



Stack Overflow Developer Survey

Chrysovalantou Smiliani
March 20,2023

Outline



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

Executive Summary



- Stack Overflow is an online platform, that helps people, find the answers they need, when they need them. It is best known, for their public Q&A platform, that over 100 million people visit every month to ask questions, learn, and share technical knowledge.
- Through the annual Stack Overflow Developer Survey dataset, it has been used a modified subset of that dataset for IBM Data Analyst Capstone Project.
- This report is based on a Survey of 11,552 Software Developers, from 135 unique Countries around the world.
- This Survey is focused, to the following programming languages:
 - C
 - C#
 - C++
 - Java
 - JavaScript
 - Python
 - Scala
 - Oracle
 - SQL Server
 - MySQL Server
 - PostgreSQL
 - MongoDB
- The mean age of the Survey participants is 30,78.
- There are 7 unique Gender values in the Survey: Man, Woman, Non-Binary-Genderqueer or Gender-non-conforming, Man- Non-Binary-Genderqueer or Gender-non-conforming, Woman- Non-Binary-Genderqueer or Gender-non-conforming, Woman-Man, Woman-Man- Non-Binary-Genderqueer or Gender-non-conforming.

Introduction



This report has been created for stakeholders, from a global IT and business consulting services firm, that is known for its expertise, in IT solutions and its team of highly experienced IT consultants and in order to keep pace with changing technologies and remain competitive, this firm, regularly analyses data, to help identify future skill requirements.

Within this initiative, it has been tasked, to collect data from various sources and identify trends for this year's report on emerging skills.

The top programming skills that are most in demand and also future desired were collected from various sources including:

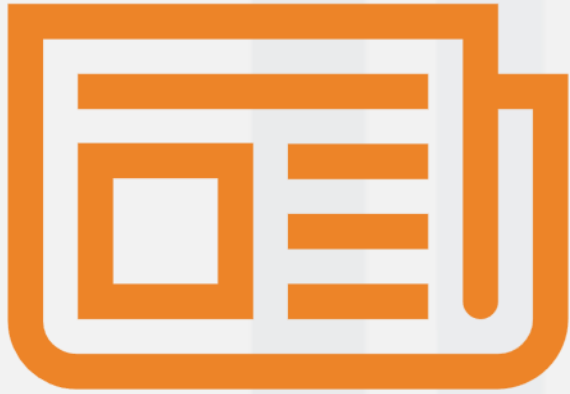
- Job postings
- Training portals
- Surveys

The data and identified insights and trends were analysed and include the following:

- What are the top programming languages in demand and also future desired?
- What are the top database skills in demand and future desired?
- What are the popular IDEs?

With IBM Cognos Dashboard Embedded have been created some Dashboards, for a better understanding.

Methodology



- Data Collection: Collecting Data Using APIs, Collecting Data Using Webscraping, Exploring Data.
 - The dataset comes from the following source:
<https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/> under a [ODbL: Open Database License](#).
- For this project have been used:
 - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv
 - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/labs/datasets/Programming_Languages.html
 - [m5_survey_data_demographics.csv](#)
 - [m5_survey_data_technologies_normalised.csv](#)
 - File: job-postings.xlsx
 - File: popular-languages.csv
- Data Wrangling: Finding Missing Values, Determine Missing Values, Finding Duplicates, Removing Duplicates, Normalizing Data.
- Exploratory Data Analysis: Distribution, Outliers, Correlation.
- Data Visualization: Visualizing Distribution of Data, Relationship, Composition, Comparison.
- Dashboard Creation: Dashboards with Cognos Dashboard Embedded(IBM).

