



Developer Survey Analysis

Afnan Ahmed Yaqub

02/11/2022

OUTLINE

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



EXECUTIVE SUMMARY

- ▶ **Data Collection**
 - ▶ Accessing APIs
 - ▶ Web Scraping
 - ▶ Exploring Data
- ▶ **Data Wrangling**
 - ▶ Dealing with Duplicates
 - ▶ Dealing with Missing Values
 - ▶ Normalizing Data
- ▶ **Exploratory Data Analysis**
 - ▶ Analyzing Data Distribution
 - ▶ Handling Outliers
 - ▶ Investigating Correlations
- ▶ **Visualizing Data**
 - ▶ Distribution of Data
 - ▶ Relationships in Data
 - ▶ Composition of Data
 - ▶ Comparisons in Data
- ▶ **Building Dashboards**
 - ▶ Cognos Analytics



INTRODUCTION

- This presentation covers the insights achieved through analysis of the **2019 Stack Overflow Developer Survey** dataset.
- The purpose of this analysis is to understand the current supply and demand for developers in the following areas,
 - Programming Languages
 - Databases
 - Web Frameworks
- The survey has collected current and future technology interests of the developers.
- The survey also acquired data regarding the developer related demographics.
- This is an analysis of the demand in the current job market for programming and the developer supply all over the world.

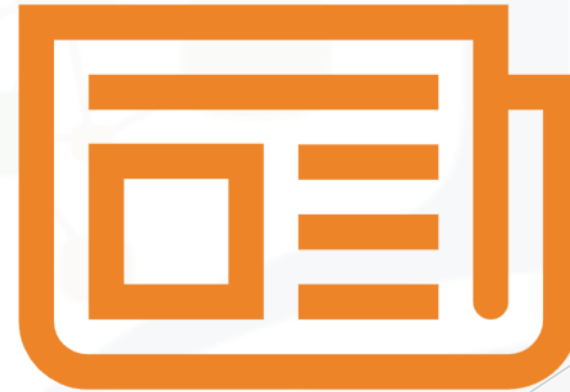


METHODOLOGY

Jobs Market Data (Demand)

► Data Collection

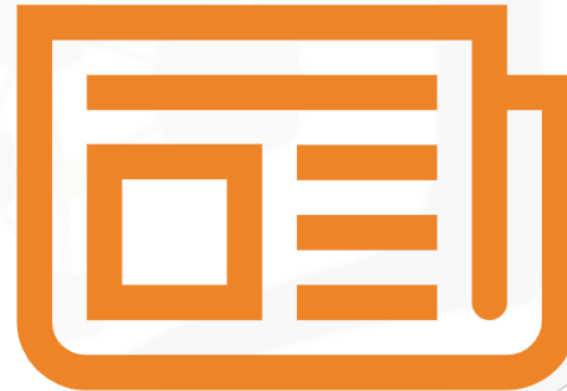
- Using APIs to collect jobs data for programming languages and location.
 - Python libraries: pandas, requests, BeautifulSoup
- Using Web Scraping to collect data on programming languages and average annual compensation.
 - python libraries: pandas
- Saving data to CSV file.



METHODOLOGY

Developer Data (Supply)

- ▶ Data collection
 - ▶ Original Dataset (CSV)
 - ▶ [Stack Overflow Developer Survey 2019](#)
 - ▶ IBM provided the following subset of the original dataset (CSV)
 - ▶ [Survey Data](#)
- ▶ Data Wrangling (python, pandas)
 - ▶ Identifying and imputing missing values
 - ▶ Finding and removing duplicates
 - ▶ Normalizing
- ▶ Exploratory Data Analysis
 - ▶ Exploring distribution
 - ▶ Exploring Outliers
 - ▶ Exploring Correlation
- ▶ Visualizing Data
 - ▶ Using python pandas plotting functions
 - ▶ Using Cognos Analytics for creating dashboard for developer data and demographics

