

# Data Analysis of developer survey taken in 2019 about changing technologies

Elisabetta De Vito-Francesco

14.12.2022

© IBM Corporation. All rights reserved.

A background image of a financial candlestick chart with a blue tint. The chart shows price movement over time with various price points labeled. A semi-transparent dark blue rectangle is centered over the chart, containing the word 'OUTLINE' in yellow.

# OUTLINE

© IBM Corporation. All rights reserved.

# OUTLINE

---



- Executive Summary
- Introduction
  - Problem
  - Questions for analysis
- Methodology
- Results
  - Answers to questions for analysis as Visualizations & Findings and implications
    - Most used vs most desired languages
    - Most used vs most desired databases
    - Developers count by age gender country
    - Compensation by age gender country
- Discussion
  - Findings & Implications
- Conclusion
- Appendix
  - Most used vs desired platforms
  - Most used vs desired web frames
  - Number of job postings
  - Popular languages
  - Work locations
- Dashboard



A background image of a financial candlestick chart with a blue overlay. The chart shows price movement over time with various numerical labels such as 4610.57, 4547.31, 4292.14, and 4487.06. A semi-transparent dark blue rectangle is centered over the chart.

# EXECUTIVE SUMMARY

© IBM Corporation. All rights reserved.

# EXECUTIVE SUMMARY

---



- **Problem:** data analysis to help identify future skill requirements
- **Questions analysis:** Which are the current technologies trends analysis? Which are the future technologies trends analysis? Which are the demographic analysis?
- **Methods:** Randomized subset of data from an online survey conducted by Stack Overflow website in 2019 (1/10<sup>th</sup> of the original size). Tools: Jupiter Notebooks, Python, SQL, IBM Cognos Analytics.
- **Results:** Visualization of top 5 current and future languages. Visualization of top 5 current and future databases. Respondents count by age, gender, country. Compensation by age, gender, country, languages.
- **Discussion:** Trends of most used vs most desired for next year languages. Trends of most used vs most desired for next year databases. Developers respondents characteristics by age gender country compensation.
- **Conclusions:** Languages to focus on: JavaScript HTML/CSS Python SQL. Databases to focus on: PostgreSQL MongoDB MySQL. Flatten the disparities between genders in number of developers. Flatten the disparities between genders in compensation of developers
- **Appendix:** Number of job postings per city in USA. Number of job postings per language in USA. Annual average compensation per language. Work location of developers in survey 2019. Most used vs desired for next year platforms. Most used vs desired for next year web frames.

A background image of a financial candlestick chart with a blue tint. The chart shows price fluctuations over time, with several data points labeled with numerical values. A semi-transparent dark blue rectangle is centered over the chart, containing the word 'INTRODUCTION' in yellow.

# INTRODUCTION

© IBM Corporation. All rights reserved.

# INTRODUCTION Problem

---



In order to keep pace with changing technologies and remain competitive, a regular data analysis to help identify future skill requirements must be performed.

# INTRODUCTION Questions for analysis

In order to keep pace with changing technologies and remain competitive, a regular data analysis to help identify future skill requirements must be performed.



- Which are the most used technologies by developers in 2019?
  - Most used languages
  - Most used databases
  - Most used platforms (Appendix)
  - Most used web frames (Appendix)
- Which are the most desired technologies by developers to be learnt for the next year?
  - Most used languages
  - Most used databases
  - Most used platforms (Appendix)
  - Most used web frames (Appendix)
- Are there any differences in numbers and compensation between developers based on age and origin?
  - How much do developers earn?
  - How old are developers?
  - Where are developers from?

Current technologies trends analysis

Future technologies trends analysis

Demographic analysis



# INTRODUCTION Questions for analysis

In order to keep pace with changing technologies and remain competitive, a regular data analysis to help identify future skill requirements must be performed.



- Which are the most used technologies by developers in 2019?
  - Most used languages
  - Most used databases
  - Most used platforms (Appendix)
  - Most used web frames (Appendix)
- Which are the most desired technologies by developers to be learnt for the next year?
  - Most used languages
  - Most used databases
  - Most used platforms (Appendix)
  - Most used web frames (Appendix)
- Are there any differences in numbers and compensation between developers based on age and origin?
  - How much do developers earn?
  - How old are developers?
  - Where are developers from?



# METHODOLOGY

© IBM Corporation. All rights reserved.

# METHODOLOGY Dataset and tools



## Dataset

- Data from an online survey conducted by Stack Overflow website in 2019.

<https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/> under a **ODbL: Open Database License**

- Randomized subset of the original dataset as 1/10<sup>th</sup> of the original size.

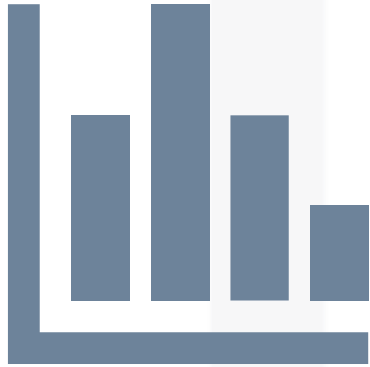
[https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1\\_survey\\_data.csv?utm\\_medium=Exinfluencer&utm\\_source=Exinfluencer&utm\\_content=000026UJ&utm\\_term=10006555&utm\\_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDA0321ENSkillsNetwork21426264-2022-01-01](https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=10006555&utm_id=NA-SkillsNetwork-Channel-SkillsNetworkCoursesIBMDA0321ENSkillsNetwork21426264-2022-01-01)

## Tools

- Jupyter notebooks → Analysis, RDMS connection, visuals
- Python → Pandas for analysis, Matplotlib for visuals
- SQL → Queries and analysis
- IBM Cognos Analytics → Dashboard visuals

# METHODOLOGY Analysis and visuals

---



- Current technologies trends for developers (2019)
  - Most used languages
  - Most used databases
  - Most used platforms (Appendix)
  - Most used web frames (Appendix)
- vs
- Future technologies trends for developers (2020)
  - Most desired languages
  - Most desired databases
  - Most desired platforms (Appendix)
  - Most desired web frames (Appendix)
- Demographic analysis
  - Disparities between genders and age for developers
    - Number of developer per gender
    - Compensation per gender
    - Number of developer per age
    - Compensation per age
    - Number of developer per origin country



A background image of a financial candlestick chart with a blue tint. The chart shows price movement over time with various price points labeled. A semi-transparent dark blue rectangle is centered over the chart, containing the word 'RESULTS' in yellow.

# RESULTS

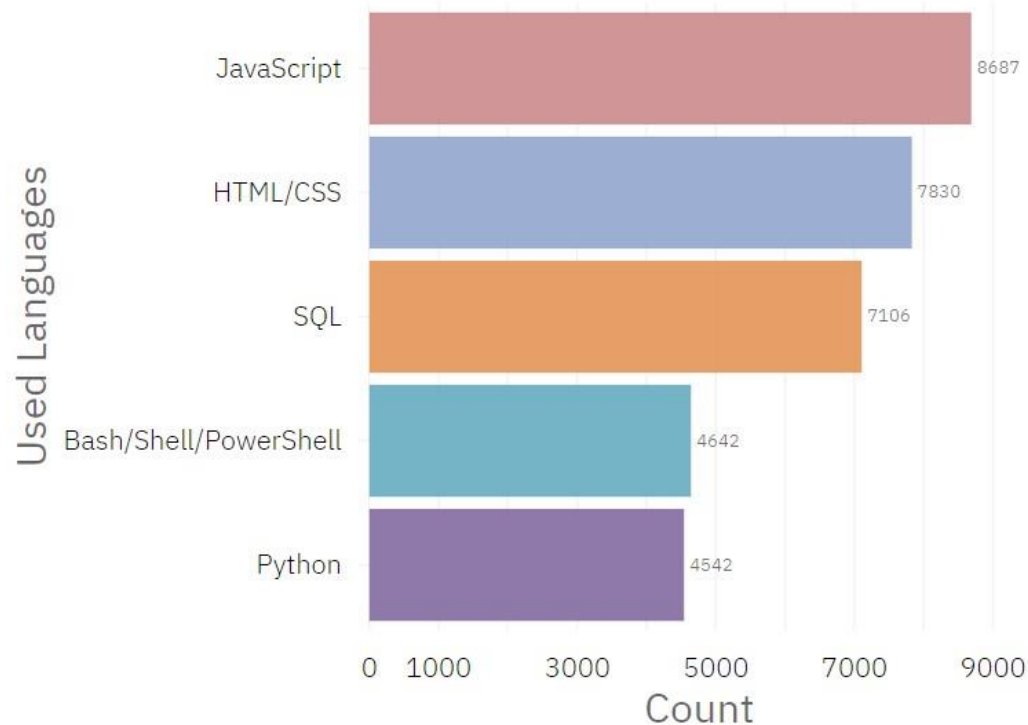
© IBM Corporation. All rights reserved.