Results

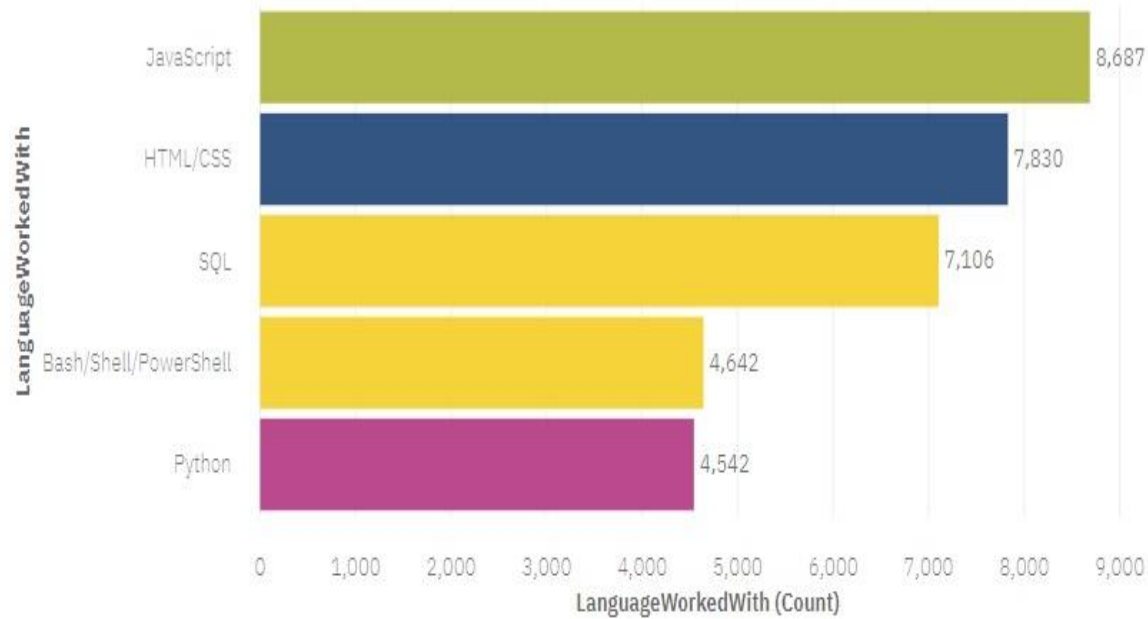


Programming Languages Trends

Top 5 Programming Languages for the Current Year

LanguageWorkedWith

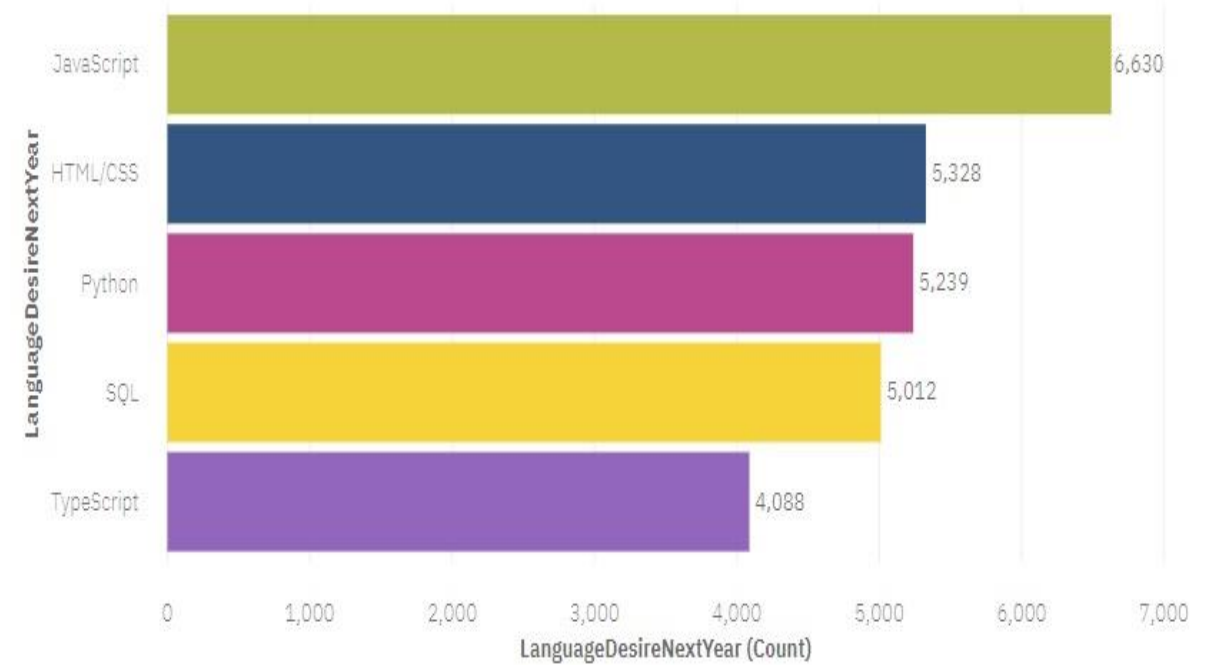
● Bash/Shell/PowerShell ● HTML/CSS ● JavaScript ● Python ● SQL



Top 5 Programming Languages for the Next Year

LanguageDesireNextYear

● HTML/CSS ● JavaScript ● Python ● SQL ● TypeScript



Programming Languages Trends - Findings and Implications

Findings

- JavaScript and HTML/CSS are the 2 most popular programming languages for the current but also for next year.
- The 3rd most popular programming language for next year is Python and it is observed a decrease to SQL, which follows in the 4th place.
- Bash/Shell/PowerShell does not seem to be between the 5 trends for next year and it seems to be replaced by TypeScript.

Implications

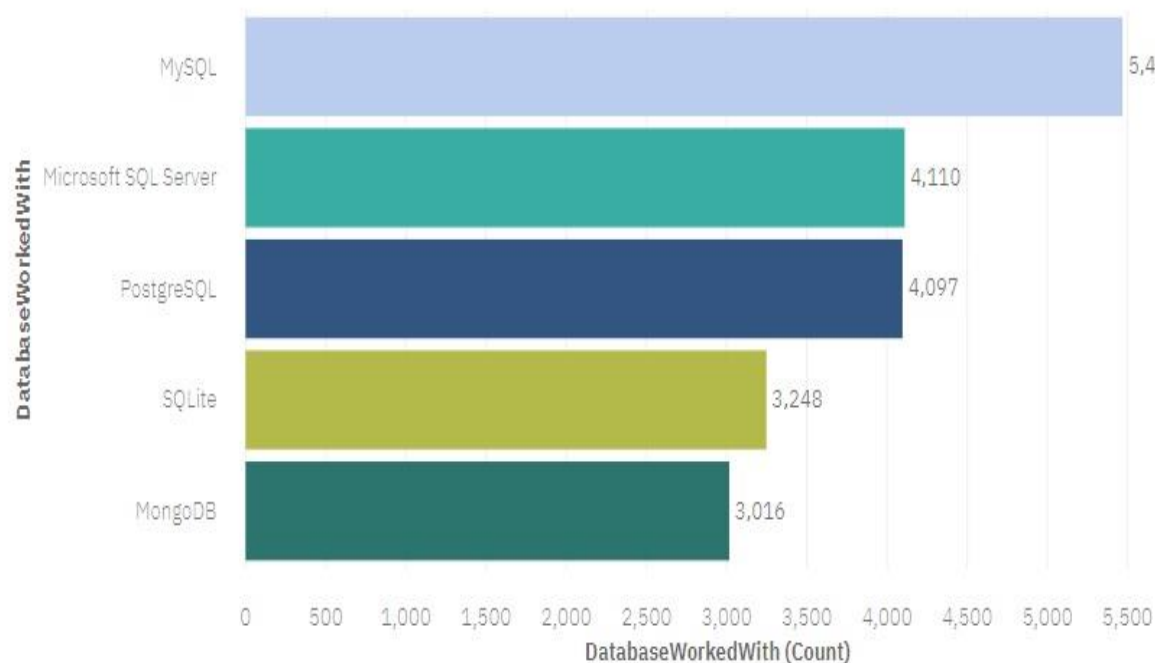
- It is suggested, continuous seminars to employees, with knowledge in JavaScript and HTML/CSS, as well as, with Python and TypeScript, in order to keep them up-to-date.
- Upon recruitment, a few weeks of training in JavaScript and HTML/CSS, that are on demand and in Python and TypeScript, that are the future trend, is recommended for new entrants, so that, they can join and follow the company's projects more quickly.
- Distribution of roles and hierarchies, based on experience and knowledge. Consequently offering guidance to juniors.

