technology Usage

Top 10 Languages Worked With



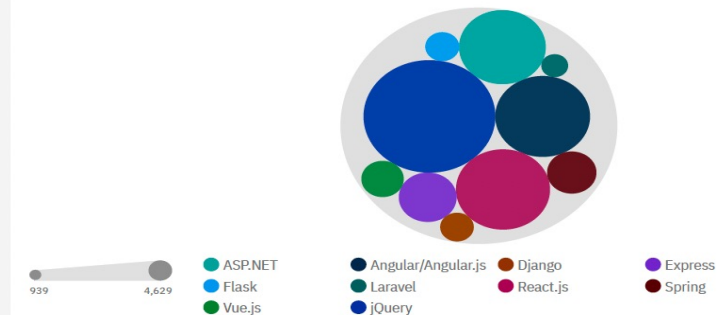
Top 10 Database Worked With



Platform Worked With



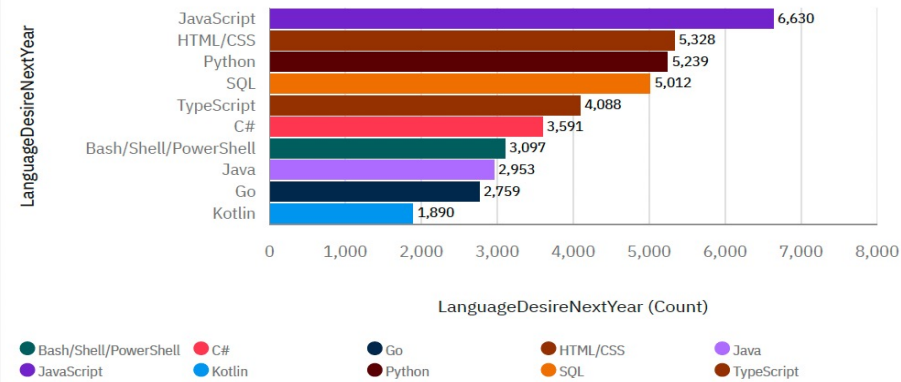
Top 10 WebFrame Worked With



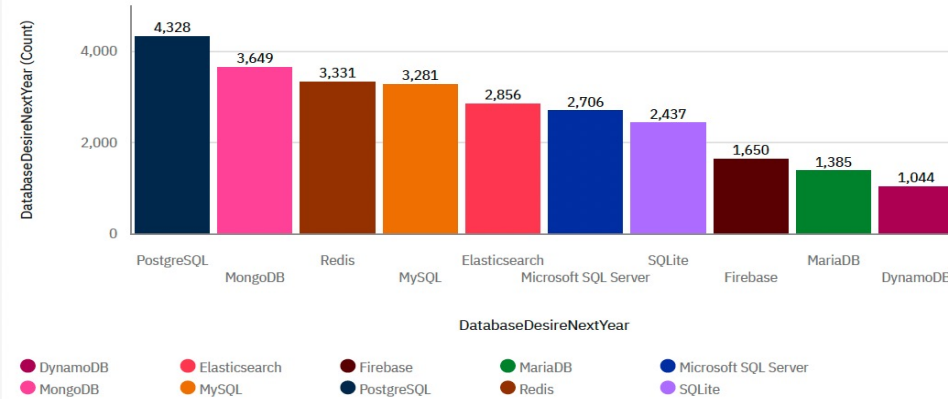
# DASHBOARD TAB 2

## B - Future Technology Trend

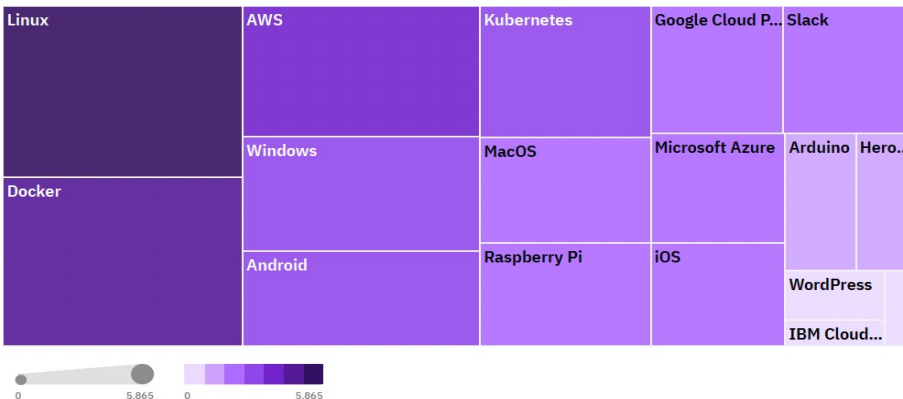
Top 10 Language Desire Next Year



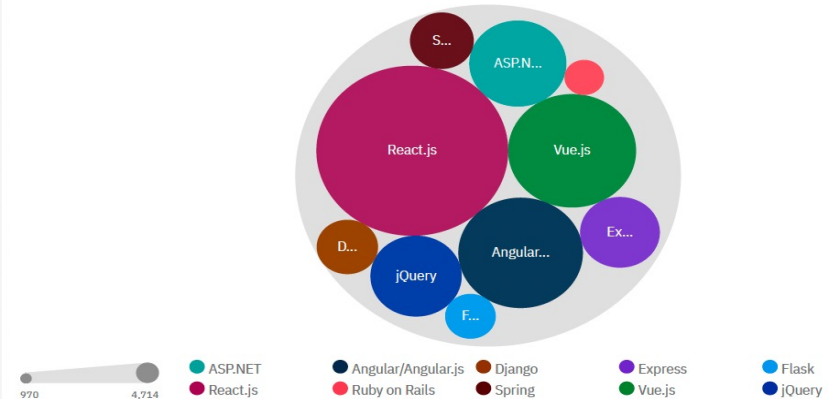
Top 10 Database Desire Next Year



Platform Desire Next Year



Top 10 Web Frame Desire Next Year

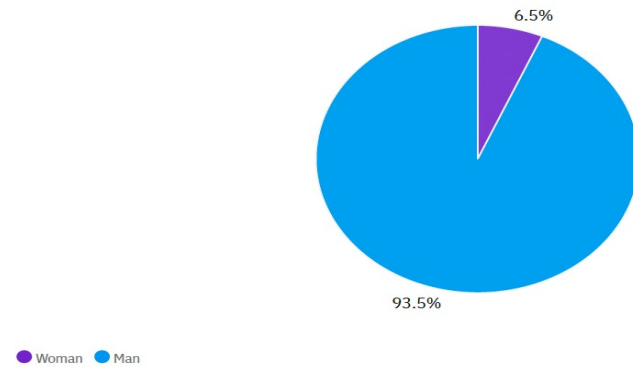




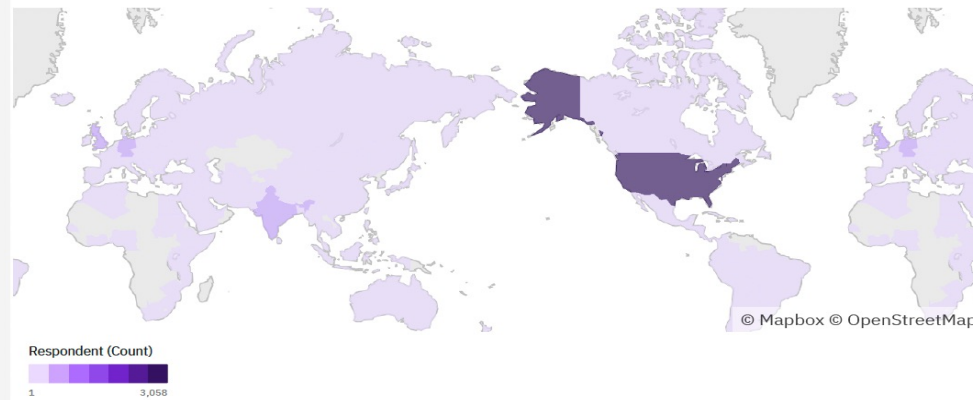
# DASHBOARD TAB 3

## C - Demographics

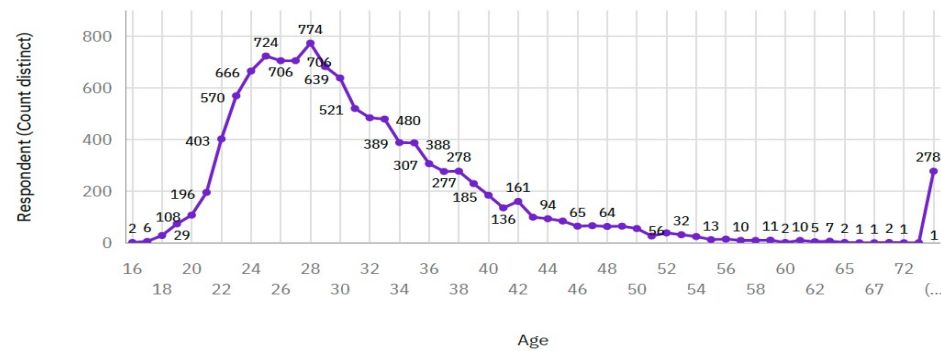
Respondent classified by Gender



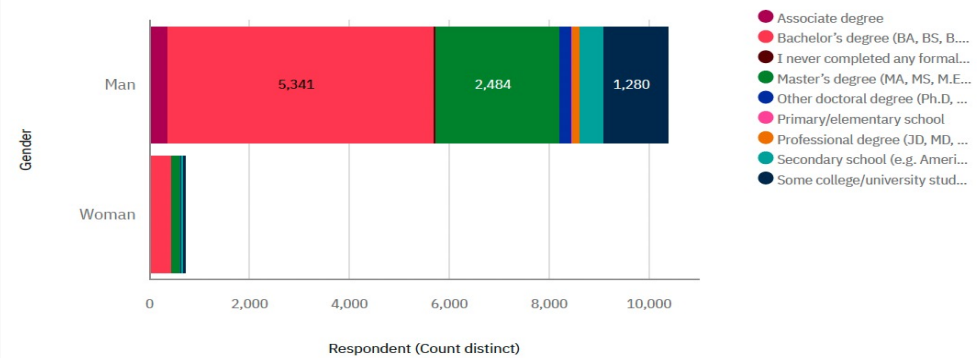
Respondent Count for Countries



Respondent Count by Age



Respondent Count by Gender by Formal Education Level



# DISCUSSION

---



- Majority of survey respondents were male (93.5%)
- Most of the respondents were between 21 and 40 years old
- The USA had the highest number of respondents
- The highest number of respondents had a bachelors degree

# OVERALL FINDINGS & IMPLICATIONS

---

## Findings

- 80 percent of programming languages, 90 percent of databases and 90 percent of Webframes are expected to remain popular next year
- 93.5 percent of the practitioners are male, whilst 6.5 percent are female
- Majority of practitioners are educated to Bachelor's degree or Masters degree level

## Implications

- C++ and PHP languages, Oracle database and Laravel IDE are expected to have reduced investment next year
- Participation levels indicate that there is still work to do to address gender parity
- Participation levels indicate that there is still work to do to address social economic parity

