



IBM Data Analyst Capstone Project

Jorge Costrá

September 05, 2023

OUTLINE



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

EXECUTIVE SUMMARY



- This report presents the top 10 programming languages and top 10 databases for the current year and next year. The data was collected from a modified subset of the 2019 Stack Overflow Developer Survey Results and using IBM Cognos Analytics for data visualization (dashboards).
- The report includes four graphs: Top 10 Language Worked With, Top 10 Language Desired Next Year, Top 10 Database Worked With, and Top 10 Database Desired Next Year. Dashboards visualizations are also presented.
- The report also discusses the most preferred programming languages and database managers. *HTML/CSS*, *Bash/Shell/PowerShell*, *JavaScript*, and *C#* are among the most preferred programming languages. *Go* is gaining popularity and *Python* is becoming increasingly popular (nearly tied with *Java*), which may indicate an increase in programmers working on these languages in the future.
- Relational database managers and NoSQL are used extensively as there is almost an equal use of them. This trend seems to continue as next year they share 50% of the preferences.
- *PostgreSQL* could be the most used relational database manager in the future as it leads the 10 most desired for next year followed closely by *MongoDB* and *Elasticsearch*.

INTRODUCTION



- Programming languages and databases are essential tools for developers who want to create and manage data-driven applications.
- To keep up with the evolving demands and opportunities of the software industry, developers need to update or expand their skills in these domains.
- A set of numerical data can be transformed with IBM Cognos Analytics for analysis and visualization, allowing an understanding of the information in terms of its trends and relationships.
- In this way, the characteristics of an event can be explained or future behaviors can be inferred.
- This paper presents the analysis of the trends resulting from surveying hundreds of developers using the *2019 Stack Overflow Developer Survey Results* and the IBM Cognos Analytics.

METHODOLOGY



- The Dataset used is a modified subset from the 2019 Stack Overflow Developer Survey Results.
(<https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>)
- This dataset was modified by *coursera* for the purpose of the course and two .csv files were selected from it:
 - *set m5_survey_data_demographics.csv*
 - *m5_survey_data_technologies_normalised.csv*

METHODOLOGY

IBM

IBM Cognos Analytics

- Three dashboards were created using IBM Cognos Analytics:
 - *Current Technology Usage.*
 - *Future Technology Trend.*
 - *Demographics.*
- For each dashboard, four graphs were built according to the course instructions.
- The graphs selected for this report are:
 - *Top 10 Language Worked With.*
 - *Top 10 Language Desire Next Year.*
 - *Top 10 Database Worked With.*
 - *Top 10 Database Desire Next Year.*

