2019 Technology Situations and Trend for 2020

- IBM Certificate Capstone Project

Chengxin Luo

December 2020

IBM Developer

SKILLS NETWORK

OUTLINE

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

EXECUTIVE SUMMARY

- Data Collection
  - APIs
  - Web Scraping
  - Public Data Resources
- Data Wrangling
  - Duplicates Handling
  - Missing Values Handling
  - Data Standardization
- Found the current situation and future trends of changing technologies
- Analyzed the correlations of different features
- Visualized data and generated Dashboards by Python and Cognos

IBM **Developer**

SKILLS NETWORK

INTRODUCTION

Situations

- Technologies are being developed rapidly
- Keep pace with changing technologies and remain competitive
  - Important and necessary to know the trend
- Stack Overflow developer surveys since 2011
  - Clear view of the developing trend of programming languages as well as databases

Our Analysis

- Current situation in 2019 and future trend in 2020
- Correlations between different features

METHODOLOGY – Data Collection

- GitHub Jobs API
  - Number of jobs currently open for various technologies
- Web Scraping
  - Programming Languages and corresponding salaries
- Open-Sourced Data
  - 2019 Stack Overflow Developer Survey

IBM **Dev**loper

SKILLS NETWORK

METHODOLOGY – Data Wrangling

- **Duplicates Handling**
  - Finding and Removing
- **Missing Values Handling**
  - Finding and Imputing
- **Data Normalization**
  - E.g. Transform all the salaries (Weekly/Monthly/Yearly) to Yearly
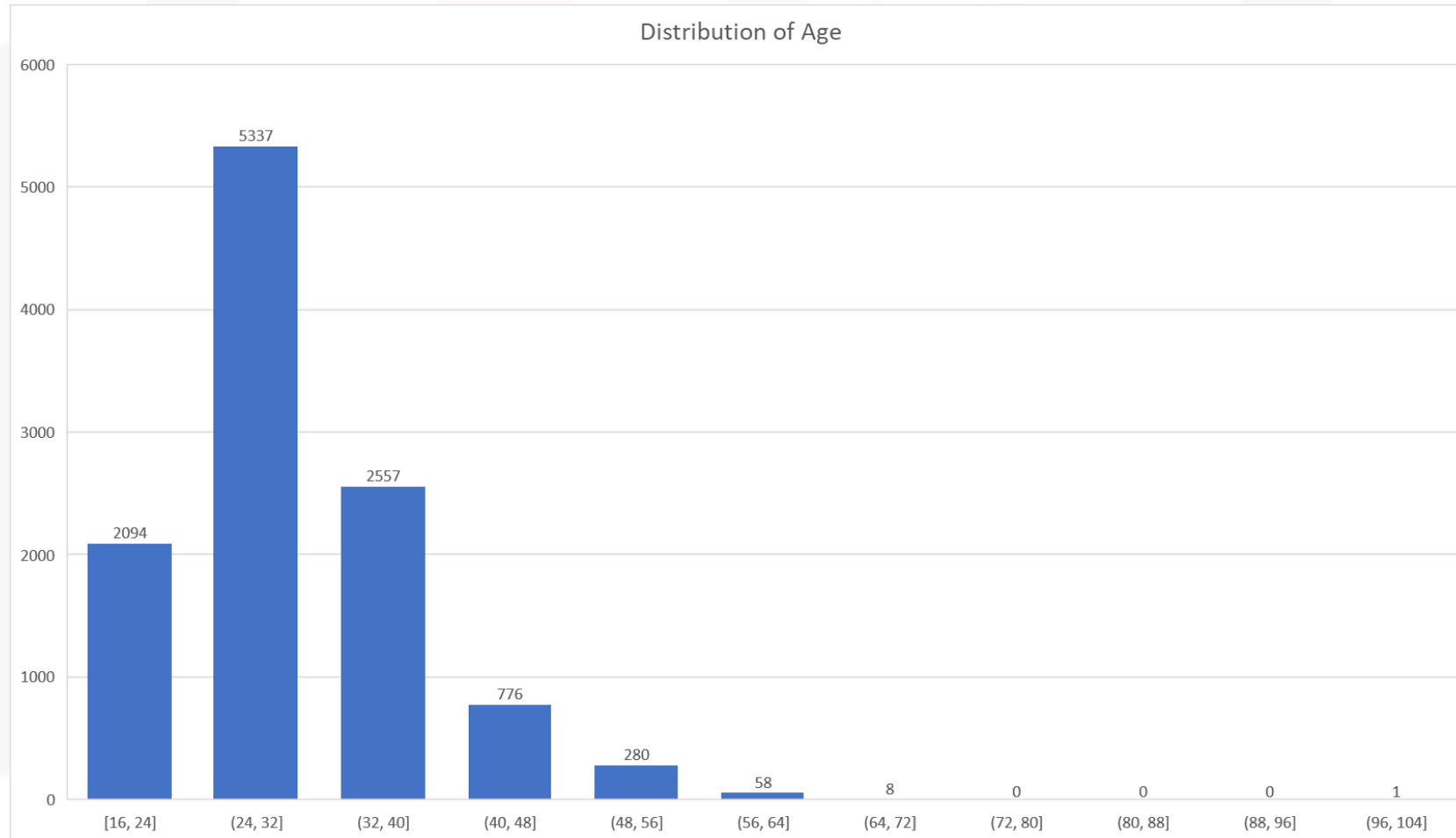
METHODOLOGY – Exploratory Data Analysis

- Determine how data is distributed
- Identify and remove outliers
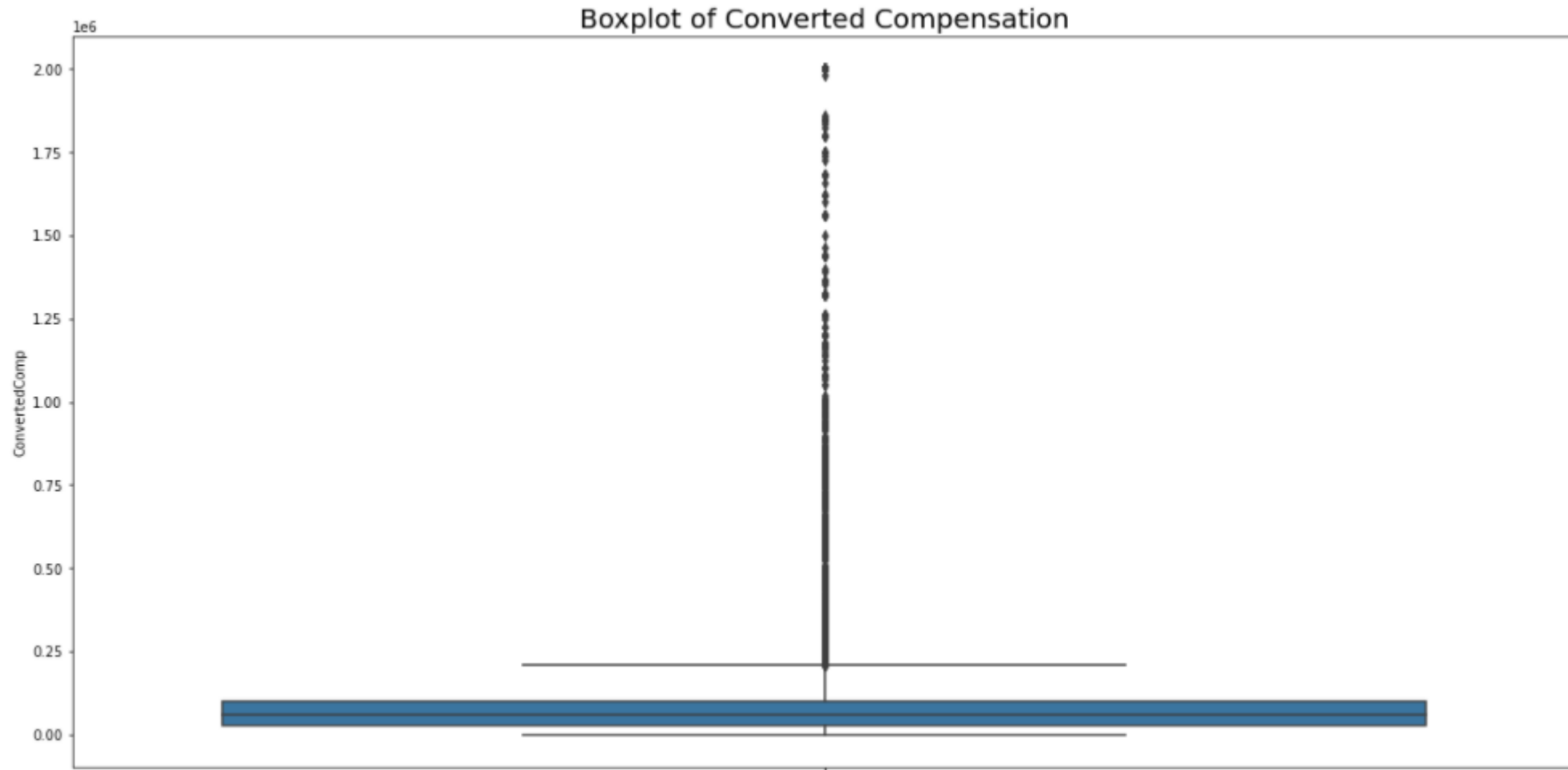- Find the correlations between different features

