

Technical Brief

DATA SKILLS REPORT

UNDERSTANDING AFRICA'S YOUNG DATA PROFESSIONALS

October 2021



data
INNOVATORS

CONTENTS

Summary	2
Introduction	3
The data skills survey	4
Insights	5
Supporting educational attainment of youth will increase the pool of future data skills	5
Education attainment	5
Build data skills from the bottom-up is needed for business and employment growth	6
Data skills levels	6
Linking education attainment and data skills	7
Data tool expertise	7
Employment of skilled data professionals	8
Employment challenges	9
Data profession aspirations	10
Actions for organisations driving data skills in Africa	11
Partner for positive change in data skills	11
Assess organisational skills gaps, design, and implement in-house data re-skilling programmes	11
Target systemic change now that will influence youth access to data skills jobs in the future	11
Methodology	12
References	13

CONTRIBUTORS

Rama Raphalalani, Teboho Letuka, Dineo Maboke, Mary-Grace Villanueva, Jerusha Govender

SUMMARY

Data and AI competencies will account for just under 20% of the new jobs to be created through the 4IR (WEF, 2020). This is increasing the demand for data skills across sectors. However, the context of data skills remains unclear as to whether young professionals are adequately skilled. The following technical brief looks at the existing gaps and need to develop data skills, particularly for African youth living through the Fourth Industrial Revolution (4IR).

In 2020, Data Innovators conducted a study to determine the current context of data skills amongst young professionals. Key questions of the study were: what is the level of qualification and type of professionals in or entering data related fields; what are the perceived level of data skills and experience of common data tools amongst data professionals, and what are barriers and aspirations for employment in data roles? The study was conducted using an online survey disseminated via social media channels. The sample of respondents included 119 individuals from 22 African countries (113 respondents from African countries), both male (72) and female (47). The age groups included 18-24 years (21), 25-34 years (73) and 35+ years (14) (1 unknown age). Using descriptive analysis, the study presents contextual findings on the state of data skills amongst young African professionals.

The key **findings** are:

- Tertiary education is the main point of data skilling amongst young professionals. More than 70% of the young professionals surveyed have degrees which include some level of data training and approximately 34% of respondents have or were undertaking courses related to data associated fields.
- Young professionals have higher level of skill in less technical data skills areas. This is evident in the decreasing percentage of those with high competence and expert level skills as we move across the competencies from data capturing (less

technical) to data visualisation (more technical). Only 5 individuals self-identified as experts in all six competencies - all possess a postgraduate degree in public health, M&E, and engineering.

- MS Excel is a tool that young professionals are required to use, and many respondents confirm is a tool they are knowledgeable in. However, 27 (23%) of respondents had no knowledge on all other data analysis software listed (Stata, R, and SPSS). Forty-five (38%) respondents had no knowledge of any of the visualisation software listed (Google Studio, Power BI, and Tableau).
- Employment in data roles remains challenging for young professionals. 42% of those aged 25-34 years were unemployed at the time of the survey, 28 of the respondents in the age group have completed tertiary education. Participants indicate access to opportunities (38%), adequate skills (31%) and experience (12%) are key barriers.

Based on the findings the **actions** proposed to increase data skills and access to data jobs are:

- Companies and organisations should partner on initiatives that build data skills, such as skills development programmes or creating internships that foster data skills.
- Organisations (private and public) must assess organisational data skills gaps, design, and implement in-house data re-skilling programmes. Professionals need to access in-depth data skills that are taught outside of tertiary education. Targeted re-skilling will allow professionals to learn new tools or push current skills on existing tools.
- Target systemic change now that will influence youth access to data skills jobs in the future. Initiatives driven by government and large companies can support increasing the number and type of data skills opportunities and imbed data skills within school and TVET education.

INTRODUCTION

As the world adapts to the developments brought on by the fourth industrial revolution (4IR), the way in which we transact, socialise, learn, and work are changing. The continuous development of technology requires that we adapt to new work roles and behaviours, and hence develop new competencies. Businesses increasingly shift their daily operations, such as, communications; data and information processing; in addition to reasoning and decision making (WEF, 2018). It is predicted that the world will have to reskill over a billion people by 2030 to meet the demand brought on by the newly created jobs, with a forecast of a change in 42% of core skills by 2021 ([WEF, 2020](#)).

