Developer Survey Result Analysis

Rishabh Gupta

20th Feb, 2021.

OUTLINE

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

EXECUTIVE SUMMARY

- A modified subset of Stack Overflow data consisting of two data sets based on the demographics and data technologies were analyzed to identify the present and future data technologies trends among people of various different age groups all across the world.
- Dashboard was created having three different tabs i.e. current data technologies ,future data technologies and demographics of respondent .IBM Cognos Dashboard Embedded (CDE) platform was used to analyze and visualize the results.
- From the Current technologies dash board it was clearly visible that
  - Java Script and MySQL were most popular languages and Databased respectively.
  - Windows and jQuery were preferred platform and webframe to work with.
- From the future trend technologies dash board it was clearly visible that
  - HTML/CSS and Elasticsearch languages and Databased respectively were most popularly desired to be worked for the next upcoming year
  - Linux and react.js were preferred platform and webframe to be worked for the next upcoming year .
- Lastly, Demographics analysis was done in the third tab and it showed that:
  - 93.7% respondent were male with majority of them being from USA and India in the age group of 21 to 43.
  - Also majority of the male population had a masters or bachelor's degree while completing this survey.
- From different visualization it was clear that data technologies filed are preferred by students pursuing bachelors and master’s degree. Seeing the popularity among youths we may see many bootcamps being organized by different universities and private entities to enhance the skillset if needed.
- Also , we may observe various MNC’s may start to diversify themselves into this field to stay competitive and entice the youth toward them.

INTRODUCTION

- To Visualize the data ,dashboards were created to emphasis on the present and future trends of data technologies.
- Problem: From such a clustered data ,we need to analyze the present data technologies and future data trends considering the various different demographics of the respondents.
- Through creating dashboards using IBM Cognos Dashboard, we would be able to visualize our data based on top 10 present and future data technology trends based on respondent gender ,formal education ,age and countries.

METHODOLOGY

- A modified subset of Stack Overflow data consisting of two data sets based on the demographics and data technologies were used to visualize the present and future data trends based on their demographics.
- As these data sets were subsets of data sets and therefore they did not require a lot of cleaning . Hence no particular methodology was used to visualize the data. We only used filtering option to ignore the nulls values from data.
- IBM Cognos dashboard was used to create different Dashboards.

RESULTS

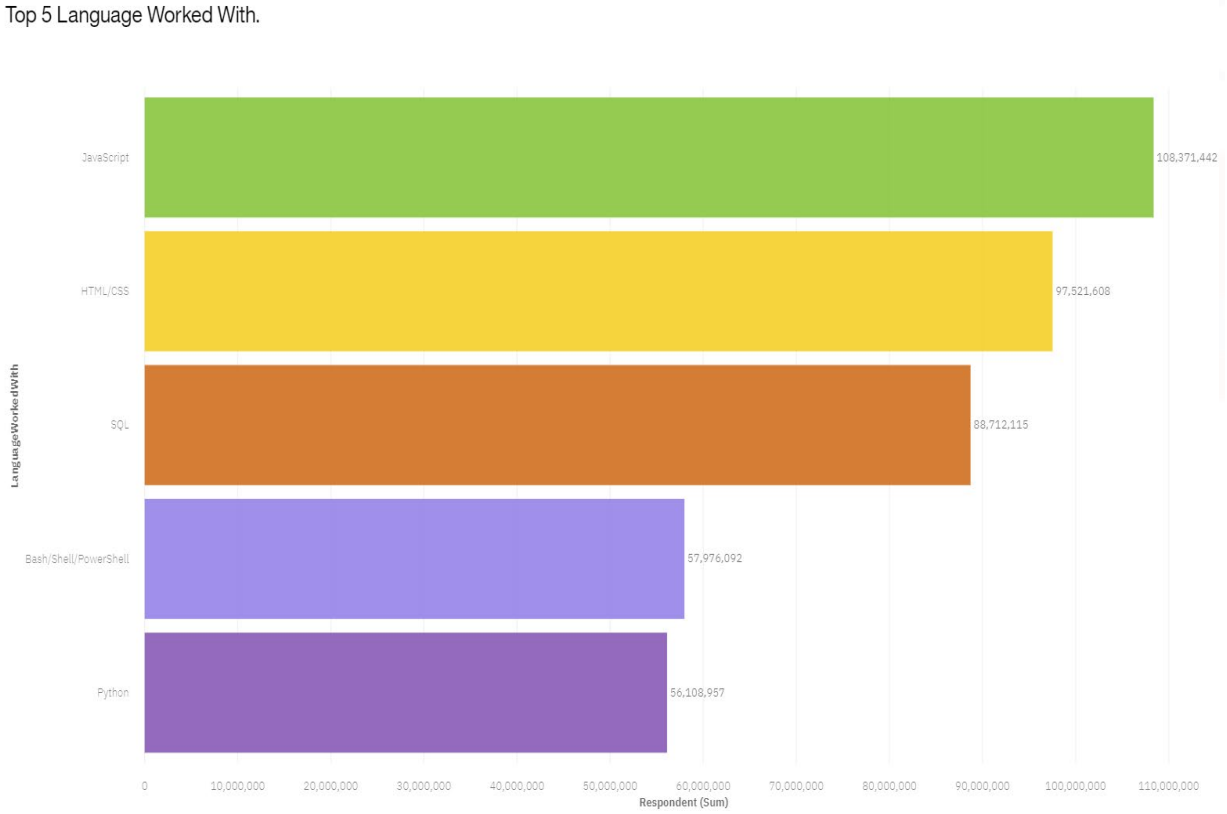
- The data was collected from stack overflow blog under ODbl: Open Database License .It was further modified and use for the analysis purpose.