# VISUALIZATION

## Programming language trends

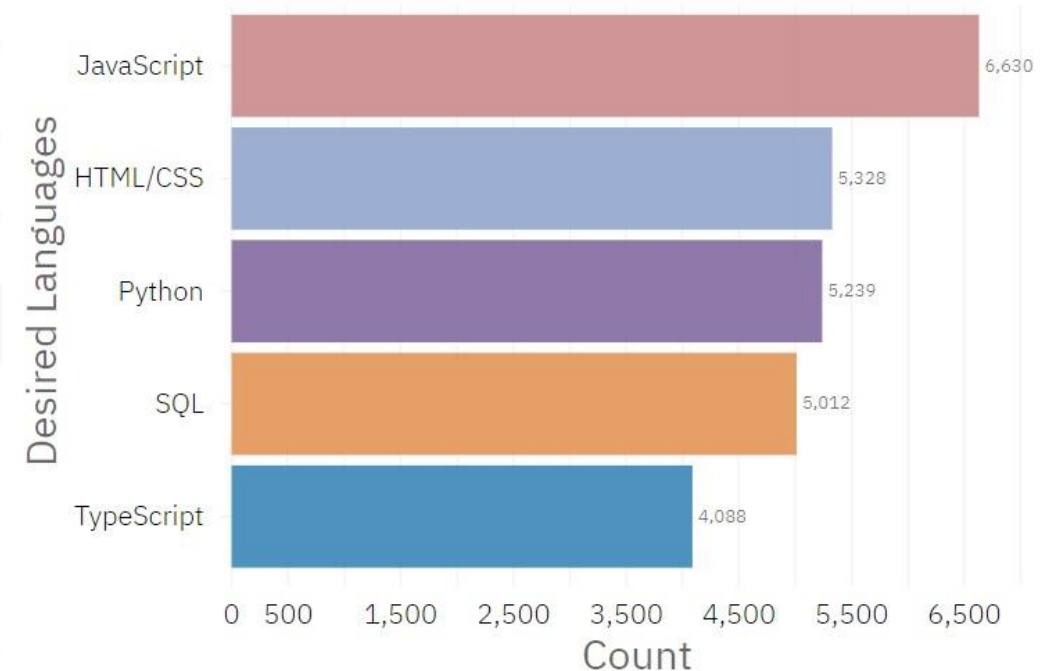
### Current Year

Top 5 most used languages by developers in 2019



### Next Year

Top 5 most desired languages for next year by developers in 2019



# FINDINGS & IMPLICATIONS

## Programming language trends

---

### Findings

- JavaScript and HTML/CSS are both the most used and most desired for next year languages
- Python is the 3<sup>rd</sup> most wanted language to be learnt for the next year (5<sup>th</sup> of most used language)
- SQL was the 3<sup>rd</sup> most used language, but decreased to 4<sup>th</sup> most desired language to be learnt
- TypeScript replaces Bash/Shell/PowerShell in the top 5 desired languages

### Implications

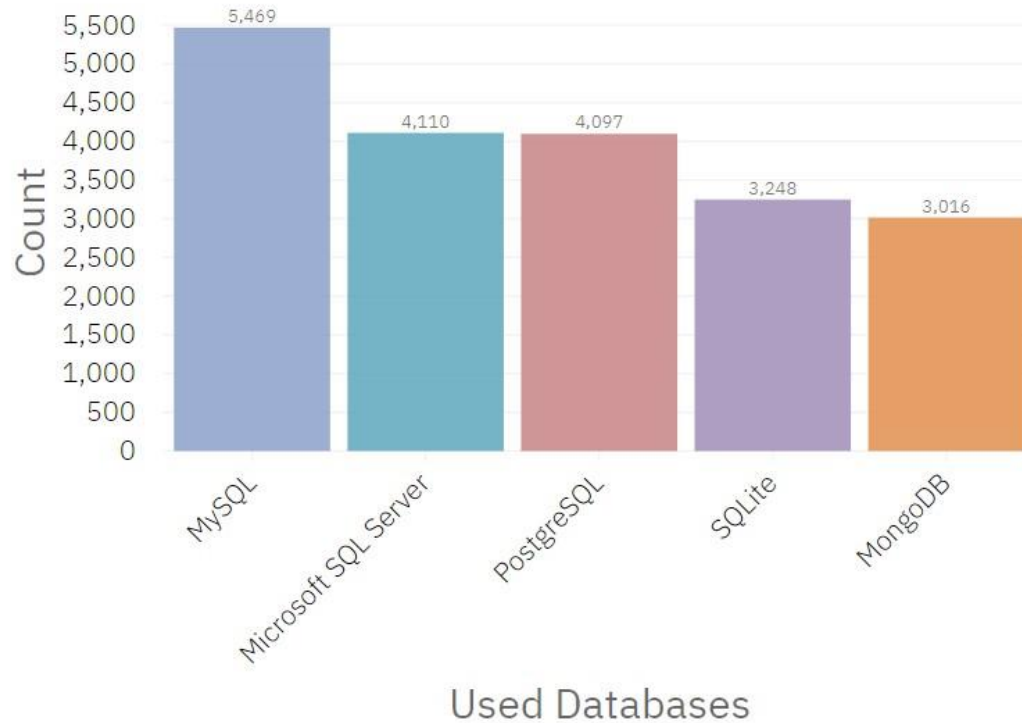
- JavaScript and HTML/CSS continue to be languages to be learnt
- Increase interest in the Opensource Python language
- Slightly decrease interest in the SQL language
- Give more focus on TypeScript rather than Bash/Shell/PowerShell

# VISUALIZATION

## Database trends

### Current Year

Top 5 most used databases by developers in 2019



### Next Year

Top 5 most desired databases for next year by developers





# FINDINGS & IMPLICATIONS

## Programming language trends

---

### Findings

- PostgreSQL and MongoDB are the most desired for next year DB, gaining better positions from the most used list
- MySQL is the most used DB and still in the top 5 most desired for next year
- Microsoft SQL server 2<sup>nd</sup> most used DB but not in the top 5 most desired for next year
- SQLite in the top 5 most used DB but not in the top 5 most desired for next year

### Implications

- Give more focus on PostgreSQL and MongoDB as DB
- Keep focus on MySQL DB
- Give more focus to Redis and Elasticsearch rather than Microsoft SQL and SQLite

# VISUALIZATION

## Demographics respondents

---

Total respondent developers

11398

Developers respondents

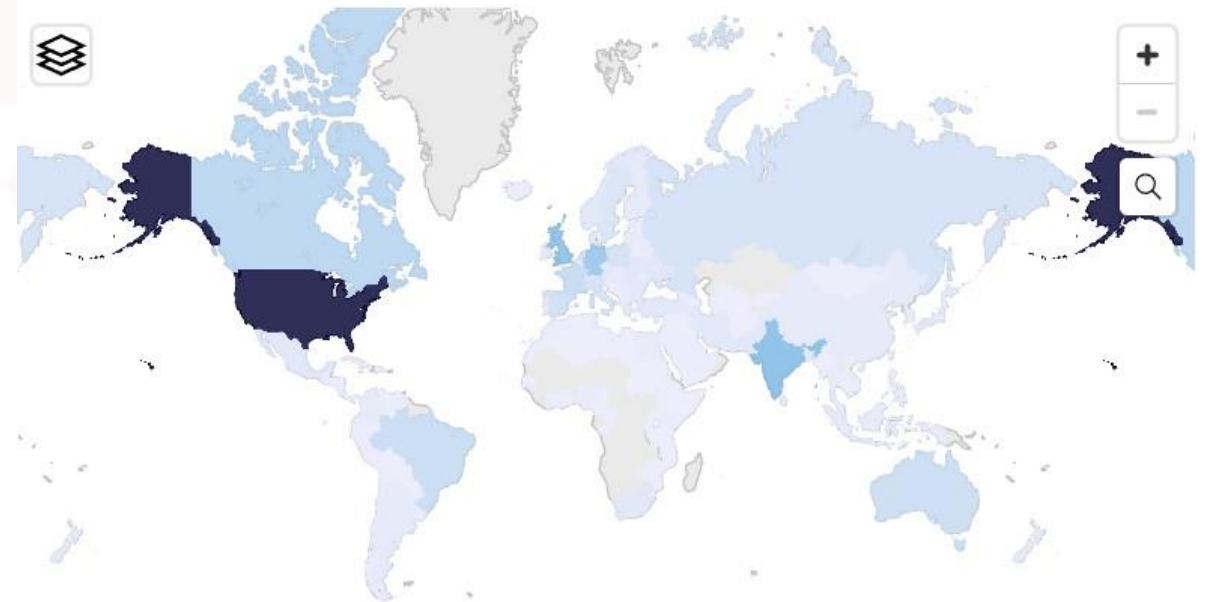
Respondent developers by country

Number of developers respondents count by country

Country (Count)