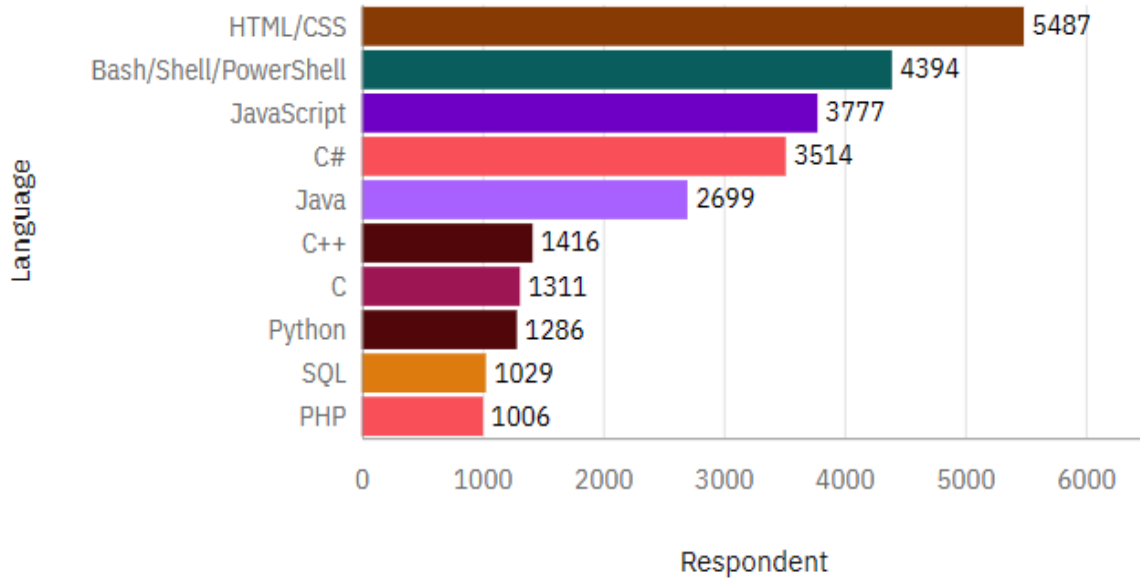
RESULTS

- *HTML/CSS* is the most frequently appearing category in both Language Worked With and Language Desired Next Year.
- *Go* is not in the top 10 of Language Worked With but is ranked fifth as a desired language next year, while *PHP* does not appear. *Python* is almost tied with *Java* as a desired language next year, while *C*, *C++* and *SQL* are at the bottom of the list.
- *MySQL* is the most commonly used database manager for the current year.
- Six relational database managers and four NoSQL database managers are being used for the current year .
- *PostgreSQL* is the most desired database manager for next year, closely followed by *MongoDB*.
- For desired databases next year, five are relational and five are NoSQL.
- *Elasticsearch* ranks third and is above *MySQL* and *MS SQL Server*. It's noteworthy that *Oracle* does not appear in the top 10 desired databases, while *DynamoDB* is in the top 10.

PROGRAMMING LANGUAGE TRENDS

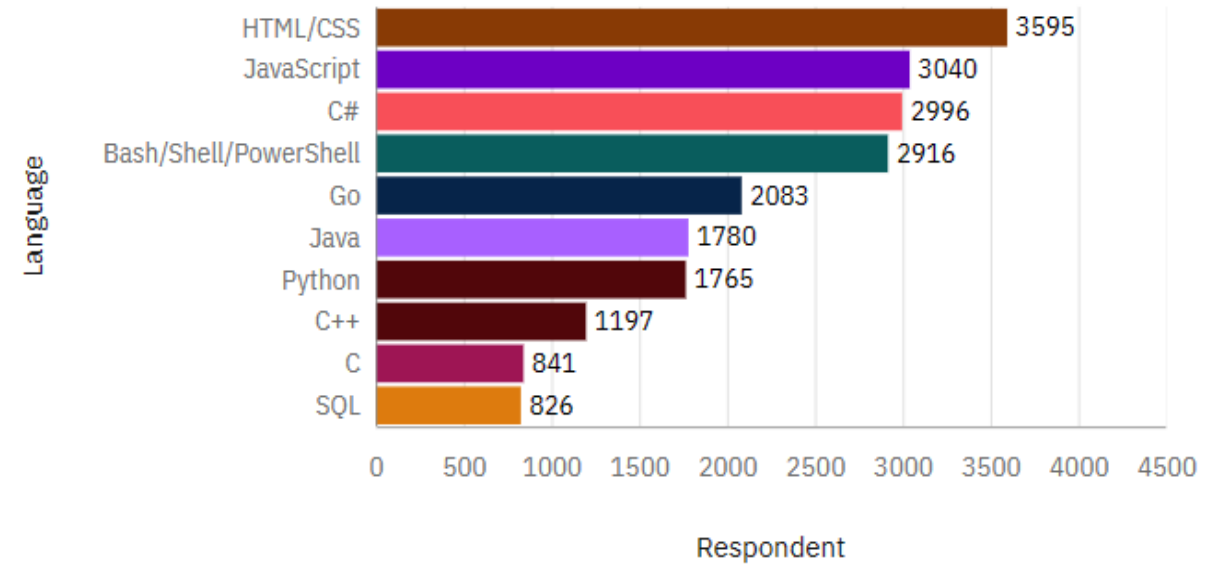
Current Year

Top 10 Language Worked With



Next Year

Top 10 Language Desired Next Year



PROGRAMMING LANGUAGE TRENDS

FINDINGS & IMPLICATIONS

Findings

- *HTML/CSS* is the most frequently appearing category both in Language Worked With and Language Desired Next Year.
- *HTML/CSS, Bash/Shell/PowerShell, JavaScript* and *C#* are on 4 top both in Language Worked With and Language Desired Next Year.
- *Go* is not in the 10 Top of Language Worked With but is on the fifth position as desired language next year, meanwhile *PHP* does not appear.
- *Python* is almost tied with *Java* as desired language next year, meanwhile *C, C++* and *SQL* are in the tail of the list.

