Statistical Commission
Fifty-first session
3 – 6 March 2020
Item 3(v) of the provisional agenda
Items for discussion and decision: Open data

Background document
Available in English only

Local-level Statistics as Open Data

Prepared by Working Group on Open Data

Local-level Statistics as Open Data

Content

Summary
National Statistical Offices, NSOs, have a role to play in providing local-level statistics as open data. The paper describes how this role can be strengthened and which steps to take regarding the choice of geographical areas, statistical content, balance between utility and privacy and how to increase the visibility and use of local-level statistics.

Local-level statistics can inform decision makers of

- geographical population patterns;
- population density in relation to risks of flooding;
- inequalities regarding education or income; or
- accessibility to various services such as schools or hospitals.

Local-level statistics support more fact-based policy-making and are crucial for follow-up and accountability. Increasing demands for timely and trustworthy local-level statistics will be hard to meet without an increased use of data sources like administrative registers and geospatial information.

There is not one solution that suits all NSOs but by developing the capability to integrate existing statistical and geospatial data sources, the amount of local-level statistics can increase even with limited resources. By applying open data standards to local-level statistics, they can be re-used for many purposes. Providing local-level statistics should be an integrated part of the NSOs production of regional statistics and undergo the same quality assurance and disclosure control.

Some country examples, coming mainly from countries represented in the Friends of the Chair Group on the Fundamental Principles of Official Statistics, FOC-FPOS, highlight good practice regarding local-level statistics.

Introduction
This background document provides guidance/advice for national statistical offices, NSOs, on steps to take when releasing local-level statistics as open data. Statistics at the local level describe the population living in geographical areas such as neighbourhoods, urban or rural areas, census districts, electoral districts or grids. The data sources for compiling local-level statistics may be the population and housing census, administrative records and/or geospatial information. Local-level statistics are always aggregates, often based on micro-data but never revealing information about single individuals, households or establishments.

The Friends of the Chair Group on the Fundamental Principles of Official Statistics, FOC-FPOS, has described in more general terms how to understand open data practices in official statistics in a background document presented at the 50th Session of the UN Statistical Commission, UNSC50[1]. A commonly used generic open data definition is that “open data and content can be freely used, modified, and shared by anyone for any purpose”[2]. Common aspects of open data also include that it is machine-readable and licensed for reuse. This can be achieved through using open data formats, making the data available for download or accessed through an API/web service, and published under an open license such as Creative Commons.

When providing local-level statistics as open data NSOs need to consider quality and privacy aspects in the same way as with all official statistics. An international seminar in Korea on open data for the Sustainable Development Goals (SDGs) in 2017[3] resulted in a description of open data as “data that are

[1] https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3c-Open-Data-guidance-and-mapping-to-FPOS-E.pdf

[2] https://opendefinition.org/

[3] https://unstats.un.org/unsd/statcom/49th-session/documents/2018-6-OpenData-E.pdf

of high quality, are well documented, respect data privacy concerns, are free and are easily accessible and usable". These criteria might be helpful when the NSO decides on which local-level statistics to open up.

The efforts in providing high quality open data also need to be user-oriented, with a fit-for-purpose perspective. There is a need to strike a balance between appropriate disclosure controls and usefulness to users in putting out such statistics as open data. While statistics describing the population at a national level can generally be presented with hardly any disclosure measures, as the risk of revealing information about single individuals or households is close to zero, statistics describing the population living in a small area almost always need disclosure control. If overly protective disclosure methods are applied, there might not be any useful value left in the local-level statistics.

The NSO as provider of open data

Citizens are entitled to public information and therefore statistical agencies should retain trust in official statistics and facilitate the correct interpretation of statistics according to the United Nations Fundamental Principles of Official Statistics (UNFPOS)[4]. Local-level official statistics can reveal local disparities beyond regional or national averages and help in building trust in public information. Local level data is important for policymaking and it often underpins the decision-making of civil society and service delivery from central and local governments. It is also helpful for better decision making by the private and charitable sectors.

