15th Meeting of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators
2 - 5 September 2024
Oslo, Norway

The use of mobile phone data for SDG ICT Indicators

Esperanza Magpantay
Senior Statistician, ICT Data and Analytics Division, ITU
UN-CEBD Mobile Phone Task Team Lead

ITU Data & Analytics data collection mandate

Supply-side data

Administrative data from telecom operators
Source: Regulators/Ministries

ICT Price Basket
Source: Regulators/Ministries & operators' website

DataHub

ITU ICT Statistics

(annual questionnaires)

Demand-side data

ICT surveys
(annual questionnaires)
Source: National Statistical Offices

Overview - Information society indicators included in the SDG monitoring framework – collected by ITU

- Target 4.4: Proportion of youth/adults with ICT skills, by type of skills
- Target 5b: Proportion of individuals who own a mobile telephone, by sex
- **Target 9c: Percentage of the population covered by a mobile network, broken down by technology**
- Target 17.6: Fixed Internet broadband subscriptions, broken down by speed
- **Target 17.8: Proportion of individuals using the Internet**

ITU Mobile Phone Big Data work

1st pilot: 2016-2017

✓ 5 countries (Colombia, Georgia, Kenya, Philippines,
✓ 16 ICT indicators (administrative data)

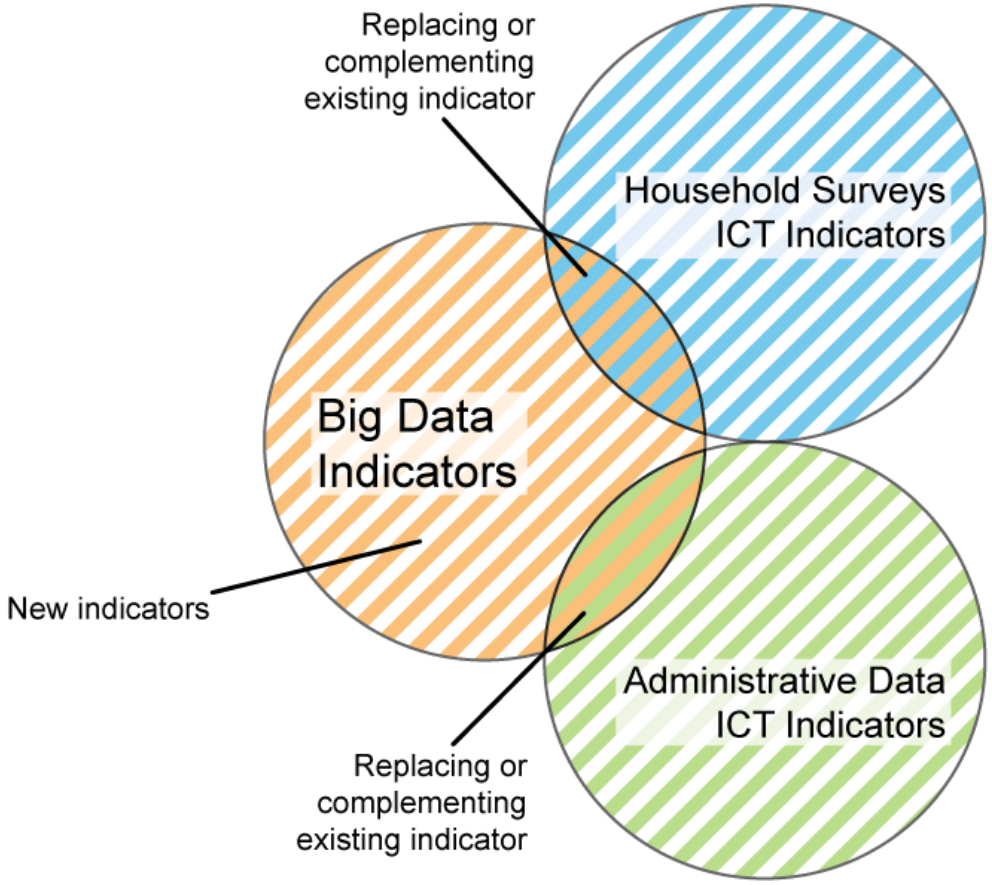
2nd pilot: 2020-2021

✓ Brazil, Indonesia

✓ 2 SDG ICT indicators
- ✓ 9.c.1 – Percentage of population covered by mobile network: 2G, 3G and 4G and above (administrative data)
- ✓ 17.8.1 – Percentage of population using the Internet (household survey data)

Ongoing: 2023-2024

Uganda, Malaysia, Mongolia, Liberia, Uruguay, Tunisia, Botswana
ITU/World Bank GDF-Mobile phone data for policy



