2021 Kaggle Machine Learning & Data Science Survey

BY : Nouf Almalki

kaggle

About Data

This year, Kaggle set out to conduct an industry-wide survey that presents a truly comprehensive view of the state of data science and machine learning and the best ways for new data scientists to break into the field. Survey data provides an overview of the sector on an aggregate scale

About Data

38 Questions

25488 Response

Data cleaning

01

Checking the nulls

02

determining the questions answered by 50% of the Kaggle community

04

Drop heterosexuals from the data

05

Abbreviation of long names

06

Adding Year column

Tools

Python and Jupyter Notebook

Numpy and Pandas

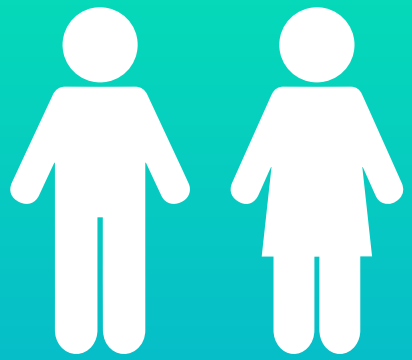
for data manipulation

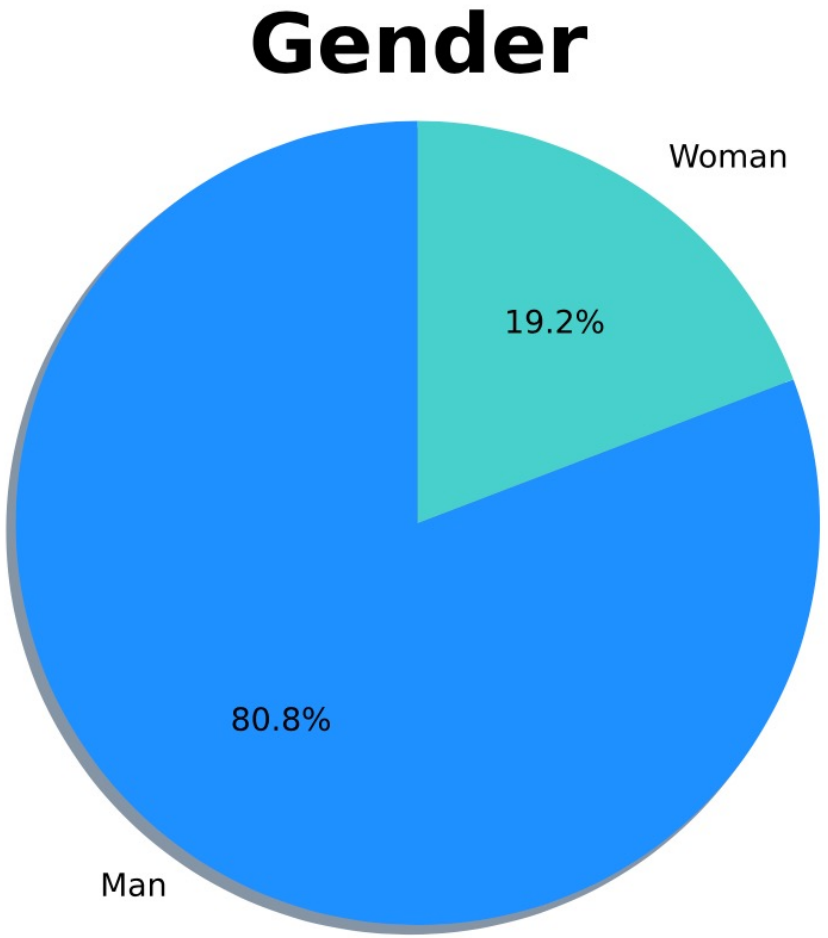
Matplotlib and Seaborn

for plotting visualization

Findings

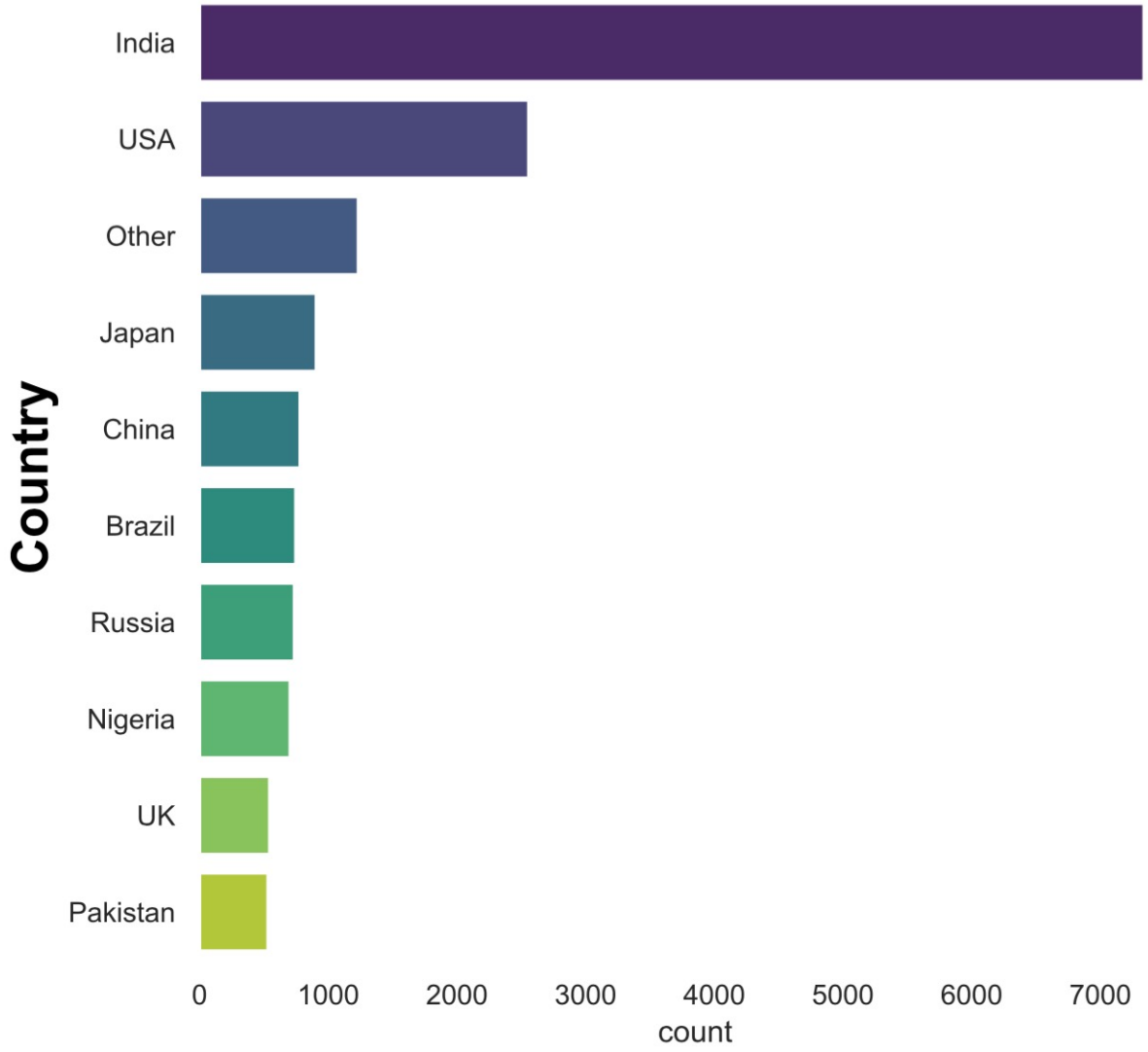
Gender





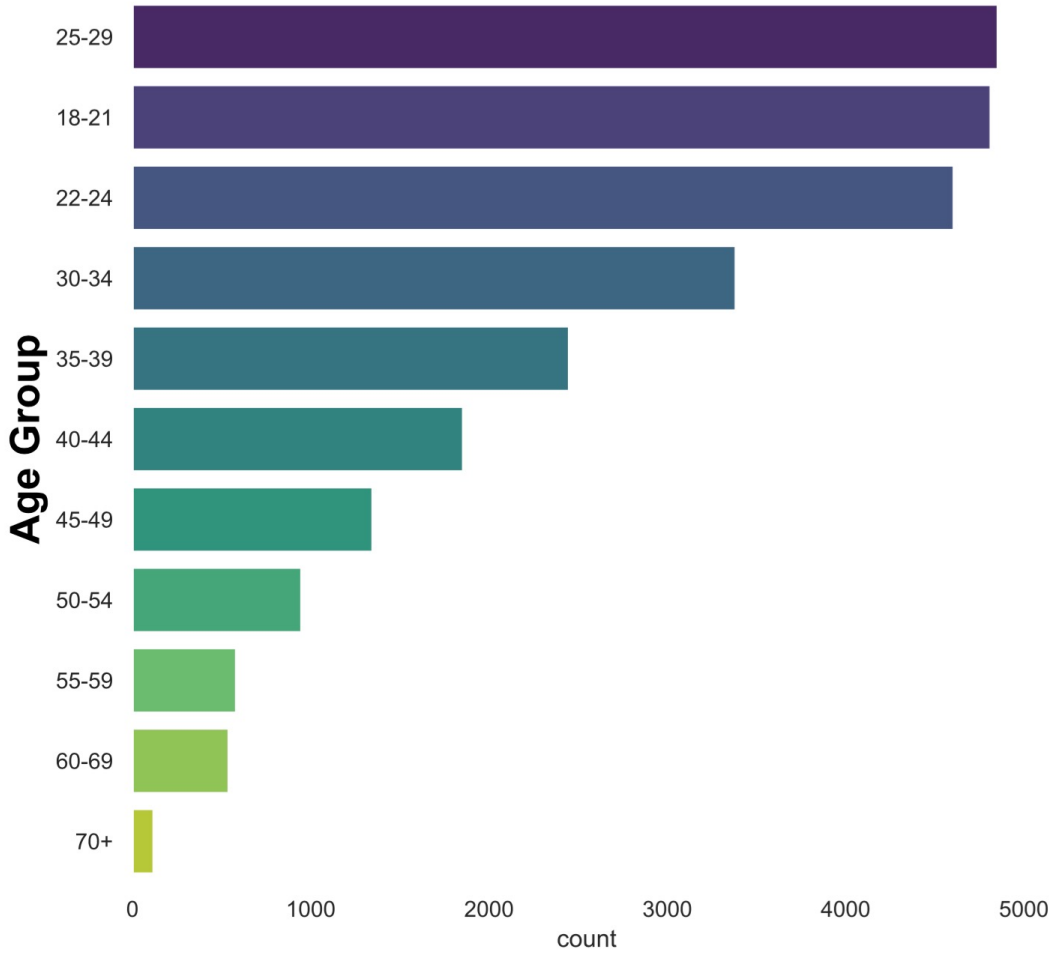
kaggle

Country



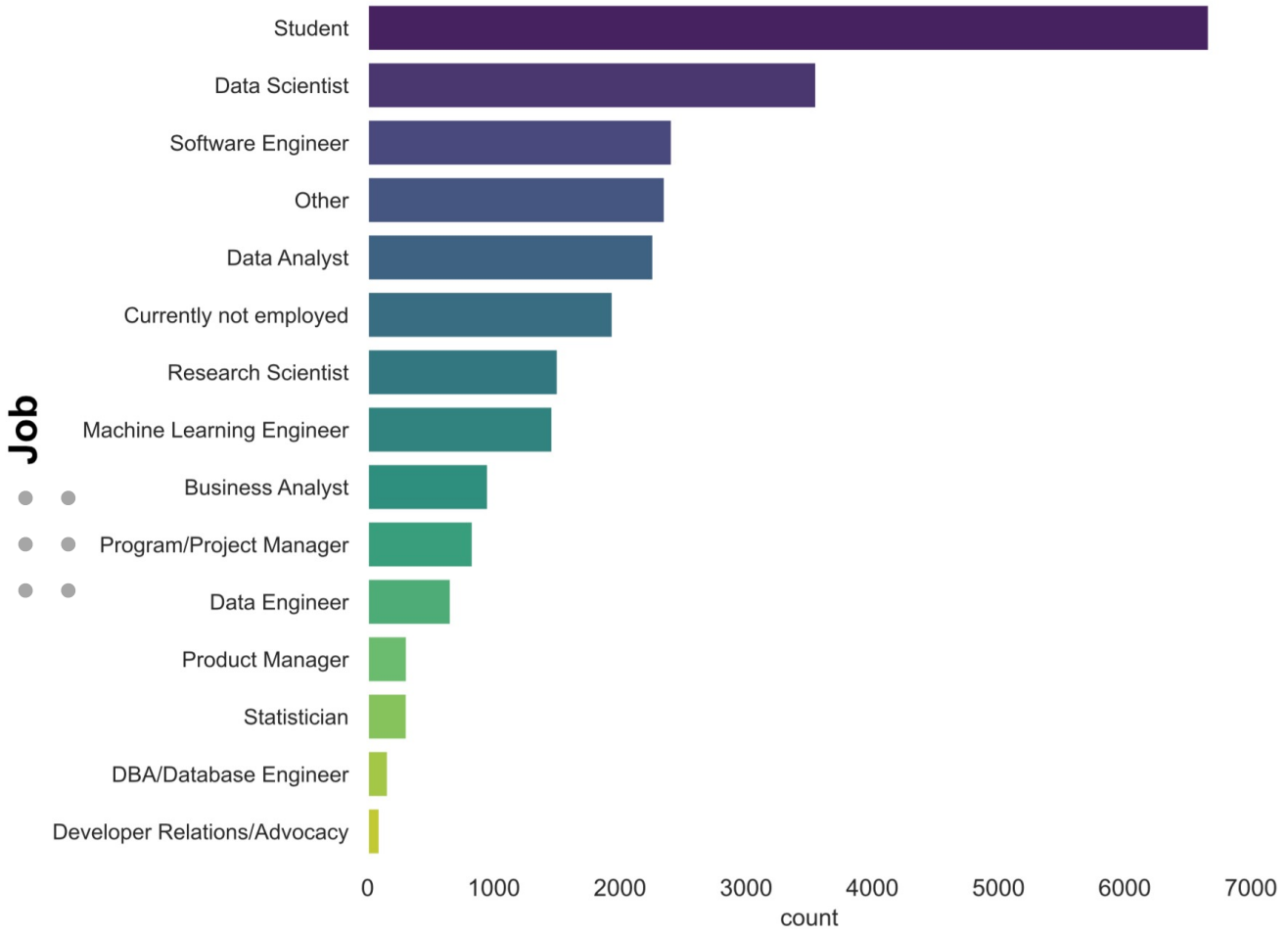
kaggle

Age Group



