Regional Workshop on Statistical Business Registers and Industrial Classifications

Session 9 - Integration of geospatial information in SBRs

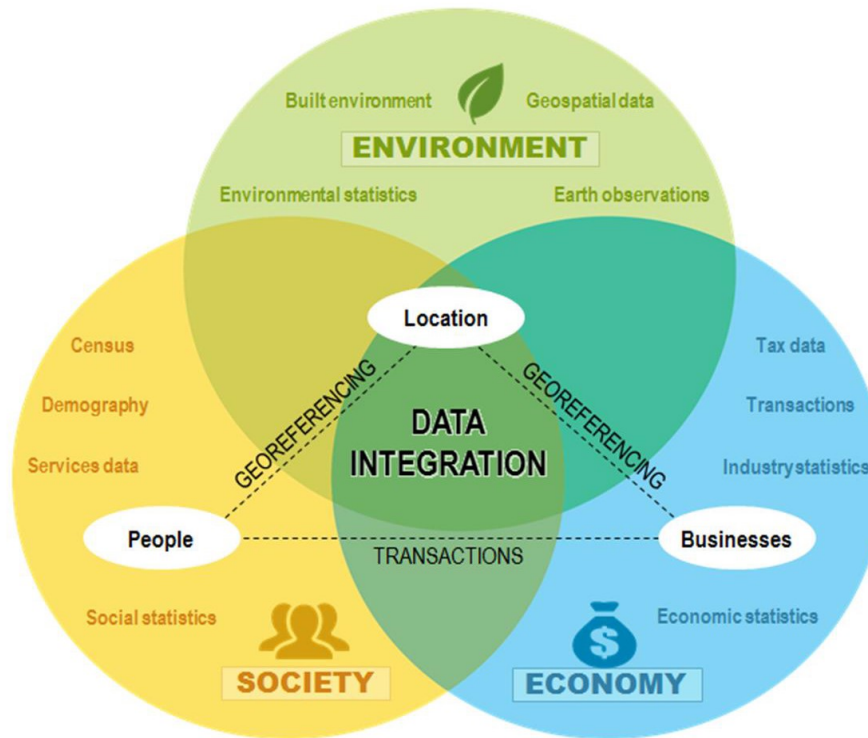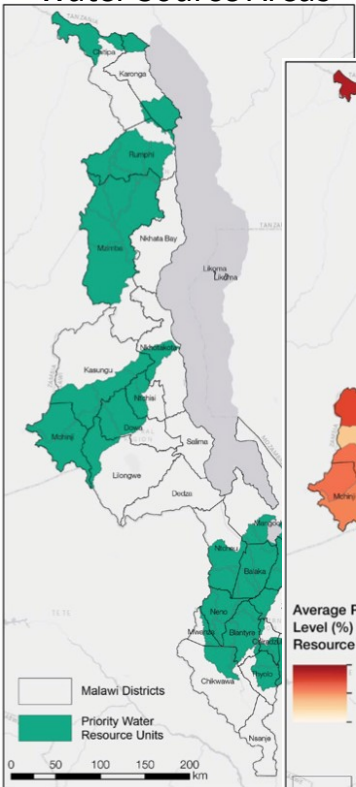Julian Chow, UNSD

Amman, Jordan, 25 – 27 June 2024

Location as a link between society, the economy and the environment

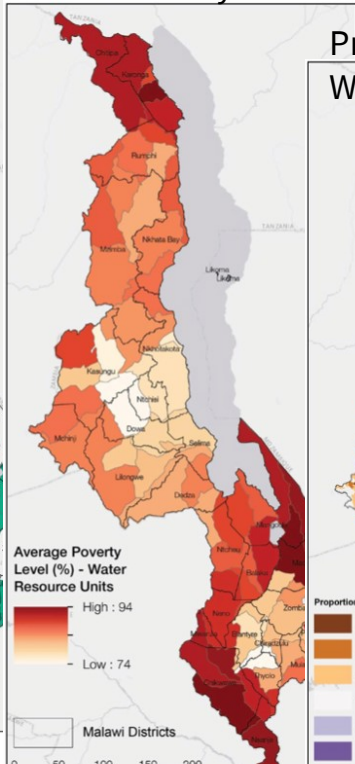# Spatial integration of ecosystem accounts with social indicators
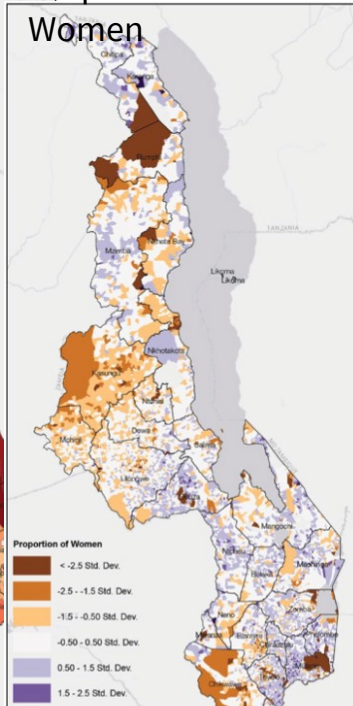
Water Source Areas

