جهـــاز التخطيـــط والإحصـــاء
Planning and Statistics Authority
State of Qatar • دولـــة قطـــر

Regional Workshop on the Changing Role of Official Statistics in the State of Qatar: Why Data Culture Matters

21st -22nd September 2022, Doha, Qatar
Pullman Hotel, Al Thuraya Ballroom - West Bay

ورشـة العمل الإقليمية حول الـدور المتغير للإحصاءات الرسـمية فـي دولة قطر: ثقافة البيانات مهمة

٢١-٢٢ سبتمبر ٢٠٢٢، الدوحة، قطر
فندق بولمان الدوحة، قاعة الثريا - الخليج الغربي

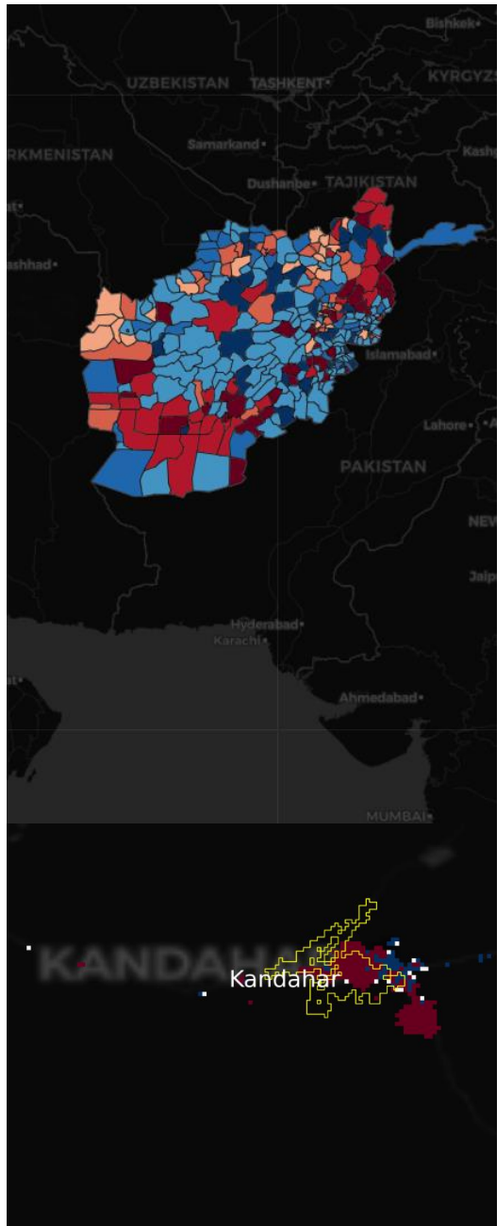
The Use of Alternative Data Sources for Rapid Socio-Economic Analyses: Lessons Learned from COVID-19

VITO INTINI, REGIONAL CHIEF ECONOMIST, UNDP ARAB STATES

Introduction

- Quantifying the socioeconomic impact after extreme shocks is important for pre-disaster planning and post-disaster assessments.
- In the past, field surveys have been the primary source of data to quantify the damages that are inflicted on people and businesses by disasters. But they often suffer from high cost, their implementation can take time, and provide snapshots in specific points in time.
- Under COVID-19, mobile phone surveys have been increasingly used: CATI, interactive voice response (IVR), SMS, web surveys.
- Recently, also satellite imagery, large-scale human mobility data (e.g., mobile phone GPS), and social media have been used to observe and analyze disasters and other shocks.

Night Light Analysis (remote sensing)



In the immediate aftermath of a disaster, it is possible to detect artificial lights, calculate how long they turned off and analyze their trends in the days that follows the event. UNDP uses this data to uncover information about areas where the electric power system has been affected. Moreover, the consumption of electric energy is used as a proxy for economic activities and, in the medium term, can tell us if the area affected by the crisis is also suffering economic consequences.

Finally, electricity consumption per capita is an indicator of wealth. We compute the Night Lights Development Index to provide localized income distribution before the crisis. Such data are used as benchmark data to understand the affected population and as input to further analysis.