UN-Committee of Experts on Big Data and Data Science for Official Statistics - Task Team on mobile phone data

Chair: ITU

6 areas of statistics:
-population, migration, tourism, information society, transport, disaster context

Objectives

Explore the use of mobile phone big data for the different areas of statistics and develop methodologies

Who

Composed of around 50 individual members/ 30 entities - international and regional agencies, countries, academia, private agencies/companies

Members

- Brazil
- Colombia
- Gambia
- Georgia
- India
- Indonesia
- Italy
- Japan
- Korea
- Malaysia

Members

- Mexico
- Netherlands
- Oman
- Qatar
- Philippines
- Romania
- Saudi Arabia
- United Arab Emirates
- Viet Nam

Members

- EU JRC
- Eurostat
- IMF
- IOM
- UNFPA
- UNGP
- UNSD
- UNESCWA
- World Bank
- OECD-ITF
- UN-ECE
- Flowminder
- GSMA
- Positium

UNBigData

ITU

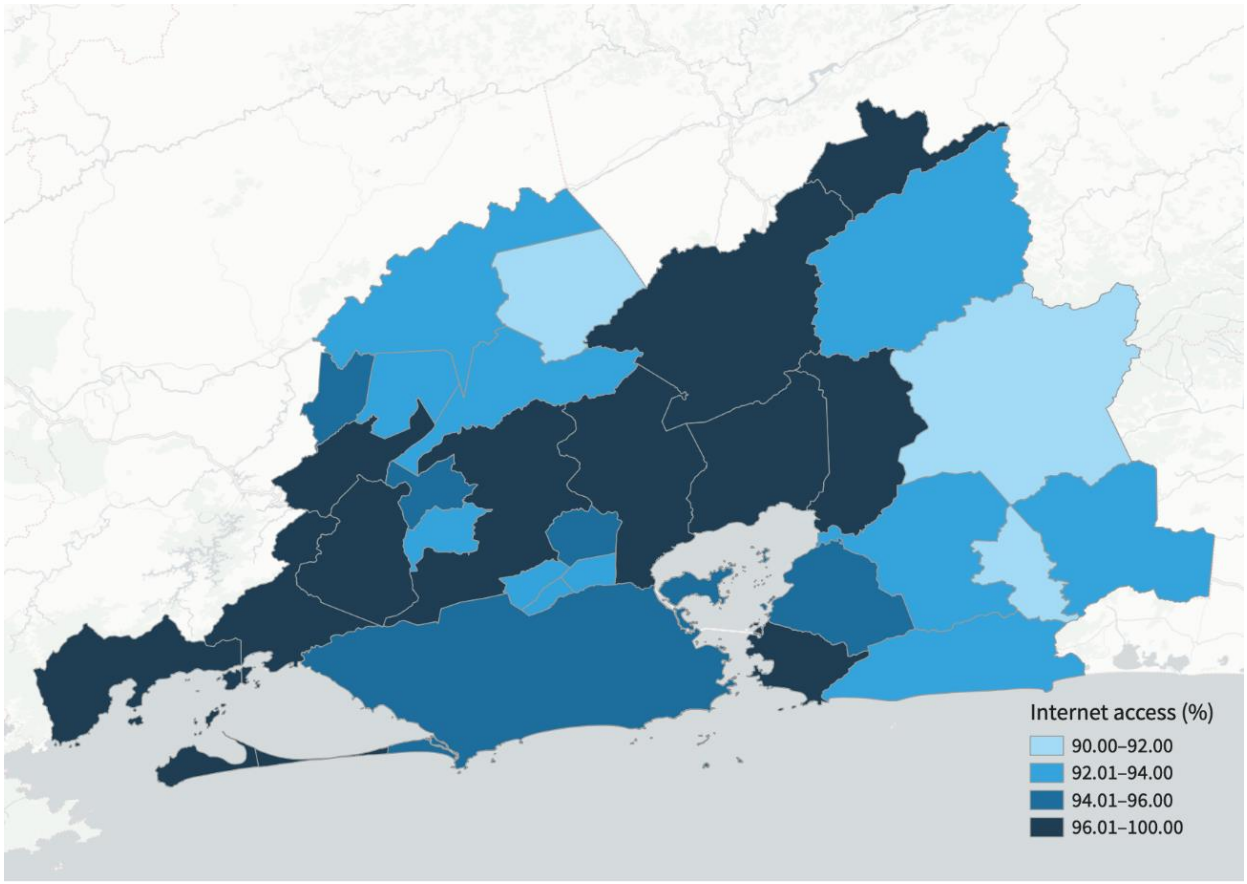
Example:
Guideline on Big Data for measuring the SDG Information society indicators (Lead: ITU)