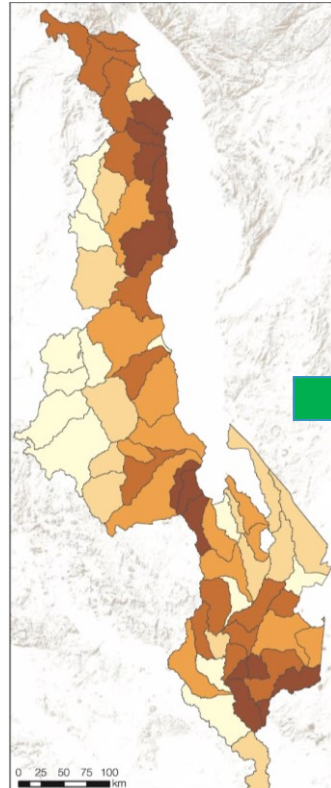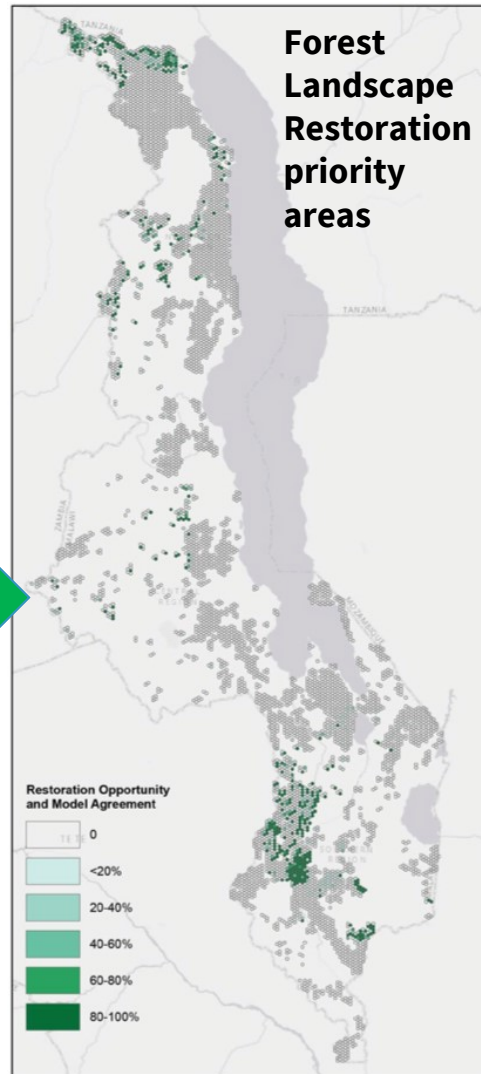**Poverty**

Proportion of Women

Ecosystem Services Benefits

Forest Landscape Restoration priority areas

Beatty et al., 2018. *Landscapes, at your service: Applications of the Restoration Opportunities Optimization Tool (ROOT)*. IUCN.

Background

- The United Nations Statistical Commission, at its 53rd session in March 2022

- Requested the Committee of Experts on Business and Trade Statistics (UNCEBTS) to develop guidelines to assist countries to integrate geospatial information into SBRs building on country experiences

- The TT-SBR included in its work plan the preparation of a report on the "Integration of geospatial information in SBRs" to be submitted to the next UNSC in March 2024

- The Global Statistical Geospatial Framework (GSGF), developed by the UN Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI), was endorsed in 2020 by the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) and the UNSC

- Close collaboration between the TT-SBR and EG-ISGI to ensure that the report is in full alignment with the GSGF

Global Statistical Geospatial Framework (GSGF)

- The GSGF provides the <u>framework to facilitate the integration of statistical and geospatial information</u>, between NSOs and National Geospatial Information Agencies (NGIAs)

- Developed by the UN Expert Group on the Integration of Statistical and Geospatial Information (EG-ISGI)

- Endorsed by United Nations Committee of Experts on Global
Geospatial Information Management (UN-GGIM) in 2020 and
by the UNSC in 2021