Database Trends

Top 5 Databases for the Current Year

DatabaseWorkedWith

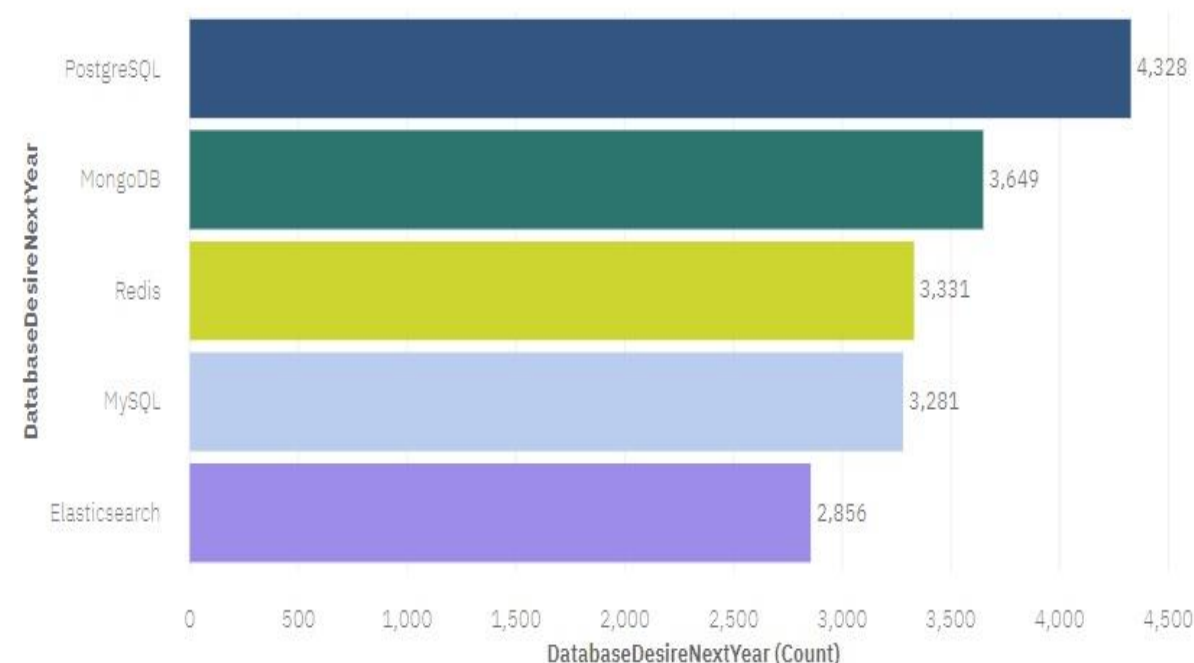
Microsoft SQL Server MongoDB MySQL PostgreSQL SQLite



Top 5 Databases for the Next Year

DatabaseDesireNextYear

Elasticsearch MongoDB MySQL PostgreSQL Redis



Database Trends - Findings and Implications

Findings

- Interest in MySQL, Microsoft SQL Server and SQLite has decreased in desired databases, in relation to the current year.
- PostgreSQL and MongoDB are the 2 most desired databases for next year, in contrast with the current year, where they are not so popular.
- There is an interest for Redis and Elasticsearch, as desired databases for next year.

Implications

- It is suggested, continuous seminars to employees, with the latest databases trends, such as, PostgreSQL, MongoDB, Redis and Elasticsearch, in order to keep them up-to-date.
- Upon recruitment, a few weeks of training in PostgreSQL, MongoDB, Redis and Elasticsearch, that are the future database trends, is recommended for new entrants, so that they can join and follow the company's projects more quickly.
- Redefinition of needs and data. Evaluation, from the business decision makers, for possible use of the databases, that are the future trends, according to the needs of the company.

Dashboard

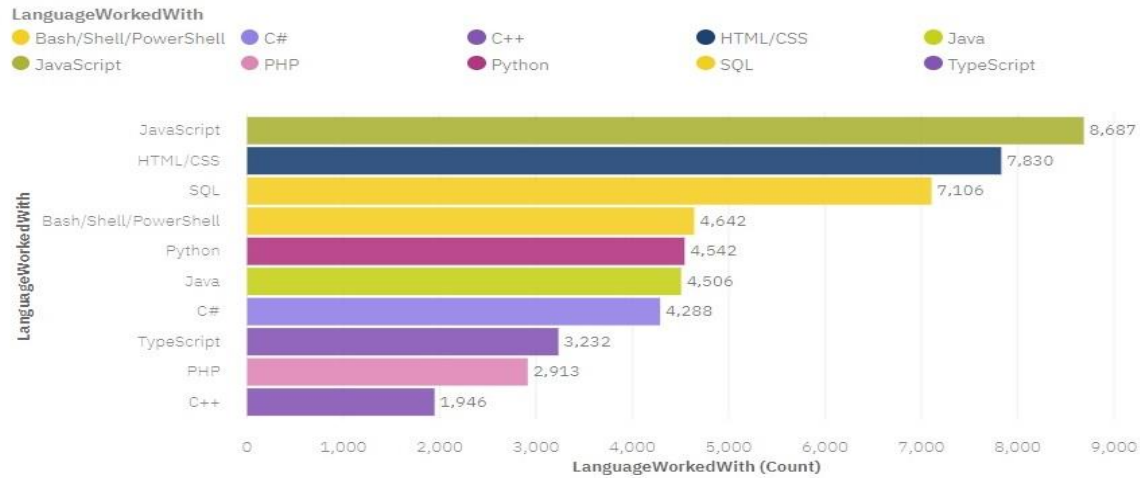


The permanent link of the read-only view of the Cognos dashboard is:

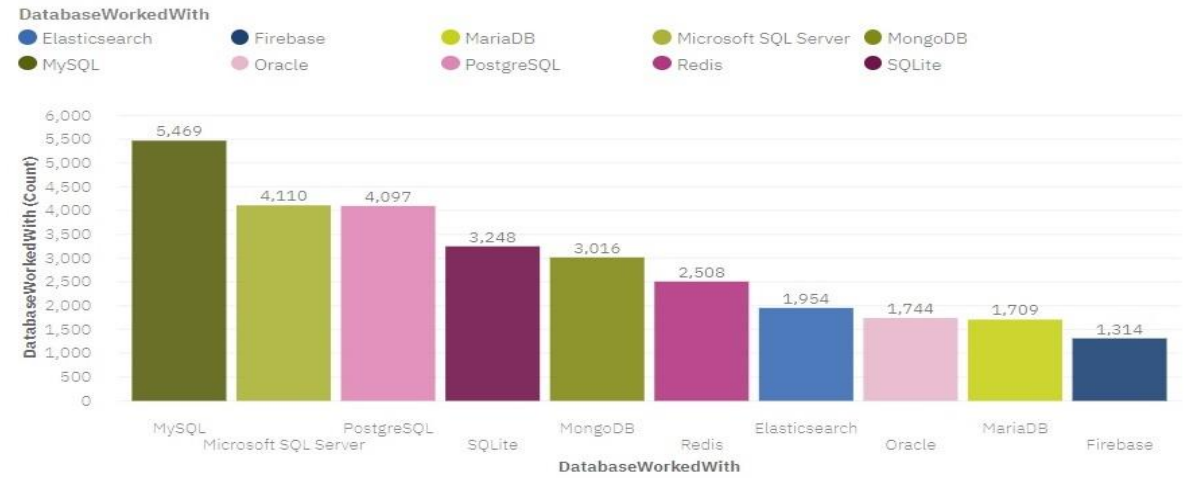
<https://eu-gb.dataplatform.cloud.ibm.com/dashboards/7c7beedf-8620-4402-ba6d-4b7dfa157b77/view/453ef50831a22bcc09abd4e407912b0f7d37715be6bbd00a888d7b490e367697a86812c2c82f185f8e105760f4bf4359ce>