1 3,127

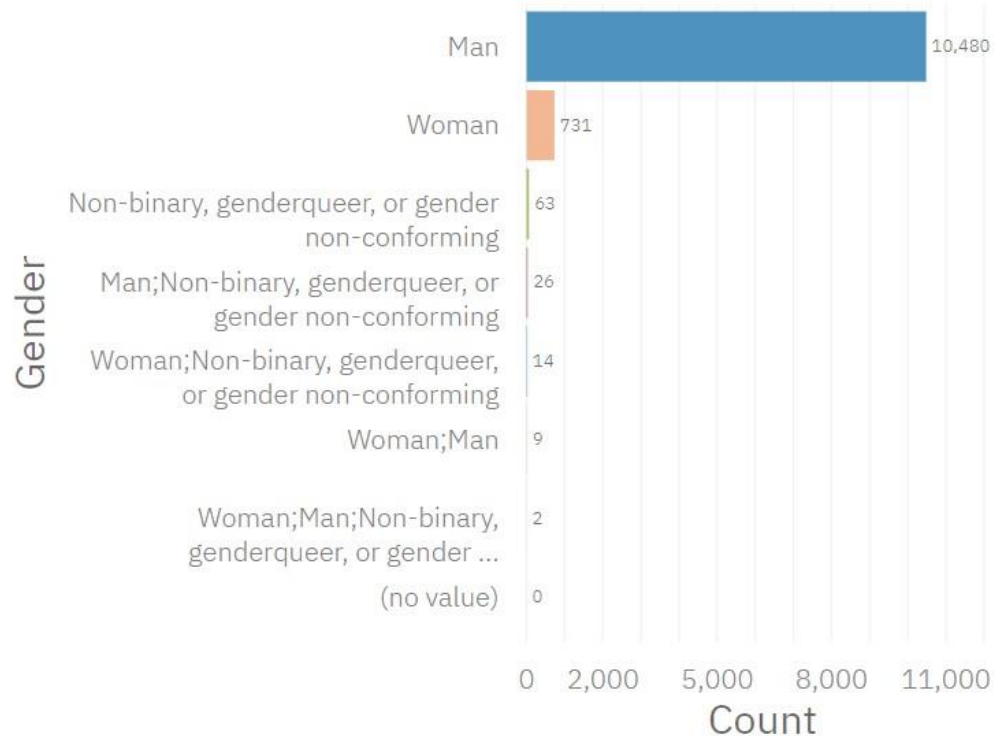


# VISUALIZATION

## Demographics Age and Gender

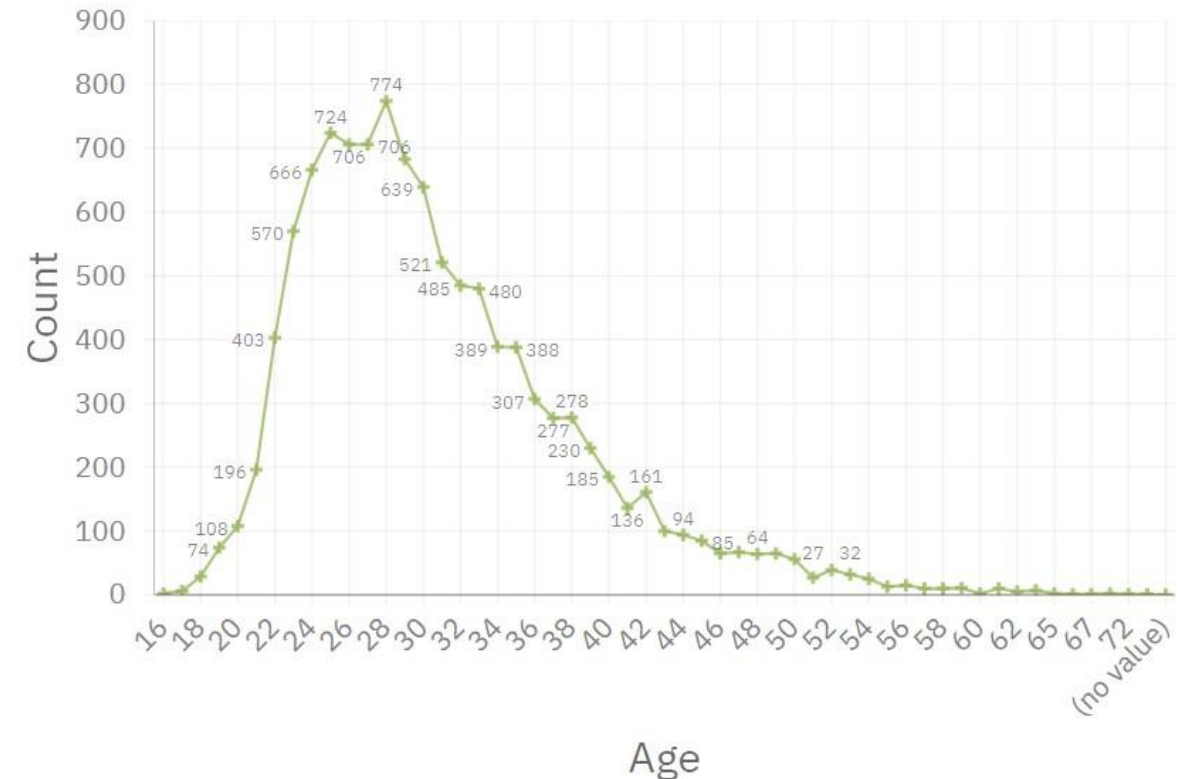
### Genders of respondent developers

Number of developers respondents count by gender in 2019



### Age of respondent developers

Number of respondents count by age in 2019



# FINDINGS & IMPLICATIONS

## Demographics Age and Gender

---

### Findings

- 92.5% of respondent developers are Men (10480)
- 6.5% of respondent developers are Women (731)
- Majority of respondents were in the 20-40 age range
- Age peak in 25-28 range

### Implications

- Not many Women developers
- Either Stock Overflow is mainly known and used by young age range
- Either developer are of young age