As part of the reskilling demand, data and AI competencies will account for just under 20% of the new jobs to be created through the 4IR (WEF, 2020). Part of this shift translates to the increase in the value organisations place on data. The demand for data skills is growing across the globe (The Royal Society, 2019; NTUC LearningHub, 2020; Nanayakkara, 2019).



There has also been an increase in the need for data skills across the board...

- Royal Society (2019)

Opportunities in the data space are vast, hence it is important to know which data skills are in demand and which will continue to be for the foreseeable future. Those who reskill and harness new and relevant data skills will be in line for the emergent data opportunities.

In 2020, Data Innovator conducted a survey to understand the status of data skills, particularly among young professionals in Africa, where unemployment rates are growing. This technical brief provides insights from the survey into the respondents' educational background, the data skills they possess, and how these two interact with their employment status.

The brief provides data and solutions aimed at supporting organisations in rethinking how to up-skill and re-skilling future and current workforces in fields of growing data demands.

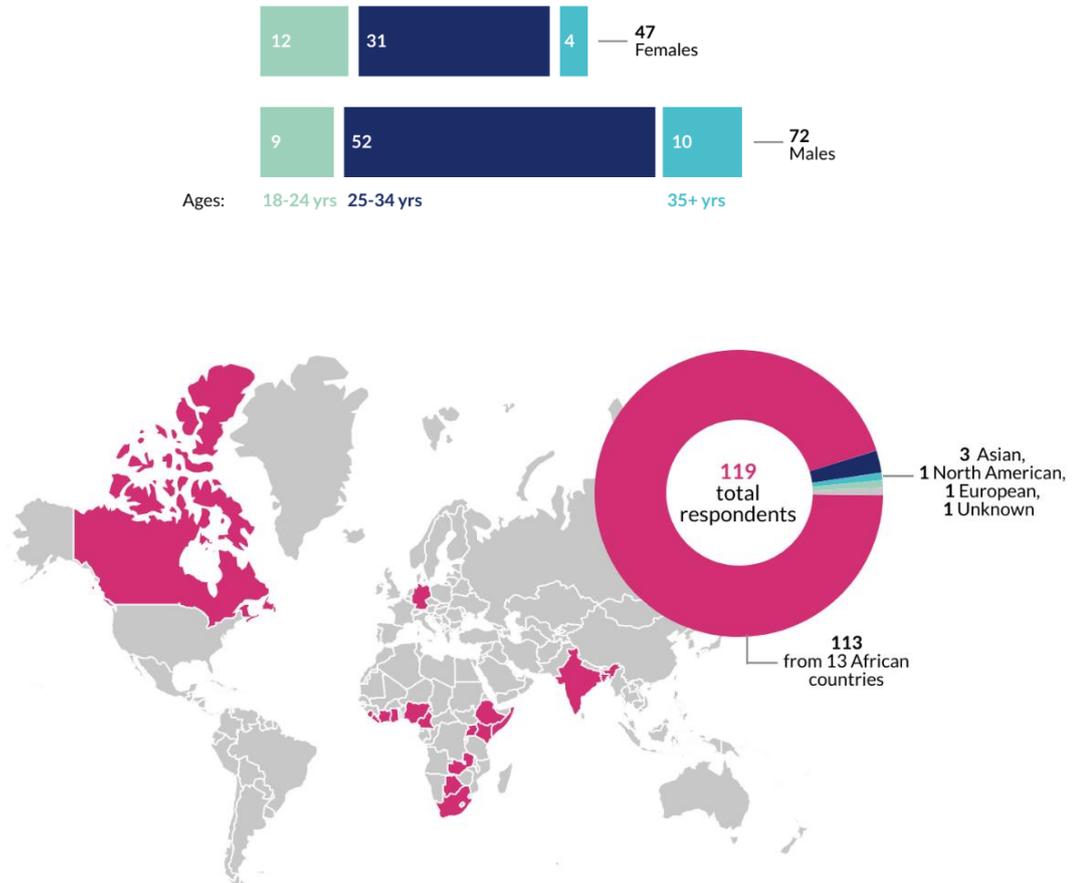


Also view the data skills dashboard for more data skills data, [click here](#).

THE DATA SKILLS SURVEY

The data skills survey is an online survey, previously conducted by Data Innovators in 2020. The survey was disseminated on social media platforms for 2 months. The profile of the respondents provides some contextual data as you review the report. This is important to note as you interpret the insights. Further details on methods on page 9.

