2021 Kaggle Survey

A focus on Nigerian Kagglers

Author

DR ADEAYO SOTAYO

Social media

Twitter - drsotayo

Github - drsotayo

LinkedIn - linkedin.com/in/adeayosotayo/

Introduction

Kaggle is an online community of students, data scientists and machine learning practitioners from around the world. Kaggle allows people to undertake courses, learn, find and publish datasets, explore and build models in a web-based data-science environment, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges.

This year (2021), Kaggle carried out its 5th Machine Learning & Data Science Survey, which received over 25,000 responses from users globally. It covered data science and machine learning questions for students and professionals across several industries, including their educational backgrounds, salaries, sectors, programming languages and techniques and many more.

Focusing on the 2021 dataset, this analysis focuses on a unique and specific subset of the dataset, i.e., Nigerian Kagglers, represented in this survey.

Introduction

This analysis aims to answer some questions related to Nigerian Kagglers, which are shown below:

What's the **uptake of Nigerians on Kaggle** and **how has it changed** over the past few years, and **how will it change** over the next few years?

In a **global and African context**, how are Nigerians performing on Kaggle?

What **educational qualifications** do Nigerian Kagglers possess?

What are the **age ranges** and **gender distributions** of Nigerian Kagglers?

How much do Nigerian Kagglers **earn**?

Introduction

This analysis aims to answer some questions related to Nigerian Kagglers, which are shown below (cont'd):

- What **jobs** do Nigerian Kagglers do, and how does the salary vary between these jobs?

- What **sectors** do they work in?

- What are the most common **programming languages** and **integrated development environments (IDEs)** used by Nigerian kagglers?

- What are the commonly used **machine learning frameworks** and **machine learning algorithms** used by Nigerian Kagglers?

- What **platforms and media sources** do Nigerian Kagglers use to learn more about data science?

Top 10 countries with the most Kagglers
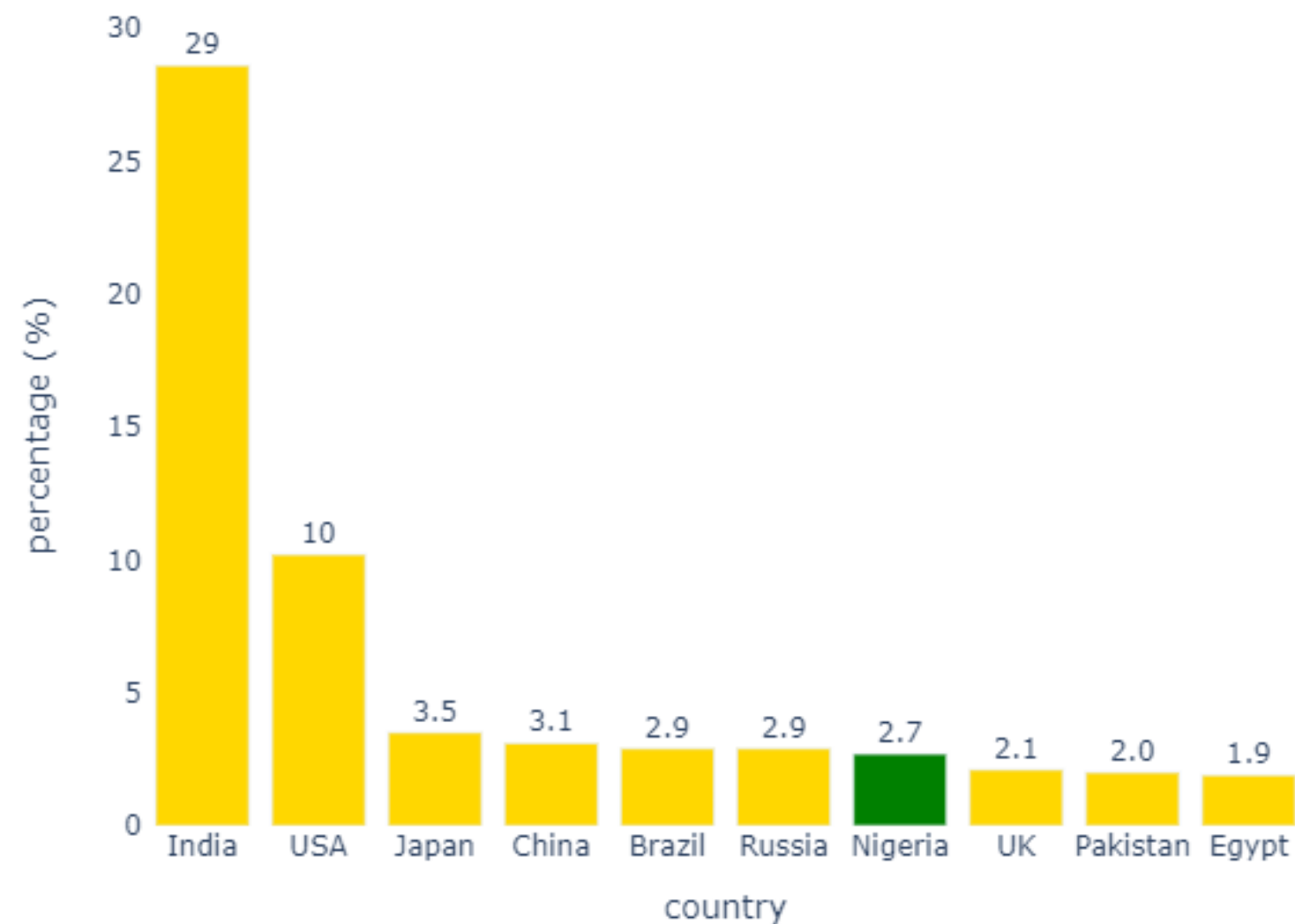
- The top three countries are India, USA and Japan.

- In comparison to other countries, Nigeria Kagglers make up about 2.7% and Nigeria ranks 7th.
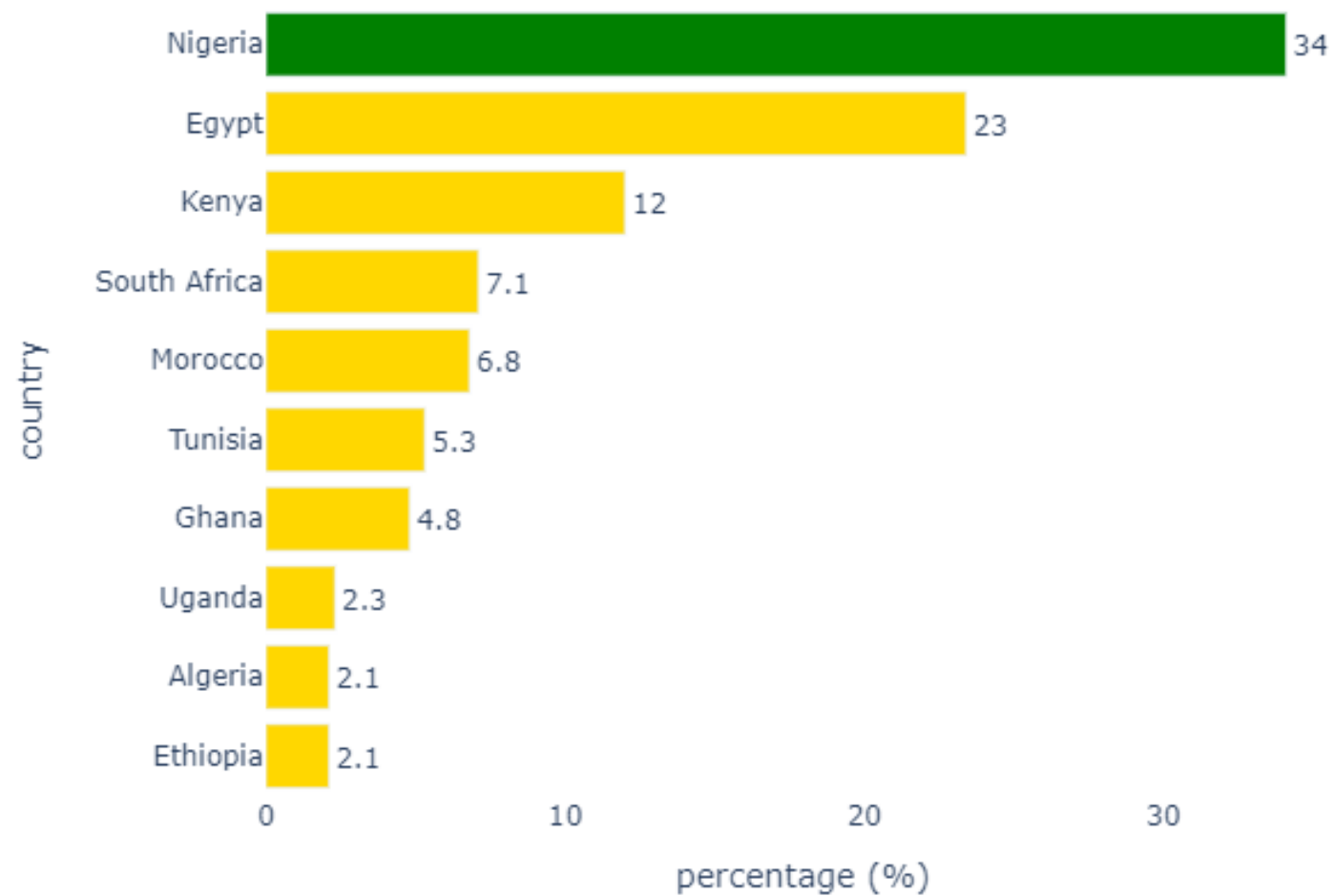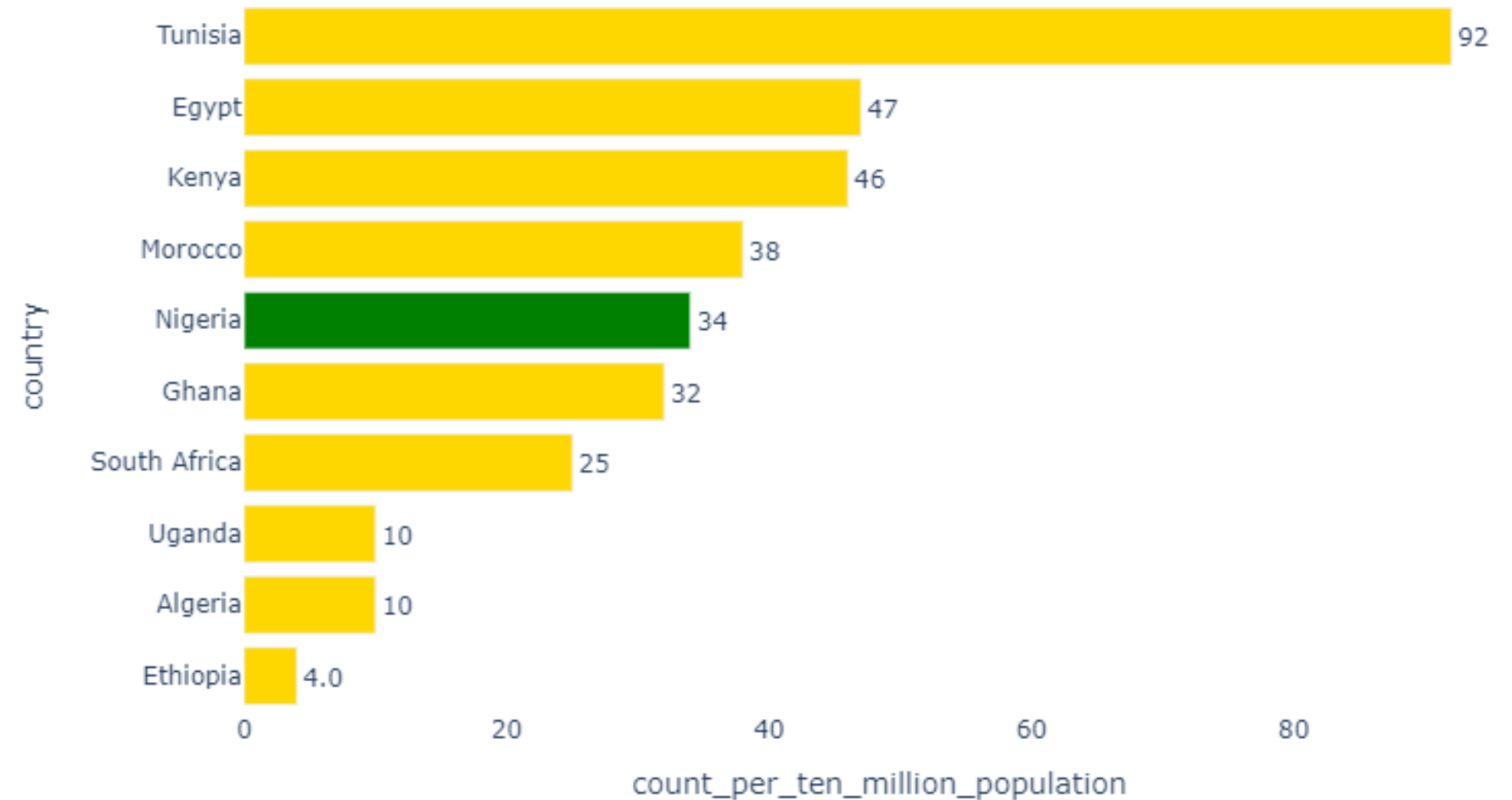
- Coincidentally, Nigeria is also ranked as the 7th most populated country in the world.

