



UN YOUTH HACKATHON

DATA ANALYSIS FOR POVERTY & COVID-19

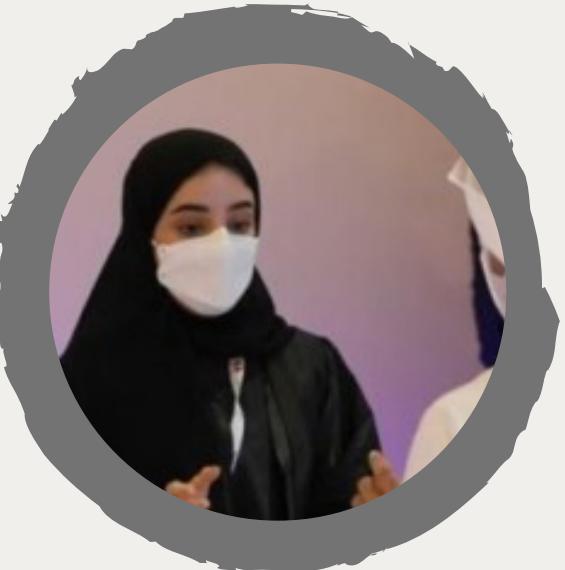
IdeaX team

who are we

MEMBERS OF IDEAX TEAM



Shamma Alsaedi
computer science
UAEU



Jawaher Alghfeli
computer engineering and AI
UAEU

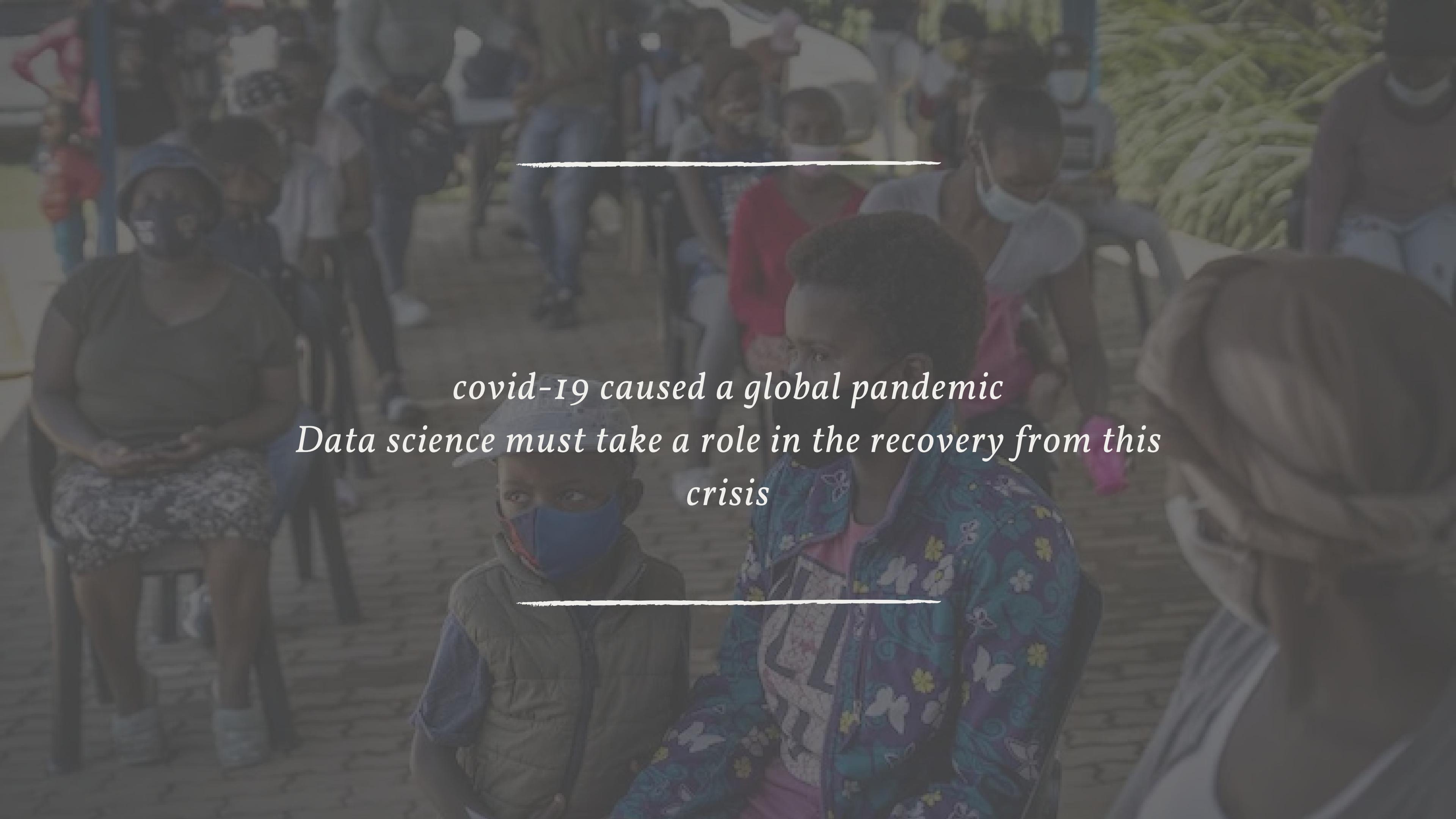


Aisha Alraeesi
information security and AI
UAEU

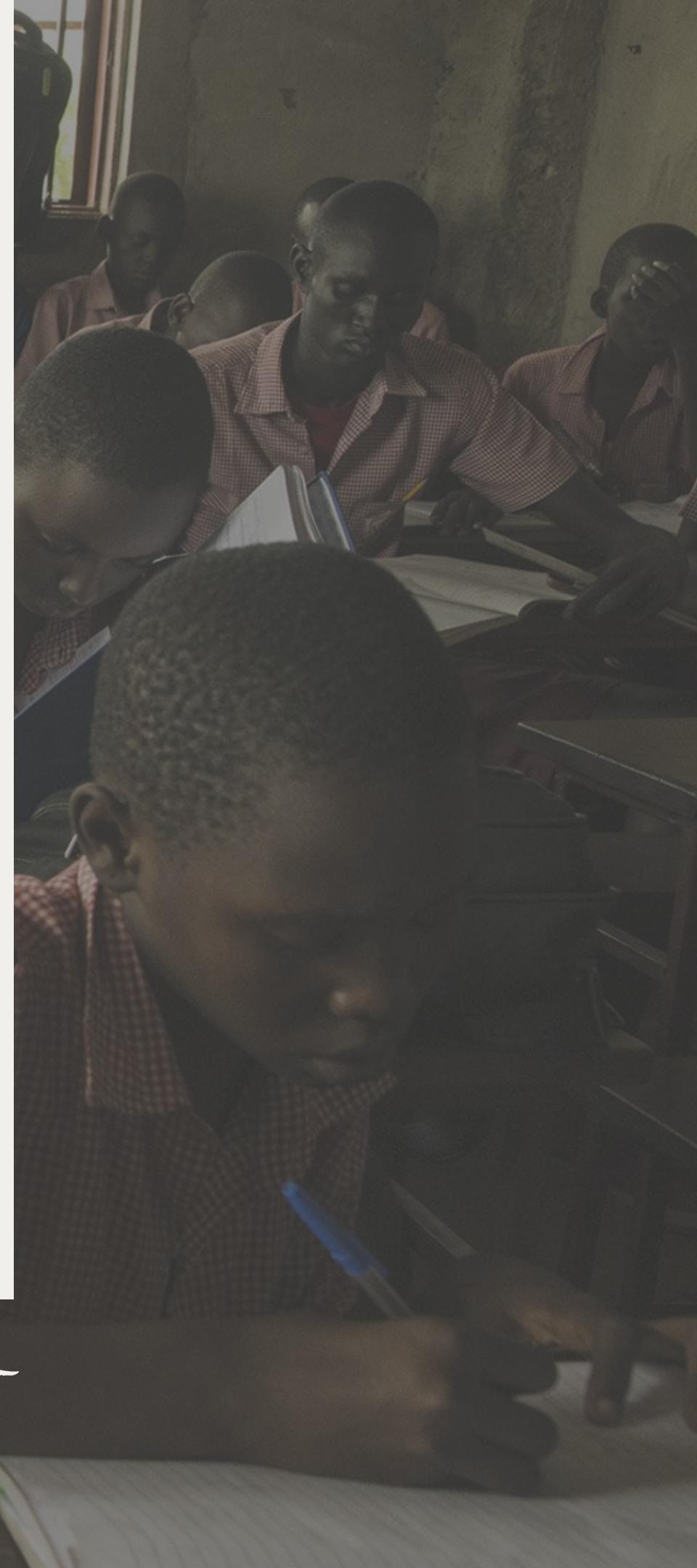


Presentation Outline

- WHY?
- TARGETED AREA
- OUR GOAL
- THE ADDRESSED POVERTY PROBLEMS
- DATASET USED
- DATA CLEANING
- VISUALIZATION
- CLUSTERING
- STUDY USAGE
- FUTURE ENHANCEMENT
- CONCLUSION
- SOURCES

A large, diverse crowd of people is gathered outdoors, all wearing face masks. The scene is set in what appears to be a park or a public square. In the foreground, a person wearing a blue patterned jacket and a mask is looking towards the camera. Behind them, many other individuals of various ages and ethnicities are visible, some in groups and some alone, all maintaining a sense of social distancing. The background shows more people and some greenery, creating a dense, textured composition.

*covid-19 caused a global pandemic
Data science must take a role in the recovery from this
crisis*