IBM **Developer**

SKILLS NETWORK

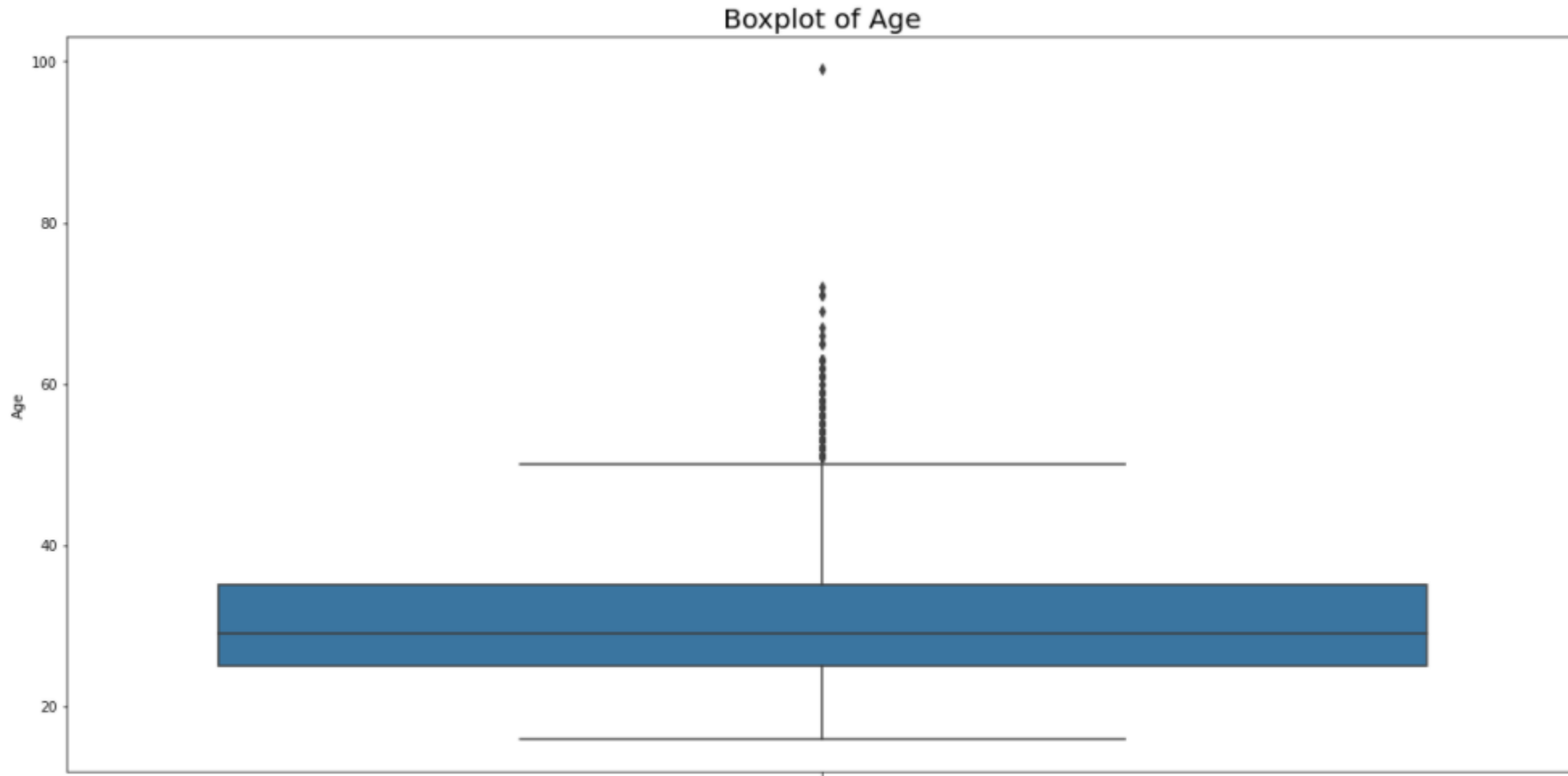# RESULTS – Distribution of Compensation
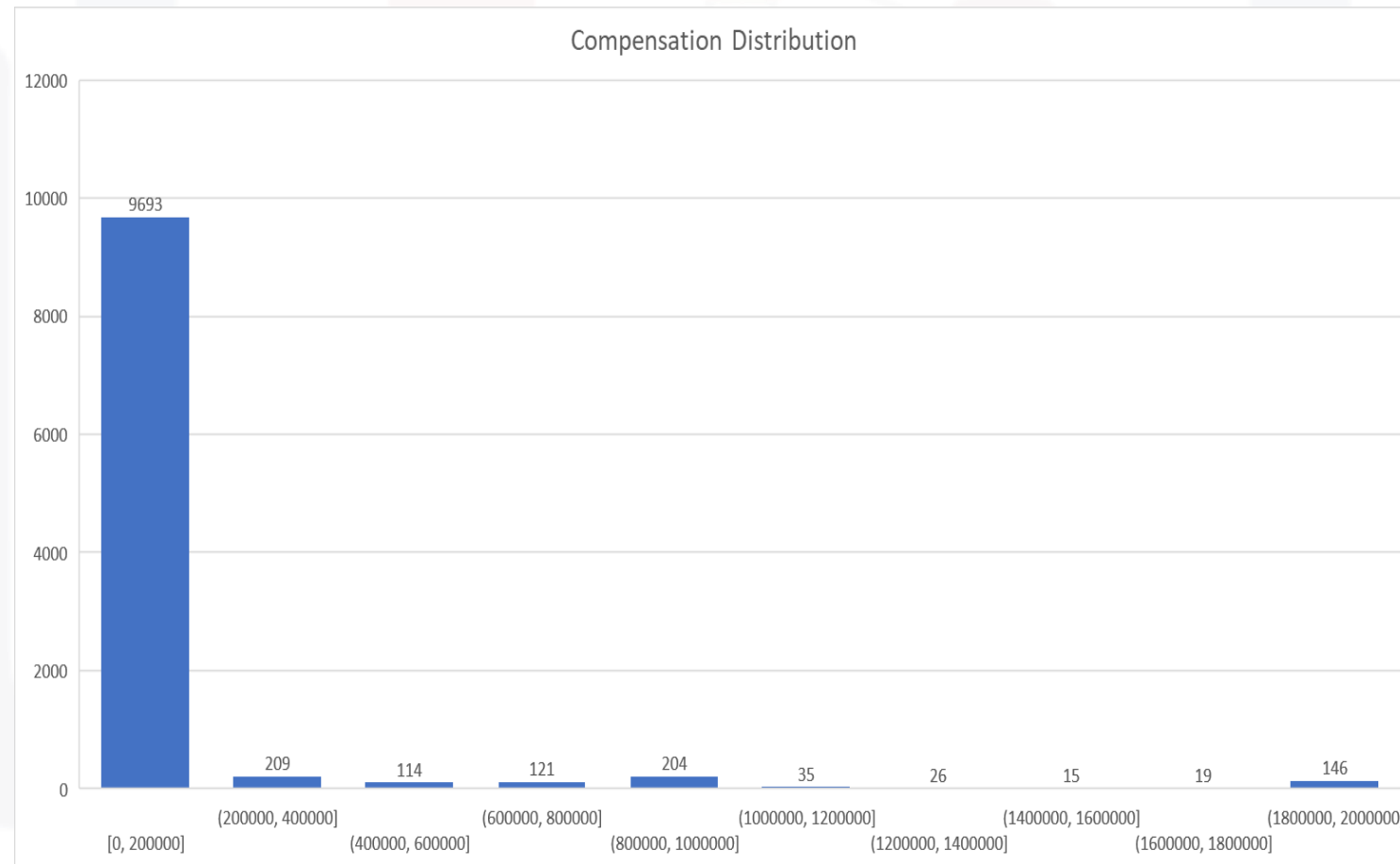
RESULTS – Distribution of Age

