Analysis of Stack Overflow Developer Survey data

Created by : MANAS
Fellow learner

© IBM Corporation. All rights reserved.

OUTLINE

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

EXECUTIVE SUMMARY

- The data is first extracted and saved
- The collected data is then cleaned and formatted
  - Removed rows with the null values
  - Filled missing data
  - Normalized the data
- Exploratory Data analysis is done
- Data visualizations are created with python
- Dashboards are built on Looker studio

INTRODUCTION

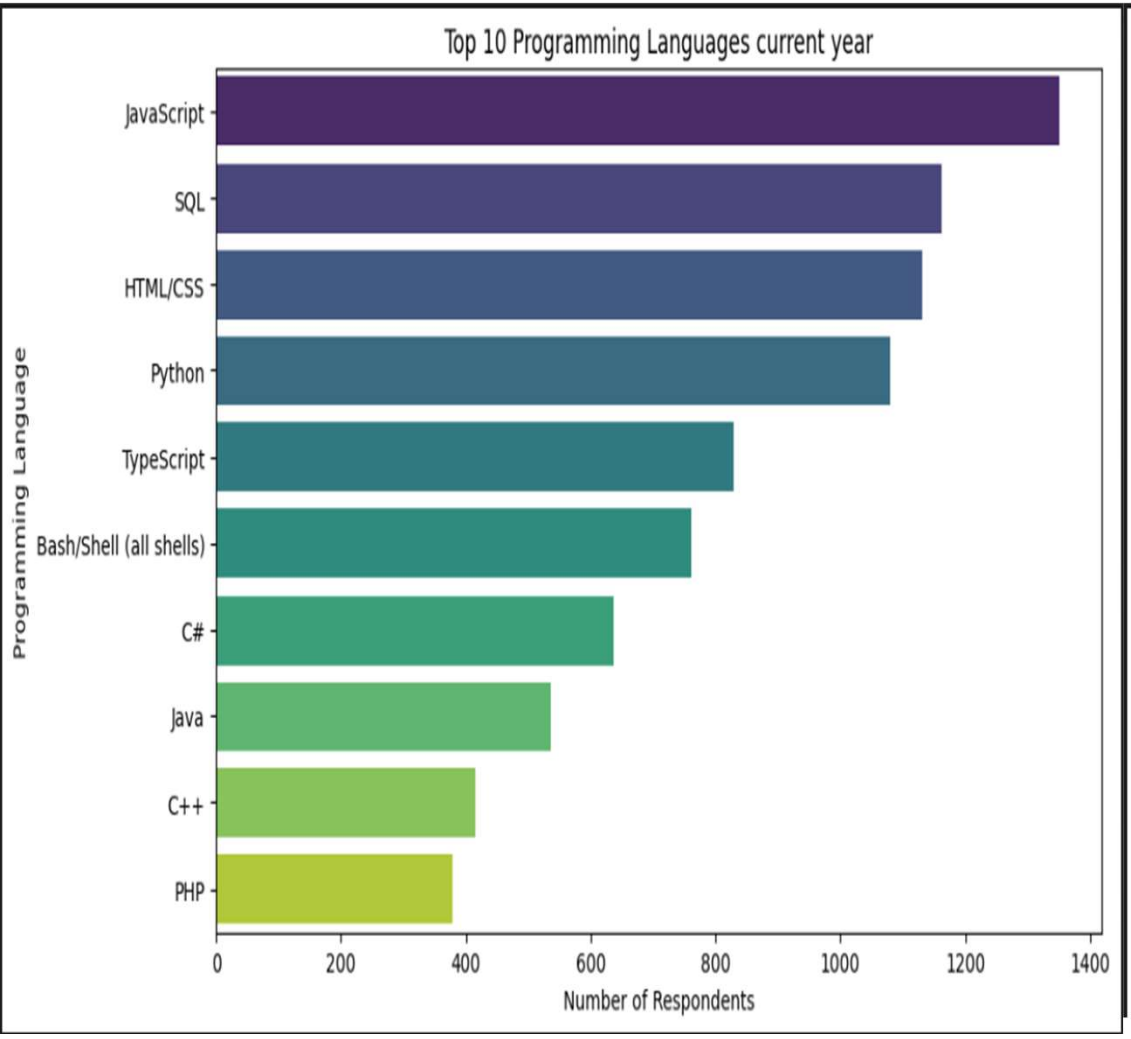
- The required data is collected by scraping internet websites, accessing APIs, and working with datasets like the latest Stack Overflow Developer Survey in various formats such as .csv files, Excel sheets, and databases.
- After gathering the data, data-wrangling techniques are applied to prepare the data for analysis.
- Once the data is prepared, statistical techniques are employed to analyze the data, identifying key trends and insights.
- Then on Looker studio dashboards are created that visualizes the results.