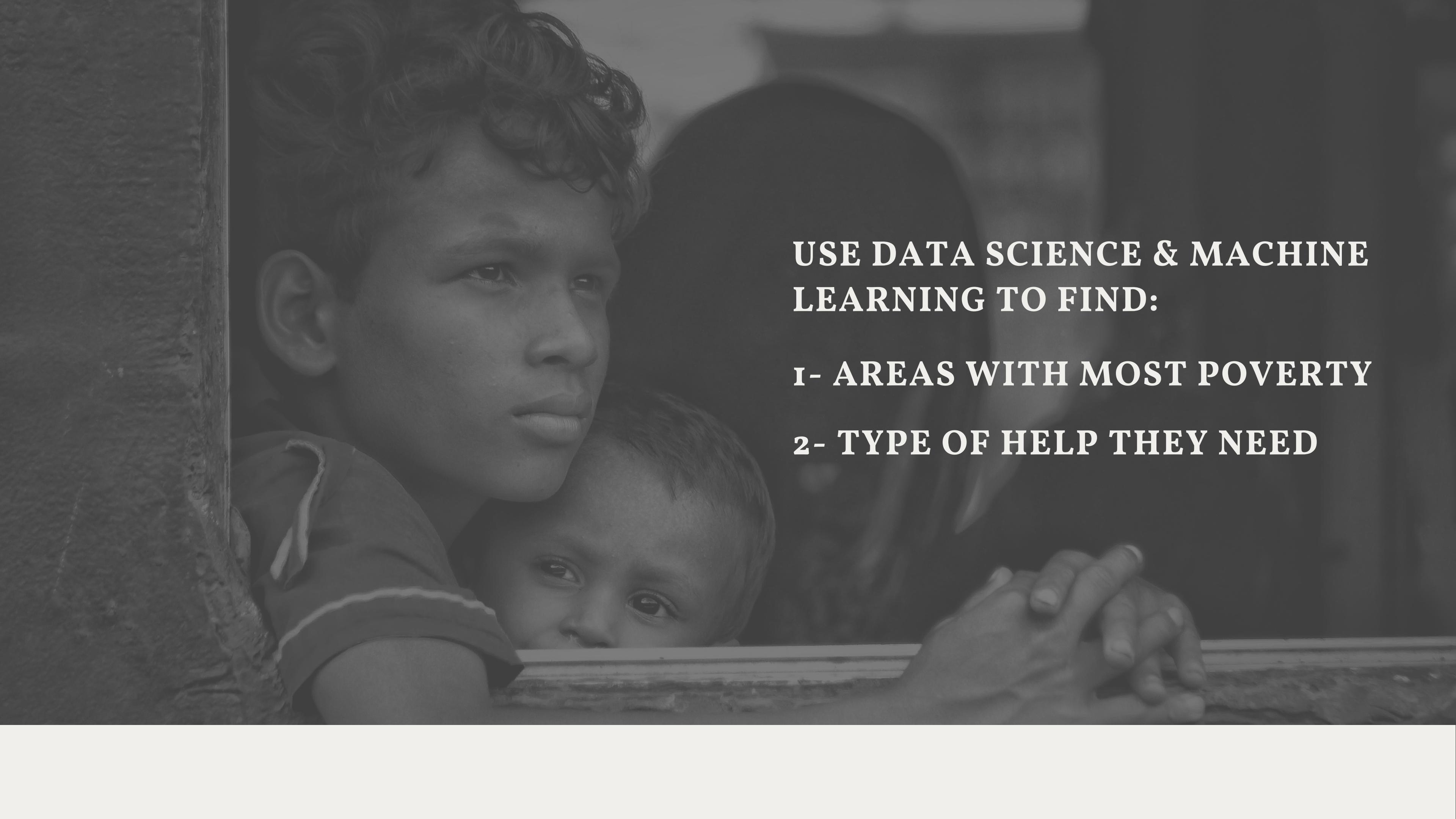
Targeted area

SDG1:

NO POVERTY

Our Goal

To get the right help to the
people in most need of it.



**USE DATA SCIENCE & MACHINE
LEARNING TO FIND:**

- 1- AREAS WITH MOST POVERTY**
- 2- TYPE OF HELP THEY NEED**

The addressed poverty problems



Food Security



Healthcare



Education



Shelter



Income

Dataset used

Socioeconomic impacts of the COVID-19 pandemic on households by the World Bank

December 6, 2021



Dataset used



data cleaning

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
row_id	code	wave	month	indicatc	indicatc	urban_r	industry	sample_	sample_	year	country	region	region_	FCS	income	lending	indicatc	indicatc	unit_me	measure	GDP_pc	In_GDP	weight
1	AFG	WAVE1		9	Demo_edu	70	National	All	5013	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Demograph	Level of edu	Percent of l	2065	8	Individual
2	AFG	WAVE1		9	Demo_edu	5	National	All	5013	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Demograph	Level of edu	Percent of l	2065	8	Individual
3	AFG	WAVE1		9	Demo_edu	16	National	All	5013	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Demograph	Level of edu	Percent of l	2065	8	Individual
4	AFG	WAVE1		9	Demo_edu	10	National	All	5013	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Demograph	Level of edu	Percent of l	2065	8	Individual
5	AFG	WAVE1		9	Demo_mal	71	National	All	5053	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Demograph	Male respo	Percent of l	2065	8	Unweight
6	AFG	WAVE1		9	Fina_access	82	National	All	5057	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Financial	Could not a	Percent of l	2065	8	Individual
7	AFG	WAVE1		9	Heal_need	58	National	All	5052	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Needed me	Percent of l	2065	8	Individual
8	AFG	WAVE1		9	Heal_reaso	88	National	All	567	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Could not r	Percent of l	2065	8	Individual
9	AFG	WAVE1		9	Heal_reaso	4	National	All	567	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Could not r	Percent of l	2065	8	Individual
10	AFG	WAVE1		9	Heal_reaso	4	National	All	567	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Could not r	Percent of l	2065	8	Individual
11	AFG	WAVE1		9	Heal_reaso	6	National	All	567	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Could not r	Percent of l	2065	8	Individual
12	AFG	WAVE1		9	Heal_reaso	0	National	All	567	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Could not r	Percent of l	2065	8	Individual
13	AFG	WAVE1		9	Heal_rece	75	National	All	2768	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Health	Received m	Percent of l	2065	8	Individual
14	AFG	WAVE1		9	Hous_ablep	36	National	All	1853	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Housing	Able to pay	Percent of l	2065	8	Individual
15	AFG	WAVE1		9	Hous_move	4	National	All	3034	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Housing	Moved sinc	Percent of l	2065	8	Individual
16	AFG	WAVE1		9	Hous_own	62	National	All	5055	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Housing	Have owne	Percent of l	2065	8	Individual