Top countries within Africa

Ranking of African countries with the most Kagglers

country	percentage (%)
Nigeria	34
Egypt	23
Kenya	12
South Africa	7.1
Morocco	6.8
Tunisia	5.3
Ghana	4.8
Uganda	2.3
Algeria	2.1
Ethiopia	2.1

Ranking of African countries with the most Kagglers (per ten million population)

| country	count_per_ten_million_population
Tunisia	92
Egypt	47
Kenya	46
Morocco	38
Nigeria	34
Ghana	32
South Africa	25
Uganda	10
Algeria	10
Ethiopia	4.0

Top countries within Africa

- Out of the 64 countries that took part in the Kaggle Survey, 10 countries are from Africa.

- There are 54 countries within Africa, which means that less than 20 % completed the Kaggle survey.

- Within Africa, **Nigeria Kagglers** make up 34% and **ranked number one in Africa**. It may be worth noting that Nigeria is also the most populated country in Africa.

- The top three African countries on Kaggle are **Nigeria, Egypt and Kenya.**

- The top three African countries (per million population) on Kaggle are **Tunisia, Egypt, and Kenya.**

- Ranking of the African countries **per ten million population** shows that **Nigeria is ranked 5th.**

- For every **ten million people in Nigeria**, only **34** of them completed the 2021 Kaggle survey.
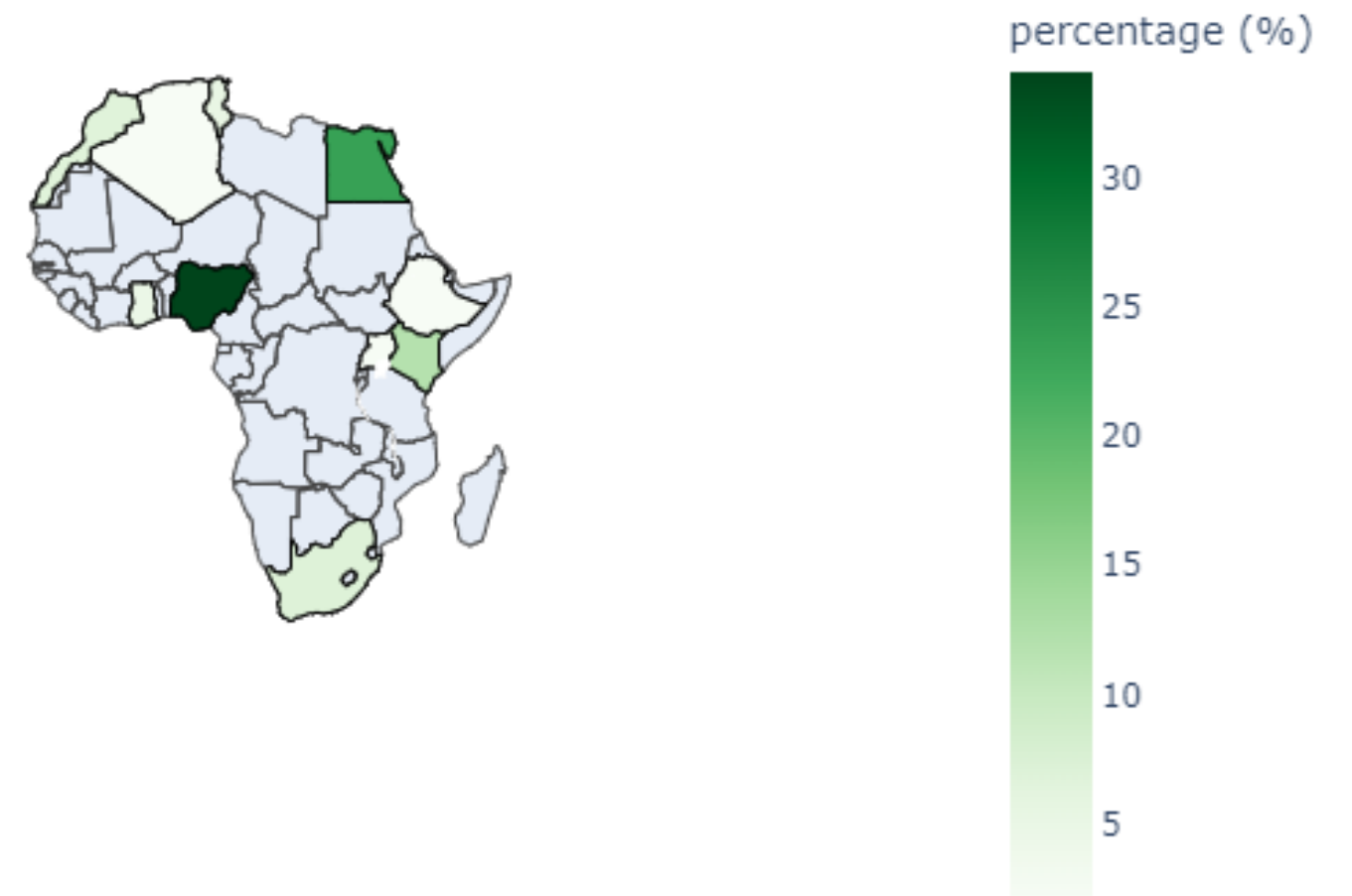
African Kagglers

Nigerian Kagglers over the past five (5) years

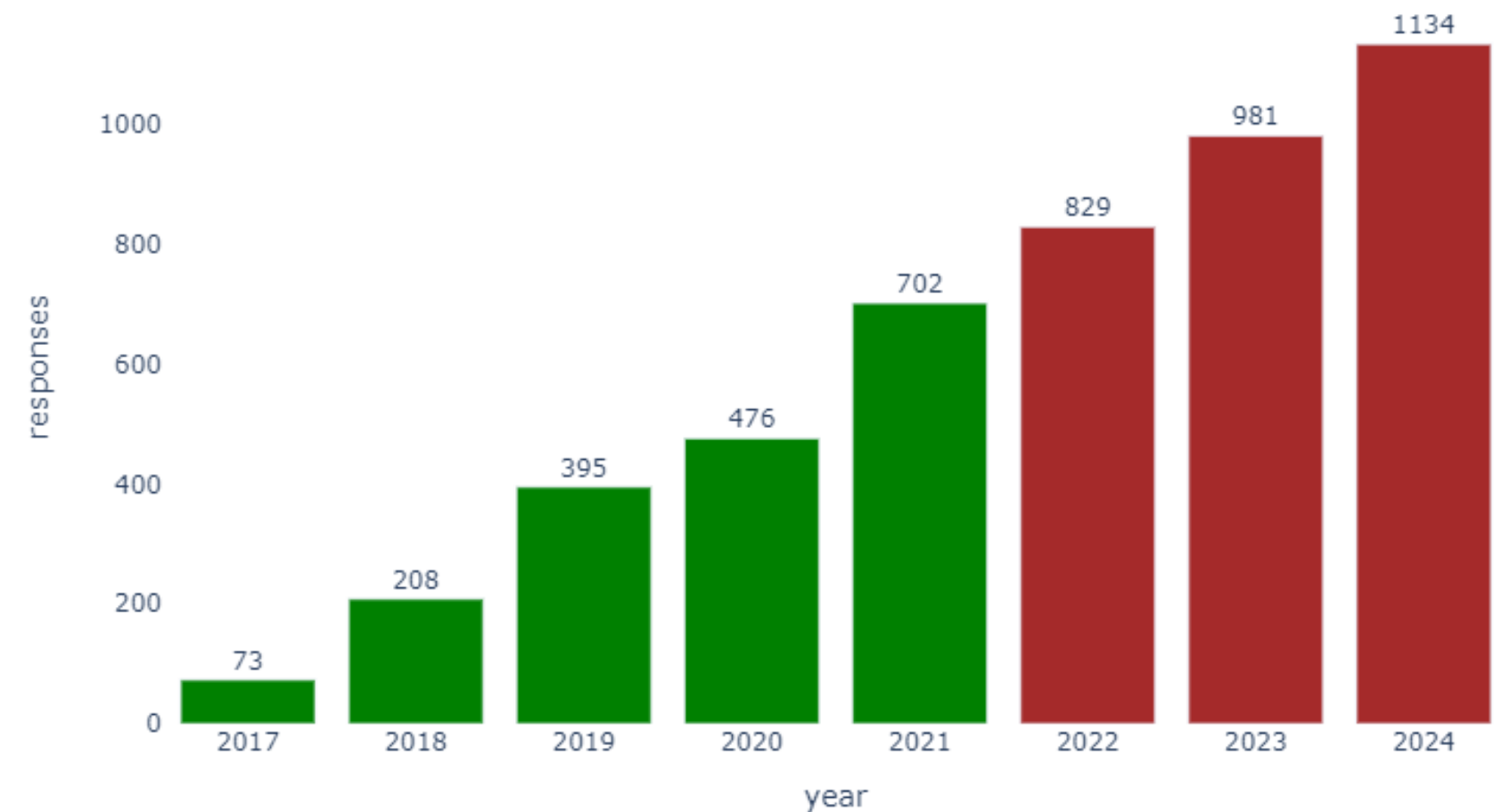
Who are Nigerian Kagglers
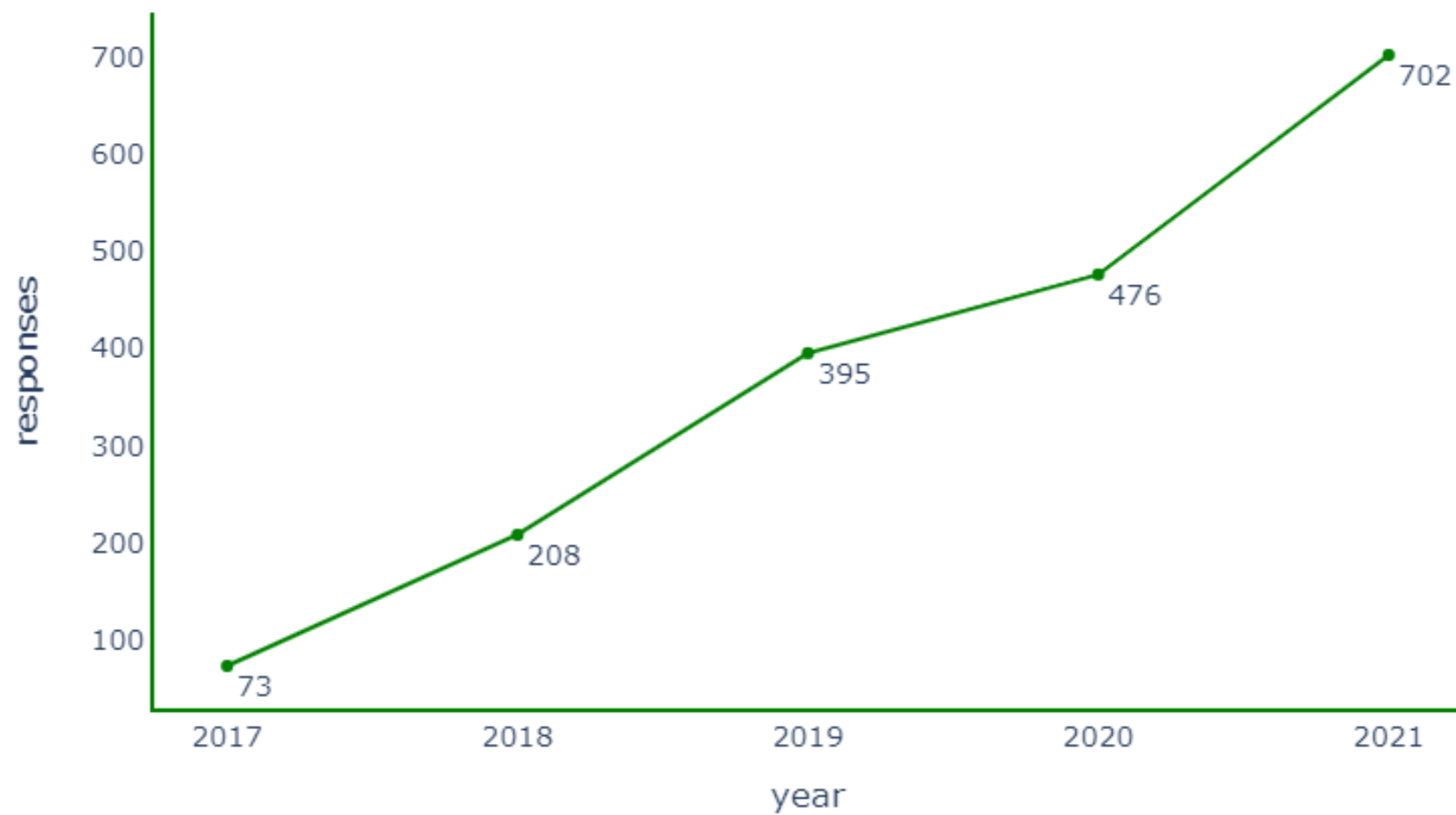
- The total number of responses for the 2021 Kaggle Survey from **Nigeria** was 702.