1. Introduction
2. Background
3. Access and preparations
4. Data sources (description of mobile operator data, quality assurance of raw data)
5. Reference data (local admin units, world population, cell data, digital elevation, household survey data)
6. Data processing (models, data protection guidelines)
7. Calculating the indicators (rationale, definition, indicators calculation, quality assurance)
8. Quality assurance
9. Conclusions

- with experiences and examples from country pilots

Example: Information society indicators

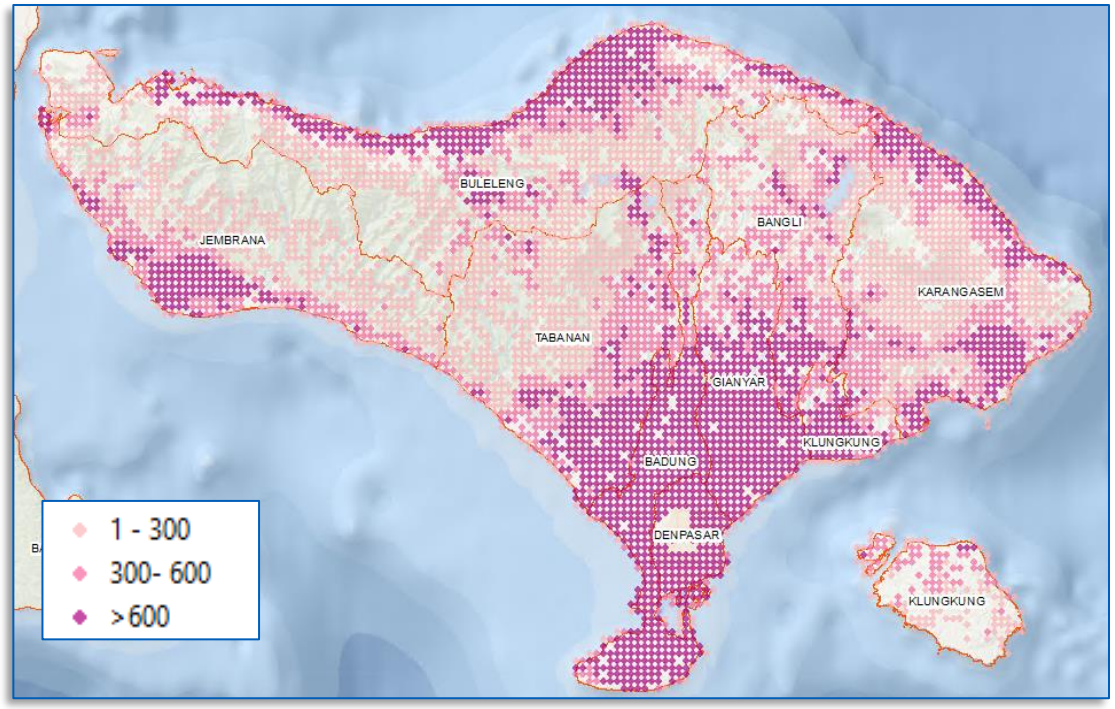
SDG indicator 17.8.1: Percentage of the population using the Internet, Rio de Janeiro, Brazil, 2021