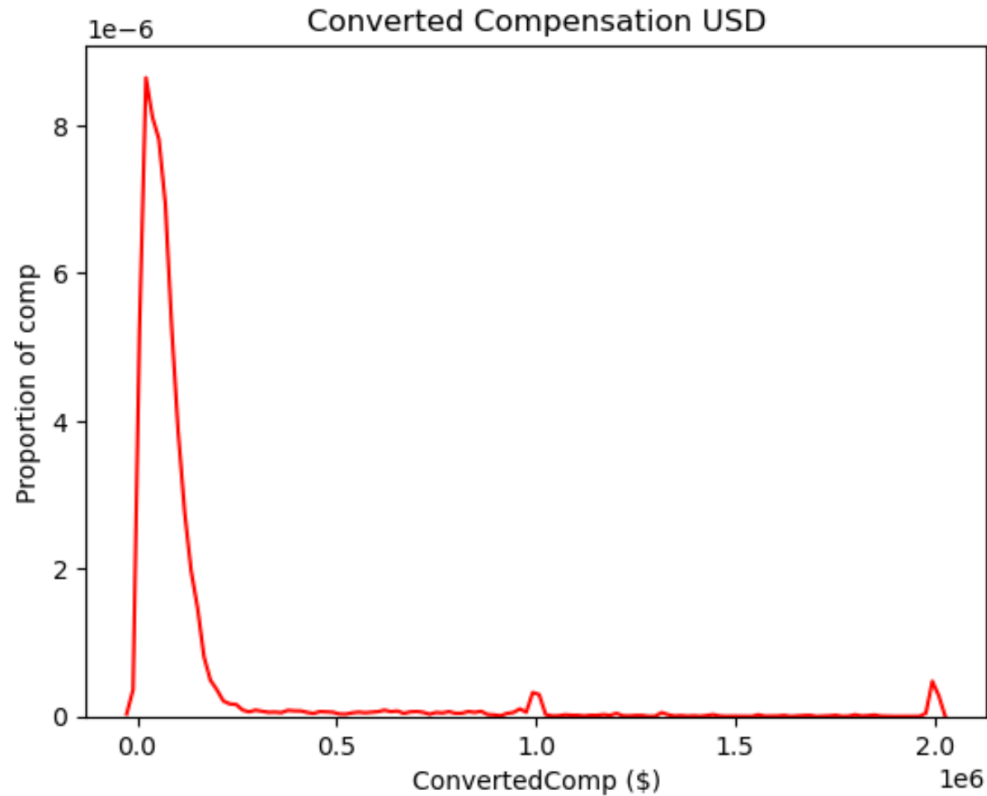
# CONCLUSION

---

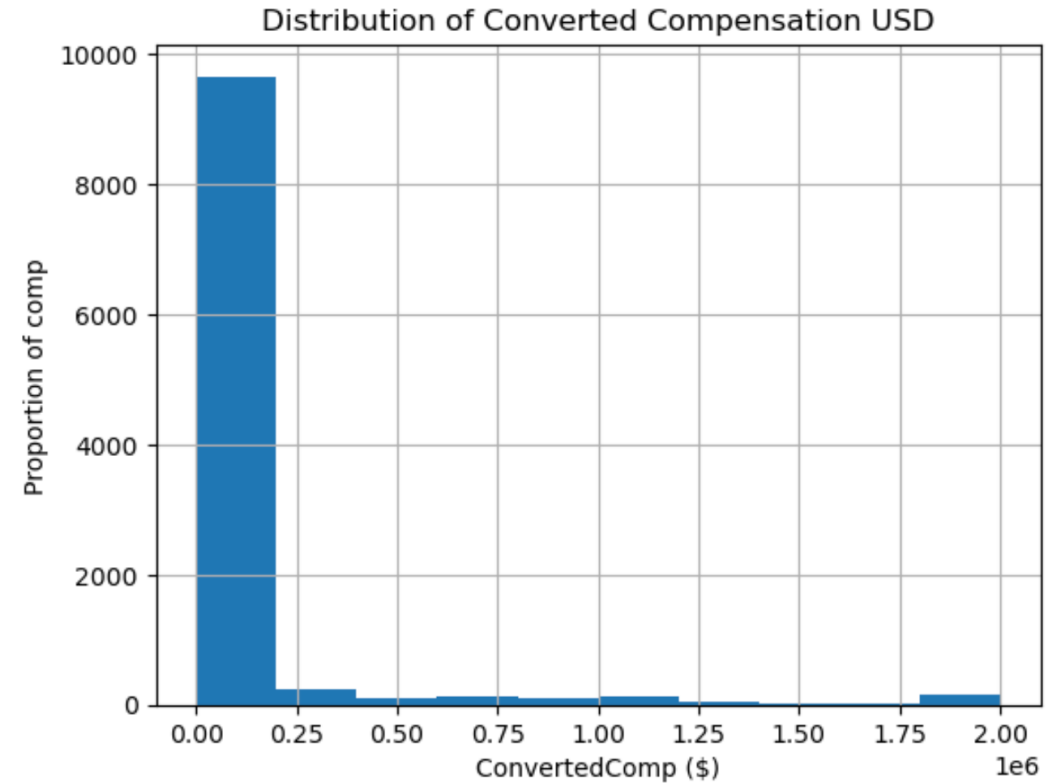


- Analysis of the collected data has led to the conclusion that:
  - Javascript is the top in-demand programming language
  - MySQL is the most used, with Postgres the top in-demand database skill
  - JQuery is the most used IDE, with React set to be the top in-demand Webframe
- The software industry is dominated by men who have at least Bachelor's degree or Master's degree
- Most active software professionals are in the age range of 21-40 years of age

# APPENDIX (1)



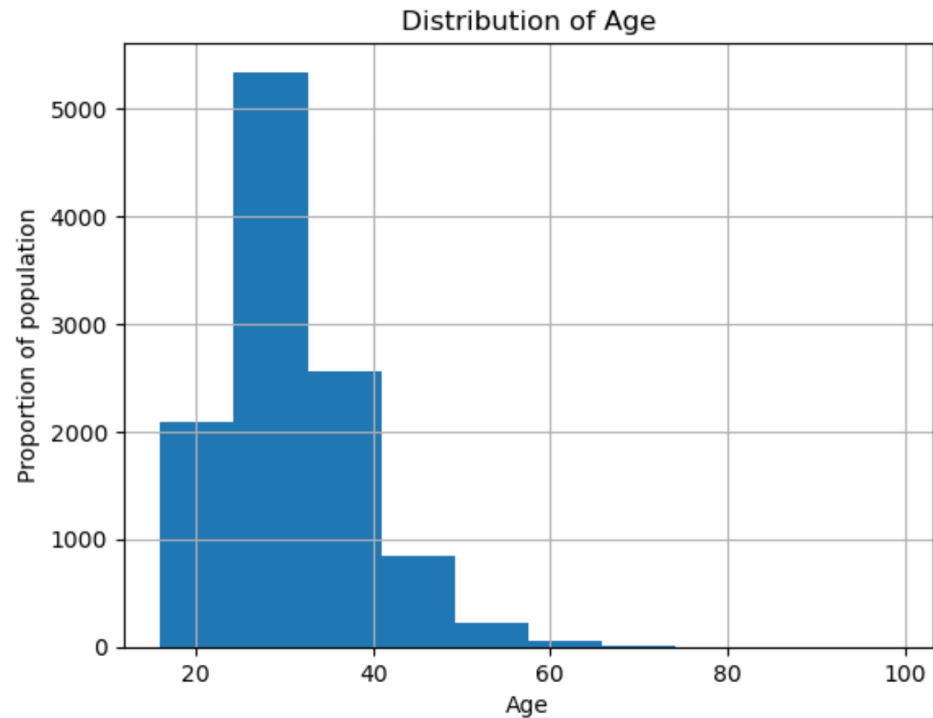
Distribution curve for the column ConvertedComp