METHODOLOGY

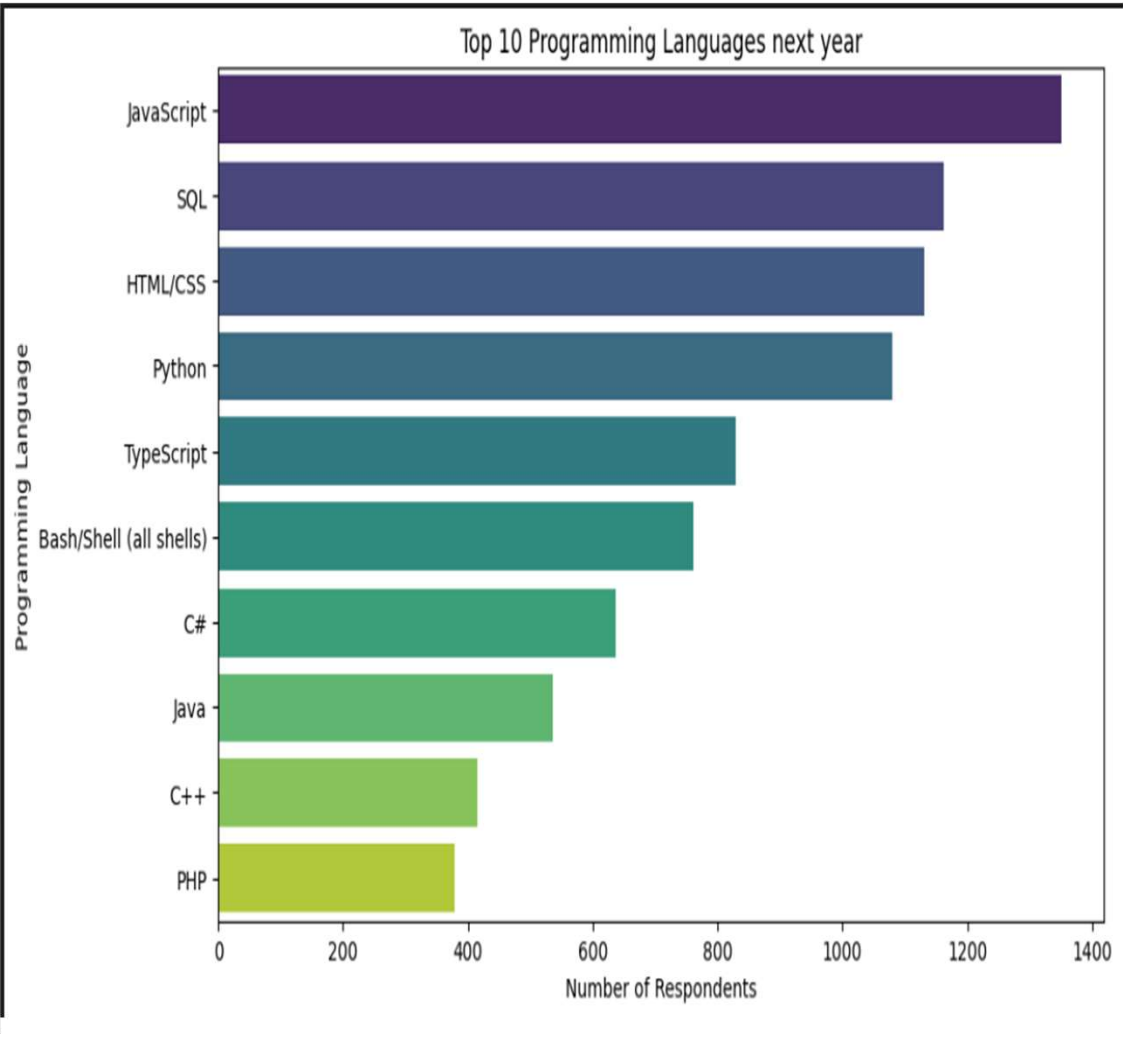
- Data wrangling, also known as data munging, involves cleaning and preparing datasets to make them ready for analysis.
  - Identify and remove duplicate rows
  - Analyze missing values
  - Examine employment status
  - Normalize data for comparative analysis
- Performed comprehensive exploratory data analysis (EDA) by integrating multiple visualization and analytical techniques.
- Built dashboard using data from the Stack Overflow survey on Google Looker Studio. It is divided into the following sections:
  - Current technology usage tab
  - Future technology trends tab
  - Demographics tab

PROGRAMMINGLANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

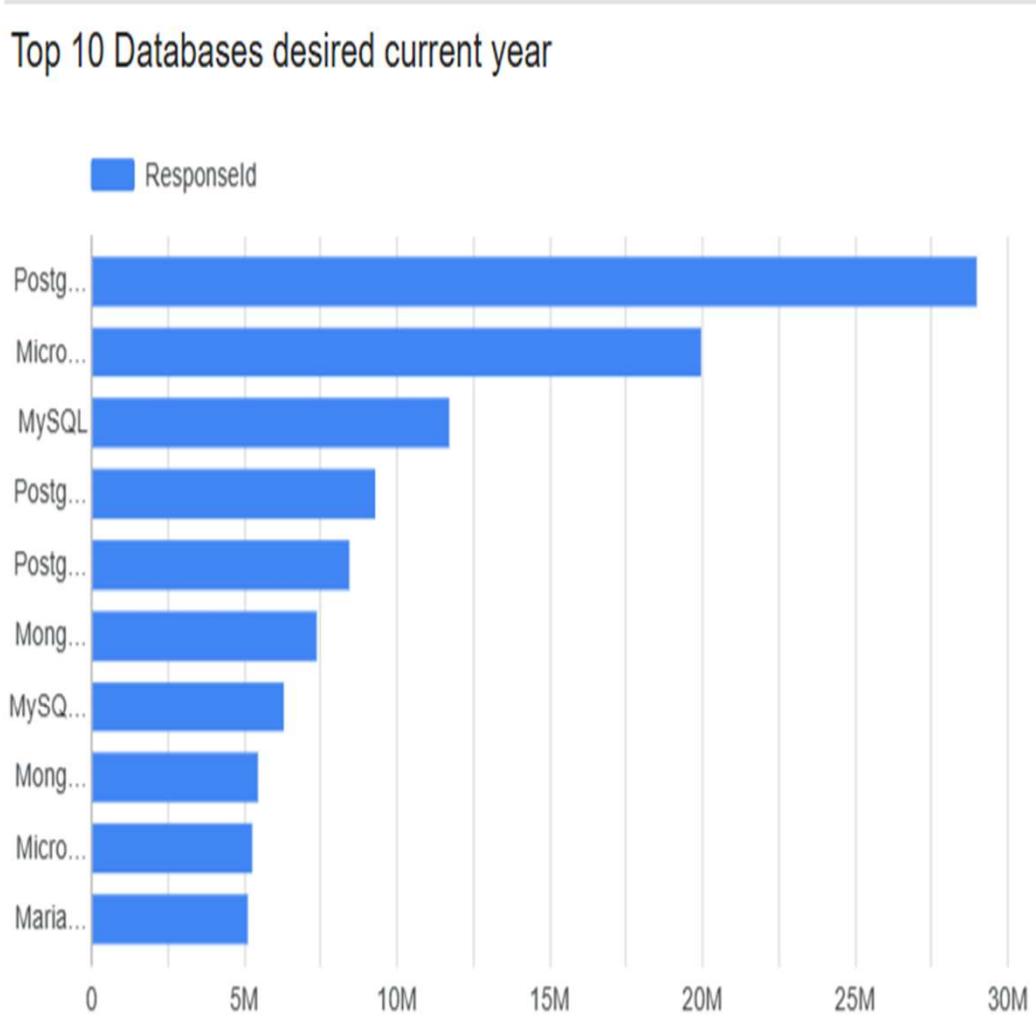
- Java script, sql, css are most used one
- There is not much change in trends
- Type script,bash might me more used in future