Making data available is an important step for building trust in public information, but not enough. This might require efforts to strengthen structural, statistical and reputational factors[5]. For example, data availability combined with effective dissemination where data connects with the populations' experience might be more powerful in building trust than only opening the data.

Giving citizens access to local-level statistics that are produced and disseminated in a consistent way for the whole country creates an opportunity for them to discover facts about their neighbourhoods and allows for comparisons between local areas. NSOs can play an important role in providing more local-level statistics[6], making them easier to access through user-friendly databases, mapping tools and using open data standards. Ideally the database should be integrated/linked to existing national/regional databases, not a separate one. Having a systematized and integrated database can help to observe and analyse a phenomenon from a bottom-up perspective.

The need for local-level statistics is also related to the 2030 Agenda for Sustainable Development[7], where targets often require action at the local level for countries to succeed in reaching the goals on the national level. The Agenda commits to 'leaving no one behind', which includes the need to measure, for example, urban and rural populations. The UN Global Geospatial Information Management, UN-GGIM, provided some early examples in 2015 on where local-level statistics might be useful to monitor sustainable development[8]. Since then a number of NSOs have started to publish SDG indicators and include local-level statistics, such as the example of Ireland[9].

The NSO needs to investigate which statistical content users are interested in, determine suitable geographies and propose measures for disclosure control, research; which tools to choose for

---

[4] https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx

[5] https://paris21.org/sites/default/files/inline-files/CRF_BackgroundNote_0.pdf

[6] Haldorson, M. (2019): Regional Statistics, Vol. 9. No. 1. 2019, High demand for local area level statistics – how do National Statistical Institutes respond? http://www.ksh.hu/docs/hun/xftp/terstat/2019/rs_haldorson.pdf

[7] https://sustainabledevelopment.un.org/post2015/transformingourworld

[8] https://un-ggim-europe.org/wp-content/uploads/2018/11/2015-10-23_UN-GGIM-Europe_Report-from-SWG-B1-on-Priority-User-Needs-ver-1.1-1.pdf

[9] https://irelandsdg.geohive.ie/

visualizations and how to set up open data access. Implementation could be carried out with a stepwise approach, interacting with users and intermediaries (open data consumers and developers) along the way. As more local-level statistics become available, the customer base will grow and diversify as they realise data, that is more useful for them in their local context, is available.

Being able to put out more statistics if the budget is limited is a challenge, so the increased openness should rely on efficient and sustainable processes – looking for ways to create more output with available data. Local-level statistics can often be produced by integrating different data sources, statistical and geospatial. There are a number of recommendations, guidance and good examples on how to do this in a report from the European project GEOSTAT 2[10]. The main message from the project is to regard integration of statistical and geospatial information as part of the regular statistical production process, supported by the Generic Statistical Business Process Model[11].

For each NSO the steps towards providing more local-level official statistics might look different, but by starting with a capability assessment the appropriate steps could be outlined in an action plan. For example, Statistics Sweden has identified strengths and weaknesses through performing an assessment of their capability to integrate statistical and geospatial information[12], which has resulted in actions towards releasing more local-level statistics as open data.

Geography: Choosing the right geography for local-level official statistics

The basics

The starting point for most NSOs will be to publish local-level official statistics as open data on geographies that already exist, for example as a result of their population and housing census. If the NSO has census data collected with geographic reference to an area (e.g. census district), then producing local-level statistics is less flexible than if data is collected with a geographic reference to a point (e.g. building coordinate).