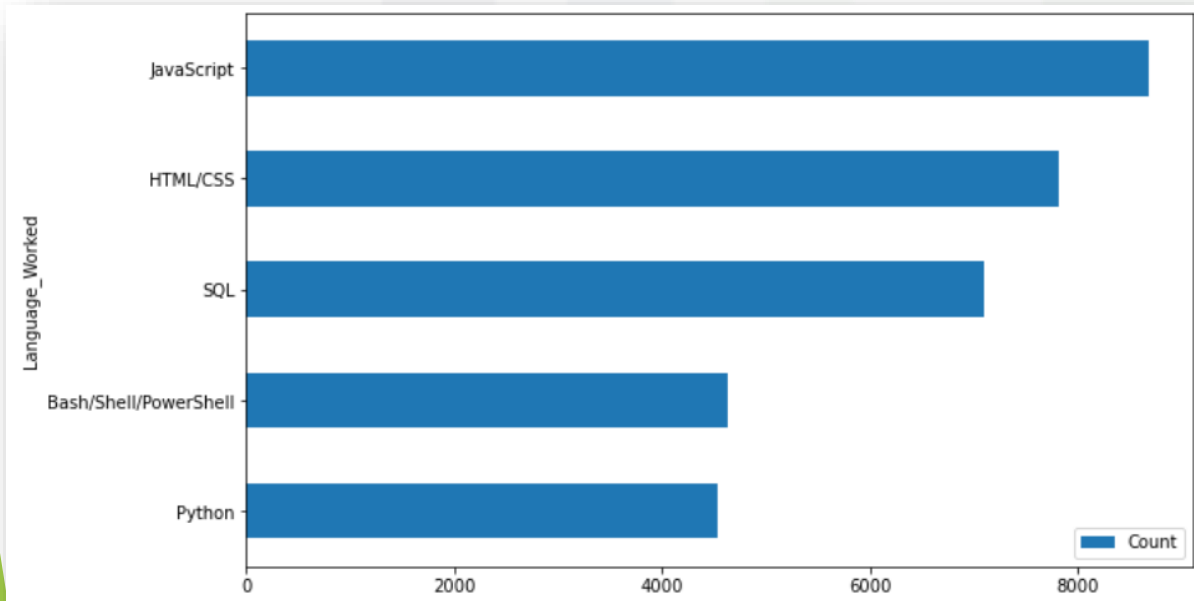


RESULTS

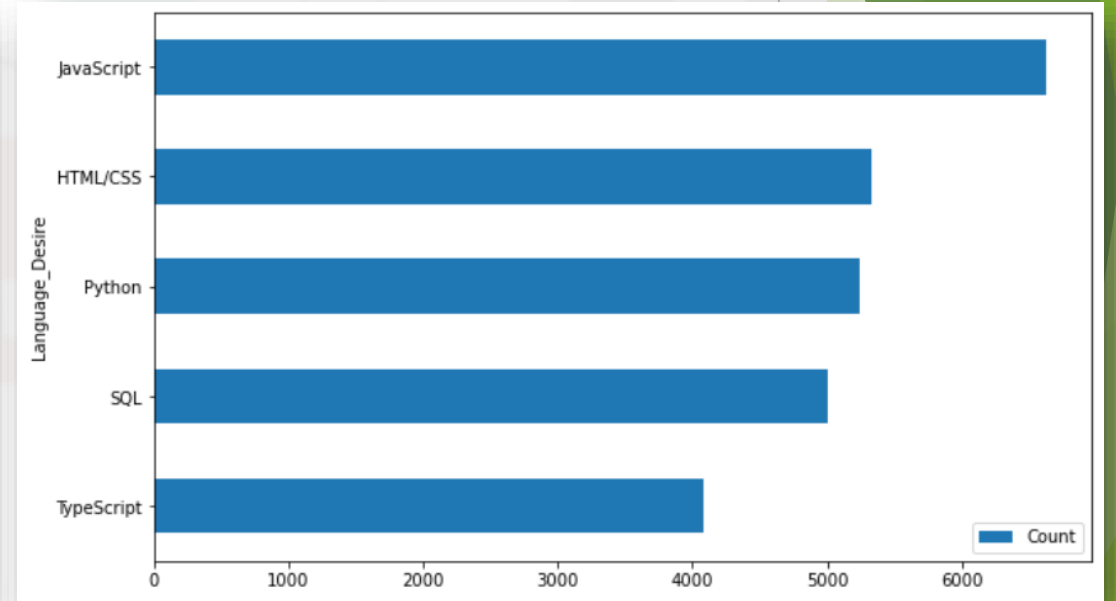
- In the following slides the results for analysis for popular programming languages in the current and next year are shared.
- The databases that programmers are mostly working in and those sought after in the next year are shown.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