Target

- Economic consequences
- Poverty

Data source

NASA *(NASA release images every week, high res monthly)*

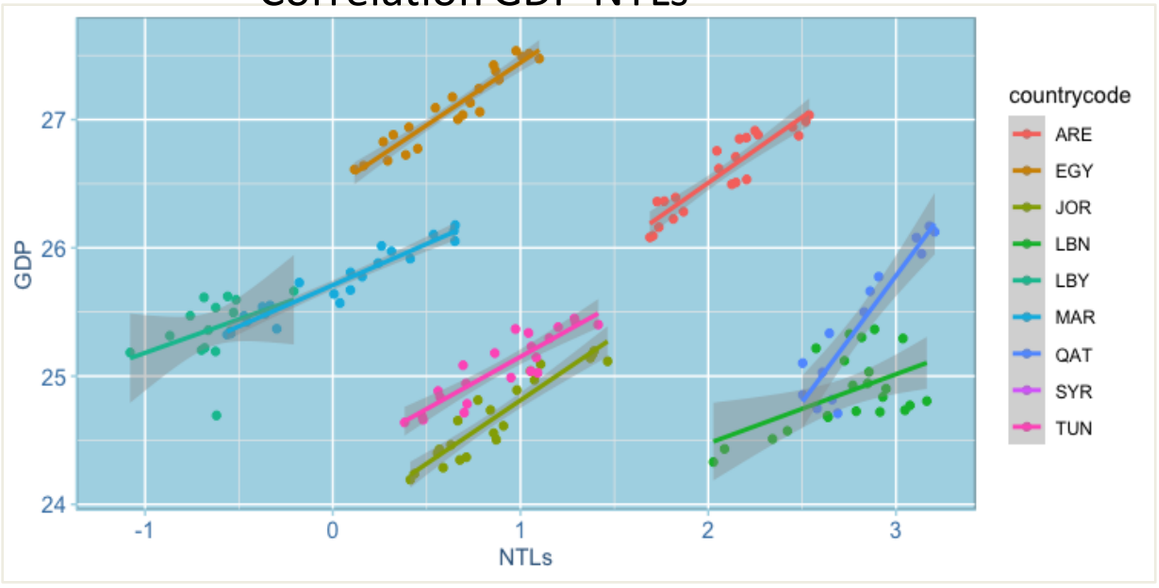
Macroeconomic Impact Using NTLs in the Arab States Region

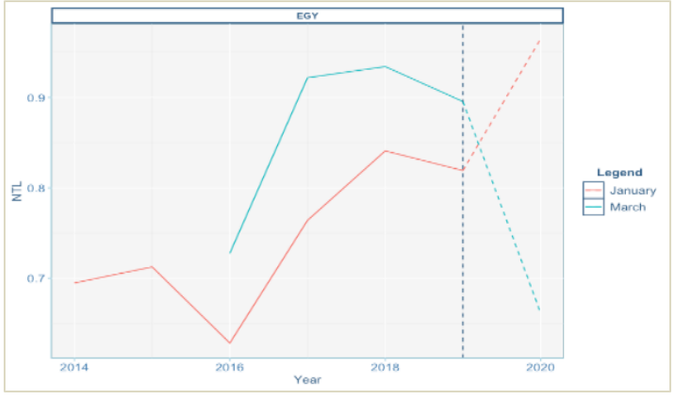
- Using satellite imagery, we obtained a measure of the elasticity of GDP to NTLs for Arab States obtained from:

$$GDP_{it} = \alpha + \beta_1 NightLights_{it} + \beta_2 X_{it} + \lambda_t + \delta_i + \epsilon_{it}$$

where: $NightLights_{it}$ is the natural logarithm of the average digital number, extracted from NOAA, X_{it} is a vector of covariates, namely population density and electricity consumption per capita, λ_t are year-specific fixed effects, δ_i are country-specific fixed effects.

Correlation GDP-NTLs





Buildings Damage Detection through satellite images (remote sensing)

- Based on satellite images, UNOSAT provides preliminary findings on buildings that are potentially damaged, severely damaged, and destroyed.

- UNDP crosses this information with baseline data on household such as the poverty or unemployment rate to quickly detect areas which might need an urgent support

Target

•Damage to buildings
•Population affected

Data source

UNOSAT building damage assessment

unitar
United Nations Institute for Training and Research
UNOSAT

Social Media Analysis

Analyze publicly social media data to gain insights into the **sentiments and needs** of people regarding the crisis, and trends within a country

Target

- Sentiment analysis
- Topic modeling
- Needs assessments

Partnership with:

Meta

QCRI
معهد قطر لبحوث الحوسبة
Qatar Computing Research Institute
جامعة حمد بن خليفة
HAMAD BIN KHALIFA UNIVERSITY

Data source

Twitter, Facebook, QCRI

Methodologies for analyses

- The causal inference procedure after a disaster is normally composed of three steps:

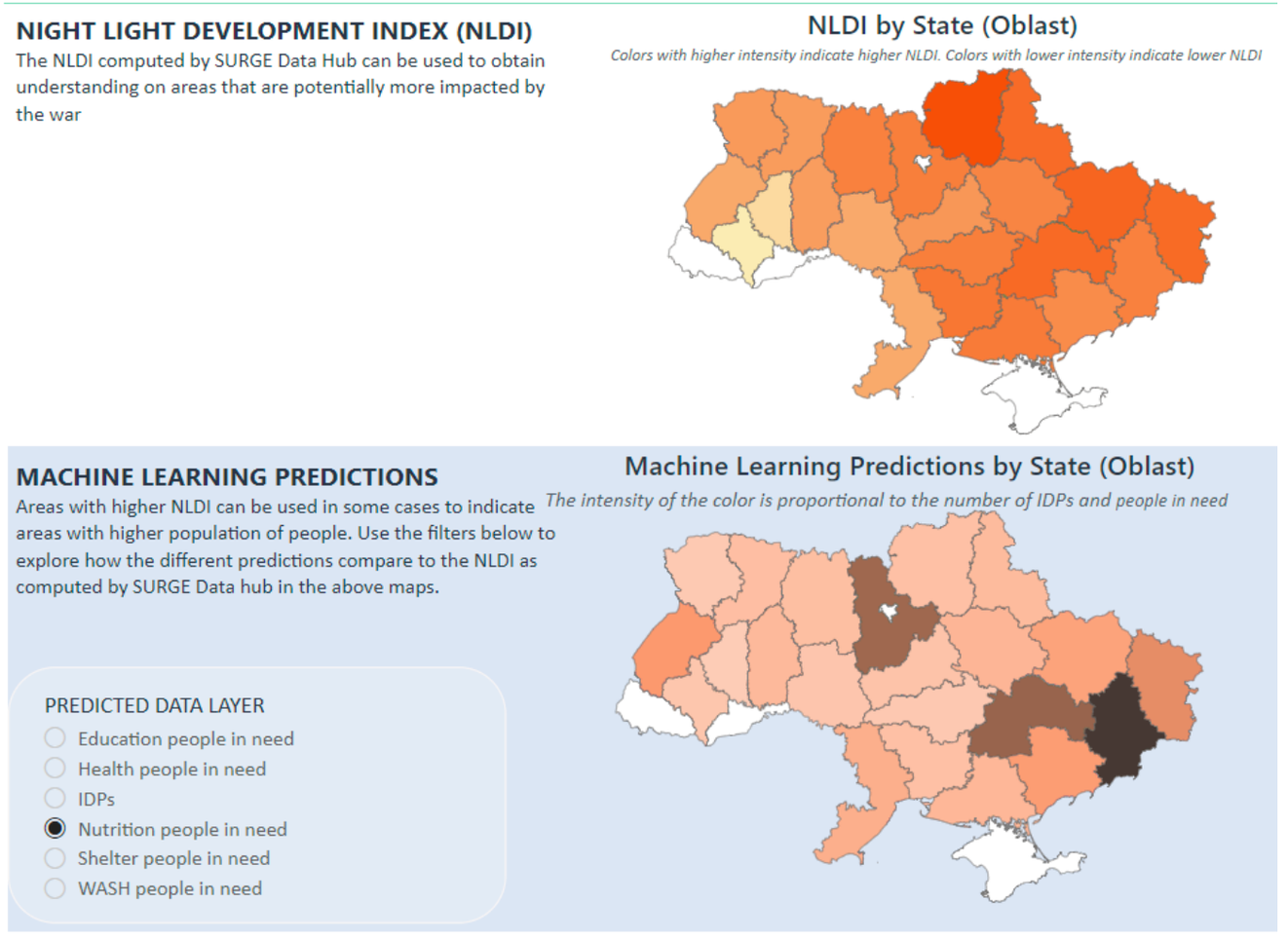
i) **Measure the impact** of the disaster on household/business i (treatment group), we first identify a similar HH/business j in another region that was not affected by the disaster (control group)

ii) **Estimate the counterfactual** ("what-if the disaster did not occur?") of i after the disaster using observed data from j

iii) **Quantify the impact** of the disaster by taking the difference between the predicted and observed impact in i.

- Studies using surveys have identified various factors that affect disasters impacts through **econometric models** (e.g. logit regression)
- However, the above **difference in differences** (DiD) method has limitations including: i) it requires that the differences between treatment and control group are invariant overtime, ii) only few points in time (i.e., pre-treatment time and post-treatment time) are normally considered. This can be implausible if the outcome of interest changes over time, such as recovery patterns after disaster
- **Bayesian time series** models, allow flexible predictions of time series data, which can be used to estimate the causal impact over time.

Machine Learning



Machine learning is at the crossroads of the previous methods. It can be used in all of them to offer more advanced analysis.
UNDP uses machine learning techniques to scan the effect of previous crises and estimate the impact of new crises on households. These techniques can be used to unpack household's humanitarian needs following a crisis.

Data source

UN OCHA, Data exchange platform, ACLED, WorldPop, NASA, etc.