17	AFG	WAVE1		9	Labo_empl	45	National	All	4883	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Currently e	Percent of l	2065	8	Individual
18	AFG	WAVE1		9	Labo_empl	58	National	All	2937	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
19	AFG	WAVE1		9	Labo_empl	65	National	Mining/Ma	530	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
20	AFG	WAVE1		9	Labo_empl	73	National	Other servi	1085	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
21	AFG	WAVE1		9	Labo_empl	19	National	Commerce	609	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
22	AFG	WAVE1		9	Labo_empl	58	National	Agriculture	434	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
23	AFG	WAVE1		9	Labo_none	42	National	All	2937	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
24	AFG	WAVE1		9	Labo_none	42	National	Agriculture	434	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
25	AFG	WAVE1		9	Labo_none	35	National	Mining/Ma	530	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
26	AFG	WAVE1		9	Labo_none	27	National	Other servi	1085	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
27	AFG	WAVE1		9	Labo_none	81	National	Commerce	609	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Employed r	Percent of l	2065	8	Individual
28	AFG	WAVE1		9	Labo_nonfa	33	National	All	5057	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Engaged in	Percent of l	2065	8	Individual
29	AFG	WAVE1		9	Labo_nono	36	National	All	364	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Labor	Unable to p	Percent of l	2065	8	Individual
30	AFG	WAVE1		9	Prev_AS_ot	98	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual
31	AFG	WAVE1		9	Prev_AnySt	98	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual
32	AFG	WAVE1		9	Prev_Sfood	96	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual
33	AFG	WAVE1		9	Prev_Sfood	97	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual
34	AFG	WAVE1		9	Prev_Sfood	97	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual
35	AFG	WAVE1		9	Prev_Sfood	94	National	All	3354	5057	2020	Afghanistar	South Asia	SAR	High-Inten	Low incom	IDA	Food Secur	Able to acc	Percent of l	2065	8	Individual

data cleaning

*Removed unnecessary
data*

such as region code and the
country lending category

region_code	lending_category	indicator_topic	indicator_display
SAR	IDA	Demographic	Level of education - No education
SAR	IDA	Demographic	Level of education - Any Primary
SAR	IDA	Demographic	Level of education - Any Secondary
SAR	IDA	Demographic	Level of education - Any Postsecondary
SAR	IDA	Demographic	Male respondents (% of respondents)
SAR	IDA	Financial	Could not access a financial service
SAR	IDA	Health	Needed medical treatment but could not afford it
SAR	IDA	Health	Could not receive medical attention
SAR	IDA	Health	Could not receive medical attention due to cost
SAR	IDA	Health	Could not receive medical attention due to cost
SAR	IDA	Health	Could not receive medical attention due to cost
SAR	IDA	Health	Received medical attention
SAR	IDA	Housing	Able to pay rent next month

data cleaning

Simplified the data

e.g. ignored the industry, area
whether urban or rural, ignored
the waves and took the