Figure 1: Survey respondents summary



INSIGHTS

SUPPORTING EDUCATIONAL ATTAINMENT OF YOUTH WILL INCREASE THE POOL OF FUTURE DATA SKILLS

Africa has historically achieved poor learning outcomes (World Bank, 2017, Bowmaker-Falconer and Herrington, 2020, van der Berg, Gustafsson and Malindi, 2020). Education systems in Africa have multiple pitfalls which will have effects on existing and future data skills, this includes:

- lack of basic cognitive skills – such as numeracy and literacy among school leavers
- low throughput of learners who start primary school to those who graduate with a tertiary degree (less than two-thirds of children complete primary school, with less than half completing secondary school)
- less than 10% of Africans have post high school qualifications

Many jobs in Africa may be displaced by digitisation, with a potential that the net gain in jobs will be positive (Magwentshu, Rajagopaul, Chui, and Singh, 2019). The caveat, however, is that most of these newly created jobs will require at least a post high school qualification.

EDUCATION ATTAINMENT

Results from the data skills survey indicate that most of young professionals that participated in the survey hold a graduate qualification at bachelor’s and a master’s degree levels. No conclusions on gender difference can be made given the gender ratio (39% female; 61% male), yet fewer women than men have indicated having a qualification. The educational outcomes of the survey respondents also paint a more positive picture among young professionals, where 71% have a post high school or more advanced qualification (table 1).

More than 70% of the young professionals surveyed have degrees which include some level of data training (figure 2). Most respondents with a tertiary qualification were from the humanities faculty. Approximately **34% of respondents have or were undertaking courses related to data associated fields** such as

demography and population science, actuarial science, information technology, biostatistics, statistics, as well as monitoring and evaluation (M&E).

Figure 2: Fields of study

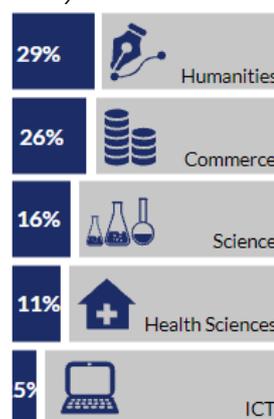


Table 1: Educational attainment by qualification level

	COMPLETE		IN-PROCESS		INCOMPLETE	
Total	93	79%	21	18%	4	3%
High school certificate	10	8%	4	3%	0	0%
Post high school certificate	19	16%	6	5%	2	2%
Bachelor's degree	27	23%	1	1%	0	0%
Honours degree	13	11%	1	1%	0	0%
Master's degree	22	19%	6	5%	1	1%
PhD	2	2%	3	3%	1	1%

BUILD DATA SKILLS FROM THE BOTTOM-UP IS NEEDED FOR BUSINESS AND EMPLOYMENT GROWTH

Employers across the continent point to inadequately skilled workers as one of the major constraints to their business' growth and sustainability (world economic forum, 2017). The skills gap will be exacerbated due to the increasing proportion of occupations that are information and communication technology intensive – which require the appropriate competencies.

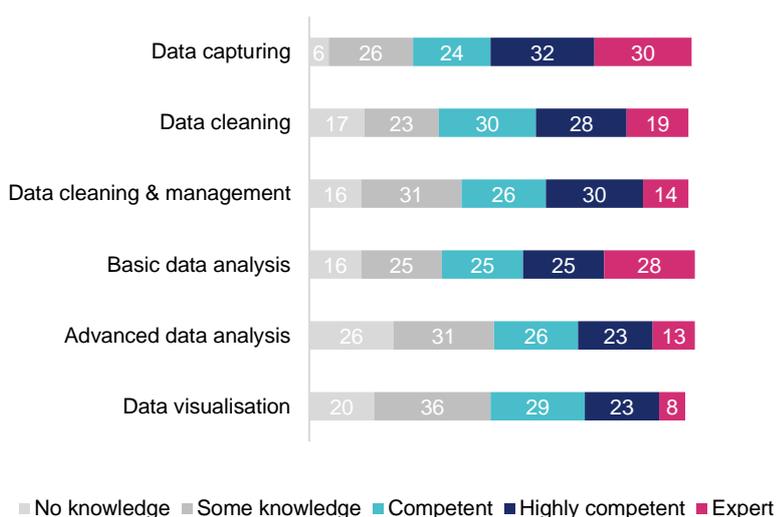
Data skills are value-add capabilities for employment and employees. Conscious efforts to upskill African youth in data skills could play a key role in decreasing the current skills gap and curb widening of the gap. These efforts should consider both downstream and upstream data competency upskilling ([Sarah nell-rodriguez, 2020](#)). Downstream upskilling is focused on creating data skills for everyone –

particularly those whose lives have been adversely affected due to a lack of data skills. Upstream upskilling is aimed at how data skills can strategically add value to organisations and communities.