Implications

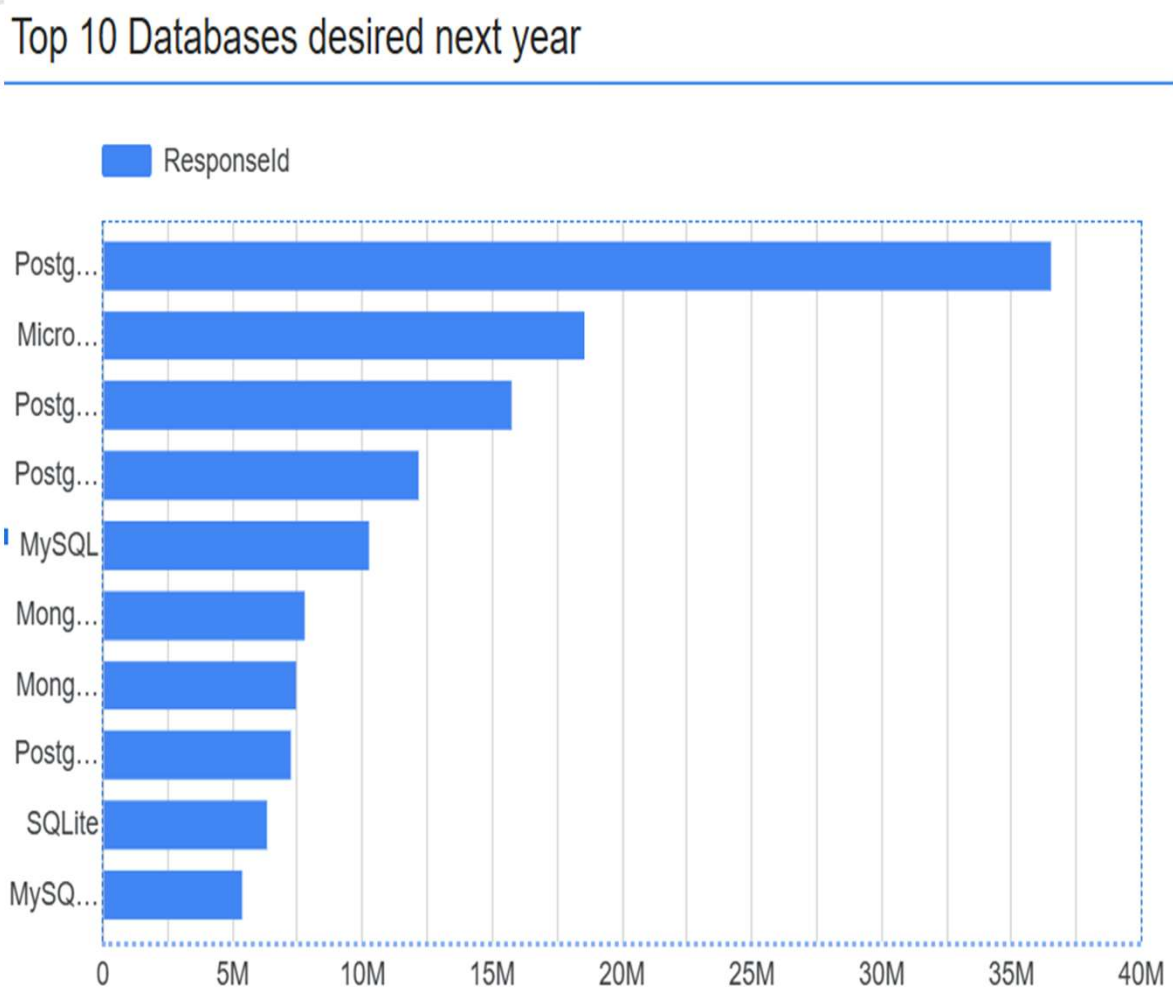
- Most of jobs are based on these languages
- Trends are similar
- There is increase in demand for type, script and bash

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Postgress is most used database, followed by Microsoft.
- The use of MySQL database trend has decreased in next years
- The Database trend is increased in next years

