Tech Nerds Data Analysis

Stack Overflow 2018 Survey

Team Members

Kuldeep

Madhukara

Shameetha

Core Message

- Each year, Stack overflow conducts a survey within the developer community about everything from their favorite technologies to their job preferences.
- Over 100,000 developers took the 30-minute survey.
- We have analyzed the data from this survey for 2018 to identify few interesting trends.
- This could be used by industry leaders to make decisions on their opportunity costs
- Help students decide on the classes that could be useful when they graduate or to identify technology leaders by country.

Questions

- Since we are developers ourselves our group found it interesting to answer the following questions:

  - What are important trends in the software industry today?
  - What languages and technologies associated with this trend?
  - How developers working in these areas are benefiting?
  - What is a typical successful developer profile look like?

Data exploration and cleanup process

- Data Source: Kaggle
- Data Size:
  - Columns: 129
  - Rows: ~ 100,000
  - File size: ~ 189 MB
  - File format: csv
- Tried to analyze a subset of data relevant to our class project.
  - So cleaned up many extraneous data columns.
    - Columns: 43
    - File size: ~ 88 MB
  - Then in our individual branches, we removed NaN type of data before starting the analysis.

Columns

```
df.columns
```

```
Index(['Respondent', 'Hobby', 'OpenSource', 'Country', 'Student', 'Employment',
       'FormalEducation', 'UndergradMajor', 'CompanySize', 'DevType',
       'YearsCoding', 'YearsCodingProf', 'JobSatisfaction',
       'CareerSatisfaction', 'JobSearchStatus', 'LastNewJob', 'UpdateCV',
       'Currency', 'SalaryType', 'ConvertedSalary', 'LanguageWorkedWith',
       'LanguageDesireNextYear', 'DatabaseWorkedWith',
       'DatabaseDesireNextYear', 'PlatformWorkedWith',
       'PlatformDesireNextYear', 'FrameworkWorkedWith',
       'FrameworkDesireNextYear', 'IDE', 'OperatingSystem', 'Methodology',
       'VersionControl', 'CheckInCode', 'AIDangerous', 'AIInteresting',
       'AIResponsible', 'AIFuture', 'Exercise', 'Gender', 'SexualOrientation',
       'EducationParents', 'RaceEthnicity', 'Age'],
      dtype='object')
```

Sample data

	Respondent	Hobby	OpenSource	Country	Student	Employment	FormalEducation	UndergradMajor	CompanySize	DevType
0	1	Yes	No	Kenya	No	Employed part-time	Bachelor's degree (BA, BS, B.Eng., etc.)	Mathematics or statistics	20 to 99 employees	Full-stack developer
1	3	Yes	Yes	United Kingdom	No	Employed full-time	Bachelor's degree (BA, BS, B.Eng., etc.)	A natural science (ex. biology, chemistry, phy...	10,000 or more employees	Database administrator;DevOps specialist;Full-...
2	4	Yes	Yes	United States	No	Employed full-time	Associate degree	Computer science, computer engineering, or sof...	20 to 99 employees	Engineering manager;Full-stack developer
3	5	No	No	United States	No	Employed full-time	Bachelor's degree (BA, BS, B.Eng., etc.)	Computer science, computer engineering, or sof...	100 to 499 employees	Full-stack developer
4	7	Yes	No	South Africa	Yes, part-time	Employed full-time	Some college/university study without earning ...	Computer science, computer engineering, or sof...	10,000 or more employees	Data or business analyst;Desktop or enterprise...

Process of data analysis and synthesizing

Developer Profile

- Demographics
- Developer Roles
- Education

Technology

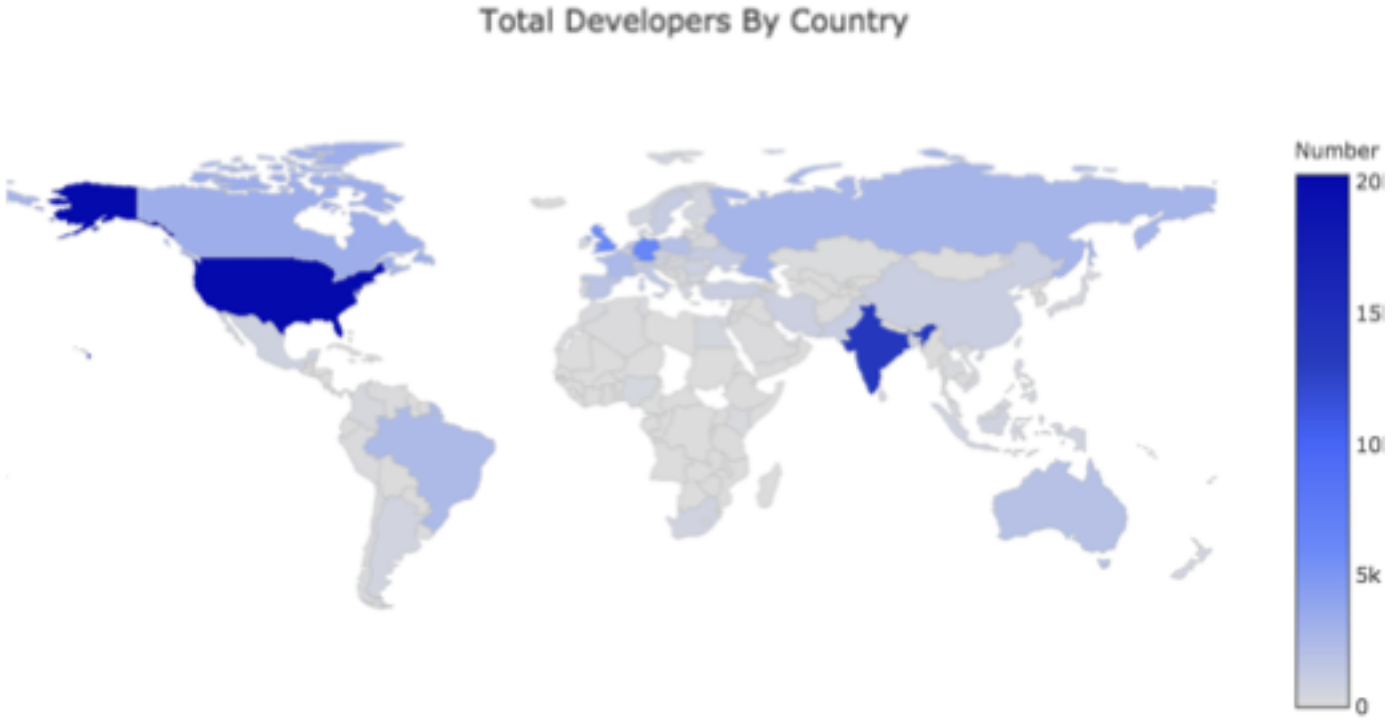
- Popular languages
- Worldwide technology development

Work

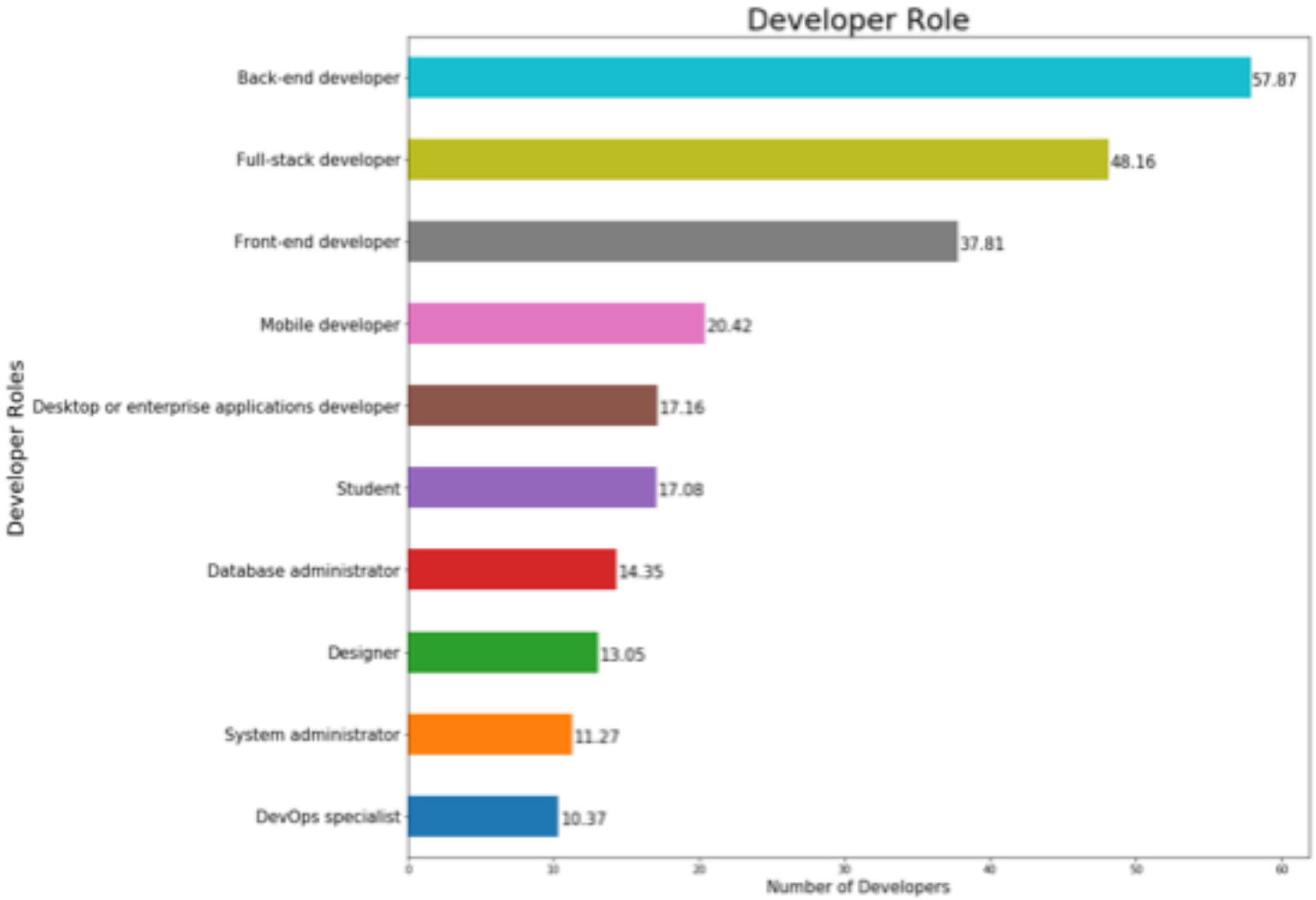
- Company size
- Career satisfaction
- Remuneration

