Data Analyst Capstone

Handy Thong

01 Feb 2023

OUTLINE

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

EXECUTIVE SUMMARY

- Trends in programming languages
- Trends in databases
- Trends in IDE/Platform
- Trends in WebFrame
- Demographics survey
- Gender gap in jobs
- Technological gap in countries

INTRODUCTION

- Technologies are always changing with fast paces
- We need to keep up with the changing pace of technologies and be competitive
- To be on top of the changes we must identify future skill requirements such as:
  - Top programming languages in demand
  - Top database skills in demand
  - Top IDE/platform used by developer
  - Top webframe used by developer

METHODOLOGY

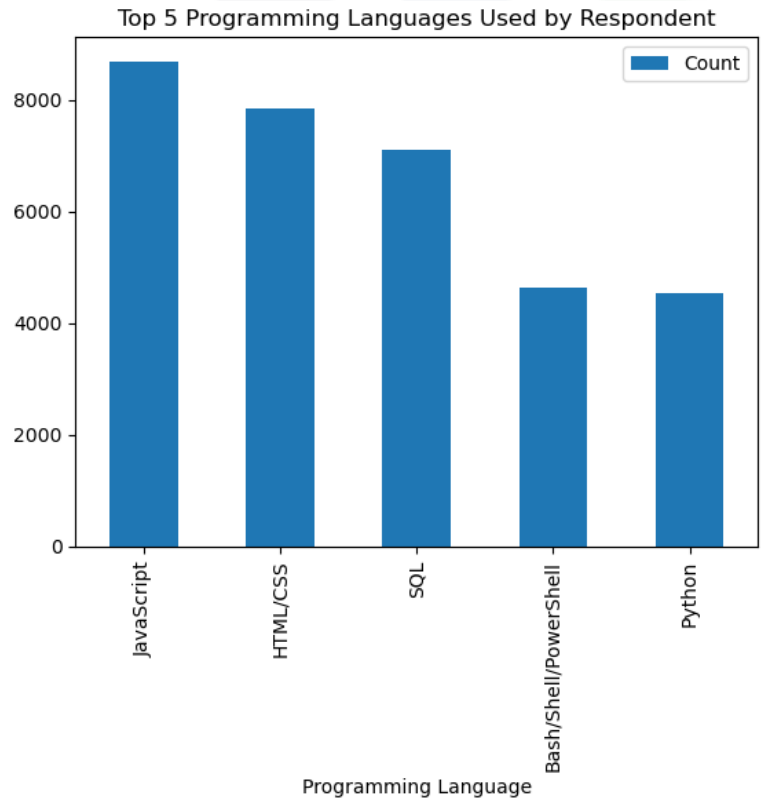
- Data Collection: Web scraping and accessing API via GitHub Jobs API.
- Data Wrangling: Cleaning of data by removing duplicates data, finding and replacing missing value, and normalizing data.
- Data Analyzing: Analyze the data distribution such as mean, median, and mode. Finding outliers inside the data. Finding correlation between the data.
- Data Visualization: Visualizing distribution of data, relationship between data, composition of data, and comparison of data.

RESULTS

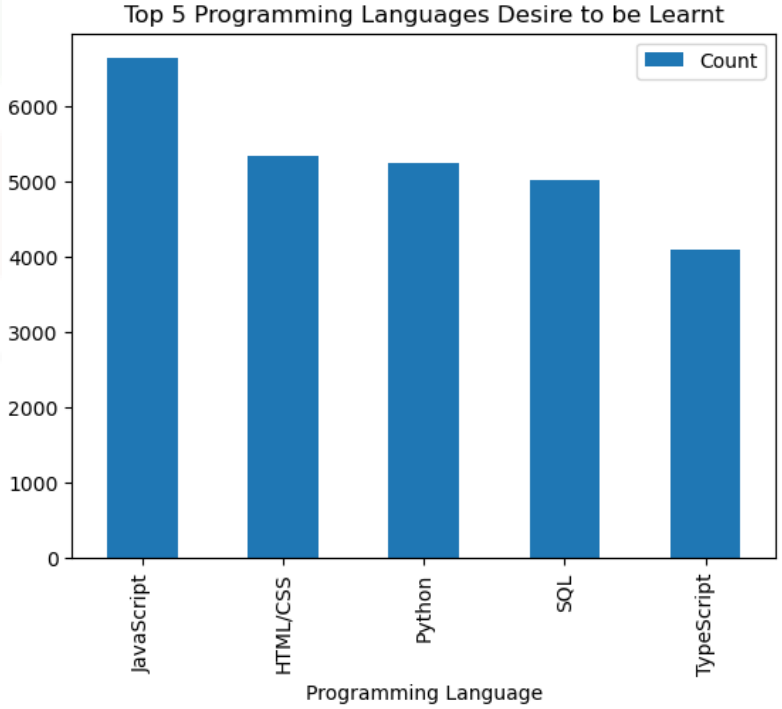
- Top Programming Language Trends
- Top Database Trends
- Top IDE/Platform Trends
- Top WebFrame Trends
- Demographic Surveys
- Interactive Dashboard