Surge Data Hub

- **Objective:** Conduct assessments offering tools for decision making at global level, with particular focus on crisis response
- **Background:** Started based on **HBDA**, transformed to train, support and advice countries to implement Digital Assessments (**SEIA**), **Social Media Analysis, humanitarian needs using Machine Learning and Night Light Variability Analysis.**
- **Mandate and Approach:**
  - Primary and Secondary data analysis
  - MVI/MPI and other indexes, SDR integration, advance analysis
- **Platforms:** Kobo Toolbox, ArcGIS, MS Power BI, Python, Azure DB, etc...

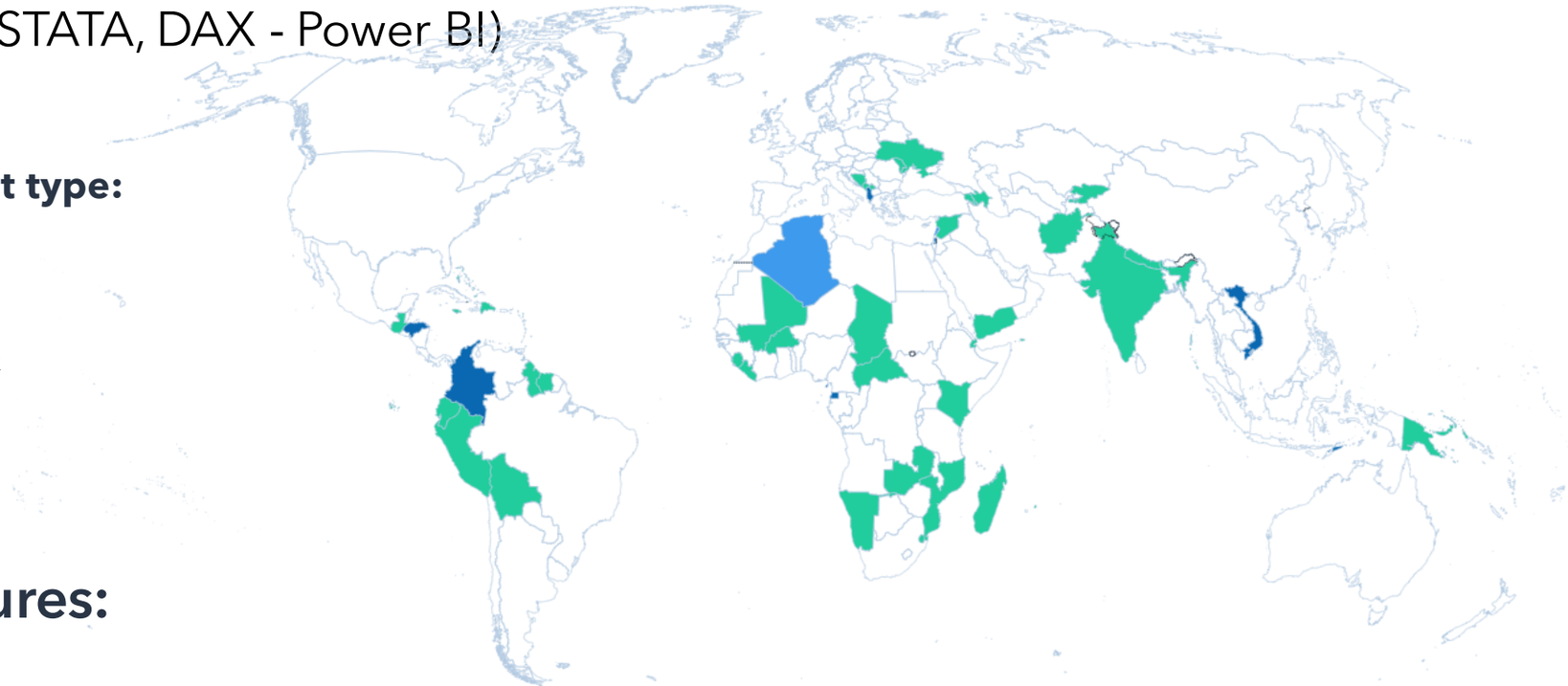
UNDP Training

- In the past 2 years we have trained Institutes/Bureaus of Statistics around the world.

- The trainings have covered:
  - data collection (Kobo Toolbox)
  - data analysis (MS Excel, SPSS/STATA, DAX - Power BI)
  - data visualization (Power BI).