- There has been a steady growth in the number of responses over the past five years.

- The number of responses from Nigerian Kagglers in **2021** is about **10 times greater** than in 2017.

Forecast of Nigerian Kagglers for the next three years using Linear Regression

- Based on the responses from the past five years, the number of Nigerian Kagglers is expected to keep growing over the next few years.

- A simple linear regression model (using scikit-learn) shows that the number of Nigerian Kagglers will exceed 800 next year.

- The model also shows that the number of responses from Nigeria Kagglers will approach and exceed 1,000 in about two-three years.

Forecast of Nigerian Kagglers for the next three years

Gender distribution of Nigerian Kagglers compared to the rest of the world
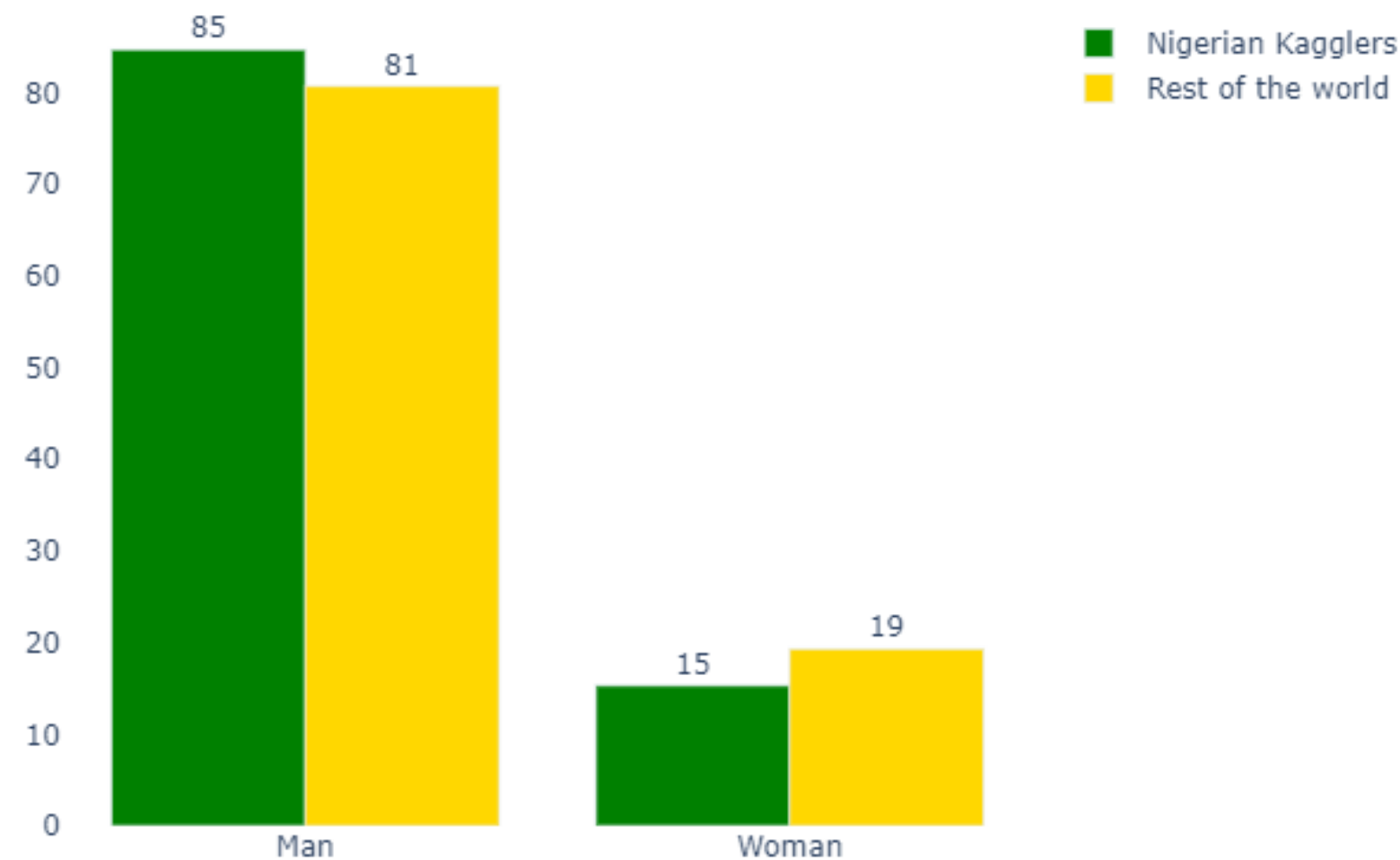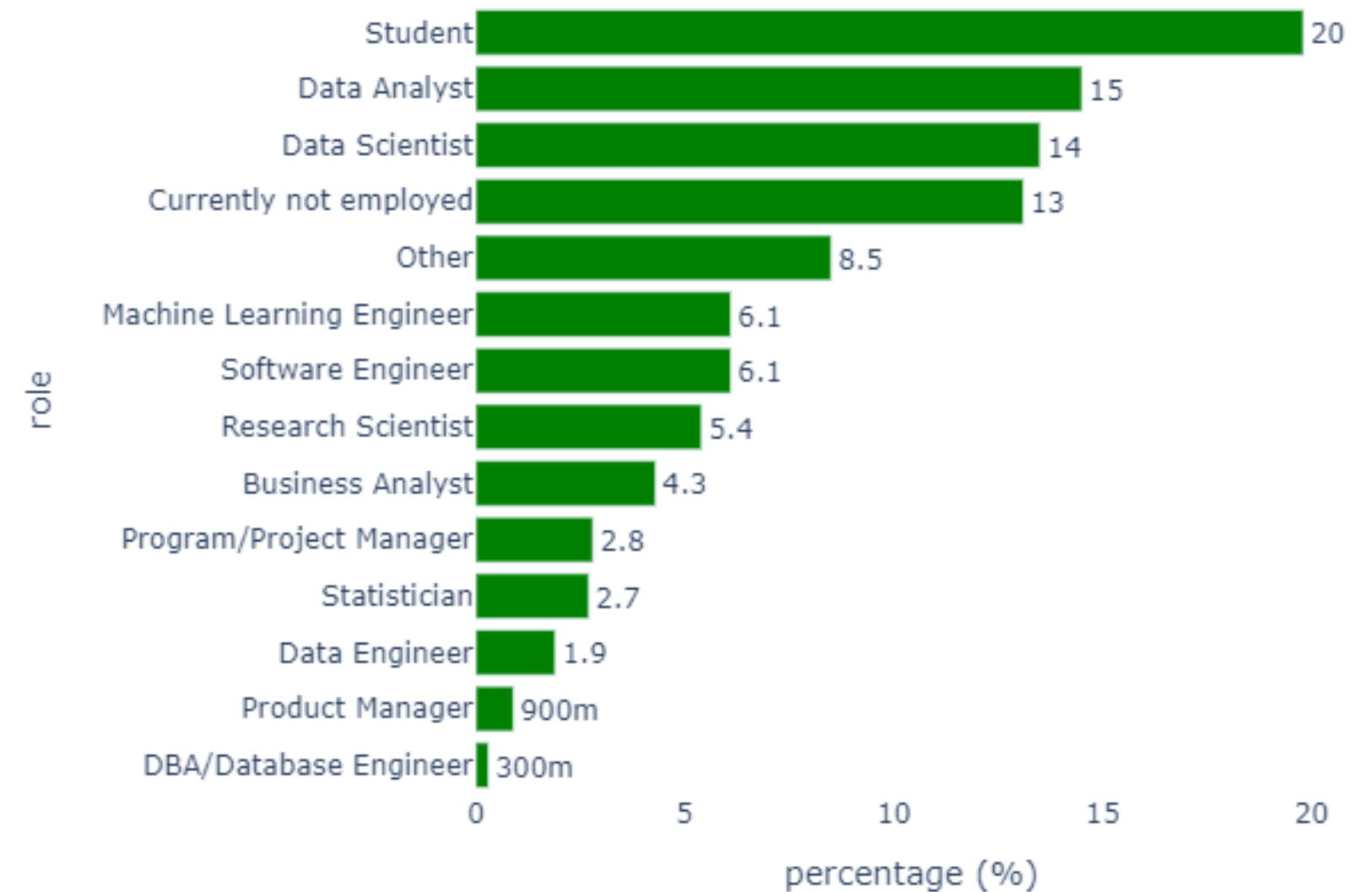
- 15% of Nigerian Kagglers are women and 85% are men. This reflects a large gender gap among Nigerian Kagglers.