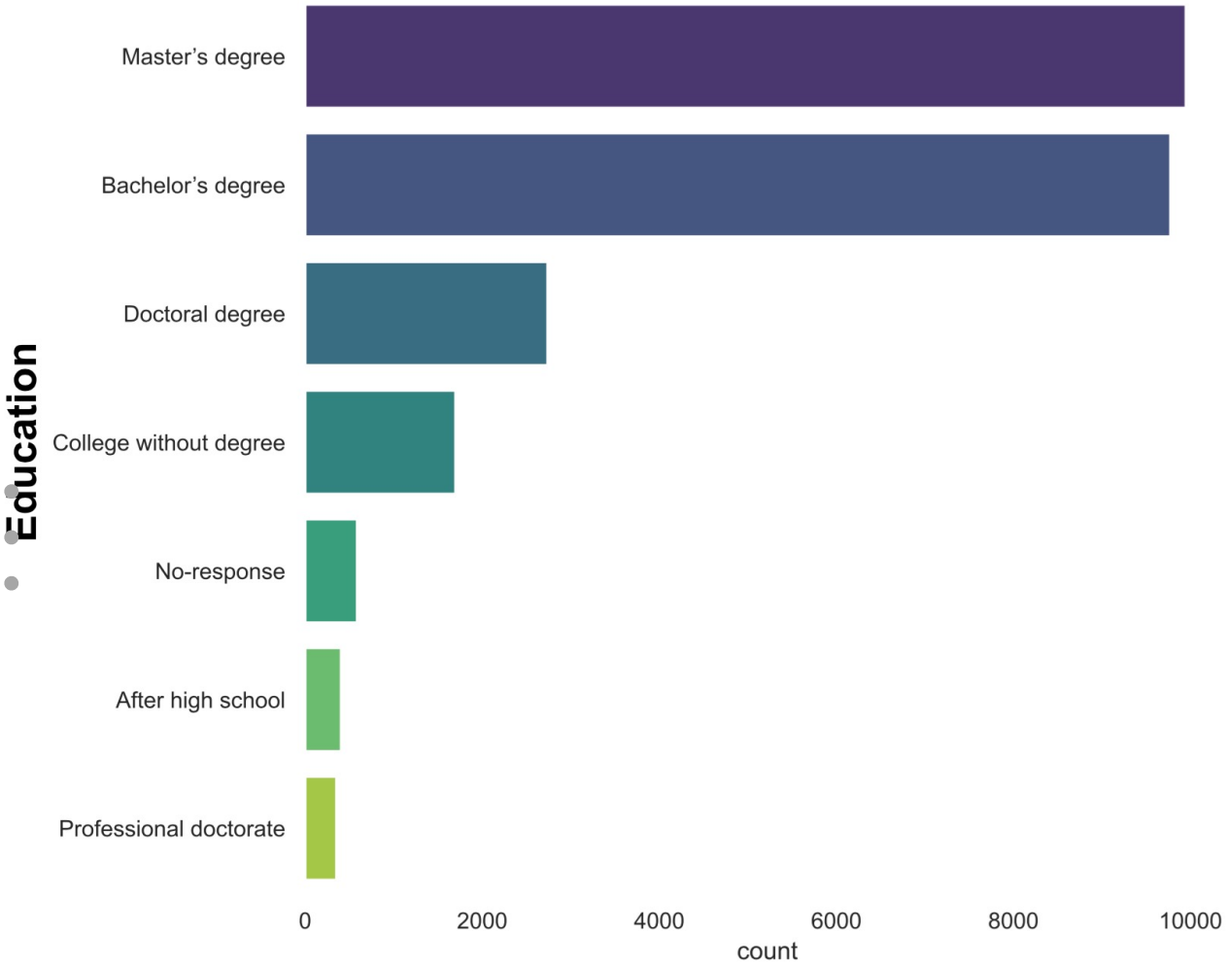
kaggle

Job


Job
Student
Data Scientist
Software Engineer
Other
Data Analyst
Currently not employed
Research Scientist
Machine Learning Engineer
Business Analyst
Program/Project Manager
Data Engineer
Product Manager
Statistician
DBA/Database Engineer
Developer Relations/Advocacy
0
1000
2000
3000
4000
5000
6000
7000
count

