DataHack4FI Season 3
Report

May 2020

Author

Robert Jones

About DataHack4FI and insight2impact

DataHack4FI is an initiative of insight2impact – a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design.

insight2impact is co-hosted by Cenfri and FinMark Trust and is funded by the Bill & Melinda Gates Foundation in partnership with The MasterCard Foundation.

For more information

Visit our website at www.i2ifacility.org.
Call us on +27 21 913 9510 or +27 11 315 9197.

Contents

1. Introduction and rationale

The availability of data in financial inclusion has grown tremendously in recent years, deepening our understanding of how access to, and the use of, financial services/products affect welfare outcomes throughout Africa. Despite the increasing availability of data, a practical disconnect exists between the decisions being made for financial and economic inclusion and the use of relevant data to inform such decisions. This gap restricts the extent to which financial services reach all segments of the population, particularly the unserved and underserved, limiting the ability of FSPs to design products/services that are inclusive.

> "DataHack4FI is a pan-African innovation competition that encourages finance, technology, and data enthusiasts to collaborate in using data and analytics to solve business challenges and extend financial services to underserved communities." (www.datahack4fi.org)

A key challenge in endorsing a data-driven approach is a distinct lack of data science skills in Africa, as well as a general lack of understanding regarding the value of innovative data analytics. The DataHack4FI initiative promotes the use of data to drive product design and business decision-making, thereby catalysing the development of new products and services driven by customer data. In doing so, the competition aims to contribute to data science skills development and the empowerment of African youths in a field that is increasingly providing opportunities for career advancement throughout Africa.

The initiative seeks to create communities of innovation by convening emerging technology companies (ETCs), aspiring data scientists, innovation hubs, development partners and other key innovation stakeholders, to stimulate the use of data for financial and economic inclusion. DataHack4FI focuses on the use of alternative data sources, big data, small data, and new data trends that are traditionally ignored by incumbents but that emerging tech companies can use to design user-centred products and services.

DataHack4FI targets young innovators and aspiring data scientists from several focus countries across Africa. The competition runs over multiple stages where teams compete to develop innovative solutions to key developmental challenges. The competition serves as a platform for participating teams to showcase their work while vying for the DataHack4FI winner's prize – technical and financial support towards continued solution development. Three seasons of the competition have been completed to date, with Season 3 seeking to build on the success of the previous iterations.

2. DataHack4FI Season 3 overview

2.1. Overview of competition

Building on the networks and experience gained from Seasons 1 and 2, Season 3 of the DataHack4FI pan-African innovation competition took place in seven African countries (Ghana, Kenya, Nigeria, Rwanda, South Africa, Uganda and Zambia) from April to August 2019. In Season 3, emerging technology companies (ETCs) were paired with data enthusiasts to collaborate around the use of data and data-driven decision-making in designing solutions to financial and economic inclusion challenges.

Throughout the course of the competition, participants were provided with both digital and in-person mentorship from experts in financial inclusion, technology, business science and data science. Participants also attended a number of skills development meet-ups and a series of webinars designed to provide information and tools relevant to the design of their solutions.

Data enthusiasts in Season 3 underwent a rigorous selection process. Aspiring data scientists who applied for the competition were afforded the opportunity to enrol in the Microsoft Professional Program (MSPP) in Data Science, an internationally accredited data science certification. Experienced data scientists were on hand to provide digital and in-person mentorship to the candidates as they worked to complete the course requirements. The top-performing candidates from each of the seven DataHack4FI countries were subsequently paired with an ETC and a data science expert for the competition. A total of 374 aspiring data scientists enrolled in the MSPP online training, with 108 candidates successfully awarded the Microsoft certification after completing the prerequisite 10 theoretical modules and a practical use-case project. The full MSPP report is available **here**.

Each of the 45 DataHack4FI teams presented their solution to a panel of expert judges at one of seven independent in-country finals, with the top two teams from each country invited to showcase their prototype solution at the DataHack4FI Grand Finale from 22 to 23 August 2019 – the main event at insight2impact's DataFest, held in Kigali, Rwanda. The DataFest convened more than 300 key stakeholders and interested contributors from across Africa, offering an exciting agenda for all in attendance: keynote speakers, panel and round-table discussions, investor panels, entrepreneurial training, an investor-readiness masterclass, and the Sendy Hackathon. Judged on a range of fundamental criteria, Botlhale AI (South Africa) was chosen as the DataHack4FI Season 3 overall winner, receiving the USD25,000 cash prize and technical assistance towards the continued development of their AI-enabled chatbot solution.

2.2. Objectives

Season 3 of DataHack4FI sought to create opportunities for data enthusiasts and ETCs, deepen the data for the financial services ecosystem, and advance financial and economic inclusion. insight2impact aimed to achieve these objectives by creating a neutral convening platform, providing mentorship and skills development training, and providing technical and financial support to emerging tech companies.

- Skills development
  - To provide data science and entrepreneurial skills development opportunities for participating data enthusiasts and ETCs

- Networking and collaboration
  - To create a convening platform for young entrepreneurs, aspiring data scientists, innovators, and public and private-sector stakeholders to promote pan-African collaboration around the advancement of financial and economic inclusion
  - To build new, and strengthen existing, networks that allow for the empowerment of youth and women to design scalable solutions

- Data-driven decision-making
  - To promote the use of data to improve financial and economic inclusion through evidence-based, client-centric solution design
  - To mobilise the data market by leveraging interest and skills of stakeholders to develop active and impactful innovation ecosystems that promote data-driven decision-making

- Financial and economic inclusion (through a combination of the previous objectives)
  - To catalyse new data-driven solutions that have a positive impact on society by serving the financially and economically excluded and underserved

2.3. Descriptive statistics

2.3.1. General overview

A total of 228 individuals participated in DataHack4FI Season 3, representing 45 teams across seven in-country competitions. Table 1 summarises the number of teams, ETC representatives, data enthusiasts and data experts at a country level.