- Compared to the rest of the world, the percentage of female Kagglers in Nigeria (15%) is lower than the rest of the world (19%).
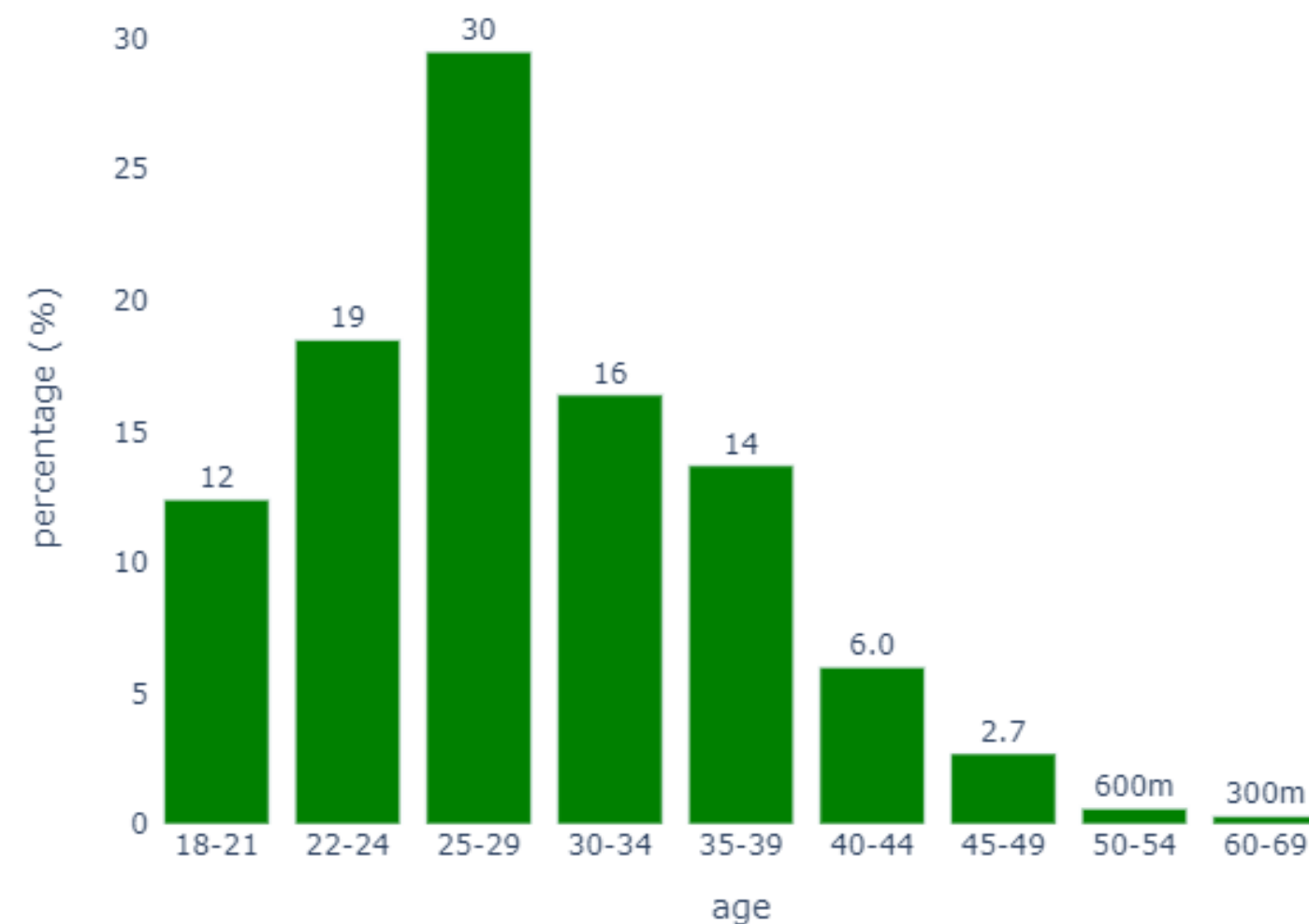
What do Nigerian Kagglers do?

- Most Nigerian Kagglers are Students, Data Analysts and Data Scientists.

- Similar to the rest of the world, Students form the majority of Nigerian Kagglers.

- Apart from students, the job titles for most Nigerian Kagglers are Data Analysts, Data Scientists and Software Engineers.

Job roles of Nigerian Kagglers

Age of Nigerian Kagglers

Age distribution of Nigerian Kagglers

- The largest proportion (30%) of Nigerian Kagglers are in the age range of 25 - 29.

- About 65% of Nigerian Kagglers are within the age of 22 - 34.

- 91% of Nigerian Kagglers are younger than 40.

- Look at the rest of the world in general, the higher the age, the lower the proportion/percentage.

- The largest proportion (19%) for the rest of the world falls between 18 and 21.

- In general, Kagglers worldwide are relatively young.
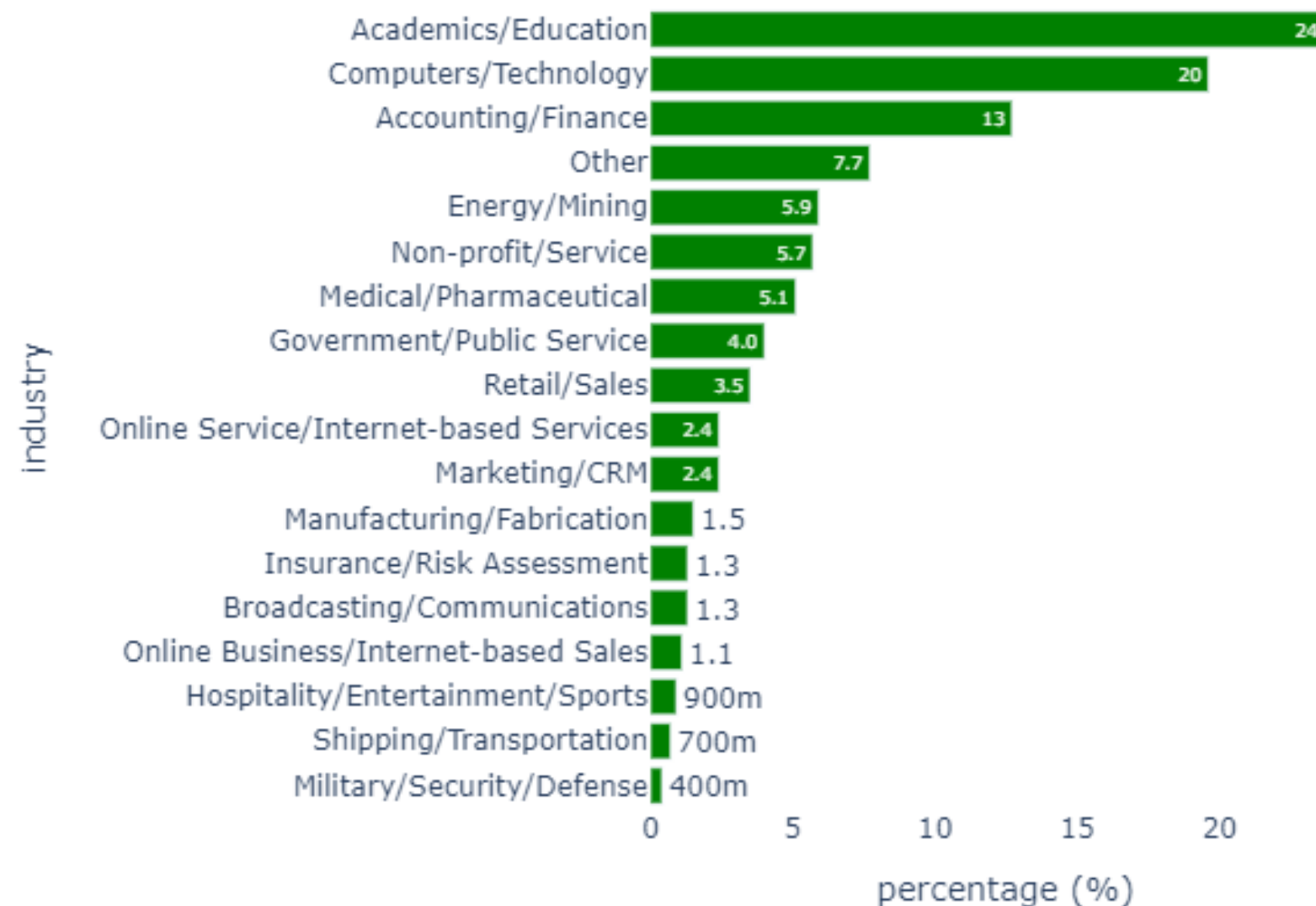
Education Level of Nigerian Kagglers

- The most common education level for Nigerian Kagglers is a Bachelor's degree (52%).

- About 93% of Nigerian Kagglers have at least a Bachelor's degree.

- This is slightly higher than the rest of the world of 89%.

Highest formal education of Nigerian Kagglers

Industry of Nigerian Kagglers
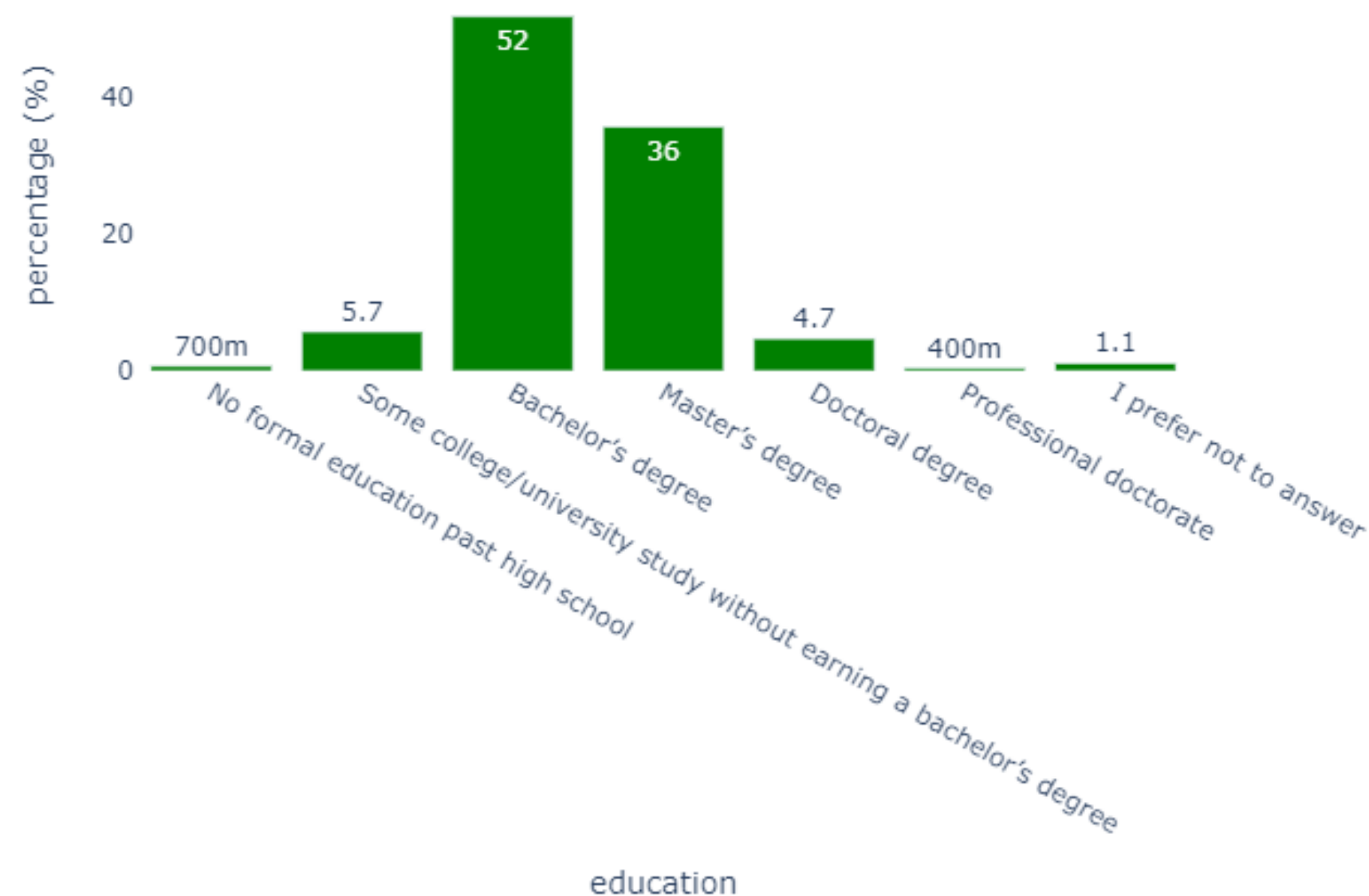
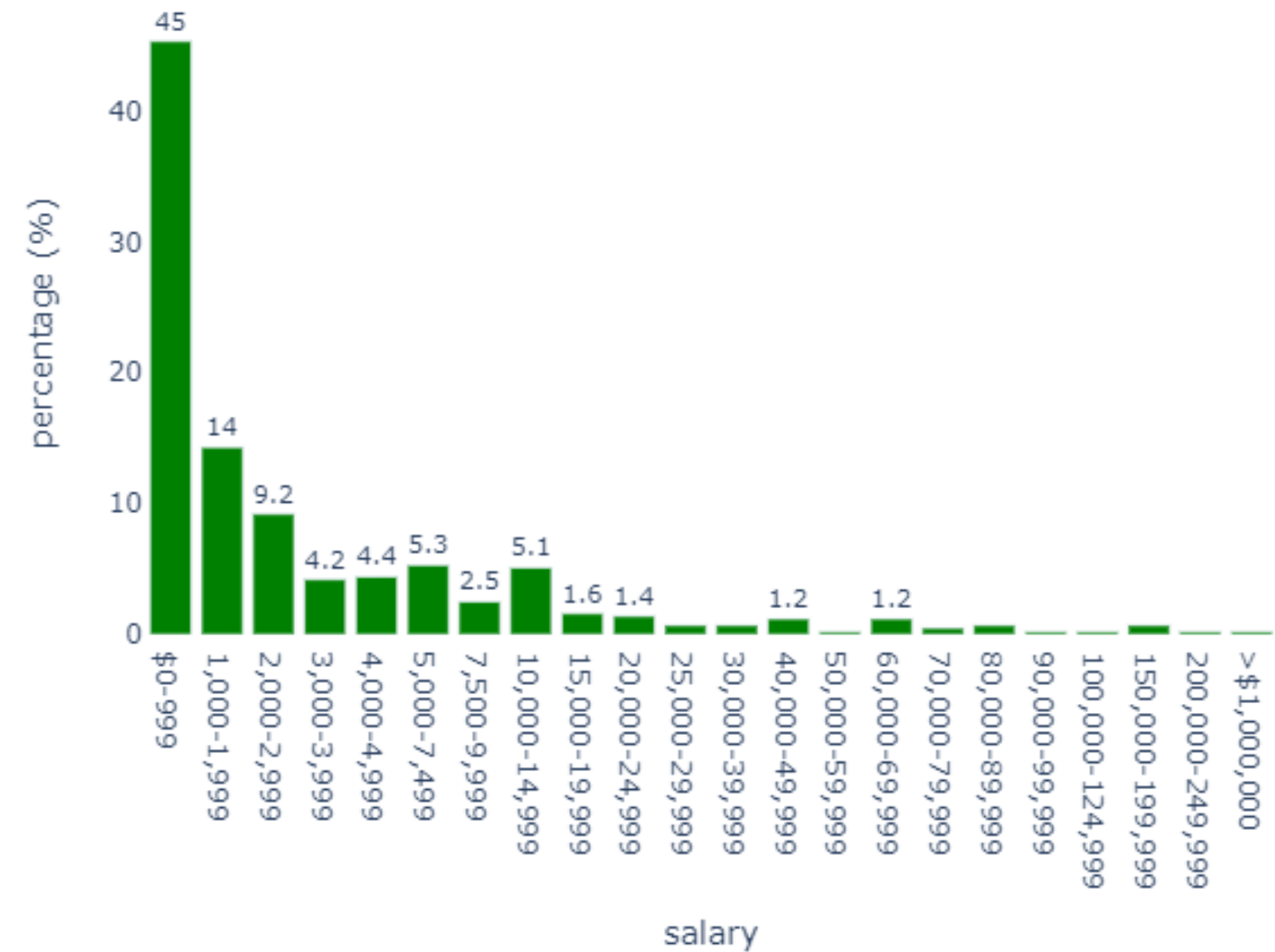# Industry of Nigeria Kagglers