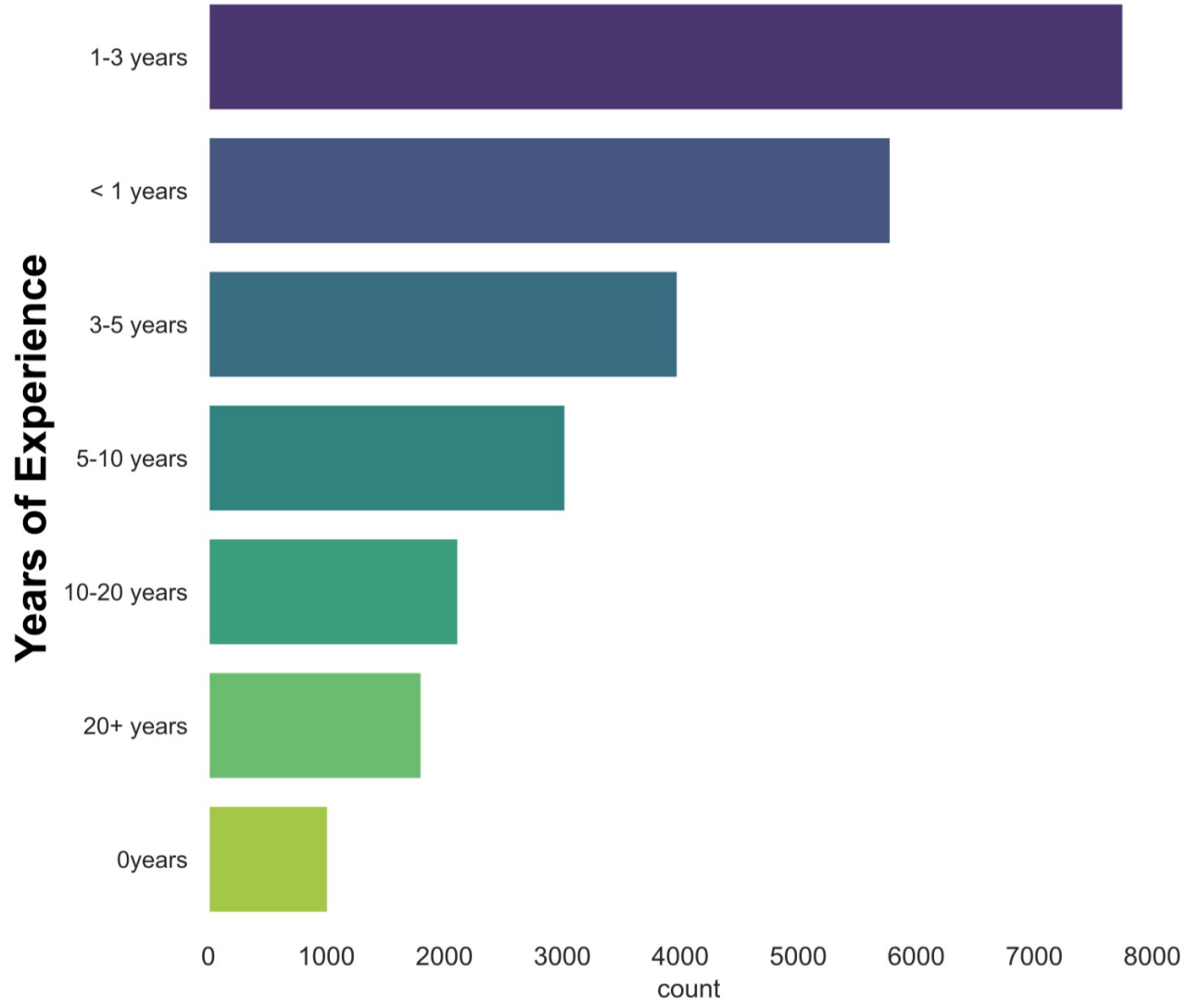
kaggle

Education



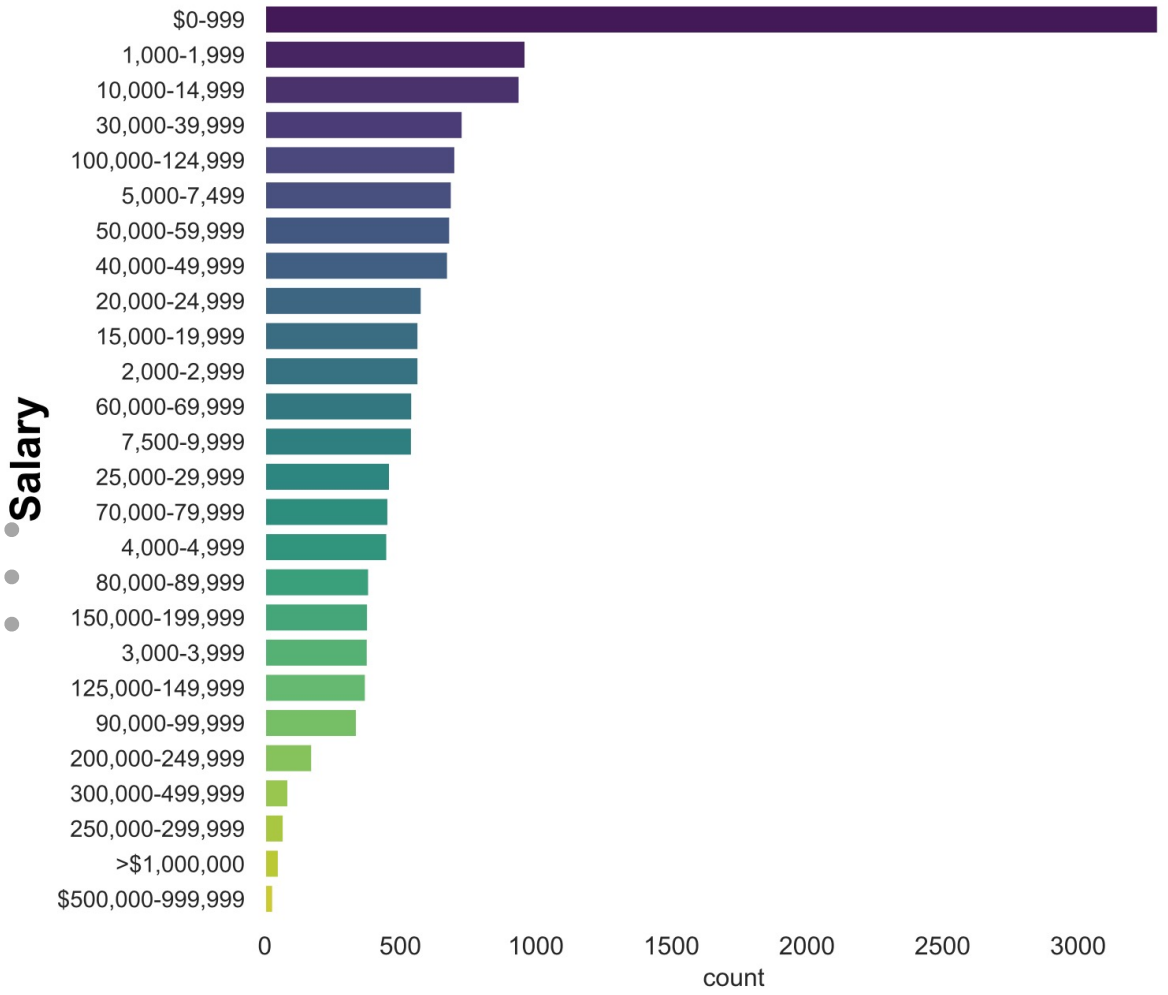
kaggle