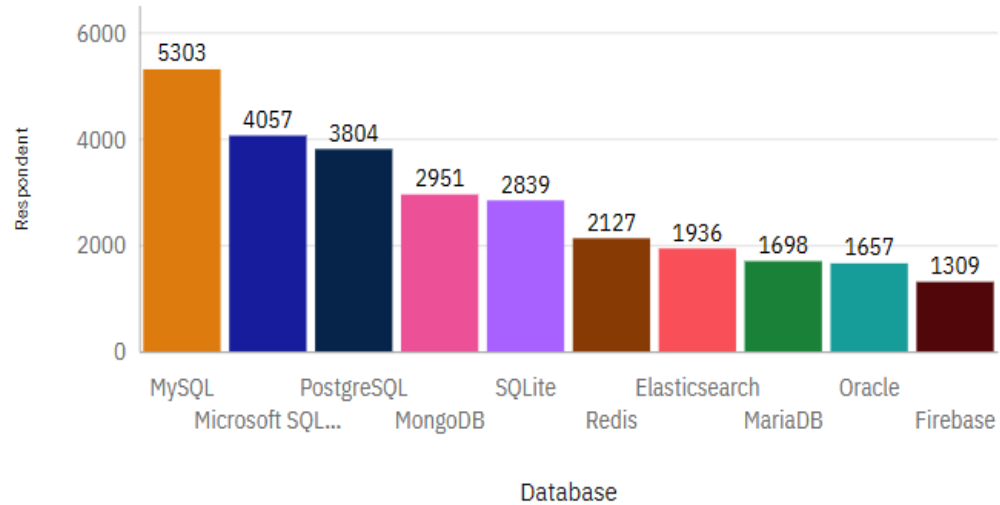
Implications

- *HTML/CSS, Bash/Shell/PowerShell, JavaScript*, and *C#* hold the highest preferences as programming languages.
- *Go* appears strongly in the preferences for the next year as *Python* increases in interest (nearly tied with *Java*); this may indicate an increase in programmers working on these languages.