Current Technology Usage

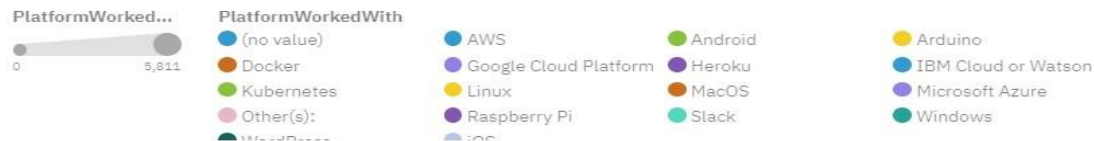
Top 10 Language Worked With



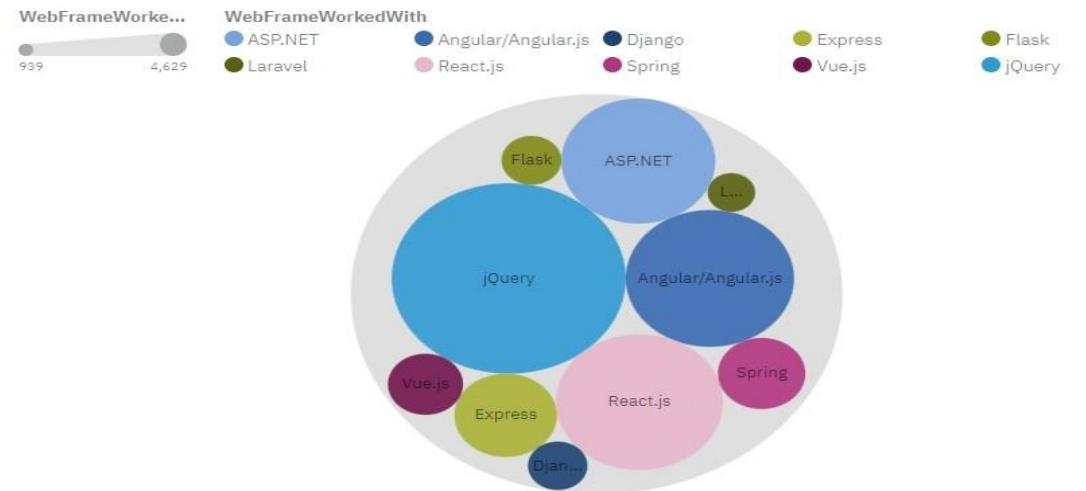
Top 10 Database Worked With



Platform Worked With

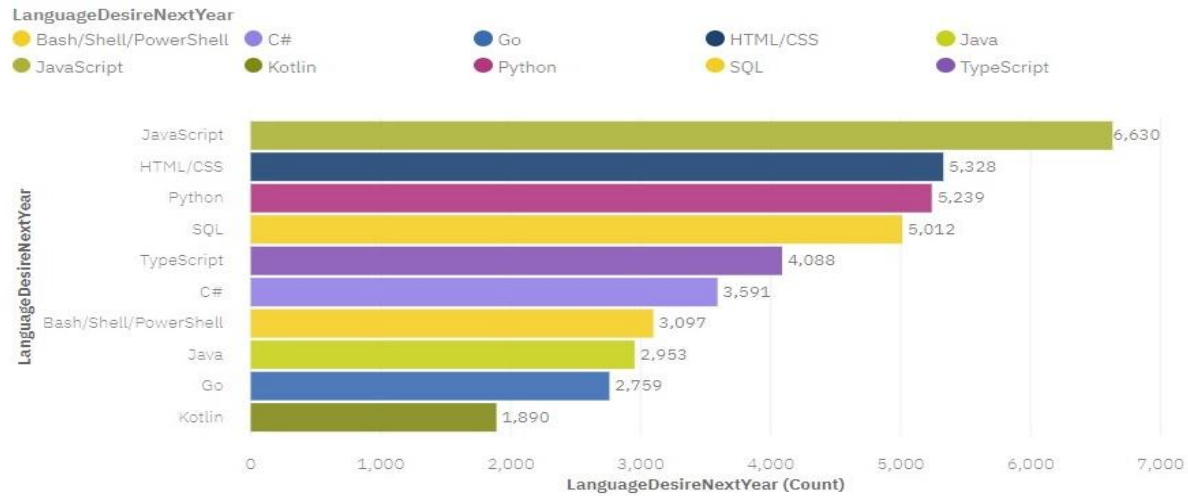


Top 10 Web Frame Worked With

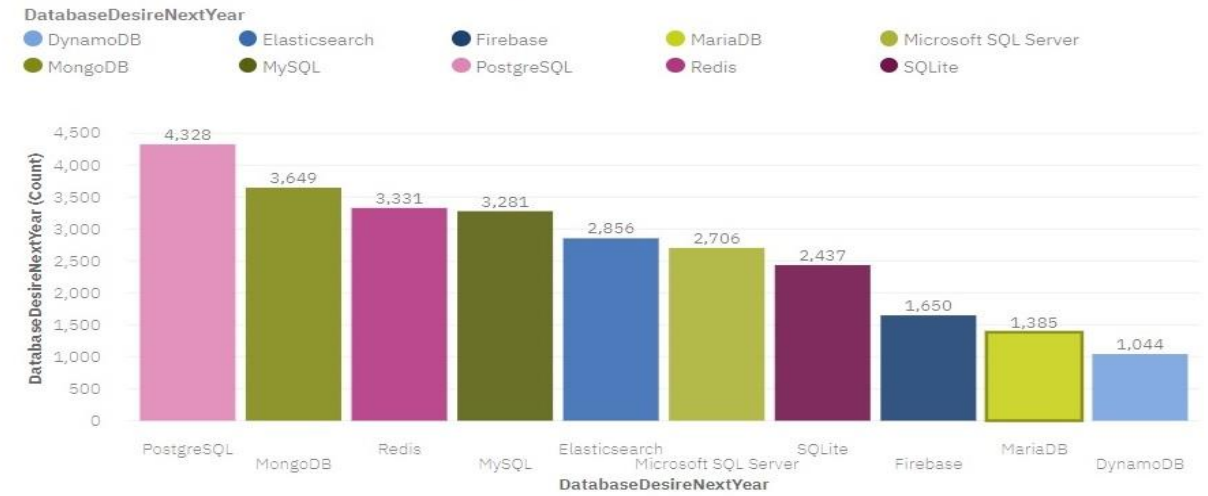


Future Technology Trend

Top 10 Language Desire Next Year