Country	No. of teams	No. of ETC representatives	No. of paired DEs	No. of data experts	Total
Ghana	6	17	6	4	27
Nigeria	7	28	7	7	42
RSA	9	25	9	7	41
Zambia	7	18	7	6	31
Uganda	7	27	7	4	38
Rwanda	3	7	3	2	12
Kenya	6	25	6	6	37
Total	**45**	**147**	**45**	**36**	**228***

Table 1: Season 3 participant overview

*This figure does not include 329 non-paired data enthusiasts who enrolled in the MSPP *(Inclusive total= 557 participants)*.

Males accounted for the majority of participants in Season 3 (82%, 187 males), with 41 females (representing 18%) actively involved in the competition. The average age of participants was 29 (±7) years old, with 81% of individuals falling into the "youth" category (18 to 35 years old[1]).

---

1   www.un.org/en/africa/osaa/pdf/au/african_youth_charter_2006.pdf

2.3.2. Emerging technology companies (ETCs)

Forty-five (45) ETCs competed in DataHack4FI Season 3, consisting of 147 representatives with an average age of 31 (±8) years old. Most companies had been in operation for one year or less (37% of ETCs), with a single ETC reporting a company maturity of more than four years (Figure 1). The average maturity of Season 3's ETCs was two years. The average number of staff members per company was six, including both full-time and part-time employees.
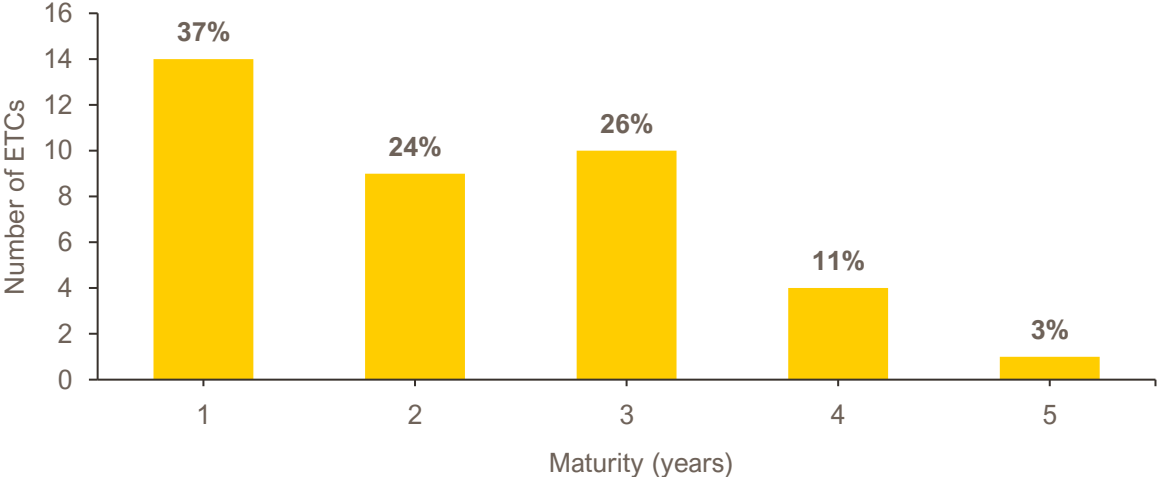

Figure 1: ETC maturity

Source: DataHack4FI MEL Survey – 38 ETC respondents

Many ETCs had participated in innovation competitions prior to DataHack4FI Season 3 (43%), 39% had completed an incubation programme, 32% had accelerator experience, and 29% had completed a bootcamp (Figure 2). Seven percent (7%) of ETCs had participated in previous editions of the DataHack4FI innovation competition. Four respondents, representing 14% of ETCs, did not have any experience of innovation-related events or programmes.
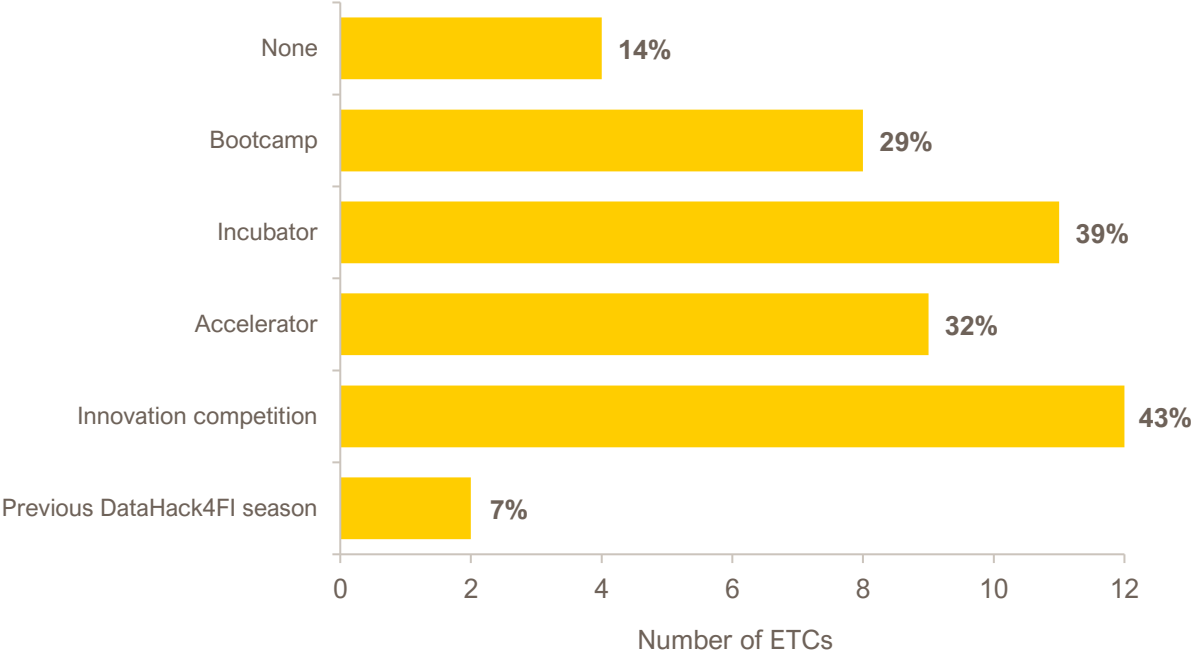

Figure 2: ETC innovation programme history

Source: DataHack4FI MEL Survey – 46 ETC respondents

Twenty percent (20%) of the solutions presented at the in-country final level were designed specifically for DataHack4FI Season 3 (six respondents), while 82% of solutions represented enhancements of existing product or service offerings (24 respondents).