Assessment type:

- HBDA
- SEIA
- SEIA & HBDA



Key figures:

45+
Assessments

180
SEIA graduates

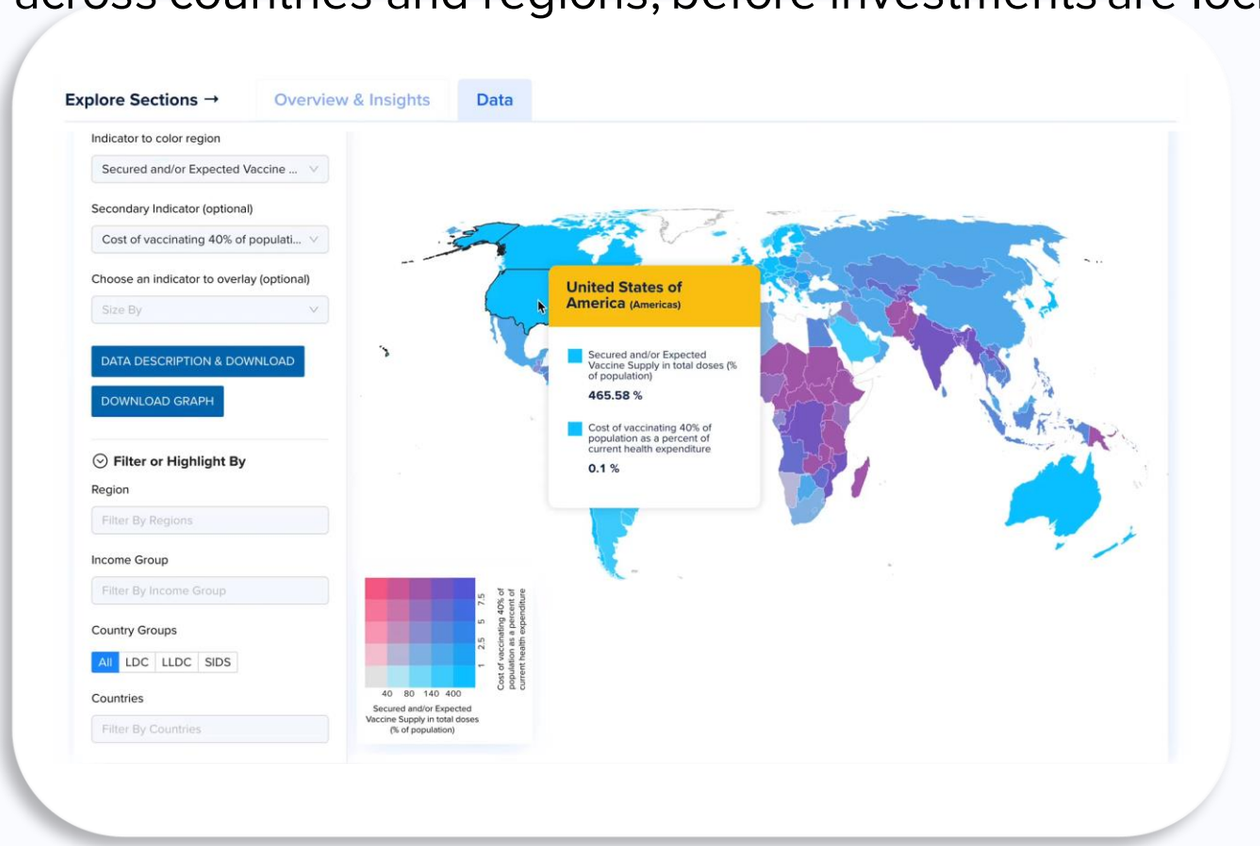
171
HBDA graduates

1K+
Community members

UNDP'S DATA FUTURES PLATFORM

An open resource that **translates multidimensional data into actionable insights** to support Governments and other development practitioners. It is a one-stop-shop for data and analytic tools that **uses problem diagnostics and the power of data** to help practitioners understand complex issues.

Operating across UNDP's work, the platform leverages interactive simulations, models, dashboards, and visualizations to support decision makers to analyze potential impacts of different combinations of policy choices, across countries and regions, before investments are locked in.



Data intelligence

Bring data together from various sources to conduct your own analysis.

Question intelligence

Explore correlation between indicators as well as experimentation and scenario assessment with country-level data using simulators.

Decision intelligence

Develop curated analyses to turn crises into an opportunity for transformation towards sustainable development in the long-term.

TOOLS, SERVICES AND SUPPORT AVAILABLE

DASHBOARDS

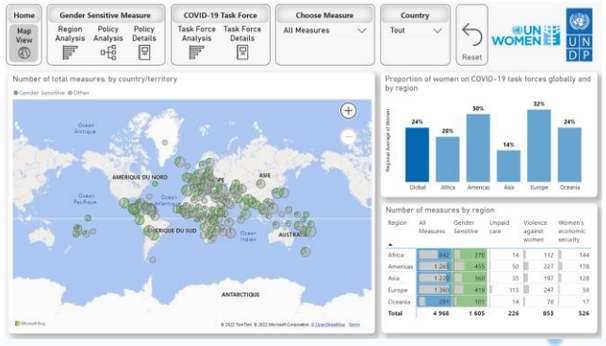
VACCINE EQUITY

Review **country and global obstacles to COVID-19 vaccination efforts** and data-driven insights and **analysis on the socioeconomic impacts** on vaccine inequity.

