The Benefits of Coordinating and Integrating Statistical and Geospatial Data within the Framework of the 2030 Agenda

Fifth Plenary Meeting of UN-GGIM Arab States

Tim Trainor
Former Co-Chair, UN-GGIM
The Value of Geospatial and Statistical Data for Data Integration

- Geospatial data provides basic geography to collect and make available statistical information. It is a geographic framework.
- Statistical data provides numbers and values for a specific geographic area on topics that include society (population), economy, environment, etc.
The Value of Data Integration

• On their own, geospatial and statistical data have value.
• When geography and statistics are joined, much more information becomes available.
• The integration can be viewed through maps and graphics.
• Trends, relationships, clusters and other observations are made possible.
Example Data Characteristics

• What is the resolution?
  – G. in metres…
  – S. by level of data collection?…
    • Enumeration area
    • Household location

• What is the accuracy – how dependable is the data?
  – G. plus or minus 8 metres for road centerlines…
  – S. 98+% accuracy based on post enumeration
Geospatial Data Sources

• Government
  – NMAs
  – City government
  – Other

• Commercial
  – Profit motive (cost)
  – Variable coverage

• Volunteer efforts
  – Variable coverage and quality
Geospatial Data Types

• Digital sources with attribution
• Remotely sensed data
  – Satellite
  – Photography
  – LiDAR
• Specialized
  – Infrastructure geospatial data below ground
    • Pipes, drains, wires
Technology

• GIS
• Mobile
• Sensors
• More satellites
• Drones
• Etc.
Need for Geospatial Requirements for Cities

• Importance of the role of location relative to attribution and statistics in measuring

• Knowing “where” leads to follow-on questions such as how much or how often or in what circumstances
  – Are there patterns of occurrence or is this an isolated instance

• Determining what level of geography is needed for effective knowledge and action

• Integrating different data types adds new dimensions and meaning

• Discovering geospatial data gaps and taking corrective steps increases the value of statistical data
Evolving Applications

• Public safety and emergency response
• Autonomous transportation
  – Including home and business delivery options
• Gaging stations to monitor:
  – Water levels
  – Environmental factors such as air quality
• Etc…
What is information is helpful in managing data?

- Definition
  - Need same understanding of terms, meaning, and usage

- Data
  - Source – who or where does the data come from?
    - National and local governments?...private sector?...other?...
    - Is it readily available or does it require a new partnership?
  - Complete or partial coverage?
  - Is it “good enough” data?

- Source

- Methodology

- Process and procedure
An example for data: The sanity check…

• What is the ideal state?
  – Full coverage of available, current and maintained, high quality, well-documented data at the needed level of geography

• What is the preferred state?
• What is minimally acceptable?
• What can be salvaged?
• What is not helpful?
Examples of Statistical Data used in Geospatial Analysis

- Census population and housing data files
- Patient files
- Soil data
- Employment/unemployment
- Animal sighting
- Cancer registry
- Police records
- School records (assignments, attendance)
### Statistical data can be overwhelming