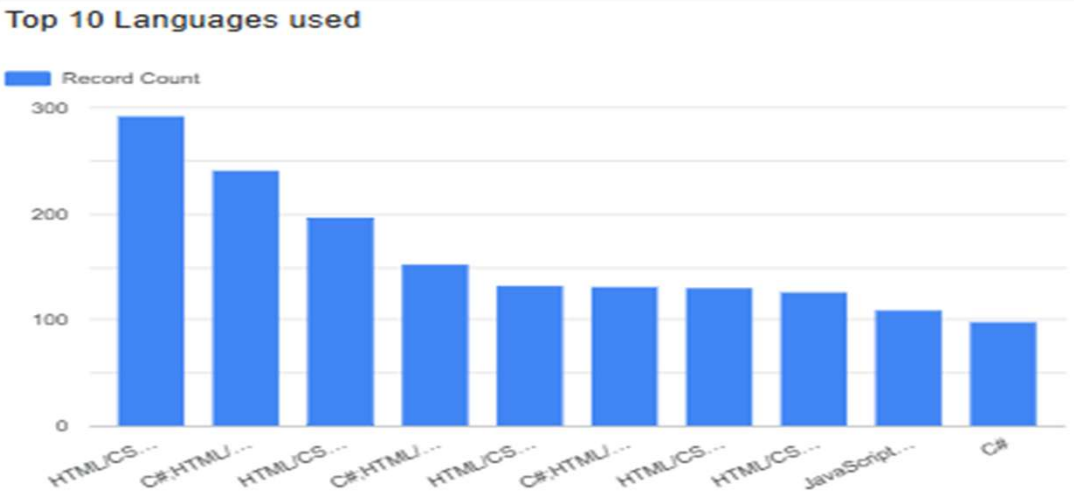
Implications

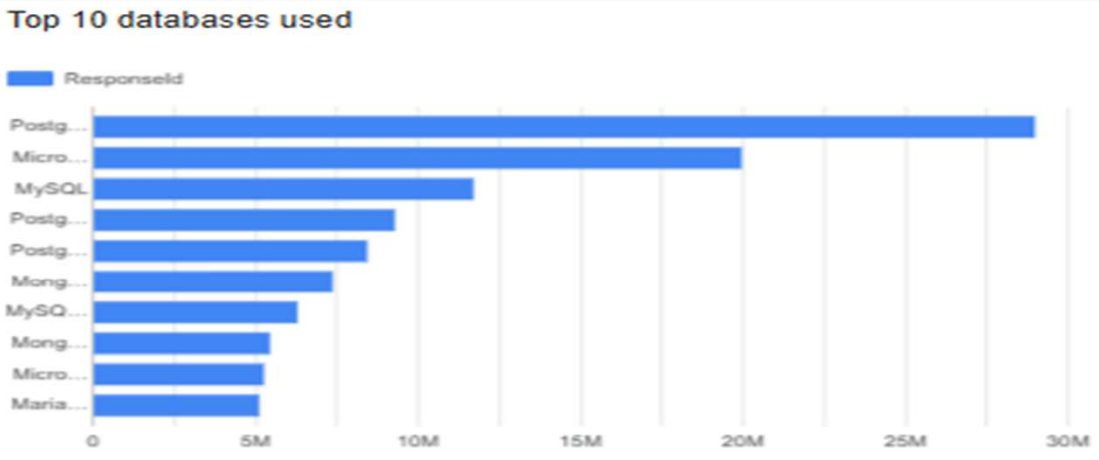
- Most of jobs seekers want to learn and work in Postgress more
- The desire to work with MySQL database seems to have reduced among jobseekers
- Shows growth in usage of databases

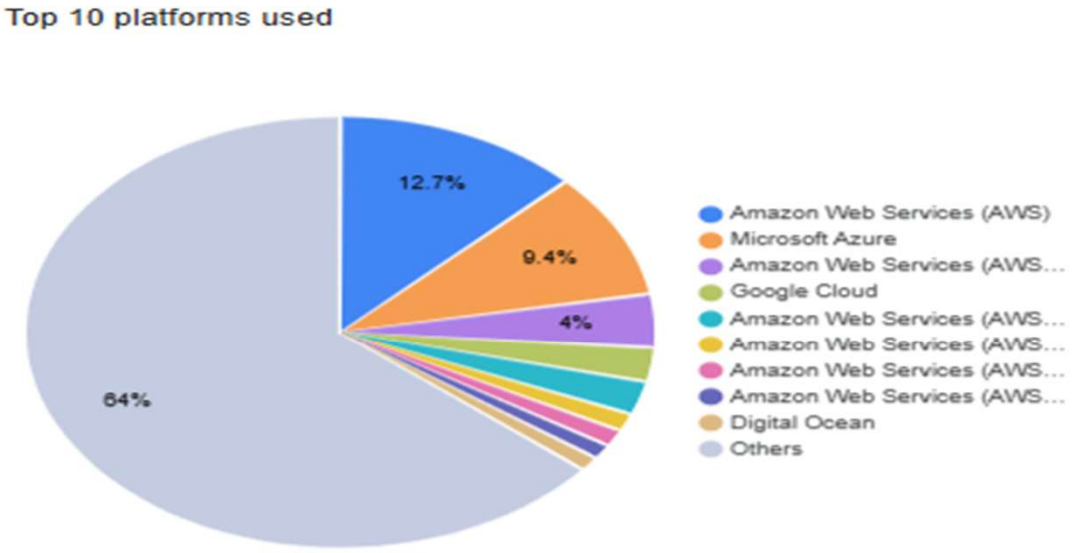
DASHBOARD

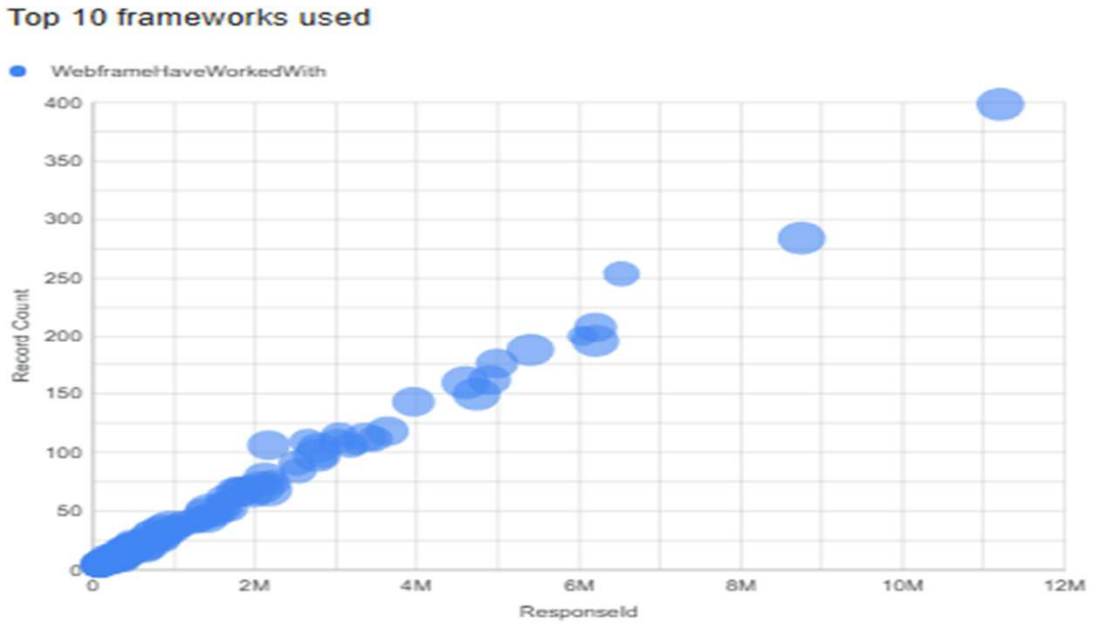
The data is uploaded to looker studio and 3 different dashboards are built for further analysis of the data.