The image shows two separate filter panels side-by-side. Both panels have a header with a red circular icon and the field name. They both include 'Sort' options (A↓ Ascending, Z↓ Descending) and 'By color:' dropdowns set to 'None'. The 'industry' panel has a 'Filter' section with a 'Choose One' dropdown and a 'Search' input field. Below these are several checkbox options: '(Select All)', 'Agriculture', 'All' (which is checked), 'Commerce', 'Mining/Manufacturing', 'Not reported', and 'Other services'. The 'All' option is highlighted with a green bar. The 'urban_rural' panel also has a 'Filter' section with a 'Choose One' dropdown and a 'Search' input field. It lists: '(Select All)', 'National' (which is checked), 'Rural', 'Urban', and '(Blanks)'. The 'National' option is highlighted with a green bar. Both panels have 'Auto Apply' checkboxes and 'Apply Filter' and 'Clear Filter' buttons at the bottom.

industry

urban_rural

Sort

A↓ Ascending Z↓ Descending

By color: None

Filter

By color: None

Choose One

Search

- (Select All)
- Agriculture
- All
- Commerce
- Mining/Manufacturing
- Not reported
- Other services

Auto Apply

Apply Filter Clear Filter

Sort

A↓ Ascending Z↓ Descending

By color: None

Filter

By color: None

Choose One

Search

- (Select All)
- National
- Rural
- Urban
- (Blanks)

Auto Apply

Apply Filter Clear Filter

data cleaning

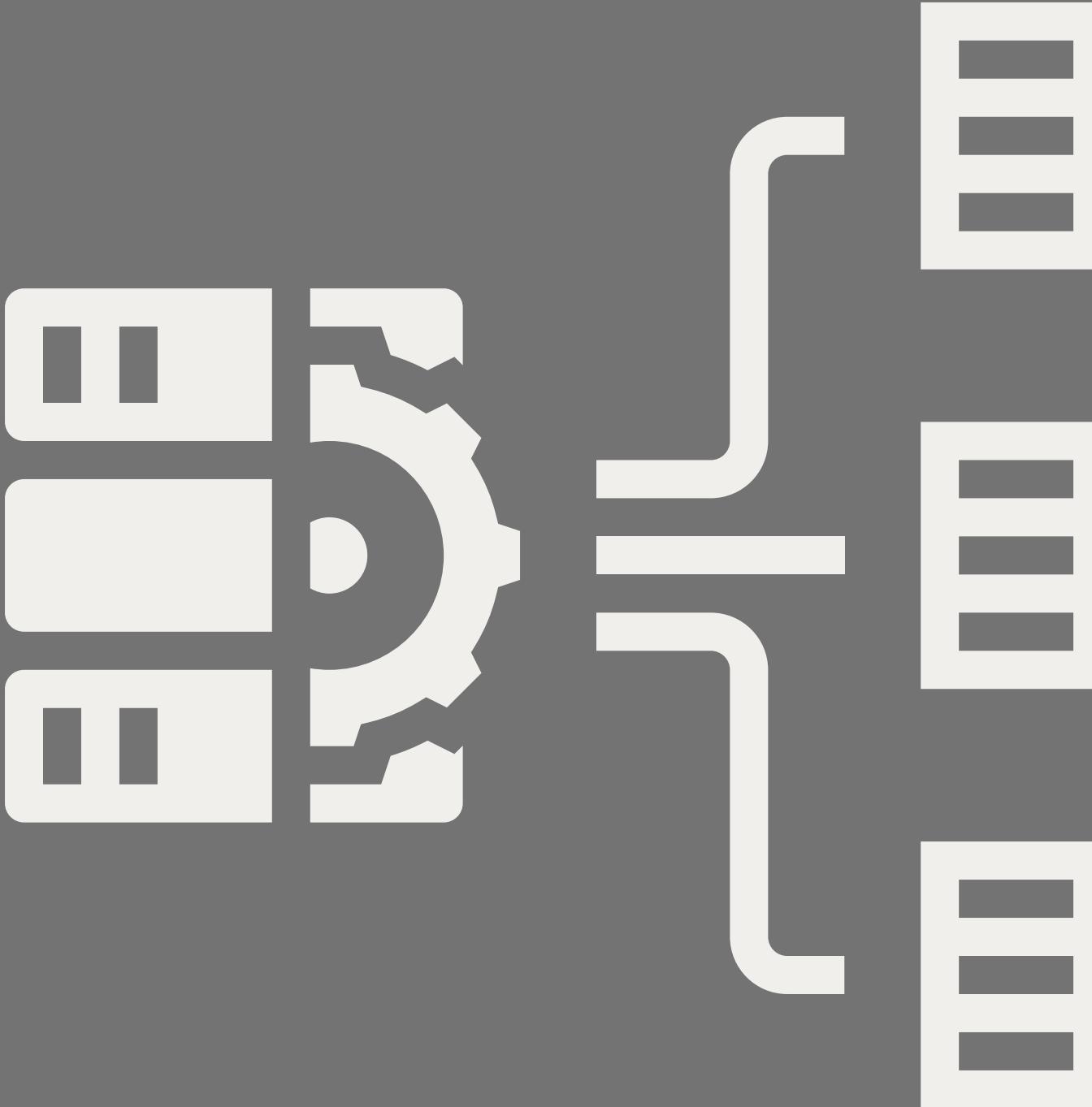
*Decreased the number
of indicators to 40

