

# IBM Data Analysis Capstone Project



<TITLE : Stack Overflow  
Developer Survey 2019 data  
Analysis>

<Presented by: Mr LIGHT>

<DATE: 6/22/2021>

# OUTLINE

---



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

# EXECUTIVE SUMMARY

---



- Point1: This work focused on a modified subset of the Stack Overflow Developer Survey 2019 data
- Point2: Some visualizations have been created and add to dashboards using Cognos Dashboard Embedded (CDE)
  - Sub Point 1: Dashboard of Current Technology Usage
  - Sub Point 2: Dashboard of Future Technology Trends
  - Sub Point 3: Dashboard of Demographics
- Point3: the studies revealed that JavaScript, HTML/CSS, SQL, Bash/Shell/PowerShell, Python, Java, C#, TypeScript, PHP and C++ are the top 10 language worked with and JavaScript, HTML/CSS, Python, SQL, TypeScript are the top 5 Language Desire Next Year
- Point4: the studies revealed that MySQL, Microsoft SQL Server, PostgreSQL, SQLite, MongoDB, Redia, Elasticsearch, Oracle, MariaDB and Firebase are the top 10 Database worked with and PostgreSQL, MySQL, Redia, MongoDB and SQLite are the top 5 Data Desire Next Year
- Point5:

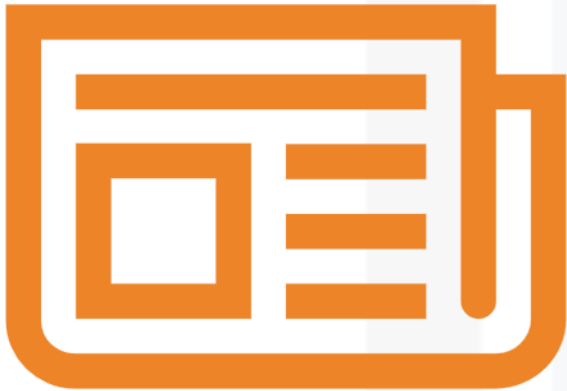
# INTRODUCTION

---



- Point1: This work focused on data analytics of a modified subset of the Stack Overflow Developer Survey 2019 data
- Point 2: This work is intended for data-driven deciders as well as all data lovers
- Point3: This work highlights gains of dashboard in data visualization
- Point4: Some visualizations have been created and add to dashboards using Cognos Dashboard Embedded (CDE)
- Point5: The Analysis have been performed by leveraging two CSV files as data assets, an m5\_survey\_data\_demographics.csv file and an m5\_survey\_data\_technologies\_normalised.csv file
- Point6
  - Sub Point1: the m5\_survey\_data\_technologies\_normalised.csv file has enabled the current technology usage and the future technology trend for the top 10 programming language, Database, Platform and Web Frame.
  - Sub Point2: the m5\_survey\_data\_demographics.csv file has enabled 4 visualizations namely, Respondent classified by Gender, Respondent Count for Countries, Respondent Count by Age and Respondent Education Level

# METHODOLOGY



- Point1: The dataset used in this work come from the following source <https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>
- under a **ODbL: Open Database License**.
- Point2: For the sake of consistency, we have been using a modified subset of that dataset ranging in 2 CSV files, m5\_survey\_data\_demographics.csv file and m5\_survey\_data\_technologies\_normalised.csv file
- Point3: The Software used in this work is the free lite plan version of Cognos Dashboard Embedded (CDE) through IBM Cloud.
- Point4: In CDE, 3 dashboards (3 separate tabs under a single dashboard) has been created as follows:
  - Sub Point1: One dashboard using the 2 x 2 rectangle areas tabbed template - rename **Technology Usage**.
  - Sub Point2: One dashboard using the 2 x 2 rectangle areas tabbed template - rename **Future Technology Trend**.
  - Sub point3: One dashboard using the 2 x 2 rectangle areas tabbed template - rename **Demographics**.