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

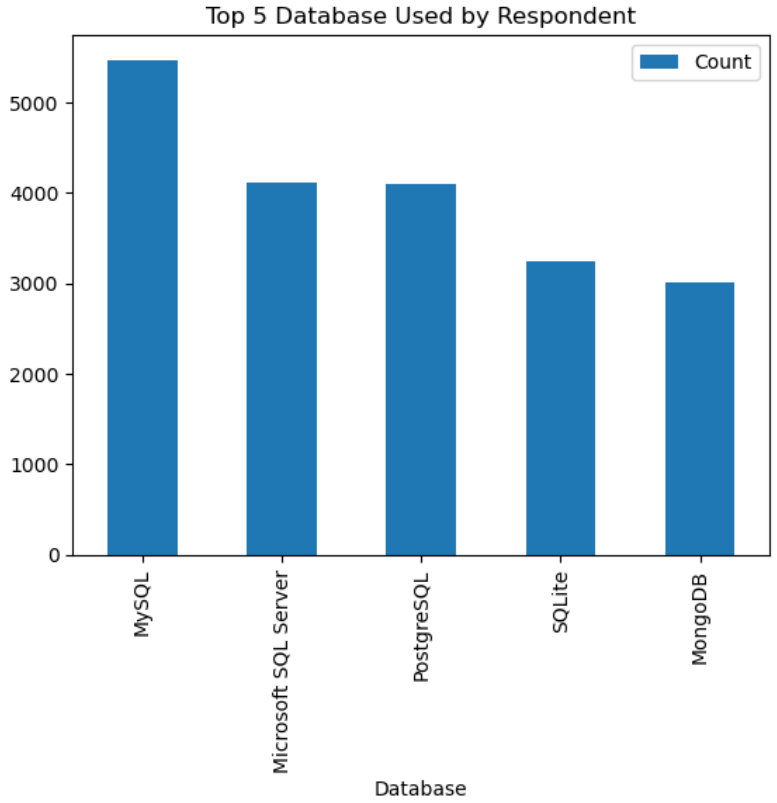
- JavaScript, HTML/CSS and SQL are top 3 in this year
- Python and Typescript are becoming popular next year
- PowerShell edged out next year, while SQL is still in top 5 next year

Implications

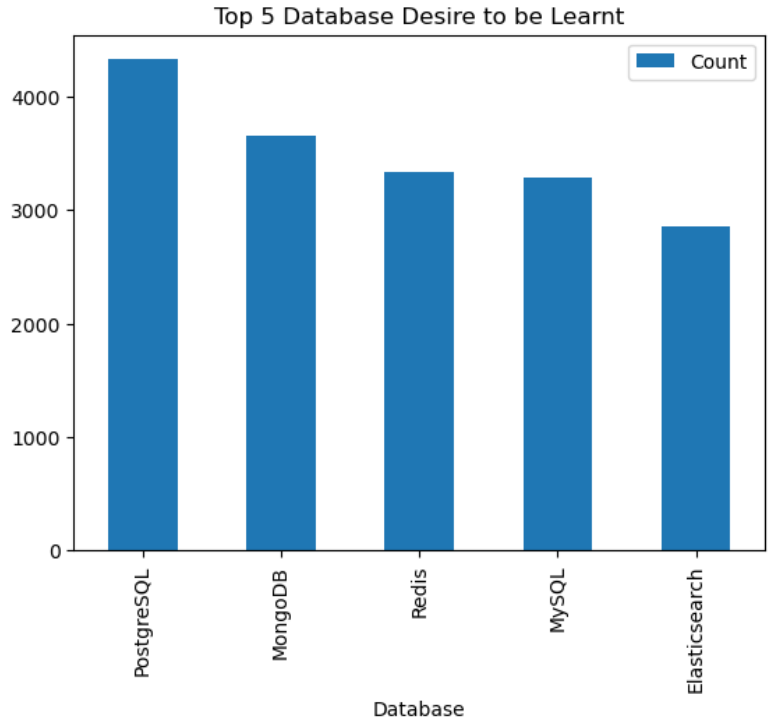
- Web developments are still in high demand
- With AI and ML rising, data analyst starts to shift to python
- SQL is still a required skills for data analyst and high in demand