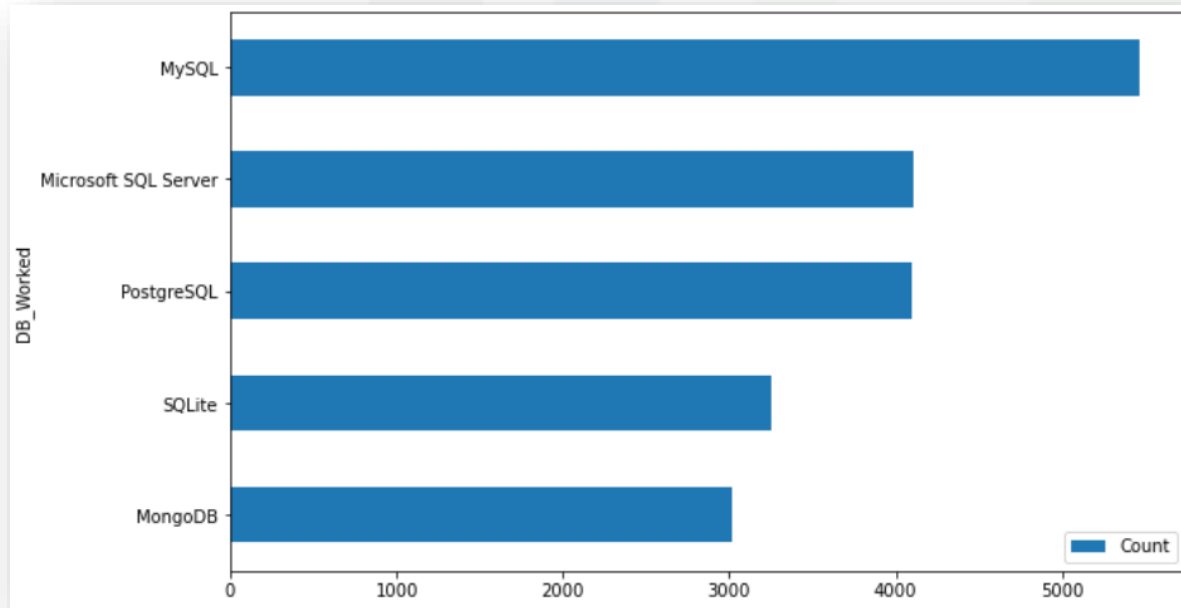
- ▶ JavaScript is and will remain the most popular programming language through the next year. HTML/CSS follow as second most popular language.
- ▶ Python and TypeScript are gaining popularity as more programmers are interested in learning these two in the next year.
- ▶ SQL is and will continue to remain popular through the next year.
- ▶ Interest is declining in Bash/Shell/PowerShell