- EG-ISGI to ensure that the report is in full alignment with the GSGF

The Global Statistical Geospatial Framework

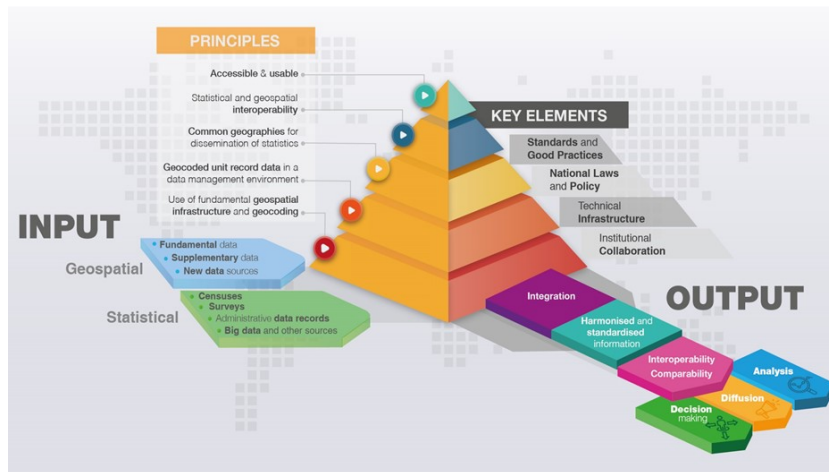# Global Statistical Geospatial Framework (GSGF)
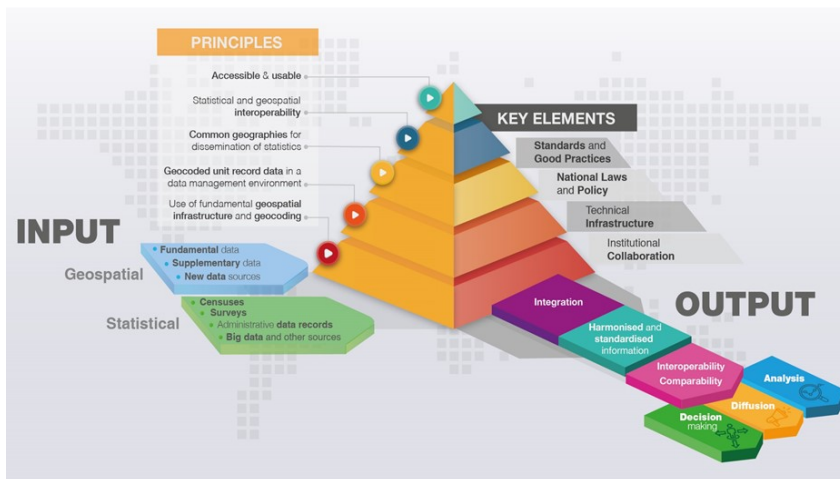
The GSGF contains five Principles:

1) Use of fundamental geospatial infrastructure and geocoding;

2) Geocoded unit record data in a data management environment;

3) Common geographies for the dissemination of statistics;

4) Statistical and geospatial interoperability; and,

5) Accessible and usable geospatially enabled statistics.

Global Statistical Geospatial Framework (GSGF)

Principle 1: Use of fundamental geospatial infrastructure and geocoding

Implementation of Principle 1 achieves the following objectives:

- Address, property, building, and location information are accurate and consistent, meeting country-level agreed standards and good practices;
- Geocoding results are as accurate and consistent as possible using common approaches or systems; and,
- Any geocoding issues are consistently managed through application of standardized approaches.

Global Statistical Geospatial Framework (GSGF)
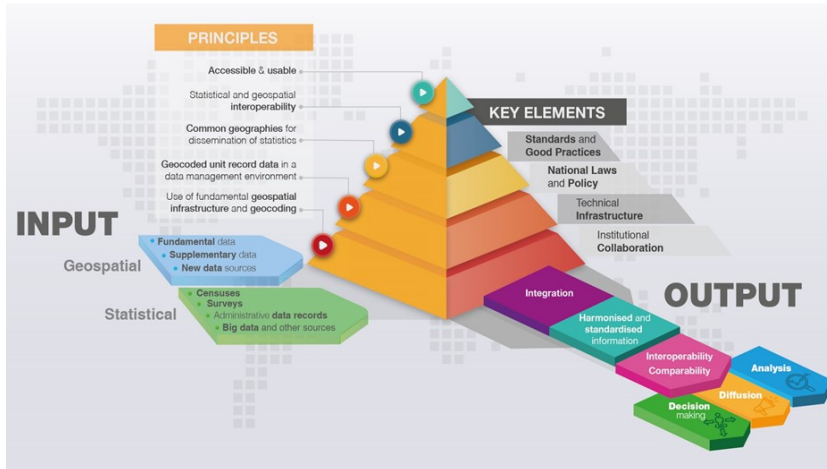
Principle 2: Geocoded unit record data in a data management environment

Implementation of Principle 2 achieves the following objectives:

• All statistical microdata is geospatially enabled for flexible use in analysis, visualization, dissemination and statistical data integration processes;

• Aggregation of data for larger geographies is simplified through storage of a unique identifier or code for a small area geography or standard grid cell for each unit record;

• Adaptation to changes to existing geographies or to allow compilation of data for new geographies is enabled;

• Data can be effectively managed, including the protection of privacy and confidentiality;

• Clear data maintenance and custodianship roles are defined; and,

• Geocoded information and metadata are consistent, interpretable and systematically maintained.