TRACKERS

GENDER EQUALITY OPPORTUNITY

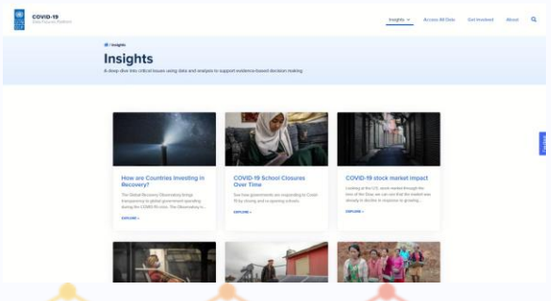
Tracking responses taken by governments worldwide from a gender lens. Analyze which of the policy measures address **women's economic and social security,** including unpaid care work, the labour market and violence against women.



SIMULATORS

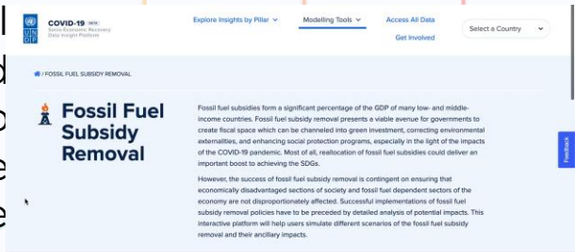
TEMPORARY BASIC INCOME

Unconditional emergency cash transfers can mitigate the immediate effects of the COVID-19 crisis on poor and near-poor households. Find the TBI amount needed to lift the vulnerable out of poverty for 132 countries, depending on policy choices.



FOSSIL FUEL SUBSIDIES

Discover how reallocating a small portion of fossil fuel subsidies could enable more students to attend school, increase electricity from renewable energies, or fund TBI for some of the world's poorest people.



To identify who is being left behind we need multidimensional data that helps us answer the following questions:

Understanding the gaps and inequity of vaccine access

Where are vaccines being administered?

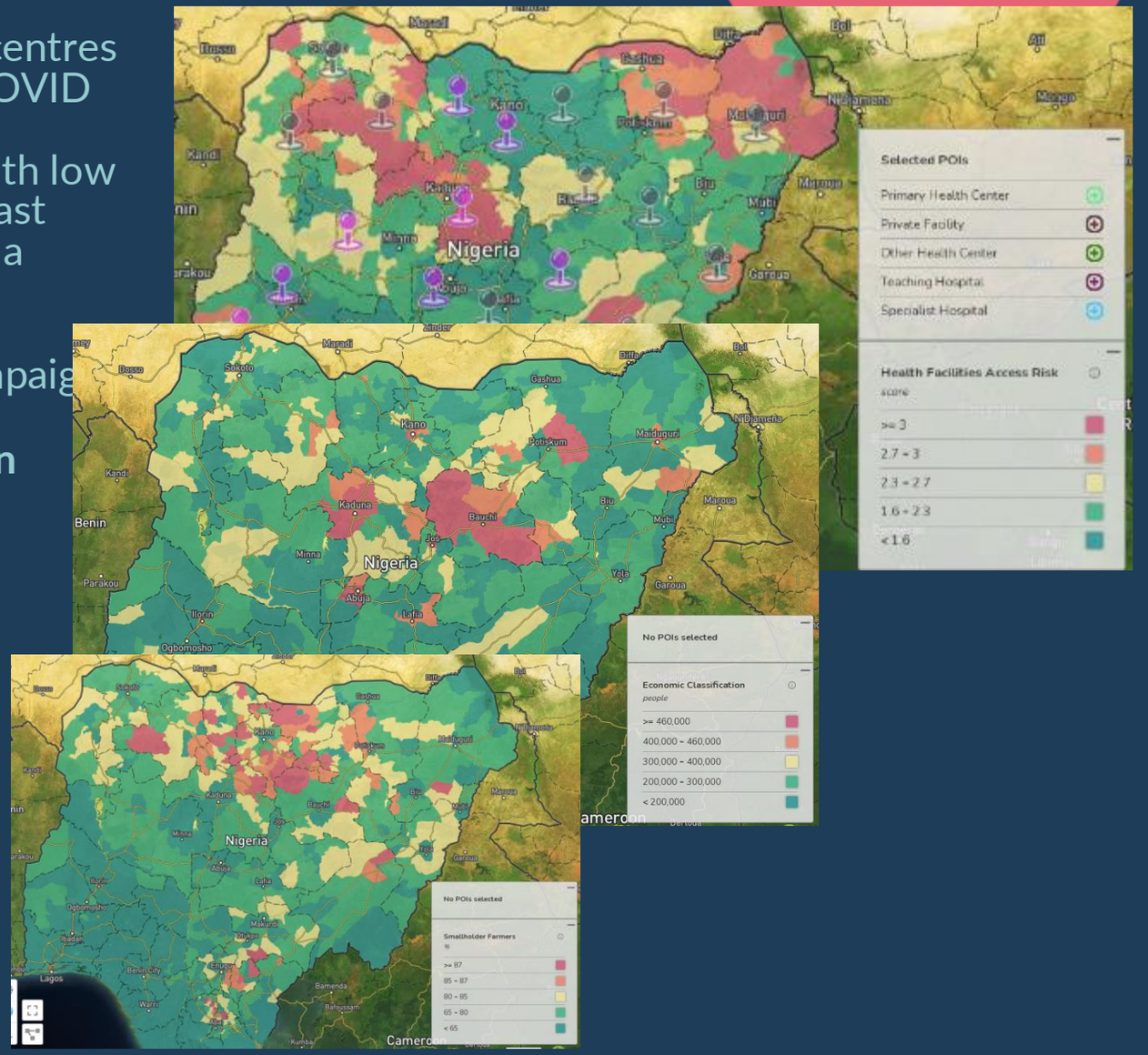
- Primary health centres providing the COVID vaccine
- Communities with low health access least likely to receive a vaccine
- Availability of information campaig