Developer Profile

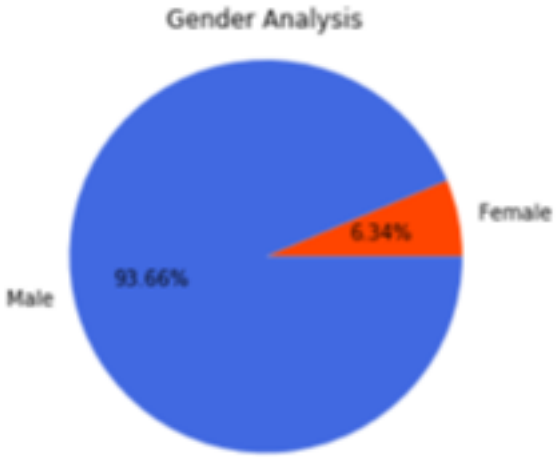
Geography

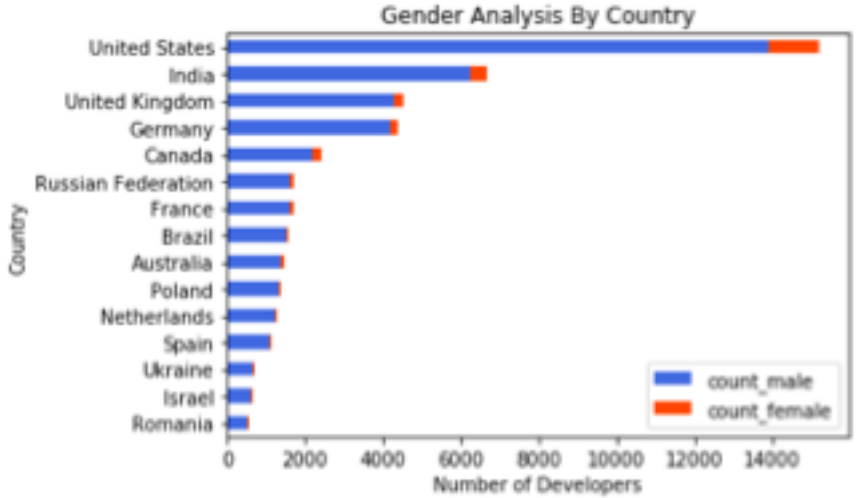


Developer Roles

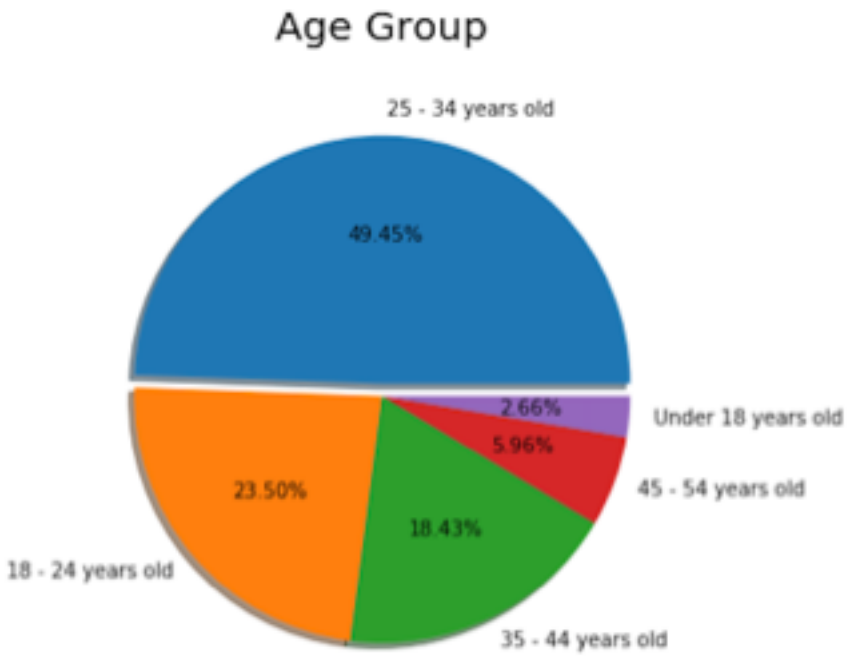


Demographics: Gender

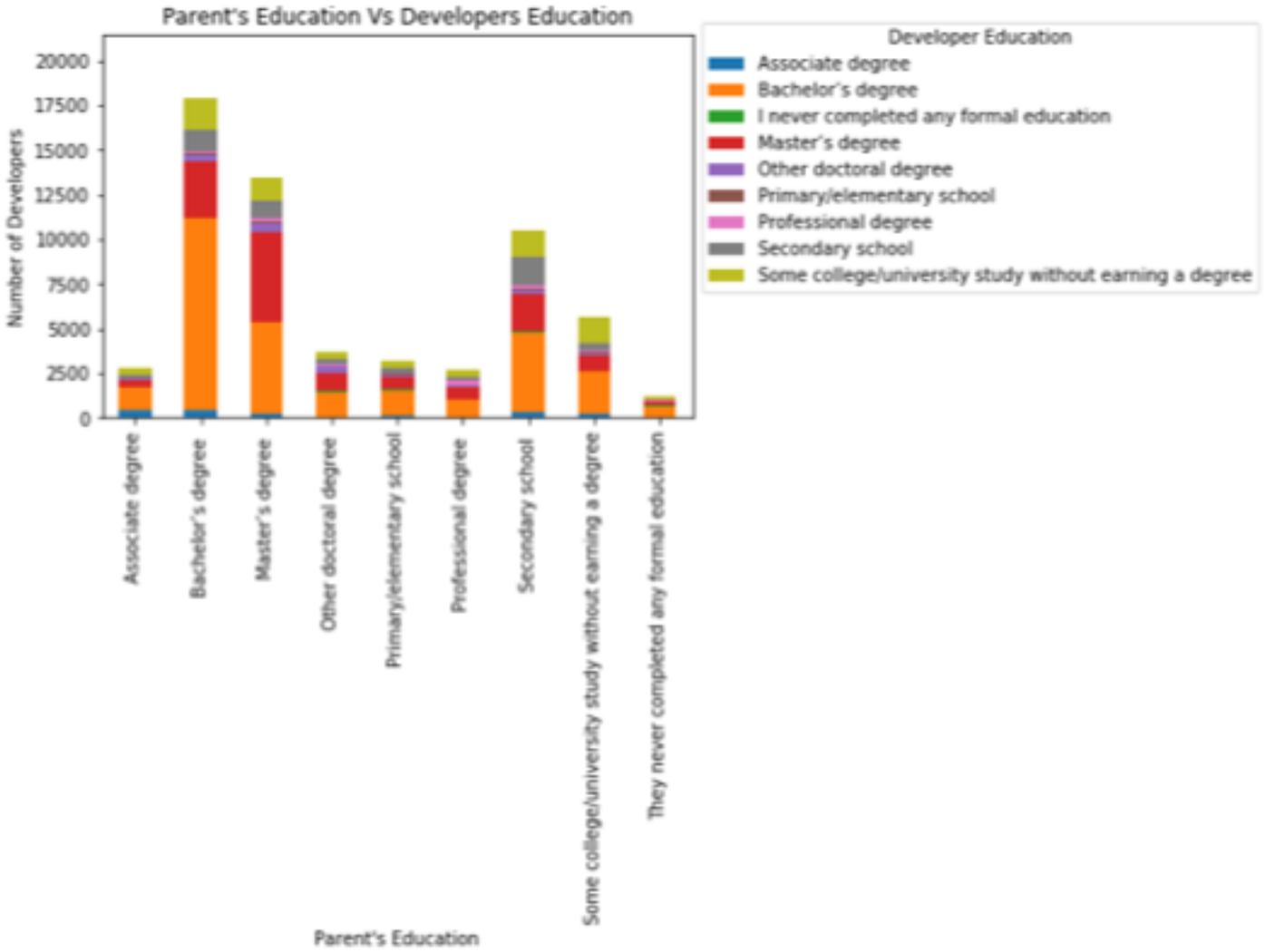




Demographics: Age

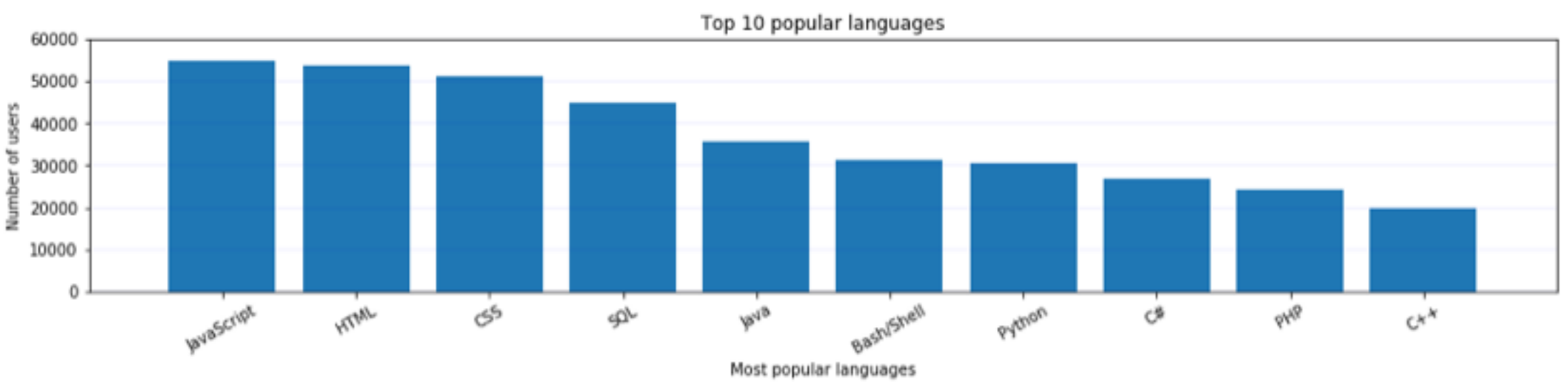