There are numerous lists which outline the specific data skills deemed as important by various people and organisations. Hence we compiled a list of six data skills that are relevant for entry to mid-level employment using critical digital and data skills lists created by [datapine](#), [Udacity](#), and [ntuc learninghub](#) (2020), and with the considerations stated above. These skills are:

- data capturing
- data cleaning
- basic data analysis
- data management
- advanced data analysis, and
- data visualisation

Figure 3: Reported data skills levels



Only 5 males (from Ethiopia, Kenya, & South Africa) self-identified as **experts in all six competencies** - all possess a postgraduate degree in public health, M&E, and engineering.

DATA SKILLS LEVELS

The data skills data presents those young professionals have higher level of skill in less technical data skills (figure 3). This is evident in the decreasing percentage of those with high competence and expert level skills as we move across the competencies from data capturing (less technical) to data visualisation (more technical). There is also an increasing proportion of those who report to having no knowledge as we move across the competencies.

The number of respondents with expert level basic analysis skills is significantly higher than advanced data analysis skills. Basic and advanced data analysis lies in that basic data analysis is confined to exploratory and descriptive analysis

whereas advanced analysis makes use of inferential and predictive statistical methods.

Using the skills level scale, we gave each respondent a composite score which is the average across the six data skills. The average self-perceived score of the cohort is two – which means that they perceive themselves as **competent** across the data skills. Women have an average of 1.73, compared to 2.18 for men. These young professionals do hold data skills; however, few hold advanced skills that would be required for roles in data science and future jobs.

LINKING EDUCATION ATTAINMENT AND DATA SKILLS

Individuals with higher levels of educational attainment have **higher self-perceived data skills** (table 2). Very few people with a postgraduate qualification have scored a zero

or one, whereas very few people without a postgraduate qualification deem themselves to be experts overall. Postgraduate degrees, specifically PhD's, require extensive research often involving data use. Those who go through postgraduate studies therefore have an avenue to hone their data skills. The table below is indicative to the fact that lower levels of education do not have robust interventions to teach data skills.

The data substantiates the need for basic data skills development at levels of basic and tertiary education. **Developing data skills mainly through higher education limits the pool of data skilled young professionals with or without qualifications.** In addition, systemic initiatives that address the education gap remains a dependency for increasing the number of young professionals with advanced data skills.

Table 2: Educational attainment and perceived overall data skills level

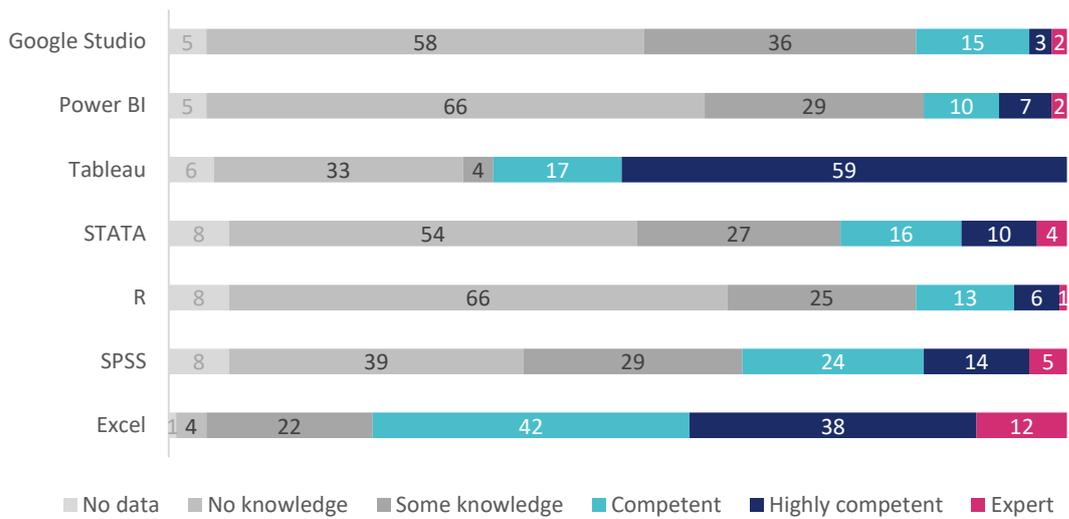
Education attained	No knowledge or some knowledge	Competent	Highly competent	Expert
High school certificate	6%	3%	3%	1%
Post high school	11%	5%	3%	4%
Bachelor's degree	8%	5%	8%	3%
Honours	0%	3%	3%	7%
Masters	3%	4%	9%	8%
PhD	0%	0%	3%	3%