2.3.3. Data enthusiasts

Forty-five (45) data enthusiasts were paired with ETCs during DataHack4FI Season 3 (representing a one-to-one ratio), consisting of 41 males (91%) and 4 females (9%). The average age of the data enthusiasts was 26 (±3) years old.

More than 80% of data enthusiasts had obtained a degree from a tertiary institution (Figure 3), with the most-held qualification being a bachelor's degree (50%). The data enthusiasts were attracted from a range of educational backgrounds, including computer science (37%), engineering (24%), mathematics (21%) and statistics (13%).
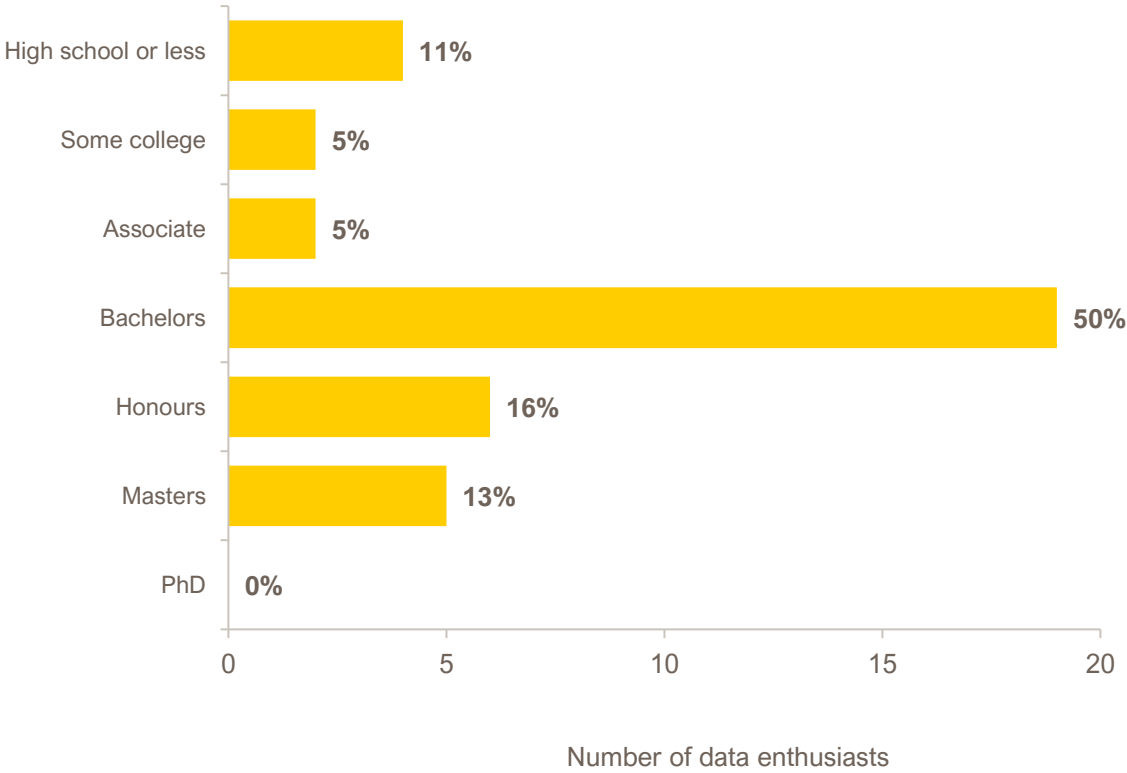

Figure 3: Data enthusiasts' highest level of education obtained

Source: DataHack4FI MEL Survey – 38 data enthusiast respondents

The majority of data enthusiasts had less than two years of data-related work experience prior to their involvement in DataHack4FI (87%, 32 respondents), 11% (4 respondents) had between two and five years of experience, and a single data enthusiast had more than five years of experience.

3. Outcomes and learnings

A mixed-evaluation approach was employed to explore key outcomes and assess the extent to which DataHack4FI Season 3 achieved its objectives. Both quantitative and qualitative data were collected from participants through targeted surveys at the in-country final stage, as well as in-person interviews at various stages of the competition.

3.1. Emerging tech companies (ETCs)

The perceived benefits for participating ETCs are displayed in Figure 4. The key areas of benefit were highlighted as "A focus on evidence-based and data-driven decisions", "An opportunity to receive skills development and mentorship", "Visibility and exposure", and "Partnerships and networking".
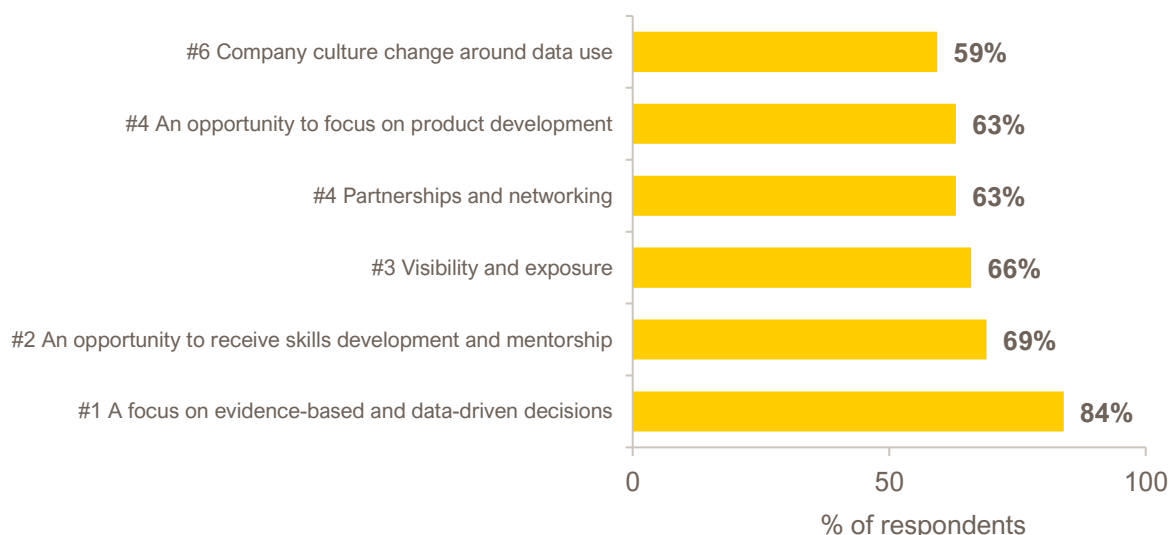