Experience



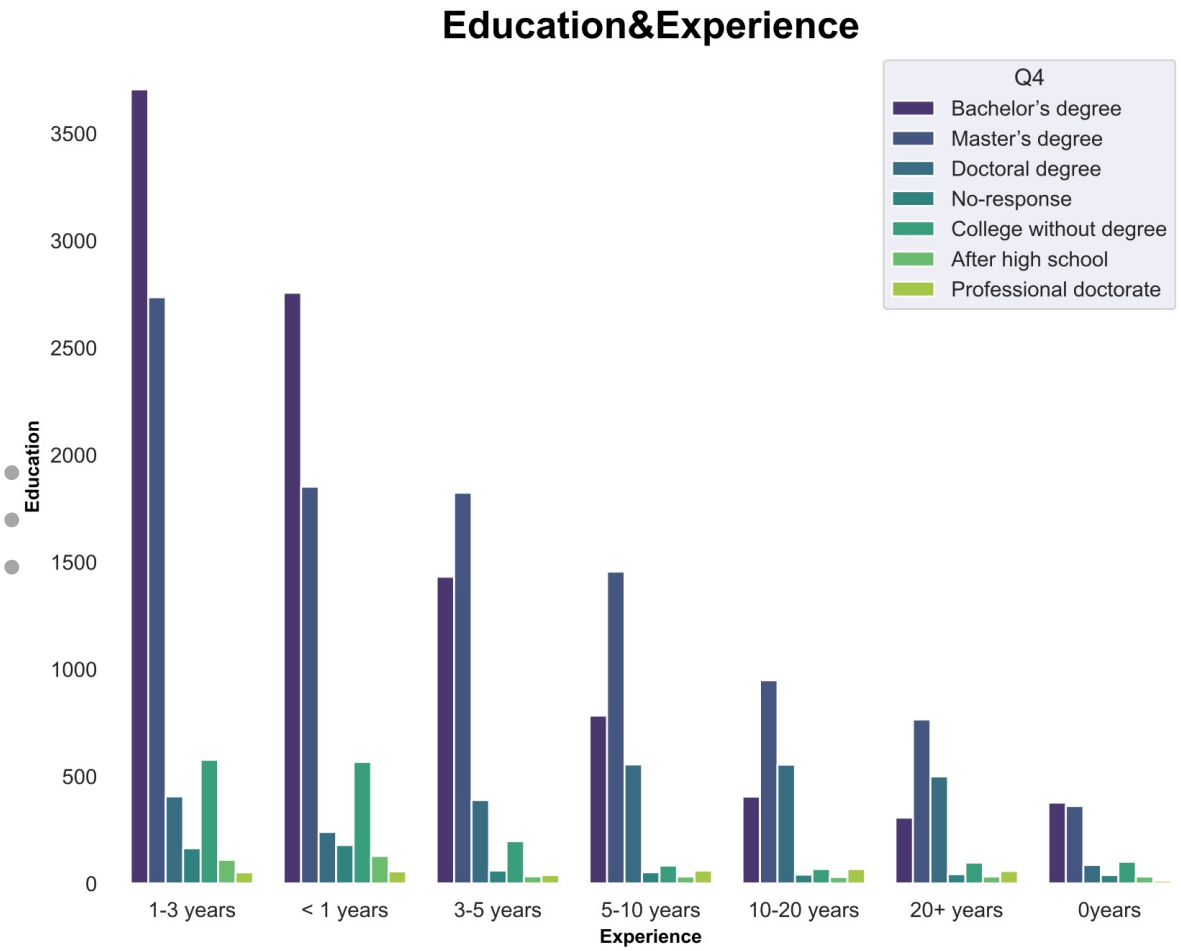
kaggle

Salary

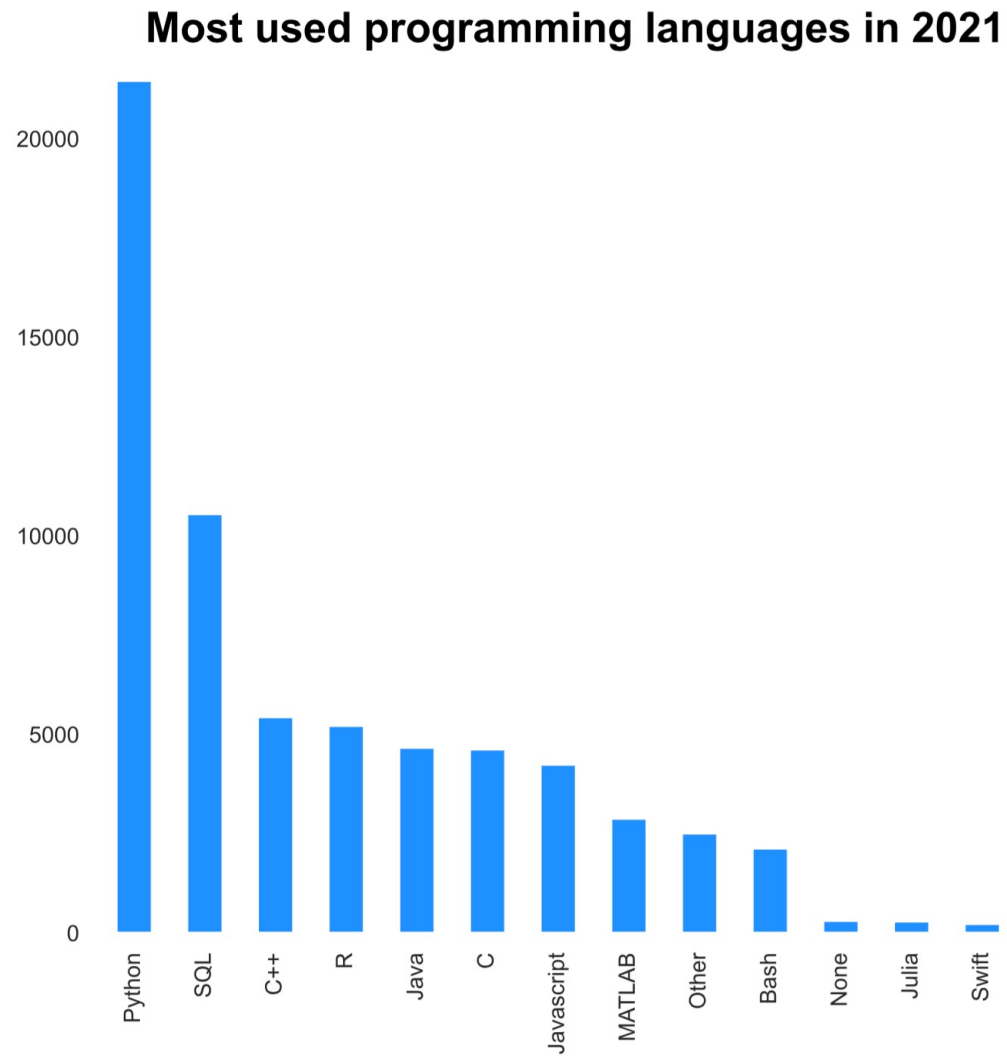