Demographics: Education

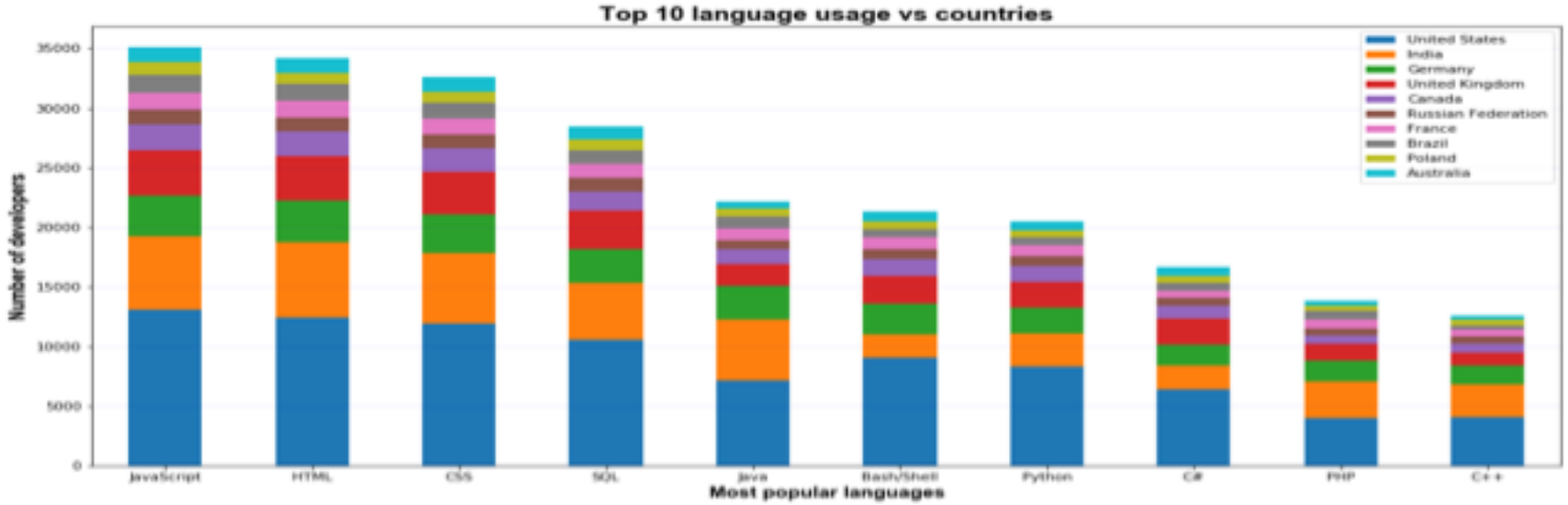


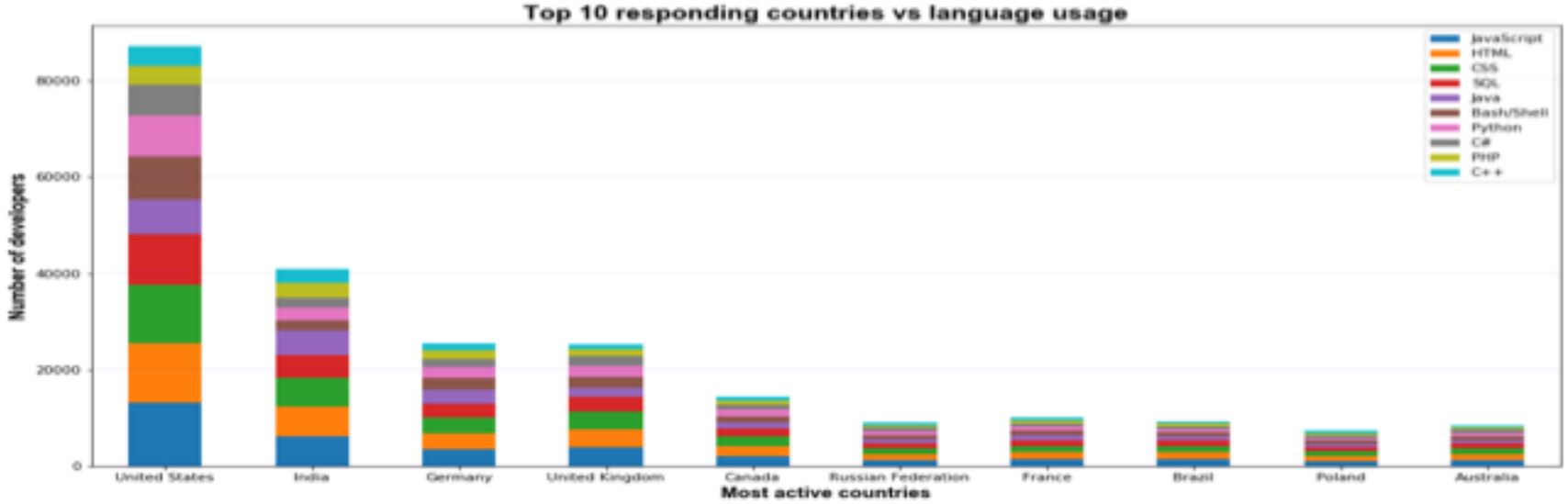
Technology

Popular programming language

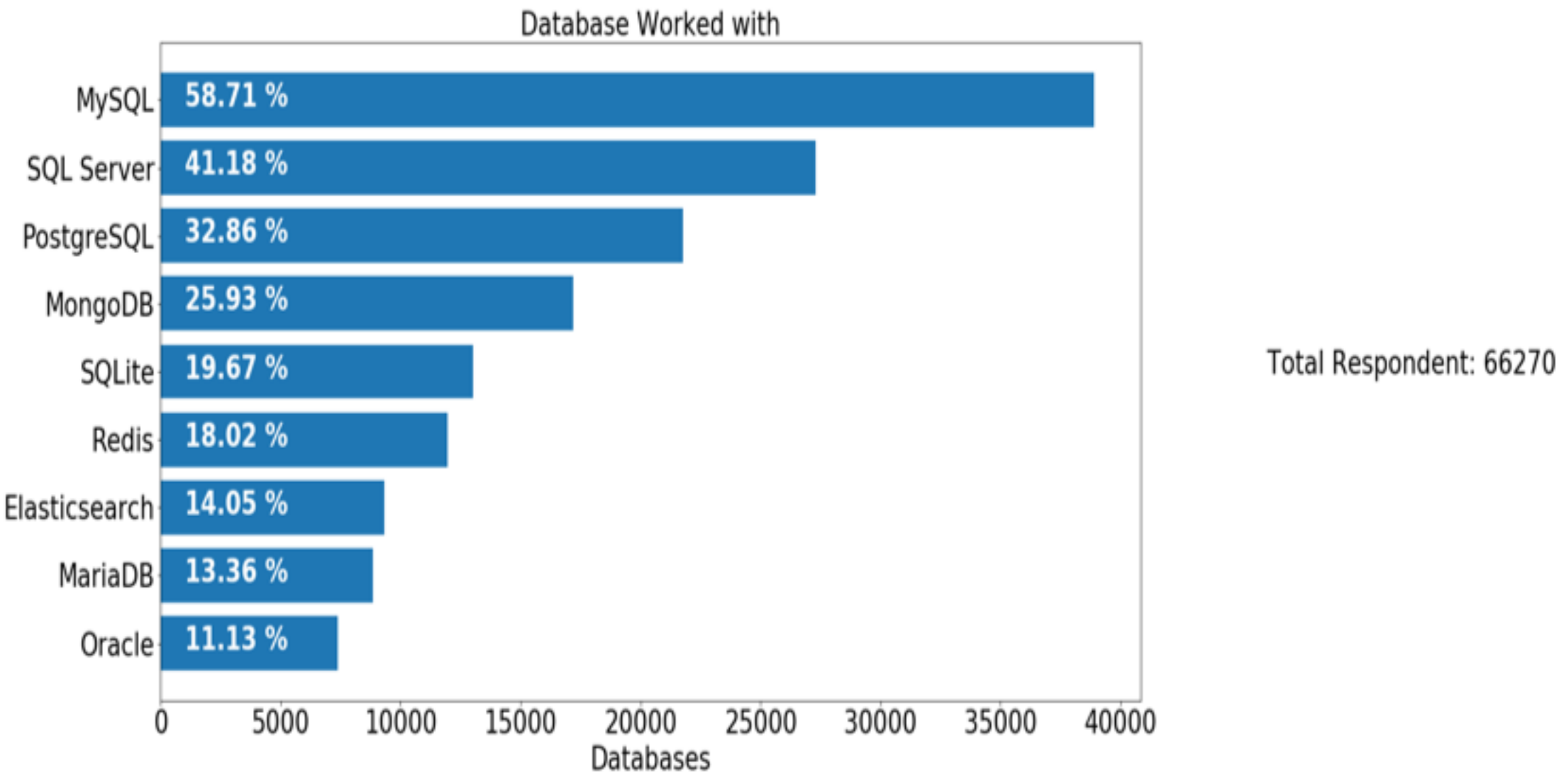


Language vs Country

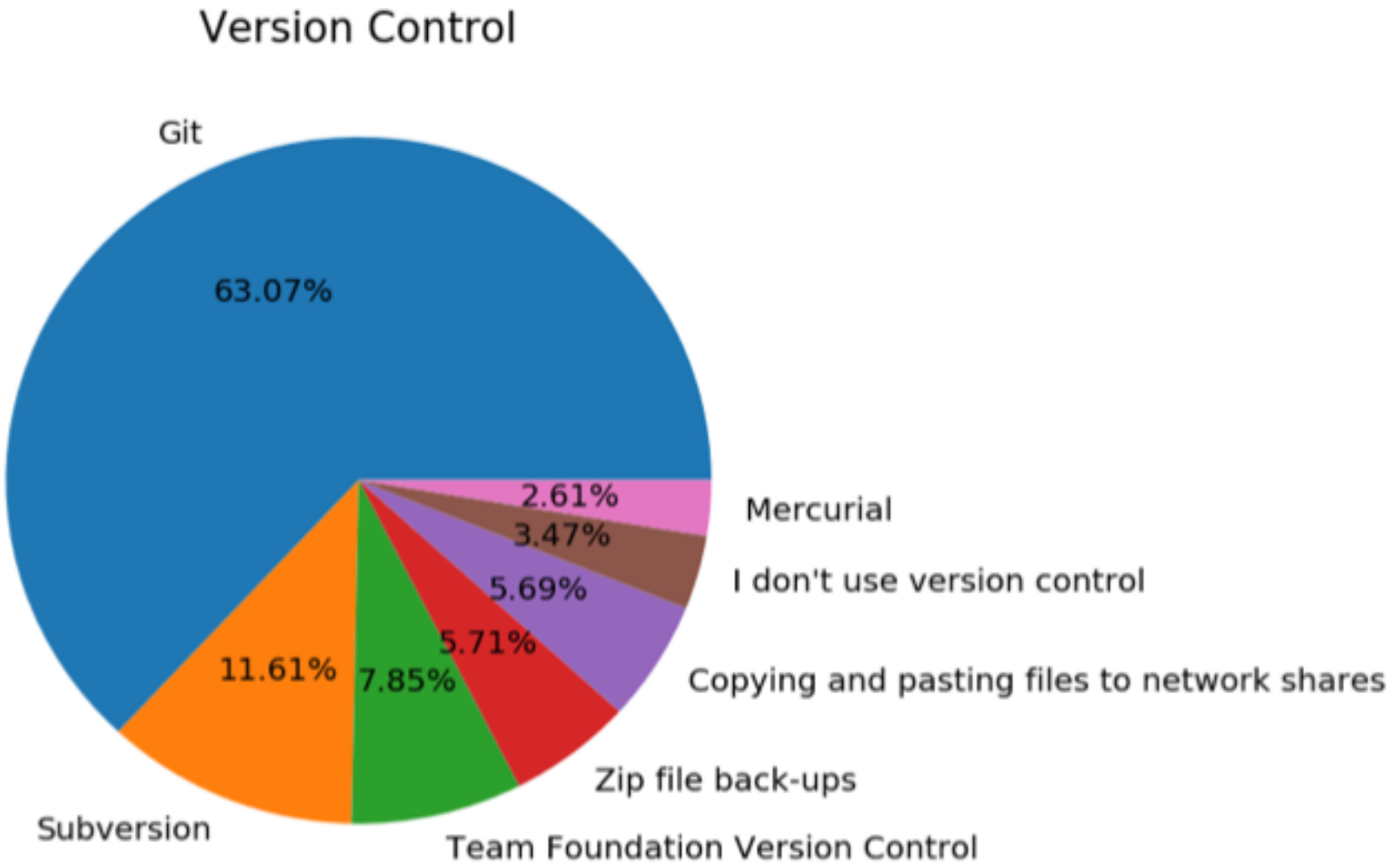




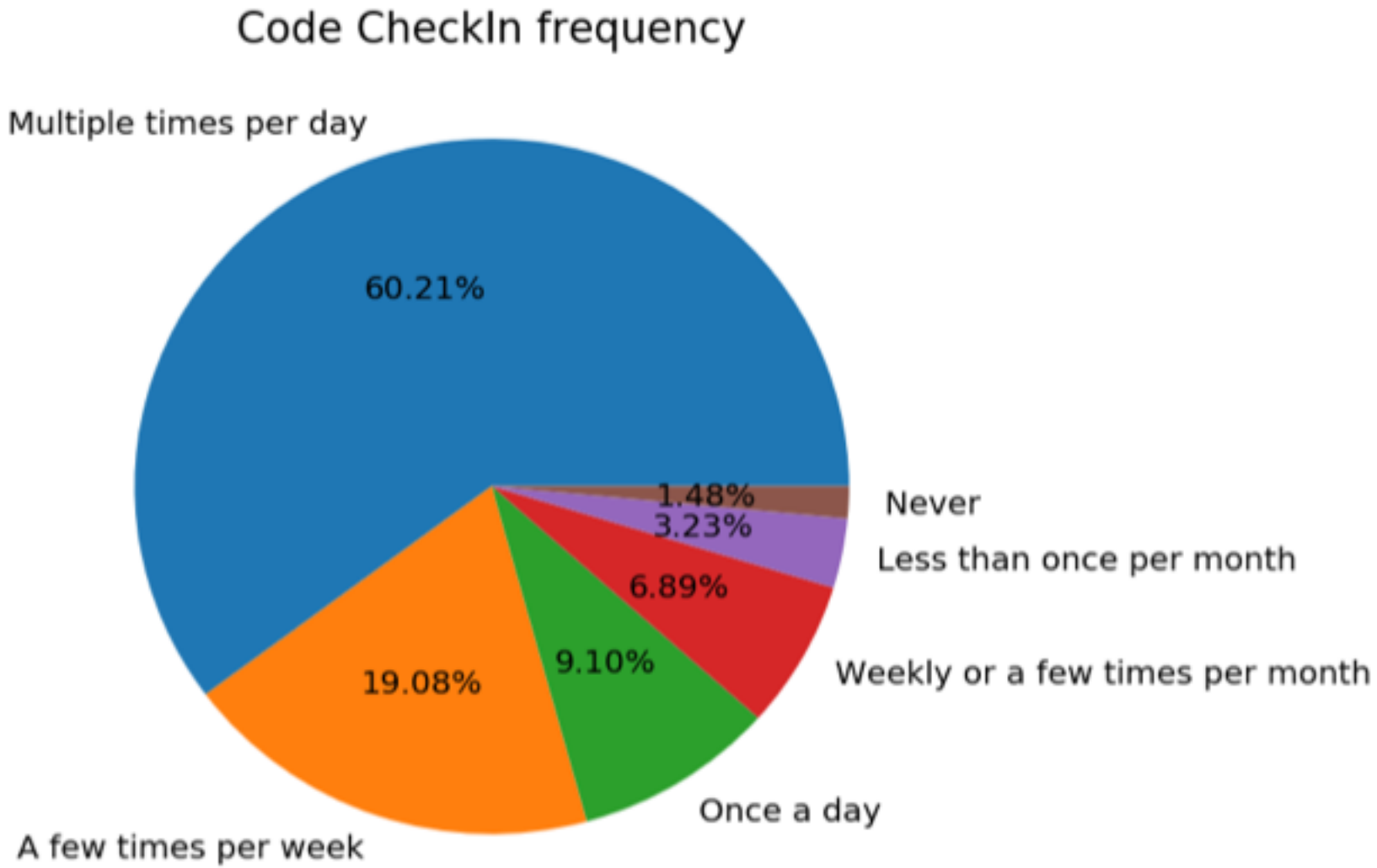
Top 10 Database