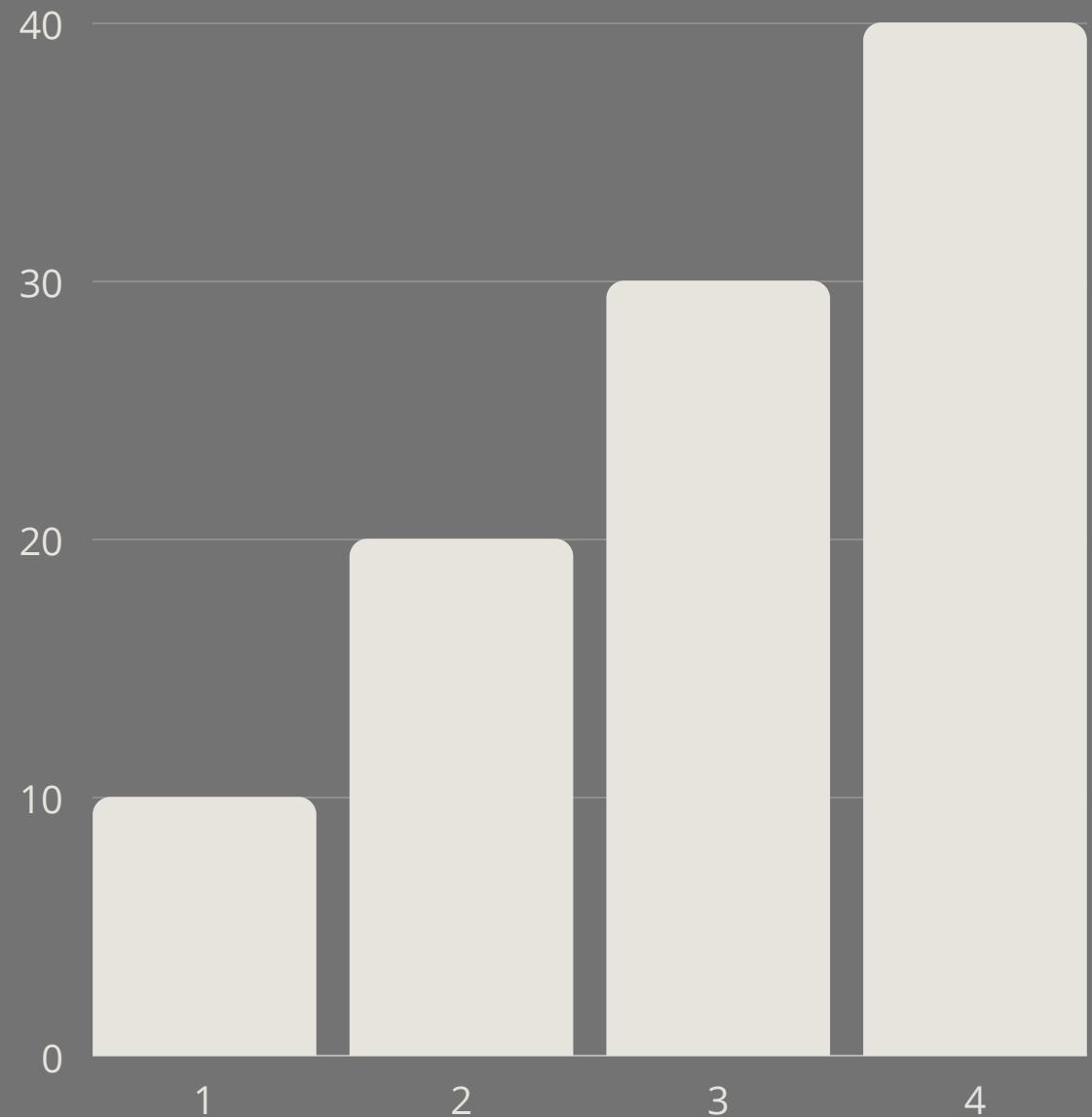
chose only those that are
related to food, health, shelter,
education and income, then
categorized them accordingly*