Implications

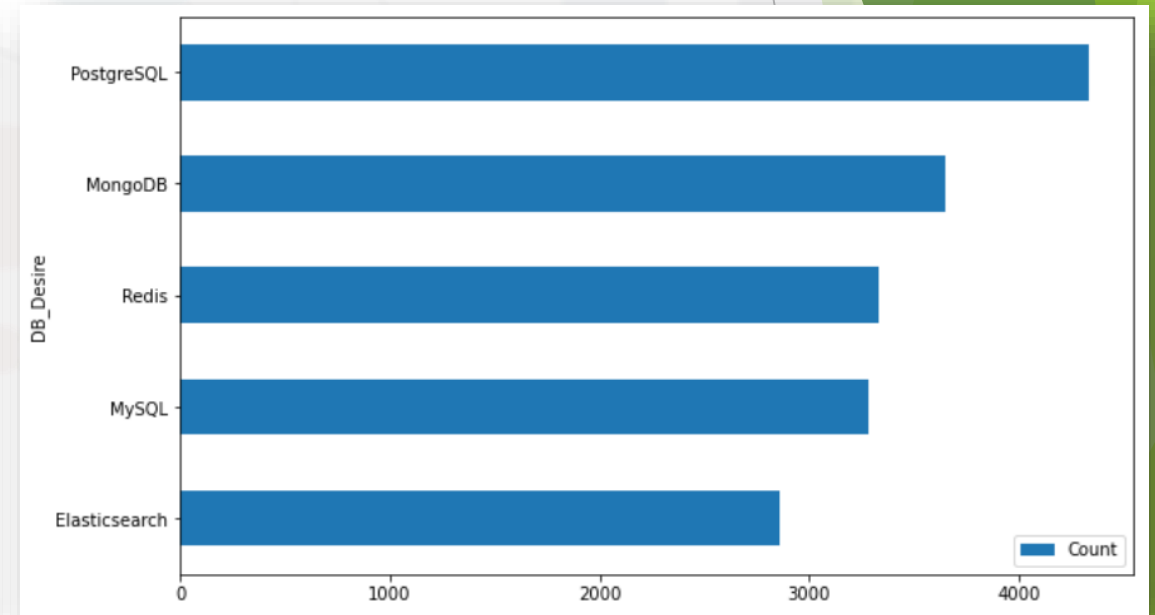
- ▶ Web Development is in high demand as dictated by the current and future trends. Requirement for support and educational resources must be adequately fulfilled.
- ▶ Python and TypeScript development and learning resources to be provided adequately.
- ▶ SQL will remain the language of choice for big data query.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- ▶ MySQL is currently the most popular database amongst programmers in 2019.
- ▶ PostgreSQL will be the most popular next year. MongoDB follows.
- ▶ Elasticsearch gaining popularity as well

Implications

- ▶ Open source database programming is gaining momentum and becoming widely popular.
- ▶ Database programming remains in high demand.
- ▶ Increase in requirements for handling unstructured data such as email, videos and image data.
- ▶ Analytics related search engines gaining popularity.