Distribution of Age

RESULTS – Boxplot of Compensation

Boxplot of Converted Compensation

RESULTS – Boxplot of Age

PROGRAMMING LANGUAGE TRENDS

Current Year

Top 5 Languages for 2019

Next Year
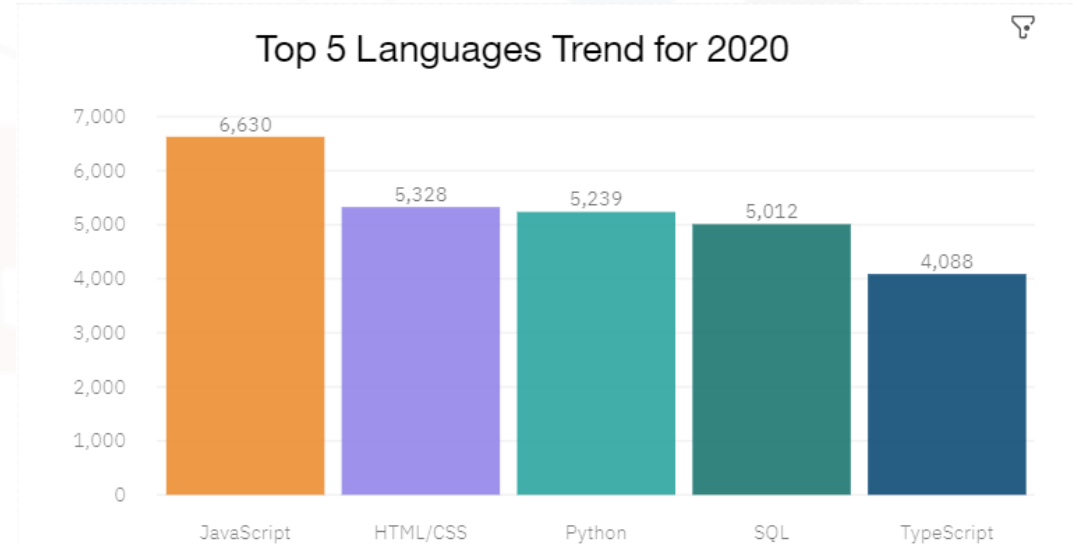

Top 5 Languages Trend for 2020
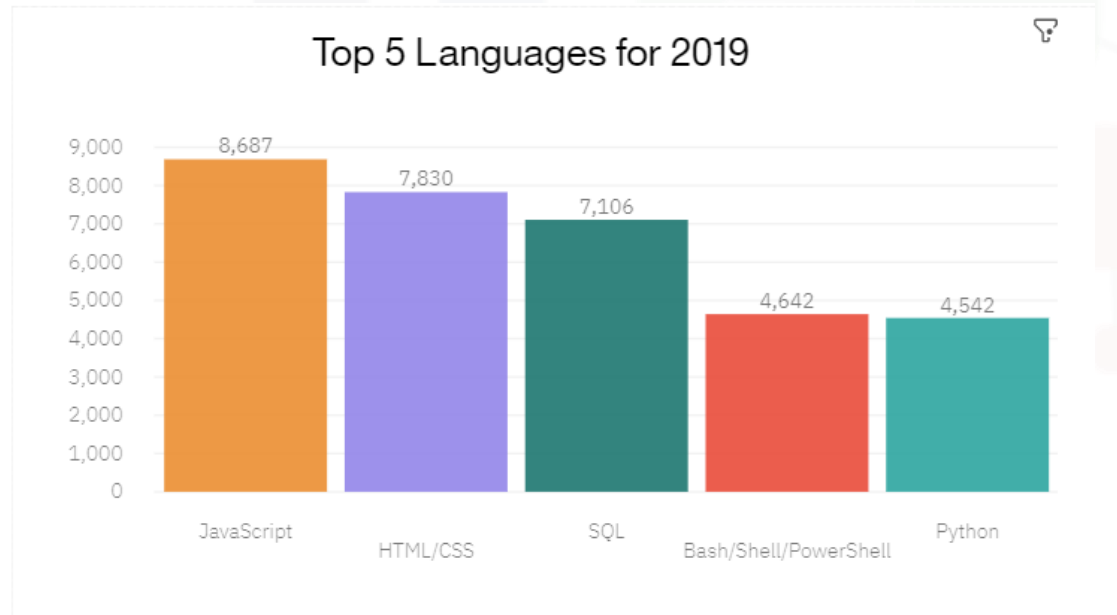
PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Average number of desires drops for top 5 languages

- Python remains in top 5 and desire is increased

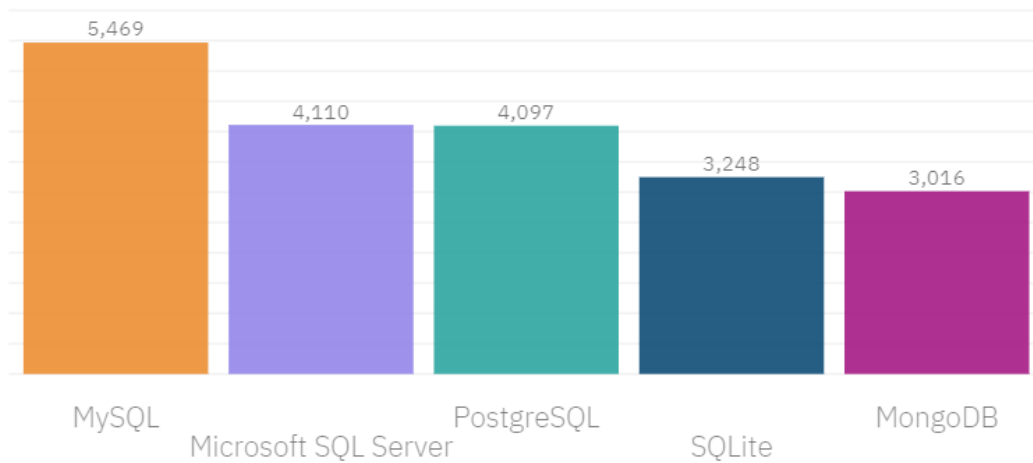- Typescript replaces Bash/Shell/PowerShell in top 5

Implications

- The non-top 5 languages has the trend to occupy more shares

- Python is becoming more competitive

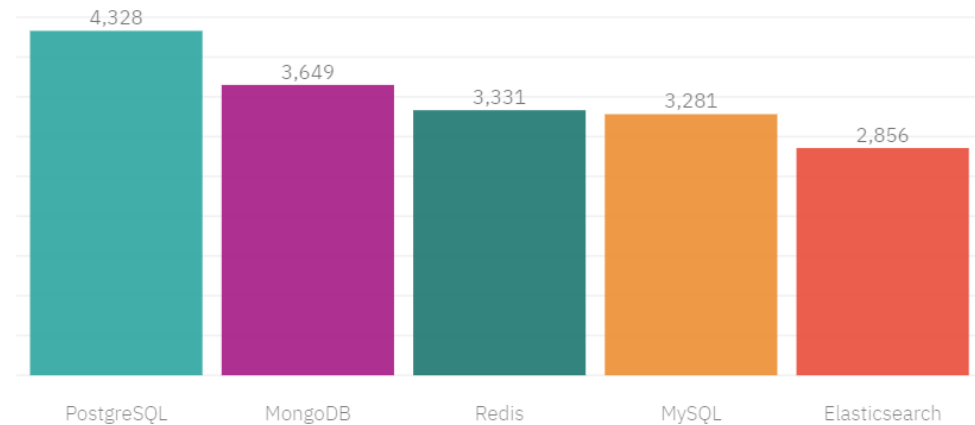- The competitive power of Shell Language has a dropping trend