Figure 4: Areas of benefit for DataHack4FI ETCs

Source: DataHack4FI MEL Survey – 32 ETC respondents

3.1.1. DataHack4FI influenced ETCs' approach to data

DataHack4FI's focus on the use of data in decision-making was identified as the main benefit for participating ETCs (Figure 4). Eighty-seven percent (87%, 28 respondents) of ETCs employed innovative analytics methods in the design of their solutions. They utilised previously unconsidered sources of data, collected new data, and applied novel analysis methods to draw more meaningful and impactful insights.

A deeper appreciation of the benefits accompanying a data-driven approach to product design and business practice has prompted ETCs to improve their internal capacity for data (Figure 5). The most commonly cited internal capacity changes include developing or improving an existing data strategy (81%, 30 respondents), hiring a consultant or data solution provider (54%, 20 respondents), as well as revamping data collection and storage processes (51%, 19 respondents).
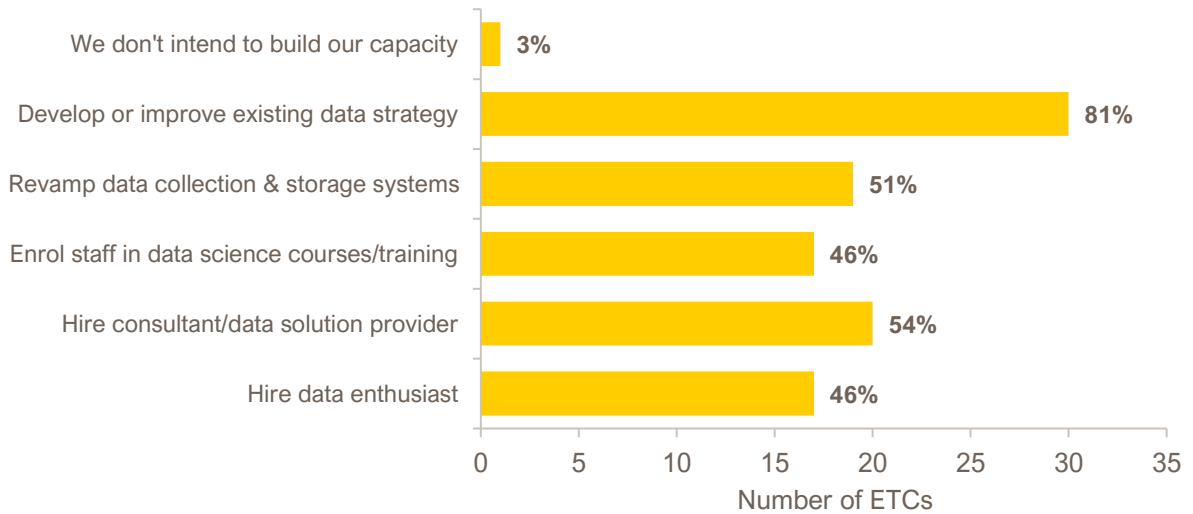
Figure 5: Investments in improving internal data capacity resulting from DataHack4FI

Source: DataHack4FI MEL Survey – 37 ETC respondents

3.1.2. ETCs benefited from the skills development opportunities

The skills development and mentorship provided throughout the competition was highlighted as the second-most beneficial element of DataHack4FI Season 3 (Figure 4). Overall, the extent to which DataHack4FI Season 3 was able to provide relevant skills development was rated 4.13 out of 5[2].
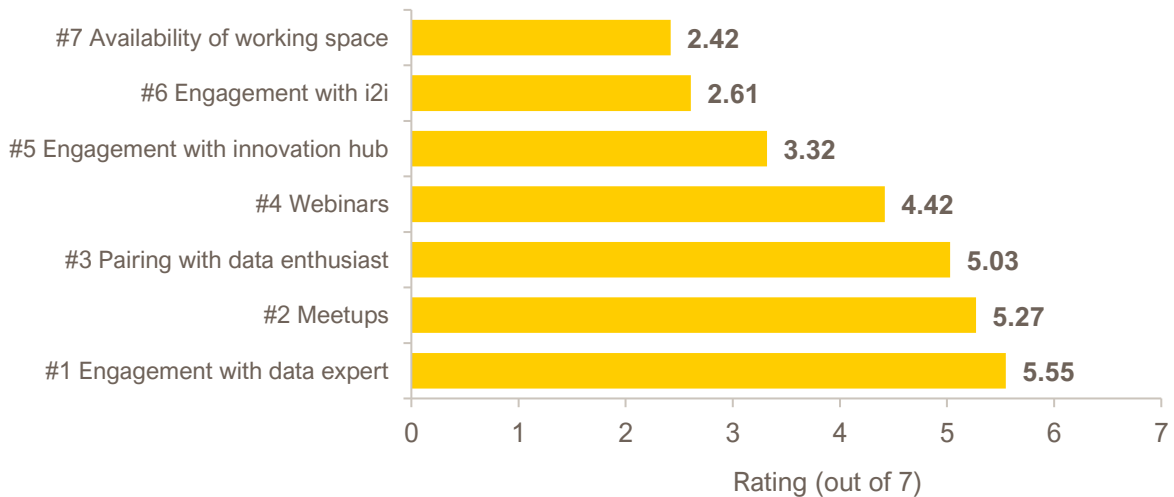