28	26	positive food	Prev_AnyStaple	% able to access [staple food item] in the past 7 days when needed? - any staple food	
29	27	Housing	Hous_moved	% of HHs moved since the outbreak and from where they moved	
30	28	Positive shulter	Hous_own	Ownership of current housing	
31	29		Hous_rooms	Number of rooms occupied	
32	30	Positive shulter	Hous_ablepayre	% of HHs able to pay rent for the next month	
33	31	Food Security	FS_worried	In the last 30 days, you or any other adult in your household were worried about not having enough food to eat because of lack of money or other resources?	
34	32		FS_unable	In the last 30 days, you or any other adult in your household, were unable to eat healthy and nutritious/preferred foods because of a lack of money or other resources?	
35	33		FS_atefew	In the last 30 days, you or any other adult in your household, ate only a few kinds of foods because of a lack of money or other resources?	
36	34		FS_skipmeal	In the last 30 days, you or any other adult in your household, had to skip a meal because there was not enough money or other resources to get food?	
37	35		FS_ateless	In the last 30 days, you or any other adult in your household, ate less than you thought you should because of a lack of money or other resources?	
38	36		FS_ranout	In the last 30 days, your household ran out of food because of a lack of money or other resources?	
39	37		FS_hungry	In the last 30 days, you or any other adult in your household, were hungry but did not eat because there was not enough money or other resources for food?	
40	38		FS_day	In the last 30 days, you or any other adult in your household, went without eating for a whole day because of a lack of money or other resources?	
41	39	Financial	Fina_access	% of HHs that could not access a financial institution (ATM, bank, mobile money) when needed	
42	40		Fina_reason1	Reasons for non-accessibility - Bank was closed	
43	41		Fina_reason2	Reasons for non-accessibility - Movement restriction	
44	42		Fina_reason3	Reasons for non-accessibility - Afraid to go out because of coronavirus	
45	43	Assets & Services	positive food	Asse_drink	% HH had sufficient water to drink in the last week
46	44		positive health living condition	Asse_wash	% HH had sufficient access to water to use to wash in the last week
47	45		positive health living condition	Asse_soap	% HH had sufficient soap to wash hands with in the last week
48	46		positive edu	Asse_intern	% HH used internet service last week
49	47	Education	positive edu	Educ_attc	% of HHs with children attending school (primary & secondary) before school closures
50	48			Educ_comp	% of HHs with children engaged in learning/education activities since school closures: Completed assignments provided by the teacher
51	49			Educ_mobi	% of HHs with children engaged in learning/education activities since school closures: Used mobile learning apps
52	50			Educ_tvprog	% of HHs with children engaged in learning/education activities since school closures: Watched educational TV programs
53	51			Educ_radio	% of HHs with children engaged in learning/education activities since school closures: Listened to educational programs on radio
54	52			Educ_teacher	% of HHs with children engaged in learning/education activities since school closures: Session/meeting with Lesson Teacher (tutor)
55	53			Educ_other	% of HHs with children engaged in learning/education activities since school closures: OTHER
56	54		positive edu	Educ_any	% of HHs with children engaged in any learning/education activities since school closures
57	55	Health	negative health	Heal_need	% of HHs needing medical treatment since COVID-19 outbreak
58	56		positive health	Heal_rece	% of HH received medical attention over HH that seek medical attention
59	57		negative health	Heal_reason1	Could not receive medical attention due to lack of money (% HHs that could not receive medical attention when needed)
60	58		negative health	Heal_reason2	Could not receive medical attention due to no medical personnel (% HHs that could not receive medical attention when needed)
61	59		negative health	Heal_reason3	Could not receive medical attention due to medical facility was full (% HHs that could not receive medical attention when needed)
62	60		negative health	Heal_reason4	Could not receive medical attention due to medical facility was closed (% HHs that could not receive medical attention when needed)
63	61		negative health	Heal_reason5	Could not receive medical attention due to hospital/clinic did not have enough supplies or tests (% HHs that could not receive medical attention when needed)
64	62			Heal_reason6	Could not receive medical attention due to afraid/concerned about catching COVID19 (% HHs that could not receive medical attention when needed)
65	63			Heal_reason7	Could not receive medical attention due to restrictions (stay-at-home orders) (% HHs that could not receive medical attention when needed)
66	64			Heal_reason8	Could not receive medical attention due to lack of transportation (% HHs that could not receive medical attention when needed)
67	65			Heal_reason9	Could not receive medical attention due to other reasons (% HHs that could not receive medical attention when needed)
68	66	Labor	positive income	Labo_empl	% of respondents currently employed/working

data cleaning

*Normalized the
indicator values*
to make them all represent
poverty





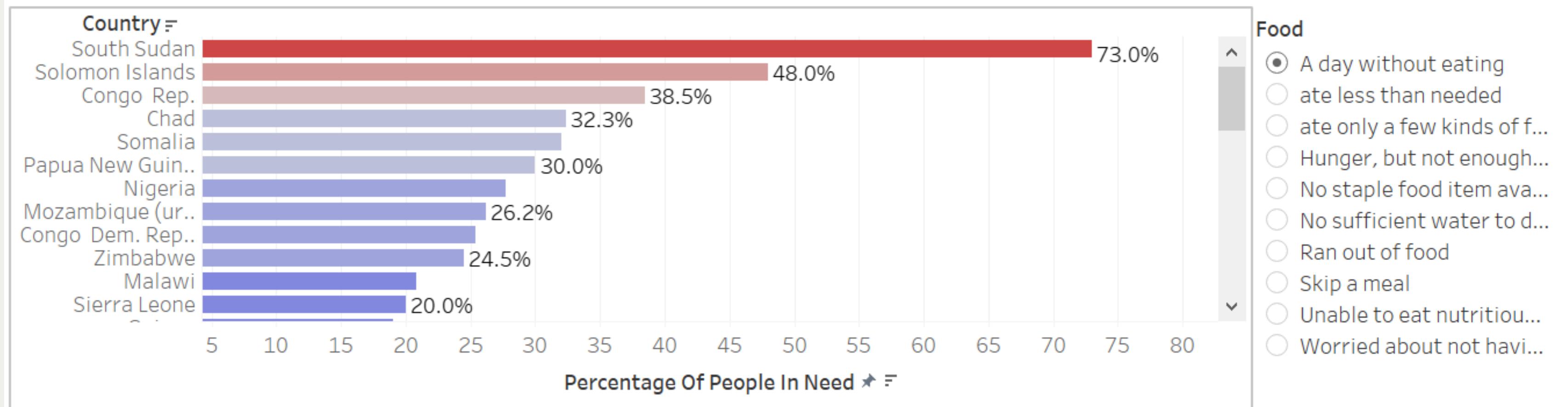
visualization

USING TABLEAU

[the link](#)