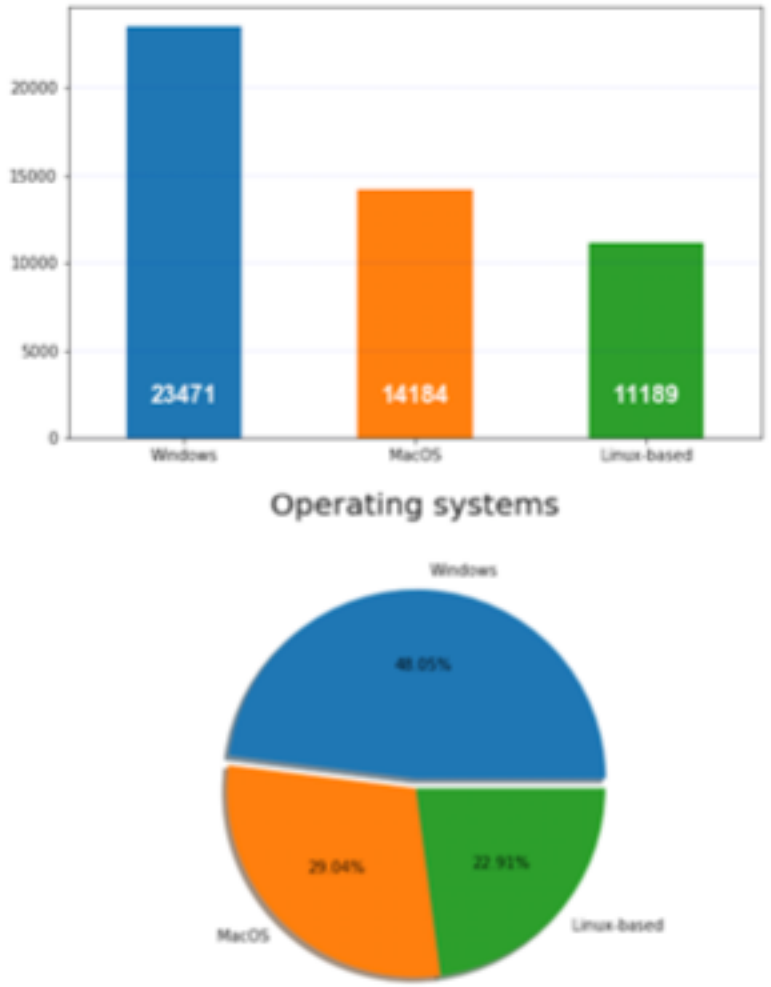
Version Control & Commit



Commit Frequency

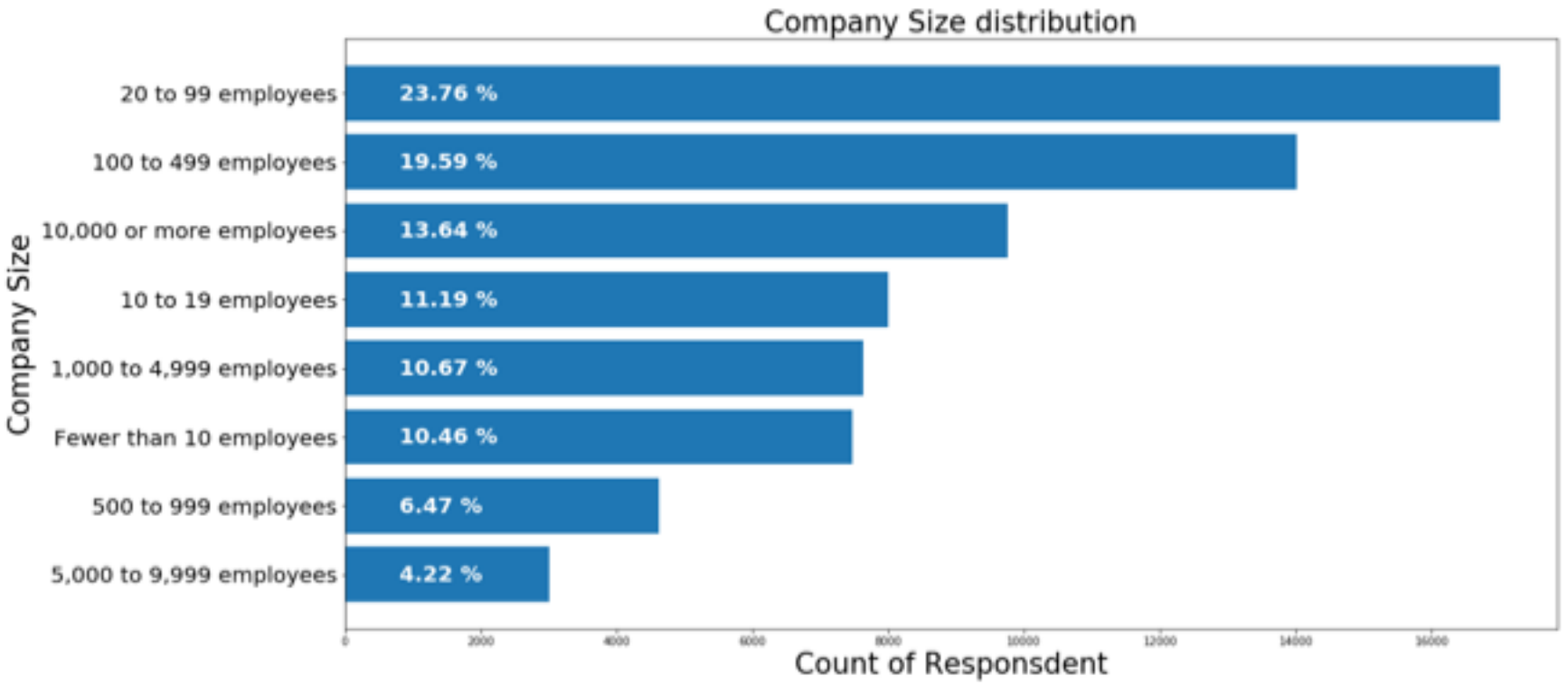


Operating System



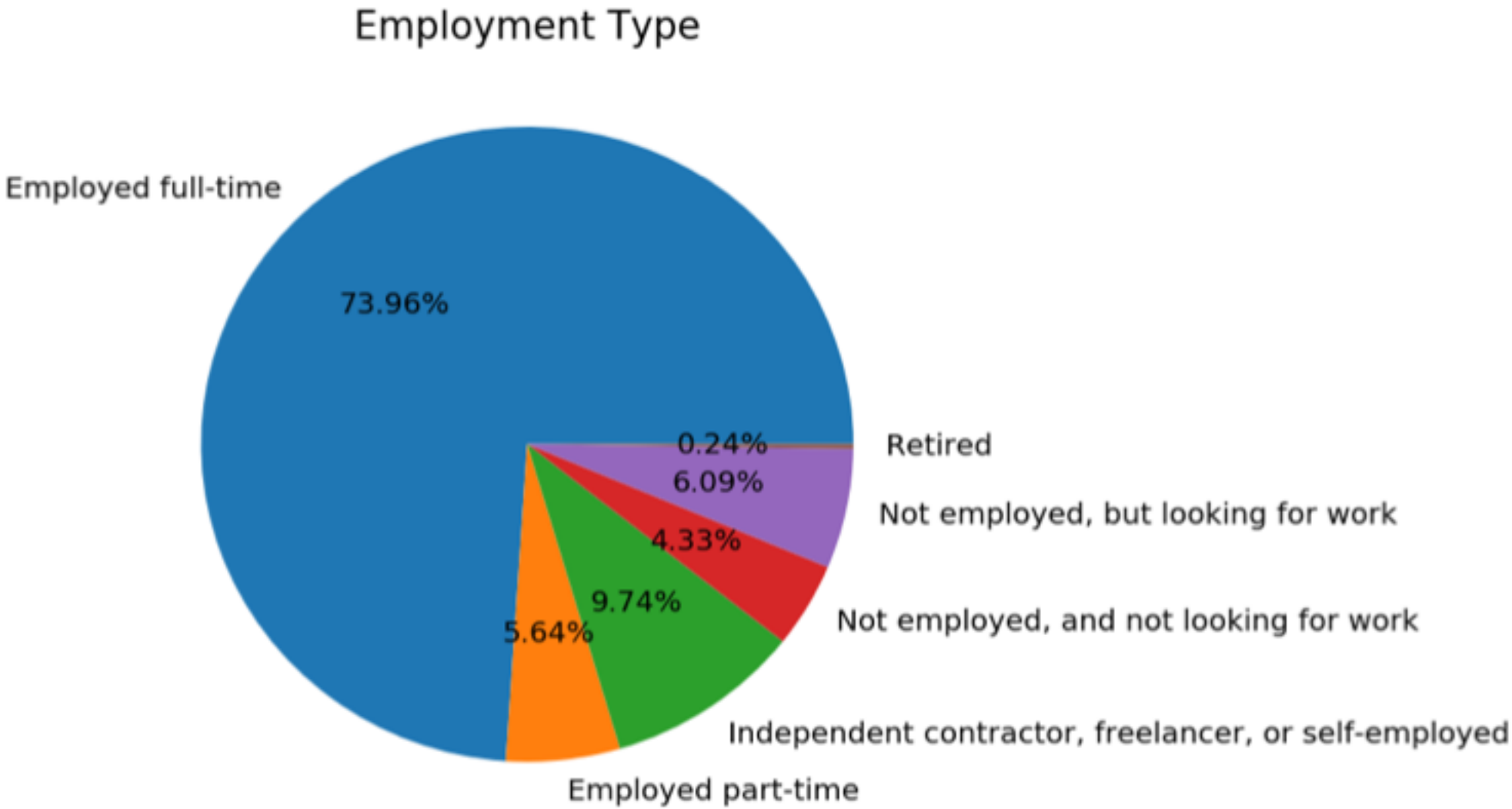
Work

Company Size

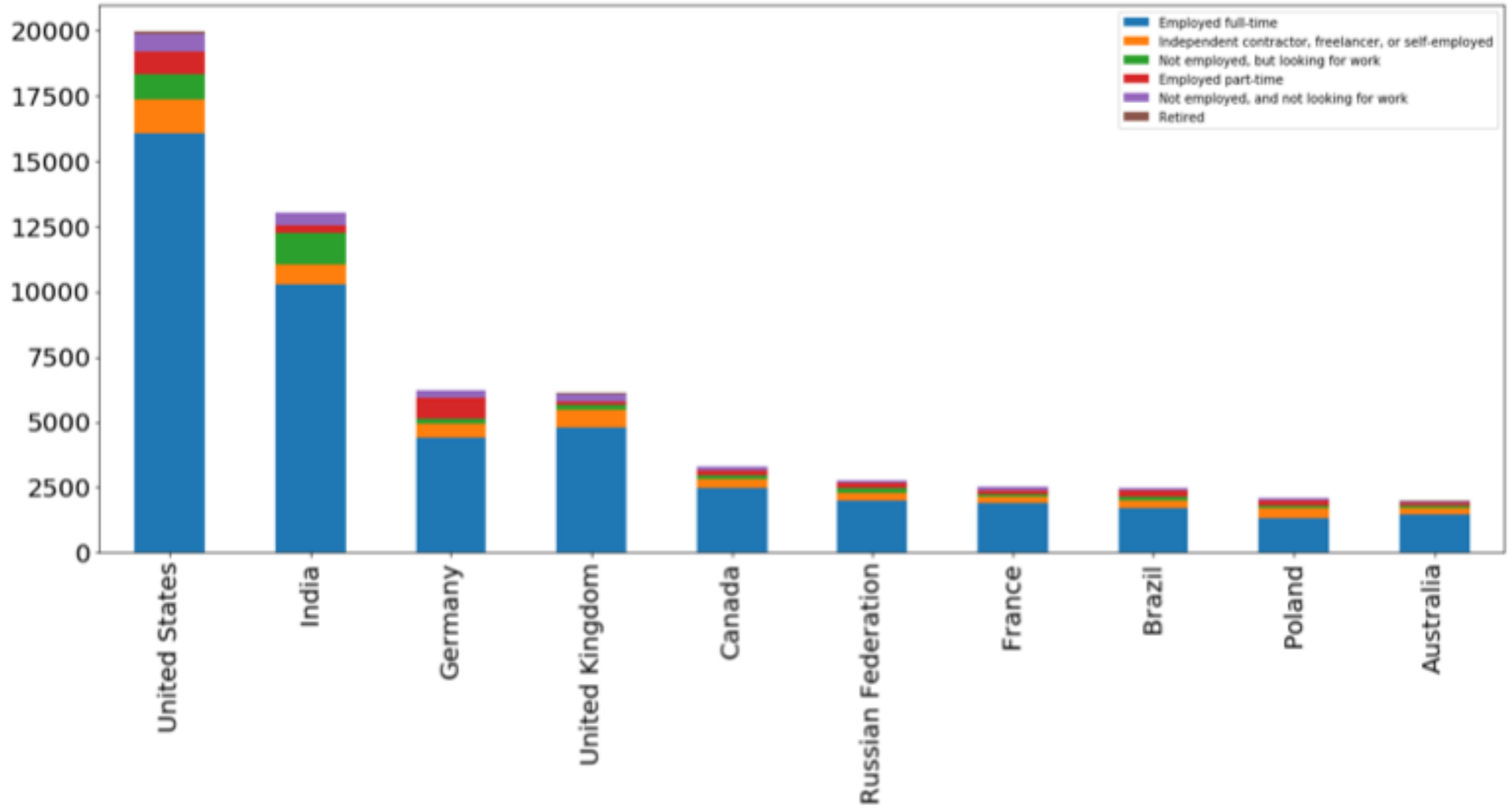


Total Respondent: 71531

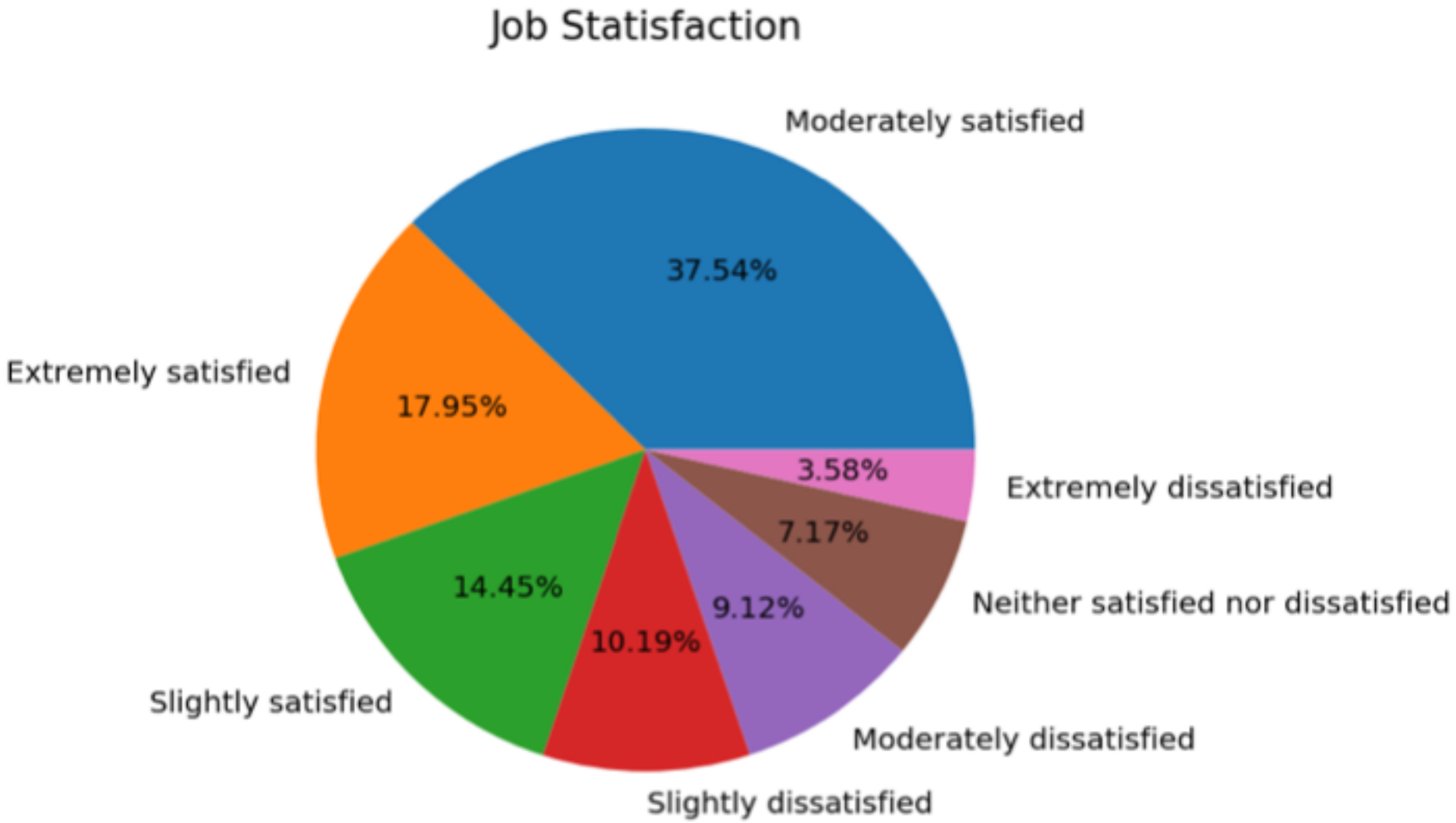