kaggle

Education&Experience



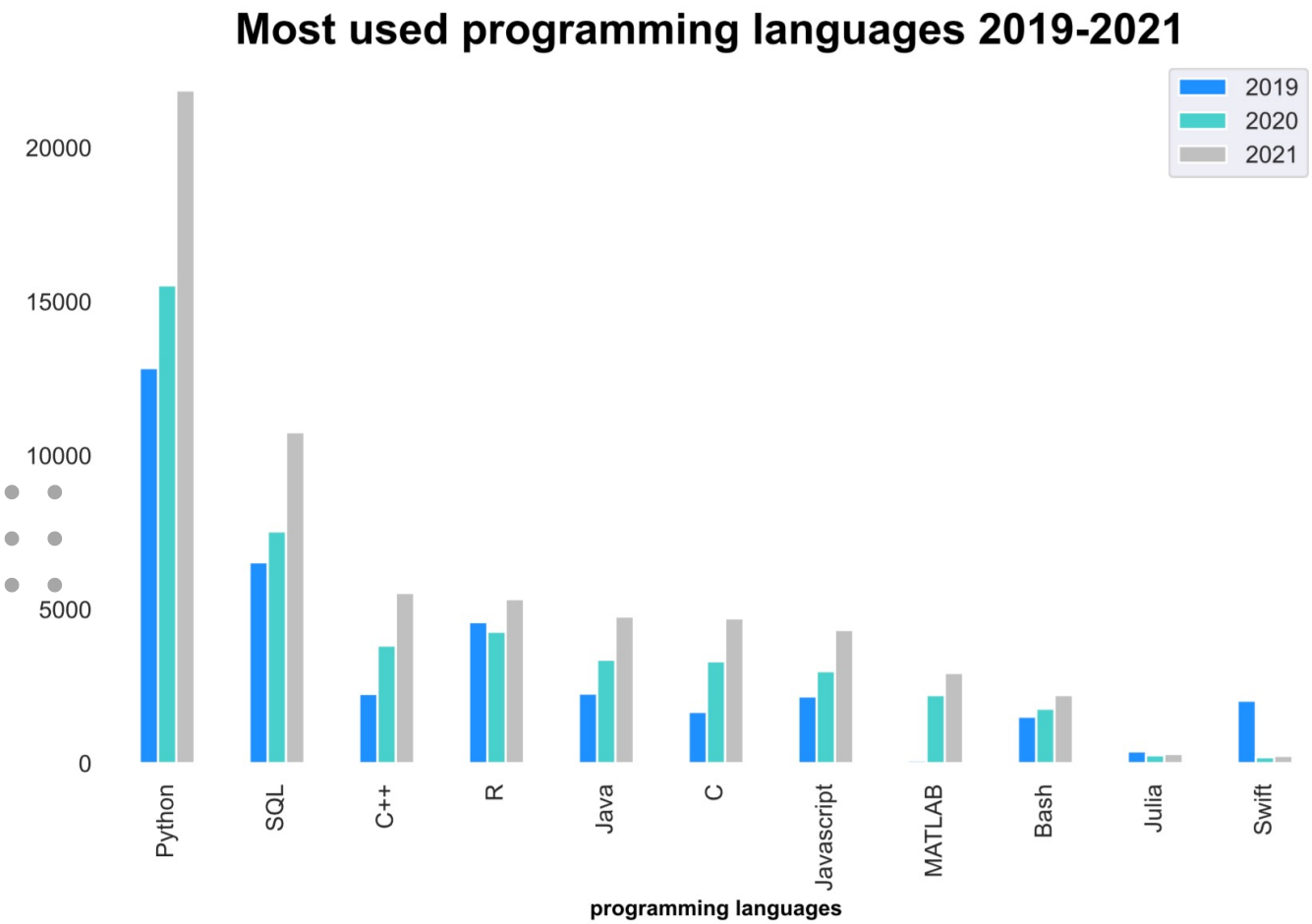
Most used programming languages for 2021