- Most Nigeria Kagglers work in the following sectors – Academics/Education, Computers/Technology, and Accounting/Finance.

- These industries are also in line with the rest of the world, however, Computers/Technology ranks as number one in the rest of the world.

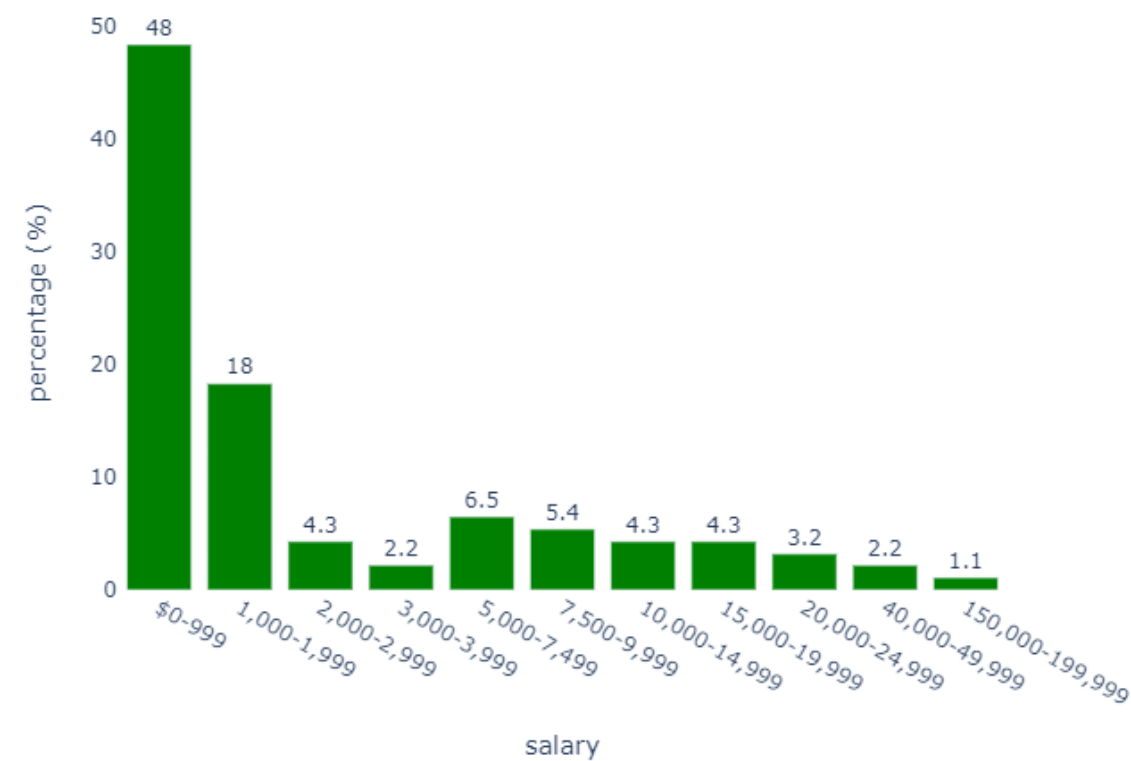# How much do Nigerian Kagglers earn?

Annual salary of Nigerian Kagglers

Breakdown of annual salary for the top three (3) professions of Nigerian Kagglers

Considering that the annual salary above takes into consideration all the job roles (including students), let's investigate the salary distribution across the three main job titles (Data Analyst, Data Scientist and Software Engineer).

Annual salary of Nigerian Kagglers that are Data Analysts

Annual salary of Nigerian Kagglers that are Data Scientists
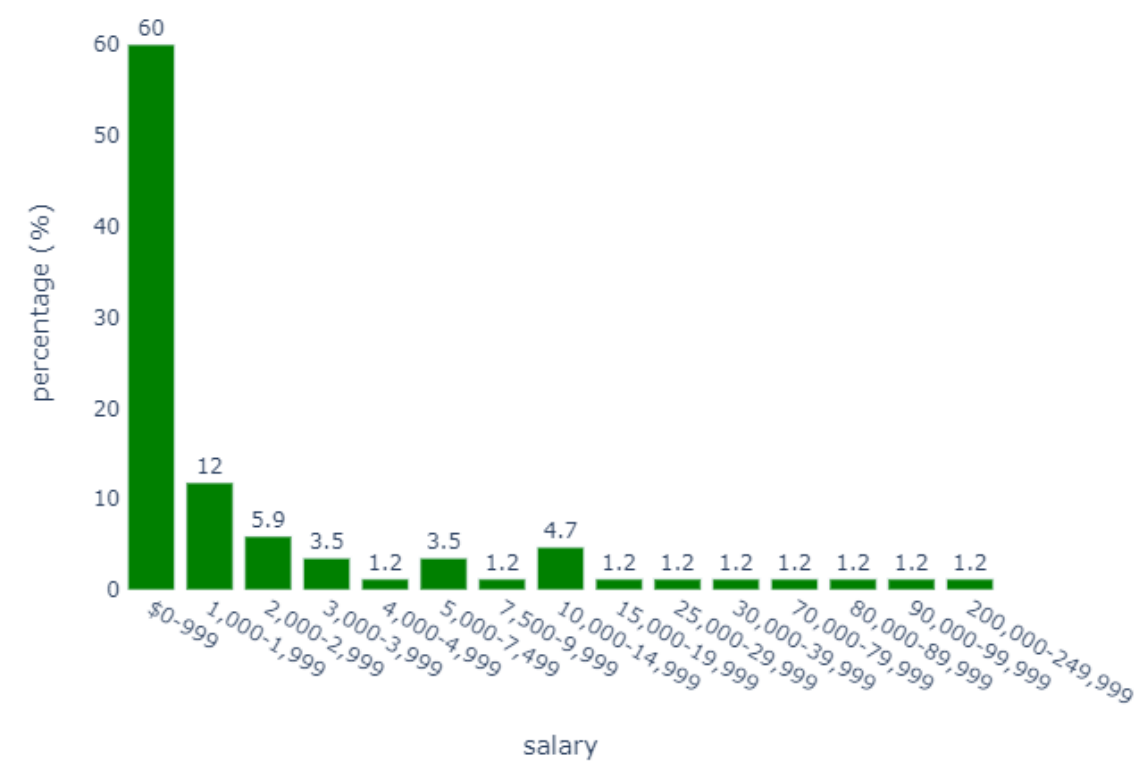
Annual salary of Nigerian Kagglers that are Software Engineers

How much do Nigerian Kagglers earn?

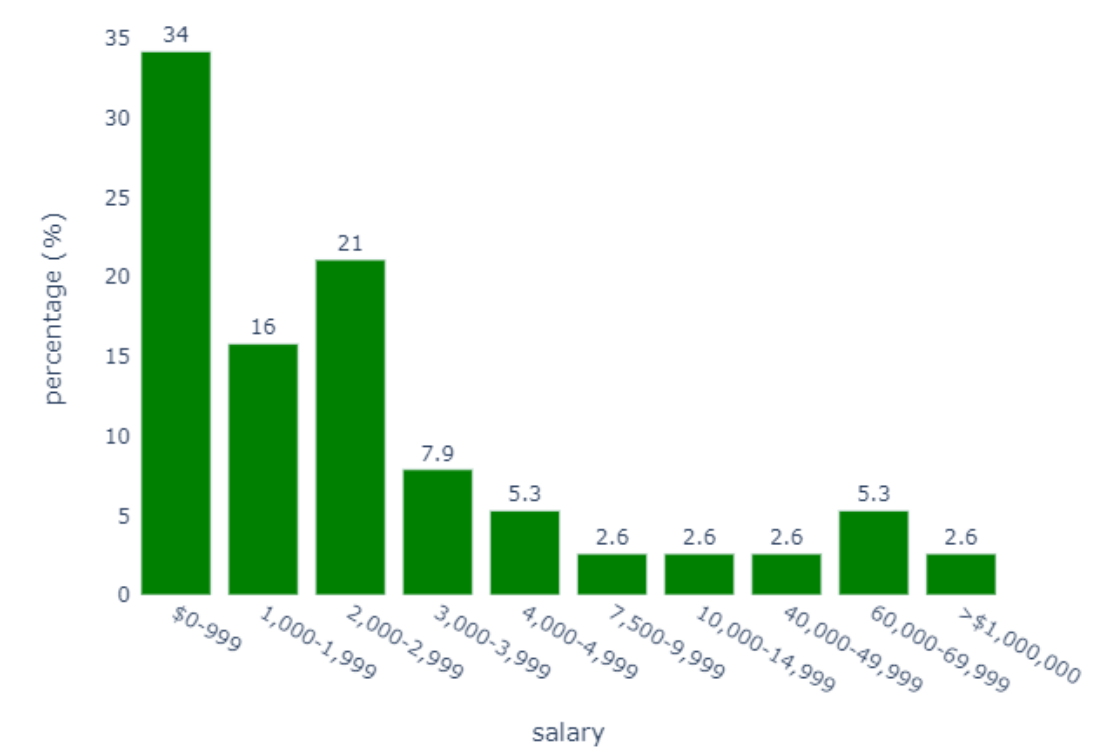Almost half of Nigerian Kagglers (45 percent) earn less than $1,000 per year.

About 60% of Nigerians earn less than $2,000 per year.

On the other hand, only about 20% of people from other nationalities earn under $1,000. Further analysis on the respective living wage and purchasing power for each country will give a more comprehensive overview.

In general, the salary distribution for Data Analysts, Data Scientists and Software Engineers (Nigerian Kagglers) are similar and mainly earn less than $3,000 per year.