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- MySQL is the most popular this year
- Microsoft SQL is second most popular this year
- MongoDB and Redis are upcoming favorites next year

Implications

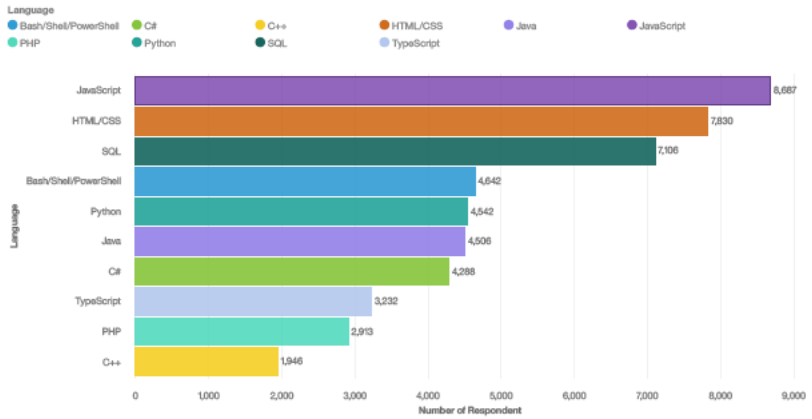
- Open-source databases are still preferable in companies
- Interest in NoSQL database are increasing
- Wide variety of data requires NoSQL database to be stored and used effectively