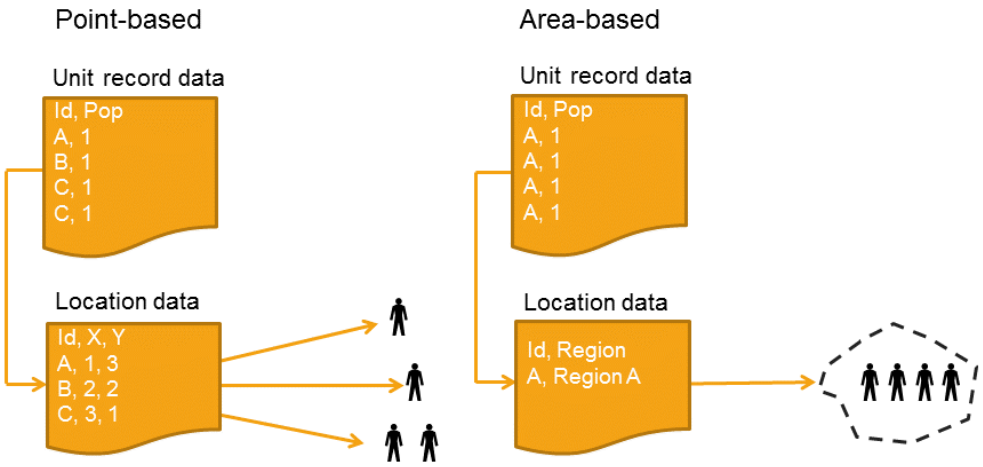

Figure 1: The difference between point-based and area-based geographic references[13].

[10] https://www.efgs.info/geostat/geostat2/

[11] https://statswiki.unece.org/display/GSBPM/Generic+Statistical+Business+Process+Model

[12] Haldorson, M.–Moström, J. (2018): Implementing the Statistical Geospatial Framework at Statistics Sweden, National Report as part of the GEOSTAT 3 project, https://www.efgs.info/wp-content/uploads/geostat/3/Impementing-the-GSGF-at-Statistics-Sweden-v0.96_Draft_for_GEOSTAT.pdf

[13] https://www.efgs.info/geostat/geostat2/

With statistical data linked to geographic points, it is possible to create a number of different output geographies taking into account their design to provide statistical content that is useful after applying disclosure control. There are also risks to consider when releasing overlapping geographies (it should not be possible to disclose information on single individuals/households when overlaying various geographies)[14].

Common geographies

"Use of a common set of geographies will ensure that all statistical data is consistently geospatially enabled and that users can discover, access, integrate, analyse and visualise statistical information seamlessly for geographies of interest." This is stated in the UN-GGIM global framework for statistical-geospatial integration, where common geographies are regarded as an important principle[15]. Geography can therefore be regarded as a tool for integrating data.

The NSO can play an important role in providing common geographies to be used by other public and private data providers. For users it is often very difficult to combine data from different sources due to a lack of common geographies. A starting point could be to create an open Geography Web Portal, such as the one provided by the **Australian Bureau of Statistics,** ABS[16], where users can find out about Australian statistical geography boundaries and more. Another example is the Open Geography portal[17] from the **Office for National Statistics,** ONS, which provides free and open access to the definitive source of geographic products, web applications, story maps, services and APIs for the UK.

Statistics New Zealand provides geographic boundaries and statistical data as open data and makes most of its geospatial and statistical data available online[18] with a licence that allows people to use, modify and share the data. Open geospatial data is also an important component of the national Spatial Data Infrastructure, SDI, because for an SDI to function properly, data must be freely accessible, machine readable and in a format that can be integrated with other data. The NSO manages their geospatial boundaries according to a national Statistical Spatial Framework, similar to the global framework developed by UN-GGIM, so that they are based on sound international principles and standards.

Statistics New Zealand has used *statistical meshblocks* (smallest geographical area bases on 30-60 dwellings or 60-120 residents) and larger area unit boundaries for some time. These boundaries are available as a dataset. These have recently been updated to "Statistical Area 1" which have an ideal size of 100-200 residents and a maximum of 500; and "Statistical Area 2" which contain 1,000-4,000 residents. The challenge with these geospatial units is that they are regularly updated as new dwellings are built making it difficult to present time-series geospatially.