A minor point is that there are more Software Engineers from the Nigerian Kaggle community who earn more (66 percent) than 1,000 USD per year compared to those that earn less (34 percent) than 1,000 USD per year.
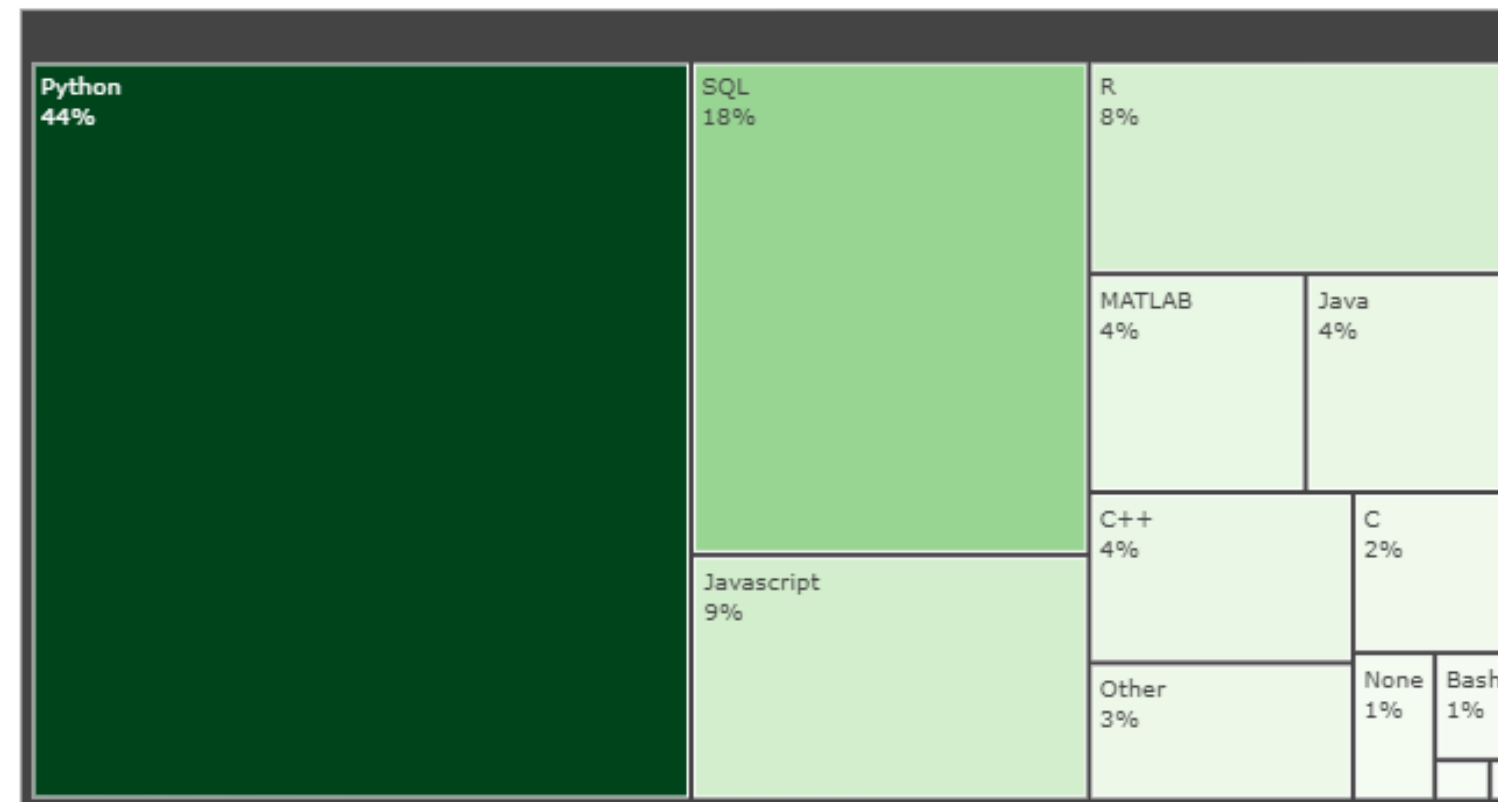
Years of programming for Nigerian Kagglers

How long have Nigerian Kagglers been programming?
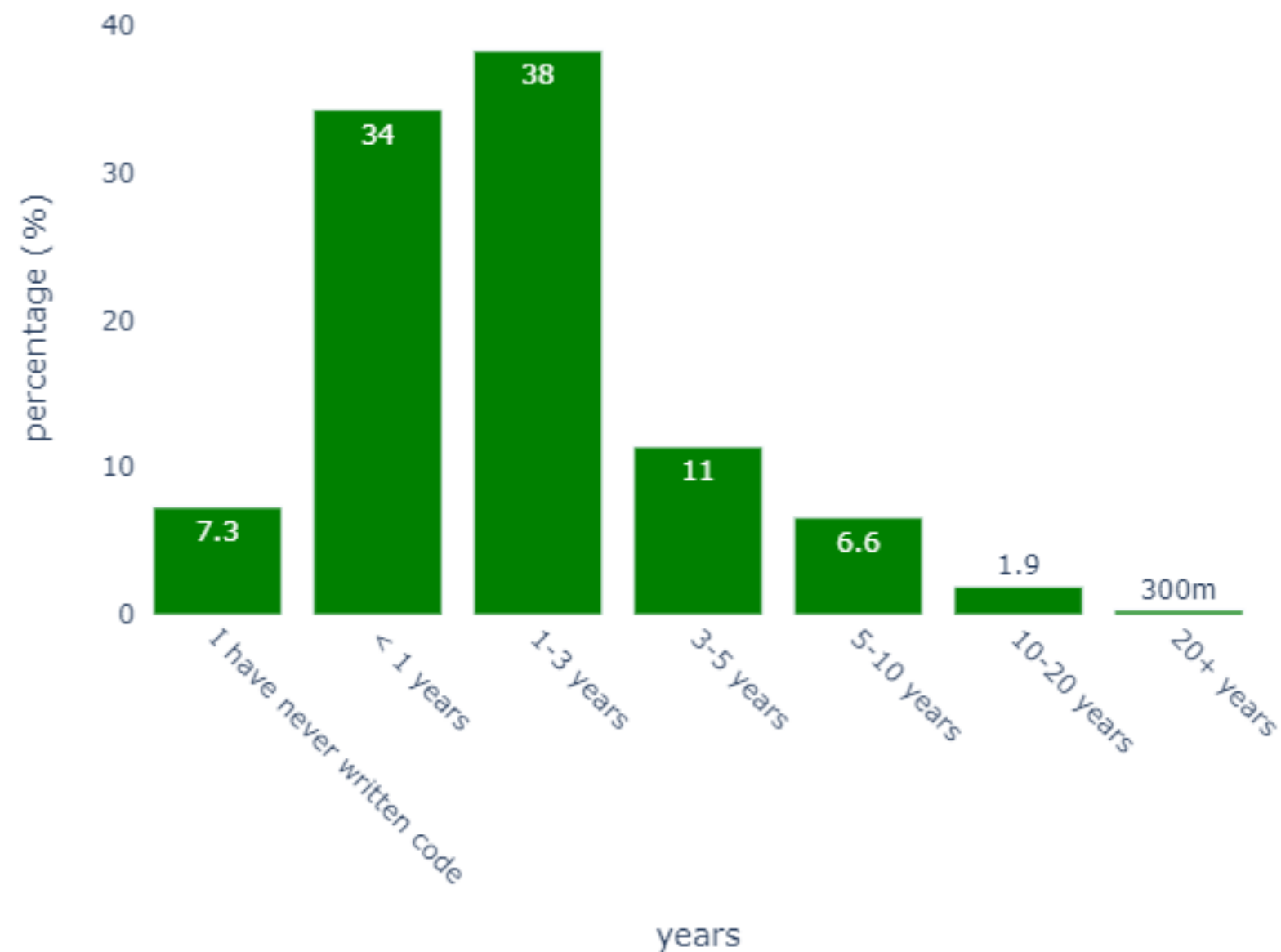
- Most Nigerian Kagglers (72%) have less than 3 years of programming experience.

- 7% of Nigerian Kagglers have never written code.

- Only about 9% of Nigerian Kagglers have over 5 years of programming experience.

What programming languages do Nigerian Kagglers use?

Programming language regularly used by Nigerian Kagglers

| Python 44% | SQL 18% | R 8% | | |
		MATLAB 4%	Java 4%	
		C++ 4%	C 2%	
	Javascript 9%	Other 3%	None 1%	Bash 1%

- The top programming language for Nigerian Kagglers is Python, which is similar to the rest of the world.

- Most Nigerian Kagglers use Python, SQL and Javascript.

- Javascript ranks third compared to C++ which ranks third for the rest of the world.

# Integrated development environment (IDE) used by Nigerian Kagglers

- Jupyter Notebook is the most popular IDE among Nigerian Kagglers and the rest of the world.

- Similarly, the three most popular IDEs among Nigerian Kagglers and the rest of the world are Jupyter Notebook, VScode and PyCharm.