DASHBOARD

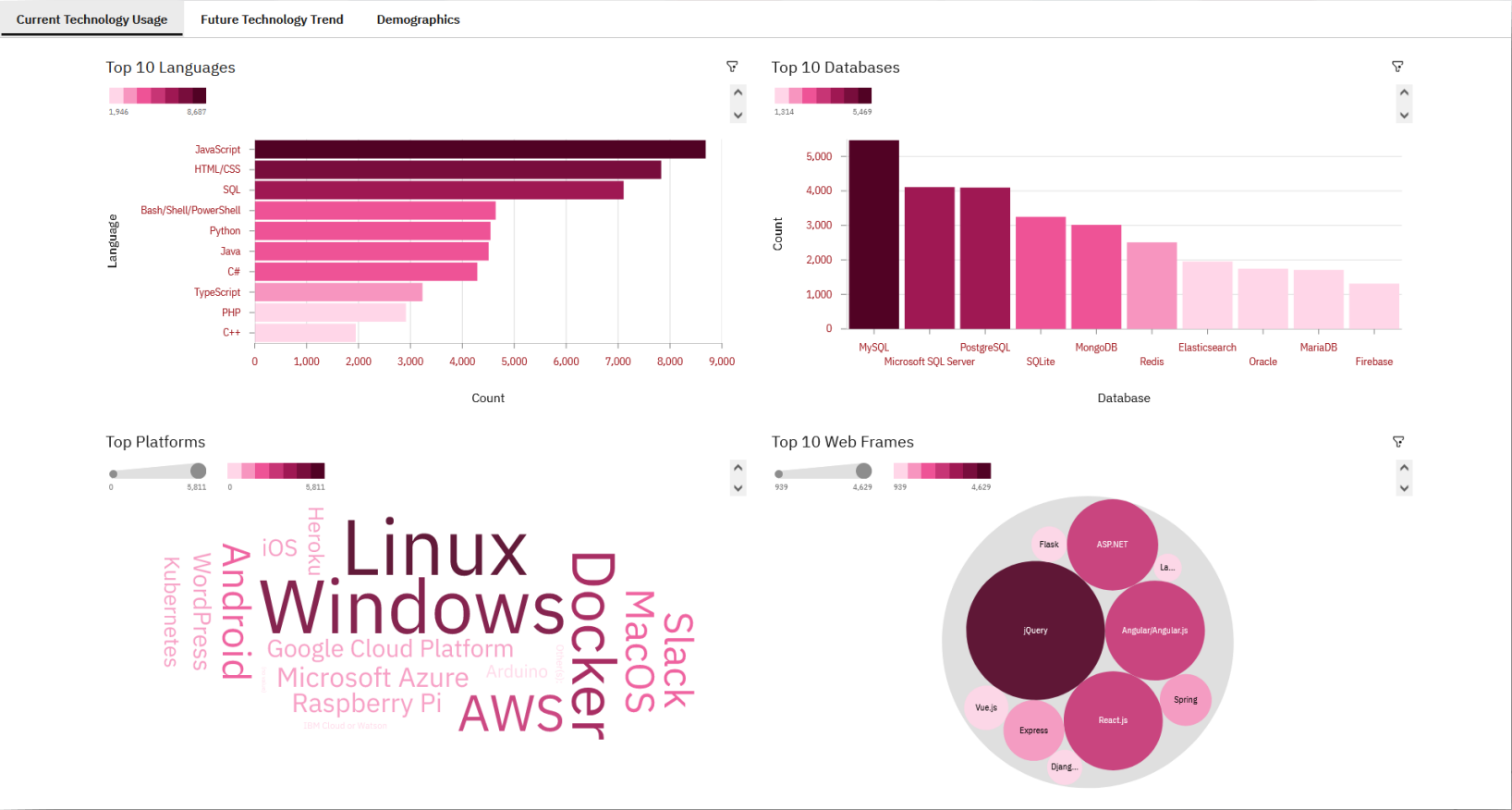
The analysis has been presented as a dashboard using Cognos Analytics platform on IBM Cloud.

Click the link below to access.