The *electoral role boundaries* are also available as a dataset-- these are based on population (census) and are also subject to change-- restricting the ability to develop time-series. *Territorial Authority boundaries* (local government) are a more stable set of boundaries and frequently used for aggregating statistics. The statistical tables, including the names of territorial authorities, can readily be combined with the boundary dataset and visualised on a map. This is useful for the likes of regional economic development organisations, but the area is too large for many community needs. *Suburb boundaries* are another popular boundary dataset (administered by local authorities), particularly for analysing property prices and rentals and civic planning.

Administration boundaries are available and different government agencies have different administrative boundaries for organising the distribution of resources and service delivery across New Zealand. For

---

[14] Costemalle, V (2019): Detecting geographical differencing problems in the context of spatial data dissemination, Statistical Journal of the IAOS 35 (2019) 559–568

[15] https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

[16] https://www.abs.gov.au/geography

[17] http://geoportal.statistics.gov.uk/

[18] www.stats.govt.nz

example, health, education, police and courts all have different area boundaries from which they draw their statistics, based on where their branch offices are located. These boundaries are all different in size and shape making correlation and analysis across sectors very difficult.

A common way for NSOs to overcome the challenges of common geographies is to use a point-based approach.

Overlapping geographies

When NSOs develop a more flexible way of producing local-level statistics, with a point-based approach as described earlier, there are risks to consider if statistics are made openly available on overlapping geographies. These risks concern geographical differencing and can be illustrated in the figure below.

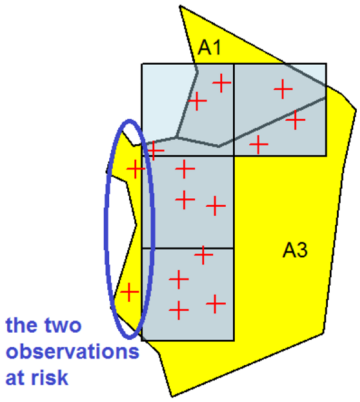

Figure 2: Example with overlapping geographies and observations at risk of disclosure[19].

In this example, the risk of disclosure is coming from the fact that the two observations at risk are part of one geography (A3) but not part of the point-based geography (the squares). Geospatial differencing refer to the two points being a remainder between the two geographies and could therefore be identified. The user may compare statistics for the four grids with statistics for the two areas A1 and A3. If the two points in the figure represent one household each, the difference between the grid-based statistics and the area-based statistics must refer to the points (geo-located households).

Before publishing local-level statistics, the NSO therefore needs to consider whether the geographies will be overlapping in a way that could reveal information about individuals. An example of how these risks were mitigated comes from **Statistics Sweden,** where new territorial geographies called *Demographic statistical areas* (DeSO) have been designed to fulfil a number of requirements. The DeSO are large enough not to be significantly affected by disclosure measures; are consistent over time and aggregate to the municipality level. Publishing statistics on DeSO is part of Statistics Sweden's open data strategy.

When assessing if other geographies should also be used for official statistics, the conclusion was that urban areas/localities (that could be interesting for users) in many places are too similar to DeSO, which could lead to a risk of disclosure. There are a few DeSO and urban areas where only a small number of households differ between the two geographies. Originally when DeSO was created, it was based on urban areas but included a buffer zone of 600 meters around the urban areas. However, there are still several DeSO where there are occasional populated properties in the buffer zone. These households can then easily be disclosed if statistics are reported for both geographies.

---

[19] Costemalle V.: Detecting geographical differencing problems in the context of spatial data dissemination, Statistical Journal of the IAOS, vol. 35, no. 4, pp. 559-568, 2019

A similar consideration has been made not to release statistics on grids together with DeSO. Statistics Sweden would like to release as many local-level statistics as possible, as part of their open data strategy, but due to the geographical differencing problem DeSO and grids are not possible to release. To solve the issue to some extent, grids or hexagons are planned to be the data source for a new geographic tool, where the user can choose an address and find out statistics for a buffer zone or a freely drawn area – as long as the statistics for the area meet some basic threshold values on minimum population. The users receive a statistical result but can't access the individual grids or hexagons.