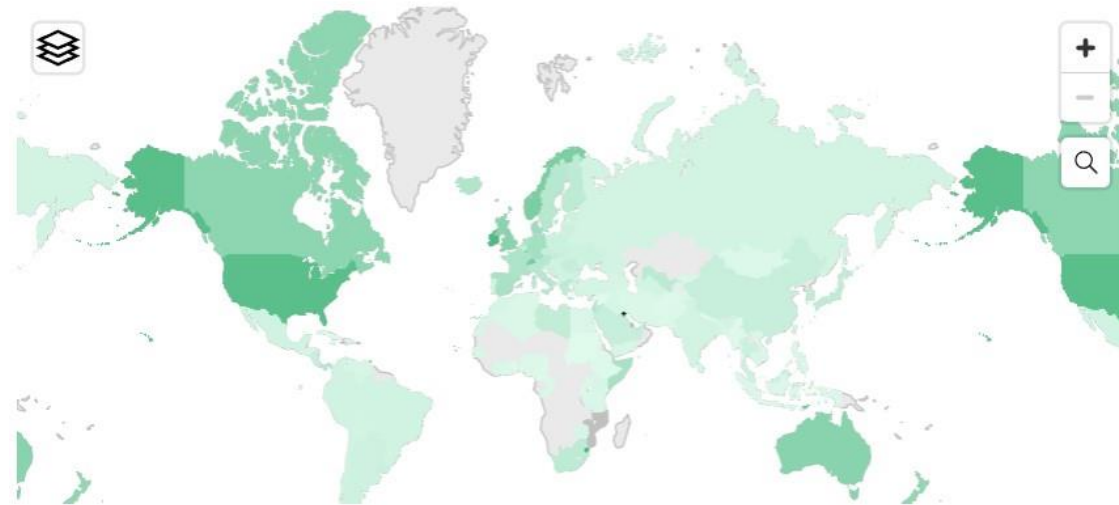
# VISUALIZATION

## Demographics Compensation by country

### Compensation (mean) by country

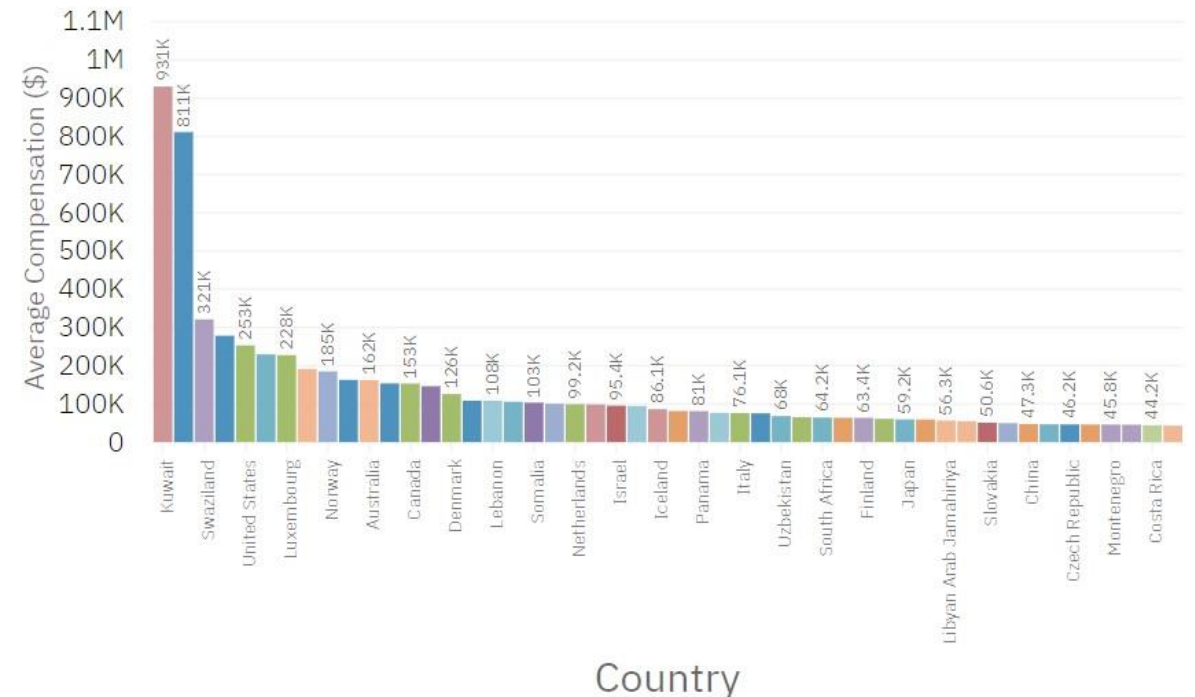
Average compensation of developers respondent by country in 2019

ConvertedComp (...)



### Compensation (mean) by country top 50

Average Compensation of top 50 countries for developers respondents in 2019

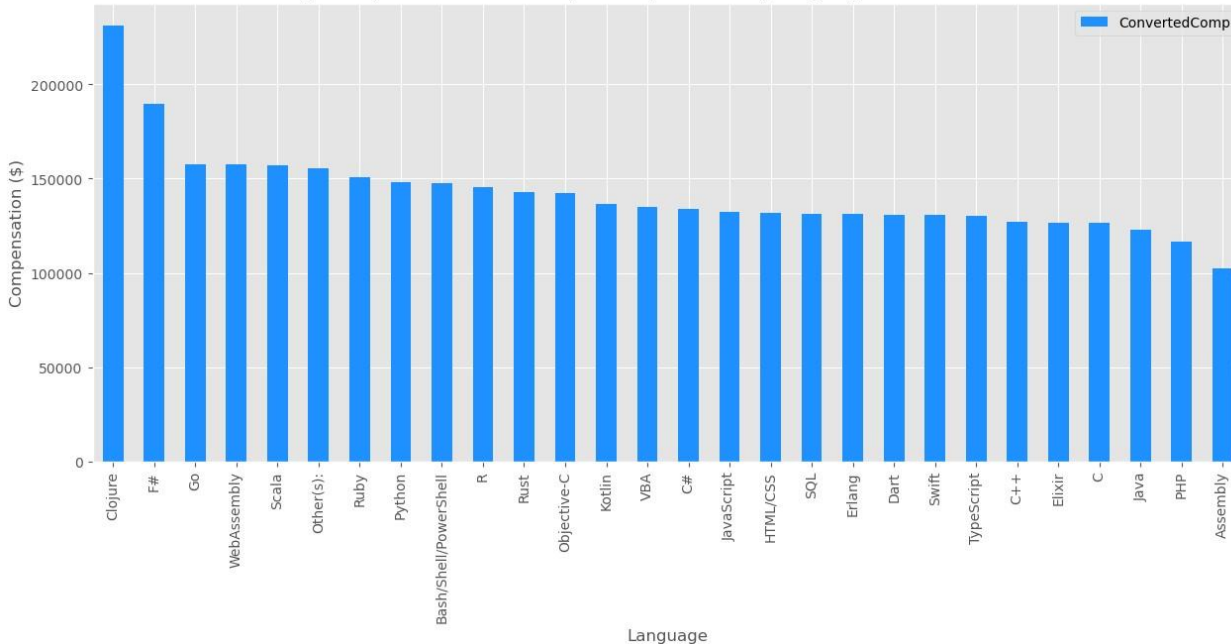


# VISUALIZATION

## Demographics Compensation by age, gender and most used language

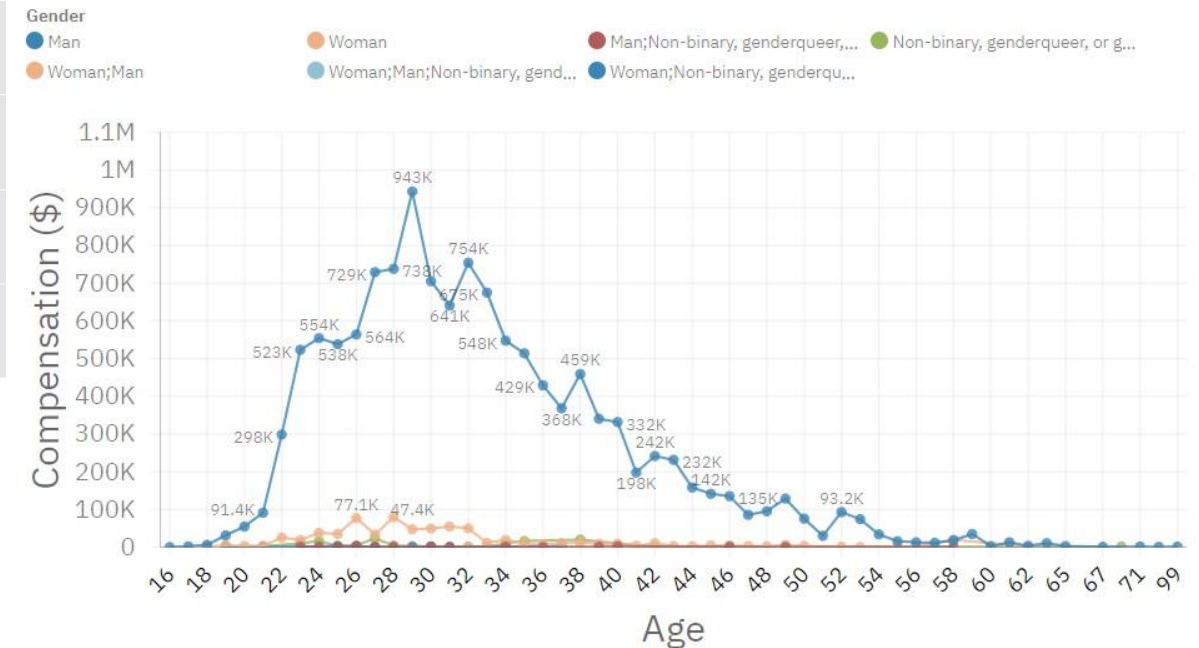
### Compensation (mean) by language

Average compensation of developers respondents by language worked with in 2019



### Compensation by age and genders

Compensation (\$) of respondent developers by age and gender groups in 2019



# FINDINGS & IMPLICATIONS

## Demographics Compensation by country age and gender

---

### Findings

- Top 3 countries with the highest average annual compensation are Kuwait, Switzerland, USA
- Most paid developers are men
- Most paid developers are between 25 - 35
- Big disparities between men and women compensation, for each age group
- Highest average compensations to developers who used Conjure and F#. Not deep disparities between the other languages