Global Statistical Geospatial Framework (GSGF)
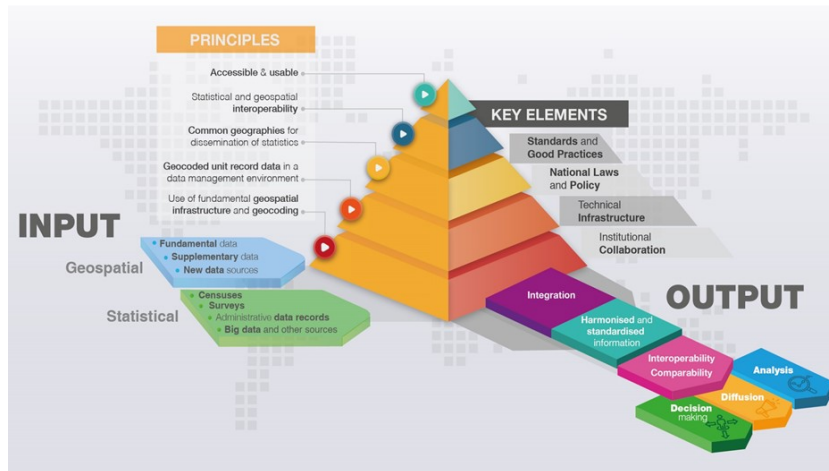
Principle 3: Common geographies for dissemination of statistics

Implementation of Principle 3 achieves the following objectives:

• Data from different sources can be integrated using a common geography;

• The visualization, analysis and interpretation of statistical and geographic information is simplified;

• Metadata supports data aggregation, integration and use;

• Identification and application of aggregation and disaggregation methods will enhance data quality and the assessment, consistency and increased use of data; and,

• The conversion of data between geographies is supported, through standard conversion mechanisms (e.g. through correspondences

Principle 4: Statistical and geospatial Interoperability

Implementation of Principle 4 achieves the following objectives:

• Greater efficiency and simplification in the creation, discovery, integration and use of geospatially enabled statistics and geospatial data;

• Ensures service-based or machine-readable access mechanisms (e.g. through APIs) are implemented to provide greater efficiency of access and use, and to allow adaptation

and evolution of uses through time; and,

• Increases the potential application of a larger range of data and technologies.

Principle 5: Accessible and usable geospatially enabled statistics

Implementation of Principle 5 achieves the following objectives:

• Data custodians can release data, with data privacy and confidentiality protected;

• Data users can discover and access geospatially enabled statistics;

• Data users can undertake analysis and visualisation;

• Web services and linked data methods enable machine-tomachine access, as well as dynamic linkage of information; and,

• Data users can know the status of modifications / changes in the data provided by data custodians through data integrity checks.

Establishment characteristics – Geographical location code

Purpose

The geographical location code complements the address and postal codes and can be used to derive classifications relating to the geographical location of units at the most detailed level, also other national classifications such as administrative regions, travel-to-work areas, health regions, and education regions.

Definition

- Countries can decide which code is most useful for their own purpose.

Proxies

- The geographical location code can refer to a classification at the most detailed level used in the country. It can refer to geocoding according to latitude and longitude points.

Sources

- Administrative sources.

Example from Mexico: Geographical coordinates in the SBR

Identification (Name and company name)

Localization (Address and geographical coordinates)

Economic (Strata of employees and economic activity code NAICS)

Contact (Telephone numbers, web page, e-mail address and social networks)

Control Identifier (Enterprise Statistical Code – CLEE)

The use of technology for updating SBR and putting geographic coordinates

- The field operations are crucial for updating the SBR. So, INEGI is using GPS in the Mobile electronic equipment of people gathering data under the Computer Assisted Personal Interviewing (CAPI)

- Likewise, through the Computer Assisted Web Interviewing (CAWI), INEGI identifies changes in the businesses informed by the respondents of Economic Surveys, and afterwards these changes are verified directly by INEGI. In the case of new establishments (births) the geographic coordinates are assigned

- In addition, linking the SBR with the registers provided by Government Agencies INEGI also assign the geographic coordinates