# RESULTS

The Result Section will be tackled as follows

---

Programming  
Language Trends

Finding and  
Implications

Database Trends

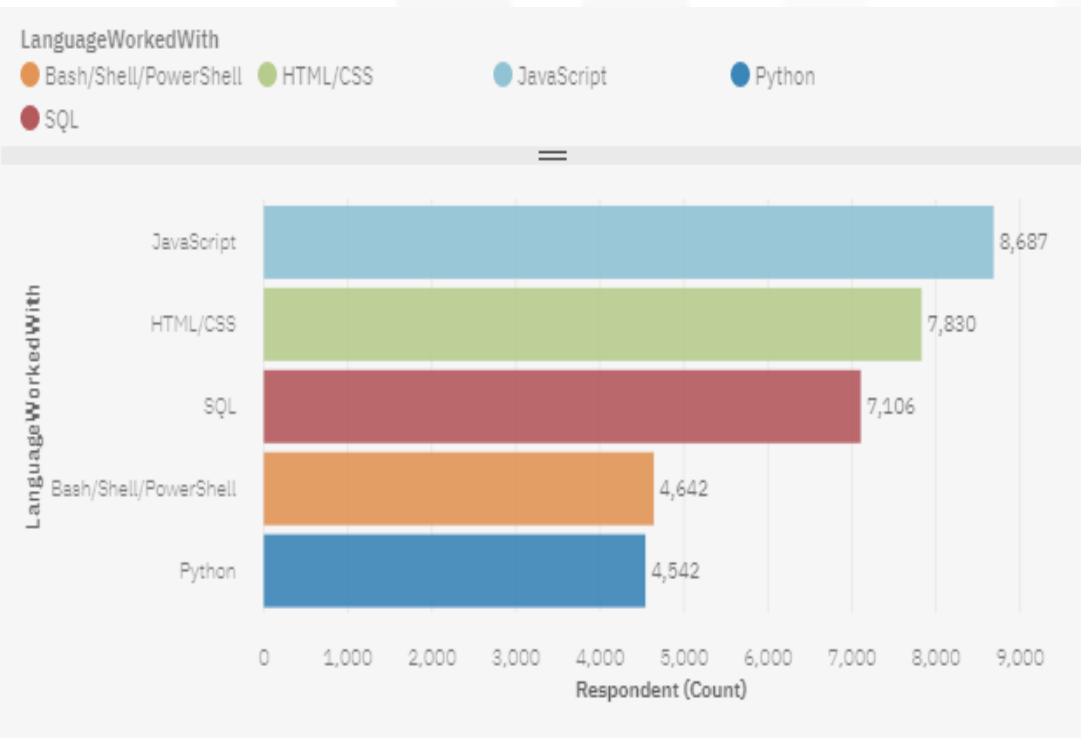
Finding and  
Implications

Shareable link of  
Cognos Dashboards  
and displaying  
screen shots

# PROGRAMMING LANGUAGE TRENDS

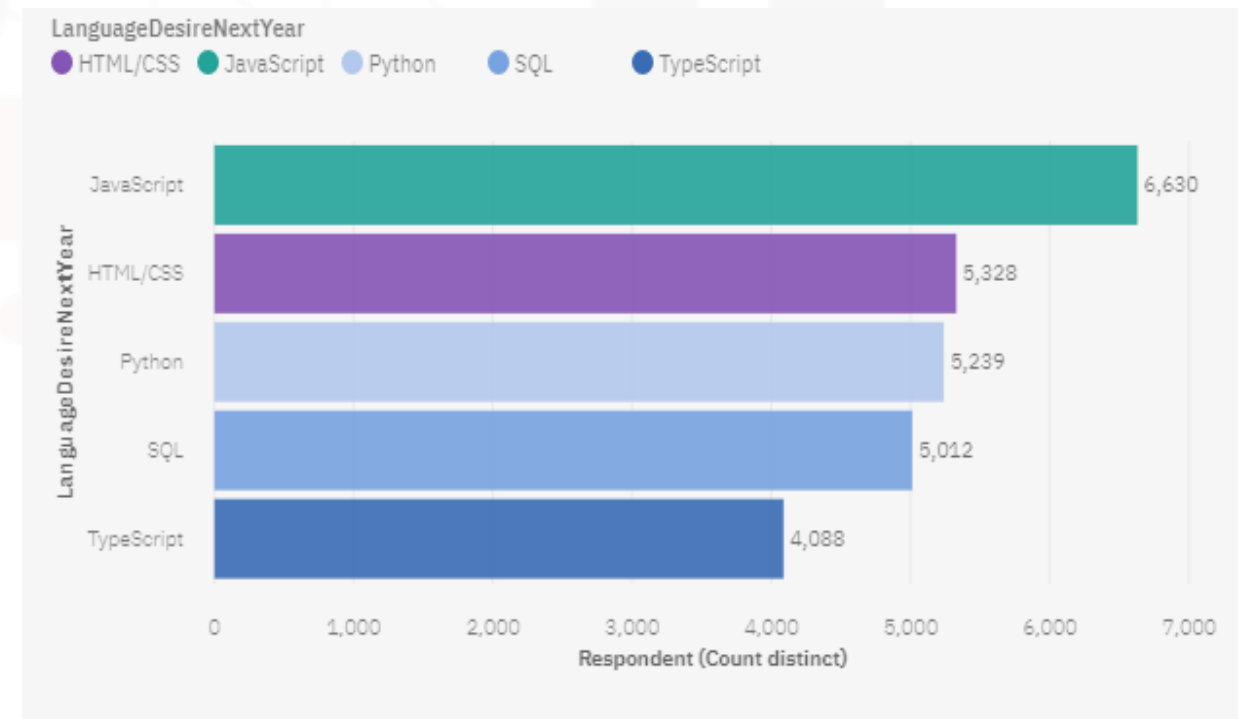
## Current Year

<Bar chart of top 5 programming languages for the current year .>



## Next Year

< Bar chart of top 5 programming languages for the next year .>



# PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

---

## Findings

- Finding 1: the Top 5 Programming Languages for the Current Year are JavaScript, HTML/CSS, SQL, Bash/Shell/PowerShell, Python
- Finding 2: the Top 5 Programming Languages for the Next Year are JavaScript, HTML/CSS, Python, SQL, TypeScript
- Finding 3 Python has evolved to the 3rd position as Programming Languages

## Implications

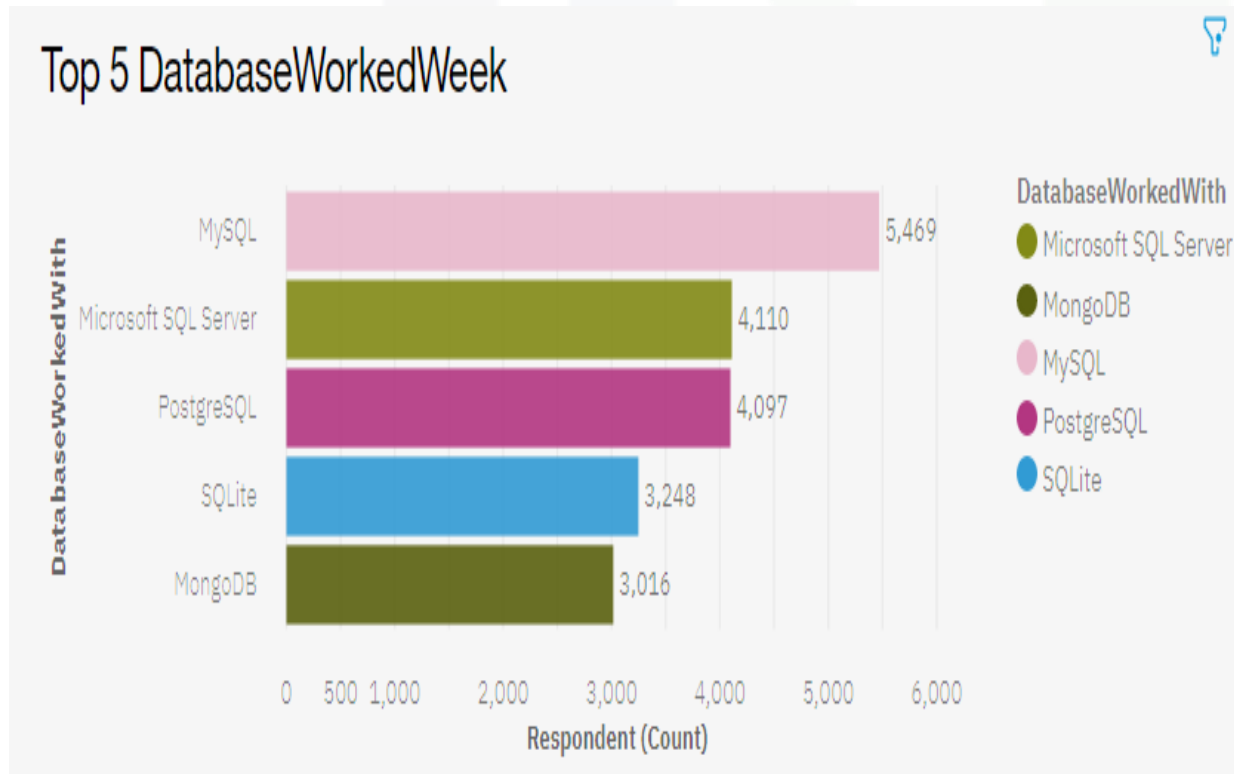
- Implication 1: Companies are willing to pay a premium for skilled in JavaScript to Python
- Implication 2: Bash/Shell/PowerShell isn't a valuable Programming language
- Implication 3: Python is a valuable Programming Language