### Implications

- Not enough respondents from the other countries
- Not enough female respondents
- Similar compensation by language used

# DASHBOARD Link

---



<https://eu-de.dataplatform.cloud.ibm.com/dashboards/0a1df9d0-958b-4f64-baa4-ad920440ebc4/view/0365e1193eaa32d067d3eee407c829047e61775db4bbd10582817b4909627297f33f15c4c8281d0ed3190532f4bf165fc0>

<https://eu-de.dataplatform.cloud.ibm.com/dashboards/ab22463e-d212-4738-b5e3-f671302f8b23/view/053cdc7760ea2bca76fec8e407c829047e61775db4bbd10582817b4909627297f33f15c4c8281d0ed3190532f4bf165fc0>



# DASHBOARD Tab 1 Current technology

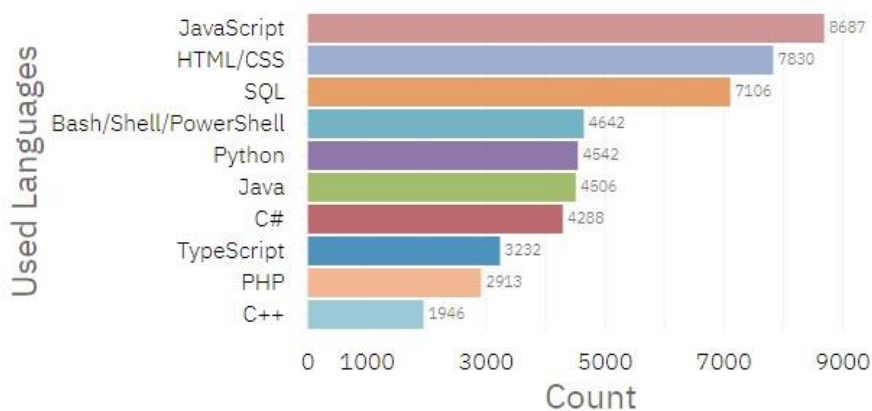
### Current Technology Usage

### Future Technology Trend

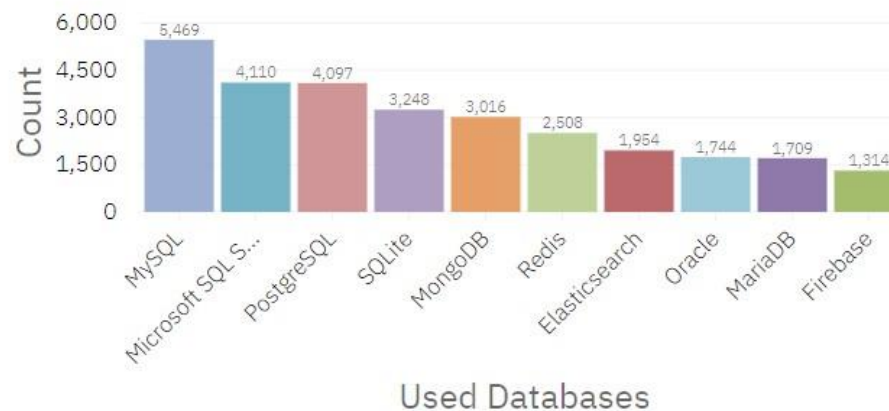
## Demographics



## Top 10 most used languages by developers in 2019



## Top 10 most used databases by developers in 2019



## Most used platforms by developers in 2019



## Top 10 most used Web Frame by developers in 2019



# DASHBOARD Tab 2 Future technology

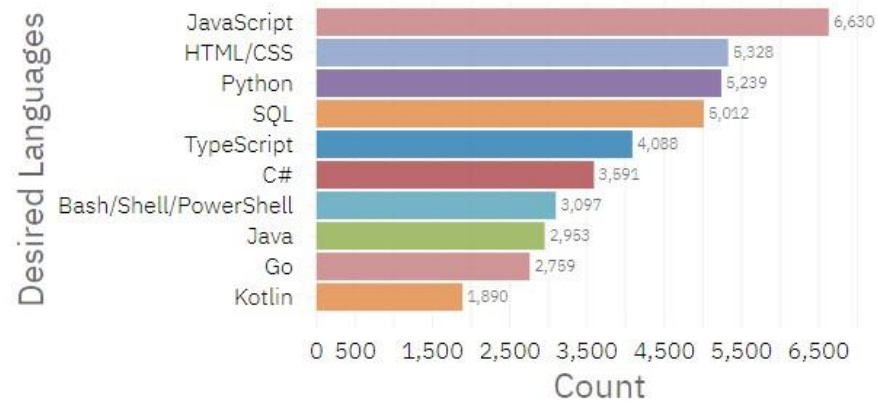
Current Technology Usage

Future Technology Trend

Demographics



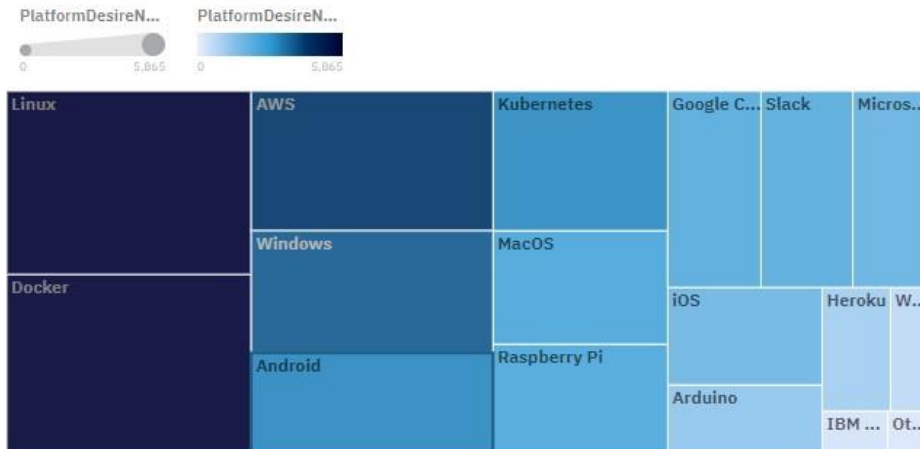
Top 10 most desired languages by developers for next year



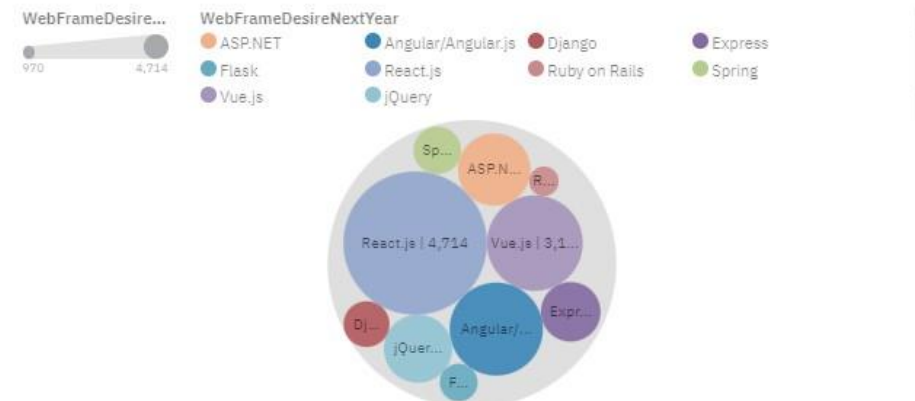
Top 10 most desired databases by developers for next year