DASHBOARD

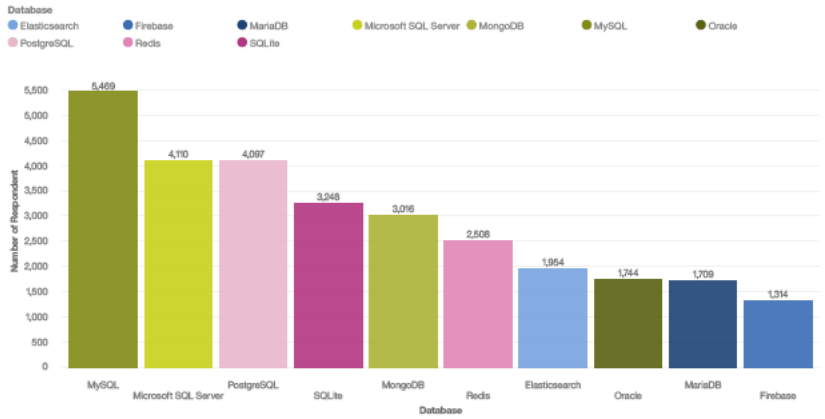
Permanent Link to Cognos Dashboard

DASHBOARD TAB 1

Top 10 Language Used by Respondent



Top 10 Database Used by Respondent

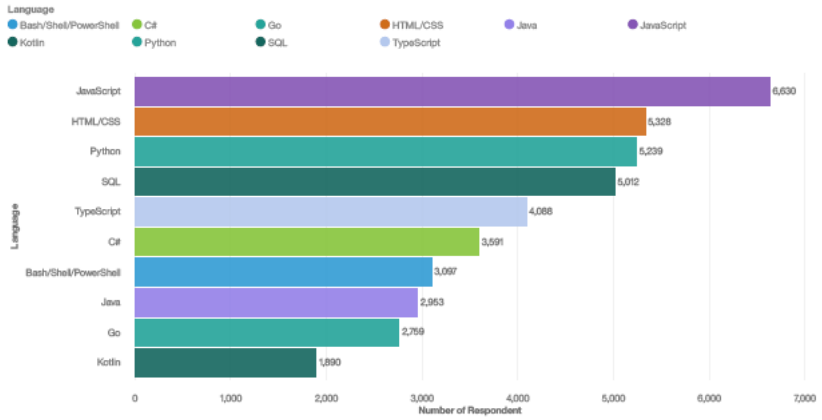


Platform Used by Respondent

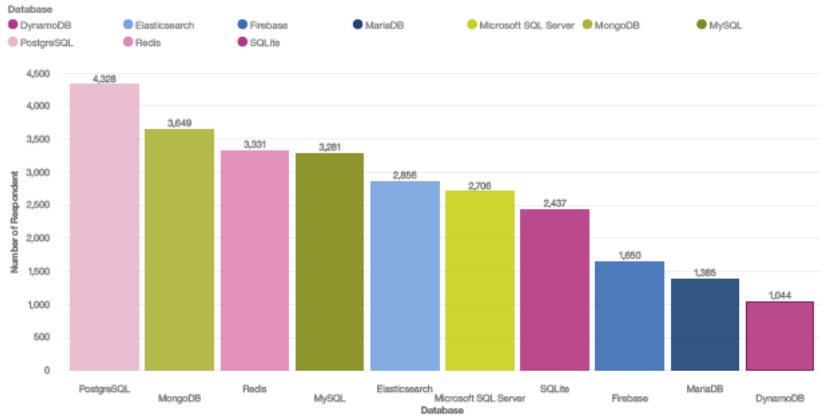
Top 10 WebFrame Used by Respondent

DASHBOARD TAB 2

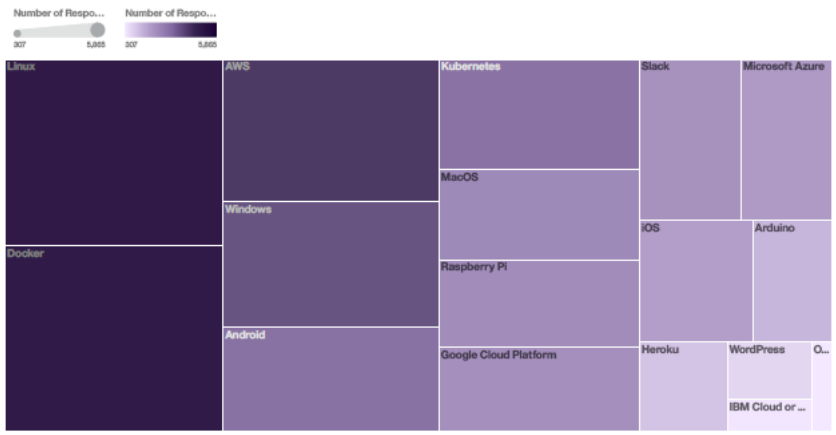
Top 10 Language Desire to be Learnt by Respondent



Top 10 Database Desire to be Learnt by Respondent



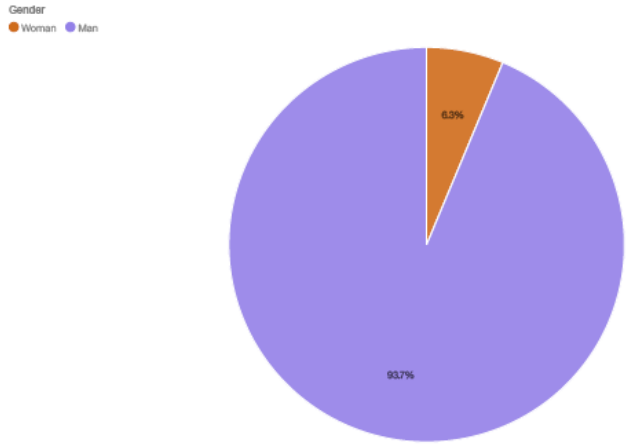
Platform Desired to be Learnt by Respondent



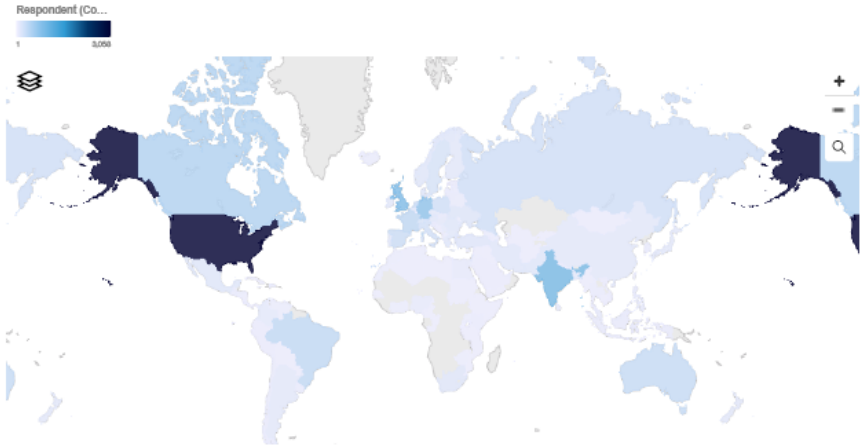
Top 10 WebFrame Desire to be Learnt by Respondent