# DATABASE TRENDS

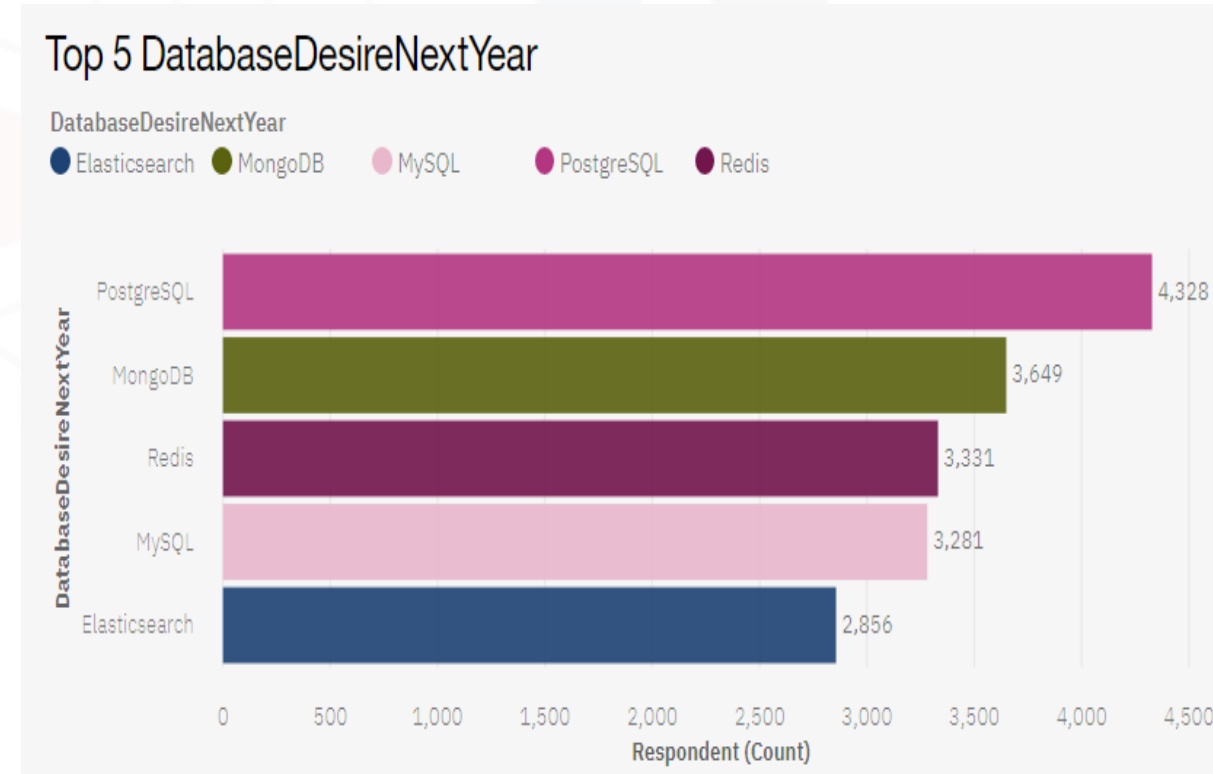
## Current Year

< Bar chart of top 5 databases for the current year >



## Next Year

< Bar chart of top 5 databases for the next year.>



# DATABASE TRENDS - FINDINGS & IMPLICATIONS

## Findings

- Finding 1: the Top 5 Database for the Current Year are MySQL, Microsoft SQL Server, PostgreSQL, SQLite, MongoDB
- Finding 2: the Top 5 Database for the Next Year are PostgreSQL, MongoDB, Redis, MySQL, Elasticsearch
- Finding 3: PostgreSQL has evolved to the 1st position as Database Desire for the Next Year

## Implications

- Implication 1: Companies are willing to pay a premium for skilled in MySQL for the current Year
- Implication 2: PostgreSQL and MongoDB are the most valuable Databases for the Next Year
- Implication 3: Companies are willing to pay a premium for skilled in PostgreSQL for the Next Year

# DASHBOARD

---

<Hint: Please right clic on the Permanent link of the read-only view of the Cognos dashboard below and choose open link in new tab option.>