<table>
<thead>
<tr>
<th>Income of Household</th>
<th>118,682</th>
<th>67,530</th>
<th>97,574</th>
<th>70,422</th>
<th>96,144</th>
<th>70,572</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under $5,000</td>
<td>4,176</td>
<td>1,249</td>
<td>2,741</td>
<td>1,243</td>
<td>2,665</td>
<td>1,246</td>
</tr>
<tr>
<td>$5,000 to $9,999</td>
<td>6,065</td>
<td>7,927</td>
<td>3,490</td>
<td>7,932</td>
<td>3,417</td>
<td>7,930</td>
</tr>
<tr>
<td>$10,000 to $14,999</td>
<td>7,081</td>
<td>12,388</td>
<td>5,397</td>
<td>12,435</td>
<td>5,309</td>
<td>12,440</td>
</tr>
<tr>
<td>$15,000 to $19,999</td>
<td>7,260</td>
<td>17,278</td>
<td>5,858</td>
<td>17,285</td>
<td>5,763</td>
<td>17,289</td>
</tr>
<tr>
<td>$20,000 to $24,999</td>
<td>6,937</td>
<td>22,165</td>
<td>5,477</td>
<td>22,186</td>
<td>5,390</td>
<td>22,181</td>
</tr>
<tr>
<td>$25,000 to $29,999</td>
<td>6,730</td>
<td>27,186</td>
<td>5,459</td>
<td>27,222</td>
<td>5,361</td>
<td>27,218</td>
</tr>
<tr>
<td>$30,000 to $34,999</td>
<td>6,148</td>
<td>32,085</td>
<td>4,990</td>
<td>32,054</td>
<td>4,910</td>
<td>32,103</td>
</tr>
<tr>
<td>$35,000 to $39,999</td>
<td>5,907</td>
<td>37,183</td>
<td>4,910</td>
<td>37,208</td>
<td>4,934</td>
<td>37,200</td>
</tr>
<tr>
<td>$40,000 to $44,999</td>
<td>5,624</td>
<td>42,013</td>
<td>4,623</td>
<td>42,047</td>
<td>4,537</td>
<td>42,049</td>
</tr>
<tr>
<td>$45,000 to $49,999</td>
<td>4,933</td>
<td>47,198</td>
<td>4,081</td>
<td>47,205</td>
<td>4,007</td>
<td>47,200</td>
</tr>
<tr>
<td>$50,000 to $54,999</td>
<td>5,088</td>
<td>51,984</td>
<td>4,222</td>
<td>52,016</td>
<td>4,175</td>
<td>52,012</td>
</tr>
<tr>
<td>$55,000 to $59,999</td>
<td>4,203</td>
<td>57,154</td>
<td>3,580</td>
<td>57,168</td>
<td>3,535</td>
<td>57,161</td>
</tr>
<tr>
<td>$60,000 to $64,999</td>
<td>4,412</td>
<td>61,941</td>
<td>3,696</td>
<td>61,935</td>
<td>3,654</td>
<td>61,935</td>
</tr>
<tr>
<td>$65,000 to $69,999</td>
<td>3,579</td>
<td>67,095</td>
<td>3,021</td>
<td>67,109</td>
<td>2,981</td>
<td>67,103</td>
</tr>
<tr>
<td>$70,000 to $74,999</td>
<td>3,769</td>
<td>72,042</td>
<td>3,166</td>
<td>72,064</td>
<td>3,134</td>
<td>72,062</td>
</tr>
<tr>
<td>$75,000 to $79,999</td>
<td>3,118</td>
<td>77,007</td>
<td>2,656</td>
<td>77,023</td>
<td>2,622</td>
<td>77,020</td>
</tr>
<tr>
<td>$80,000 to $84,999</td>
<td>3,143</td>
<td>81,979</td>
<td>2,709</td>
<td>82,015</td>
<td>2,671</td>
<td>82,016</td>
</tr>
<tr>
<td>$85,000 to $89,999</td>
<td>2,860</td>
<td>87,142</td>
<td>2,339</td>
<td>87,141</td>
<td>2,303</td>
<td>87,136</td>
</tr>
<tr>
<td>$90,000 to $94,999</td>
<td>2,516</td>
<td>92,009</td>
<td>2,187</td>
<td>91,954</td>
<td>2,162</td>
<td>91,999</td>
</tr>
<tr>
<td>$95,000 to $99,999</td>
<td>2,110</td>
<td>97,155</td>
<td>1,867</td>
<td>97,165</td>
<td>1,840</td>
<td>97,163</td>
</tr>
<tr>
<td>$100,000 to $104,999</td>
<td>2,498</td>
<td>101,830</td>
<td>2,186</td>
<td>101,875</td>
<td>2,170</td>
<td>101,872</td>
</tr>
<tr>
<td>$105,000 to $109,999</td>
<td>1,778</td>
<td>107,162</td>
<td>1,523</td>
<td>107,145</td>
<td>1,504</td>
<td>107,141</td>
</tr>
<tr>
<td>$110,000 to $114,999</td>
<td>1,782</td>
<td>111,973</td>
<td>1,541</td>
<td>111,997</td>
<td>1,529</td>
<td>111,993</td>
</tr>
<tr>
<td>$115,000 to $119,999</td>
<td>1,480</td>
<td>117,204</td>
<td>1,271</td>
<td>117,228</td>
<td>1,255</td>
<td>117,223</td>
</tr>
<tr>
<td>$120,000 to $124,999</td>
<td>1,470</td>
<td>121,842</td>
<td>1,263</td>
<td>121,860</td>
<td>1,246</td>
<td>121,851</td>
</tr>
<tr>
<td>$125,000 to $129,999</td>
<td>1,243</td>
<td>127,026</td>
<td>1,105</td>
<td>127,030</td>
<td>1,082</td>
<td>127,028</td>
</tr>
<tr>
<td>$130,000 to $134,999</td>
<td>1,236</td>
<td>132,066</td>
<td>1,074</td>
<td>132,127</td>
<td>1,061</td>
<td>132,122</td>
</tr>
<tr>
<td>$135,000 to $139,999</td>
<td>1,050</td>
<td>137,202</td>
<td>937</td>
<td>137,218</td>
<td>921</td>
<td>137,226</td>
</tr>
<tr>
<td>$140,000 to $144,999</td>
<td>974</td>
<td>141,989</td>
<td>828</td>
<td>142,055</td>
<td>823</td>
<td>142,058</td>
</tr>
<tr>
<td>$145,000 to $149,999</td>
<td>763</td>
<td>146,959</td>
<td>664</td>
<td>146,955</td>
<td>668</td>
<td>146,955</td>
</tr>
<tr>
<td>$150,000 to $154,999</td>
<td>983</td>
<td>151,788</td>
<td>849</td>
<td>151,794</td>
<td>843</td>
<td>151,792</td>
</tr>
<tr>
<td>$155,000 to $159,999</td>
<td>671</td>
<td>157,072</td>
<td>590</td>
<td>157,028</td>
<td>583</td>
<td>157,020</td>
</tr>
<tr>
<td>$160,000 to $164,999</td>
<td>636</td>
<td>162,151</td>
<td>537</td>
<td>162,125</td>
<td>530</td>
<td>162,121</td>
</tr>
<tr>
<td>$165,000 to $169,999</td>
<td>541</td>
<td>167,076</td>
<td>474</td>
<td>167,064</td>
<td>464</td>
<td>167,073</td>
</tr>
<tr>
<td>$170,000 to $174,999</td>
<td>655</td>
<td>171,948</td>
<td>466</td>
<td>171,981</td>
<td>461</td>
<td>171,986</td>
</tr>
</tbody>
</table>
Statistical Data