DASHBOARD TAB 3

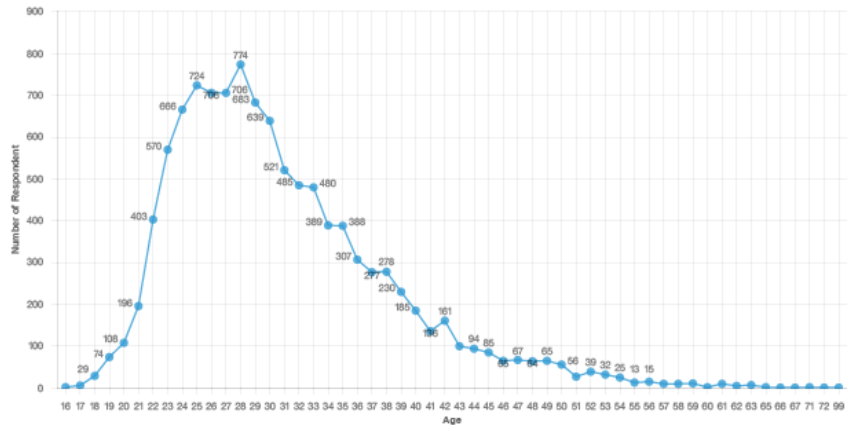
Respondent Gender Demographics



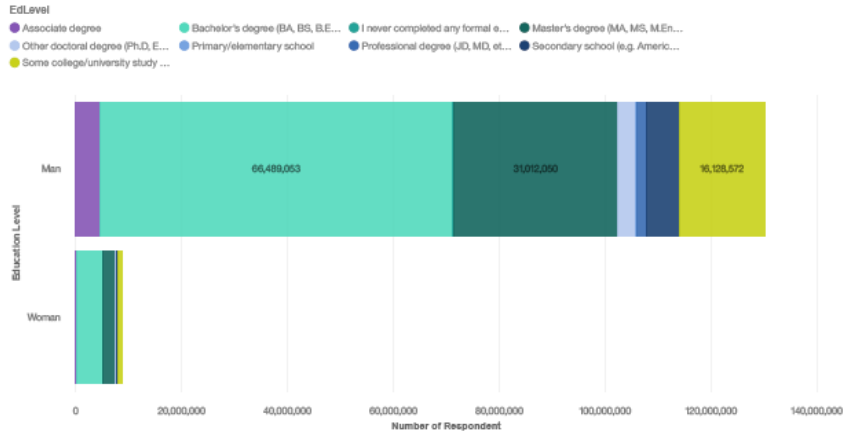
Respondent Country Demographics



Respondent Age Demographics



Respondent Education Level Demographics



DISCUSSION

- Technology trends now and future
- Training and reskilling workers
- Female participation in technology field
- Bridge divide of technology gaps in developing countries
- Eliminate age and education discrimination in employment

OVERALL FINDINGS & IMPLICATIONS

Findings

- Fast changing technology every year
- Concentration on countries like USA and India
- Gender gap in technology jobs
- Platforms like Dockers and AWS are growing