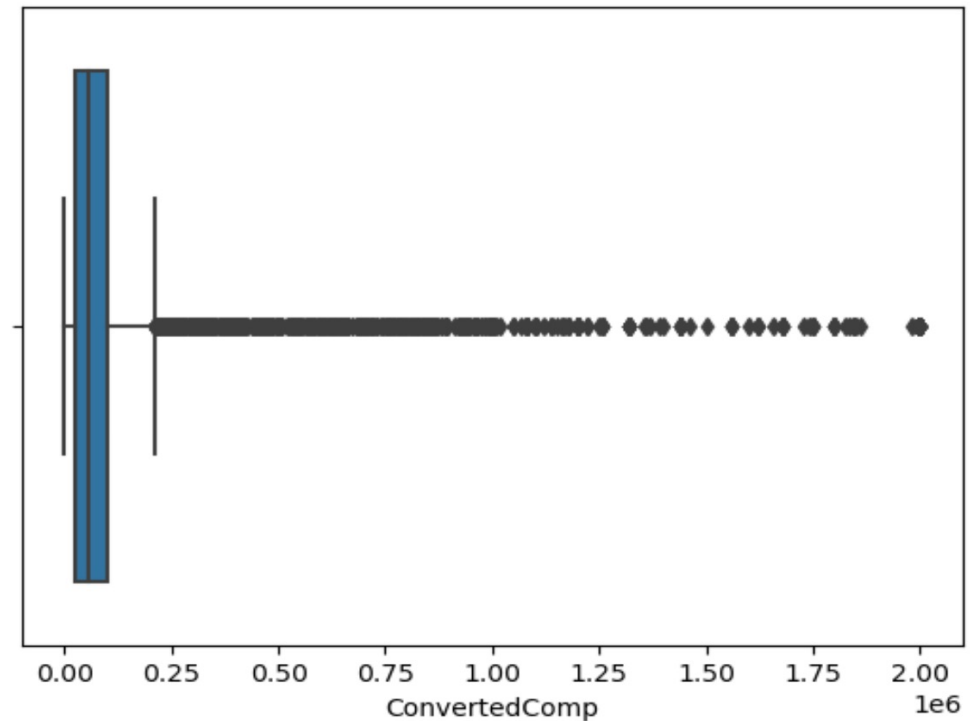
Histogram for the column ConvertedComp

# APPENDIX (2)

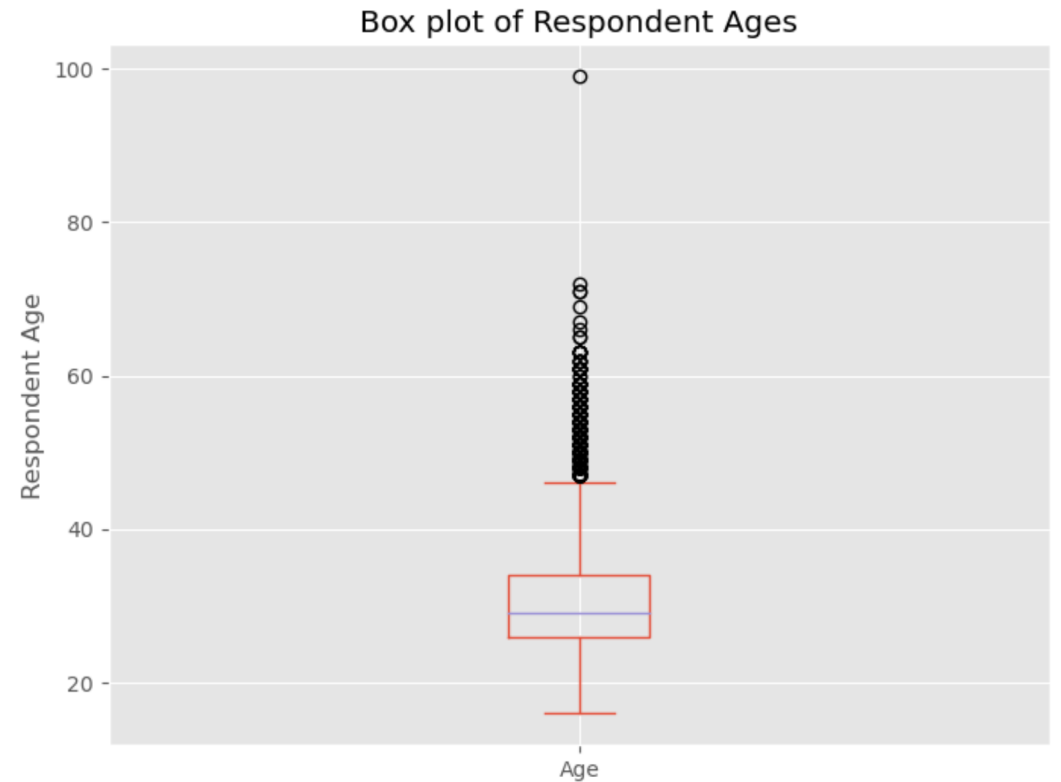
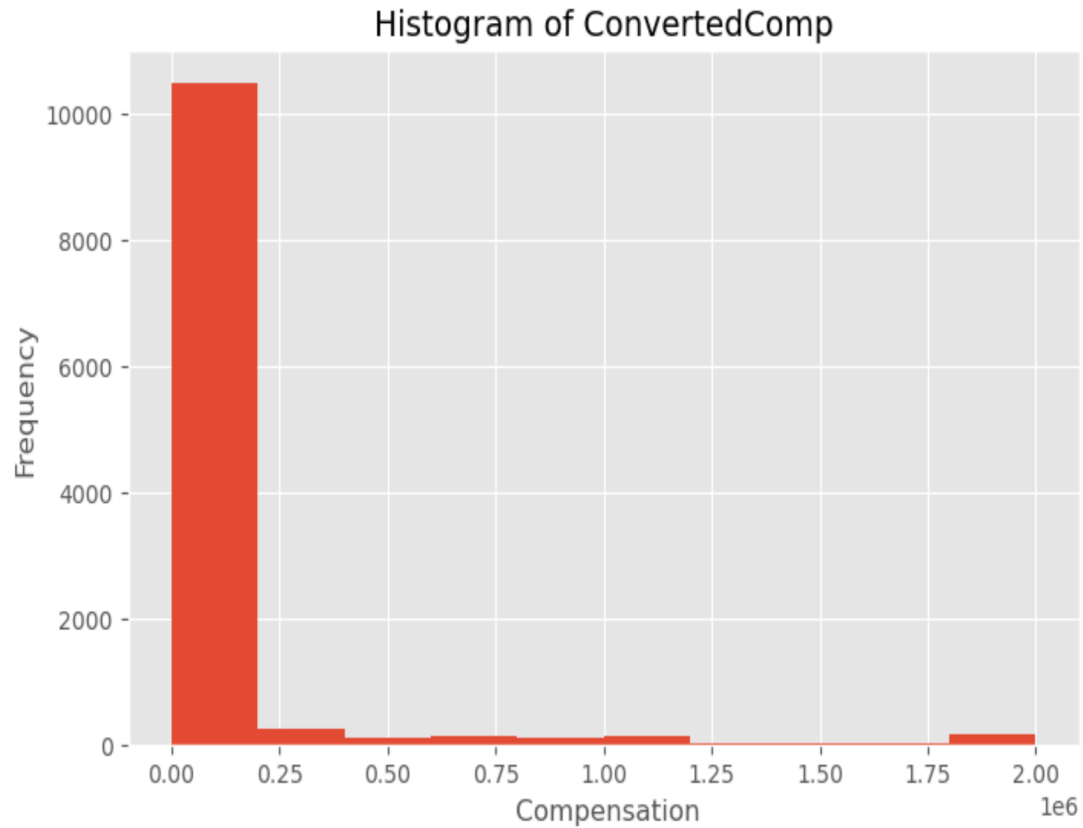
Distribution for Age