DATA TOOL EXPERTISE

Using the skill level scale, we asked participants to rank their skills for various data software. The software platforms were divided into two categories – data analytics and data visualisation software. MS Excel is a tool that young professionals are required to use, and many respondents confirm is a tool they are knowledgeable in (figure 4). However, **27 (23%) of respondents had no knowledge on all other data analysis software listed (Stata, R, and SPSS)**. Forty-five (38%) respondents had no knowledge of any of the visualisation software listed (Google Studio, Power BI, and Tableau).

R and Tableau tool experience are considered among the top 5 data skills in high demand ([University of Massachusetts, 2019](#)). The demand for skills in analysis tools beyond MS Excel and knowledge of dashboard tools, particularly in fields requiring data analysis, is increasing. In addition, the technical skills need to be matched with soft skills such as communication, critical thinking, and problem solving. Young professionals must be able to make appropriate interpretations for the specific field of use to draw true value from the data.

Figure 4: Common data tool skills levels



EMPLOYMENT OF SKILLED DATA PROFESSIONALS

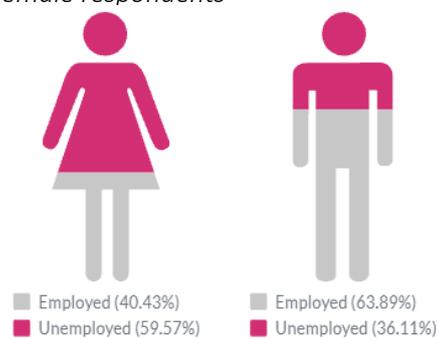
Unemployment in South Africa alone has consistently been above 20% for the past decade, with women and youth being the most vulnerable (StatsSA, 2019). Low levels of employment can be extrapolated across the continent.

Almost half of the respondents in the data skills survey are currently not employed, the majority of those not employed are women (figure 5). Respondents aged 18-24 years have the highest proportion of unemployed individuals. Seventy-one percent of those aged 18-24 years are currently not employed as compared to 42% of those aged 25-34 years. Results indicate that 7 respondents aged 18-24 who have completed their studies were unemployed while 28 respondents aged 25-34 with completed education were unemployed. Seven people without a tertiary degree have been unemployed for more than 3 years.

Young professionals with at least a bachelor’s degree are less likely to report unemployment. There is an unequal distribution across the state of employment when considering the level of educational qualifications acquired. Combined, individuals who do not possess at least a bachelor’s degree account for 35% of

the sample. These individuals, however, make up 59% of those who are unemployed. This translates into 75% of persons without a degree being unemployed. This is an incredibly high number considering that only 29% of those with at least a bachelor’s degree are unemployed. It should be noted that having a higher degree does not make one immune to unemployment as seven out of 29 Master’s degree holders and two out of six the PhD holders were not employed at the time of the survey. An interesting finding is that of the seven who are self-employed, only one of them hold a postgraduate degree.

Figure 5: Employment status male and female respondents



Unemployed young professionals indicate that lack of opportunities and experience are barriers in employment

The field of study one goes through is arguably as important as the level of education acquired when considering employment opportunities. It was previously mentioned that 34% of those with a tertiary education studied a field that is directly associated with data, more than 70%

When further disaggregating employment status by data skills levels, those who hold an employment position have an average score of 2.12. Those who are currently unemployed scored 1.87. Both scores are related to an average level of competence across data skill

are in fields with any level of data related training. Many of these people are currently engaged in some sort of employment role. Out of the 25 employed people, 20 of them stated that their current employment corresponds with the qualification they have obtained.

areas. There are 31 people whose composite scores put them at high competence and 14 whose scores correspond with the expert level. Out of these people, 19 of them were unemployed.