What are the socio-economic characteristics of those communities with low levels of vaccinations?

- **Economic classification** (asset wealth, income and education)
- **Vulnerability:** disadvantaged age groups; ethnic minorities

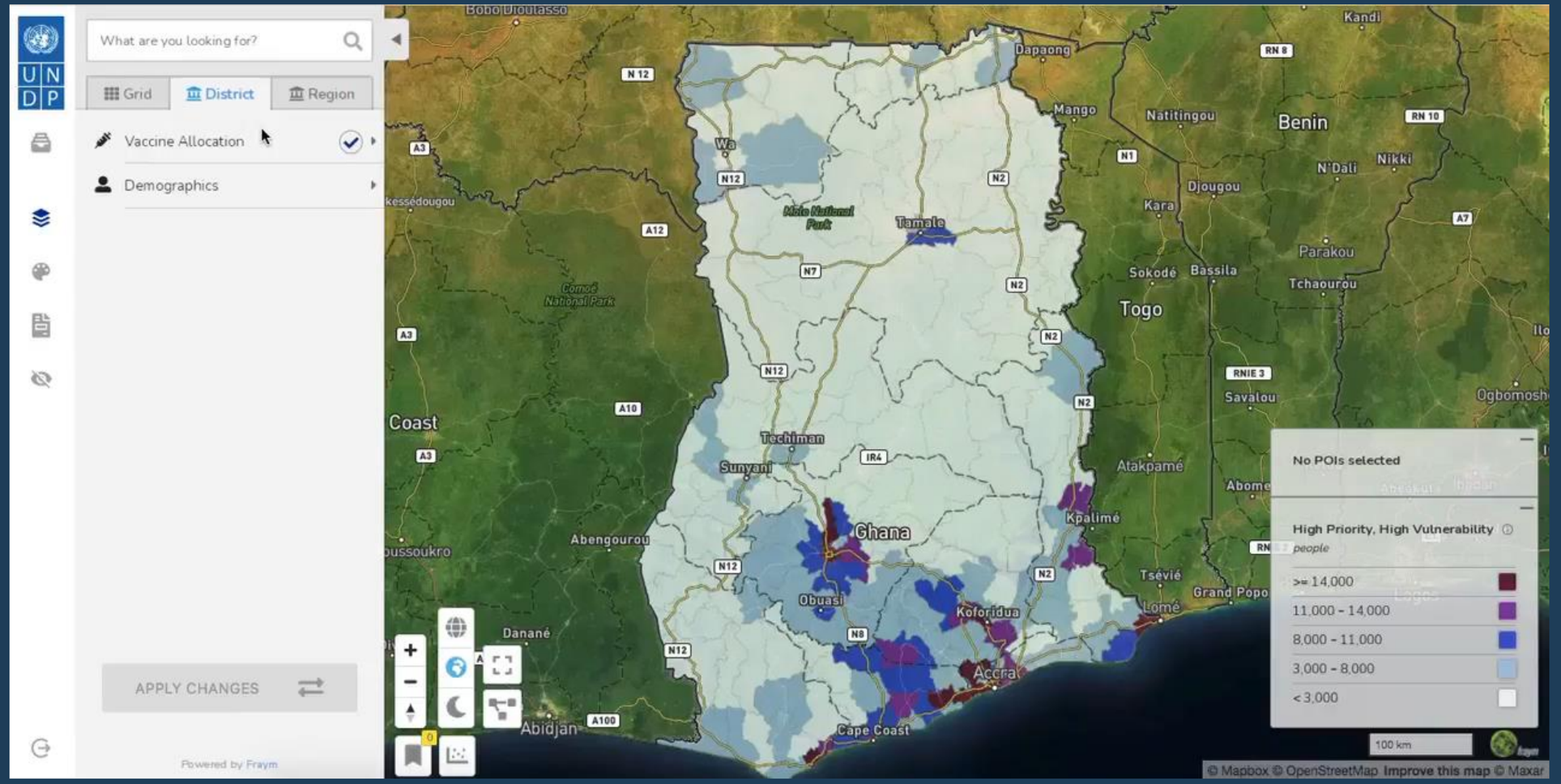
Socioeconomic impacts of COVID-19

- Changes in employment
- Changes in economic activity
- Occupations (i.e. smallholder farmers)



Microplanning - actionable tool and data for decision making

2. Using data for prioritization

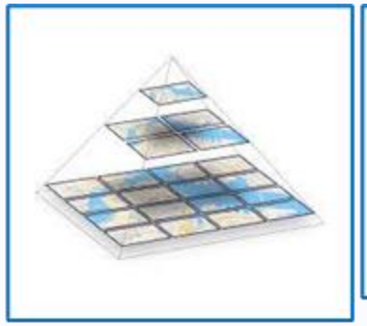


Bringing together this multiple indicators can be the basis for an easy-to-use decision-making tool that allows policymakers to identify at-risk groups.

It can serve as a foundation for logistics and awareness microplanning to effectively prioritise the populations at higher risk of being left behind.

DATA AGGREGATION
GEO HUB

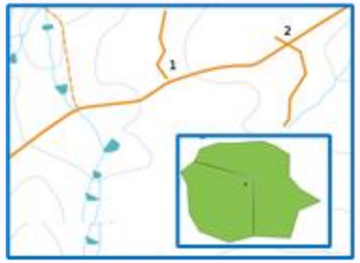