Employment Status



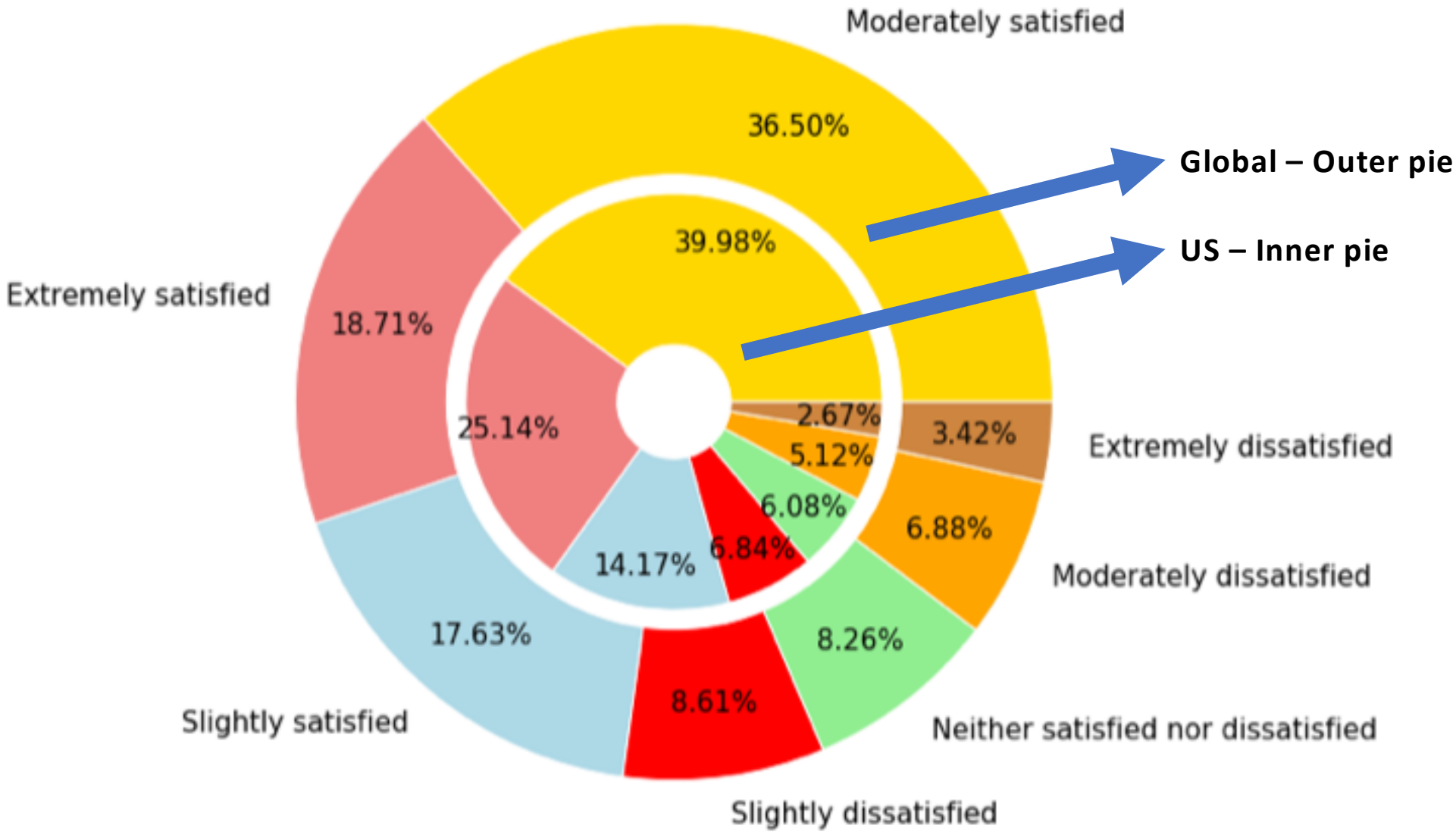
Employment Status by Country



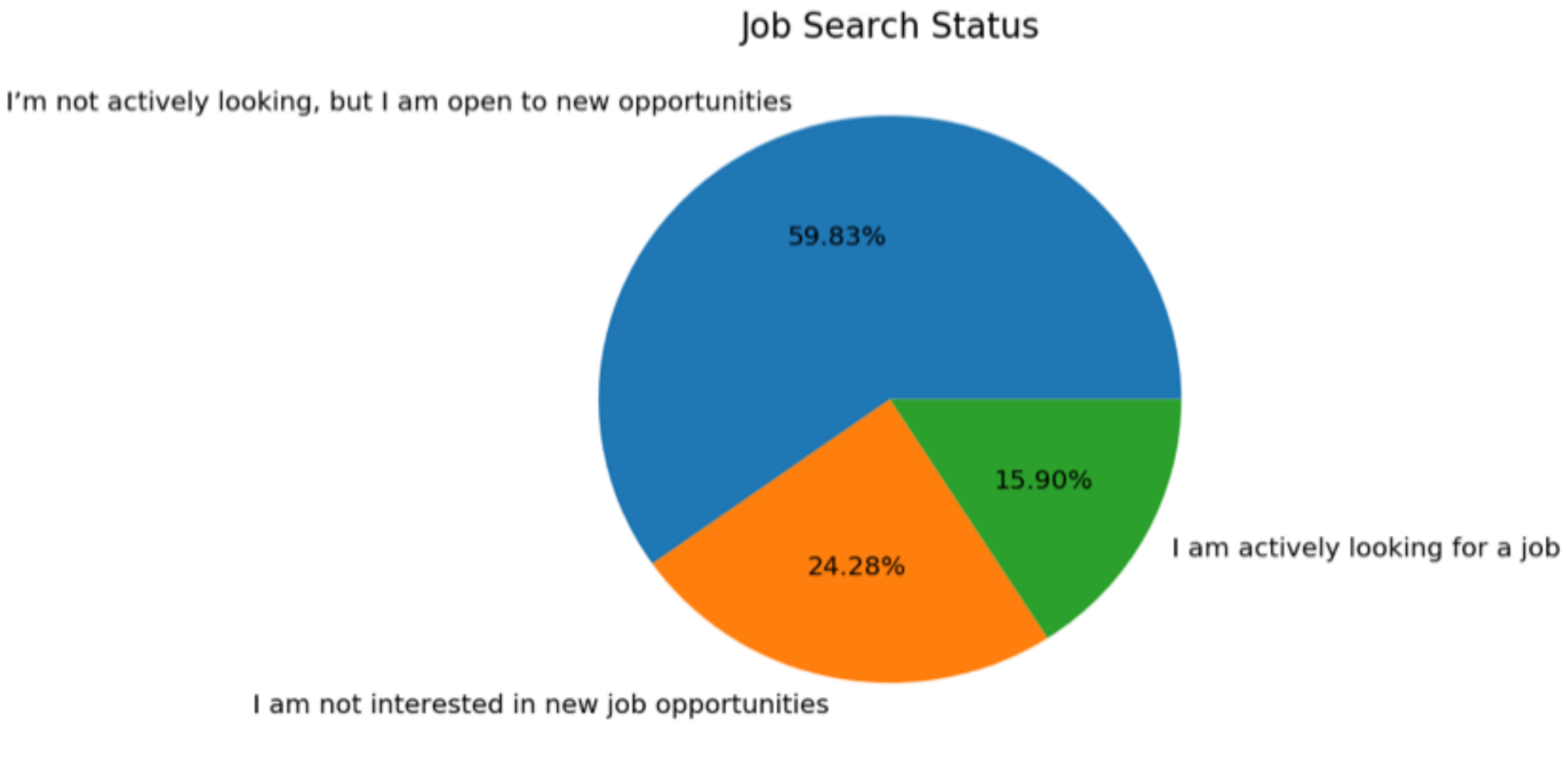
Job Satisfaction



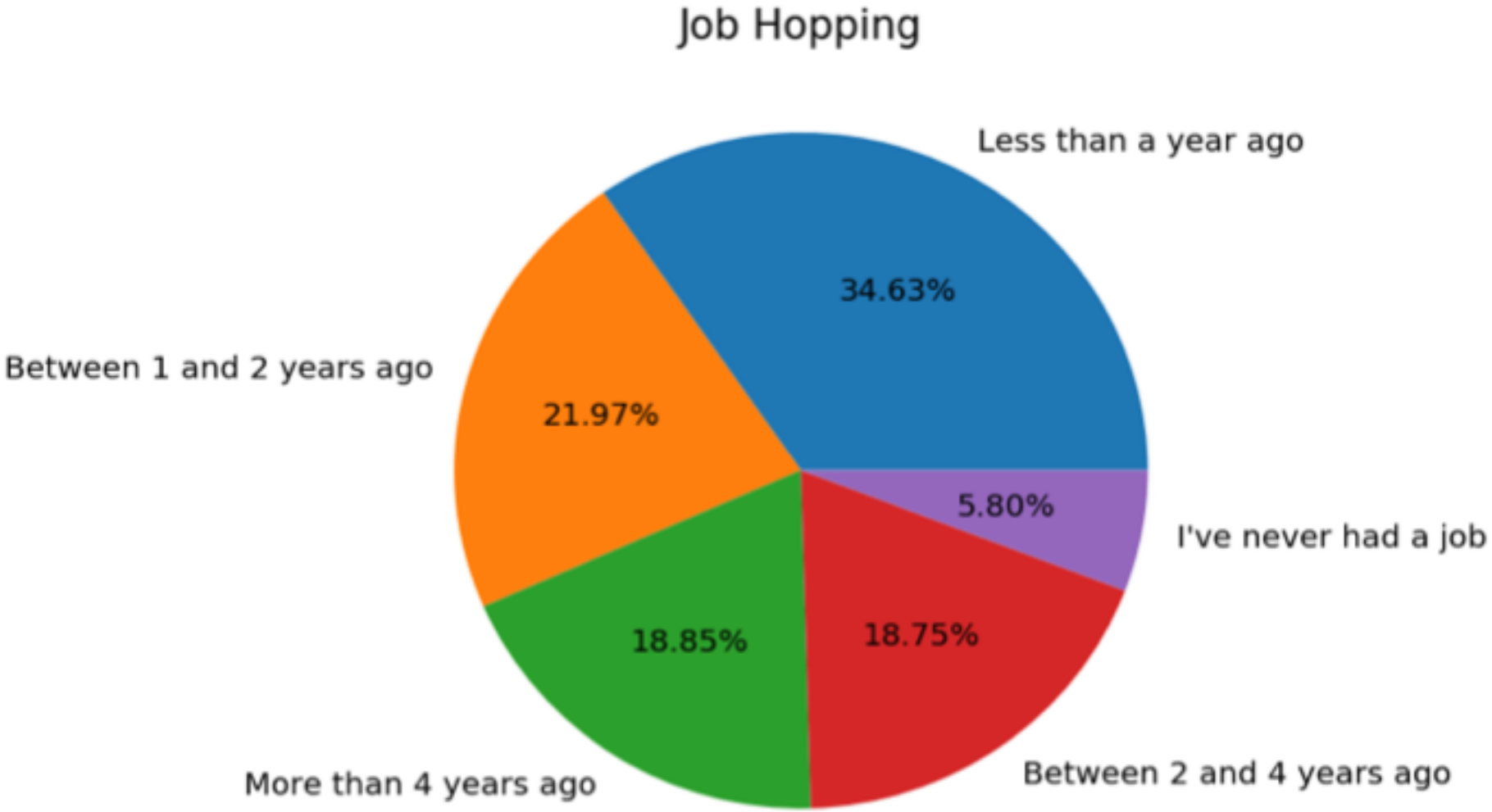
Career Satisfaction – Global vs US



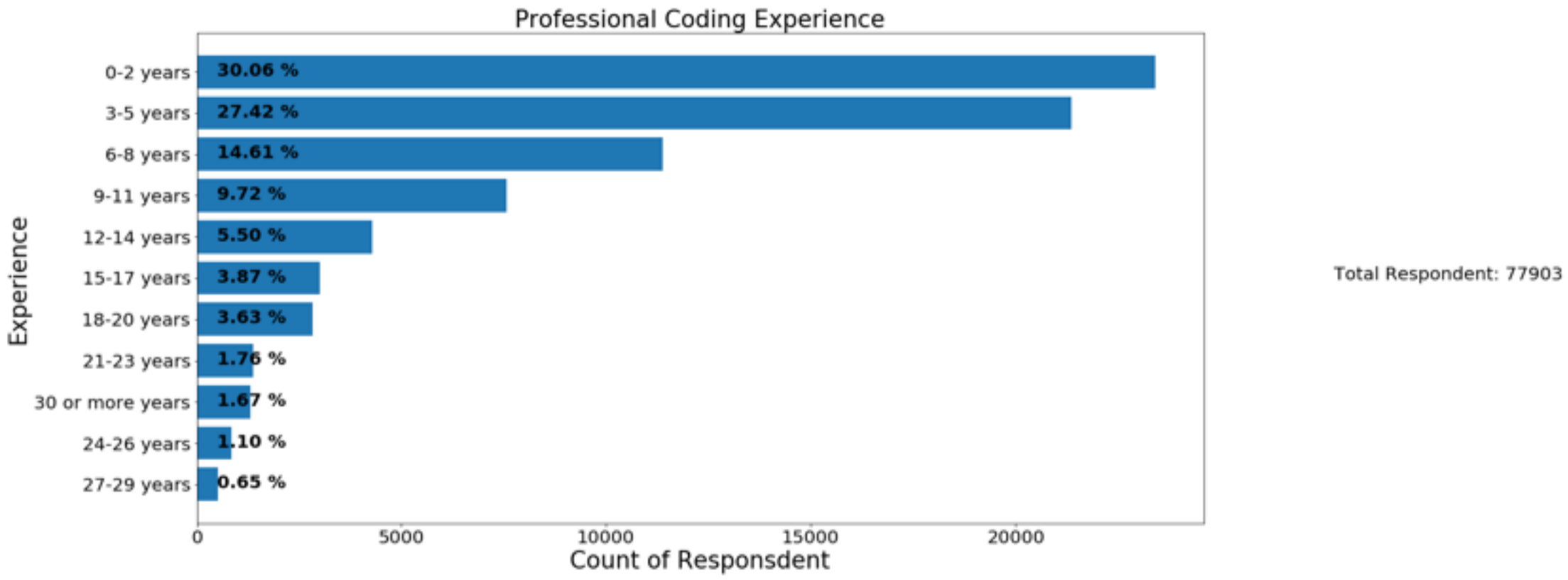
Job Search Status