Boxplot for ConvertedComp

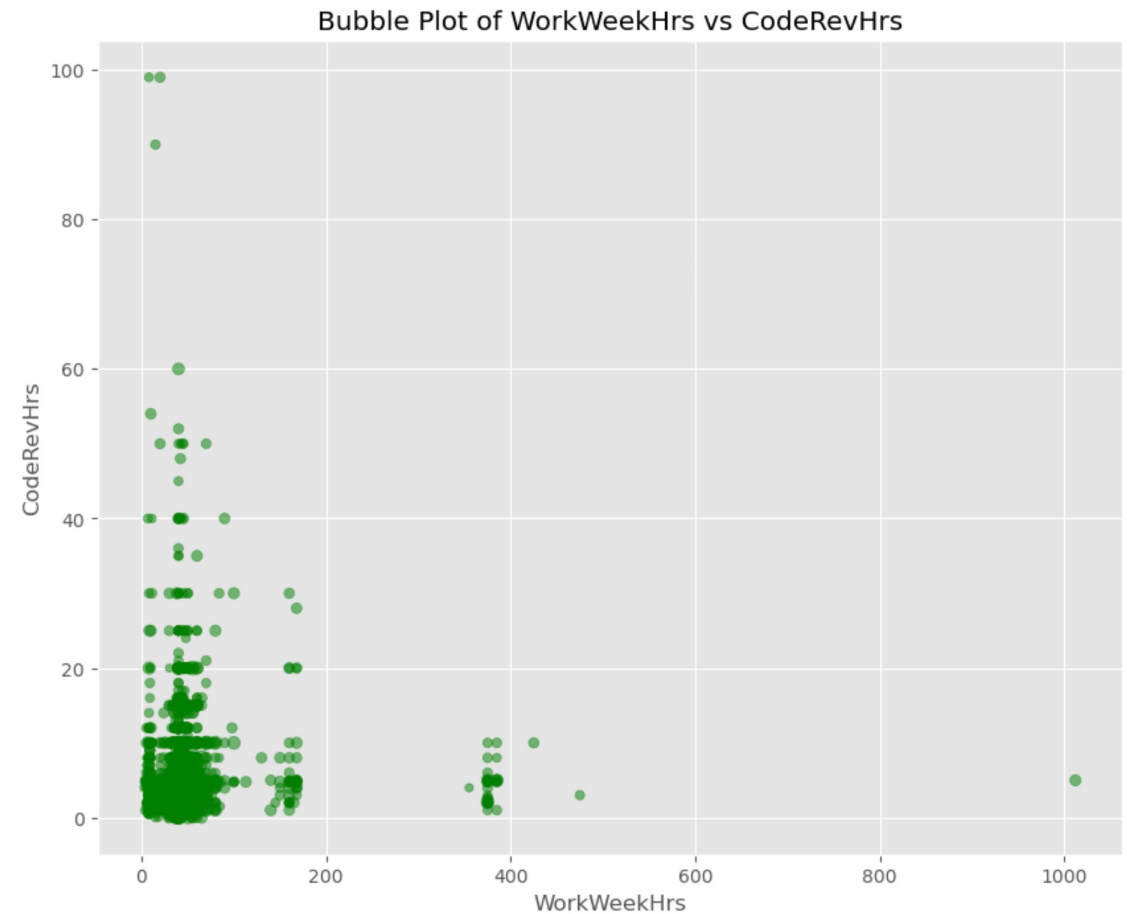
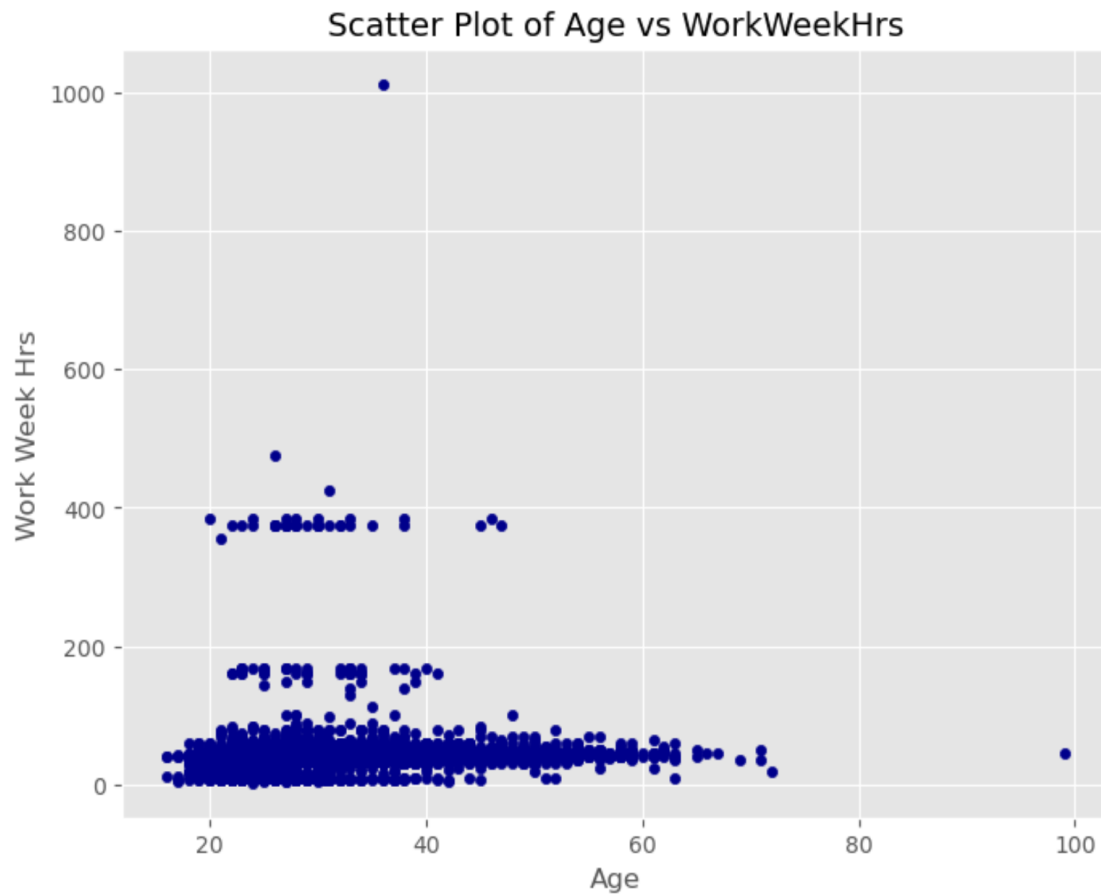