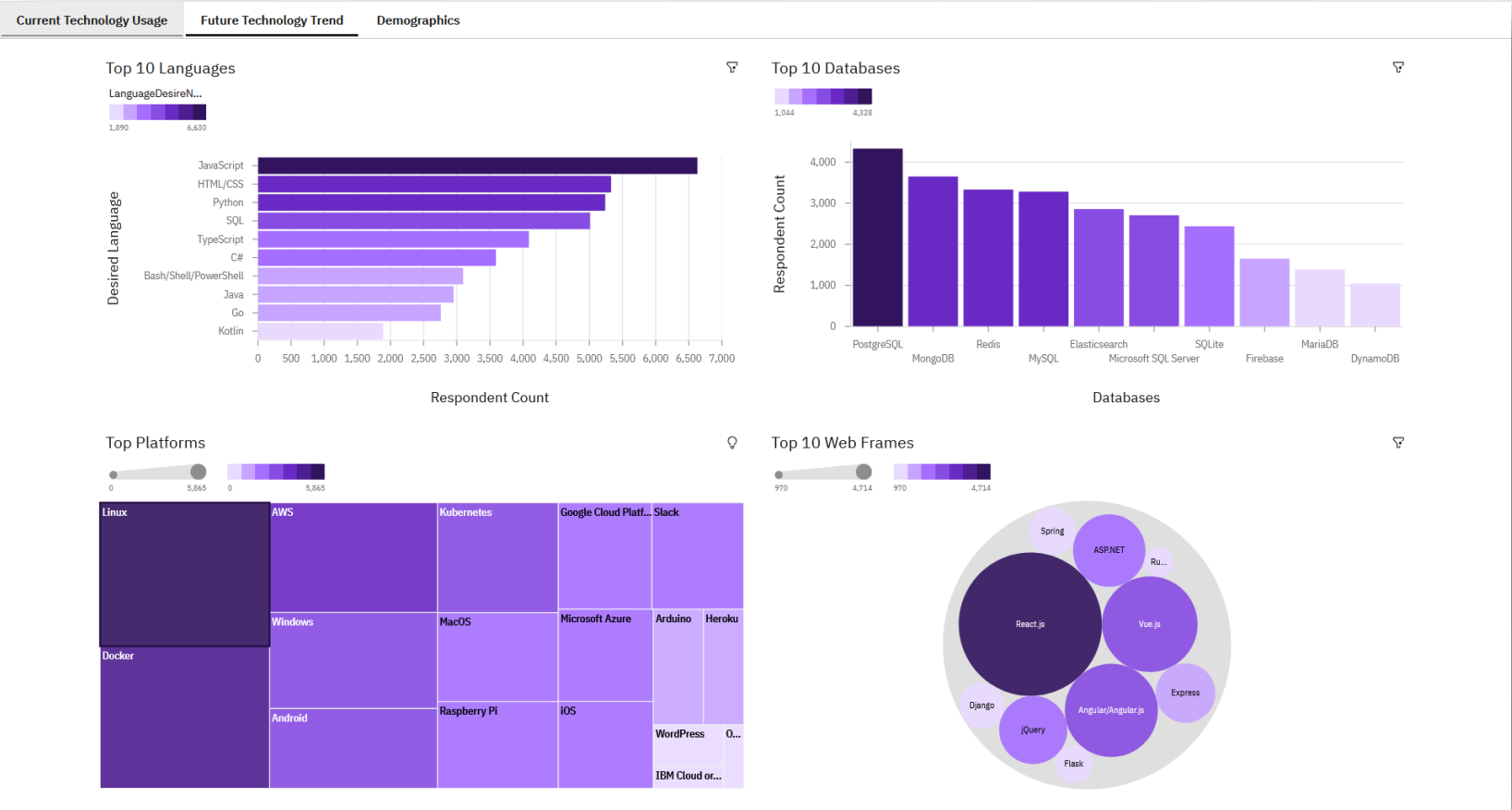
[Developer Survey Dashboard](#)