IBM **Developer**

SKILLS NETWORK

DATABASE TRENDS

Current Year

Top 5 Database Worked With

5,469 — MySQL
4,110 — Microsoft SQL Server
4,097 — PostgreSQL
3,248 — SQLite
3,016 — MongoDB

Next Year

Top 5 Database Desired Next Year

4,328 — PostgreSQL
3,649 — MongoDB
3,331 — Redis
3,281 — MySQL
2,856 — Elasticsearch

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- SQL DB occupies 4 of top 5 in 2019 but only occupies 2 of top 5 in 2020 trend

- MySQL/PostgreSQL/Mongo DB remains top 5

- Redis and Elasticsearch step in top 5 trend

Implications

- SQL DB is losing power

- Open-sourced SQL DB is still competitive

- Cloud-based/No-SQL DB becoming more competitive
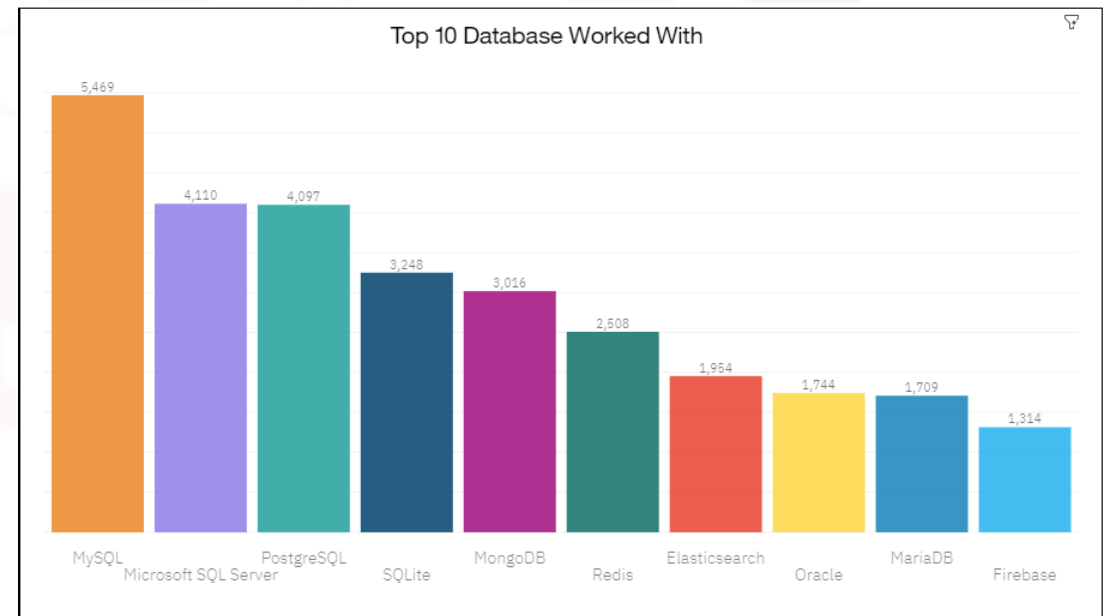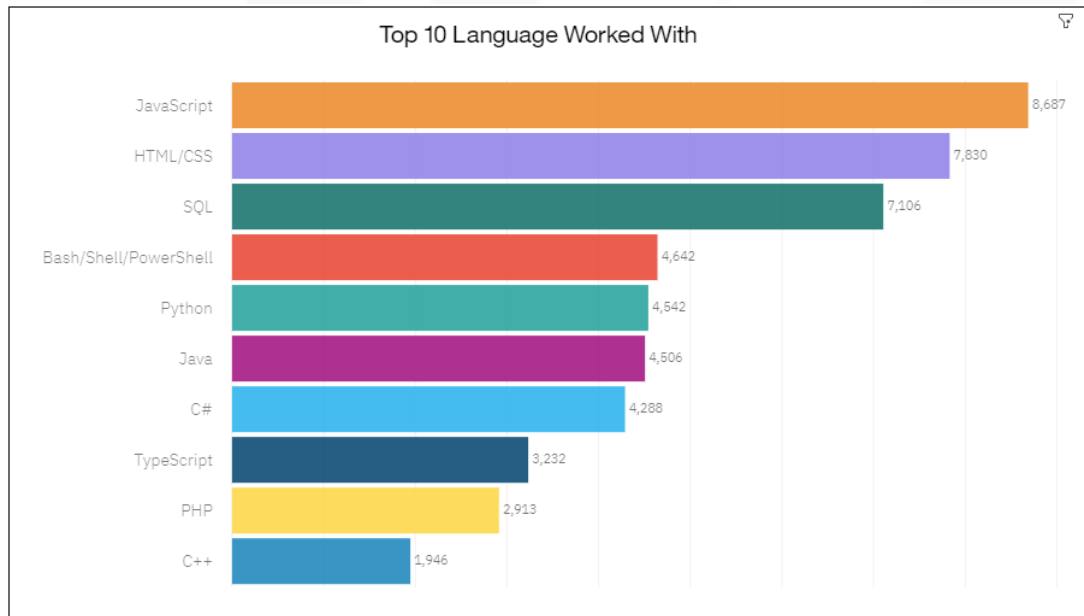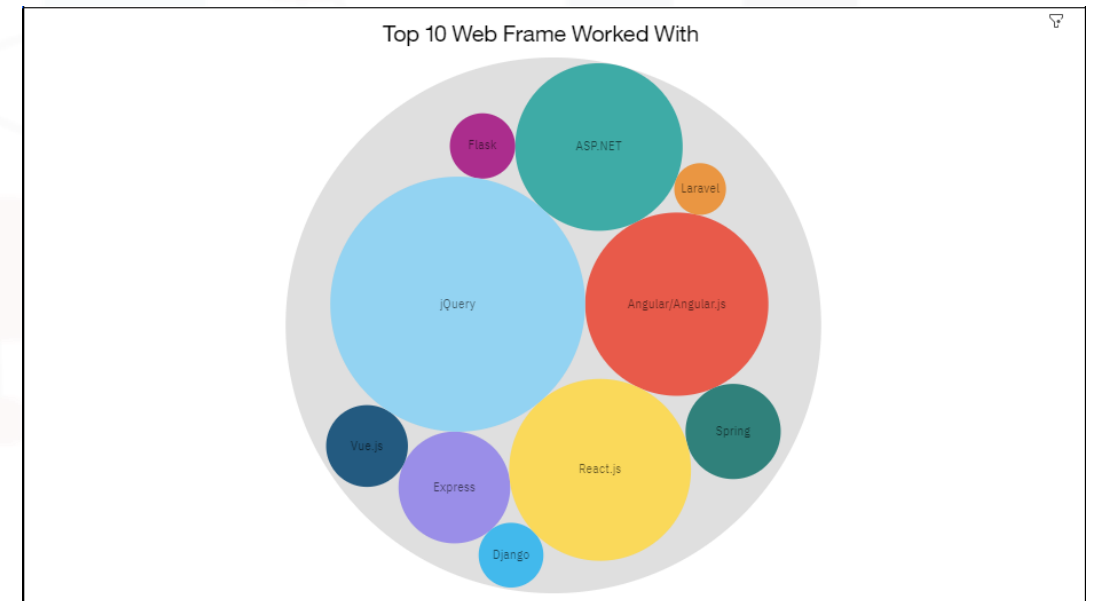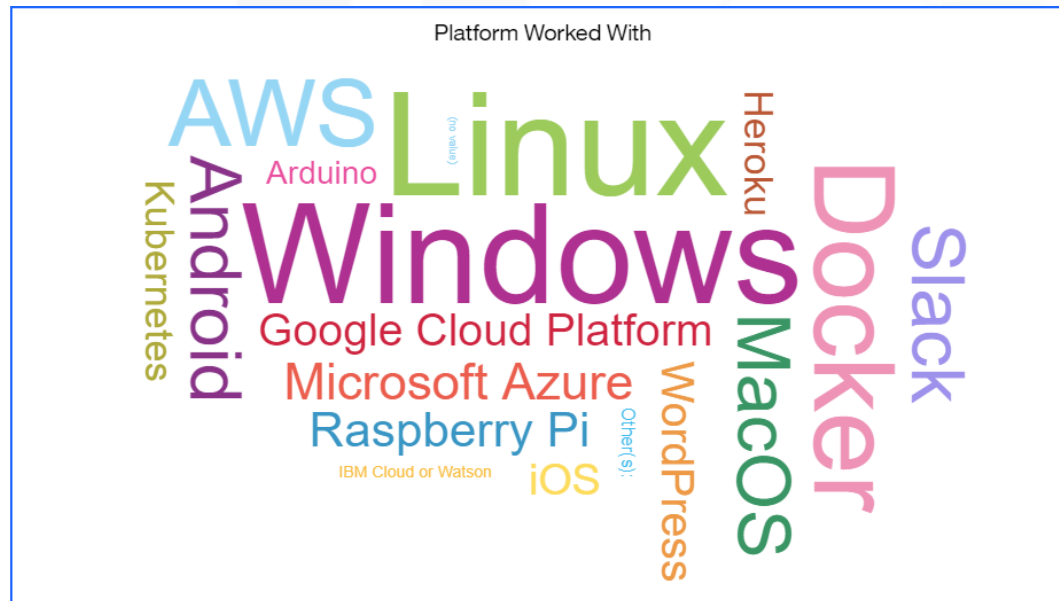
IBM **Dev** loper