kaggle

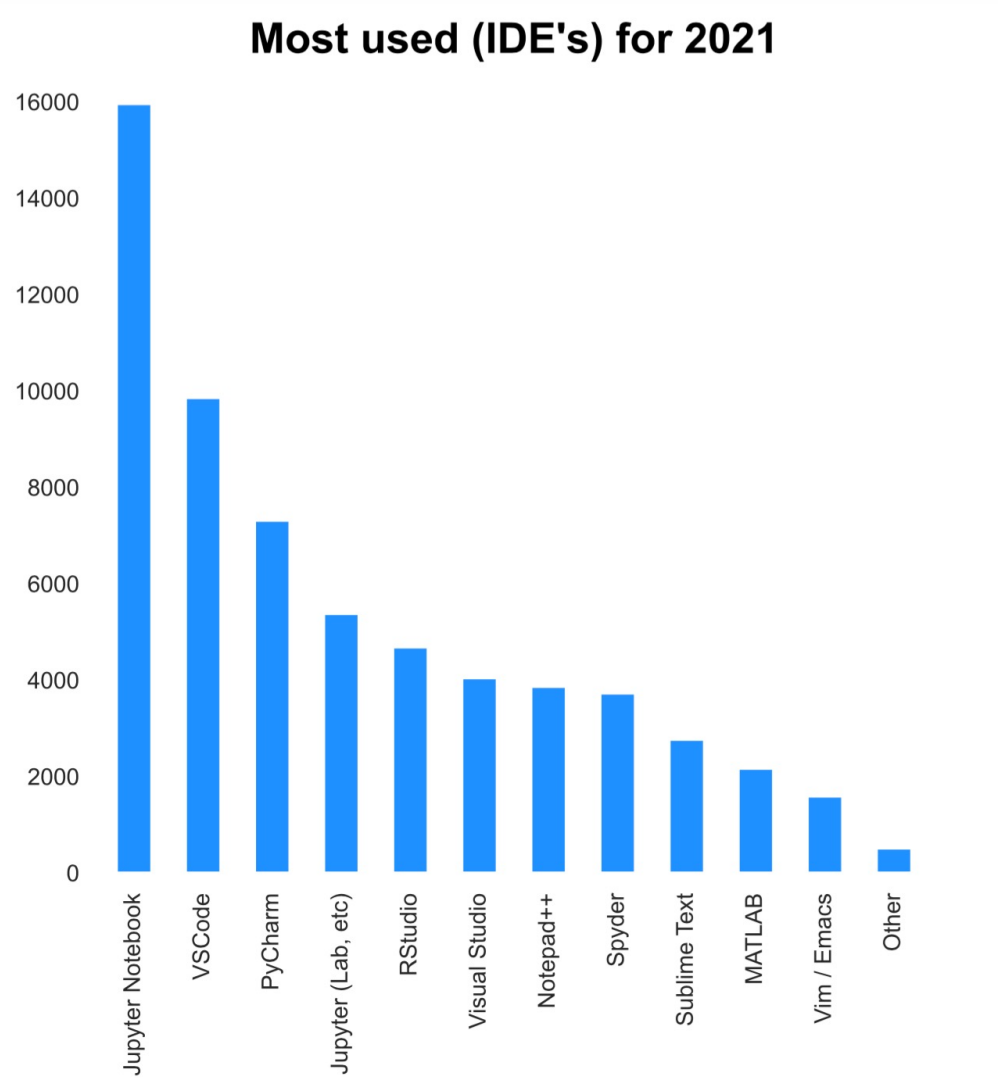
Most used programming languages 2019-2021

2019-2021

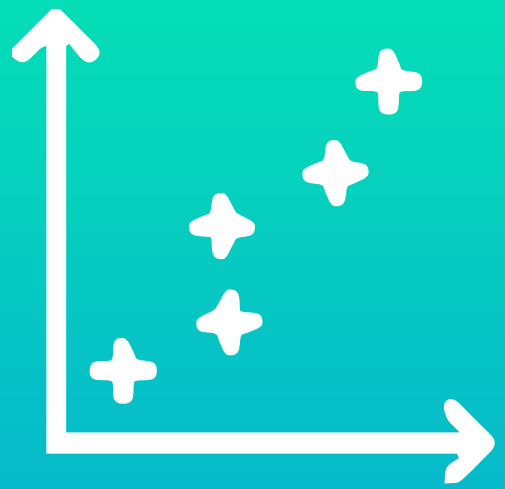


kaggle

Most used (IDE's) for 2021



kaggle



Correlation between experience and salary

0.193

kaggle

Thanks!