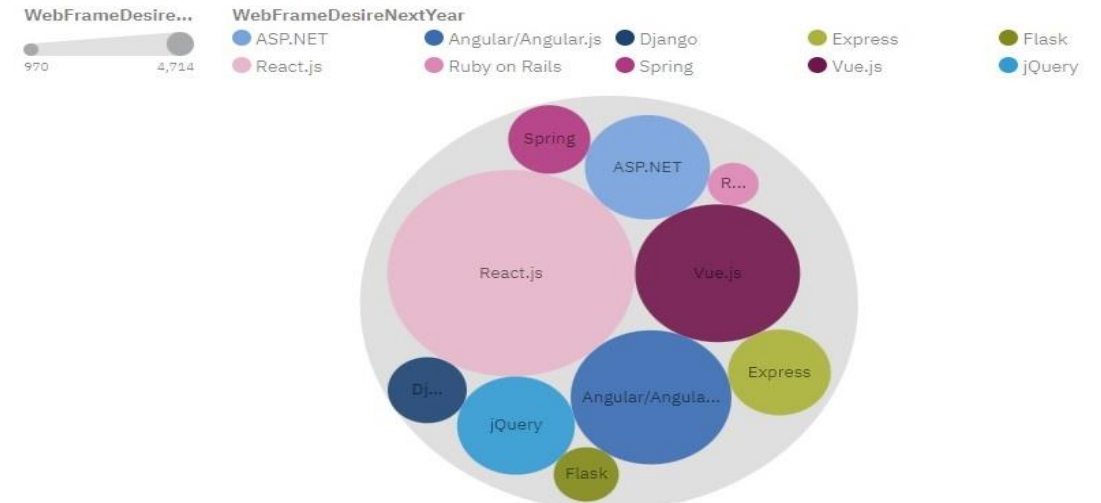
Top 10 Database Desire Next Year



Platform Desire Next Year



Top 10 Web Frame Desire Next Year

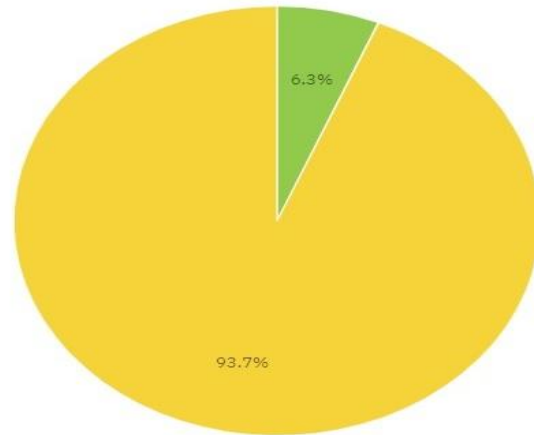


Demographics

Respondent classified by Gender

Gender

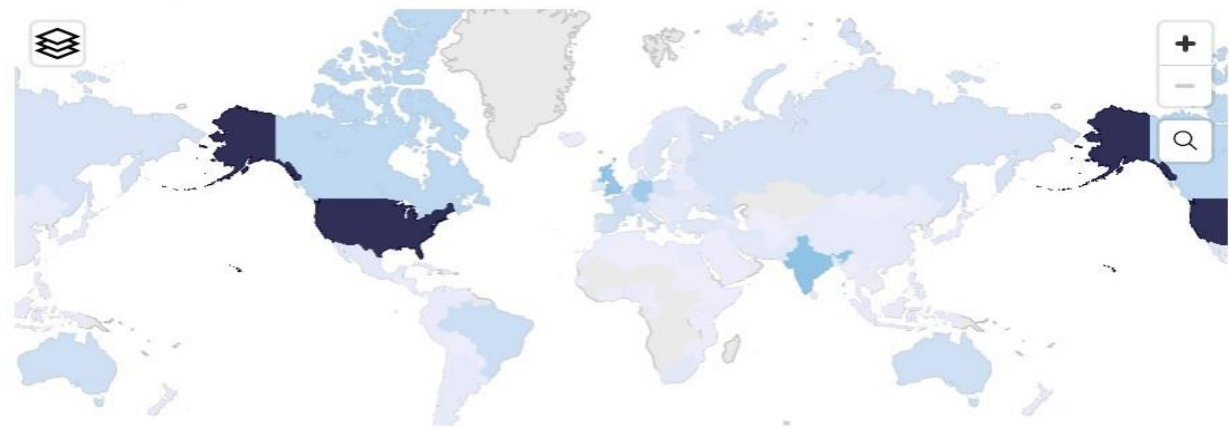
Woman Man



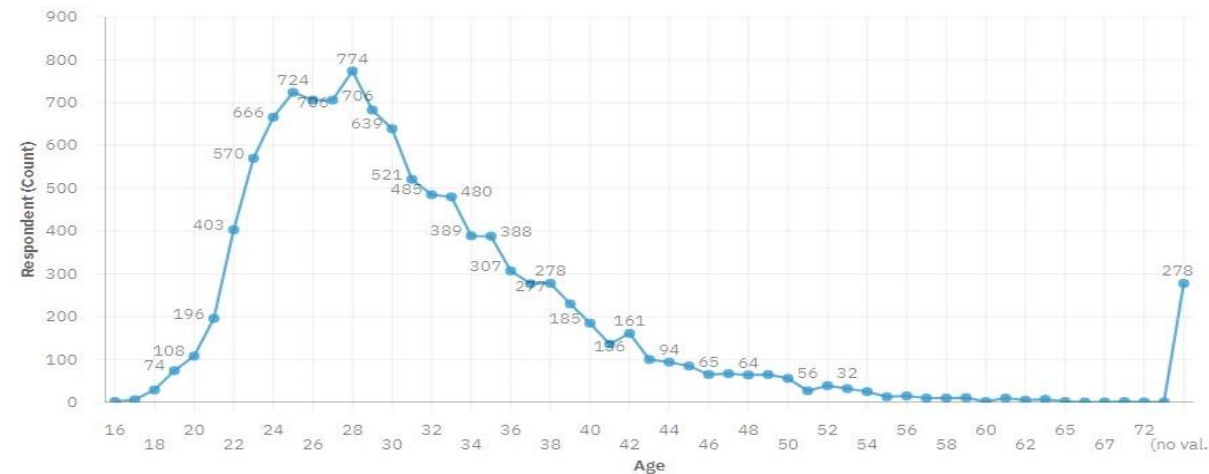
Respondent Count for Countries

Respondent (Cou...)