- It consisted of two data sets :
  - Survey data technology normalized consisted of approximately 75K respondents querying them about the present and future data technologies interest . Then using visualization tools such as bar, columns graphs along with tree maps and hierarchy trees were used to analysis the top trending present and future technologies trends.
  - Survey data demographics was more clustered data set a which involved around 11K respondents answering about various demographics questions such as country , age , education level, ethnicity, gender etc. The results of these finding were visualized in form of bar ,line charts along with maps.

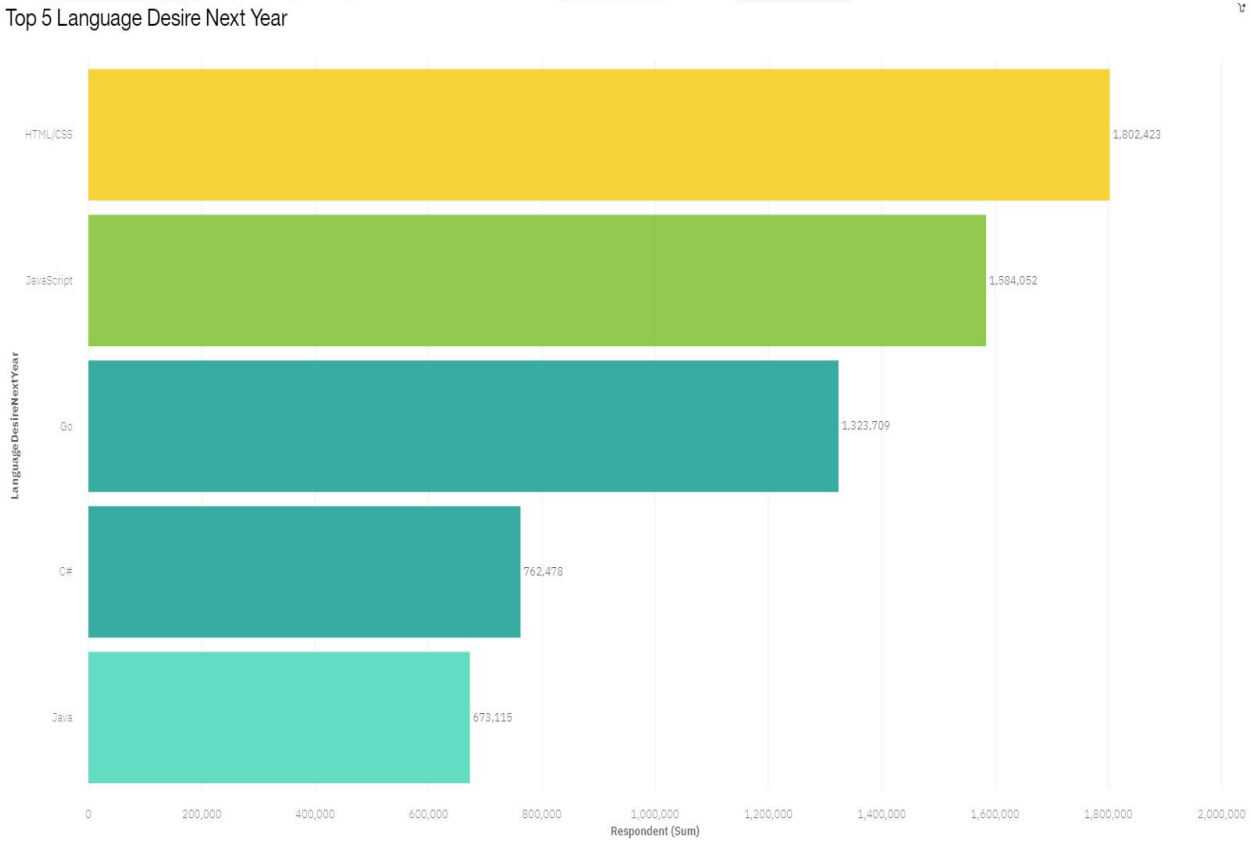
- In next few slides we will be discussing the crucial findings from the data sets for present and future technologies trends