Content: Choosing relevant statistical content

The basics

Data sources appropriate for creating local-level statistics are mainly censuses, administrative registers and other data sources where the data already have a geographic reference. The reference should be to a geographic coordinate (building, address, real estate) or to a small area (neighbourhood, census district etc.). Small areas might sometimes also be administrative areas, where there is a wealth of statistical information available already. But if that is not the case, which statistics should have high priority to release?

The ninth session of the UN-GGIM Committee of experts endorsed a global set of **fundamental data themes** presented in a background paper[20]. One theme covers population distribution: Geographical distribution of people, including population characteristics. The theme is considered fundamental because "it is vital to understand the spatial distribution of the population and its characteristics, as well as how population impacts urbanisation, regional development or sustainability".

Without data on population distribution it is not possible to assess change over time from an urban/rural perspective. Local-level statistics on where the population live and where people are at a given time is also highly in demand when preparing for natural disasters or making accessibility assessments. The recommendation is for countries to have population statistics on any geographies available (census tabulation area, city, postal code area, grid etc).

The UN-GGIM paper states that "ideally 5-year age cohorts by gender would exist globally, which allow for groupings to include specific at-risk populations, such as elderly, family composition, children, indigenous population, immigrant/ethnic origin, and education. In the absence of such detailed cohorts, estimates of these at-risk populations are needed. Additional demographics include rates of employment, personal income, household income, and living conditions, i.e., the types of materials used for dwelling units".

Country examples

The **Department of Statistics Malaysia** (DOSM) publishes statistics at various levels, specifically at the national, states and even administrative district (local) level, which are all accessible at DOSM's Portal. The statistics are downloadable and can be viewed in an interactive visualisation and geographical layers. As of now, the local-level statistics include socio-economic data such as Population, Housing, Expenditure, Health, Basic amenities and Agriculture. As a commitment and support to the implementation of the SDGs, DOSM had managed to ensure that the majority of the SDG goals are available at the state level. Nevertheless, DOSM is continuously committed to further enrich the local-level statistics produced and is in the midst of improving its compliance to data openness by enhancing the accessibility of the data.

Statistics Sweden has a long tradition of producing local-level statistics as commissioned services, but has not provided any statistics on a more detailed level other than municipality as official statistics. That is now changing as part of their open data strategy, including local-level statistics as official statistics and re-

[20] http://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/Fundamental_Data_Publication.pdf

allocating appropriations to fund the production. Based on the statistics that have been most in demand as commissioned services they have selected 20 tables to release as local-level official statistics and open data based on Demographic Statistical Areas, DeSO. There will be one-dimensional tables, due to confidentiality, where the user will receive one variable (e.g. population by age, sex, civil status, citizenship, country of birth or Swedish/foreign background etc) presented on DeSO areas. The statistics can be used, for example, to describe inequalities between different local areas, to design national policies targeting areas at risk for increased crime and segregation.

Finding a balance between utility and privacy

The basics

NSOs should play a role as providers of open data; firstly, because NSOs are supposed to be a trustworthy protector of the confidentiality of identifiable information, and secondly because NSOs are expected to be a neutral space for statistical information. i.e. provide statistical information free of political interference and private interests.

NSOs need to find a balance between openness and privacy when starting to release more local-level statistics as open data. There are ways to maintain the balance by choosing a suitable geography for the local-level statistics and finding a suitable way to present statistics (one-dimensional tables for example). Since each country operates under a unique combination of legislative and cultural settings, tolerance on privacy protection can vary greatly.

The more granular the geospatial data is, the more useful it is for communities, but at the same time the risk to privacy is greater. This is acknowledged in a report by the French National Institute of Statistics and Economic Studies, INSEE:

> *"Disclosure control experts dealing with spatial data face a paradox. On the one hand, such data need more protection because they permit more identifications, but on the other hand they offer many possibilities of analysis that users do not want to distort too much."*[21]

This paradox between the need to maintain confidentiality and protect privacy within spatial datasets on the one hand; and their accessibility and utility to statisticians and non-statisticians on the other, is at the heart of the question of how openness and privacy should be balanced in the case of sub-national, granular data. Finding the balance between confidentiality and utility is a question of both governance and methodology.