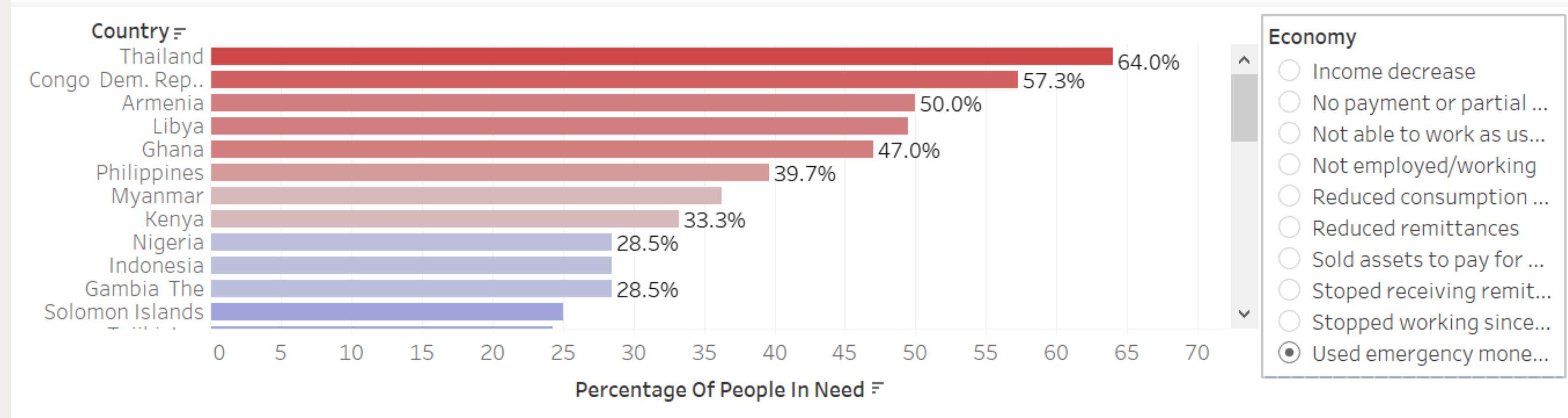
visualization

Food Security



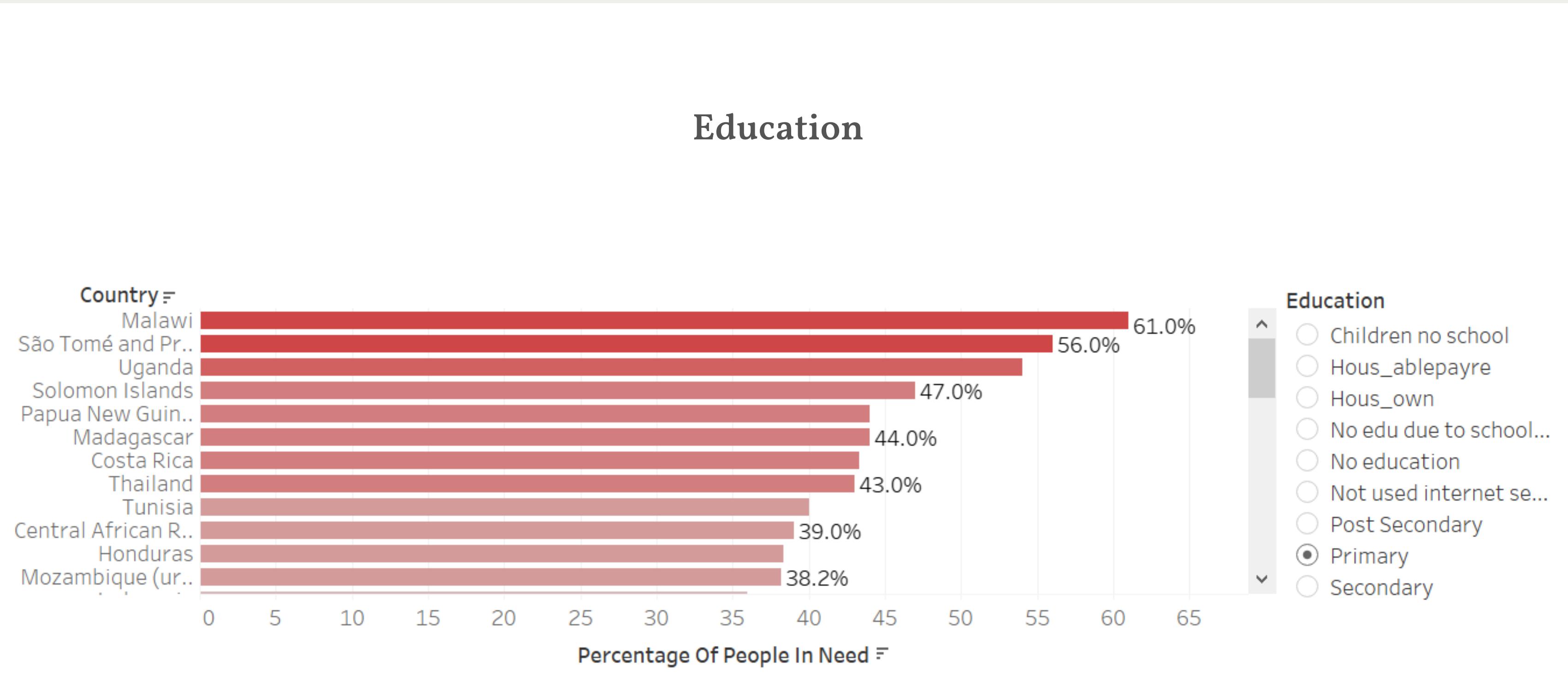
visualization

Economy



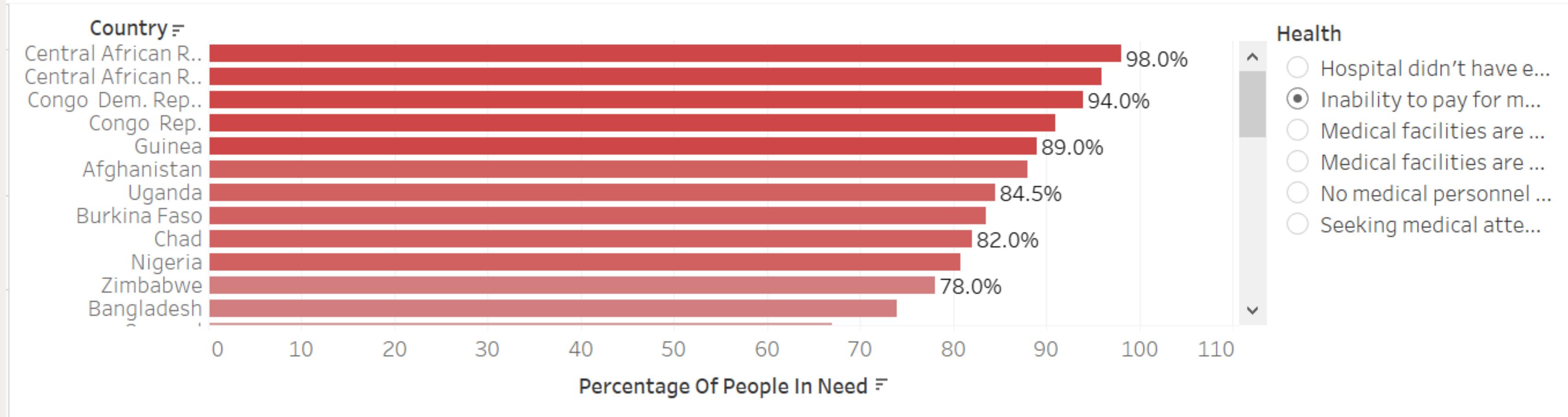
visualization

Education



visualization

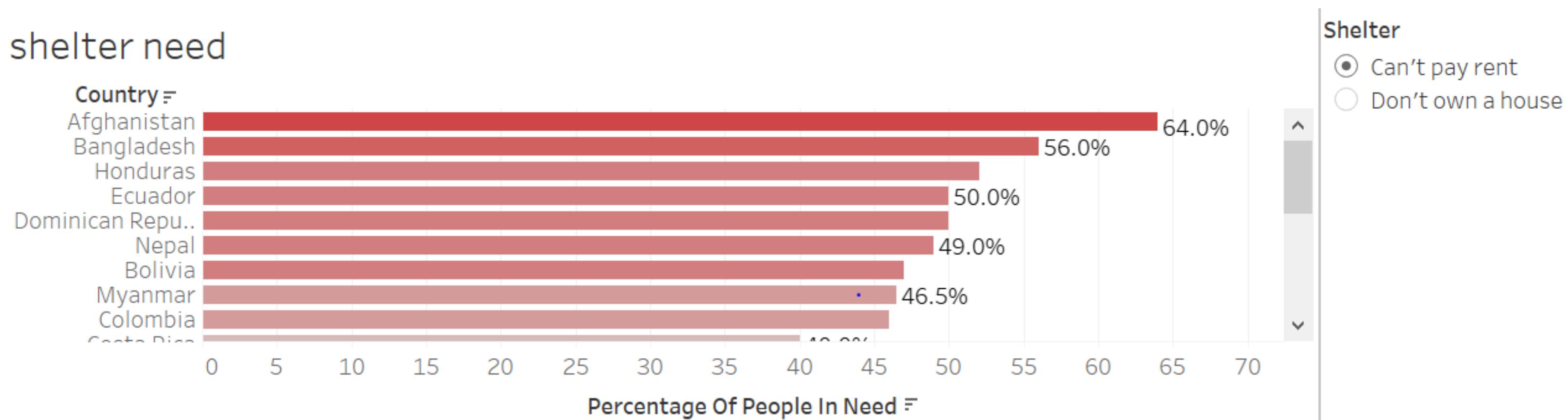
Healthcare



visualization

Shelter

shelter need



Machine learning approach

Clustering (unsupervised learning)
To understand hidden patterns

Problem

not all indicators are available for all countries, couldn't find common

```
In [154]: with pd.option_context('display.max_rows', None, 'display.max_columns', None):
    print(df)
```

	Afghanistan	Argentina	Armenia	Bangladesh	\
Asse_drink	
Asse_intern	
Asse_soap	
Asse_wash	
Copi_reduce	
Copi_sold	
Copi_used	
Demo_Dratio	
Demo_educ1	70.0077	33.0649	
Demo_educ2	4.7486	19.6034	
Demo_educ3	15.523	41.7285	
Demo_educ4	9.72066	5.60316	
Demo_hsize
Demo_male
Demo_male_sample	70.5126	69.3168	
Educ_any	
Educ_atte	
Educ_comp	

solution

- 1- Take the average for each category
- 2- Add another dataset

A	B	C	D	E
Country	food	health	Education	Income
Afghanistan	2	30.25	62.25	55
Argentina	9.8	15.2	30.9473684	34.1764706
Armenia	27.8	23	21.3636364	48.75
Bangladesh	5	24.25	51.25	61
Bhutan	3	1	58.5	49
Bolivia	17.6666667	23.5555556	33.2631579	43.3333333
Burkina Faso	25.0149254	22	22.5714286	15.6923077
Cambodia	18.8275862	0.4	20.9	34.5238095
Central Afric	27.6666667	34.75	56.25	29.4
Central Afric	30	48.25	31.25	36.25
Chad	67.8333333	27.75	86	33.5555556
Chile	7.06666667	12.2	27.1578947	34.3529412
Colombia	24.0666667	14.4	34.5789474	42.4117647

study usage

For Humanity

Poverty Donate

“ Hand by hand, we can reduce the impact of Covid 19 on poverty