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



IBM Developer

SKILLS NETWORK

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

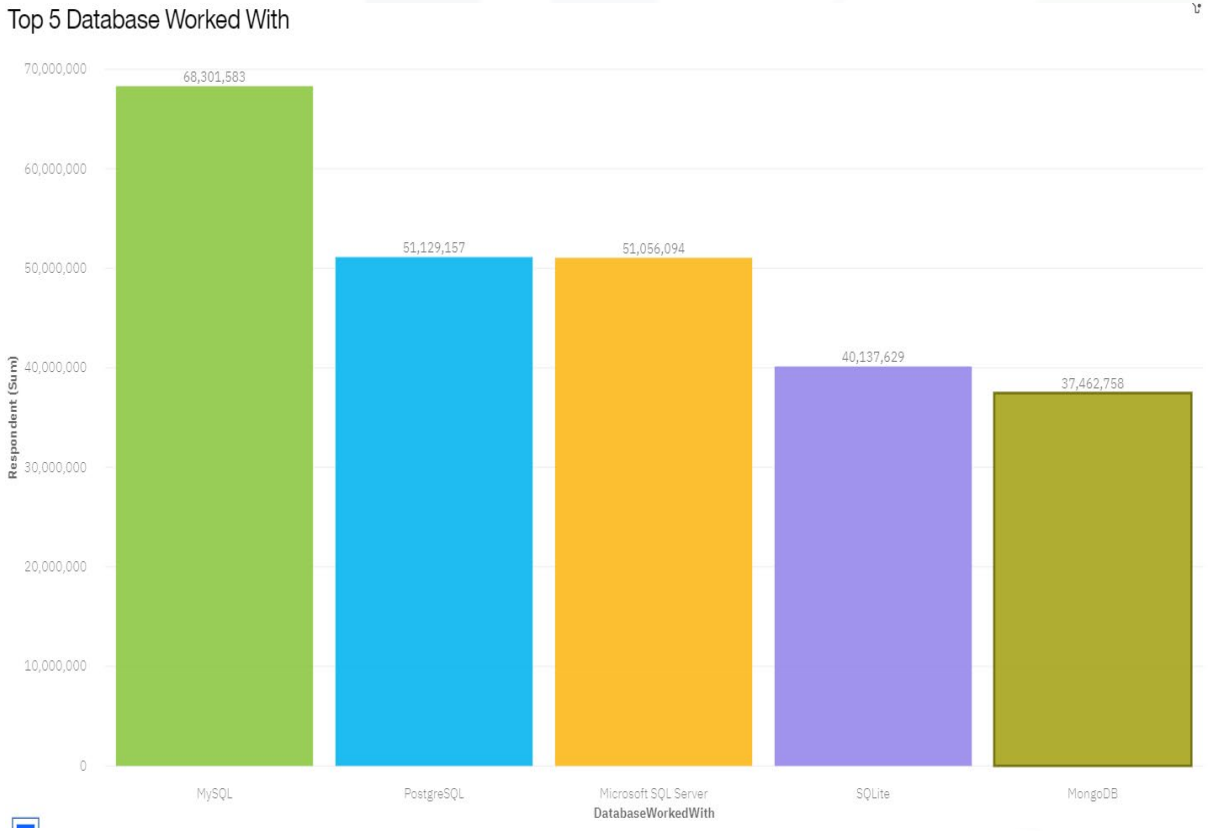
- JavaScript and HTML were the languages that were preferred and will continue to dominate in future too.
- They were further followed by SQL, Bash, and Python for current trends but they may not be the next go to languages in future.