DENUE offers statistical information referred to a geographic area, which increases the variety of products, the interrelation of data and their analysis when visualizing them on digital cartography
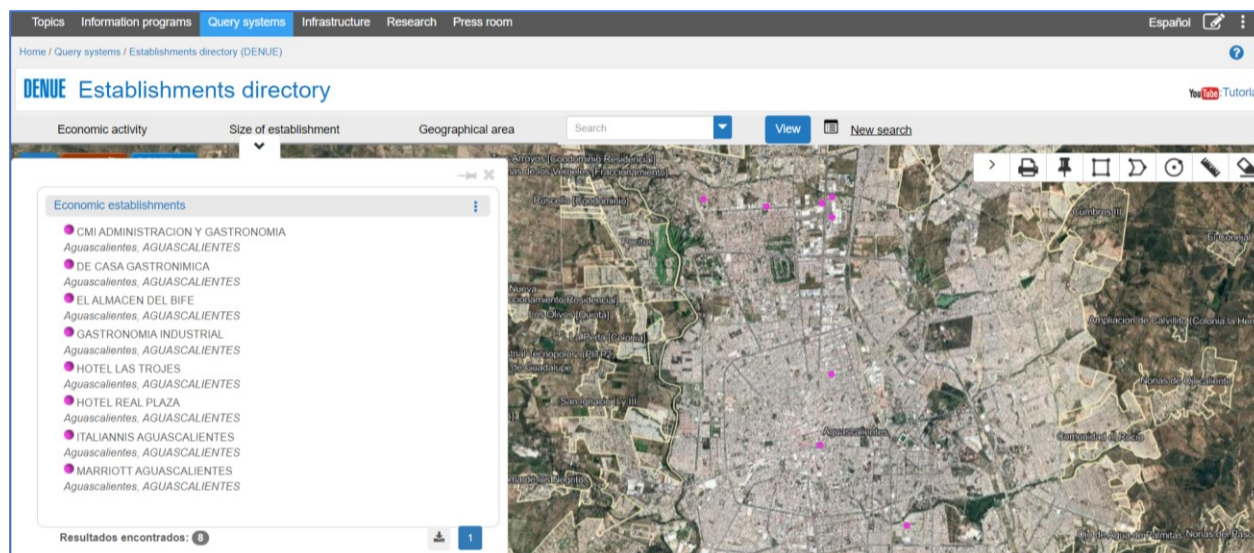
In Mexico, INEGI has developed some exercises institutionally and others in coordination with some public and private institutions

DENUE is available in: https://www.inegi.org.mx/app/mapa/denue/default.aspx

Evolution of an enterprise of convenience stores
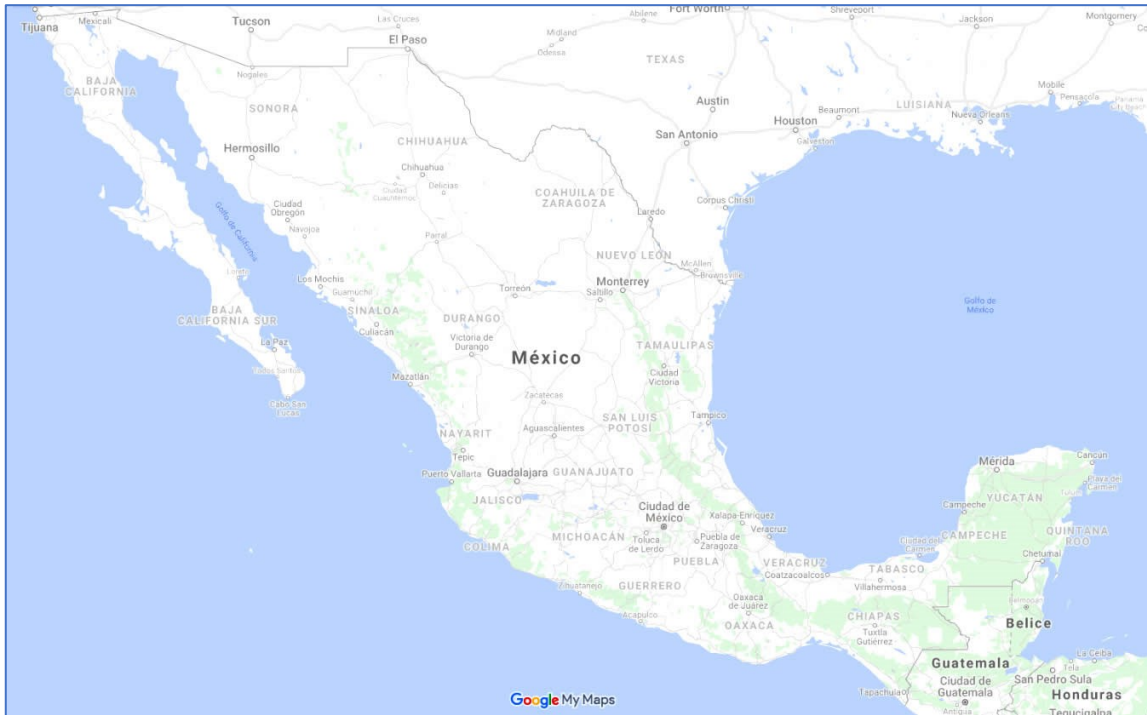
● **2008**