Most desired platforms by developers for next year



Top 10 most desired Web Frame by developers for next year



# DASHBOARD Tab 3 Demographics

Current Technology Usage

Future Technology Trend

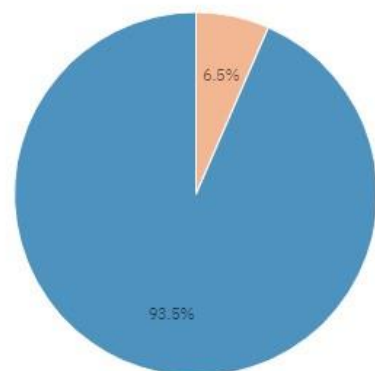
Demographics



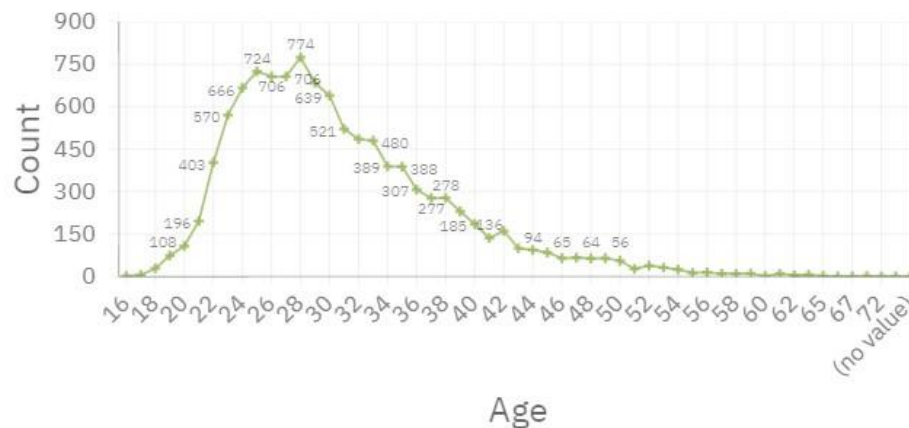
Number of respondents classified by gender in 2019

Gender

Woman Man



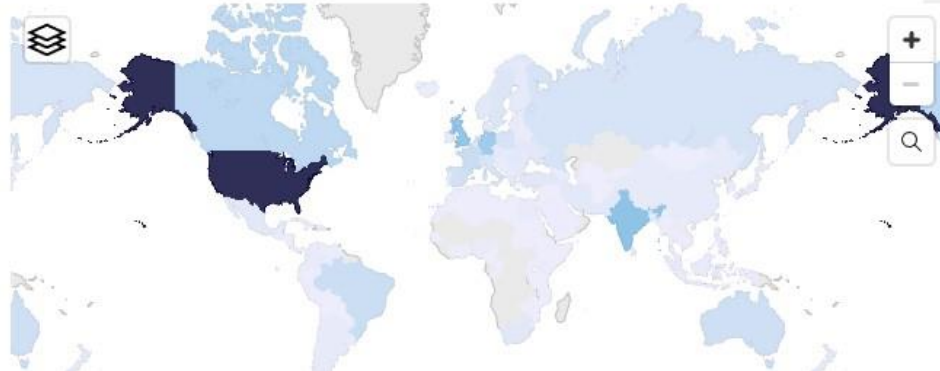
Number of respondents count by age in 2019



Number of respondents count by country in 2019

Country (Count)

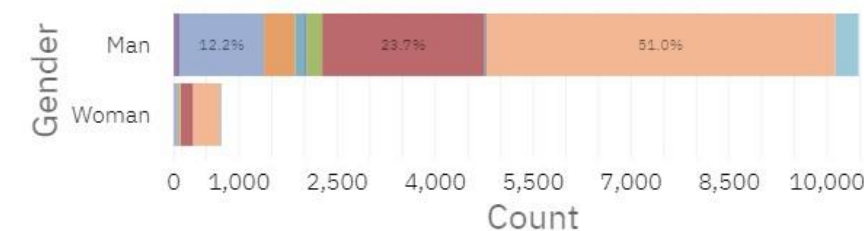
1 3,058



Number of respondents count by gender and classified by education level in 2019

EdLevel

(no value) Some college/university study ... Secondary school (e.g. Americ...  
Professional degree (JD, MD, e... Primary/elementary school Other doctoral degree (Ph.D, E...  
Master's degree (MA, MS, M.E... I never completed any formal ... Bachelor's degree (BA, BS, B.E...  
Associate degree



A background image of a financial candlestick chart with a blue tint. The chart shows price fluctuations over time, with several data points labeled with numerical values. A semi-transparent dark blue rectangle is centered over the chart, containing the word 'DISCUSSION' in yellow.

# DISCUSSION

© IBM Corporation. All rights reserved.



# DISCUSSION

---



- Trends of most used vs most desired for next year languages
- Trends of most used vs most desired for next year databases
- Developers respondents characteristics by age gender country compensation



# FINDINGS & IMPLICATIONS

## Programming language trends

---

### Findings

- JavaScript and HTML/CSS are both the most used and most desired for next year languages
- Python is the 3<sup>rd</sup> most wanted language to be learnt for the next year (5<sup>th</sup> of most used language)
- SQL was the 3<sup>rd</sup> most used language, but decreased to 4<sup>th</sup> most desired language to be learnt
- TypeScript replaces Bash/Shell/PowerShell in the top 5 desired languages

### Implications

- JavaScript and HTML/CSS continue to be languages to be learnt
- Increase interest in the Opensource Python language
- Slightly decrease interest in the SQL language
- Give more focus on TypeScript rather than Bash/Shell/PowerShell

# FINDINGS & IMPLICATIONS

## Programming language trends

---

### Findings

- PostgreSQL and MongoDB are the most desired for next year DB, gaining better positions from the most used list
- MySQL is the most used DB and still in the top 5 most desired for next year
- Microsoft SQL server 2<sup>nd</sup> most used DB but not in the top 5 most desired for next year
- SQLite in the top 5 most used DB but not in the top 5 most desired for next year

### Implications

- Give more focus on PostgreSQL and MongoDB as DB
- Keep focus on MySQL DB
- Give more focus to Redis and Elasticsearch rather than Microsoft SQL and SQLite

# FINDINGS & IMPLICATIONS

## Demographics Age and Gender

---

### Findings

- 92.5% of respondent developers are Men (10480)
- 6.5% of respondent developers are Women (731)
- Majority of respondents were in the 20-40 age range
- Age peak in 25-28 range

### Implications

- Not many Women developers
- Either Stock Overflow is mainly known and used by young age range
- Either developer are of young age

# FINDINGS & IMPLICATIONS

## Demographics Compensation by country age and gender

---

### Findings

- Top 3 countries with the highest average annual compensation are Kuwait, Switzerland, USA
- Most paid developers are men
- Most paid developers are between 25 - 35
- Big disparities between men and women compensation, for each age group
- Highest average compensations to developers who used Conjure and F#. Not deep disparities between the other languages

### Implications

- Not enough respondents from the other countries
- Not enough female respondents
- Similar compensation by language used

A blurred background image of a financial candlestick chart. The chart features white candlesticks on a dark blue grid. A white line graph is overlaid on the chart, showing an upward trend. Several numerical values are visible on the chart, including 4610.57, 4547.31, 4292.14, and 4487.06. The overall image has a blue tint.

# CONCLUSIONS

© IBM Corporation. All rights reserved.



# CONCLUSION

---



- Languages to focus on: JavaScript HTML/CSS  
Python SQL
- Databases to focus on: PostgreSQL  
MongoDB MySQL
- Flatten the disparities between genders in  
number of developers
- Flatten the disparities between genders in  
compensation of developers

A background image of a financial candlestick chart with a blue tint. The chart shows price fluctuations over time, with several data points labeled with numerical values. A semi-transparent dark blue rectangle is centered over the chart, containing the word 'APPENDIX' in yellow.

# APPENDIX

© IBM Corporation. All rights reserved.