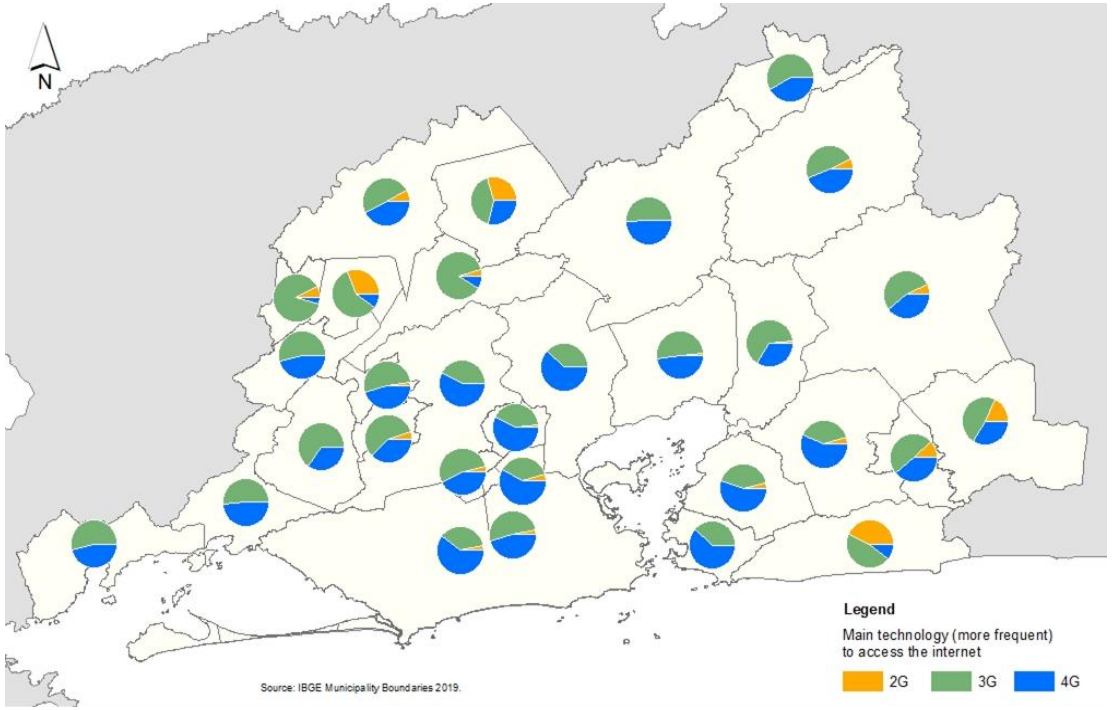
Mobile phone data in line with household survey results

Results: Percentage of population covered by mobile signal (SDG 9.c.1)

Indonesia



Brazil



ITU Jupyter Notebooks for information society

Field Name	Type	Mode	Description
msisdn	String		Hashed subscribers identifier
datetime	Timestamp		Transaction date (date and hour)
cell_id	String	NULLABLE	Hashed cell identifier
latitude	Float		Latitude of Base Transceiver Station (BTS)
longitude	Float		Longitude of Base Transceiver Station (BTS)
data_type	String		Data source, can be CDR/CHG or IPDR/UPCC
service	String		Transaction service (4G/ 3G/ 2G)

Field name	Type	Mode	Description
msisdn	String		Hashed subscribers identifier
age	Int		Subscribers age from registration data
gender	String		Subscribers gender (M/F) from registration data

```
# detect duplicate rows
df_duplicates = df.groupBy(df.columns).count().filter("count > 1")
print(f"number of duplicate rows: {df_duplicates.count()}")
df_duplicates.show()
```

```
number of duplicate rows: 539
[Stage 11:>                                                         (0 + 3) / 3]
+------+-------------------+-------+--------+---------+---------+-------+----------+-----+
|msisdn|           datetime|cell_id|latitude|longitude|data_type|service|      date|count|
+------+-------------------+-------+--------+---------+---------+-------+----------+-----+
|     8|2024-08-09 03:10:00|  164.0|  43.277|   -3.163|     IPDR|     3G|2024-08-09|    2|
|     8|2024-10-28 05:19:00|  164.0|  43.277|   -3.163|     IPDR|     3G|2024-10-28|    2|
```

```
# Print the number of records in the DataFrame
print("Number of records before deduplication: {}\n".format(df.count()))

# Drops the duplicate rows from the dataframe
df = df.dropDuplicates()

# Display the first five rows of the DataFrame in a tabular format
df.show(5)

# Print the number of records in the DataFrame
print("Number of records after deduplication: {}".format(df.count()))
```

```
Number of records before deduplication: 536851
+------+-------------------+-------+--------+---------+---------+-------+----------+
|msisdn|           datetime|cell_id|latitude|longitude|data_type|service|      date|
+------+-------------------+-------+--------+---------+---------+-------+----------+
|     0|2024-05-11 17:19:00|  419.0|  43.217|   -3.122|     IPDR|     3G|2024-05-11|
|     0|2024-05-19 03:01:00|  746.0|  43.327|   -3.089|     IPDR|     4G|2024-05-19|
|     0|2024-05-25 07:00:00|  873.0|  43.321|   -3.077|     IPDR|     3G|2024-05-25|
|     0|2024-05-29 11:13:00|  655.0|  43.327|     -3.1|     IPDR|     3G|2024-05-29|
|     0|2024-06-01 23:11:00|  786.0|  43.318|   -3.084|      CDR|     2G|2024-06-01|
+------+-------------------+-------+--------+---------+---------+-------+----------+
only showing top 5 rows

[Stage 20:==================>                                       (1 + 2) / 3]
Number of records after deduplication: 536310
```

Additional resources

- UN Big data task team on mobile phone data
- ITU Big Data pilots
- Online training course on mobile phone data

Or contact us at: indicators@itu.int

Thank you!

indicators@itu.int