● **2013**

● **2018**

Evolution of an enterprise of convenience stores

2008 — 5,596 Establishments
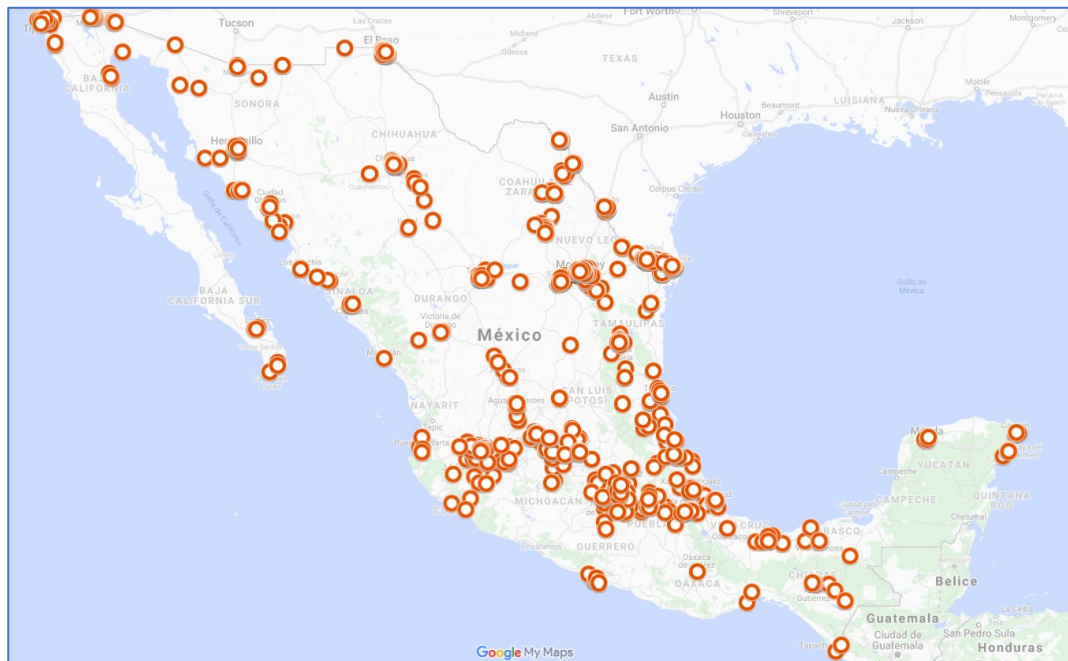
2013

2018

TOP 2008
1. Nuevo León
2. Tamaulipas
3. Baja California
4. Mexico City

Evolution of an enterprise of convenience stores

2008

2013   9, 648 Establishments
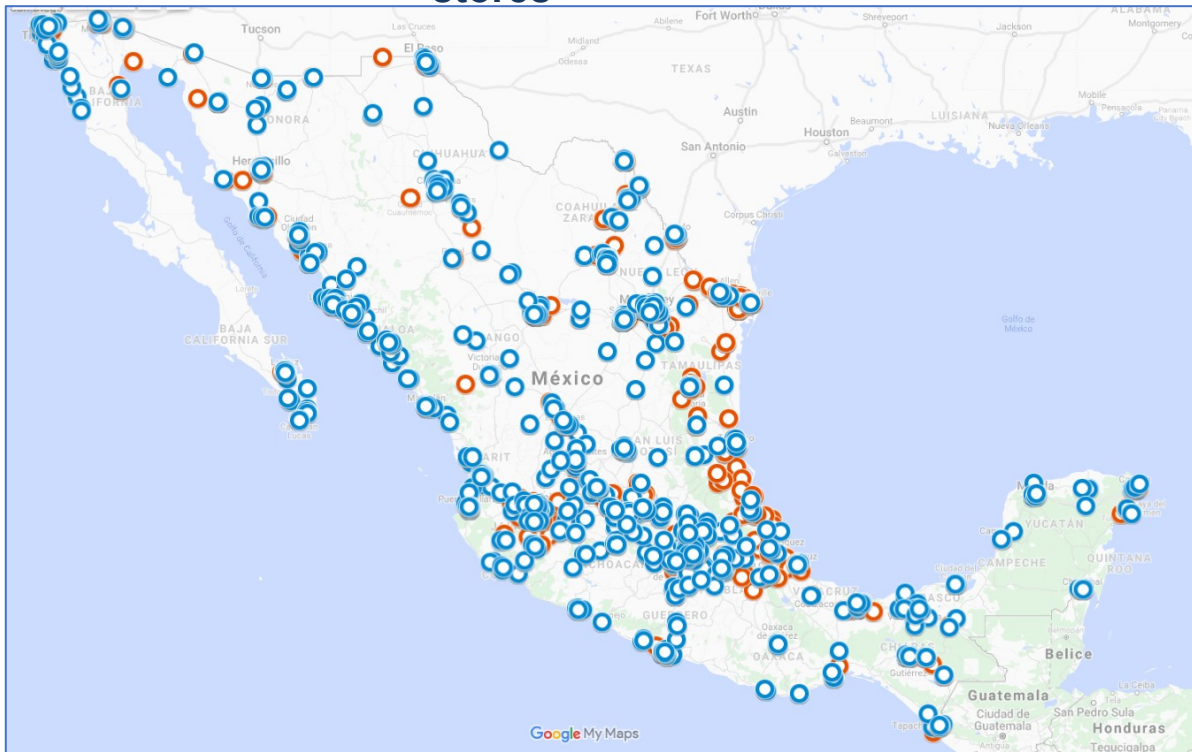
2018

TOP 2013
1. Mexico City
2. Nuevo León
3. México
4. Jalisco

Evolution of an enterprise of convenience stores

2008

2013

2018   15, 241 Establishments
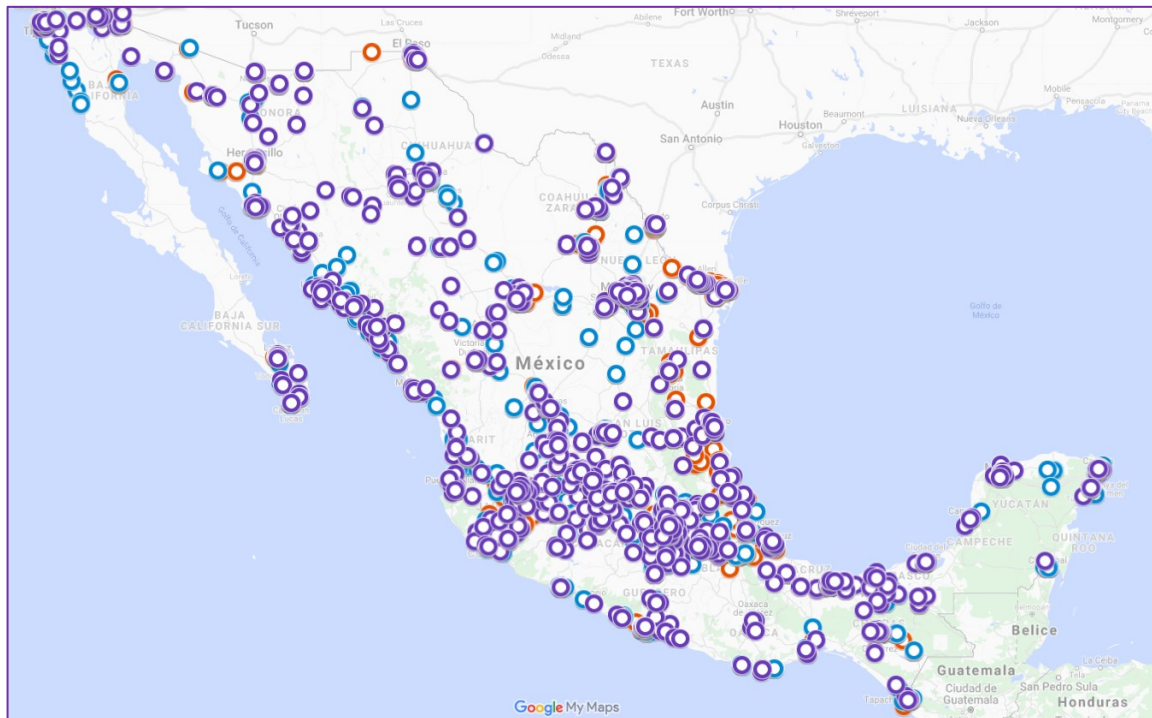
TOP 2018
1. Mexico City
2. Nuevo León
3. México
4. Jalisco

Work on developing guidelines on the integration of geospatial information in the SBRs

- The main objective of the report is to provide guidelines

- It mainly address the following issues:

  - What is geospatial information and benefits of integrating it in the SBR

  - What it means integrating spatial information in the SBR

  - How is geospatial information maintained

  - How to integrate geospatial information in the SBR

  - Country examples of the benefits of the integration

- Initial note:
  https://unstats.un.org/UNSDWebsite/statcom/session_55/documents/BG-4b-Note_on_Statistical_Business_Registers_and_Geospatial_Information-E.pdf

Some technical questions raised during the Wiesbaden Group meeting

- Linking business, location and people registry

- The approach in converting establishment/local units into x,y coordinates

- When the size of the establishment is large, the approaching in determining the appropriate address for the conversion into x,y coordinates

- Cases where addresses from the administrative source does not correspond to the location of economic activities

- Release at the disaggregate level - Confidentiality issue

Progress since November 2023

- An initial effort was undertaken in late 2023 to collect national case studies and identify technical issues to be addressed in order to facilitate the development of development of the methodological approach for the integration of geospatial information into the SBR

- Canada, Mexico and Sweden contribute by sharing the relevant cases in their country in integrating geospatial information into the SBR.

- Initial response received from the Mexico and Sweden to the technical question posed

- Updated note incorporated the above contributions drafted. The note will be submitted as a background document to the UNSC 2024. The note serves as input into the ongoing and future work on integrating geospatial information into the statistical business registers for the Task Team on SBRs and the EG-ISGI.