A black and white photograph showing a group of approximately eight children of African descent sitting on a rough, textured concrete ledge. Some of the children are wearing simple cloth face masks. They are dressed in various patterns of clothing, including a leopard-print top and a checkered shirt. The background shows a weathered wall with a small opening. A yellow circular button with a downward-pointing arrow is located in the bottom left corner of the image.

Future Improvement

CAN BE APPLIED
TO AREAS IN A
CERTAIN
COUNTRY TO
PRIORITIES HELP

USE MACHINE
LEARNING
CLUSTERING
APPROACH TO
UNDERSTAND
HIDDEN AREAS
VULNERABLE TO
COVID-19 INDUCED
POVERTY

USE MACHINE
LEARNING
APPROACH TO
PREDICT FUTURE
VULNERABLE
AREAS THAT
COVID 19 WOULD
AFFECT THE
MOST

A group of five diverse children are smiling and interacting in a classroom setting. In the foreground, a young boy with dark hair and bangs looks directly at the camera with a neutral expression. Behind him, a girl with short blonde hair is laughing heartily, her hands clasped near her chest. To her right, another girl with long dark hair and a nose piercing is looking towards the camera with a slight smile. In the background, a boy with a shaved head is laughing and looking towards the camera. On the far right, a boy with curly brown hair is holding a green lollipop and looking slightly to the side. The background shows shelves with various items and a yellow wall.

Conclusion

Sources

- <https://www.weforum.org/agenda/2020/04/coronavirus-pandemic-half-a-billion-people-into-poverty/>
- <https://hub.officialstatistics.org/dataset/socioeconomic-households>
- <https://datacatalog.worldbank.org/search/dataset/0037769/Harmonized-COVID-19-Household-Monitoring-Surveys>
- <https://www.nature.com/articles/d41586-020-01393-7>
- <https://www.sciencedirect.com/science/article/pii/S004896972100718X>
- <https://github.com/ngaiyin1760/Factor-analysis-Cluster-analysis-on-countries-classification>