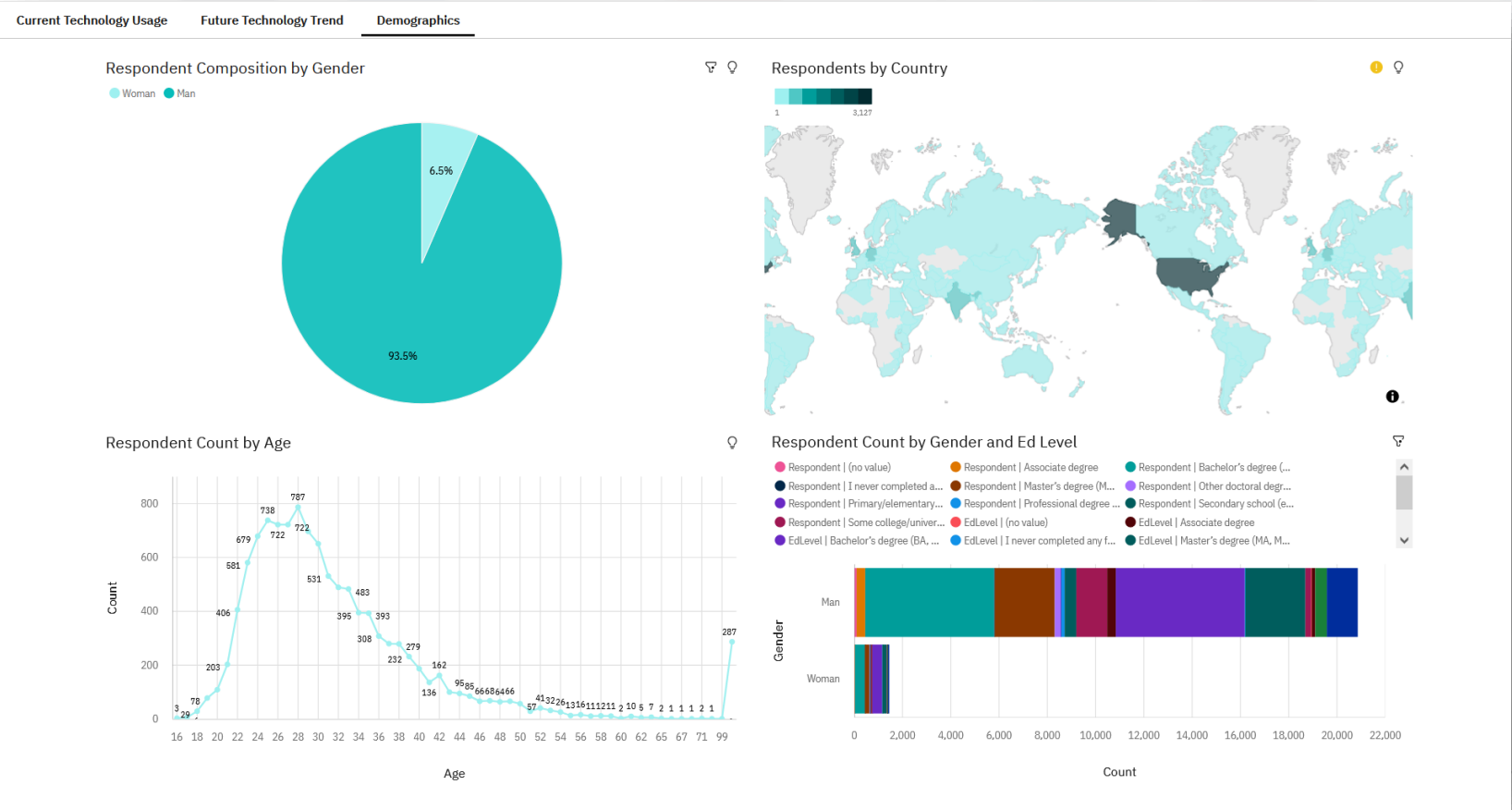
DASHBOARD TAB 1 - Current Technology Usage



DASHBOARD TAB 2 - Future Technology Trend



DASHBOARD TAB 3 - Demographics



DISCUSSION



OVERALL FINDINGS & IMPLICATIONS

Findings

- ▶ Most popular languages for web development are JavaScript and HTML/CSS. TypeScript also gaining momentum.
- ▶ Python is gaining popularity for scientific computing and analytics.
- ▶ Open source database programming is in high demand.

Implications

- ▶ More business interests in E-Commerce and other online activity.
- ▶ E-learning will be the most sought after learning source.
- ▶ Data analytics and scientific computing will gain momentum as the most important job categories.
- ▶ Handling of unstructured data will be on the rise as more and more people get involved in social networking and online communications.