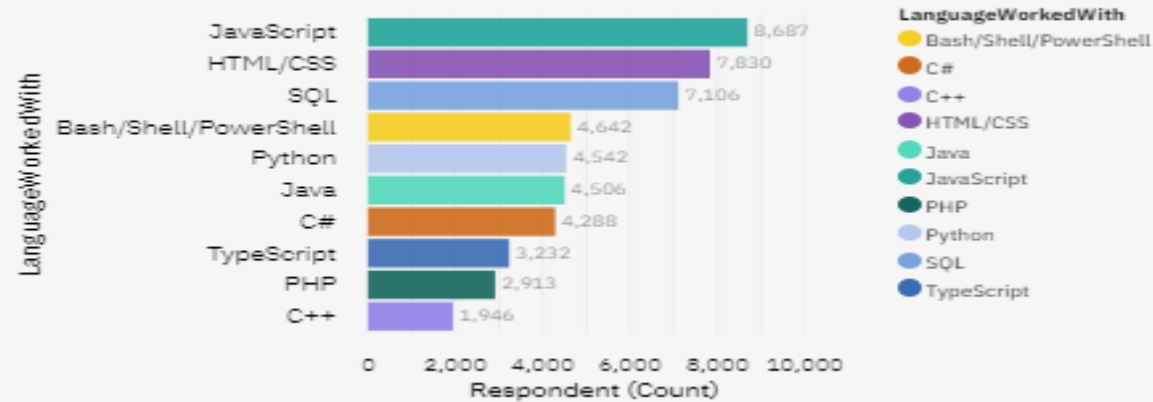


[Permanent Link of the read-only view of the Cognos dashboard](https://dataplatfom.cloud.ibm.com/dashboards/e6b826a5-e92e-48aa-8e19-83e204c35274/view/4521c639149275d670c2c4e407cb2c0e7a622d0cb1bb8550d6807b495c677097f33d41c3c82e4c0b89195660a1ea4458c8)

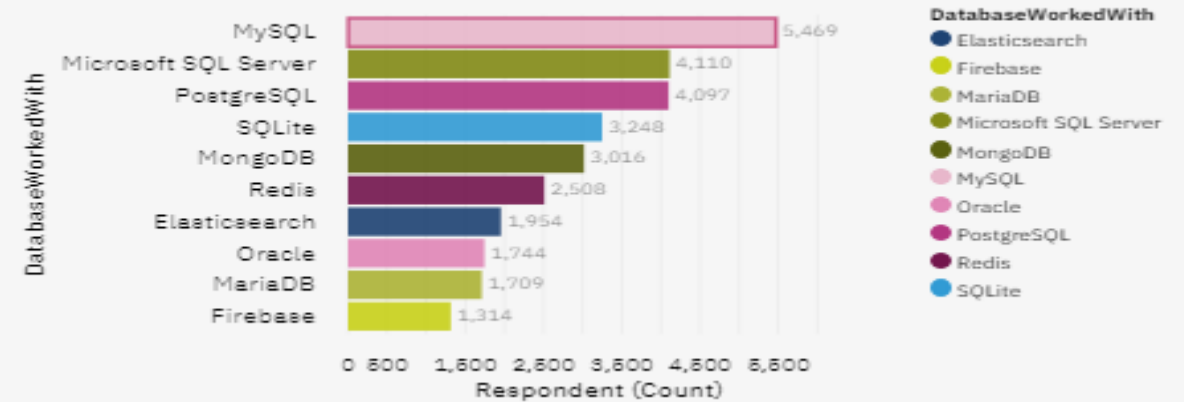
<https://dataplatfom.cloud.ibm.com/dashboards/e6b826a5-e92e-48aa-8e19-83e204c35274/view/4521c639149275d670c2c4e407cb2c0e7a622d0cb1bb8550d6807b495c677097f33d41c3c82e4c0b89195660a1ea4458c8>