EMPLOYMENT CHALLENGES

Data skills survey participants provided some challenges they have faced regarding employment opportunities, more specifically in the data field. Most challenges reported cited the lack of opportunities, limited skills or the need for further education/training and lack of level of experience required Figure 6.

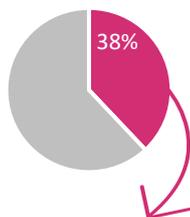
Other reasons for employment challenges include internet/data access, coronavirus lockdowns, no qualification in a data field, lack of resources, scams in job advertisements, no call-backs, and low salaries offered.

Table 3: Employment status of respondents with education on data fields

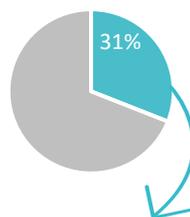
Data related fields	Unemployed	Employed
Computer science	1	0
Data science	1	0
Demography	1	3
ICT	1	0
M&E	1	9
Public Health	2	7
Statistics	1	6
Total	8	25

Figure 6: Challenges in accessing data role employment

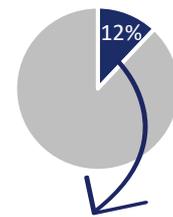
Lack of opportunities



Limited skills



Level of experience



<i>"It is often difficult to find opportunities within the social research space"</i>	<i>"I do not possess the requisite skill in the use of data. Effort to acquire these skills were hampered by the fact that classes where these things are taught, are expensive"</i>	<i>"Most companies, that have responded to my inquiry, have questioned my level of experience within the field"</i>
<i>"[There are] no opportunities listed in my area"</i>	<i>"Acquiring the right skills for the job demand"</i>	<i>"I think one of the challenges I normally get is lack of maximum experience required."</i>

DATA PROFESSION ASPIRATIONS

Young professionals remain positive about their future in data roles. 38 of 119 are currently in data related jobs and 58 aspire employment in a data related role.

"I would like to work as a BI (Business) Analyst, or in a Consulting role with emphasis on Data Visualisation"

"I would like to become a data analysis/monitoring and evaluation specialist within the research field."

"I have four years' experience in data collection management and analysis for different organization, I aspire to be a renowned data scientist."

"I am looking for a role that explores a combination of my data skills. My preference is a role that allows me to work on different projects for a specified time that employ a different set of data skills from time to time."



ACTIONS FOR ORGANISATIONS DRIVING DATA SKILLS IN AFRICA

Digitisation is one of the drivers of the 4IR and has already shown how much value it can add to society. In Africa's context, it can assist in unlocking developmental outcomes across the board. Improving the general skills levels in data competencies in African youth is key. Digitisation and data competencies, however, are generally taught to those enrolled in higher levels of post high school education. This is an obstacle given that less than 10% of Africans have post high school qualifications. **So how can we act to ensure we support changes in data skills within school, after school and increasing access to jobs associated to the imminent rise in data skills demands?** Here are the actions we propose:

1. PARTNER FOR POSITIVE CHANGE IN DATA SKILLS

Private, public, and civil society need to join forces to improve the data skills African youths possess. This will have a knock-on effect on jobs in the future as data jobs demand increases. Creating links between data skills providers and demands will support individuals in accessing data skills jobs and identifying what specific skills to hone. Given the need for data skills to be developed at various levels will require the engagement of NGOs, schools, and government departments to compliment curriculum changes that allow for individuals that do not complete high school or enter tertiary education to have basic data skills. This should include data management, advanced data analysis, and data visualisation, which are good foundational skills, while leaving people with a great platform to launch into more advanced data competencies.

2. ASSESS ORGANISATIONAL SKILLS GAPS, DESIGN, AND IMPLEMENT IN-HOUSE DATA RE-SKILLING PROGRAMMES

Entities in both the private and public sectors need to make conscious efforts to reskill their staff. Organisations can identify which digital competencies will improve their efficacy and efficiency. This together with a skills audit of their staff will allow companies to identify their skills gaps and start putting interventions in place to close these skills gaps. Some of this may not require high-cost investment but leveraging of CSO and open-source learning programmes, as well as creating in-house projects for individuals to test data skills.