Implications

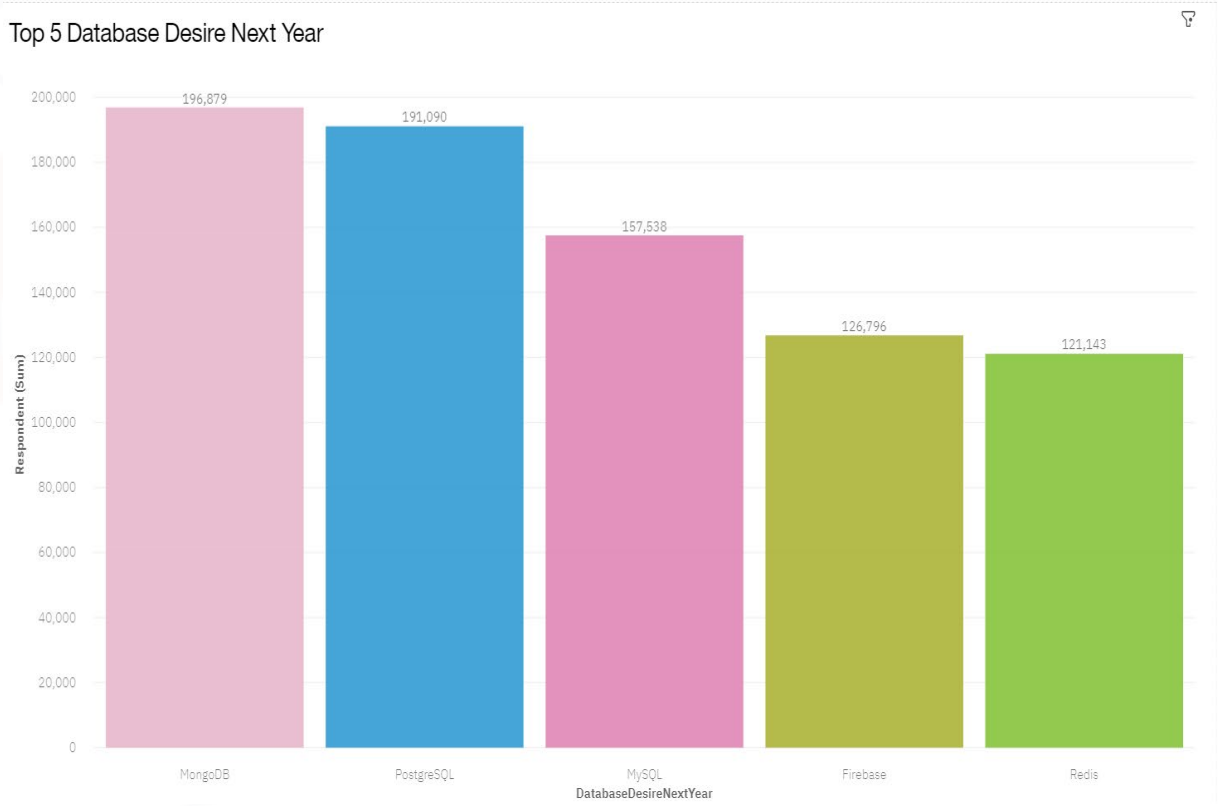
- The job market would see a higher influx of JavaScript and HTML engineer.
- This may decrease the present high demand of such developer and we may see a saturation point in coming years.

DATABASE TRENDS

Current Year



Next Year



IBM Developer

SKILLS NETWORK

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- My SQL dominates the Database platform preferred by developers but in next few years we may see a steep decline in its preferences.
- PostgreSQL was the next database platform preferred by developer and see a steep increase in its preference to be worked on in future too.
- MongoDB would be the database platform which would be most preferred Database platform to be preferred in future.