Figure 6: Rating of the most useful elements that assisted ETCs to develop their solutions

Source: DataHack4FI MEL Survey – 36 ETC respondents

---

2   Source: DataHack4FI MEL survey – 32 ETC respondents

The in-person meetups (#2) and webinars (#4) were of particular benefit (Figure 6), providing opportunities for the participants to access content and tools relevant to the development of their solutions, while receiving guidance from experts in various thematic areas, including financial inclusion, data science and entrepreneurial/business. Figure 7 illustrates the perceived impact of the skills development elements on ETCs.
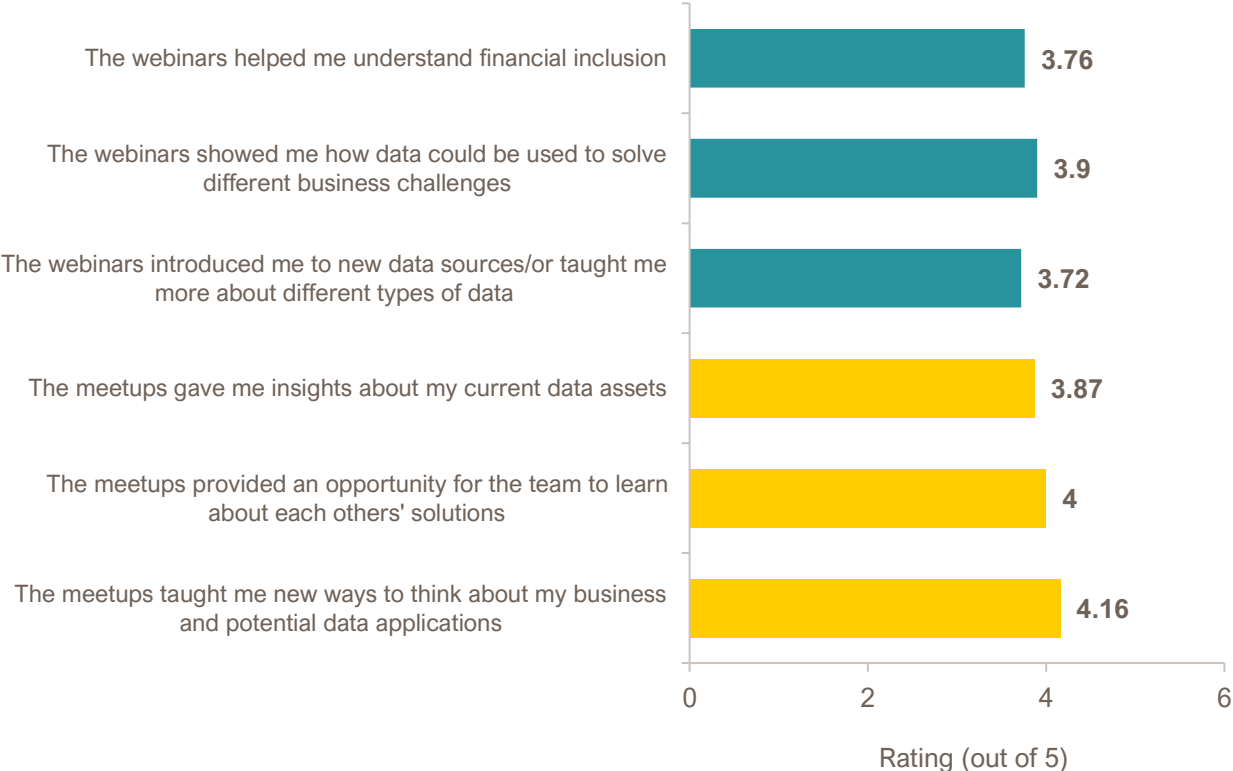

Figure 7: ETC's rating of meetup and webinar impact

Source: DataHack4FI MEL Survey – 32 ETC respondents

3.1.3. ETCs valued the collaboration and networking opportunities

The DataHack4FI initiative promotes networking and collaboration among key stakeholders within the innovation ecosystem, rated as the fourth-most valuable aspect of the competition by ETCs (Figure 4).

As illustrated in Figure 6, the opportunity to collaborate with data science experts (#1) and data enthusiasts (#3) was considered particularly valuable in assisting each team to develop their DataHack4FI solution. The influence of the data science expert and data enthusiast is acknowledged as the key proponent in advancing the ETC's understanding of the implications of a data-driven approach to business decision-making and product design (Figure 8).
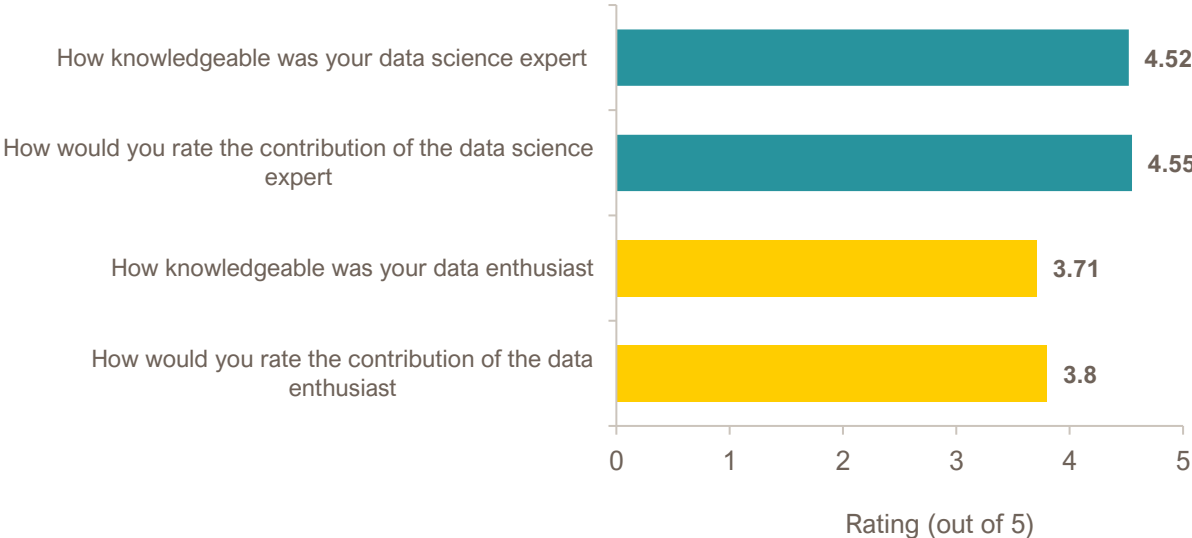