# APPENDIX (3)

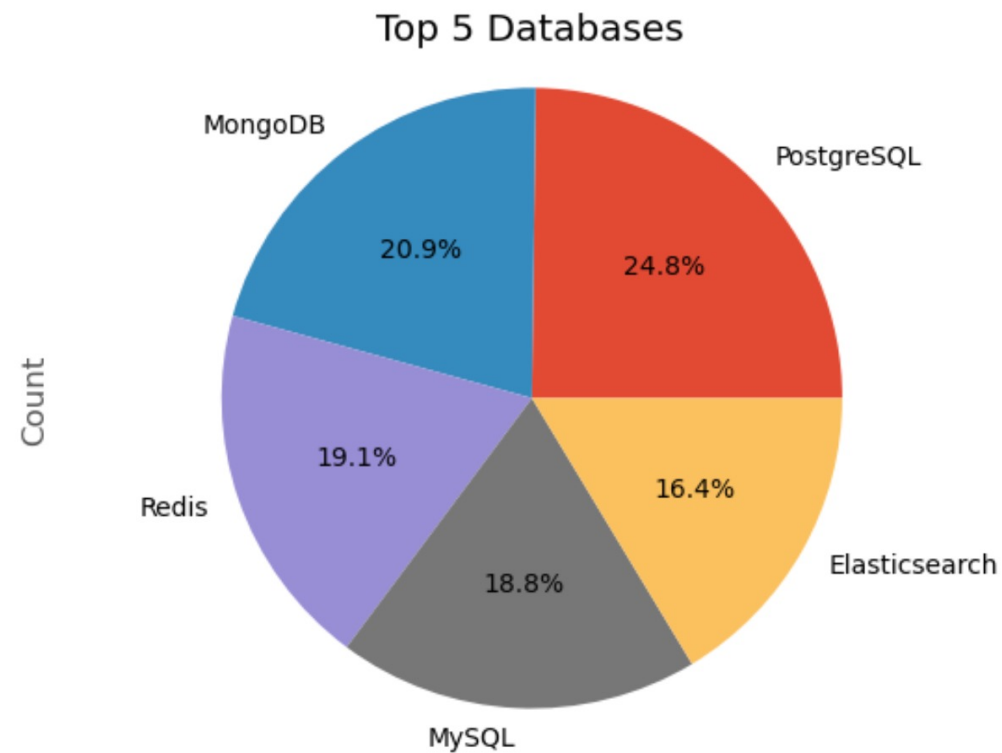




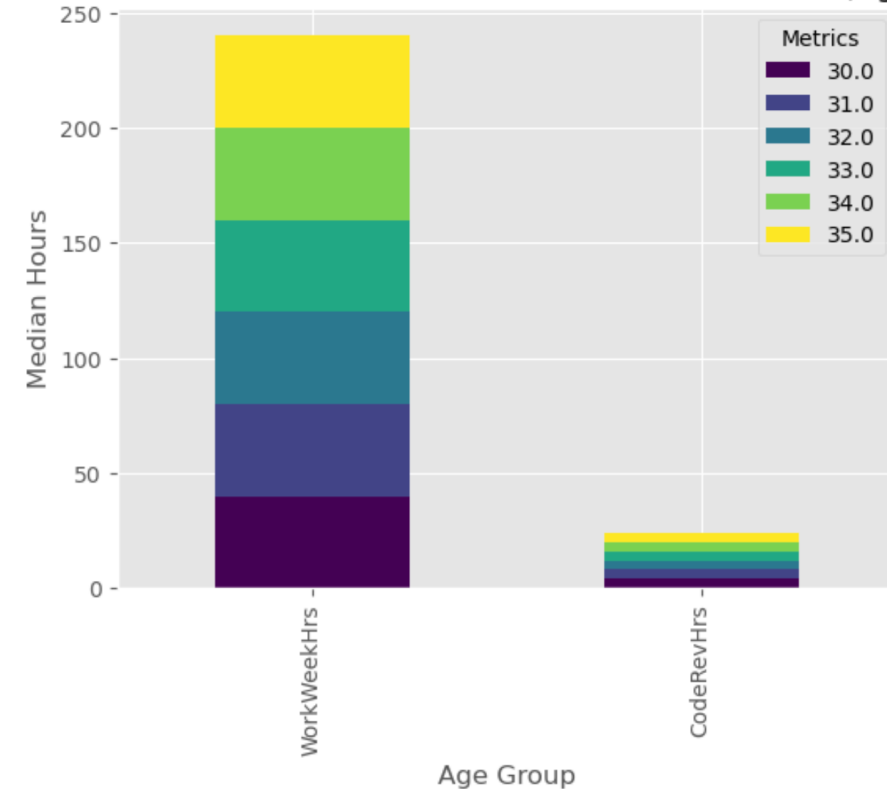
# APPENDIX (4)