# APPENDIX

---



- Most used vs desired for next year platforms
- Most used vs desired for next year web frames
- Number of job postings per city in USA
- Number of job postings per language in USA
- Annual average compensation per language
- Work location of developers in survey 2019

# VISUALIZATION

## Platforms trends

### Current Year

Most used platforms by developers in 2019



### Next Year

Most desired platforms by developers for next year

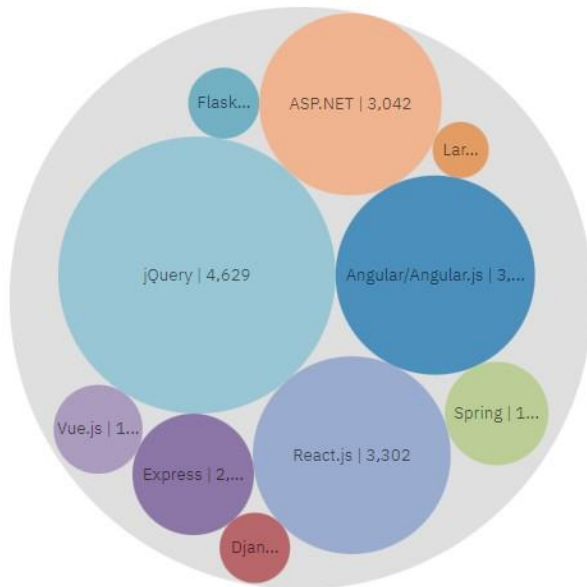
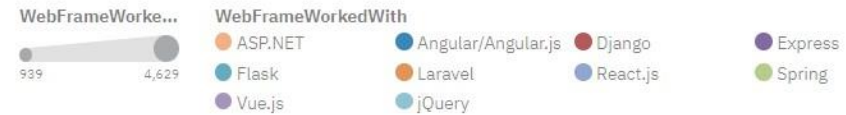


# VISUALIZATION

## Web frames trends

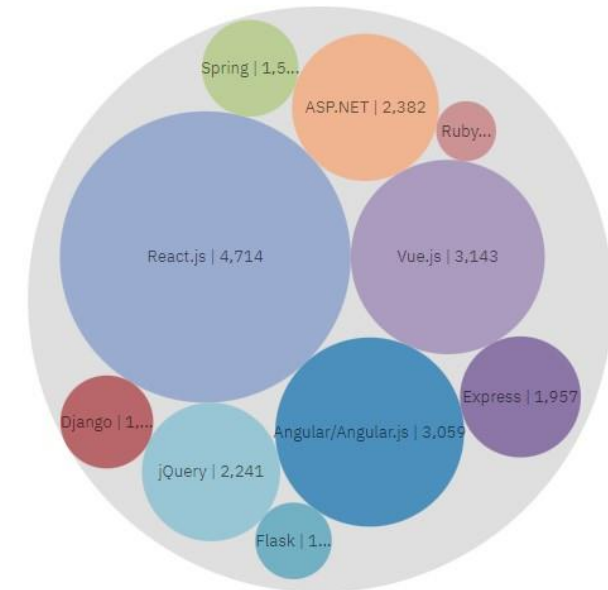
### Current Year

Top 10 most used Web Frame by developers in 2019



### Next Year

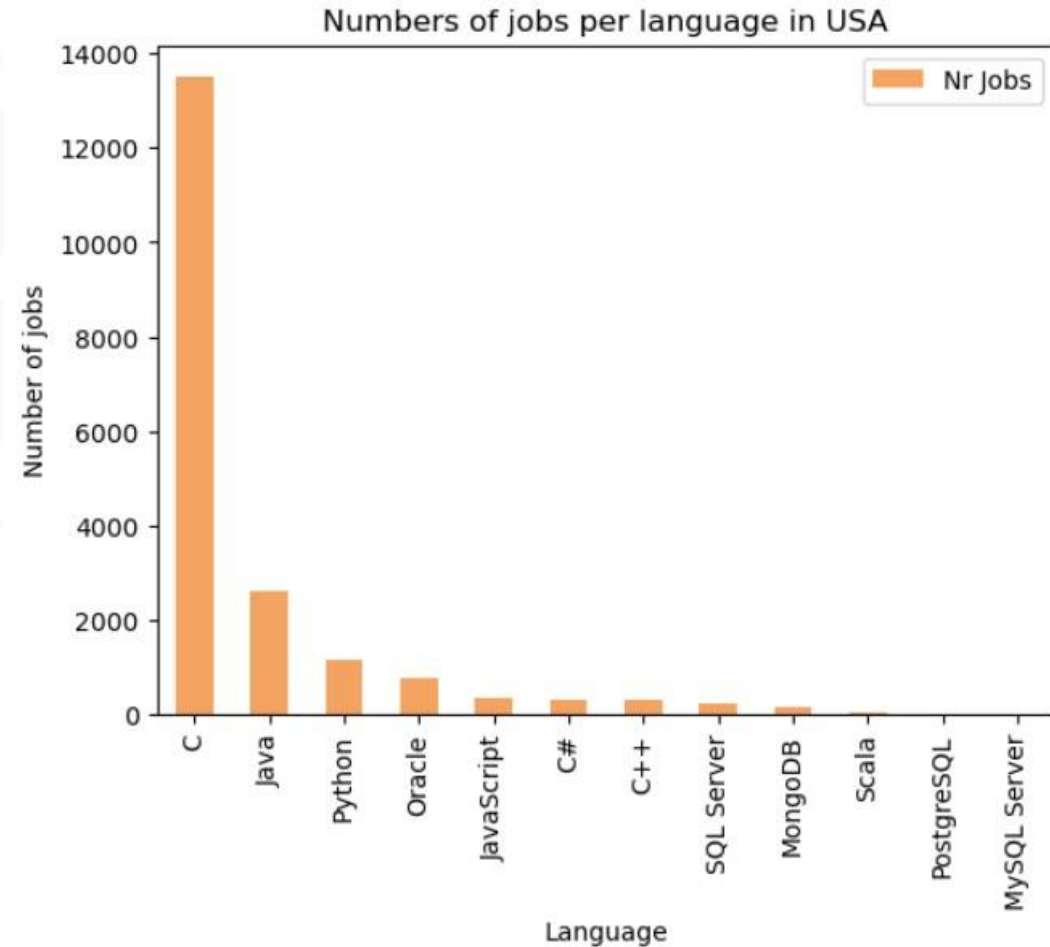
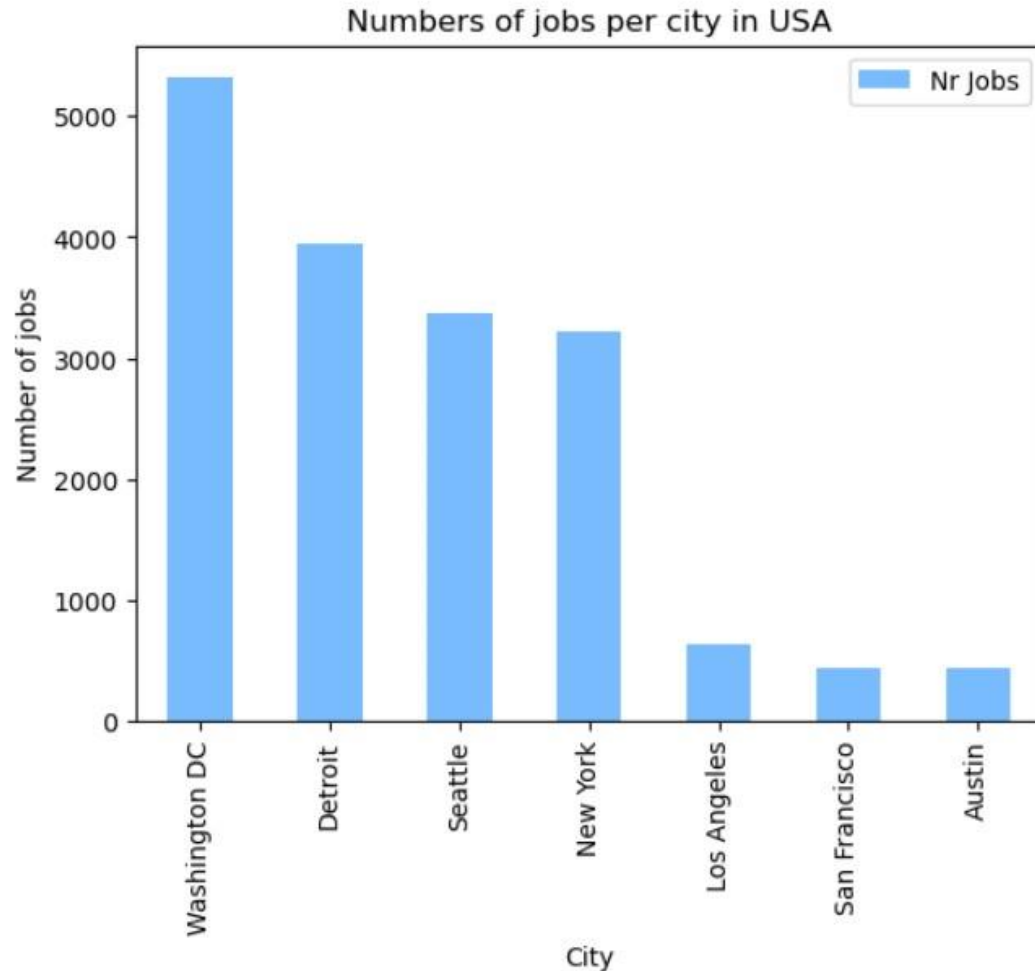
Top 10 most desired Web Frame by developers for next year