SKILLS NETWORK

DASHBOARD

IBM Cognos Dashboard Link:

https://dataplatform.cloud.ibm.com/dashboards/93db381f-225d-4824-8d16-d477f609dfb8/view/771fd77f358b368d76ecc4e407ca7a542e32205cb2bb8152d5807b490e637897a83f14c7c87a490cd9420c31a0ea410ac0
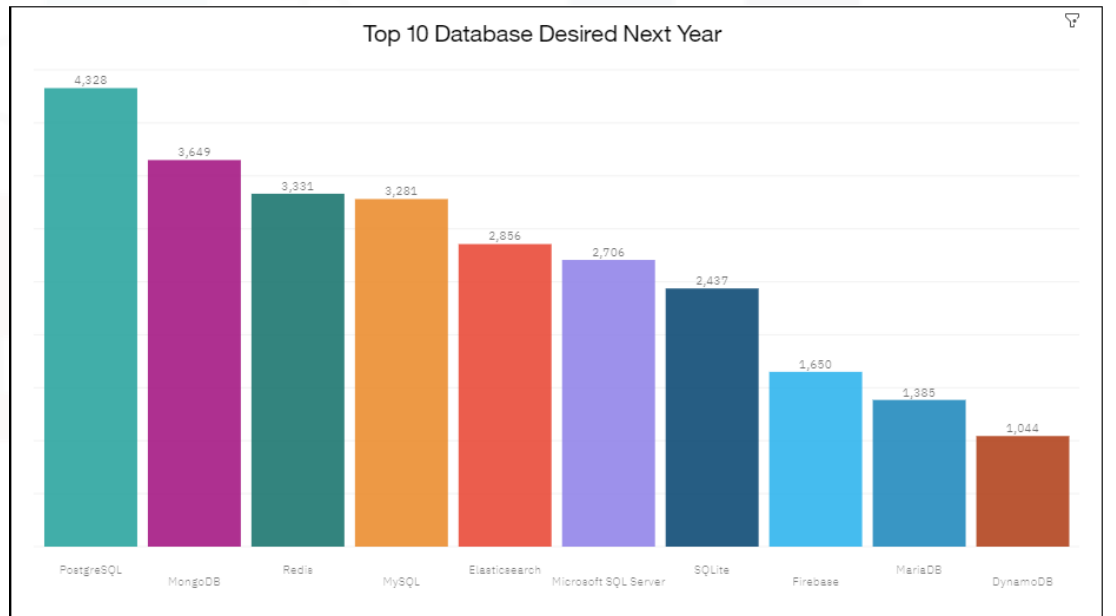
IBM **Developer**

SKILLS NETWORK

DASHBOARD – Current Technology Usage

Top 10 Language Worked With

- JavaScript: 8,687
- HTML/CSS: 7,830
- SQL: 7,106
- Bash/Shell/PowerShell: 4,642
- Python: 4,542
- Java: 4,506
- C#: 4,288
- TypeScript: 3,232
- PHP: 2,913
- C++: 1,946

Top 10 Database Worked With

- MySQL: 5,469
- Microsoft SQL Server: 4,110
- PostgreSQL: 4,097
- SQLite: 3,248
- MongoDB: 3,016
- Redis: 2,508
- Elasticsearch: 1,954
- Oracle: 1,744
- MariaDB: 1,709
- Firebase: 1,314

IBM Developer

SKILLS NETWORK

DASHBOARD – Current Technology Usage

Platform Worked With

Top 10 Web Frame Worked With

DASHBOARD – Future Technology Trend

Top 10 Language Desired Next Year

Language	Value
JavaScript	6,630
HTML/CSS	5,328
Python	5,239
SQL	5,012
TypeScript	4,088
C#	3,591
Bash/Shell/PowerShell	3,097
Java	2,953
Go	2,759
Kotlin	1,890

Top 10 Database Desired Next Year

Database	Value
PostgreSQL	4,328
MongoDB	3,649
Redis	3,331
MySQL	3,281
Elasticsearch	2,856
Microsoft SQL Server	2,706
SQLite	2,437
Firebase	1,650
MariaDB	1,385
DynamoDB	1,044

IBM Developer

SKILLS NETWORK

DASHBOARD – Future Technology Trend
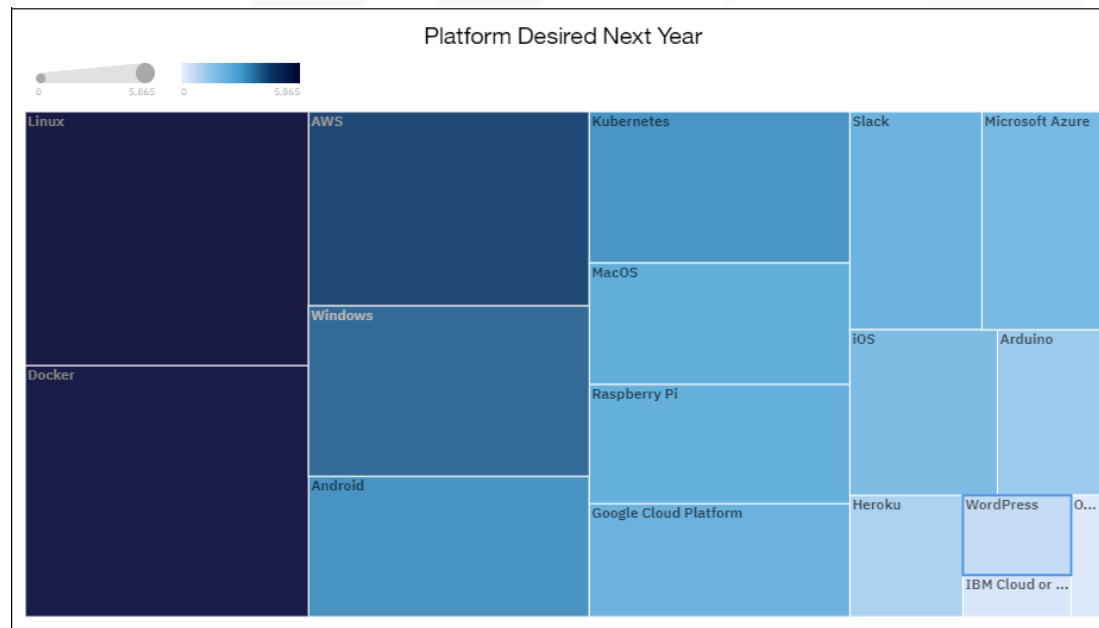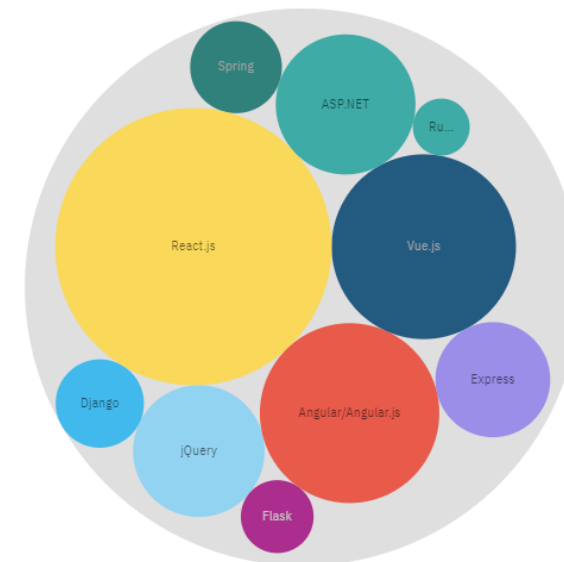
DASHBOARD - Demographics