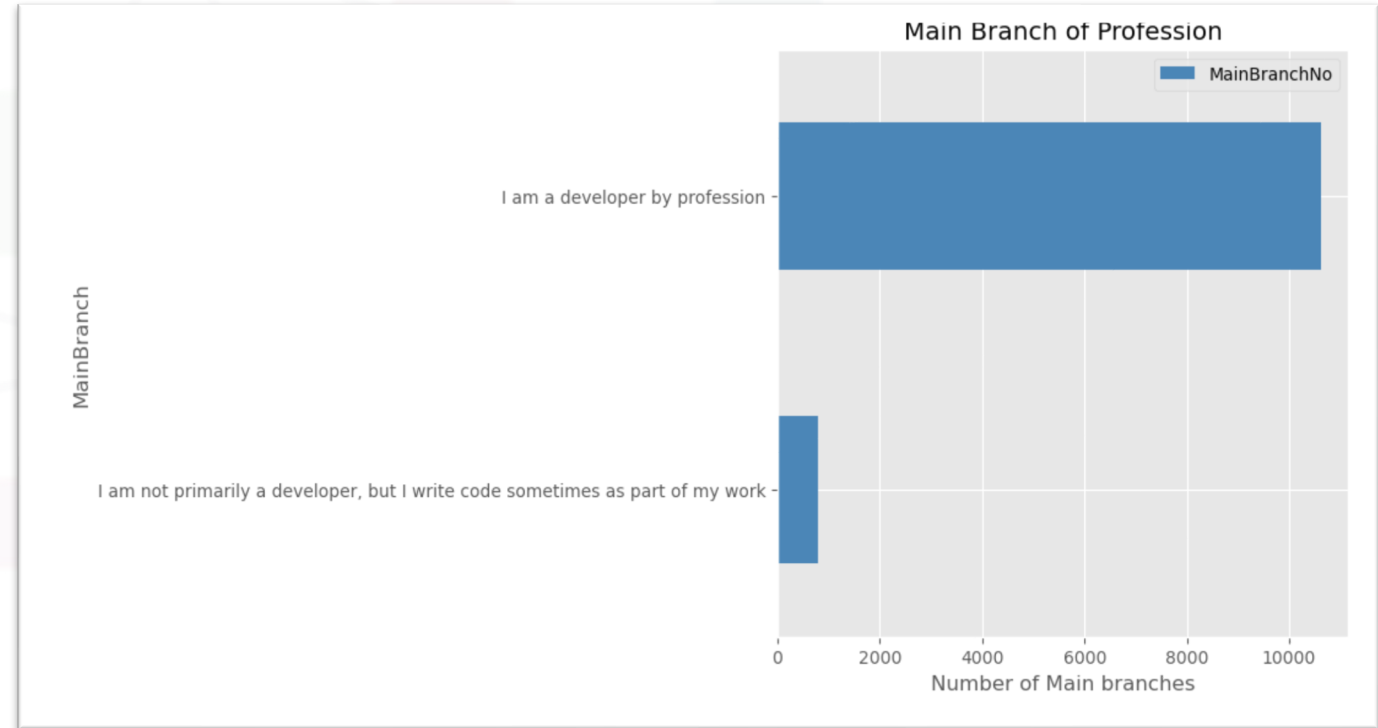
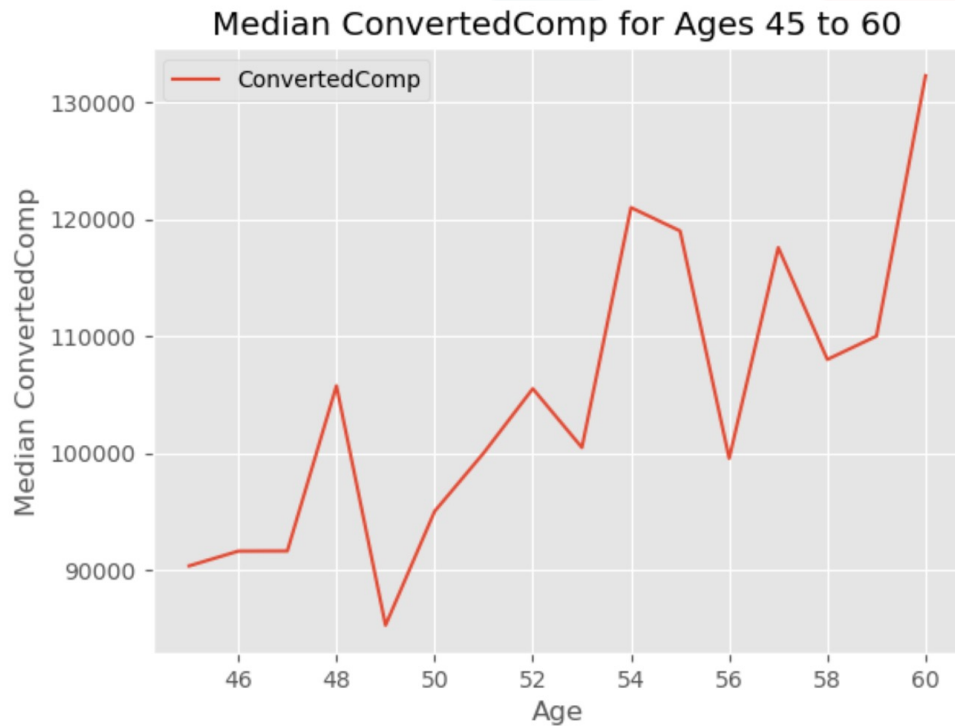
# APPENDIX (5)