1 3,058



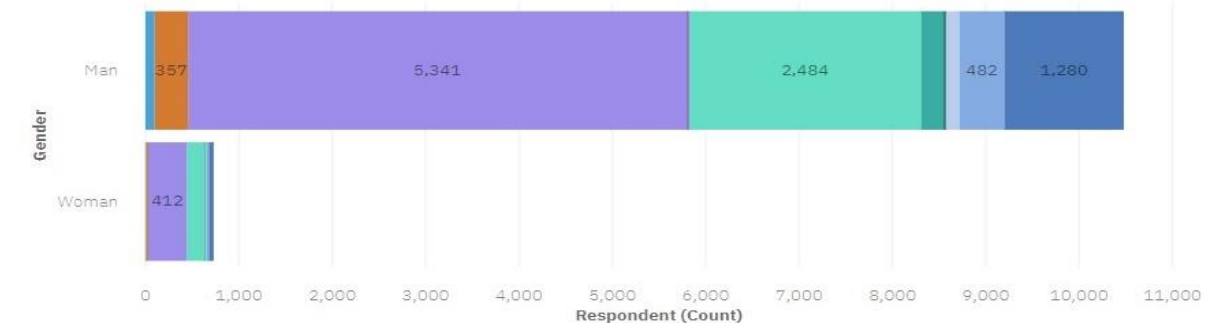
Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level

EdLevel

(no value) Associate degree Bachelor's degree (BA, BS, B.E...
I never completed any formal ... Master's degree (MA, MS, M.E... Other doctoral degree (Ph.D, E...
Primary/elementary school Professional degree (JD, MD, e... Secondary school (e.g. Americ...
Some college/university study ...



Discussion



Overall Findings and Implications

Findings

- JavaScript, HTML/CSS and SQL continue to be between the most desired programming languages among the others, but also Python and TypeScript are gaining a significant interest.
- PostgreSQL and MongoDB databases are gaining a lot of interest for the next year. Meanwhile there is a new interest for Redis and Elasticsearch, in contrast to the interest for MySQL and Microsoft SQL Server that is decreasing.
- Linux, Docker and AWS platforms keep a high interest. Regarding to Windows and Slack, it is observed a significant interest decrease.
- Interest for React.js and Vue.js web frames is growing rapidly, in comparison with the interest for jQuery that is being decreased.

Implications

- It is recommended, continuous training and updating of employees and new entrants, through seminars, in JavaScript, HTML/CSS, and SQL, as well as, in Python and TypeScript, that are gaining more and more interest.
- Redefinition of needs and data. Evaluation, for possible use of new databases, that are the future trends, according to the needs of the company.
- Regarding to platforms, it would be recommended for the company to evaluate, that there is no more interest for Windows and Slack, in comparison with Linux, Docker and AWS and accordingly, come to an agreement, with new companies, that are offering their platforms, which are blooming in the labor market.
- Training in new web frames, such as, React.js and Vue.js, that are gaining interest, in order to keep up-to-date.

Conclusion



- Investing money, for the recruitment of new staff, with skills required by the job market, as well as, to train them, for their quick integration into the company.
- Investing money, in training those already employed, to acquire the new skills and to be informed about the new trends.
- Investing money, to acquire new, more helpful and modern platforms and to remove the outdated ones.

Appendix



Any relevant additional charts, or tables that have been created during the analysis phase, follow below:

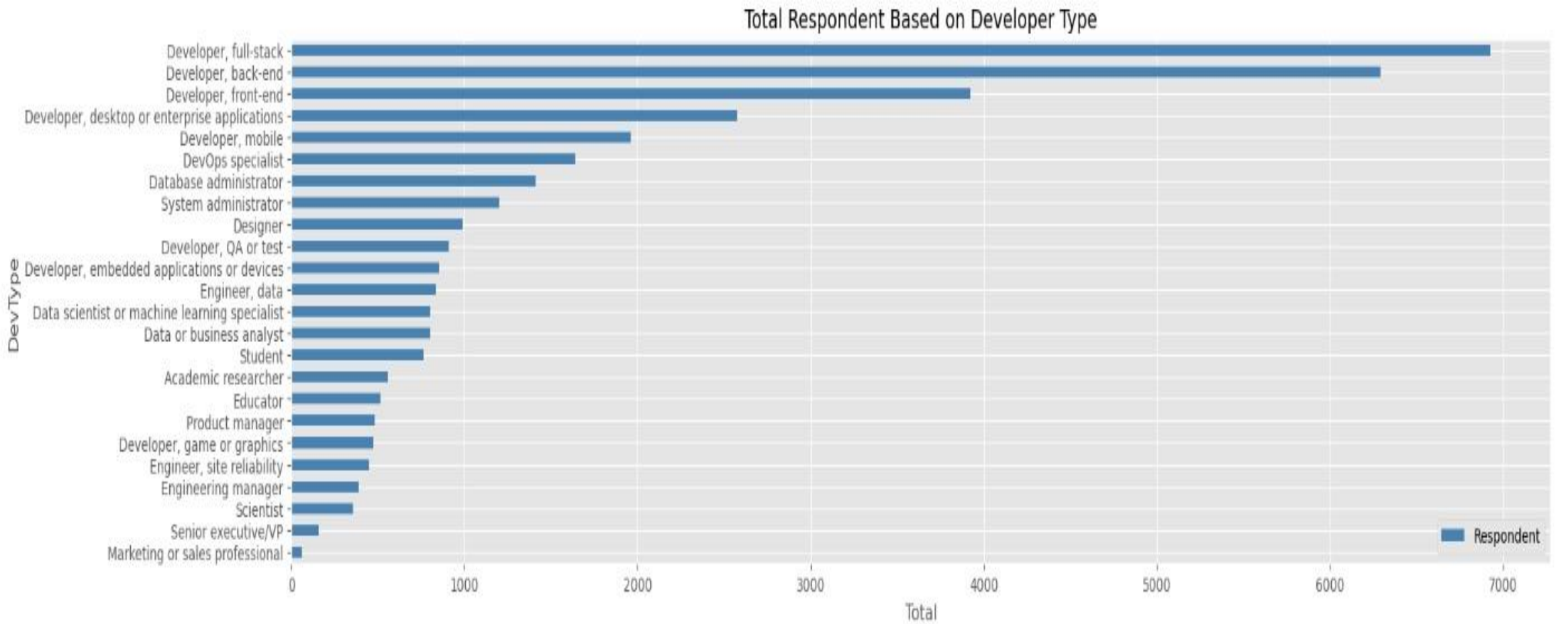
Correlation Info

Here, we have a correlation, between the Age of the Respondents and some other variables of our dataset.

- Negative correlation, between Age and the time that a Respondent devote to the code review(CodeRevhrs).
- Highest correlation, between Age and the Annual Compensation of the Respondents(ConvertedComp).

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

Total Respondent by Developer Type

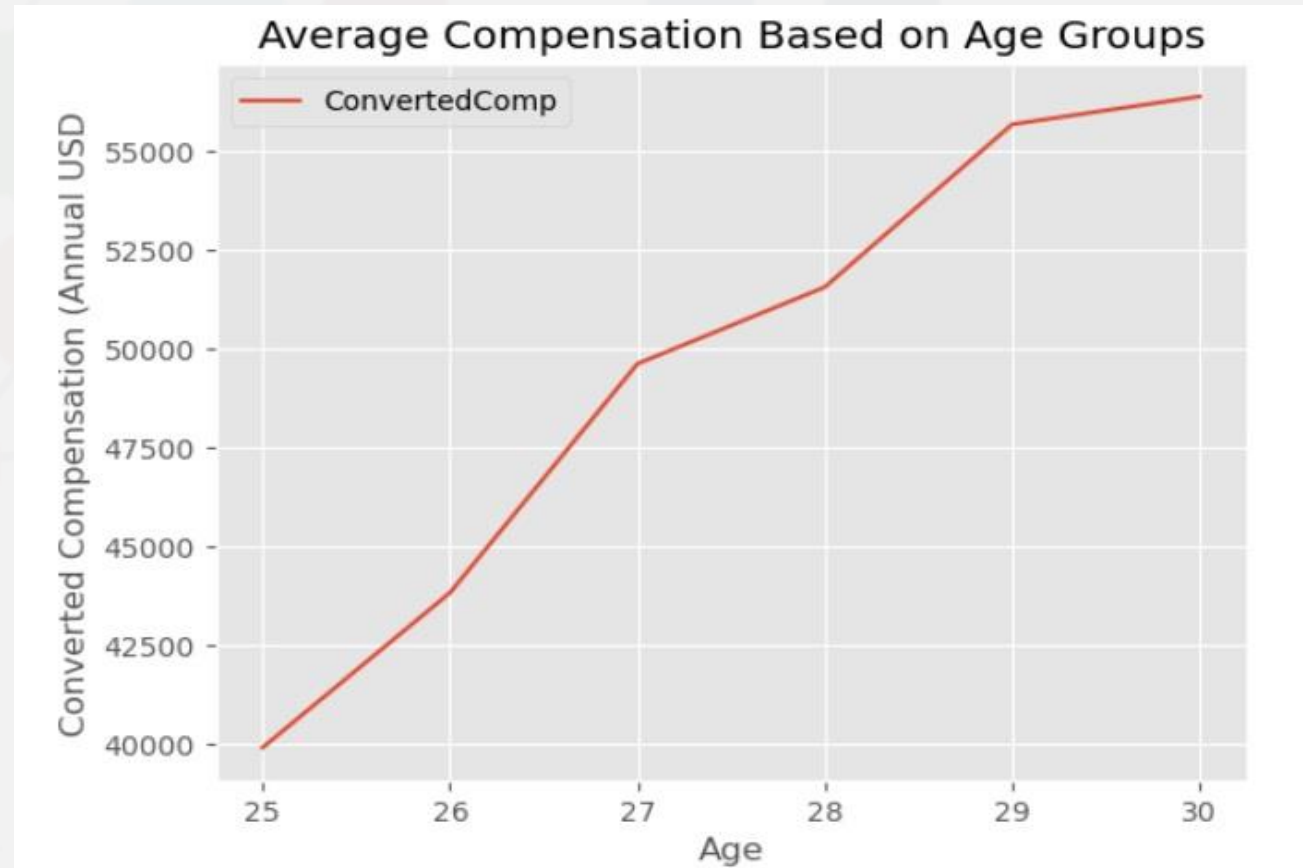


Total Respondent Findings

- Majority of the Survey Responders are the Developers full-stack. The Developers back-end follow and in the third place, with fewer responses, are the Developers front-end.
- Marketing or sales professional seems to be the type, that gave the minimum of the responses.