# DASHBOARD TAB 1

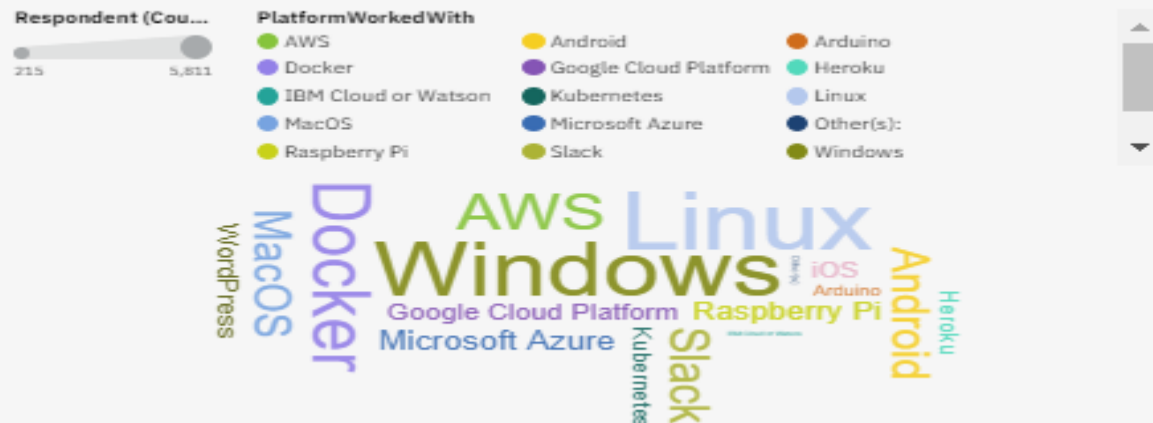
Top 10 LanguageWorkedWith



Top 10 DatabaseWorkedWith



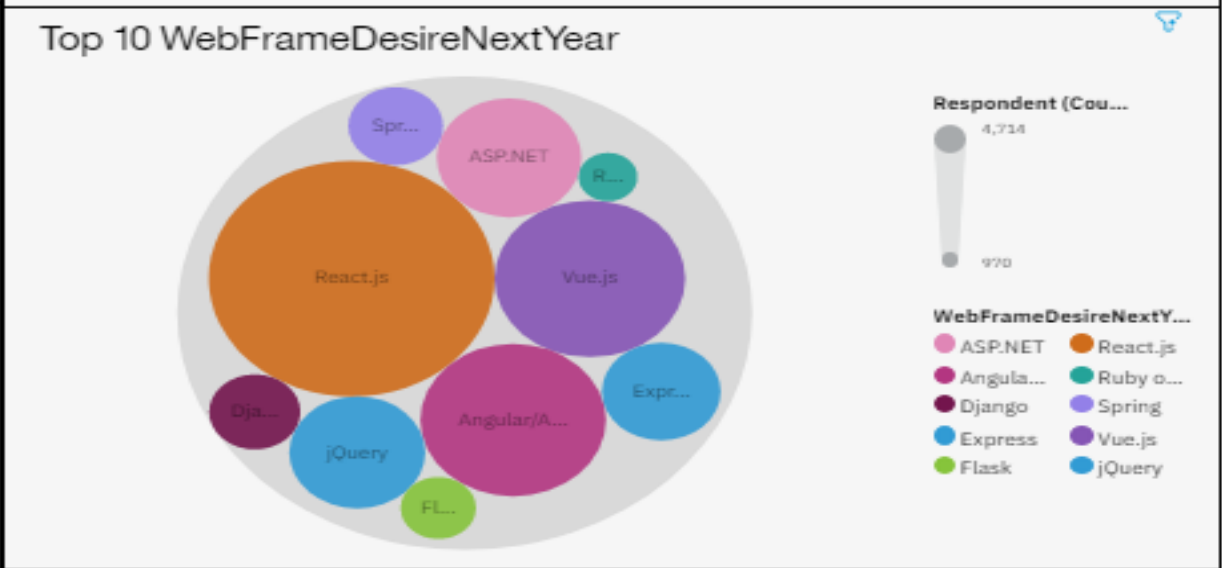