Implications

- We may see more organization preferring to work with MongoDB and PostgreSQL, as the popularity of them may increase in coming years
- We may see MongoDB and PostgreSQL as the next essential databases skills to have in resume
- MySQL may reach a saturation point and wouldn’t be sufficient enough to give you an edge over others in term of skills.

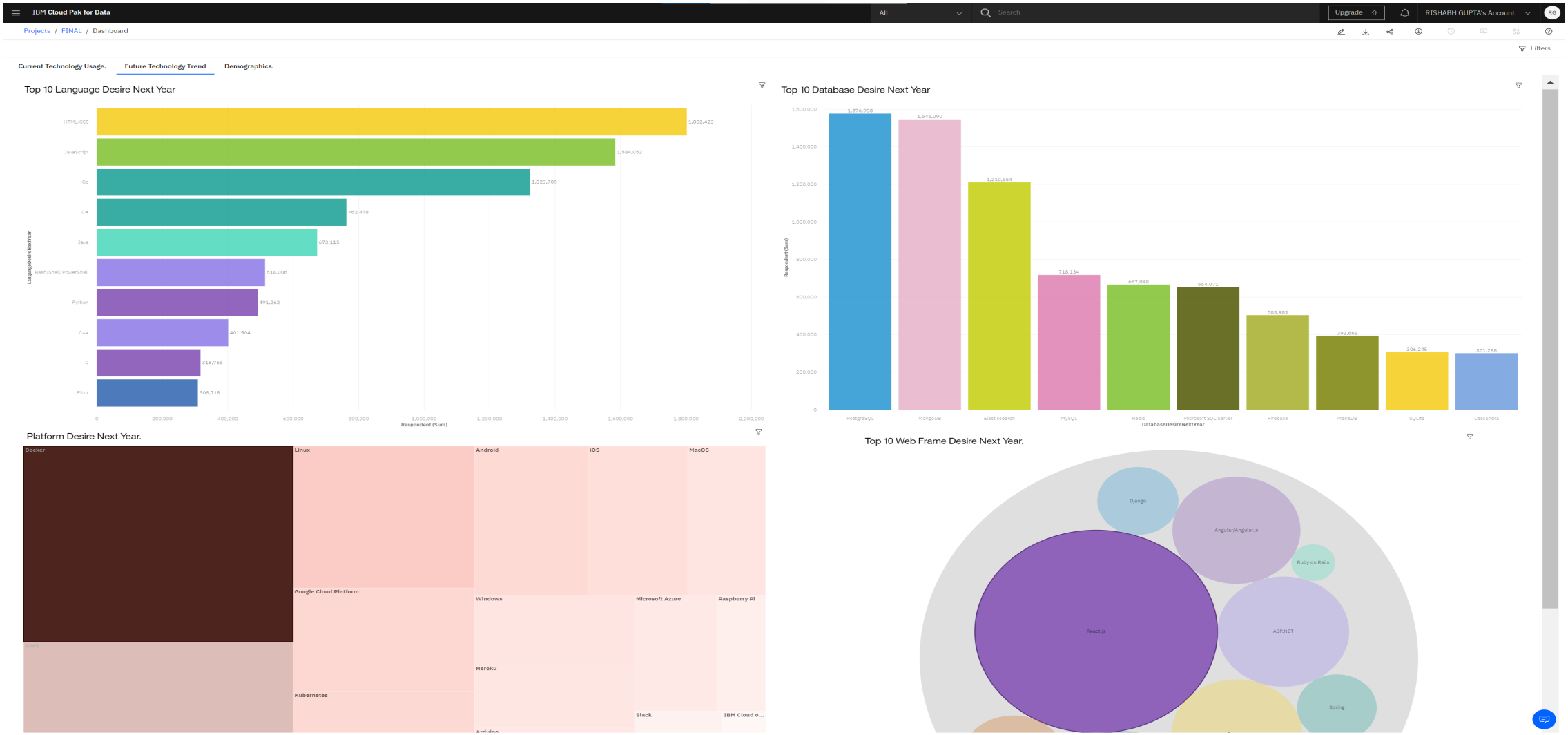
DASHBOARD

Link:

DASH BOARD

DASHBOARD TAB 1

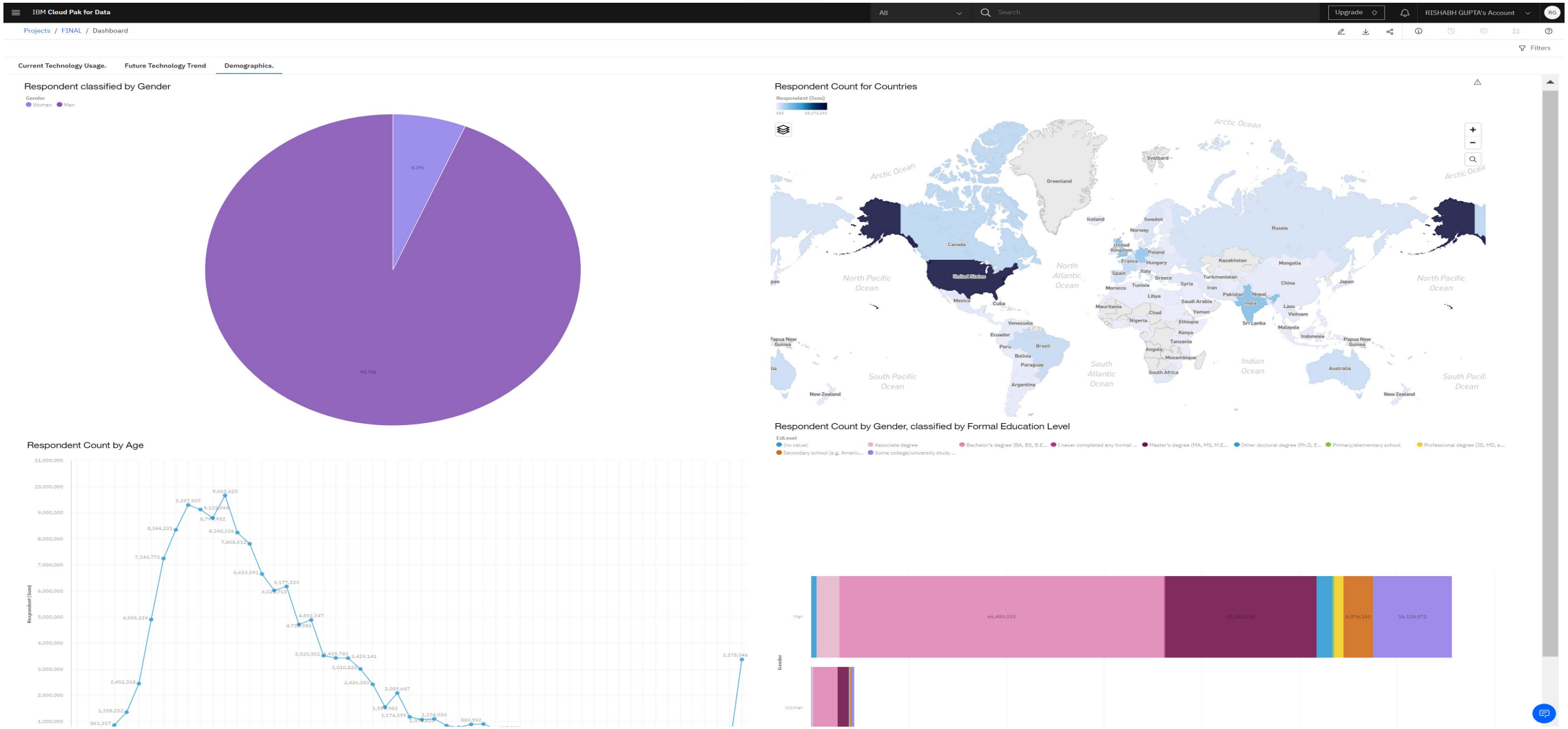
DASHBOARD TAB 2



IBM Developer

SKILLS NETWORK

DASHBOARD TAB 3



DISCUSSION

- Languages:
  - Will SQL and Python tend to loose it dominance in upcoming years?
  - Will other languages such as Go, C# would be the new emerging languages of future?
- Database:
  - Would MongoBD and PostgreSQL would be the new MYSQL in emerging times
  - Why is the popularity of MySQL declining so drastically?