Figure 8: Contribution of data science experts and data enthusiasts

Source: DataHack4FI MEL Survey – 31 ETC respondents

These connections are expected to continue to assist the ETCs as they progress their businesses and develop solutions in the future. ETCs expressed their interest in maintaining the relationship with their data enthusiast, either by hiring them formally (48%, 13 respondents) or in an informal capacity (42%, 15 respondents). ETCs were also eager to maintain their engagement with the data science expert in a formal manner (65%, 20 respondents), while 29% were keen to maintain an informal working relationship (9 respondents).

3.2. Data enthusiasts

Note: This section includes insights based on data collected from the paired data enthusiasts, as well as survey responses from data enthusiasts who enrolled in the MSPP.

3.1.4. Data enthusiasts advanced their data science knowledge during DataHack4FI

The data enthusiasts benefited from the skills development opportunities offered during DataHack4FI, rating this element of the competition (on average) 4.42 out of 5[3]. Data enthusiasts were able to advance both their theoretical data science knowledge and practical capabilities by engaging with their paired ETC and data science expert, attending the meetups and webinars, and enrolling in the Microsoft Professional Program in Data Science (MSPP).

Data enthusiasts experienced positive working relationships with their team, providing unique skills that are otherwise lacking within their paired ETCs (Figure 9). This relationship allowed the data enthusiasts to apply their knowledge to a practical use case, thereby solving real-world problems through data analytics. The data science experts played a pivotal role in developing the data enthusiasts' data science capabilities, with both in-person and digital mentorship assisting them as they designed data-driven solutions alongside their ETC (Figure 9). The benefit of this association was such that 90% of data enthusiasts (28 respondents) expressed intent to maintain contact with their data science expert[4].

Figure 9: Data enthusiasts' rating of team dynamics

Source: DataHack4FI MEL Survey – 31 data enthusiast respondents

---

3   Source: DataHack4FI MEL survey – 31 data enthusiast respondents

4   Source: DataHack4FI MEL survey – 31 data enthusiast respondents

The in-person meetups and webinars were highlighted to be of value in imparting data science knowledge (Figure 10), providing access to data-related content and tools, while receiving guidance from data science experts.
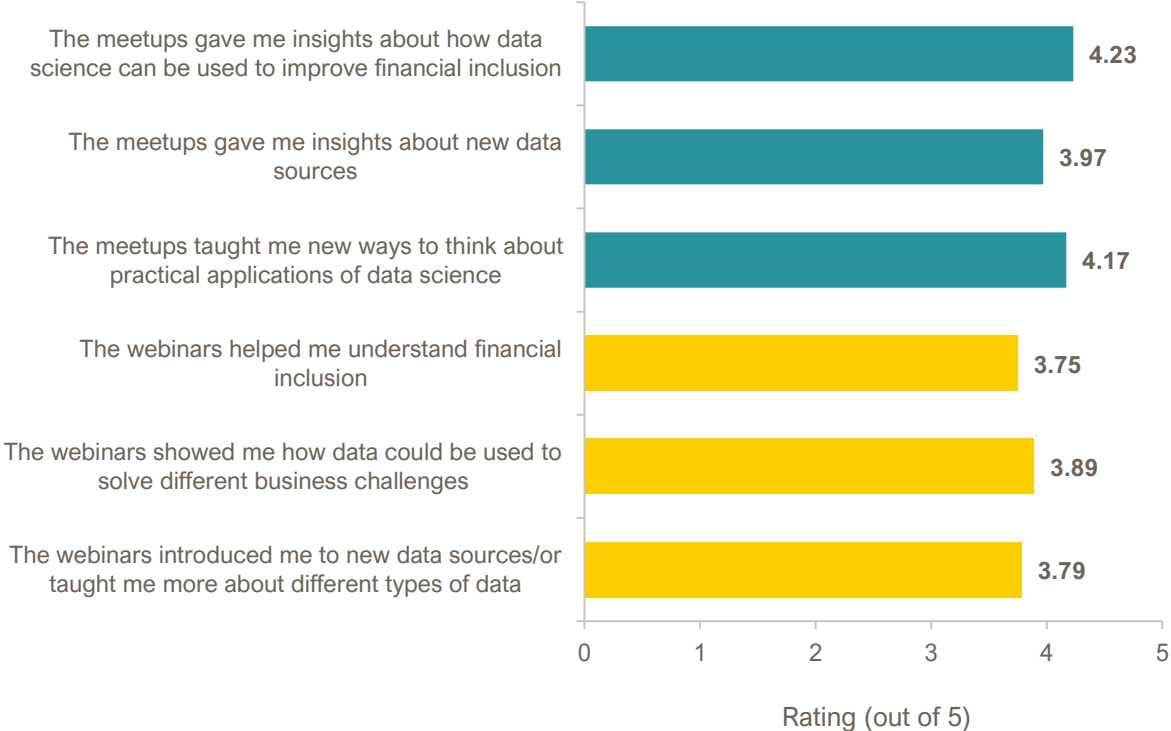