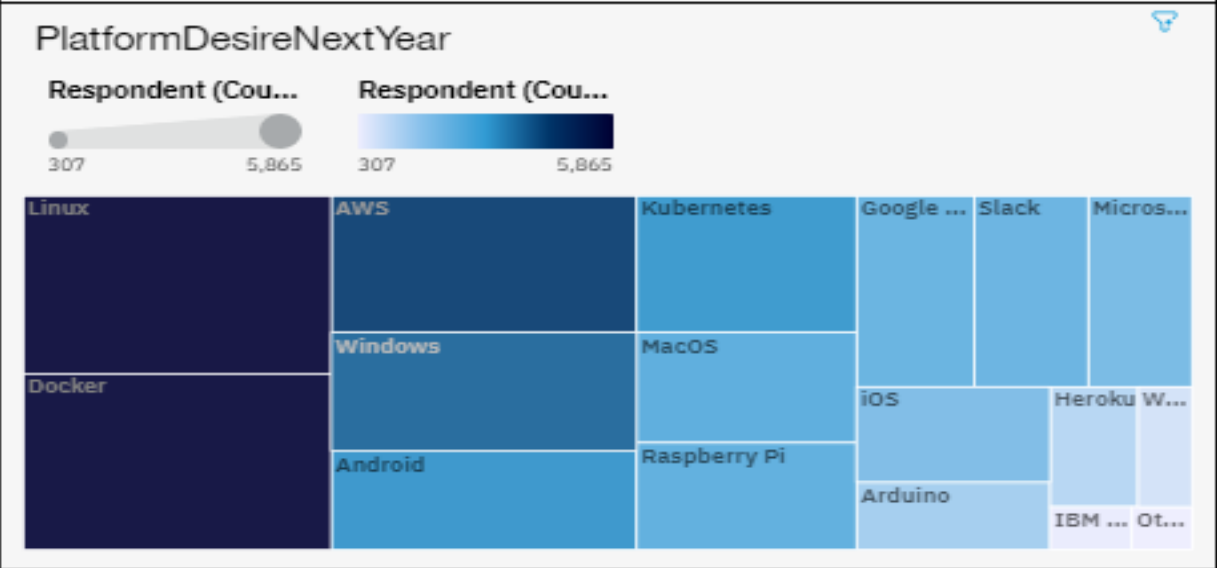
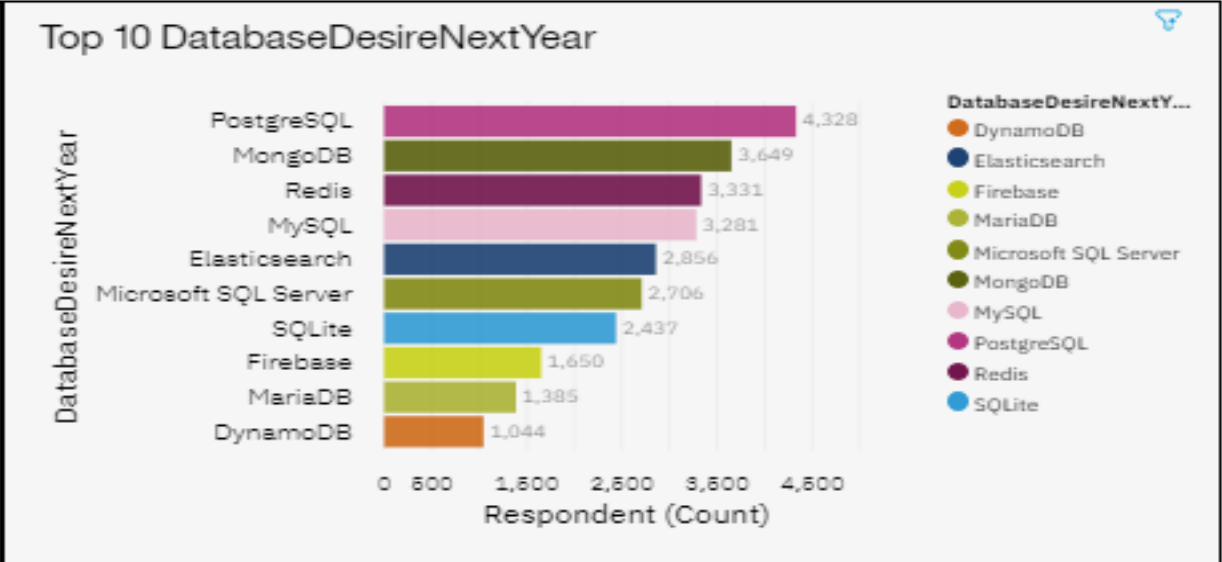
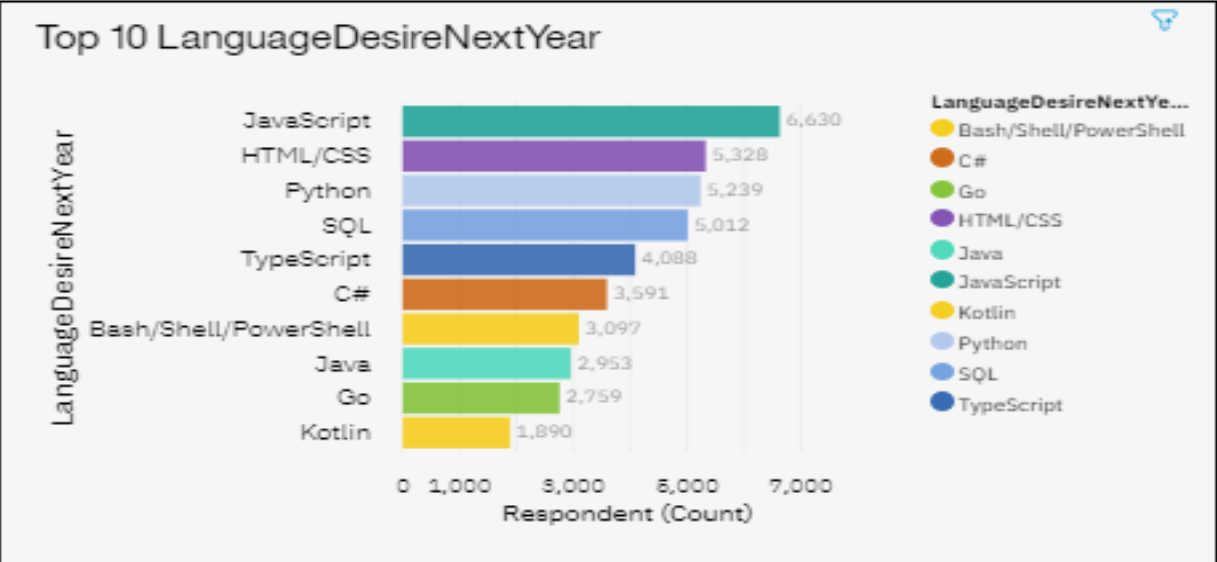
PlatformWorkedWith