Spatial Data

<table>
<thead>
<tr>
<th>County</th>
<th>2000 Population</th>
<th>2010 Population</th>
<th>Percent Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altoona</td>
<td>3907</td>
<td>1935</td>
<td>-49.8</td>
</tr>
<tr>
<td>Athens</td>
<td>2042</td>
<td>2077</td>
<td>1.7</td>
</tr>
<tr>
<td>Athens county</td>
<td>4000</td>
<td>3900</td>
<td>-2.5</td>
</tr>
<tr>
<td>Athens county</td>
<td>12083</td>
<td>15800</td>
<td>31.4</td>
</tr>
<tr>
<td>Athens county</td>
<td>501</td>
<td>1450</td>
<td>190</td>
</tr>
<tr>
<td>Athens county</td>
<td>1245</td>
<td>1450</td>
<td>190</td>
</tr>
<tr>
<td>Athens county</td>
<td>22619</td>
<td>30313</td>
<td>34.4</td>
</tr>
<tr>
<td>Athens county</td>
<td>17347</td>
<td>67011</td>
<td>292%</td>
</tr>
<tr>
<td>Athens county</td>
<td>15300</td>
<td>20534</td>
<td>34.4</td>
</tr>
<tr>
<td>Athens county</td>
<td>3952</td>
<td>2873</td>
<td>-27.4</td>
</tr>
<tr>
<td>Athens county</td>
<td>2567</td>
<td>1016</td>
<td>-60.7</td>
</tr>
<tr>
<td>Athens county</td>
<td>639</td>
<td>146</td>
<td>-77.6</td>
</tr>
<tr>
<td>Athens county</td>
<td>594</td>
<td>430</td>
<td>-27.5</td>
</tr>
<tr>
<td>Athens county</td>
<td>8704</td>
<td>470</td>
<td>-94.9</td>
</tr>
<tr>
<td>Athens county</td>
<td>354</td>
<td>146</td>
<td>-59.1</td>
</tr>
<tr>
<td>Athens county</td>
<td>1076</td>
<td>1076</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Ohio - 2010 Census Results