Figure 10: Data enthusiasts' rating of meetup and webinar impact

Source: *DataHack4FI MEL Survey – 30 data enthusiast respondents*

The paired and non-paired data enthusiasts benefited considerably from their enrolment in the MSPP in Data Science, rating the course usefulness as 4.9[5] and 4.4[6] (out of 5) respectively. Prior to enrolling, the majority of data enthusiasts considered their data science knowledge to be at a "beginner" (32%) or "intermediate" (50%) level. The MSPP resulted in a substantial increase in data science understanding among the cohort, with 54% (40 candidates) considering themselves "experienced" data scientists following the programme (Figure 11 11).
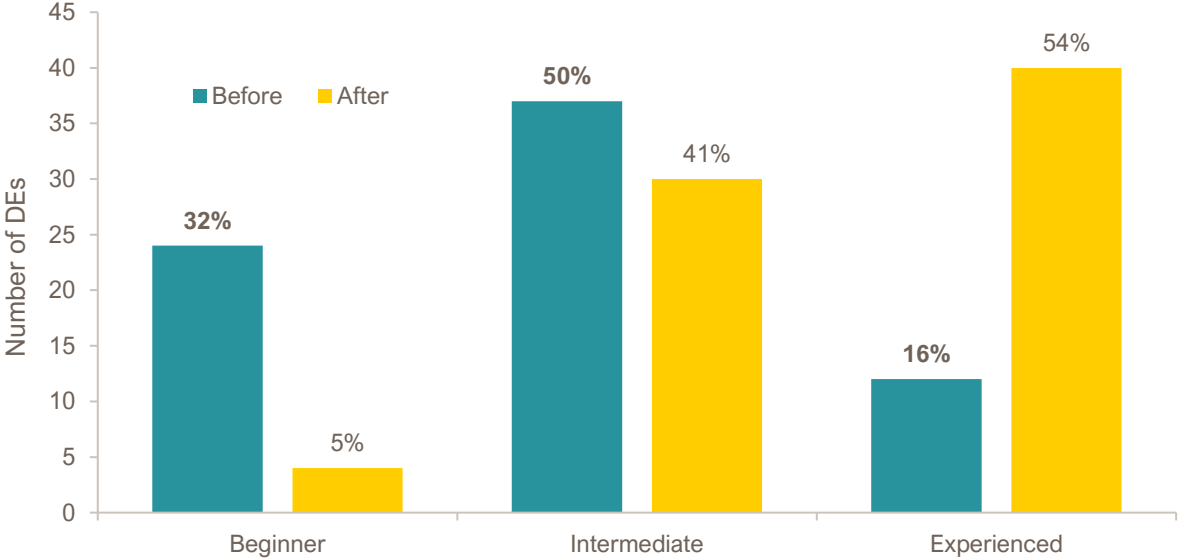

Figure 11: Data science knowledge before and after the MSPP

Source: *DataHack4FI MSPP Survey – 76 respondents*

In addition to increasing the theoretical and practical data science capabilities of aspiring data scientists, the skills development components of DataHack4FI motivated data enthusiasts to pursue further data science education following the completion of the competition; where 100% of respondents indicated their intent to engage in further data science training[7].

---

5   Source: DataHack4FI MEL survey – 30 data enthusiast respondents

6   Source: DataHack4FI MSPP survey – 76 respondents

7   Source: DataHack4FI MSPP survey – 72 respondents

3.1.5. DataHack4FI positively affected data scientists' career prospects

The DataHack4FI initiative directly affected the career prospects of participating data enthusiasts (Figure 12). The competition, particularly the MSPP training, resulted in a marked increase in the data scientists' perception of their own employability, with 85% of data enthusiasts (63 respondents) expressing an improvement in employability[8].
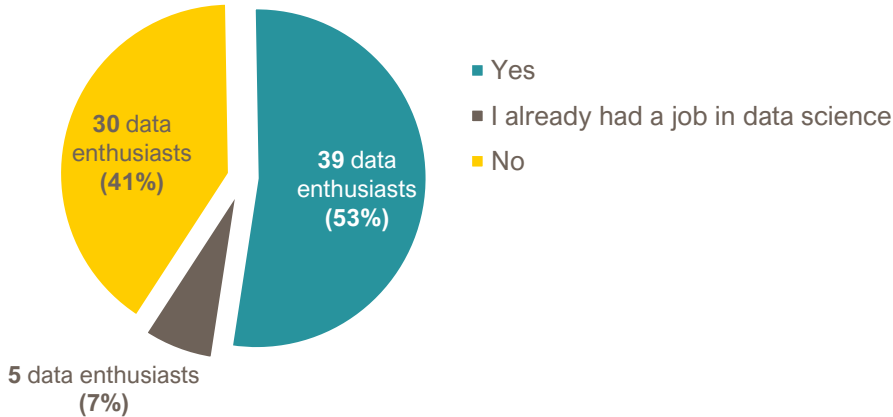

Figure 12: Distribution of data enthusiasts entering data science-related jobs post-DataHack4FI