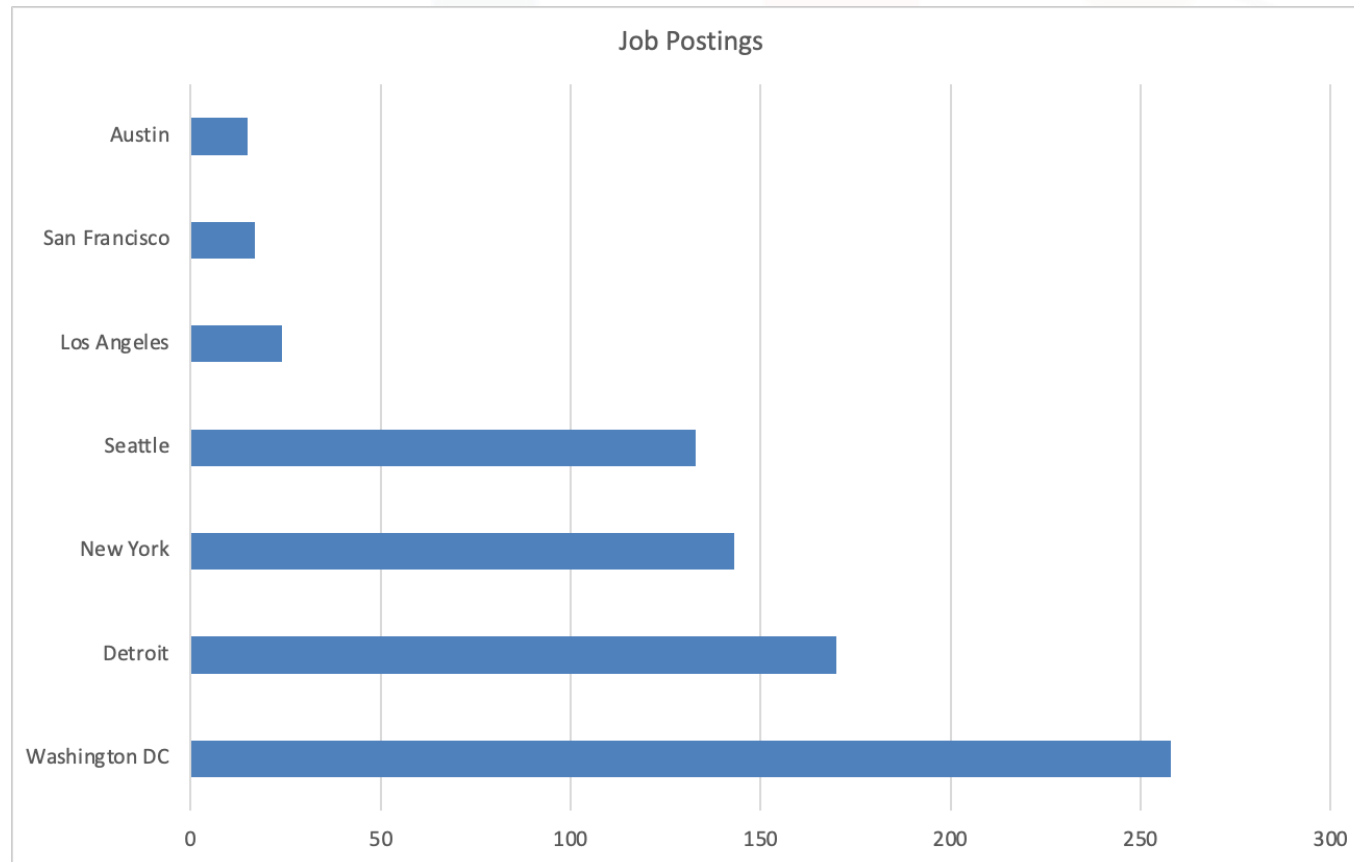
Stacked Chart of Median WorkWeekHrs and CodeRevHrs (Age 30 to 35)



# APPENDIX (6)



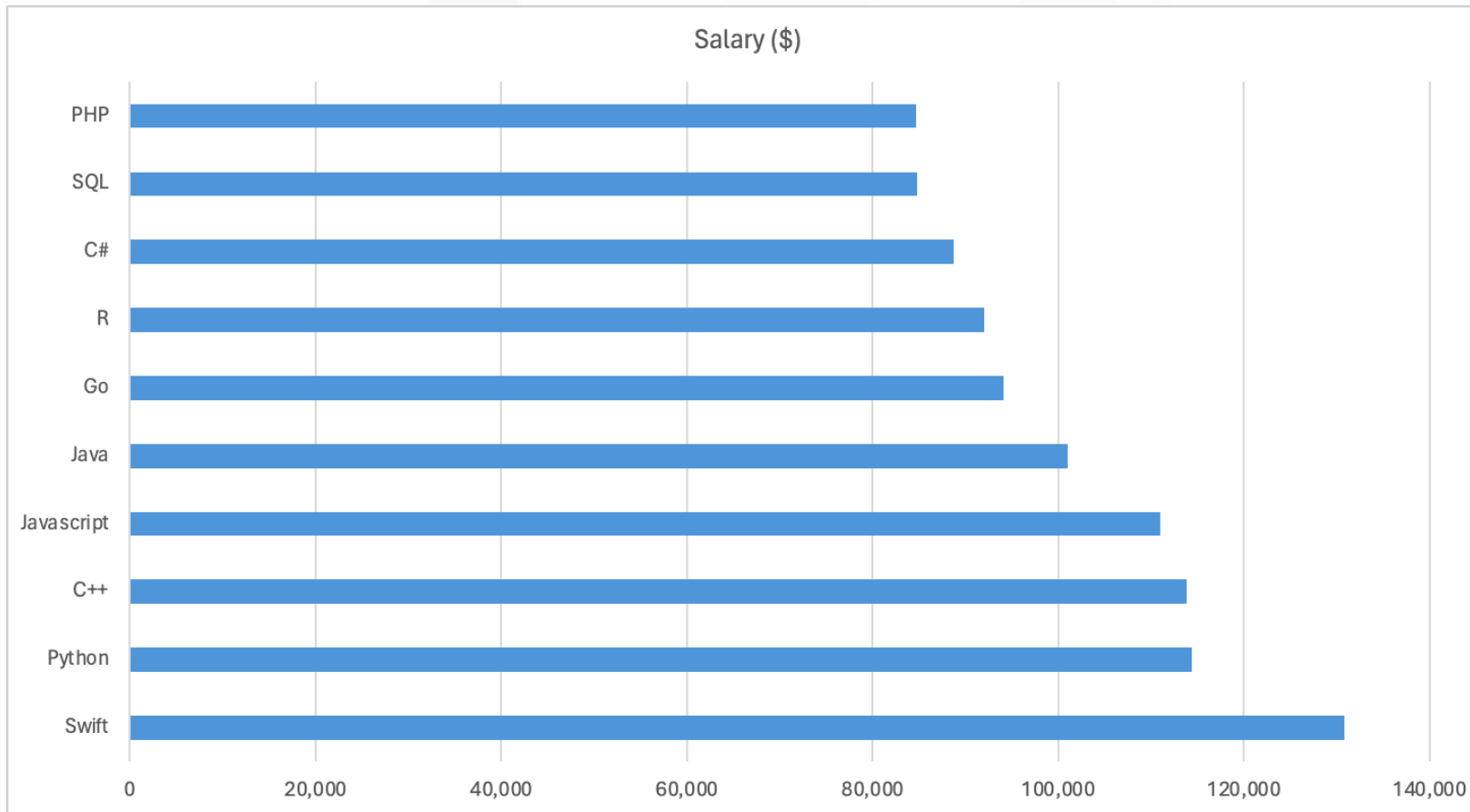
# JOB POSTINGS



Survey was done on the number of jobs currently open for various technologies across Los Angeles, New York, San Francisco, Washington DC, Seattle, Austin and Detroit.

Data was collected using the API for the **number of job postings** for the above locations

# POPULAR LANGUAGES



Data was extracted from a given web site.

Scraped data included the name of the **programming language** and **average annual salary**.