There is appetite for data related roles as approximately half of the respondents aspire to have a data related job. The two suggestions above will work on empowering current and future workforces. We do, however, require interventions aimed at making sure that there is alignment between employers (both private and public sectors), and education and skilling interventions to make sure people are empowered with relevant skills for now and the future. Furthermore, this alignment can create better linkages between employers, education, and skilling interventions, and those looking for employment. This will improve transparency and thus improve access to work opportunities.

3. TARGET SYSTEMIC CHANGE NOW THAT WILL INFLUENCE YOUTH ACCESS TO DATA SKILLS JOBS IN THE FUTURE

The long-term solution, particularly for Africa, is educational reform. Government and supporting organisations must realign on what critical learning outcomes are and have curricula that is relevant for youth entering 4IR jobs. Shifts in curriculum and policy which addresses education gaps must remain at the fore. With this, focus should be placed in integrating data skills within STEAMIE and 4IR interventions at grassroots levels.

METHODOLOGY

This survey had 119 respondents, who reside in 22 countries. The survey was administered online, open to anyone willing to take it and was disseminated through Data Innovator's social media and network base.

To understand participants' data capabilities, they were asked to rate their level of skill for the six skills we are assessing for – data capturing, data cleaning, basic data analysis, data management, advanced data analysis, and data visualisation. They could describe themselves as having either no knowledge, some knowledge, basic competence, highly competent, or expert level skills. We assigned the following score to the skills level scale: Expert - 4, Highly competent - 3, Basic competence - 2, Some knowledge - 1, and No Knowledge - 0.

We have identified three main limitations to the study:

Selection bias - Respondents were convenience sampled, this limits the representivity of the sample across countries, but provides a diverse global perspective. This will affect some of the results such as the levels and fields of study, and to a certain extent, the level of data skills.

Social desirability bias - With respondents having to rate their own level of skills, this survey is prone to social desirability bias. The implication of this is that the data skills score we have calculated could be elevated relative to their true skill levels

Sample size - With further editions of this survey, emphasis will be placed on increasing the sample size and more targeted sampling for representativity.

REFERENCES

- Bowmaker-Falconer, A., & Herrington, M. (2020). *Igniting start-ups for economic growth and social change*. Stellenbosch: Global Entrepreneurship Monitor.
- Magwentshu, N., Rajagopaul, A., Chui, M., & Singh, A. (2019). *The future of work in South Africa: Digitisation, productivity, and job creation*. Johannesburg: McKinsey & Company.
- Nanayakkara, W. (2019, November 8). Data, the Driving Force of the Fourth Industrial Revolution (4IR). Sri Lanka.
- NTUC LearningHub. (2020). *Closing the data skills gap in the Covid-19 era and beyond*. Singapore: NTUC LearningHub.
- The Royal Society. (2019). *Dynamics of data science skills: How can all sectors benefit from data science talent?* London: The Royal Society.
- van der Berg, S., Gustafsson, M., & Malindi, K. (2020). *Education and skills for the economy and links to labour markets in South Africa*. Stellenbosch: Resep, Stellenbosch University.
- World Bank. (2017, October). An Analysis of issues shaping Africa's future. *Africa's Pulse*.
- World Economic Forum. (2017). *The Future of Jobs and Skills in Africa*. Geneva: World Economic Forum.
- World Economic Forum. (2018). *The Future of Jobs Report*. Geneva: World Economic Forum.
- Zahidi, S. (2020). We need a global reskilling revolution - here's why. *World Economic Forum Annual Meeting*. Geneva: World Economic Forum.



ABOUT US

Helping organizations push data boundaries to create positive futures.

Data Innovators is a social enterprise led by a team of young Africans who are passionate about making systemic change in development through data. Data Innovators Group was founded in 2021, borne from the partnership between Advisory Services at New Leaders Foundation NPC and Data Innovator PTY Ltd. Our core services include Data Driven Solutions, Monitoring, Evaluation, Research and Learning Services, Capability Building and Project Implementation.

Three ways the Data Innovators aims to help organisations overcome data boundaries for a positive future are by:

- Innovating and strengthening impact measurement in across sectors globally
- Collaborating in efforts to drive strategic business intelligence in the development sector broadly.
- Providing thought leadership and reach new frontiers in setting up data and MERL systems.



Contact:

The Education Hub, 6 Blackwood Ave, Parktown, Johannesburg, 2193, South Africa | www.datainnovators.co | info@datainnovators.co | [@data_innovators](https://twitter.com/data_innovators)