Empower the average user to harness the power of geospatial data integrated from disparate projects globally

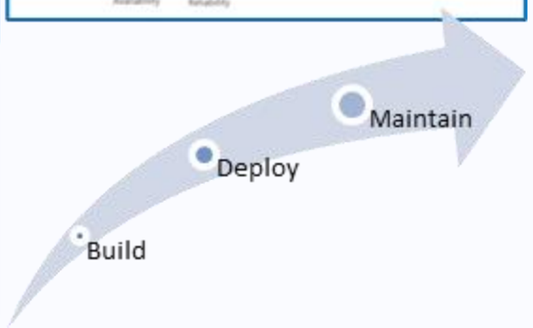


The hub will allow us to

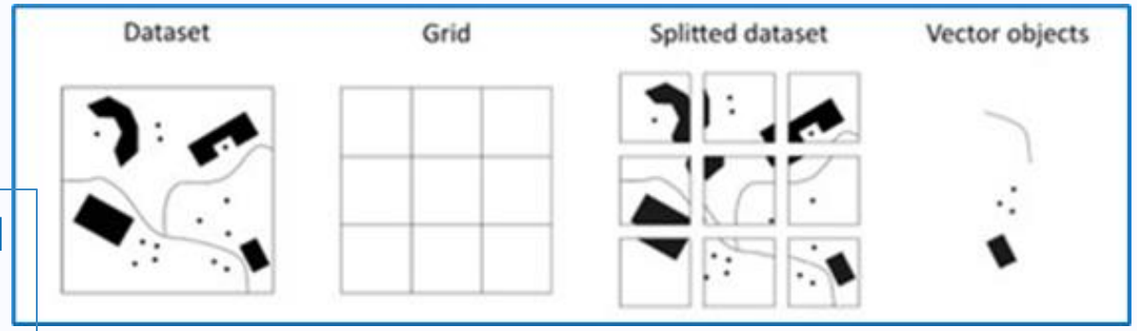
Bring all relevant geospatial data together in one centralized location

- Create simulation based scenarios
- Identify trends in space and time at subnational levels





- Users can merge datasets, assisted by built-in tools and programming languages. Users must still employ data cleaning, manipulation to integrate the datasets.
- Barriers to entry: (knowledge of the ArcGIS interface/basic GIS principles, and data cleaning for integration with other datasets).



- Example: the Hub built **dynamic global heatmap of electricity access** allowing policy makers to perform hot-spot analysis to identify areas that lack electricity access at the local level.

Thank you