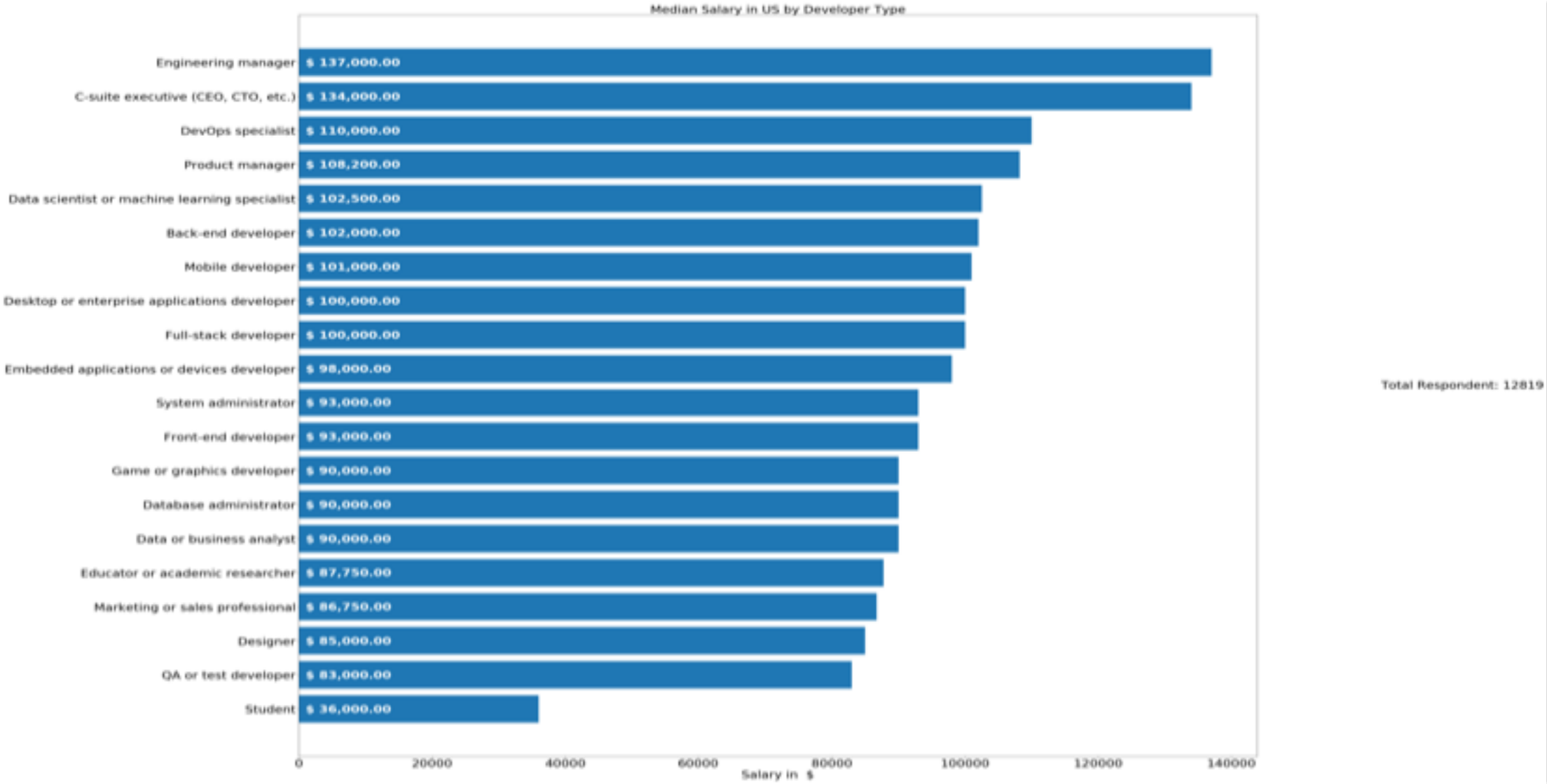
Job Hopping


Job Hopping
Less than a year ago
34.63%
Between 1 and 2 years ago
21.97%
5.80%
I've never had a job
18.85%
18.75%
More than 4 years ago
Between 2 and 4 years ago

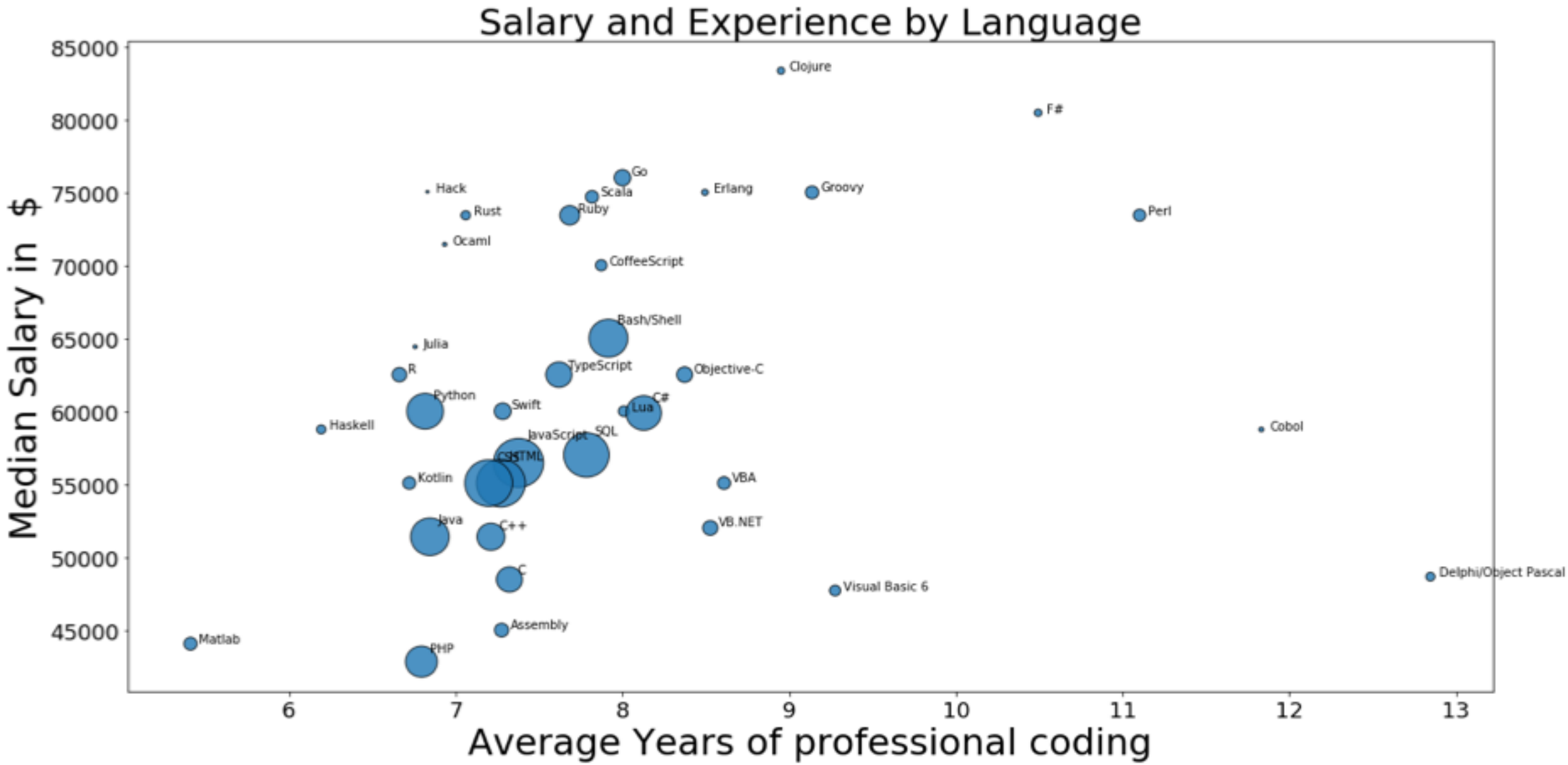
Professional Coding Experience



Salary by Job Title



Salary and Experience by Language: Global



Conclusion

- Emerging trends in technology:
  - Popular age group: 24- 35 years
  - Web-Technologies: Javascript, HTML and CSS are most popular languages
  - MySQL is the most popular database engine
  - Git is an extremely popular version control system worldwide
  - United States and India emerged as two top most technology countries, while UK and Germany are following.
  - Popular platforms – Windows and Mac while Linux is not way behind
  - Career satisfaction – Moderately satisfied both globally and in US.
  - Most developers preferred to switched job within 1 year.

Thank you!

Q & A?