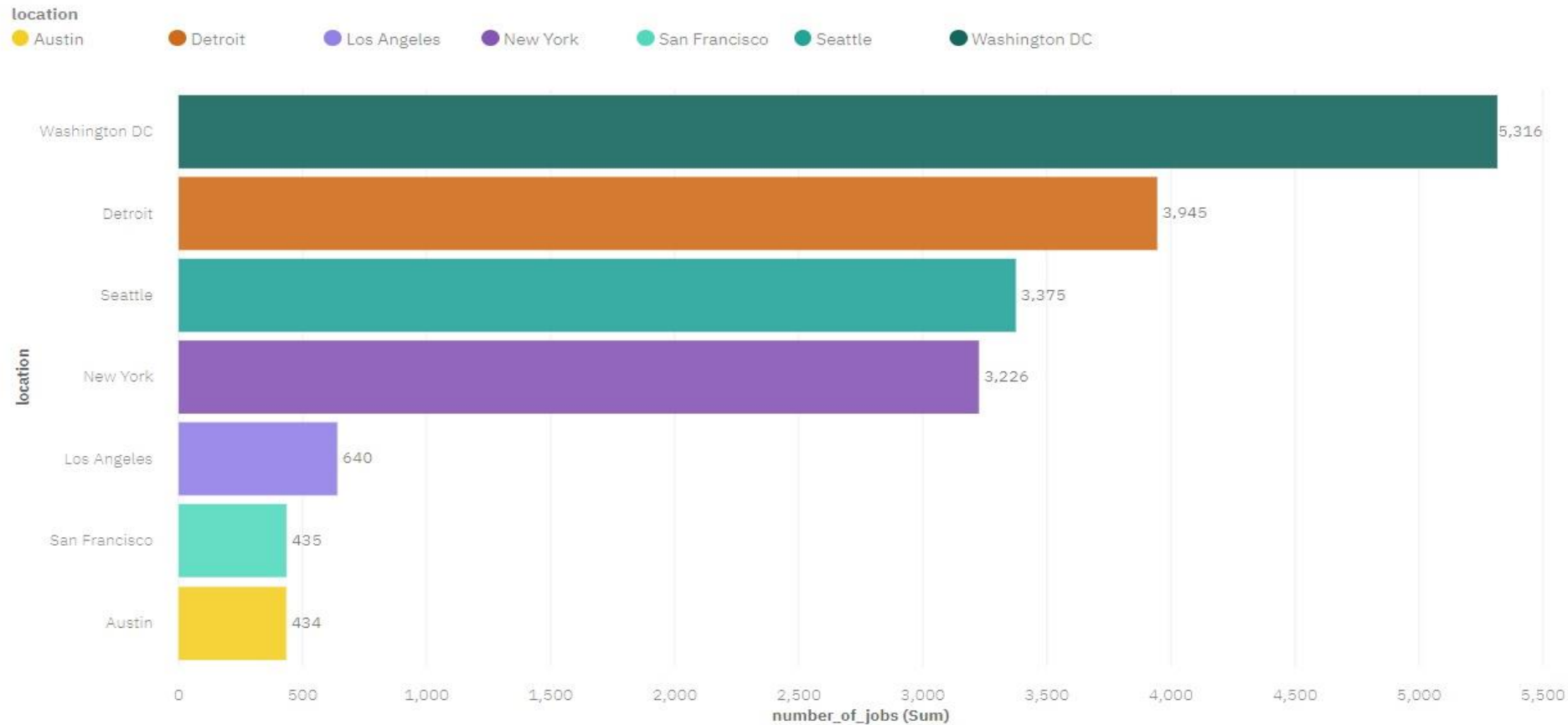
Average Compensation by Age

- With an increase of the Age, there is an increase in the Annual Compensation of the Respondent.
- The correlation here is positive.



Job Postings

Number of Jobs by Location

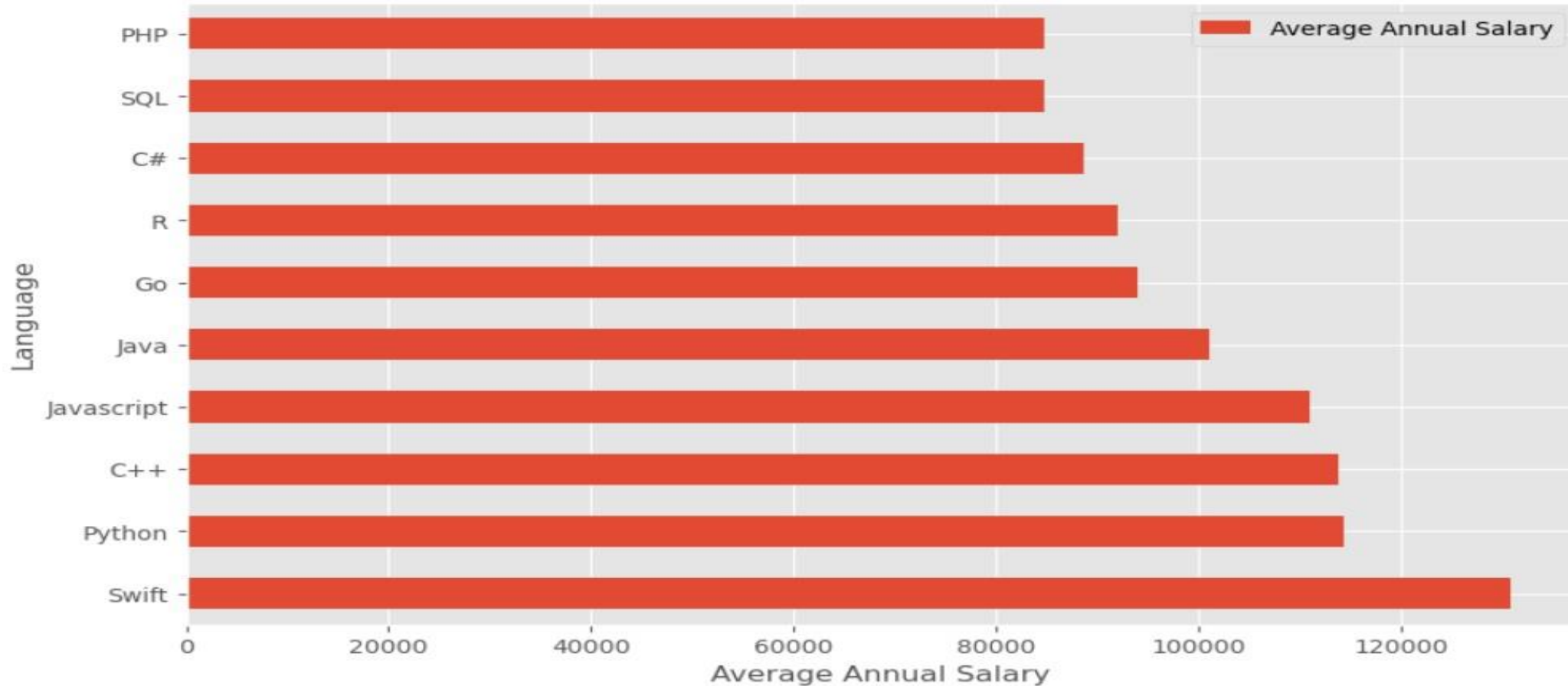


Job Postings Findings

- There is a significant difference in the number of job postings, between Washington DC(5316) and Detroit(3945). They are being followed by Seattle, New York and Los Angeles.
- In comparison with San Fransisco(435) and Austin(434), that they do not appear a remarkable difference between them.

Popular Languages

Average Annual Salary by Programming Language



Popular Languages Findings

- Among the most popular languages, the most well-paid annually programming language is Swift with Python, C++ and JavaScript to be close behind.
- Among the least popular languages, the worst-paid are, with no remarkable differences between them, SQL and PHP.