In terms of governance, NSOs need to balance the requirements set out in the UNFPOS to both produce useful statistics that are trusted by citizens and to uphold confidentiality. Adopting an 'open by default' approach[22] to publications can help to do this. Adopting this approach simply means that "public data should be disclosed unless there is a legitimate reason for it not to be"[23]. In practice this requires NSOs to proactively disclose data unless there is a legitimate, lawful reason not to. Upholding confidentiality is a widely recognised legitimate reason and the methodological mechanics of how to do so in relation to spatial statistical micro-data is touched upon later in this section of the paper.

To apply an 'open by default' approach to the publication of local-level official statistics, NSOs need to consider their governance frameworks – from the level of the statistical law governing their activities, down to the internal business processes that are employed within the NSO itself. In terms of legislation, NSOs should consider adapting and adopting clauses from UN Economic Commission for Europe,

---

[21] https://www.efgs.info/wp-content/uploads/informationbase/introduction/Handbook_of_Spatial_Analysis_INSEE_EUROSTAT_2018.pdf, p 351

[22] See Principle 1 of the Open Data Charter: http://opendatacharter.net/

[23] https://www.sdsntrends.org/research/2019/3/4/maximizing-access-public-data#intro

UNECE's Generic Law on Official Statistics[24], in particular clauses under Article 20 that pertain to confidentiality. UNECE's official guidance on the Generic Law is also a helpful source of detailed guidance and Chapter 7 in particular covers emerging issues related to statistical legislation, including the publication of open data[25].

In terms of governance structures within NSOs, if government-wide Open Data Policies exist, NSOs should ensure that they adequately reflect their needs and approaches. If no such policies exist, NSOs should consider drafting a policy that covers their needs. The Sunlight Foundation's *Open Data Policy Guidelines*[26] may provide a helpful starting point in setting-out and explaining the various components that should be included in an open data policy.

The next step in the open by default approach will be for NSOs to consider the implications of protecting confidentiality in localised spatial data on their data collection and compilation methodologies. Chapter 14 on Confidentiality of spatial data within the Handbook of Spatial Analysis[27] provides a very helpful starting point from which tailored approaches can then be further elaborated and developed.

Country examples

The **Department of Statistics Malaysia** (DOSM) is entrusted to safeguard the privacy and confidentiality of the data collected from respondents. Hence, the disclosure of the data collected is bound to the Micro-Data Dissemination Policy which was established by DOSM in 2009. It is in accordance with the Statistics Act 1965 (Revised-1989) and other regulations stipulated by the DOSM, particularly pertaining to the confidentiality of respondents' information. This policy enables DOSM to disseminate statistics while ensuring that the confidentiality of respondents' information is protected.

DOSM has implemented certain procedures to ensure the confidentiality of the data is safeguarded. For example, where there is individual unit data with less than three establishments, it will never be revealed. Thus, the aggregation procedure for certain unit data will be developed to ensure only data with at least three or more establishments can be disclosed.

INEGI, Mexico benefits institutionally from the integration of statistics and geography in a single institution; a major advantage in matters that involve linking both areas. These institutional settings and integrative tools designed to fulfil national priorities allow for georeferencing statistical data, for example, with the production of a National Geostatistical Framework, as a basis for georeferencing each observation in censuses and surveys.

Local-level official statistics can better reveal disparities beyond more aggregated information, which is useful for decision-makers and may help in building trust in public information, for example by revealing local patterns of inequality. However, maintaining this trust also concerns finding a balance between openness and privacy. By law, INEGI protects the privacy of data providers and guarantees the confidentiality of the individual information provided and its use for statistical purposes. In addition to regulations addressing information security, INEGI has put in place organizational, administrative, physical and technological measures for guaranteeing information security. On this note, output geographies are designed to provide useful content after applying rigorous disclosure control. The lowest level geography (report unit) is a city block, which is applied to some statistical information programmes like the Population Census.