Source: DataHack4FI MSPP Survey – 74 respondents

Thirty-nine (39) candidates (53%) who enrolled in the MSPP training have taken up either part-time or full-time data-science-related employment since the end of DataHack4FI Season 3. The distribution of candidates who have obtained data-related jobs is shown in Figure 13; 17 have secured full-time employment, 15 have secured part-time employment, and six have entered unpaid internships or tutoring.
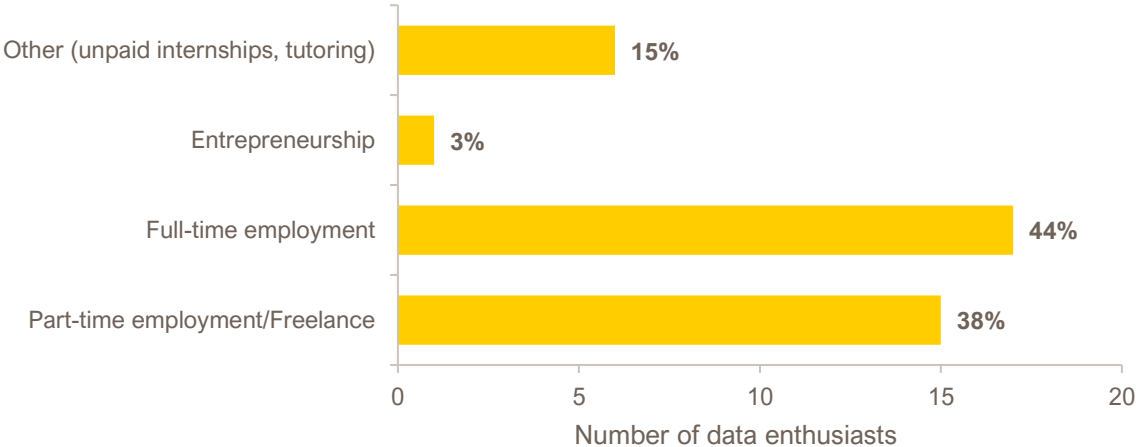

Figure 13: Distribution of type of work engaged in post-DataHack4FI

Source: DataHack4FI MSPP Survey – 39 respondents

---

8   Source: DataHack4FI MSPP Survey – 74 respondents

4. Conclusion

The third Season of the DataHack4FI pan-African innovation competition involved 228 participants, representing 45 teams from across seven African countries. The inclusion of a further 329 aspiring data scientists who enrolled in the MSPP in Data Science but not formally paired with a team for the competition, results in a grand total of 557 participants. The initiative catalysed the development of 45 innovative data-driven solutions, either enhancements of existing products/services or entirely new solutions, with the potential to solve key developmental challenges faced by the financially and economically excluded or underserved in Africa. Furthermore, the initiative resulted in the certification of 108 African data scientists through the MSPP data science training.

DataHack4FI Season 3 sought to create opportunities for aspiring data scientists and emerging tech companies to collaborate in the design of data-driven innovations that advance financial and economic inclusion, while accessing relevant support from experts, mentors and other key innovation ecosystem stakeholders. The initiative set out to achieve four core objectives: (1) skills development, (2) networking and collaboration, (3) data-driven decision-making, and (4) financial and economic inclusion.

4.1.   Skills development

Season 3 was successful in providing market-relevant skills development opportunities, targeted at both ETCs and aspiring data scientists. The competition delivered in-person meetups, webinars, entrepreneurial training, pitch training, and an investor masterclass, as well as digital and in-person mentorship from experts in financial inclusion, technology, data science and business science.

DataHack4FI's aspiring data scientists developed their theoretical and practical capabilities by engaging with data science experts, by enrolling in the MSPP in Data Science, and by applying their knowledge to real-world use cases as they worked to design solutions alongside their team. The initiative was able to motivate data enthusiasts to pursue further data science education, while increasing their employability, with several participants entering data-related jobs after the competition. In this way, DataHack4FI contributed to the empowerment of African youths in a field that is increasingly providing opportunities for career advancement.

4.2.   Networking and collaboration

DataHack4FI was able to convene key innovation ecosystem stakeholders across the seven target countries, with a common objective of solving key financial and economic inclusion challenges through collaboration. The competition assembled emerging tech companies, aspiring data scientists, innovation hubs, experts and mentors with a range of relevant expertise, financial service providers, investors and development partners.

Teams competing in Season 3 (consisting of an ETC, their paired data enthusiast and a data science expert) experienced positive working relationships that fostered collaboration, learning and the design of data-driven solutions that solve key developmental challenges. Ultimately, insight2impact was able to broker partnerships among key stakeholders that serve to realise these solutions, while facilitating the continued development of enabling innovation

ecosystems. Multi-stakeholder networking and collaboration presents a realistic avenue to reach the financially and economically underserved.

4.3. Data-driven decision-making

DataHack4FI positively affected the extent to which emerging tech companies understand and appreciate the benefits of adopting a data-driven approach to product design and decision-making. By collaborating with data enthusiasts and data science experts, the way data was utilised by the team evolved during the competition. ETCs employed previously unconsidered sources of data, collected new data, and applied novel analytical methods to draw more impactful inferences. This approach has several benefits for both the product/service provider and the end-user, allowing ETCs to:

- better understand their customers, enabling the design of solutions that meet market needs
- improve the extent to which their models can accurately predict an output for a given use case
- gain additional insight that influences the strategic direction of their business, or that reaffirms their current strategy, and
- identify new market segments to target new innovations

Their involvement in the competition also prompted ETCs to enhance their capacity for data collection and analysis by developing or improving an existing data strategy, hiring data science human capital, or revamping internal data collection and storage processes.

DataHack4FI Season 3 was effective in conveying the importance of adopting a data-driven and evidence-based approach, thereby bridging the gap between the decisions being made for financial and economic inclusion and the available data, catalysing the development of customer-centric products/services.

4.4. Financial and economic inclusion

Season 3 of DataHack4FI successfully stimulated the use of data by participating ETCs, ensuring the solutions developed during the competition are underpinned by data and designed to meet the needs of their target population. Products/services with a foundation in data have a greater potential to reach scale due to a deeper understanding of the market need and the unique context in which the financial or economic inclusion challenge exists. During the competition, teams were able to clearly identify and articulate who their end customer is, understand the nature of their particular use case, and design innovative solutions that address these unique challenges. This approach further enhances the likelihood of these solutions reaching market and increases the prospects for emerging tech companies to develop into sustainable businesses.

Ultimately, DataHack4FI Season 3 catalysed the development of 45 data-driven solutions, designed to solve key developmental challenges faced by the financially and economically underserved. Such products have the potential to affect the quality of life for millions of Africans who make up the base of the financial market segment pyramid.

How to find us

Get involved. Contact us.

Robert Jones
robert@i2ifacility.org

+27 11 315 9197
i2ifacility.org