Response to the technical question posed

1. **What does it mean to integrate geospatial information (geographical coordinates) in SBRs? Does it mean that geospatial information has to be maintained in the SBR? Or does it mean that the SBR is linked to geospatial information (maintained outside the SBR)?**

Responses

- keep and maintain (correct) geographical coordinates outside the SBR (e.g. in an authoritative address register), to be linked to local units by means of the address string; establish an efficient, automated address cleaning and geocoding service for incoming address data

- The geospatial information in the SBR is included from the origin when data is collected. The basic statistical unit in the SBR is the establishment. Geographic data is automatically associated with establishments during fieldwork using computing devices supported by INEGI's infrastructure. Thus, geospatial information is maintained in the SBR which is fully updated every five years with the Economic Census and annually through business surveys, administrative registers, and information provided by informants through an interactive updating tool.

Response to the technical question posed

2. What are the additional advantages from integrating geospatial information in the SBR vs just having addresses?

Responses

- Geographical coordinates should ideally be maintained outside the SBR, and linked to local units whenever needed by means of address matching. We would suggest that it is rather an advantage of having SBR and geospatial information separated in different repositories. Having only the address of the local unit in the SBR is not a disadvantage if the criteria mentioned above can be met.

- The geographic location coordinates in the SBR enables the publication of information on digital maps or satellite images, which promotes the combination of geospatial data with business statistics, enriching the analysis of the distribution of economic activity in a specific territory (locality, state, regions and basic statistical areas

Response to the technical question posed

3. Can addresses be easily converted to geographical coordinates? If so, how?

Responses

- Yes, if the address information assigned to each local units is compliant with the agreed address standard, i.e. the authoritative address register (which contains geographical coordinates and is of good quality), the address information of each local units can easily be converted to geographical coordinates (or geospatially enabled).

- Use computing devices/tool (using a GPS) with which points (X,Y coordinates) are assigned based on address data, including in-between streets to avoid errors in the location of the points where the establishments are located. In this sense, the assignment of geographic coordinates comes from the field operations, supported by internal analysis if it is required.

Response to the technical question posed

4. If there is a building and dwelling register in the country with geocoordinates, could this be exploited to integrate geospatial information in the SBRs?

Responses

- Yes, if the address information assigned to each local units is compliant with the agreed address standard, i.e. the authoritative address register (which contains geographical coordinates and is of good quality), the address information of each local units can easily be converted to geographical coordinates (or geospatially enabled).

- Use computing devices/tool (using a GPS) with which points (X,Y coordinates) are assigned based on address data, including in-between streets to avoid errors in the location of the points where the establishments are located. In this sense, the assignment of geographic coordinates comes from the field operations, supported by internal analysis if it is required.

Response to the technical question posed

7. The approach in converting establishment/local units into x,y coordinates

Responses

- Geocoding strategy leveraging GeoDepot which is a geospatial database of Statistics Canada was developed by Statistics Canada's then Geography Division. If a street is matched to a Blockface that is one side of a street between two consecutive features intersecting that street, then the longitude and latitude coordinate values of the Blockface's representative point are assigned to that address on the SBR. Sometimes, when a match to a Blockface is not achievable due to inevitable data imperfection, the centroids of Statistics Canada's census dissemination blocks or dissemination areas or Canada Post's Postal Codes are applied to those addresses on the SBR.

- By using the address information of the local unit and matching with equivalent address information in the authoritative address register, which contains x, y coordinates.

Response to the technical question posed

8. When the size of the establishment is large, the approaching in determining the appropriate address for the conversion into x,y coordinates

Responses :

- Statistics Sweden approach is as follows:

  - 1. Convert address information to x, y coordinates by matching with the authoritative address register. 2. If needed:

  - 2a. Calculate the spatial footprint of the actual establishment of a certain local unit (for example: mines, etc.).

  - 2b. Make use of additional information (land use polygons, etc.).

  - Link x, y coordinates with additional information.

Response to the technical question posed

9. Cases where addresses from the administrative source does not correspond to the location of economic activities

Responses :
- In most cases, the location of the address location corresponds well with the actual location of the economic activity. However, there are some exceptions. One exception is local units of agricultural holdings. The local unit of an agricultural holding is typically assigned to the address of the owner/farmer dwelling. In many cases this is equivalent to the location of the holding, but there are also cases where the owner/farmer does not live at or close to the holding, which could result in an incorrect location of the economic activity. To handle this problem, the Board of Agriculture maintains a separate register on agricultural holdings. This register is linked to the SBR but contains also additional information on the location of the holding/economic activity.

Discussion

- Do your countries geocode your SBR? Could you share your experience

- What is the business case to justify the conversion of business location into geographical coordinates for statistical purposes?

- The approach in converting establishment/local units into x,y coordinates

- What does it mean to integrate geospatial information (geographical coordinates) in SBRs? Does it mean that geospatial information has to be maintained in the SBR (e.g. update through census) ? Or does it mean that the SBR is linked to geospatial information (e.g. maintained outside the SBR such as address register)?

Thank you!