DASHBOARD - Demographics

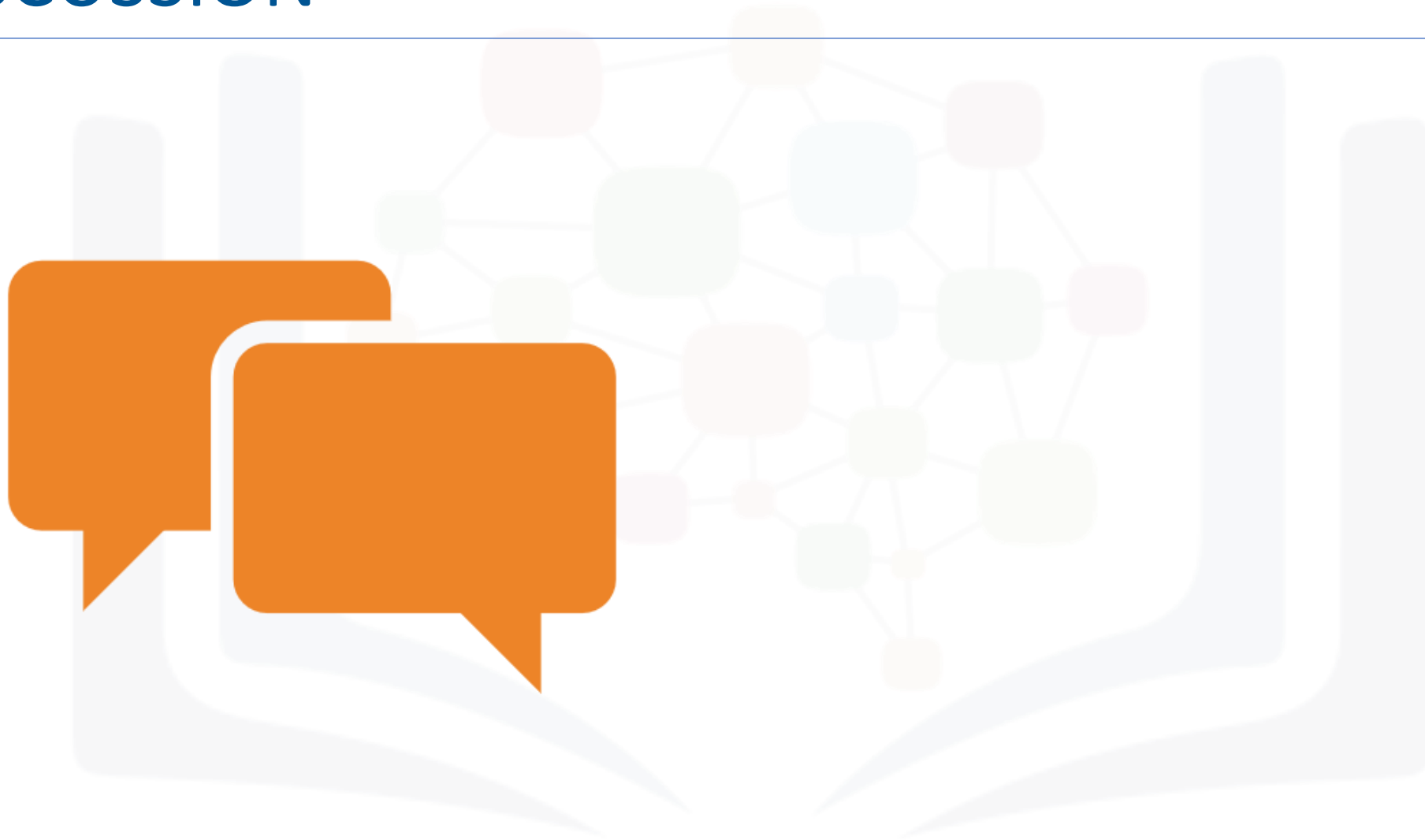Respondents Count for Countries

DASHBOARD - Demographics

Respondents Count by Age
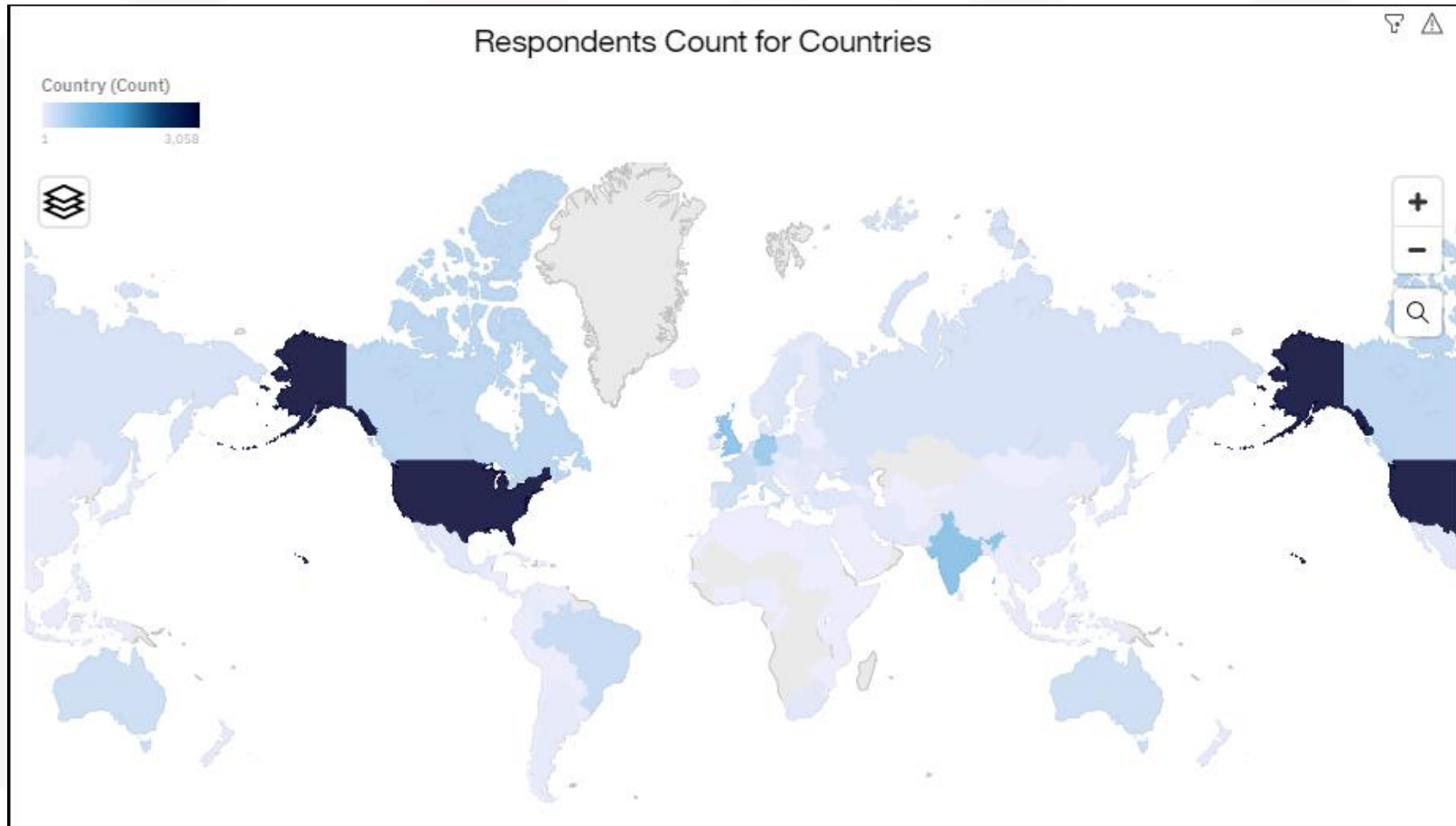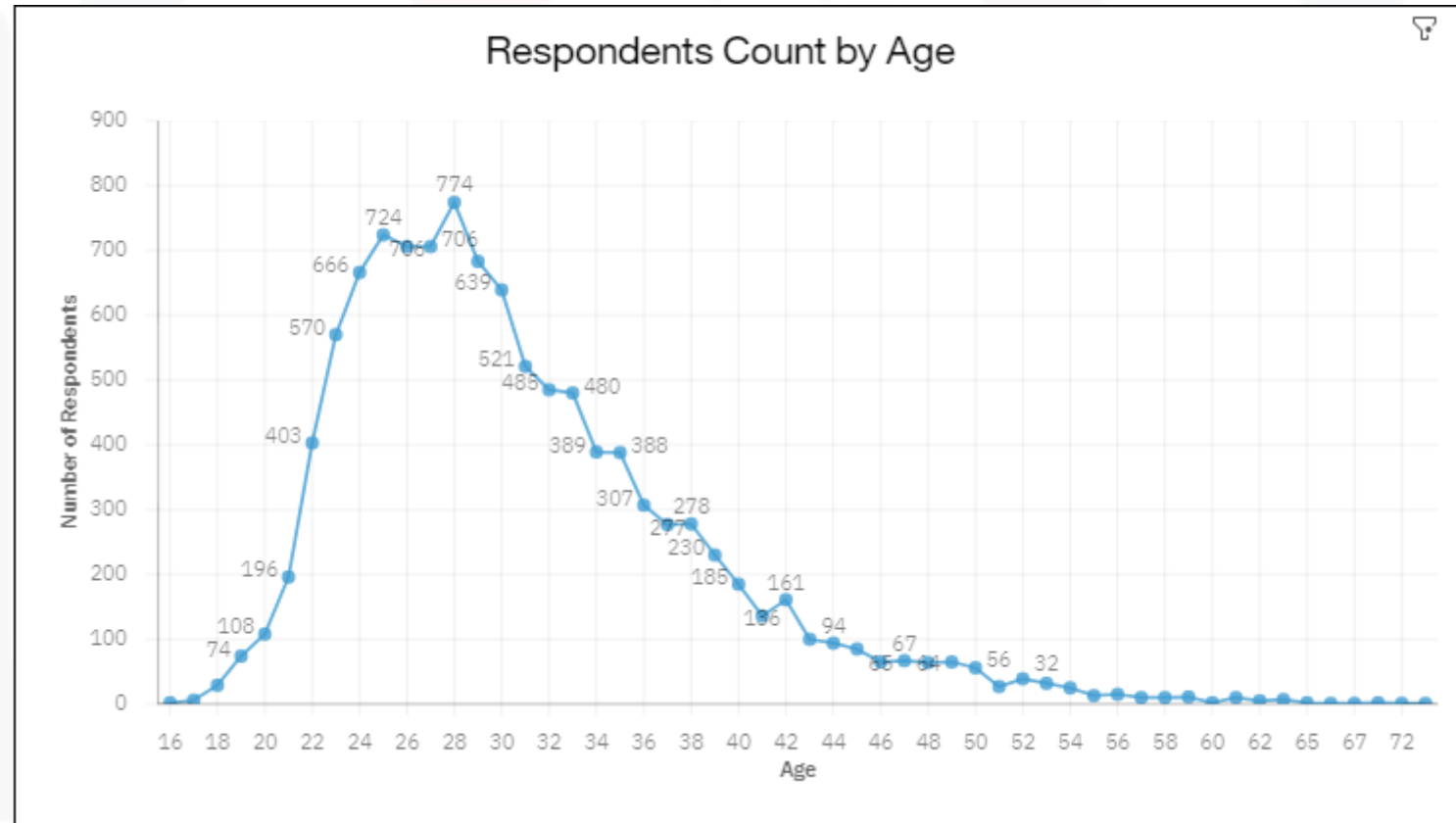
DISCUSSION

OVERALL FINDINGS & IMPLICATIONS

Findings

- Men forms a large proportion of coding industry