IDES used by Nigerian Kagglers
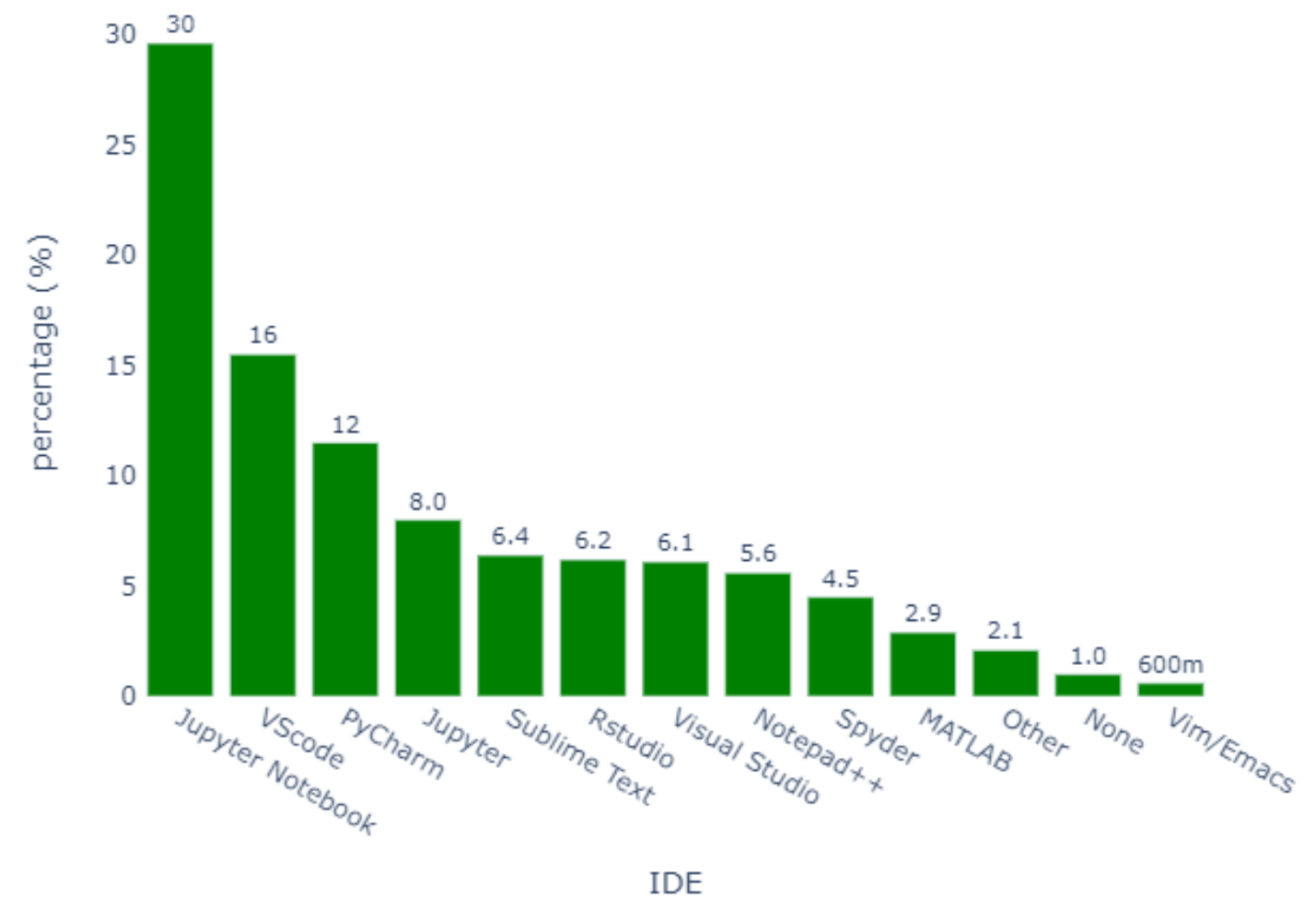
How long have Nigerian Kagglers been using Machine Learning methods?
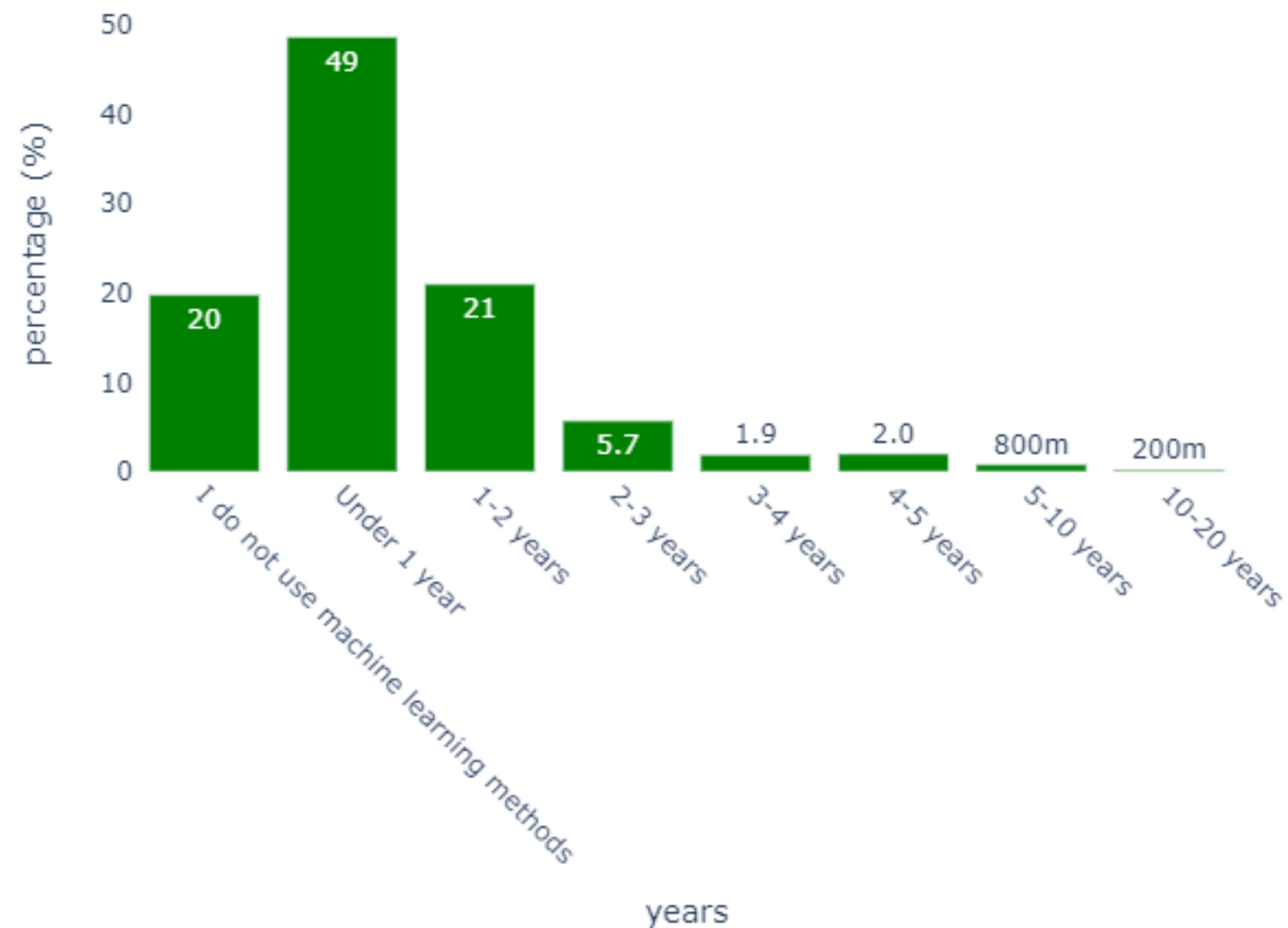
- 70% of Nigerian Kagglers have been applying Machine Learning methods for under two years.

- 20% of Nigerian Kagglers do use Machine Learning methods.

Machine Learning frameworks used by Nigerian Kagglers

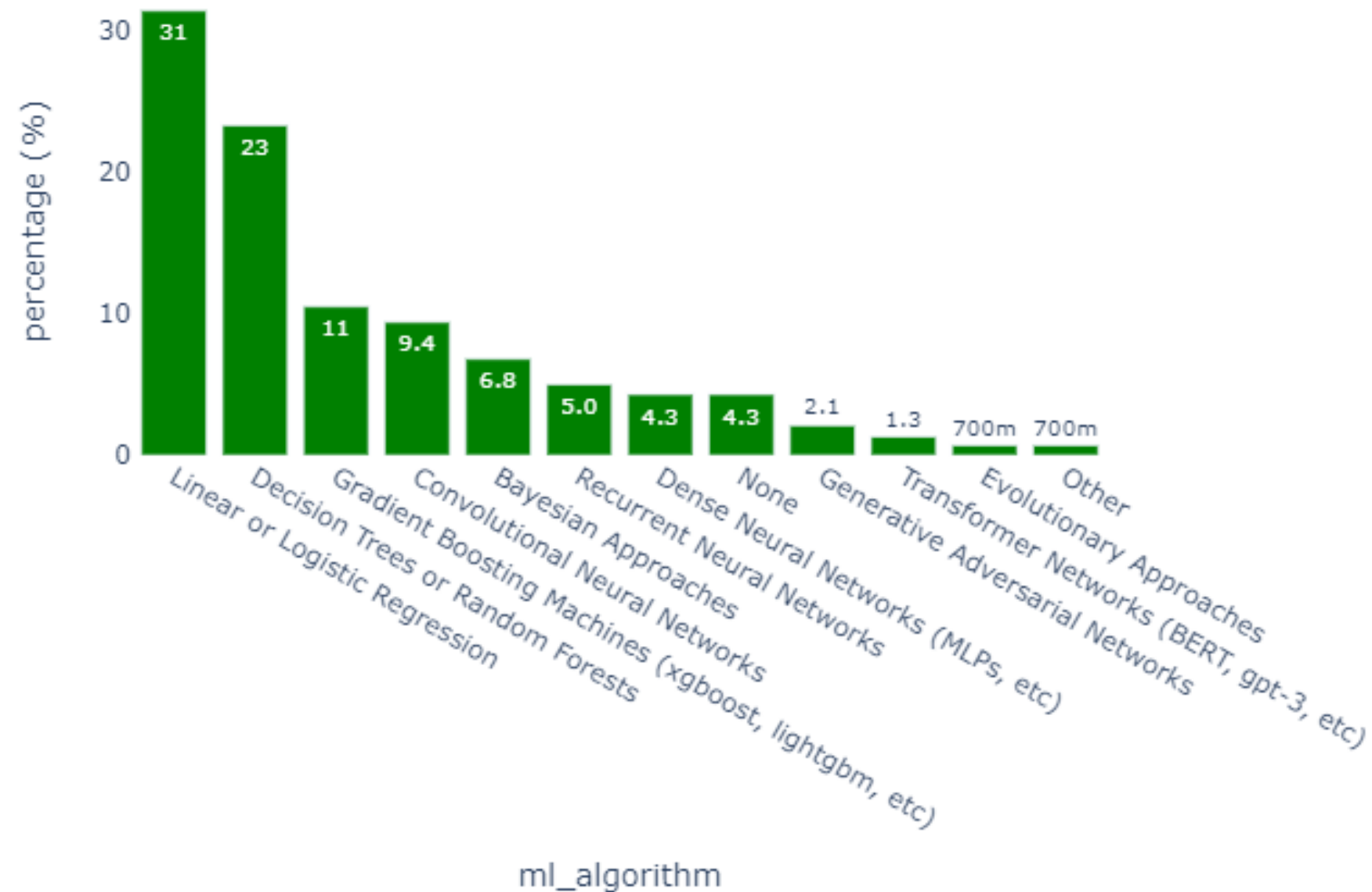Machine Learning frameworks regularly used by Nigerian Kagglers
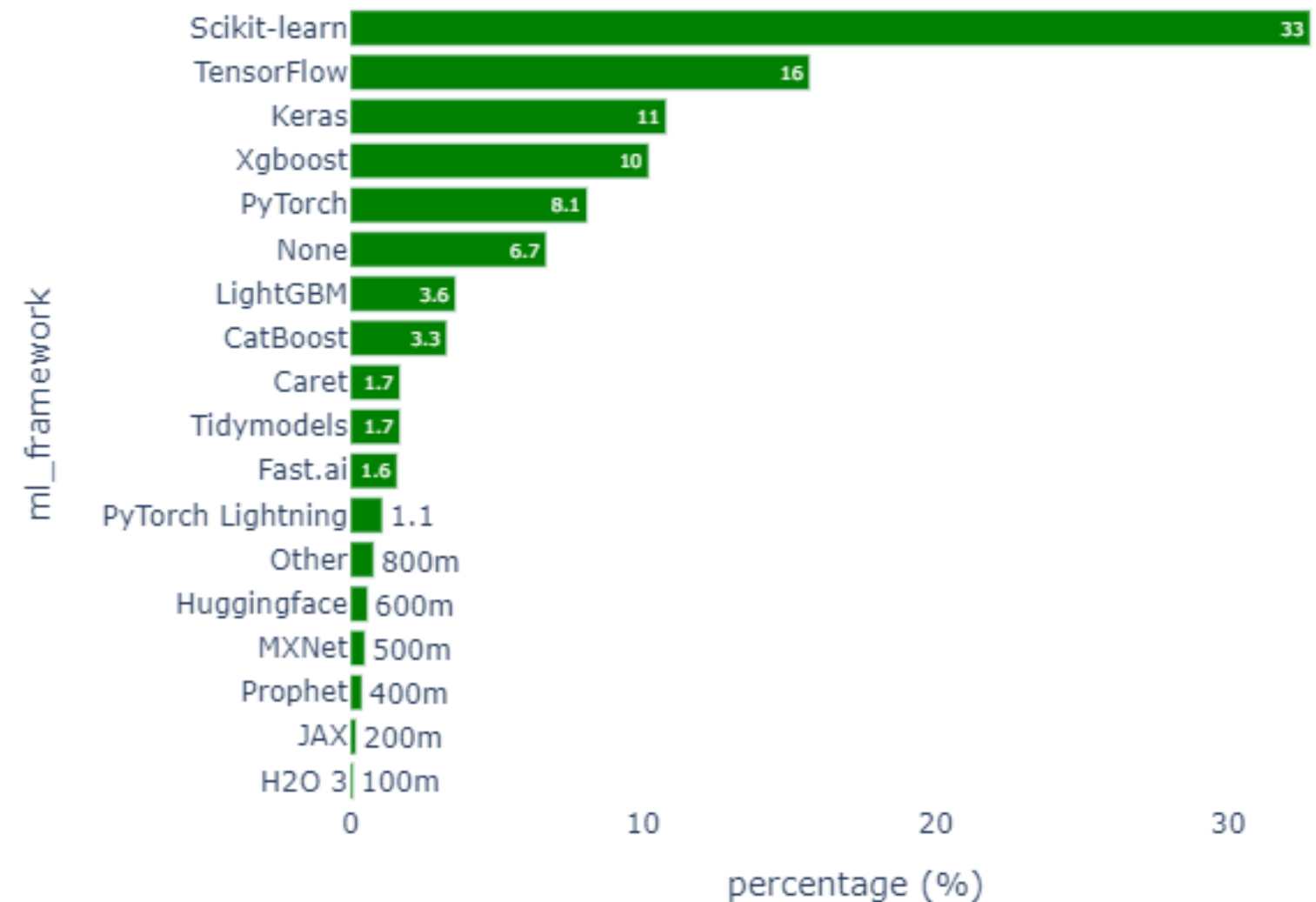
- Scikit-learn, TensorFlow and Keras are most popular Machine Learning frameworks used regularly by Nigerian Kagglers, this is also the same with the rest of the world.