DATABASE TRENDS

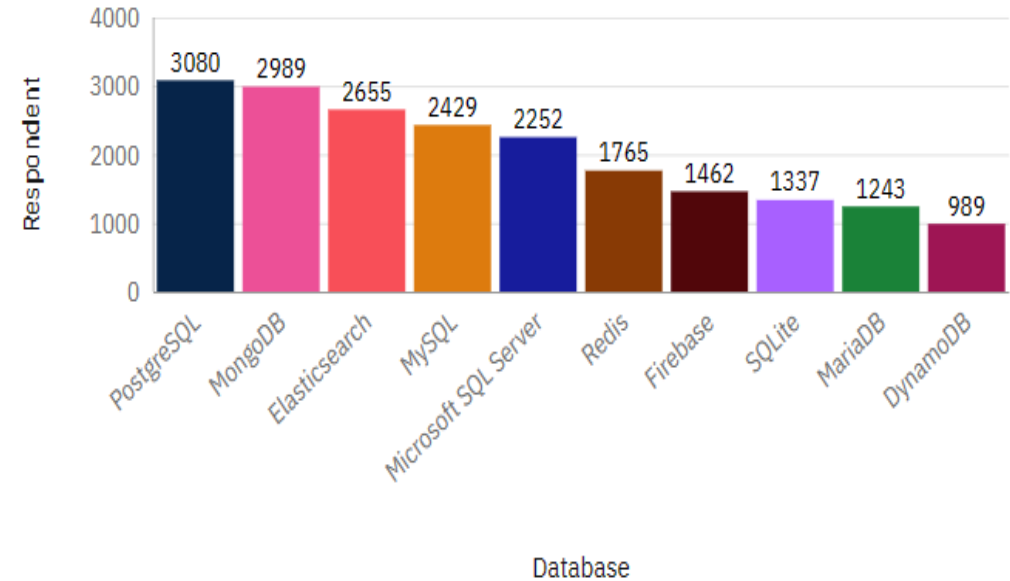
Current Year

Top 10 Database Worked With



Next Year

Top 10 Database Desired Next Year



DATABASE TRENDS

FINDINGS & IMPLICATIONS

Findings

- In the graph for the current year, it can be seen:
- *MySQL* appears as the most worked with database manager.
- 6 of the database managers are of the relational type and 4 NoSQL.
- In the graph for the next year, it can be seen:
- *PostgreSQL* is the most database manager desired, closely followed by *mongoDB*.
- 5 of the databases are relational and 5 NoSQL.
- *Elasticsearch* ranks third and above *MySQL* and *MS SQL Server*.
- *Oracle* does not appear in the top 10 databases desired and *dynamoDB* get in to the 10 top.

Implications

- Relational database managers and NoSQL are used extensively as there is almost an equality use of them.
- This seems to continue as next year they share 50% of the preferences.
- *PostgreSQL* could be the most used relational database manager in the future as it leads the 10 most desired for the next year followed closely by *MongoDB* and *Elasticsearch*.