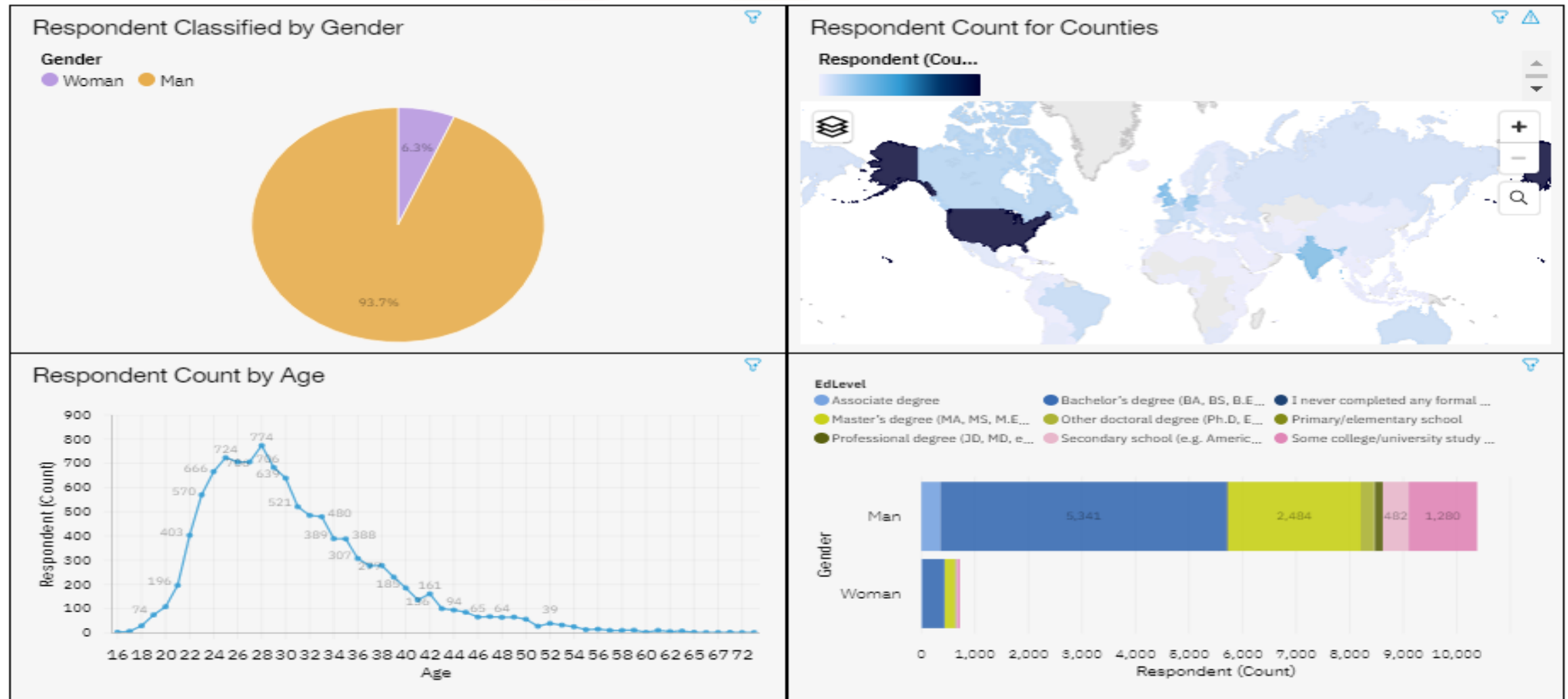
Top 10 WebframeWorkedWith



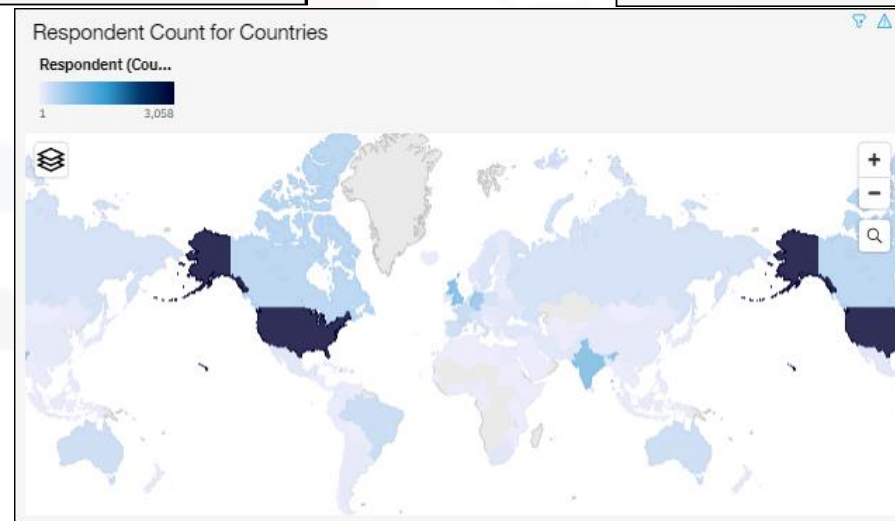
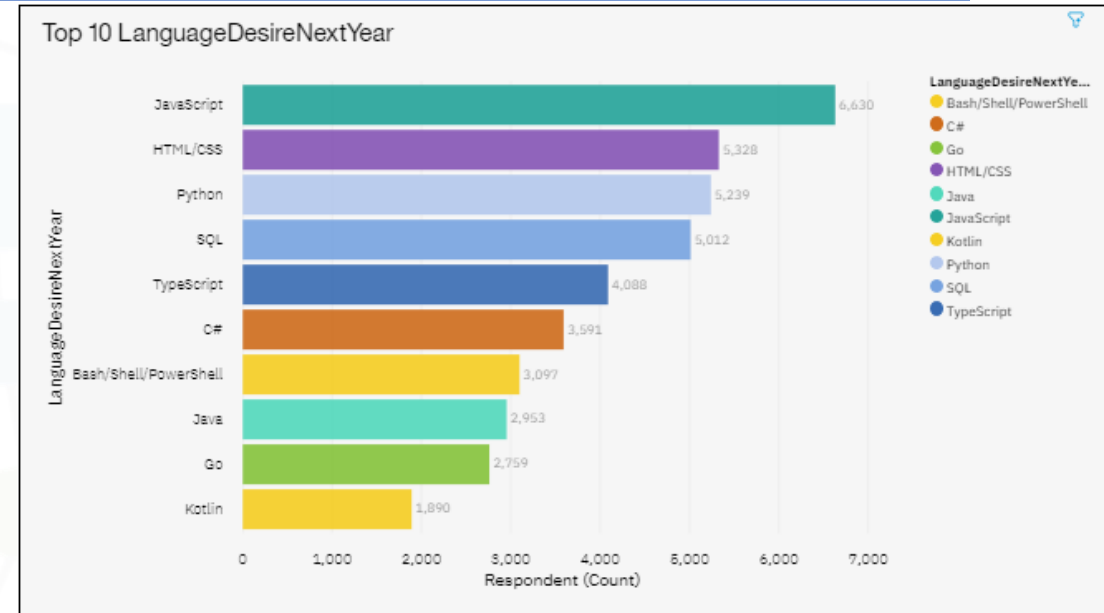
# DASHBOARD TAB 2



# DASHBOARD TAB 3



# DISCUSSION



# OVERALL FINDINGS & IMPLICATIONS

---

## Findings

- Finding 1: The Dashboard TAB1 highlights Linux, Windows and AWS as the most widely used Platform Worked With
- Finding 2: The Dashboard TAB2 Warns that Respondent are more likely to learn Python
- Finding 3: The Dashboard TAB3 informs that a large majority of Respondent are from US

## Implications

- Implication 1: Linux-Windows-AWS offers some advantages in terms of Lower price – Higher performance but also for the ease of using
- Implication 2: Python is very suitable for Web Development, Data Science including Machine Learning-Data Analytics and Data Visualization, Scripting
- Implication 3: This is due to the state-of-the-art equipment out there, a persistent worldwide leadership of American Tech Companies