Percent Change in Population by County: 2000 to 2010

2010 Census: Ohio Profile

Population Density by Census Tract

United Nations Secretariat
Global Geospatial Information Management

ggim.un.org
Geography adds value to data

- Connects statistics to geographic areas
- Reveals patterns, relationships and trends
- Simplifies big data
- Generates hypotheses and questions
- Turns data into information
- Tells a story
Assumptions

• Geospatial data is core to the 2030 Sustainable Development Agenda
• Statistics are the facts that measure compliance to the indicator framework
• Location information offers perspective, greater understanding and a view of the data through a geographic lens
• Geospatial data complements statistical information by telling a story that supports planning, programs, and decision-making
Statistical Geospatial Framework

- Accessible & usable
- Interoperable data & metadata standards
- Common geographies for dissemination of statistics
- Geocoded unit record data in a data management environment
- Use of fundamental geospatial infrastructure and geocoding
Establishing a Geospatial Framework for Statistical Data

- What geographic data are needed?
- What level of accuracy is required?
- What is the timeframe?
- How frequently are the data utilized?
- What geospatial technologies are available?
- What are the benefits and costs?
- Who are the stakeholders?
Grid-based data (global to local) vs Geographic Areas

10,000 km window/100 km grids (Global scale)

1,000 km window/10 km grids (International regions)

100 km window/1 km grids (National regions)

10 km window/100m grids (Urban Districts)

1 km window/10 m grids (Urban neighborhoods)

100 m window/1 m grids (Urban blocks)

AIANNH Areas (American Indian, Alaska Native, Native Hawaiian Areas)

United Nations Initiative on Global Geospatial Information Management
Data collection is tabulated by Administrative Areas at varying levels, namely:

- Global
- Regional
- National
- Subnational
- Cities
Geographic Areas and Boundaries

Cities and Human Settlements:
Example enumeration boundaries for places in the US
- Counties
- Census tracts
- Census blocks
- Census Designated Places
- Minor Civil Divisions/Towns

Public ownership and use:
- Parcel and land records
So where are these urban areas?
Observations

• No international agreement or practice on urban/rural by NSOs

• Urban/rural serves different purposes
  – Some intentional
    • Economic development (Urbanized Areas to Metropolitan Areas)
  – Some unintentional
    • Program implementation via laws
      – Rural health care and housing
    • Normally no control but sometimes has unintended consequences (“it’s because of the Census Bureau…”)

United Nations Initiative on Global Geospatial Information Management

Un-Ggim un.org
Realities

• More than one type of use of urban/rural designations
  – Functional use – population based
  – Physical observable - for example land use planning

• Challenge with urban/rural applies to both developed and developing countries

• The impact of this exercise is not limited to supporting the SDGs: it enables capacity development.
Considerations

• Temporal issues are important
  – Development and movement of population occurs over time
    • Increase and decrease based on events and conditions
  – How are temporal data accounted?

• How to react to special circumstances and anomalies within a Member State
  – The U.S.
    • American Indian Areas
    • Colonias
Cities and Smart Use of Data
Houston and Hurricane Harvey
Is there a plan?

Top 10 metro areas most affected by storm surge risk
Number of homes at risk of hurricane damage

1. Naples: 183,090
2. New York City: 723,183
3. Bradenton: 254,900
4. New Orleans: 391,004
5. Virginia Beach: 388,349
6. Cape Coral: 313,955
7. Houston: 283,380
8. Jacksonville: 171,189
9. Tampa: 459,275
10. Miami: 784,773

Source: CoreLogic
What data exists?