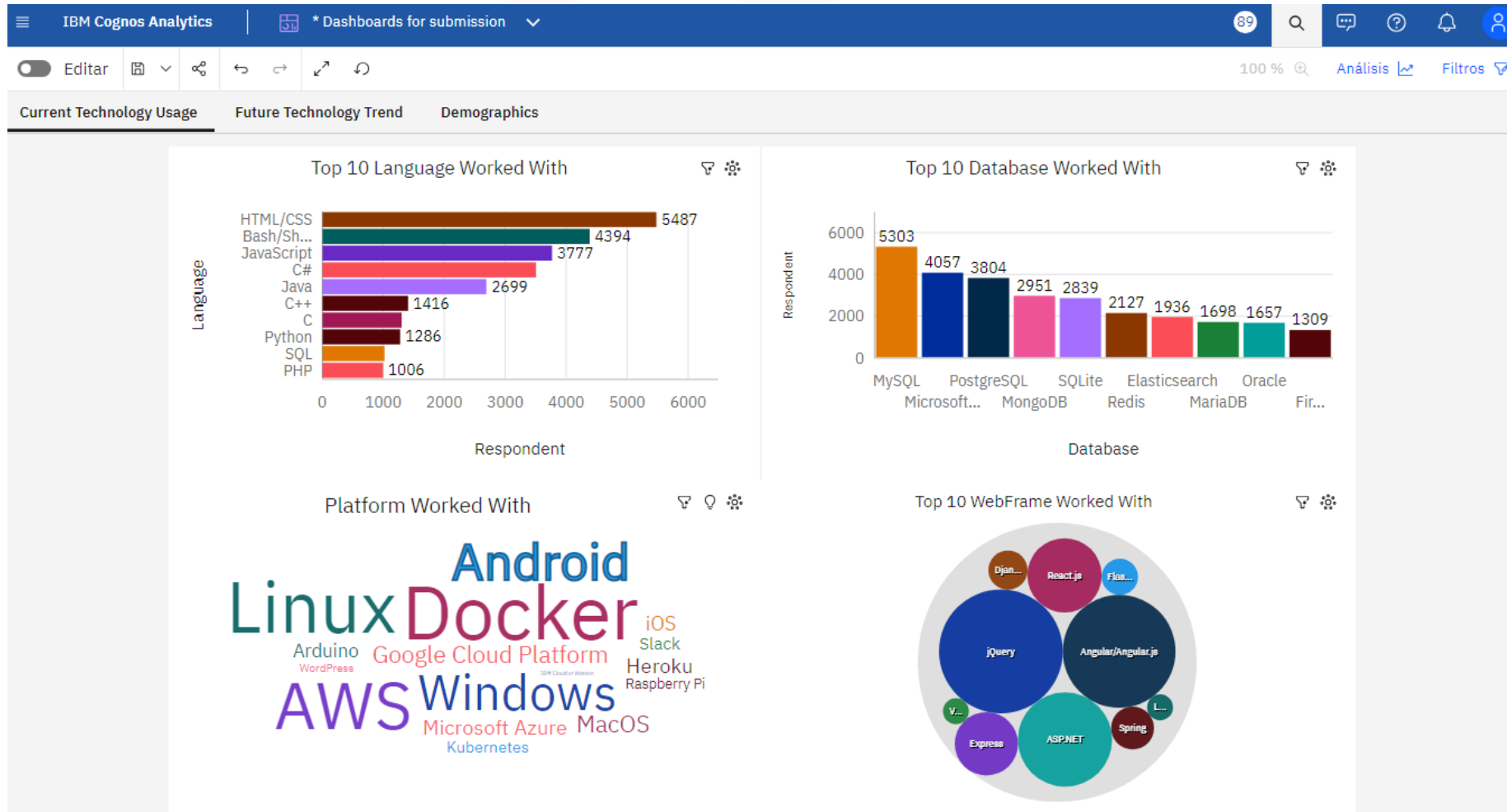
DASHBOARD



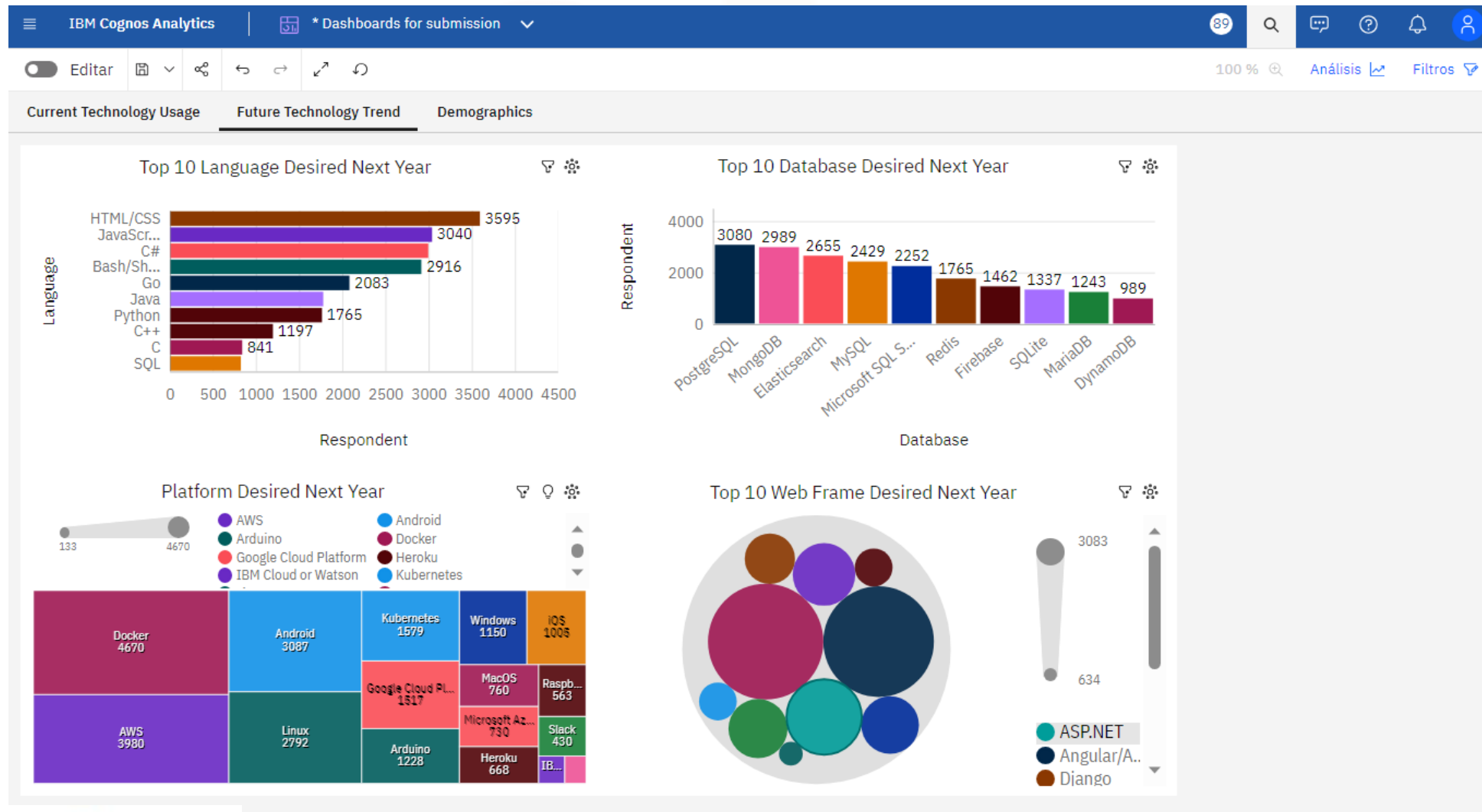
Permanent link of the read-only view of the Cognos dashboard:

https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCapstone%2BProject%2FDASHBOARDS%2FDashboards%2Bfor%2Bsubmission&action=view&mode=dashboard&subView=model0000018a5cdeb619_00000002

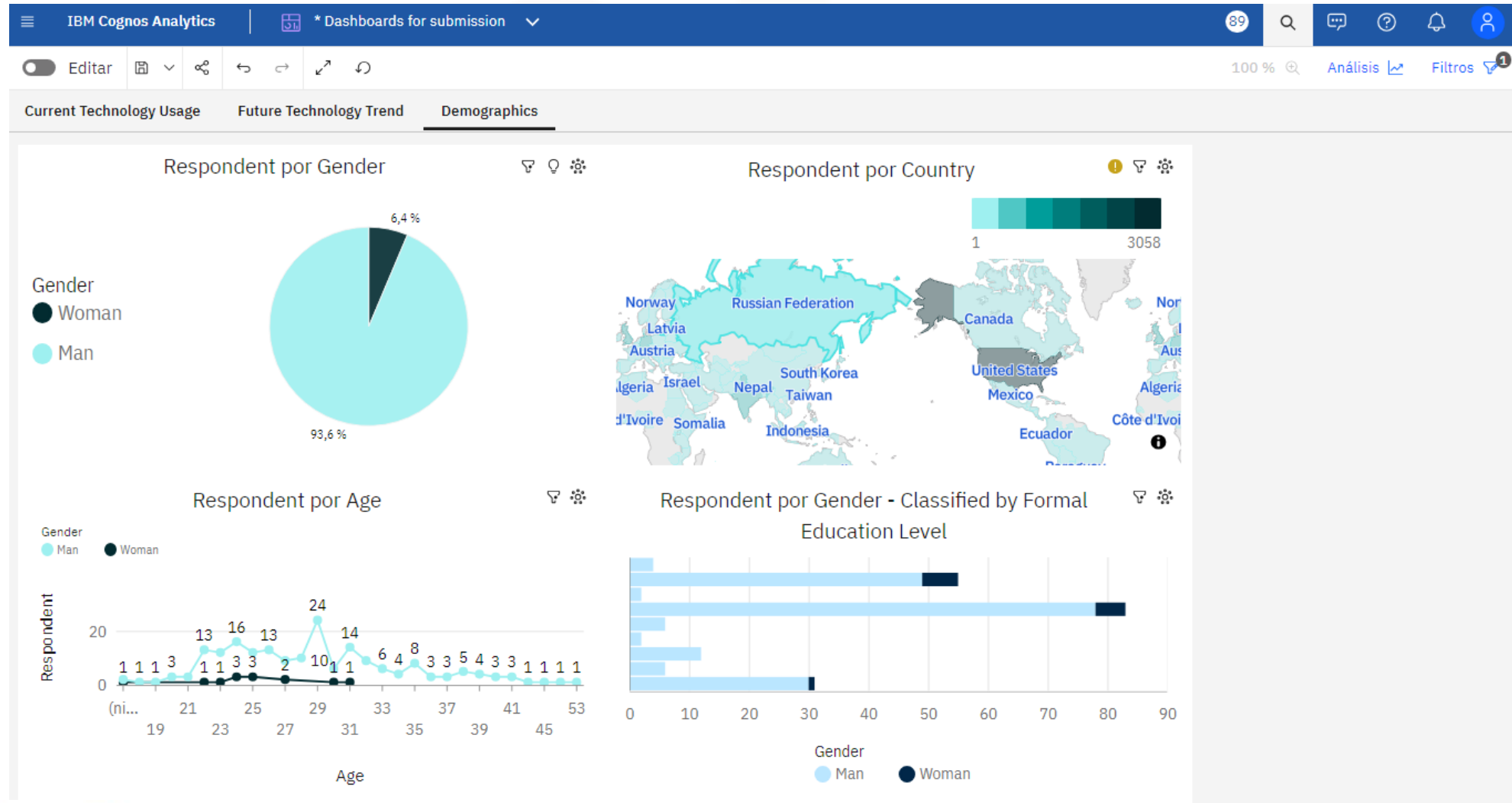
DASHBOARD TAB 1



DASHBOARD TAB 2



DASHBOARD TAB 3



DISCUSSION



- *HTML/CSS* is currently the most widely used programming languages.
- *Go* is experiencing a surge in popularity and *Python* is gaining ground, nearly on par with *Java*. This suggests that more programmers may be working with these languages in the coming years.
- Relational and NoSQL database managers are both extensively used, with approximately equal usage. This trend is expected to persist, with both technologies holding 50% in the 10 top of the desired programming languages and data bases for the next year.
- *MySQL* is currently positioned as the most widely used relational database manager in the market.
- *PostgreSQL* has been ranked as the top choice in the list of the most desirable database managers for the upcoming year, with *MongoDB* and *Elasticsearch* following closely behind, therefore more programmers may be working with these databases in the near future.

OVERALL FINDINGS & IMPLICATIONS

Findings

- *HTML/CSS* is the most frequently worked with and desired language category.
- For the next year, *PostgreSQL* is the most desired, followed closely by *mongoDB*.
- *Go* is not currently in the top 10 languages worked with, but it is the fifth most desired language for the next year.
- *Python* is almost tied with *Java* as the desired language for the next year, while *C*, *C++*, and *SQL* are lower on the list.
- In terms of database managers for the current year, *MySQL* is the most worked with, with 6 relational and 4 NoSQL databases.
- There are 5 relational and 5 NoSQL databases, with *Elasticsearch* ranking third.
- *Oracle* does not appear in the top 10 desired databases, but *dynamoDB* enters the list.

Implications

- The most preferred programming languages include *HTML/CSS*, *Bash/Shell/PowerShell*, *JavaScript*, and *C#*.
- There is a growing interest in *Go* for the upcoming year, while *Python* is on the rise and almost on par with *Java*.
- This suggests an increasing number of programmers working with these languages.
- Both relational and NoSQL database managers are widely used, with a nearly equal distribution between the two types. This trend is expected to continue, with both types comprising 50% of the preferences for the next year.
- *PostgreSQL* is positioned to become the most utilized relational database manager in the future, leading the top 10 desired list for the next year. It is closely followed by *MongoDB* and *Elasticsearch*, indicating their popularity among developers.

CONCLUSION



Based on the data provided and their analysis, these are the conclusions:

- *HTML/CSS* is currently the most widely used programming language.
- *Go* is experiencing a surge in popularity and *Python* is gaining ground, nearly on par with *Java*. This suggests that more programmers may be working with these programming languages in the coming years.
- Relational and NoSQL database managers are both extensively used, with approximately equal usage. This trend is expected to persist, with both technologies holding 50% in the 10 top of the desired programming languages and data bases for the next year.
- *MySQL* is currently positioned as the most widely used relational database manager in the market.
- *PostgreSQL* has been ranked as the top choice in the list of the most desirable database managers for the upcoming year, with *MongoDB* and *Elasticsearch* following closely behind. Therefore, more programmers may be working with these databases in the near future.

APPENDIX A – Source of data

Developer Survey Results 2019:

https://insights.stackoverflow.com/survey/2019?utm_source=so-owned&utm_medium=blog&utm_campaign=dev-survey-2019&utm_content=launch-blog

Full Data Set:

<https://info.stackoverflowsolutions.com/rs/719-EMH-566/images/stack-overflow-developer-survey-2019.zip>

Dashboards:

https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCapstone%2BProject%2FDASHBOARDS%2FDashboards%2Bfor%2Bsubmission&action=view&mode=dashboard&subView=model0000018a5cdeb61900000002

Popular languages:

https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/labs/datasets/Programming_Languages.html

APPENDIX B – Links of interest



Cognos Analytics:

<https://www.ibm.com/products/cognos-analytics>

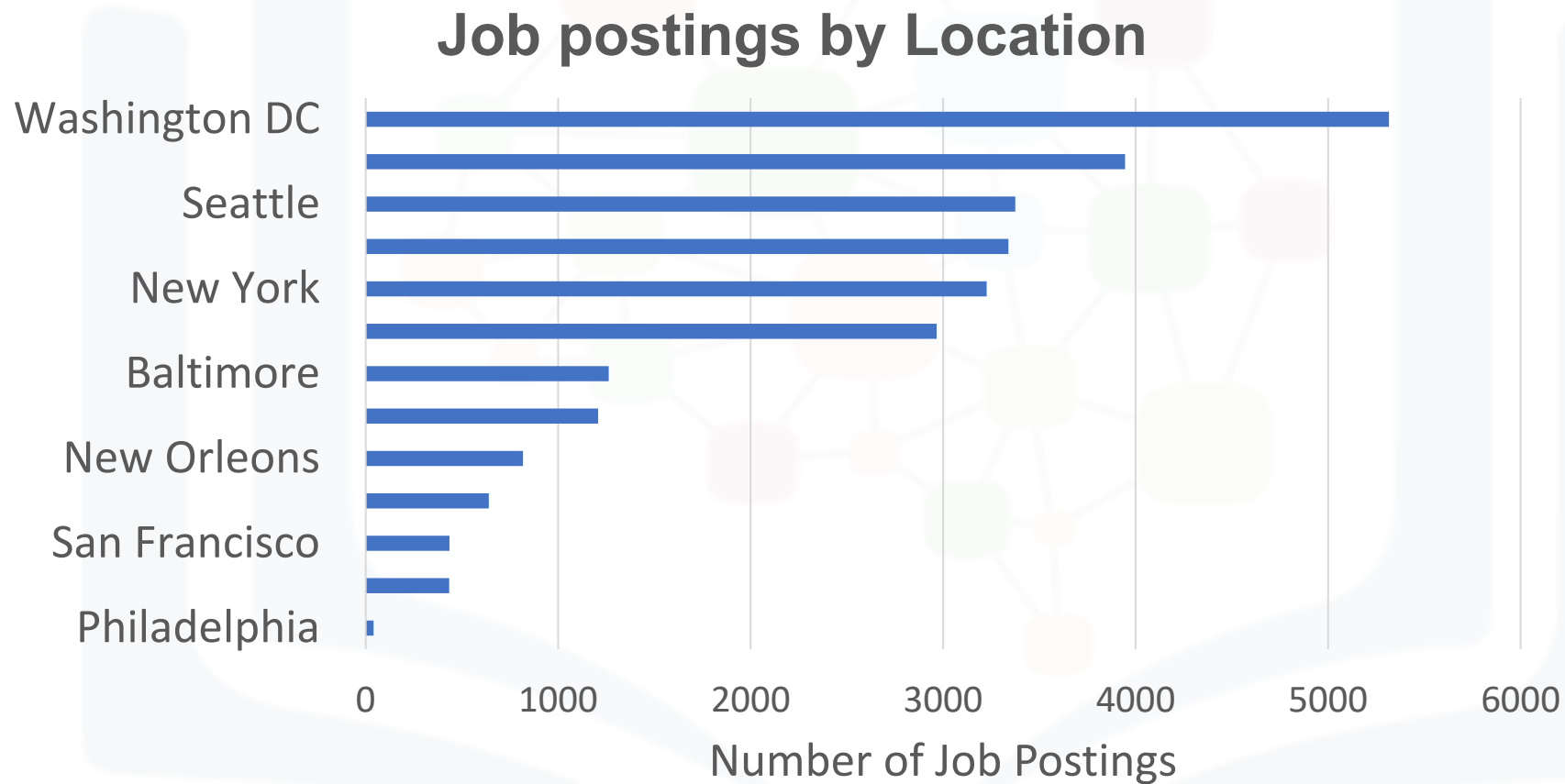
Stack overflow:

<https://stackoverflow.com>

MS Office on line (must have an account):

<https://www.office.com/>

JOB POSTINGS



POPULAR LANGUAGES