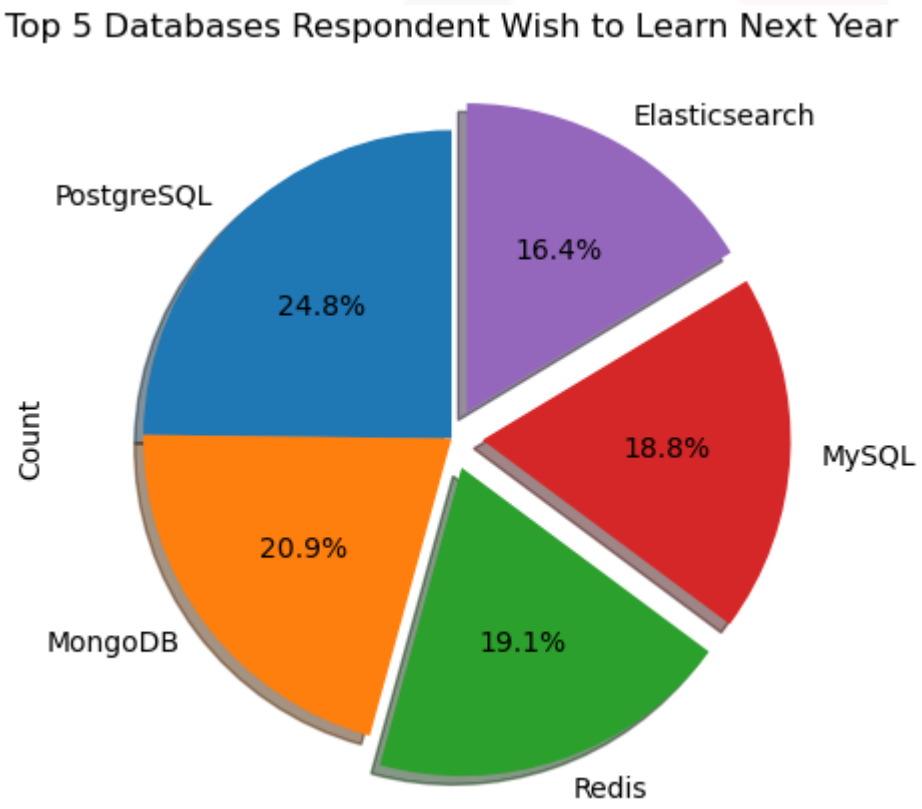
Implications

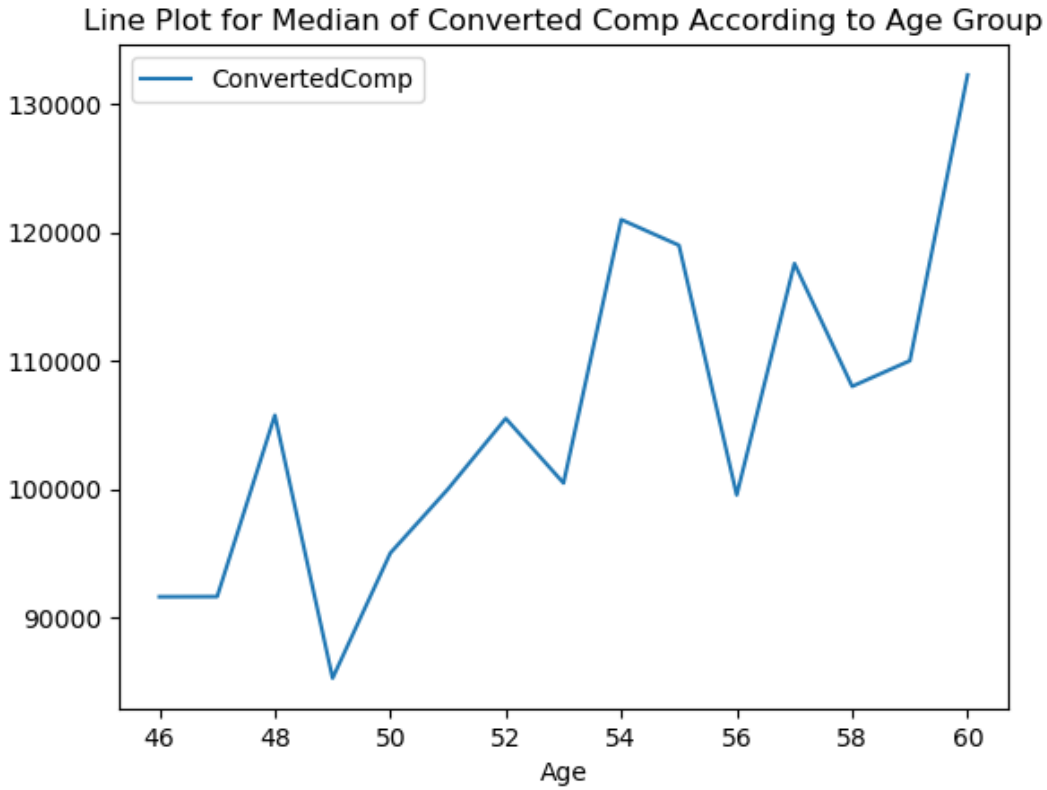
- Company need to be flexible and adaptable to rapid changes
- Spreading out of technology to different countries
- Technology need to be promoted to female
- Shift to faster app deployments and cloud services in future