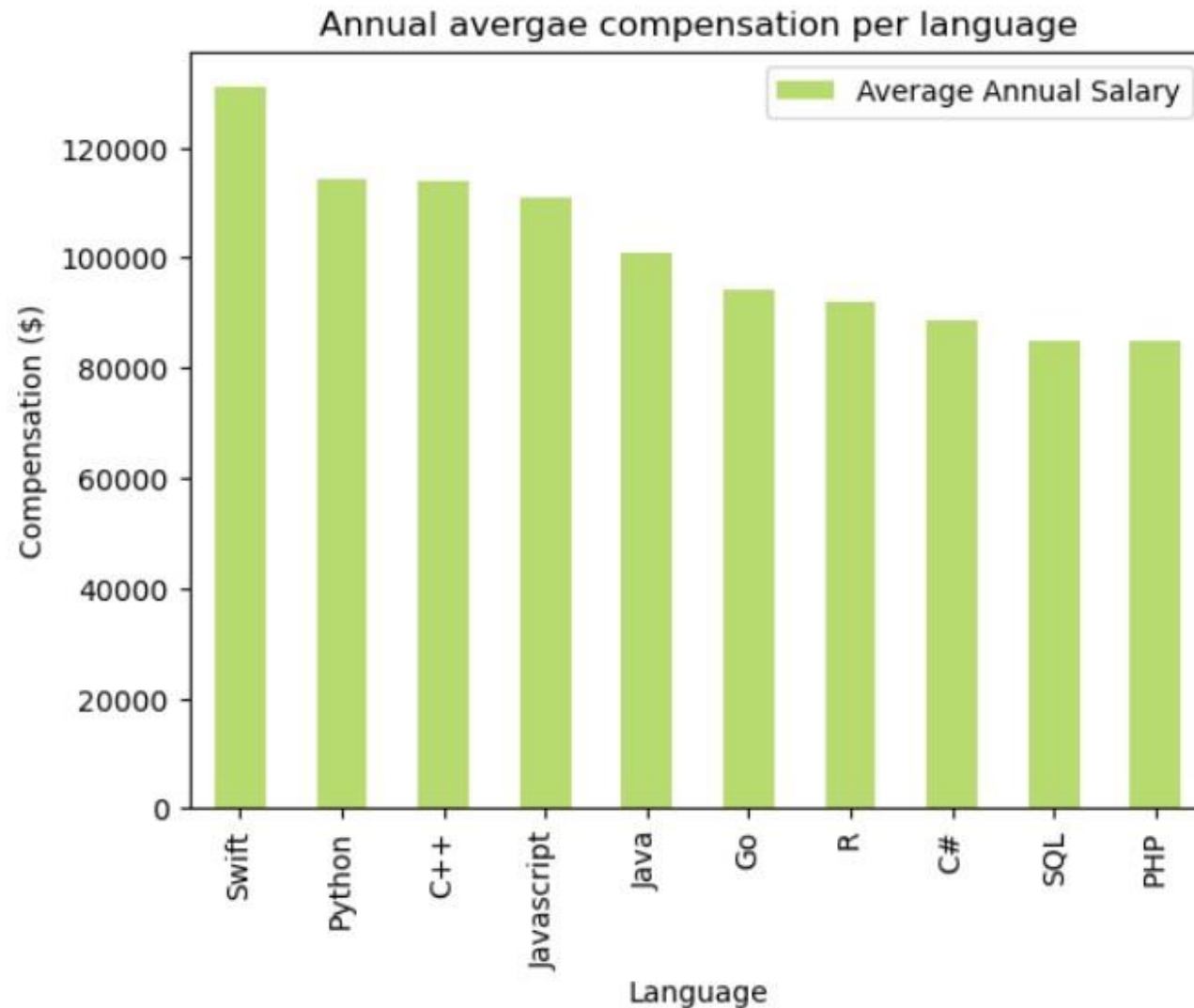
# VISUALIZATION

## Job postings



# VISUALIZATION

## Annual average compensation per language

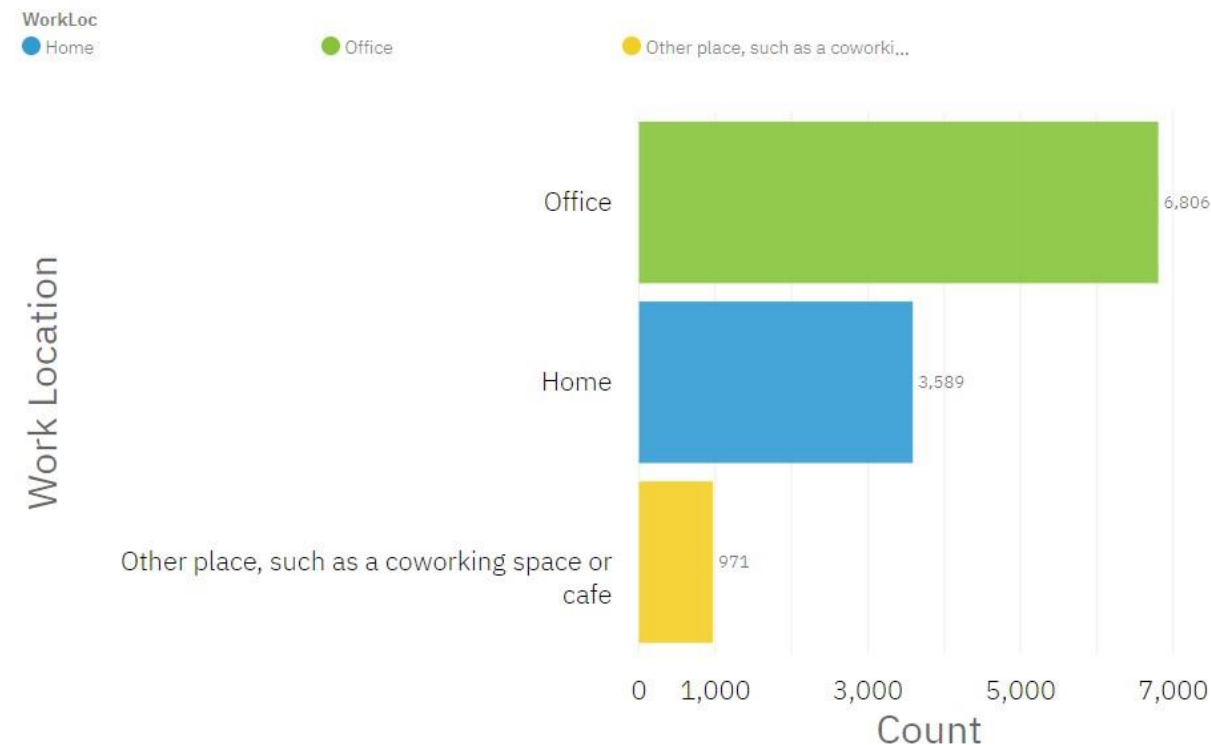


# VISUALIZATION

## Work location

### Work location of developers respondents

Work location of developers respondents in 2019



### Work location and use of developers respondents

Work location and frequency of use of developers respondents in 2019