CONCLUSION

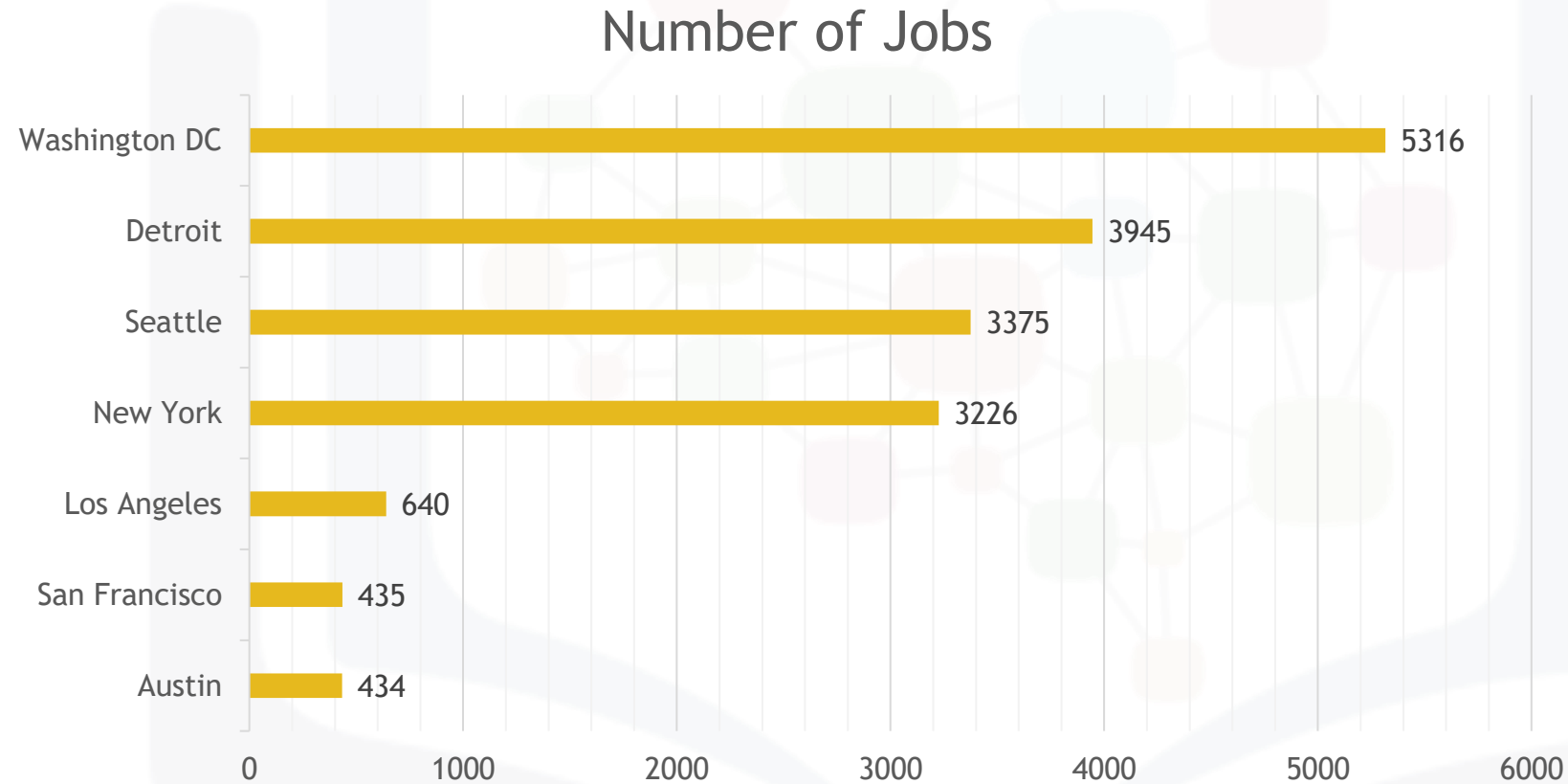
- ▶ With the increase in businesses moving online the requirements for programming resource for web development is going to see new heights.
- ▶ Open source programming platforms have really gained popularity and improvements in platform management has attracted huge programmer interests.
- ▶ As more and more data gathers due to social networking, there is an increasing need for handling data and so has risen the demand for database programming languages and platforms both offline and on cloud.
- ▶ Programmer interest shows that increase in number of online learning platform and improved teaching methods will be required to balance the demand and supply.
- ▶ Online business opportunities are going to increase in future.



APPENDIX



JOB POSTINGS



POPULAR LANGUAGES

Popular Languages with Average Salary