CONCLUSION

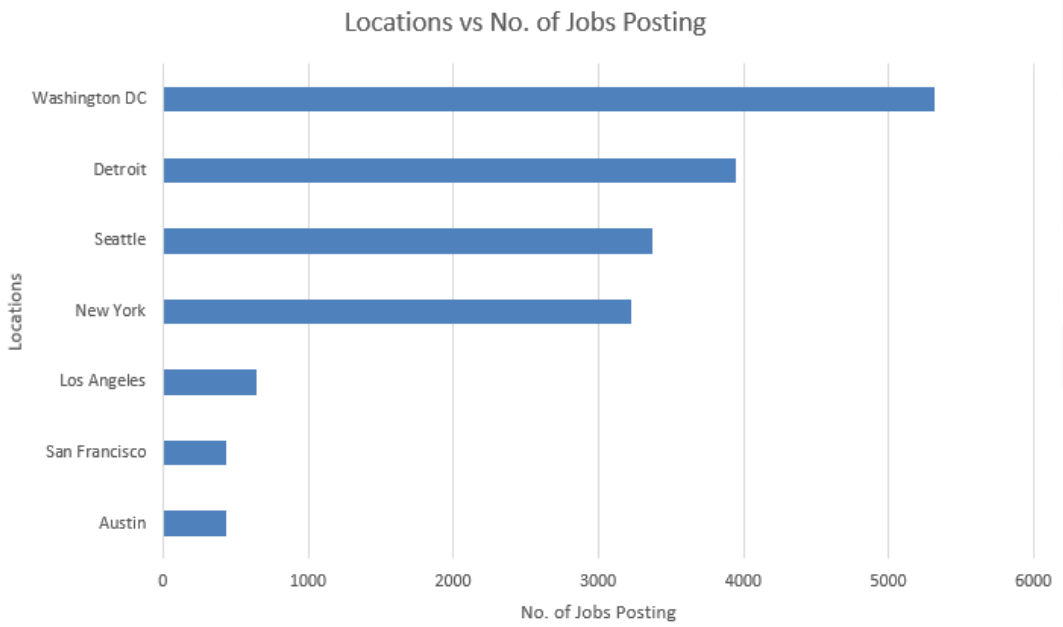
- Technology trend in the coming year
- Prominent gender gap within technology job
- Concentration of technology job in certain countries such as USA and India
- Clouds becoming increasingly more popular in the coming years
- AI & Machine learning is a preferred method for data analyst in the coming year