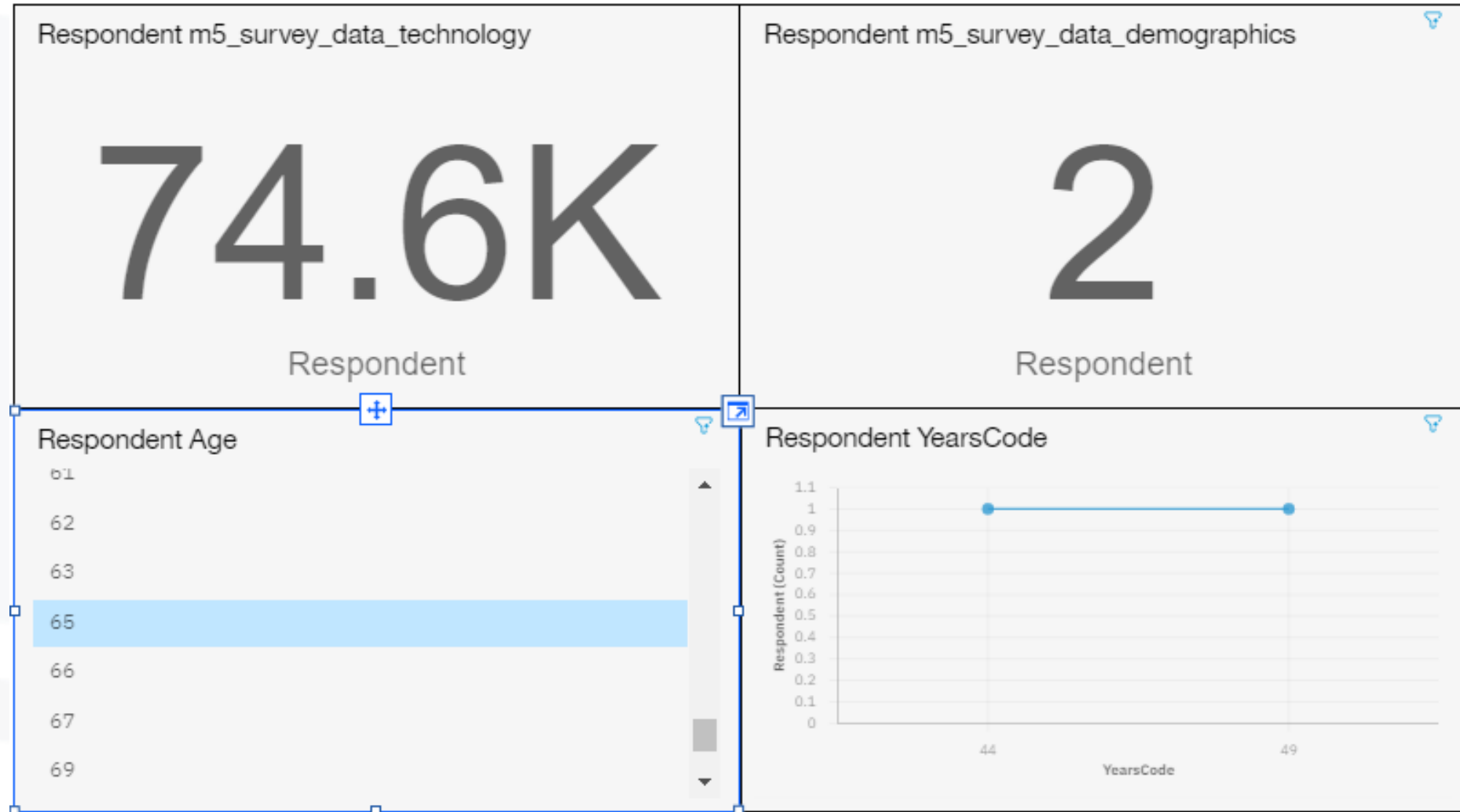
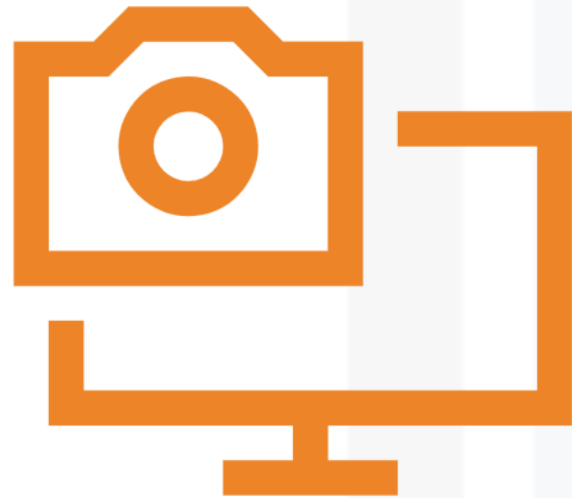
# CONCLUSION

---



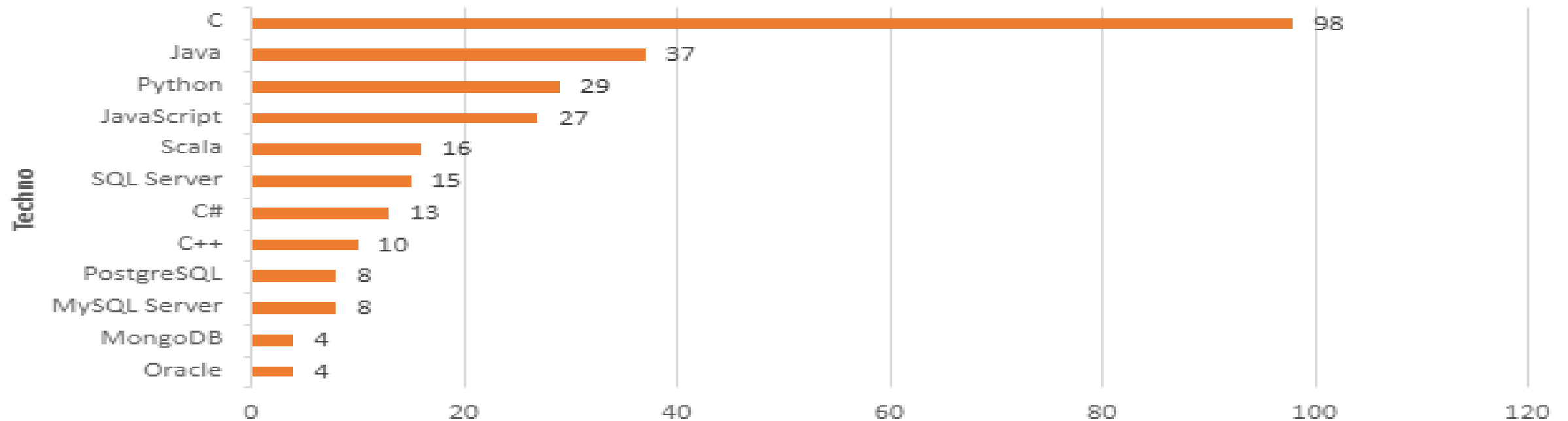
- Point 1: In the rapidly expanding technological world of today, the use of data to run business is the new standard
- Point 2: data analytics is crucial for mapping company's performance
- Point 3: data analytics is crucial for improving brand's customer experience
- Point 4: data analytics is crucial for making decisions quicker, solve problems faster
- Point 5: data analytics is indispensable for measuring success of company and employees
- Point 6: data analytics is indispensable for understanding users, market and the competition

# APPENDIX



# GITHUB JOB POSTINGS

Number of Job Posting per Techno



	Oracle	MongoD B	MySQL Server	PostgreS QL	C++	C#	SQL Server	Scala	JavaScrip t	Python	Java	C
Job	4	4	8	8	10	13	15	16	27	29	37	98

Number of Job Posting

# POPULAR LANGUAGES

Average Annual Salary per Language