- Most developers are between 20 to 40 years old(especially 24~32)

- Most developers hold a Bachelor's or Master's degree
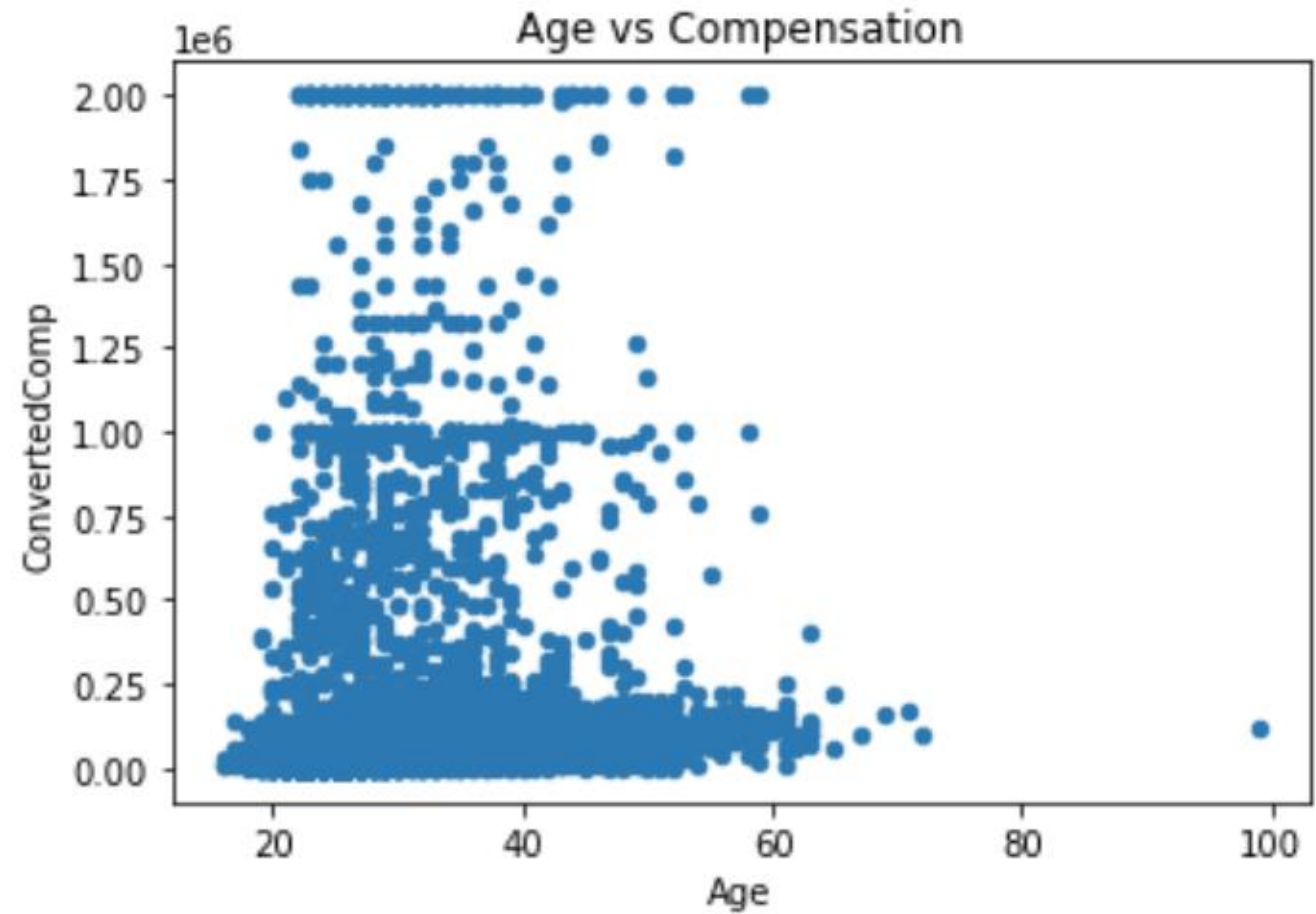
Implications

- Industry should make some efforts to balance the gender

- A good age for coding is less than 40 years old, step into management level after 40