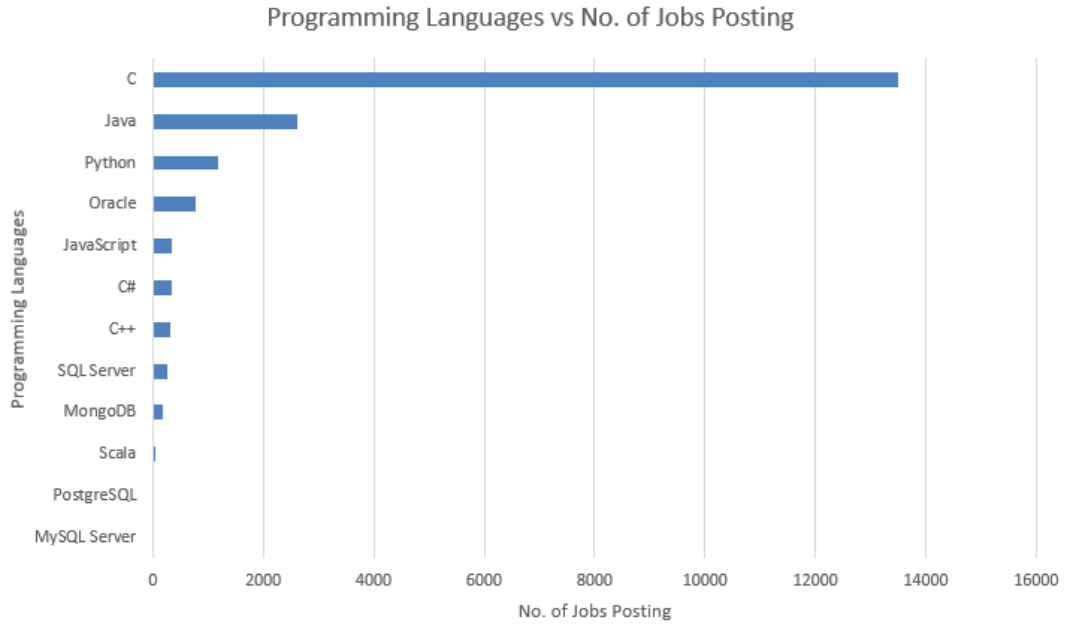
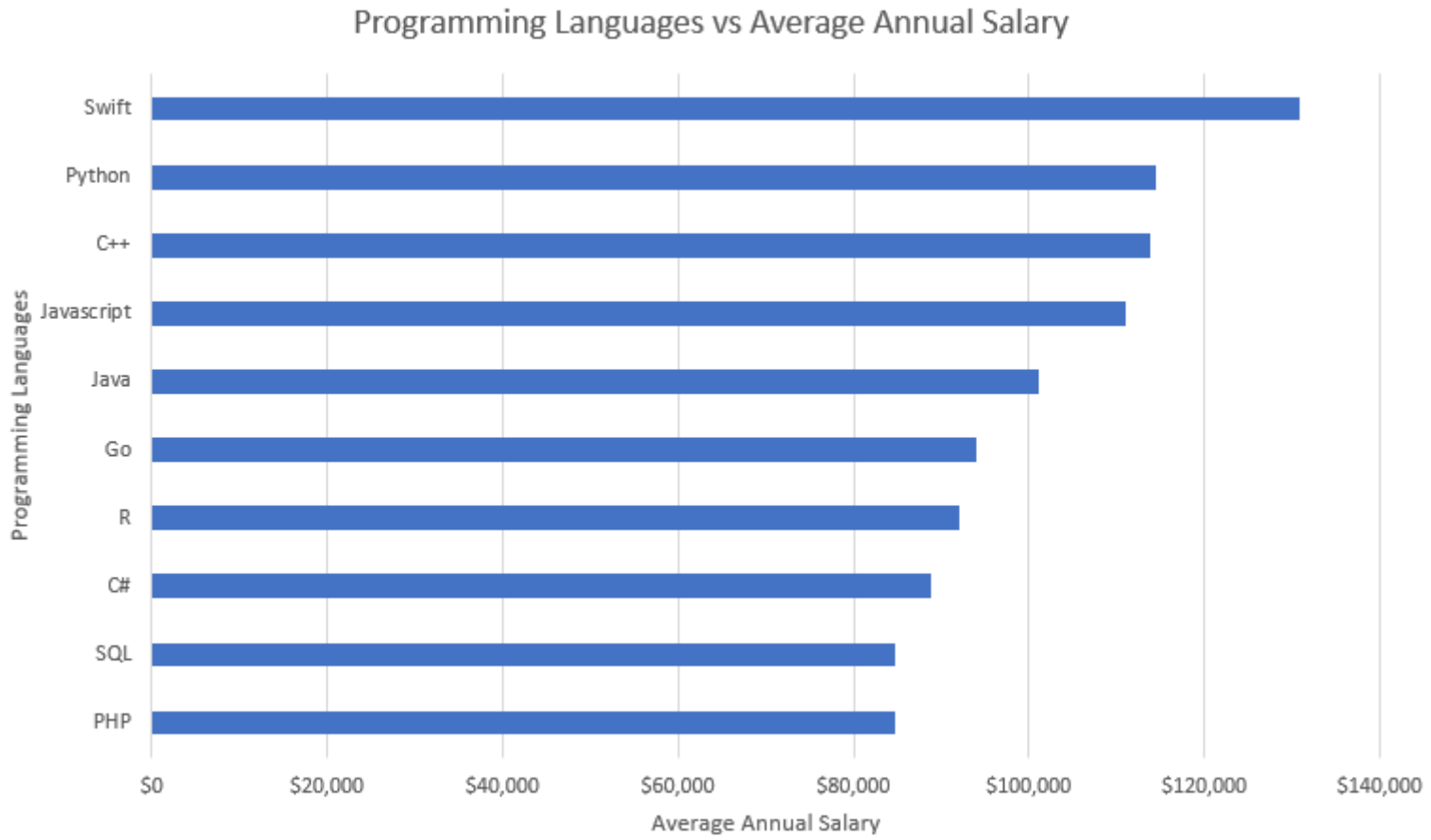
APPENDIX


Top 5 Databases Respondent Wish to Learn Next Year
Elasticsearch
PostgreSQL
16.4%
24.8%
18.8%
MySQL
Count
20.9%
MongoDB
19.1%
Redis


Line Plot for Median of Converted Comp According to Age Group
ConvertedComp
130000
120000
110000
100000
90000
46
48
50
52
54
56
58
60
Age

JOB POSTINGS





POPULAR LANGUAGES