OVERALL FINDINGS & IMPLICATIONS

Findings

- The database technologies surveys results are dominated by male population, majority of them being USA respondents.
- Bachelor's and Master’s students highly prefer doing specialization in data technology.

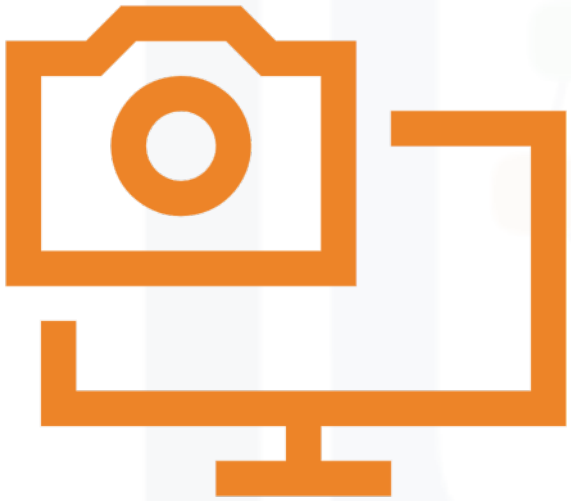
Implications

- Seeing the popularity amongst youths, we may see more bootstrap camps being set up to enhance their skillsets.
- The popularity trends may be similar in other countries too and we may see various MNC’s diversifying into this sector.