Current Technology Usage

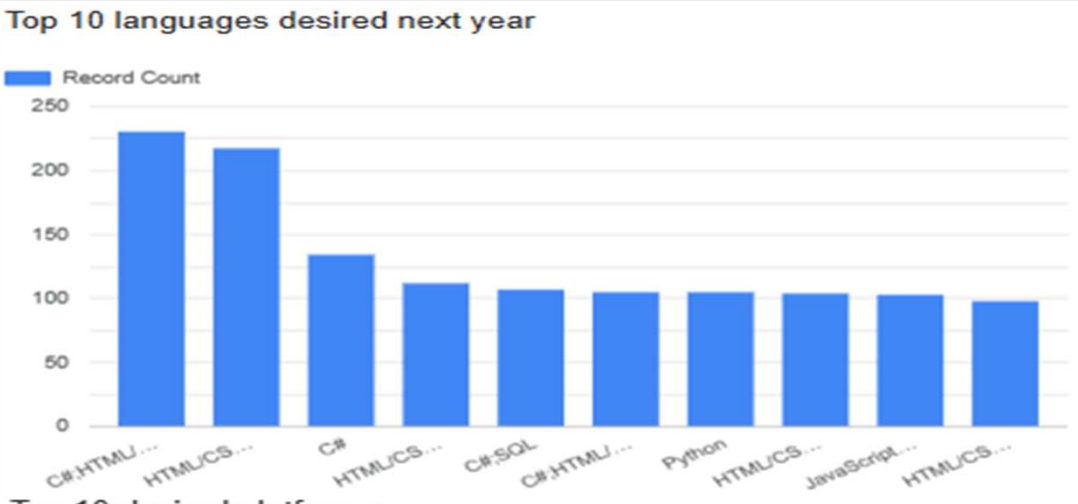


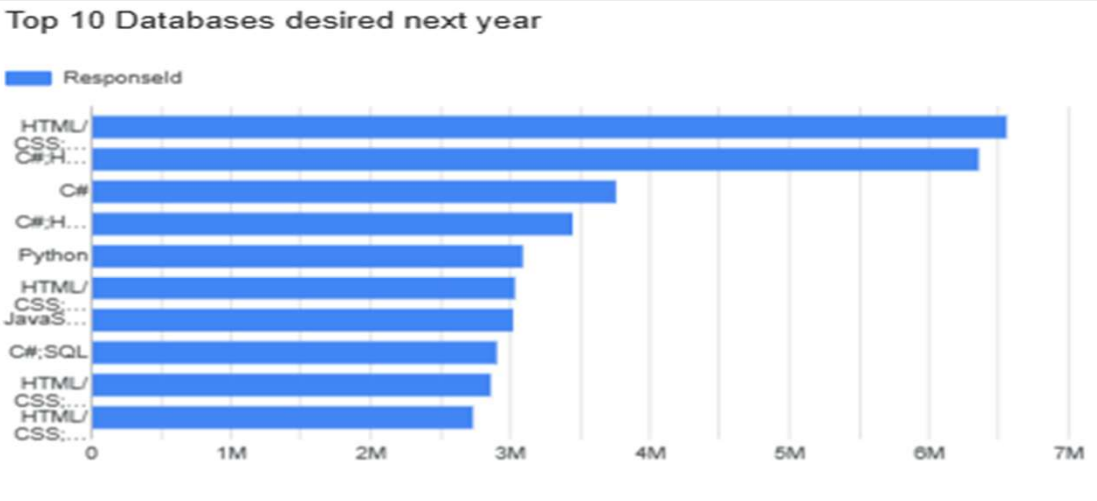


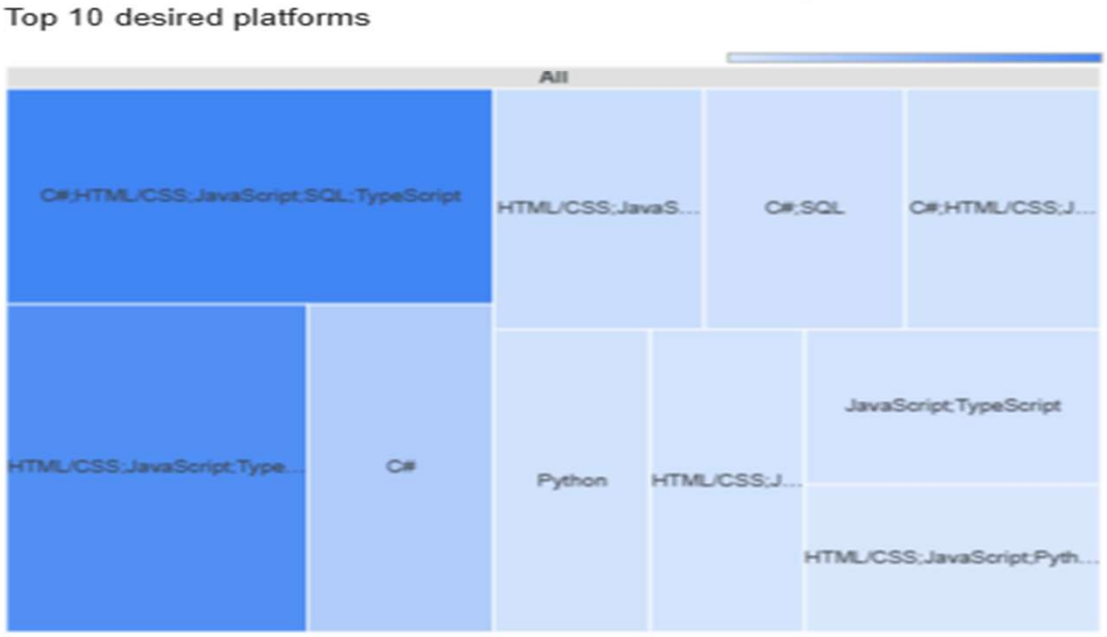


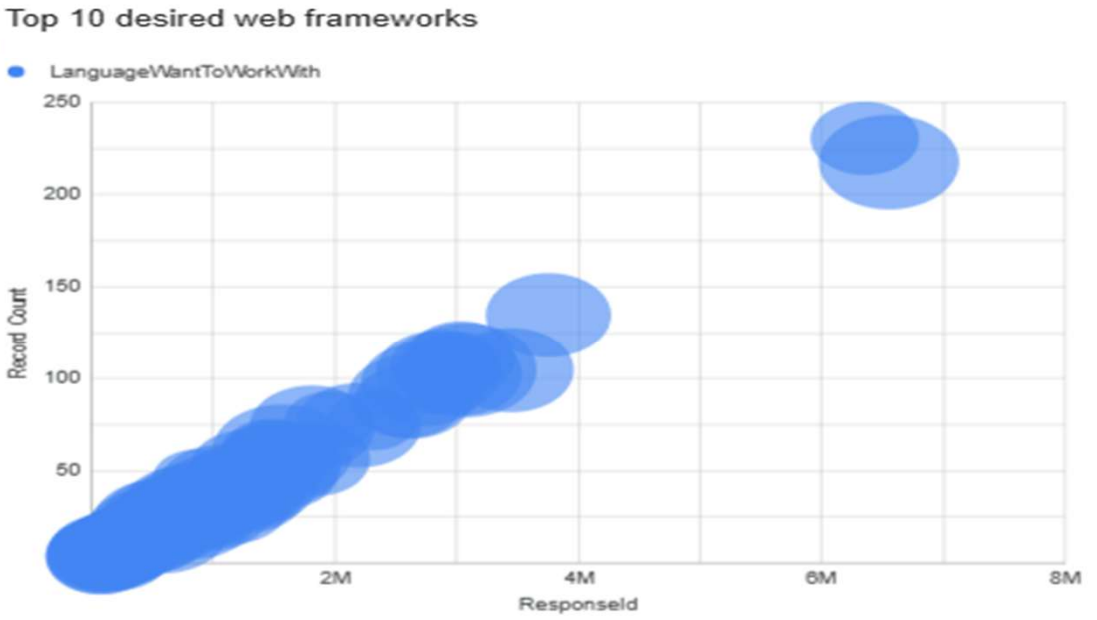


Future Technology Usage



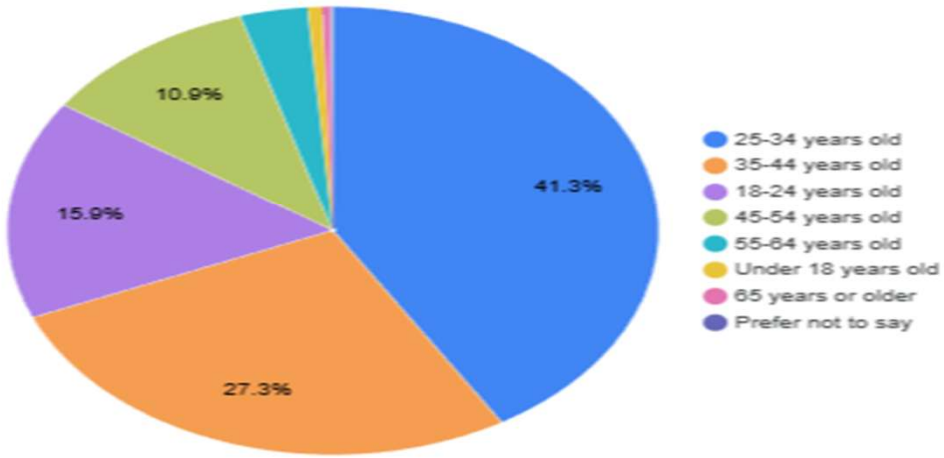




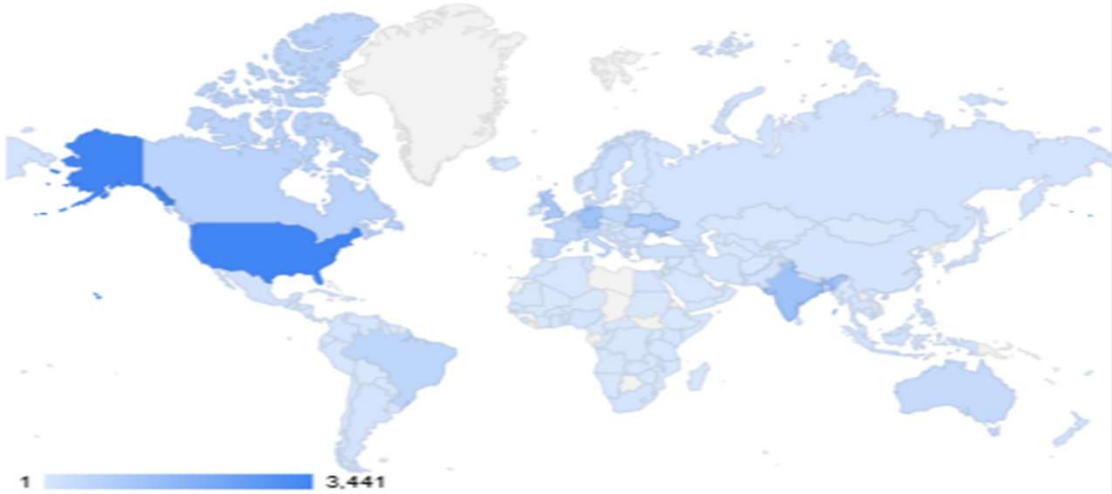


Demographics Report

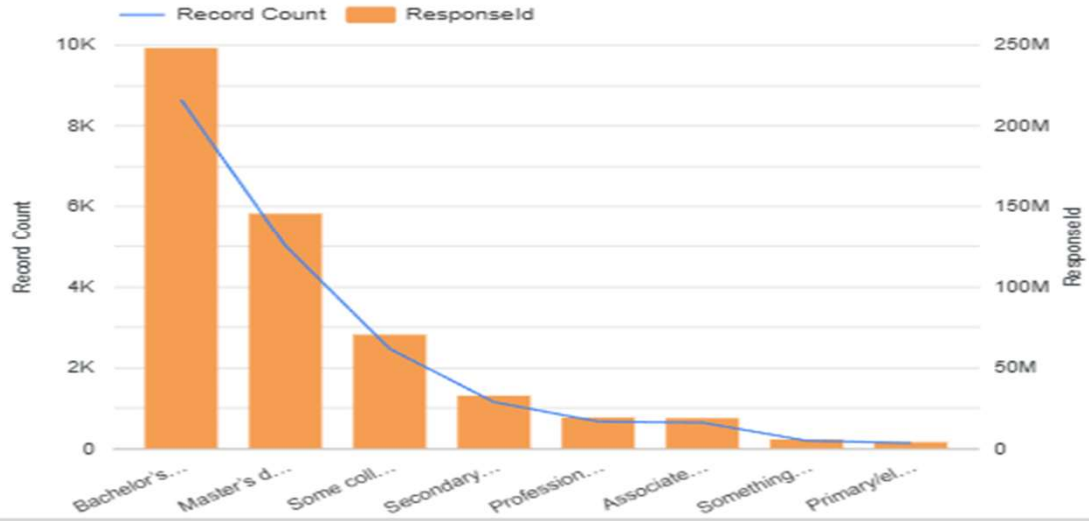
Respondents by Age



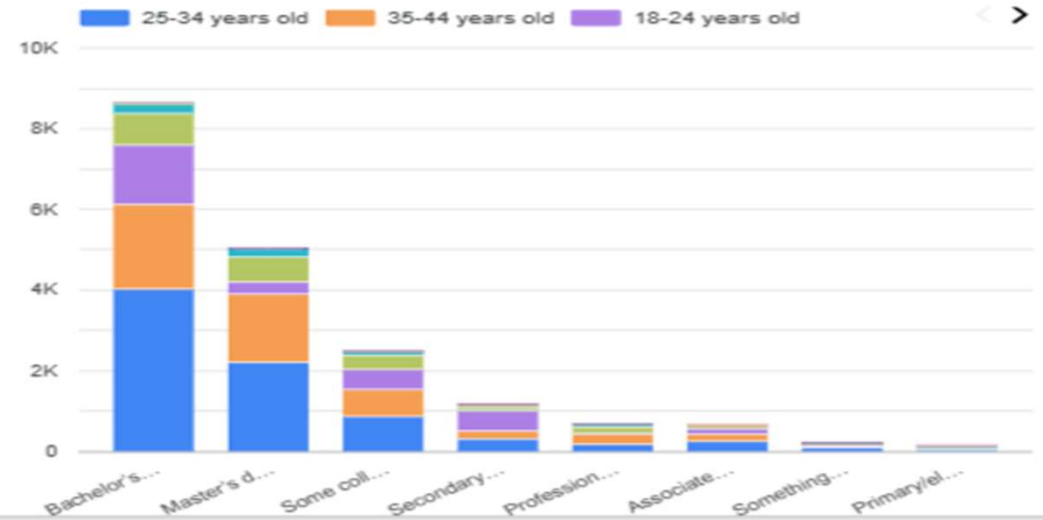
Respondents count by Country



Respondents distribution by Education level



Respondents count by Age, Classified by education level



DISCUSSION

- On the Google Looker Studio , the dashboard with 3 different report findings is created .
- We can see the different trends among the jobs listed and skills used and needed for these jobs.
- There are many teenagers as well getting jobs with their skills.

OVERALL FINDINGS & IMPLICATIONS

Findings

- In current technology usage : HTML/CSS is most used programming language, postgress is most used database and AWS is most used platform.
- There are more of young people working in these jobs .
- most of job listings are from USA and most of them are young people with bachelors degree .

Implications

- Most of job seekers need to learn HTML/CSS , Postgress and AWS skills
- Young people need to earn Bachelors degree to get more jobs
- Getting job in USA has relatively higher chance due to more job oppurtunities