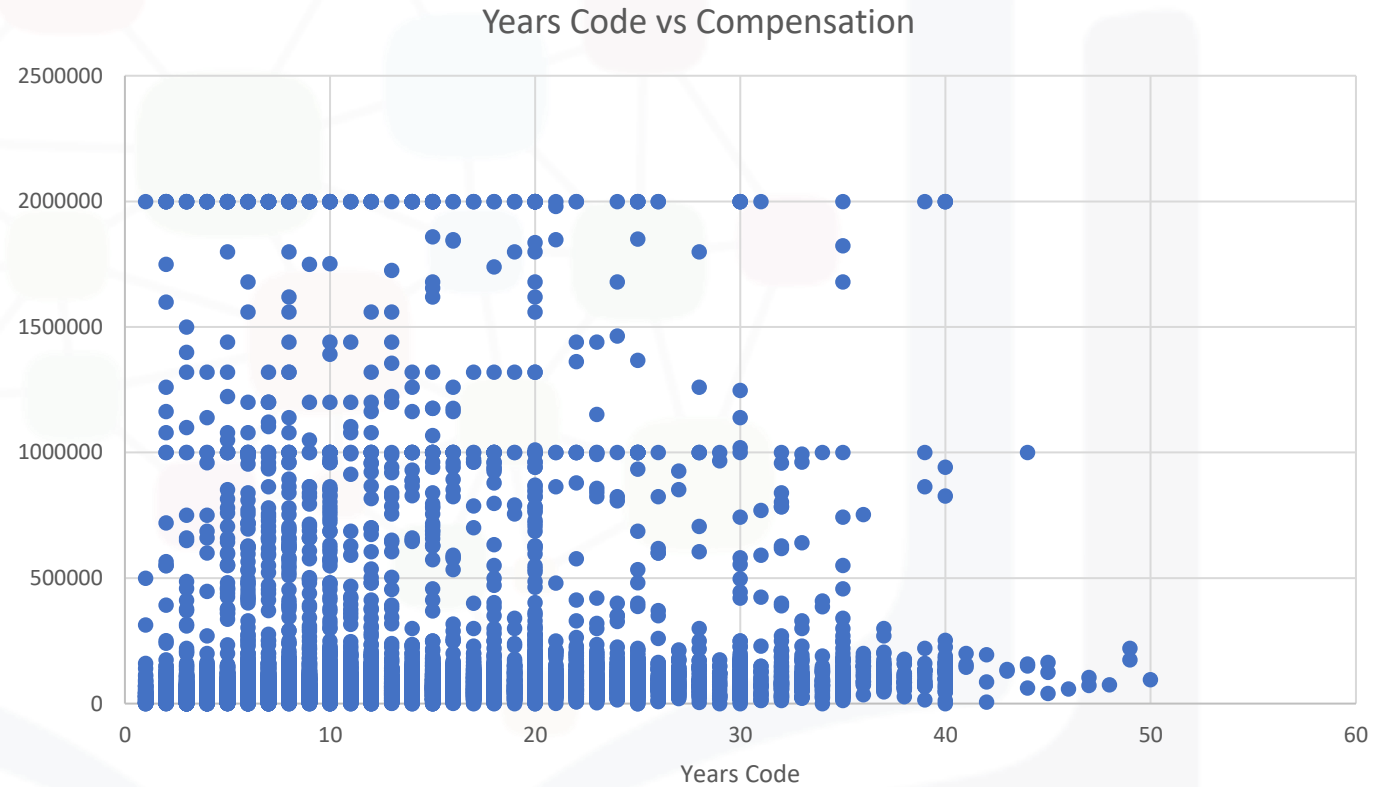**IBM Developer**

SKILLS NETWORK

CONCLUSION

- Python as an open-sourced, easy-to-use language is becoming more competitive

- Cloud based database is turning desirable because of efficiency and mobility

- No-SQL seems more important as data is not just in tabular form

- Windows and Linus are still the most important platform for development
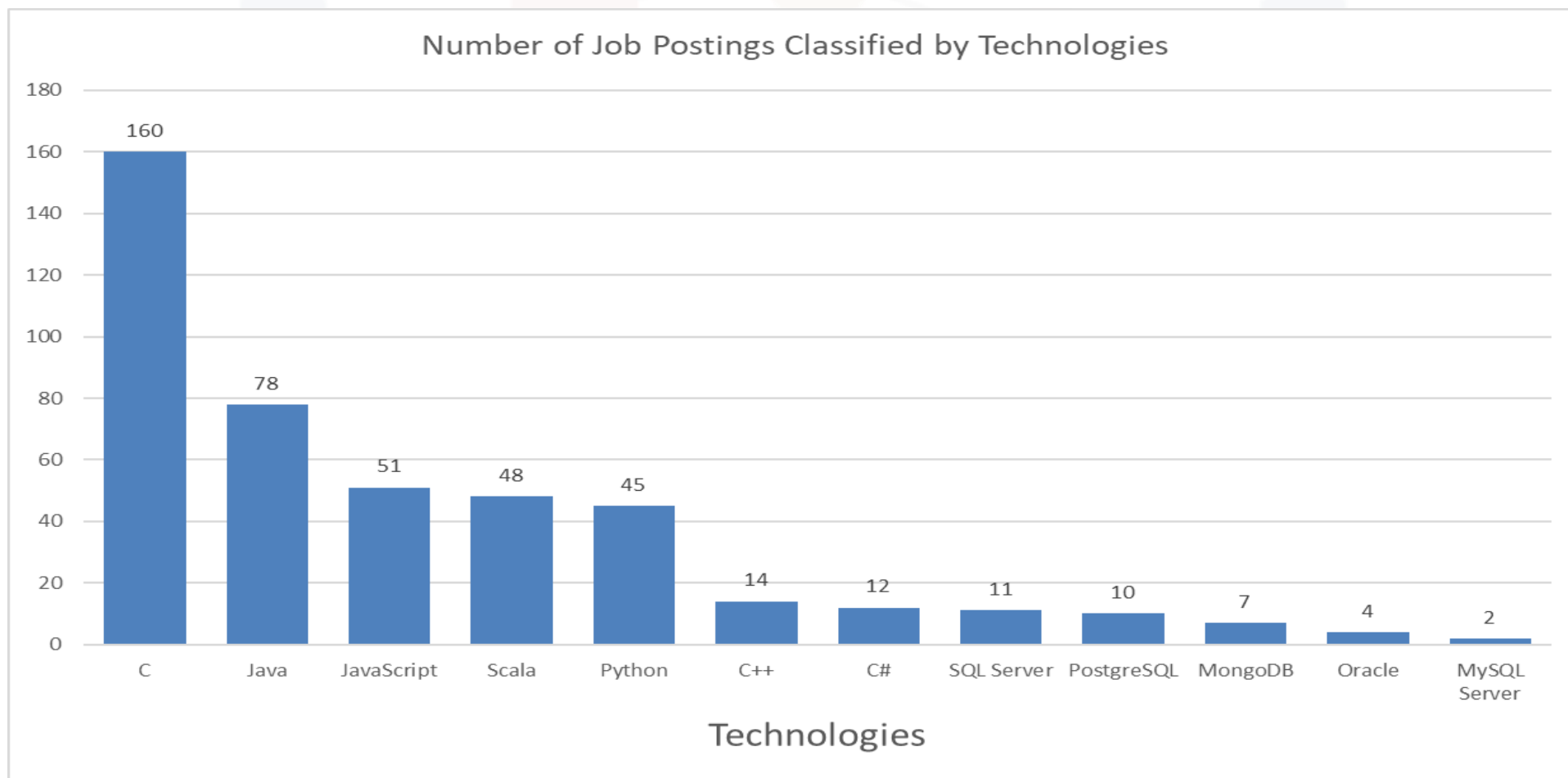
APPENDIX

Age vs Compensation

APPENDIX

Years Code vs Compensation

GITHUB JOB POSTINGS

Number of Job Postings Classified by Technologies

POPULAR LANGUAGES

Annual Average Salary By Programming Languages

Language	Salary
Swift	$130,801
Python	$114,383
C++	$113,865
Javascript	$110,981
Java	$101,013
Go	$94,082
R	$92,037
C#	$88,726
SQL	$84,793
PHP	$84,727

IBM Developer

SKILLS NETWORK