Machine Learning algorithms used regularly by Nigerian Kagglers

Machine Learning algorithms used by Nigerian Kagglers

- Linear/Logistic Regression, Decision Trees/Random Forests and Gradient Boosting Machines(XGBoost, lightgbm, etc) are the most popular Machine Learning algorithms used regularly by Nigerian Kagglers, this is also the same with the rest of the world.

Database used by Nigerian Kagglers

Database used by Nigerian Kagglers

- The most common databases used by Nigerian Kagglers are MySQL, PostgreSQL, and Microsoft SQL Server. This is similar to the rest of the world.
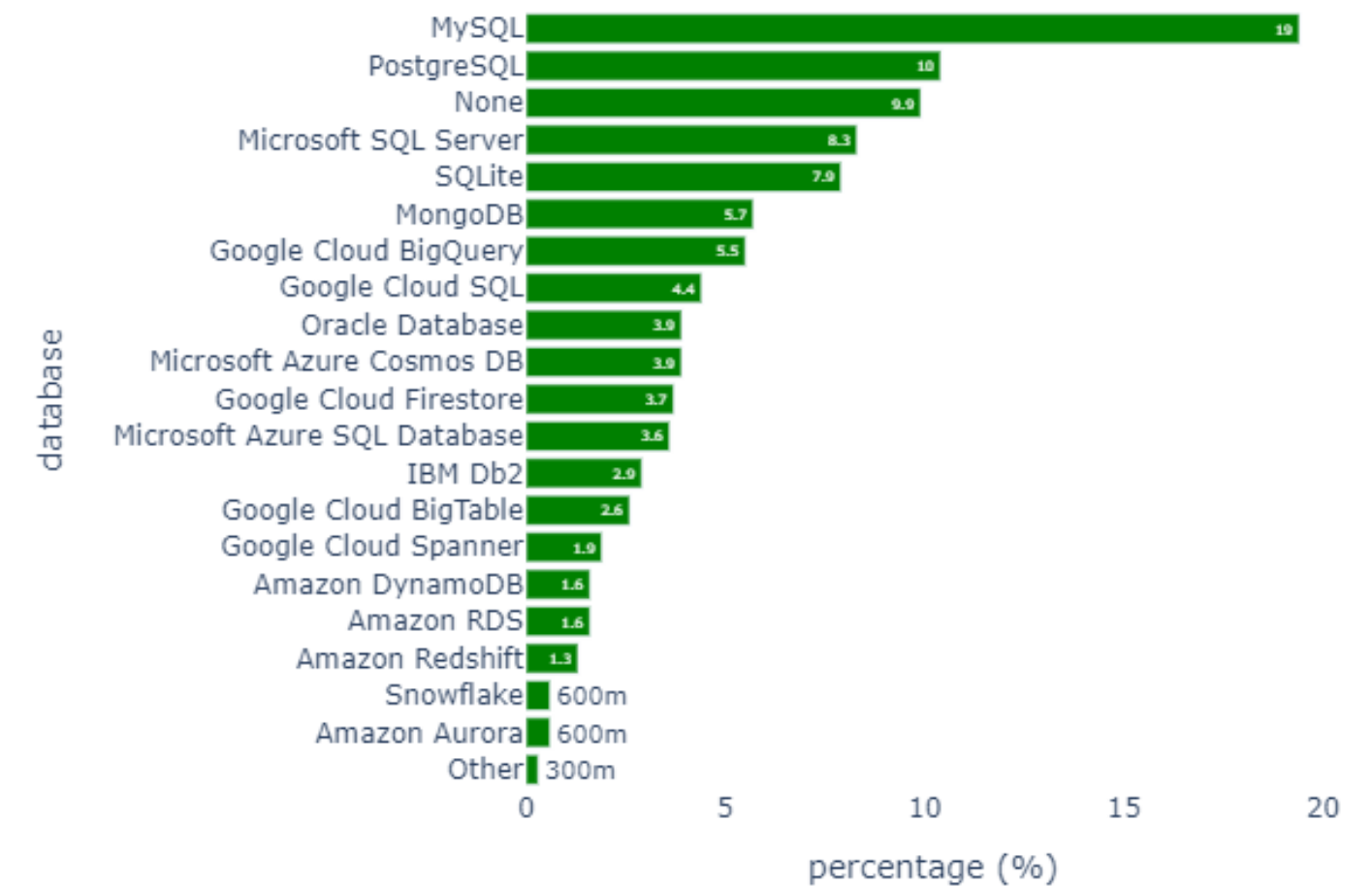
Favourite media source on data science topics

Favourite media source on data science topics for Nigerian Kagglers
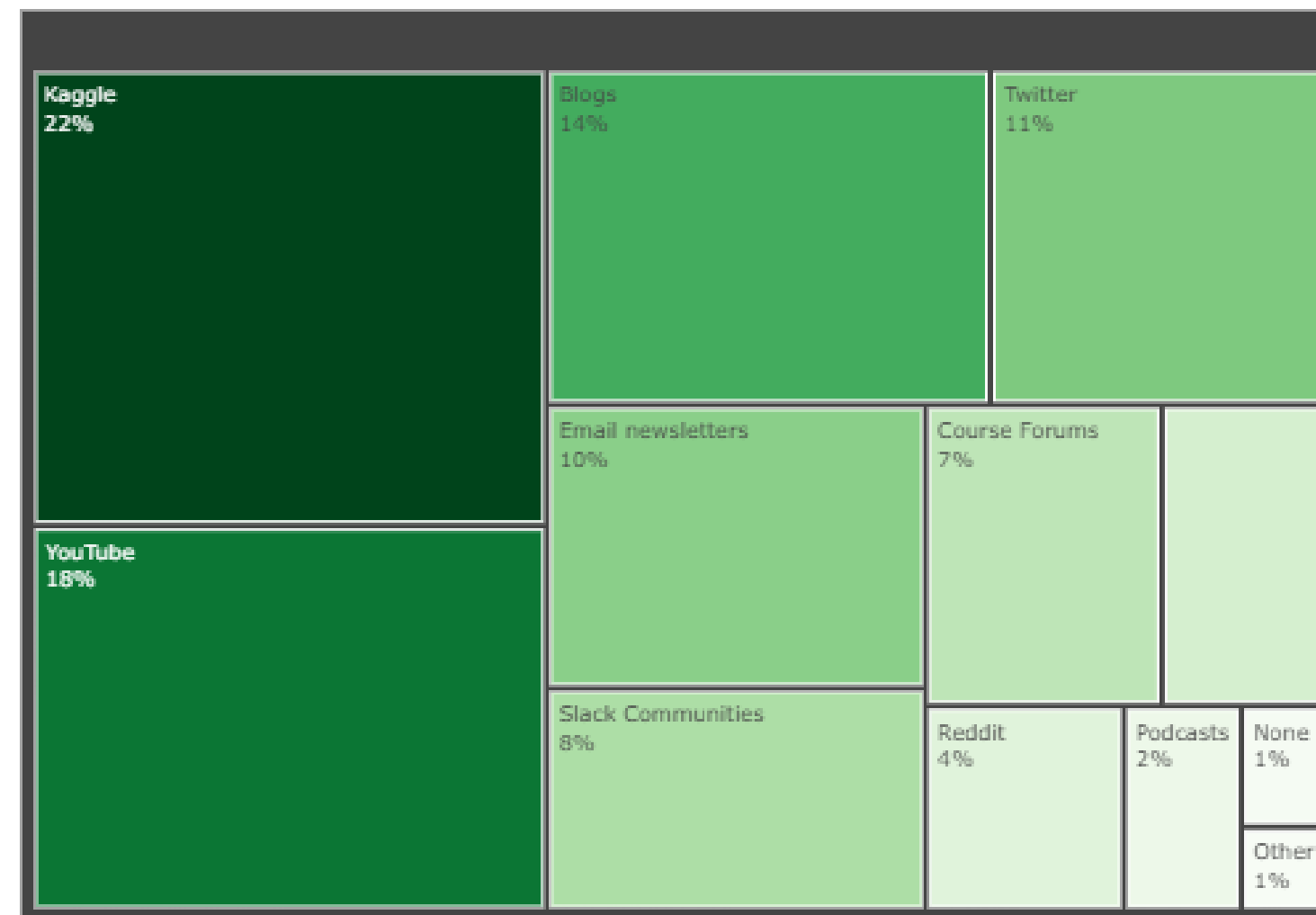
Kaggle 22%	Blogs 14%	Twitter 11%
YouTube 18%	Email newsletters 10%	Course Forums 7%
	Slack Communities 8%	Reddit 4% / Podcasts 2% / None 1% / Other 1%

- The top three favourite media sources on data science topics by Nigerian Kagglers are Kaggle, YouTube and Blogs.

Concluding Remarks

This notebook covers the analysis of a unique and specific subset of the dataset, i.e., Nigerian Kagglers, represented in the 2021 Kaggle Machine Learning & Data Science Survey. Some interesting questions were posed around Nigerian Kagglers, which were answered with some data exploration, narrative texts and visualisations.

A summary of the concluding remarks are shown below:

- In comparison to **64 countries**, Nigeria Kagglers make up about **2.7%** and is **ranked 7th.**

- There has been **steady growth** over the past five years, and the total number of responses for the 2021 Kaggle Survey from **Nigeria was 702**. Also, this number for **2021** is about **10 times greater** than in **2017**. A regression model shows that the number will approach and **exceed 1,000 in two-three years.**

- Within Africa, **Nigeria Kagglers** made up 34% and **ranked number one**. Furthermore, the top three African countries on Kaggle are **Nigeria, Egypt and Kenya**. However, ranking of the African countries **per ten million population** shows that **Nigeria is ranked 5th.**

- The analysis of **gender distribution** shows that **15% of Nigerian Kagglers are women and 85% are men**. This reflects **a significant gender gap** among Nigerian Kagglers.

Concluding Remarks

Apart from students, most Nigerian Kagglers are **Data Analysts, Data Scientists and Software Engineers.**

The **largest proportion (30%)** of Nigerian Kagglers are in the age range of **25 - 29**. Also, **91%** of Nigerian Kagglers are **younger than 40**. This shows that most Nigerian Kagglers are reasonably young.

The most common education level for Nigerian Kagglers is a **Bachelor's degree (52%)**. About **93%** have **at least a Bachelor's degree**.

Most Nigeria Kagglers work in the **Academics/Education, Computers/Technology and Accounting/Finance sectors.**

Almost **half of Nigerian Kagglers (45 per cent) earn less than USD 1,000 per year.** In general, the salary distribution for **Data Analysts, Data Scientists and Software Engineers (Nigerian Kagglers)** are similar and mainly **earn less than USD 3,000 per year.**

The **top programming language** for Nigerian Kagglers is **Python**, which is similar to the rest of the world. Most Nigerian Kagglers also use **Python, SQL and Javascript regularly.**

Concluding Remarks

The three most popular IDEs among Nigerian Kagglers and the rest of the world are **Jupyter Notebook, VScode and PyCharm.**

Scikit-learn, TensorFlow and Keras are the most popular Machine Learning frameworks.

Linear/Logistic Regression, Decision Trees/Random Forests and Gradient Boosting Machines(XGBoost, lightgbm, etc.) are the most popular Machine Learning algorithms used regularly by Nigerian Kagglers.

The most common databases used by Nigerian Kagglers are **MySQL, PostgreSQL, Microsoft SQL Server.**

The top **three** favourite media sources on data science topics by **Nigerian Kagglers are Kaggle, YouTube and Blogs.**

I hope you enjoyed reading this notebook

You can connect with me on:

- Twitter - drsotayo

- Github - drsotayo

- LinkedIn - Linkedin.com/in/adeayosotayo/

Thanks to Kaggle for publishing the dataset for the 2021 machine learning and data science survey.

A copy of the source code for this work is available here.

DR ADEAYO SOTAYO