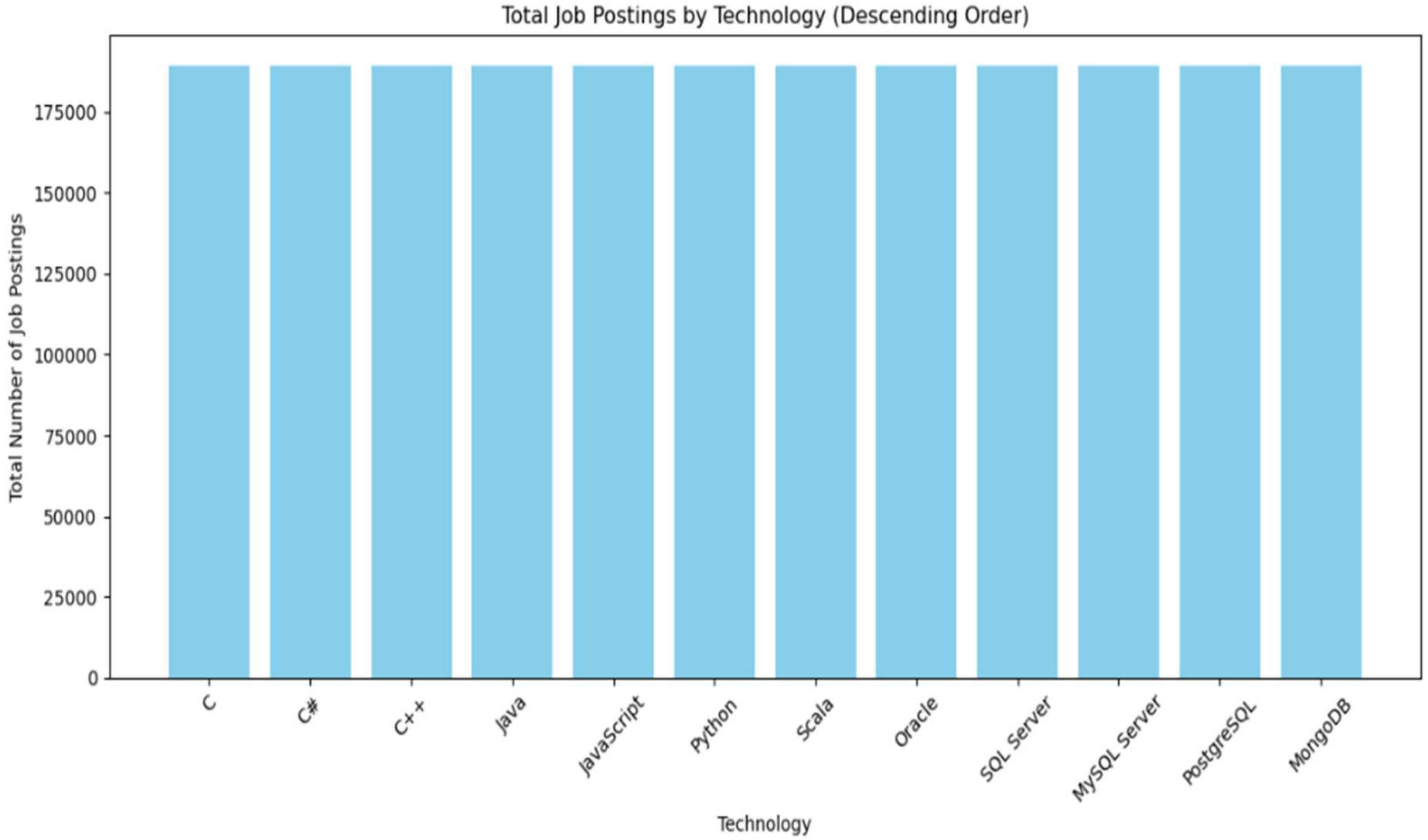
CONCLUSION

- HTML/CSS is most used and desired programming language.
- Postgress is most used and desired database
- AWS is most used platform
- Job seekers must have minimum of Bachelors degree and have most of these skills.
- With experience and additional degrees the salary will be increased so much

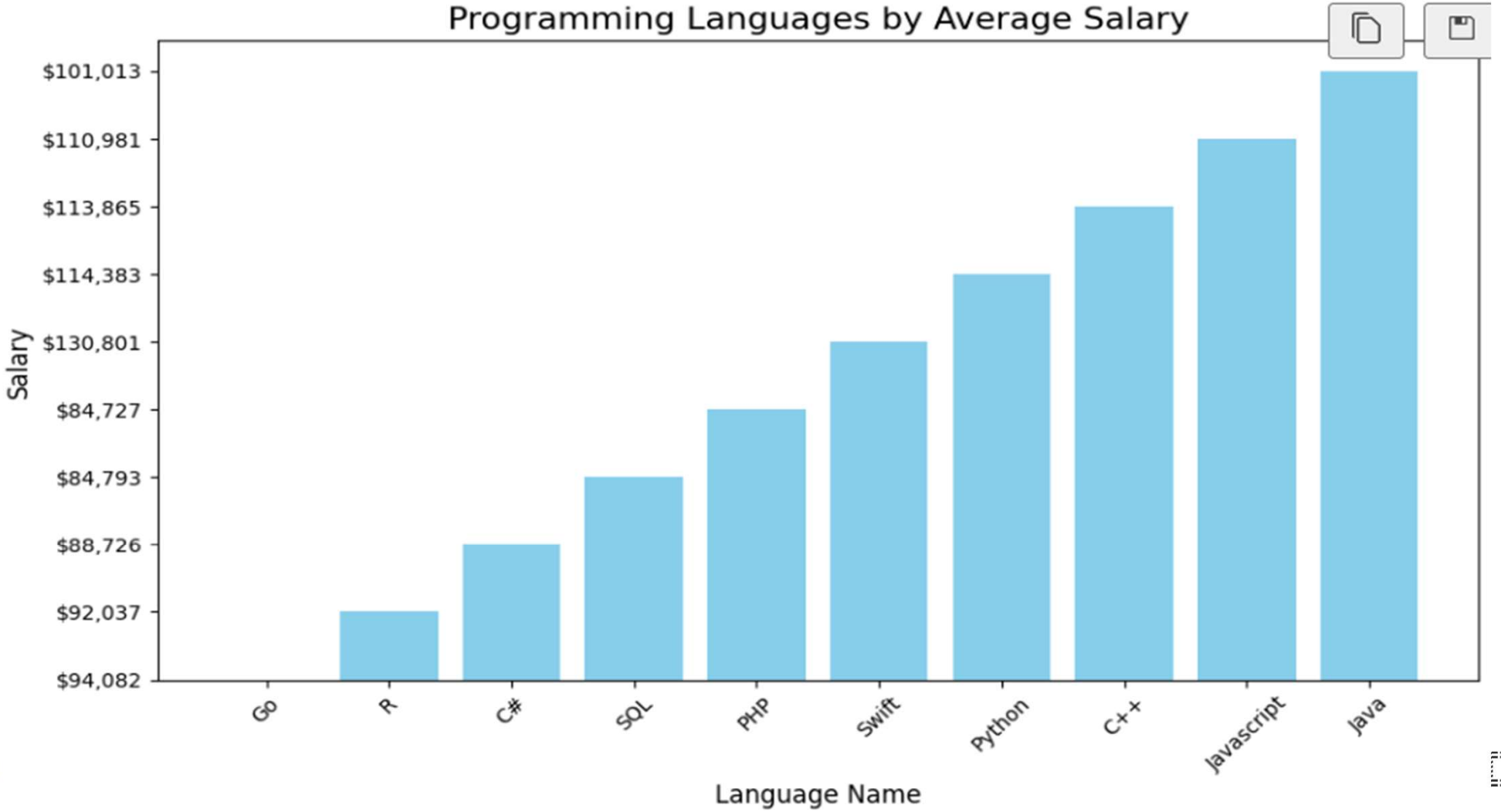
APPENDIX

- Included two more relevant additional charts on job postings and popular language distribution, that have created during the analysis phase.

JOB POSTINGS



POPULAR LANGUAGES



Thank You!

For your attention