We have building locations

We have a water inundation index

Potential Storm Surge Flooding*

- Inter tidal Zone/Estuarine Wetland
- Greater than 1 foot above ground
- Greater than 3 feet above ground
- Greater than 6 feet above ground
- Greater than 9 feet above ground
- Leveed area
- Consult local officials for flood risk

*Displayed flooding values indicate the water height that has about a 1-in-10 (10%) chance of being exceeded.
What capabilities are available?

We have digital elevation models.

We have a local government flood mapping tool.
Planning for different culprits in a disaster

Heavy deluge of rain

Exceeding infrastructure storm drain capacity
The aftermath and recovery – are we ready?
A quick review…and some simple steps

- Re-evaluate building codes and zoning
- Before rebuilding, use existing data
  - Use DEMs and 3DEP combined with water inundation index to determine likelihood of flooding for each building
  - Determine rate of storm surge from storm drain capacity
  - Measure the flood probability and apply flood insurance rate proportionately
Cities and Example SDGs
By 2020, halve the number of global deaths and injuries due to traffic accidents…

Total numbers for a nation are telling for a national perspective.

To take action requires more information:
- Operator error (drunken driving, seat belt use…)
- Road conditions (sharp curves, pot holes…)
- Traffic safety aids (speed signs, traffic lights…)

Knowing location of traffic events is required for next steps.
…Reforms giving women equal rights to economic resources as well as access to ownership and control over land…

How does geospatial data contribute?

Land parcel and cadastre records are needed
Parcel size and extent (boundary)
Land use (agriculture, residential, economic)
Ownership by characteristics including sex
…building resilient infrastructure, promoting inclusive and sustainable industrialization and fostering innovation…

“What proportion of rural population who live within 2 km of an all-season road”

What are the geospatial implications?
How to differentiate rural and urban populations?
  Accepted definitions are needed
Locations of housing units
Existence of a geospatial detailed maintained road network
Goal 11: Make cities and human settlements inclusive, safe, resilient, and sustainable.

Target 11.7: By 2030, provide universal access to safe, inclusive and accessible, green and public spaces, particularly for women and children, older persons and persons with disabilities.

Indicator: The average share of the built-up areas of cities in open space in public ownership and use.
Goal 11: Getting Started

- **An assumption** – there is public accessibility to diverse datasets and GIS tools on a national and global scale

- **An urban geography framework** - small area geography to merge target group statistics, and land use/land classification data

- **Diverse datasets** – demographic data, earth observation data, crime data, land use/land classification data, open areas, or protected areas

- **A geospatial methodology** - an integrated solution to geospatial problem; see previous studies listed below

---


Case Study for Goal 11:
Pittsburgh, PA

Data Sources:

- **Base layers** (Boundaries/roads/DEMs)
- **Target Population Data** (Population, age/sex, crime statistics)
- **Accessible Open Space Layer** ((Protected Areas Database), Open Street Map, Parcel data)
- **National Land Cover Database (NLCD)**
- **Additional Gridded Datasets**
  - Landscan - Oak Ridge National Labs
  - CIESIN/SEDAC - NASA/Columbia University
Data Integration Model: A GIS Solution for Goal 11

- Assess data quality and select appropriate **small area geography**; e.g., block group, census tract, or gridded polygons;

- Extract **access points** to open space or protected areas; e.g. parks, recreation areas;

- Link **target population data** to small area geography or gridded polygons; e.g. demographic, economic, health, crime statistics;

- Create “**isochrone/isodistance**” maps (time/distance to access points);

- Develop a “**proximity index**” for each city (weighted (average) time and/or distance to the areas of interest) to allow comparison to other cities.
A Path Forward

• Close collaboration is needed by:
  – National Mapping and Geospatial Agencies
  – National Statistical Organizations
  – Those who have begun this process are realizing tangible successful outcomes

• 2020 Round of Censuses benefit from these collaborations where new data and new data types are identified and planned for in support of the SDGs

• New synergies across organizational boundaries have longer term benefits beyond meeting the SDGs
Conclusion

• Geospatial data is core to the 2030 Sustainable Development Agenda
• Statistics are the facts that measure compliance to the indicator framework
• Location information offers perspective, greater understanding and a view of the data through a geographic lens
• Geospatial data complements statistical information by telling a story that supports planning, programs, and decision-making