CONCLUSION

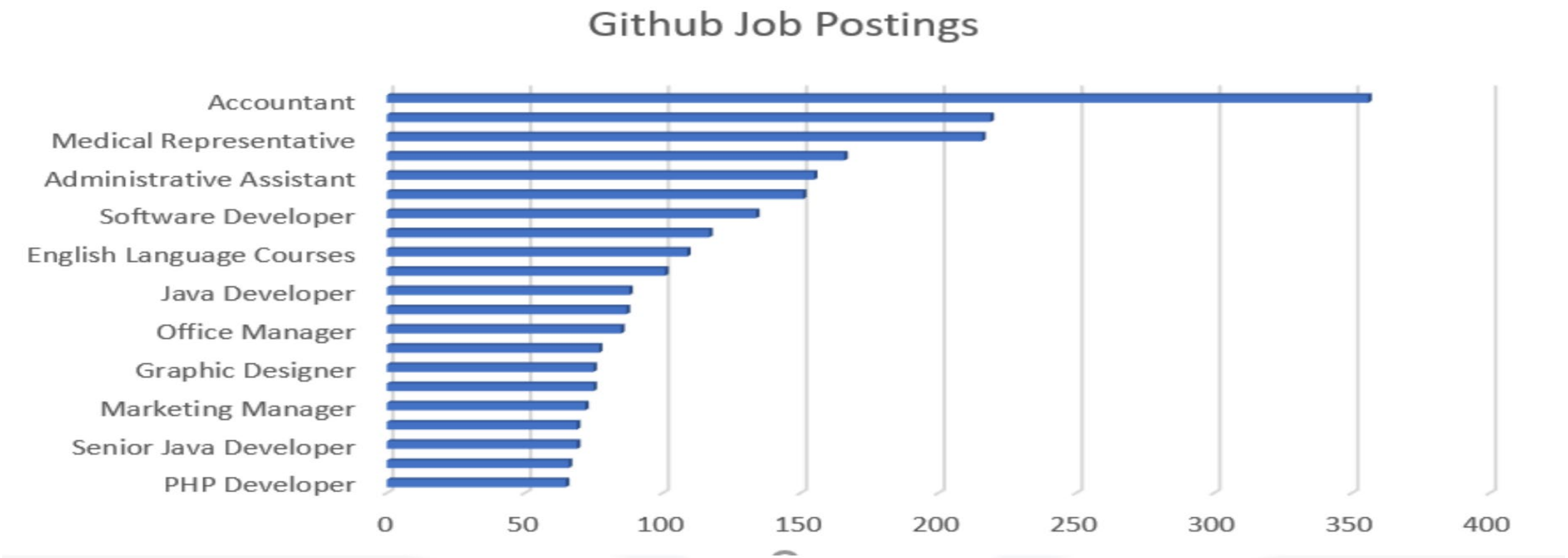
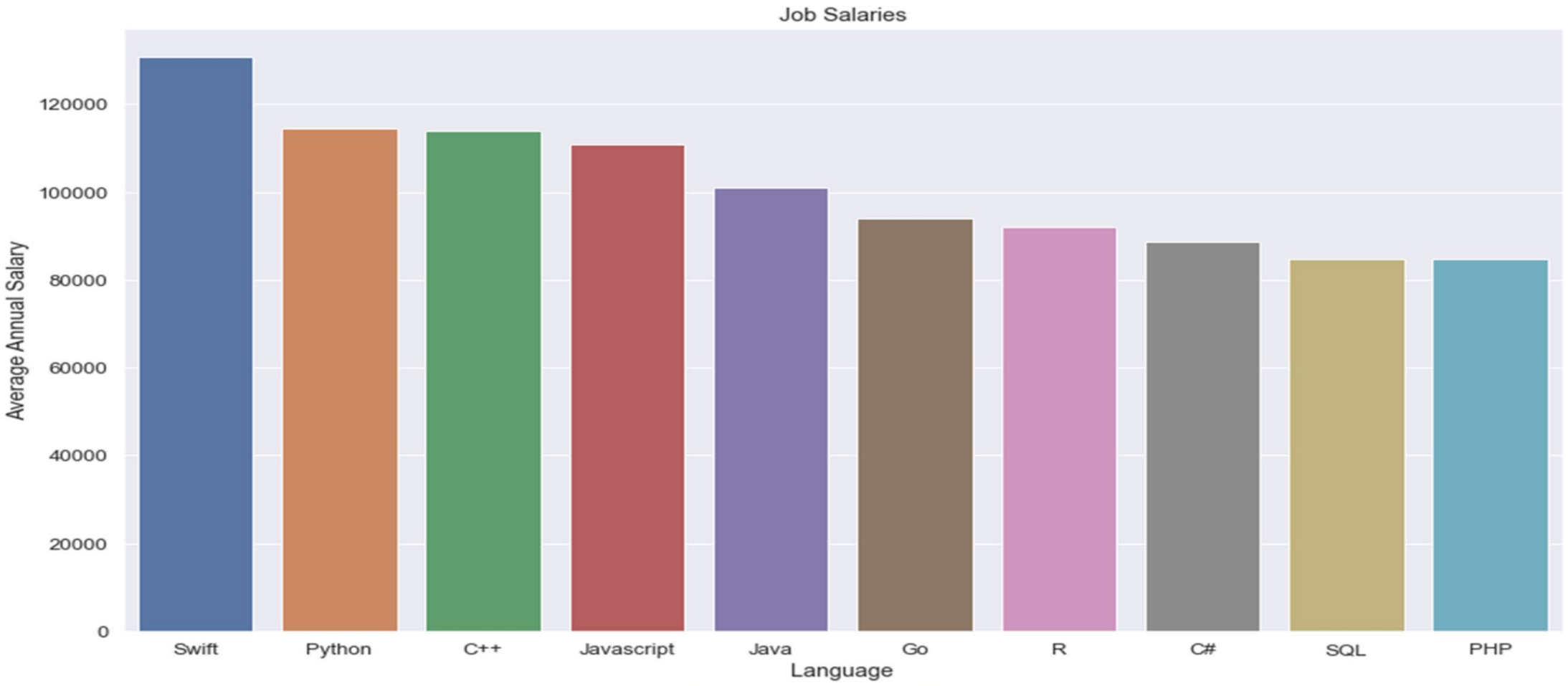
- We wanted to analyze the dataset for the preferred current and future trends in data technology while also considering the demographics of the respondent.
- From the data visualization it is clearly visible that the data technologies trends are preferred by population b/w 21 to 36 years
- It has a wider acceptance among male as compared to women.
- Among the languages , HTML and Javascript will remain the most preferred languages to be worked with.
- Among the Database, we may see the popularity of SQL decreasing and which may be replaced by MongoDB and PostgreSQL.
- We may see an increase in number of bootcamps for enhancing skills and may come across many MNC's diversifying into data technology to keep up the pace. Further analysis on same can be done next to validate this.

APPENDIX



- Resources used:
  - Database taken from:Stack overflow blog under ODbl: Open Database License
  - IBM Cognos dashboard were used for creating Dashboards
- Acknowledgement:
  - Coursera for allowing the access to this course

GITHUB JOB POSTING



POPULAR LANGUAGES