INEGI is also experimenting with design of output geographies that address privacy concerns by a rearrangement of the data (proportionally) into equal area cells, forming a grid. This approach also

---

[24] https://www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20163_E.pdf

[25] http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2018/CES_6_Common_elements_of_statistical_legislation__Guidance__for_consultation_for_upload.pdf

[26] https://sunlightfoundation.com/opendataguidelines/

[27] https://www.efgs.info/wp-content/uploads/informationbase/introduction/Handbook_of_Spatial_Analysis_INSEE_EUROSTAT_2018.pdf

favours the easier integration with other data structures like raster, which is more commonly used in the report of geographical information (i.e. elevation, vegetation, weather, etc). Equal area units also allow for the application of standardized methodologies to bigger areas.

Statistics Sweden has designed their new Demographic Statistical Areas, DeSO, to protect confidentiality. Sweden is divided into 290 municipalities and about 6,000 DeSO areas. On average the DeSO population is about 9 times smaller than a municipality. The median DeSO 2018 had just over 1,700 inhabitants, while the average municipality had just over 16,000. inhabitants. The lowest number of inhabitants in a DeSO is approximately 650, a quarter of the 2,450 who lived in Bjurholm (Sweden's smallest municipality). The smallest DeSO to the surface covers about three blocks and is about 150 x 150 meters. In summary, this means that Statistics Sweden could not publish statistics on DeSO to the same extent as they do for municipalities.

There is a need to limit different combinations of data. This is to ensure that the population base is not too small. The population is reduced rapidly if users get cross-tabulated data, which then increases the risk of identifying specific individuals. At the DeSO level cross-tabulations are only allowed by gender, and in some cases where suitable, by broad age categories.

Statistics based on administrative registers, as in the Swedish case, don’t have the same uncertainty and, consequently, have less protection than sample-based statistics have. A user may have knowledge of which objects are included in the production of the statistics, which cannot be assumed in the same way for sample-based statistics. Geographical location generally facilitates identification.

Local-level statistics should not be less protected than the currently published municipal statistics. Since statistics for DeSO sum up to the municipality level, it is possible to recalculate the values. To allow for DeSOs to be summed up, a method of Controlled Tabular Adjustment (CTA) is preferred. CTA changes the values of so-called sensitive table cells and also values of other table cells to keep the values of each DeSO possible to sum up on municipality level. It remains to put the method in practical operations supported by a suitable tool.

Increased visibility for local-level statistics – increased use

The basics

Official statistics on administrative geographies can often be disseminated without support of a vizualisation tool, such as a map. The name of the administrative area might give the user enough information to interpret the numbers. For local-level statistics, the user very likely needs to look at the boundaries of different geographies to be able to use the statistics effectively.

For NSOs to release local-level statistics as open data, the minimum requirement could be to release the data as machine-readable data and let others use and re-use the data in various services for users. This implies using existing platforms for data dissemination and not creating parallel and special solutions for publishing statistics at a local level. Otherwise it is easy to end up with separate solutions where files are prepared, manually packaged and published on the NSO’s website. This is inefficient both from a resource point of view and from the point of view of use.

Many NSOs have higher ambitions and provide various mapping and visualisation tools themselves. It is recommended to investigate suitable tools or collaborate with a Geospatial Partner (public or private) that can add such tools for increased use and visibility to the local-level statistics.

The variety of tools available ranges from the simplest visualisation tool-- where the user can see the boundaries, coloured by, for example, population density, but with no additional geographical background information – to the most advanced tool where the user has a detailed map with a lot of geospatial layers (buildings, roads, water etc) and uses geospatial queries to access and interrogate a database.

The Global Statistical Geospatial Framework[28] suggests enabling the release of geospatially enabled statistical information in a usable and accessible form. It specifically promotes the use of standard web services and linked data methods to provide dynamic, machine-readable access to data with the necessary assurances regarding integrity.

Country examples

Statistics New Zealand has illustrated their customers' use of Stats NZ services:

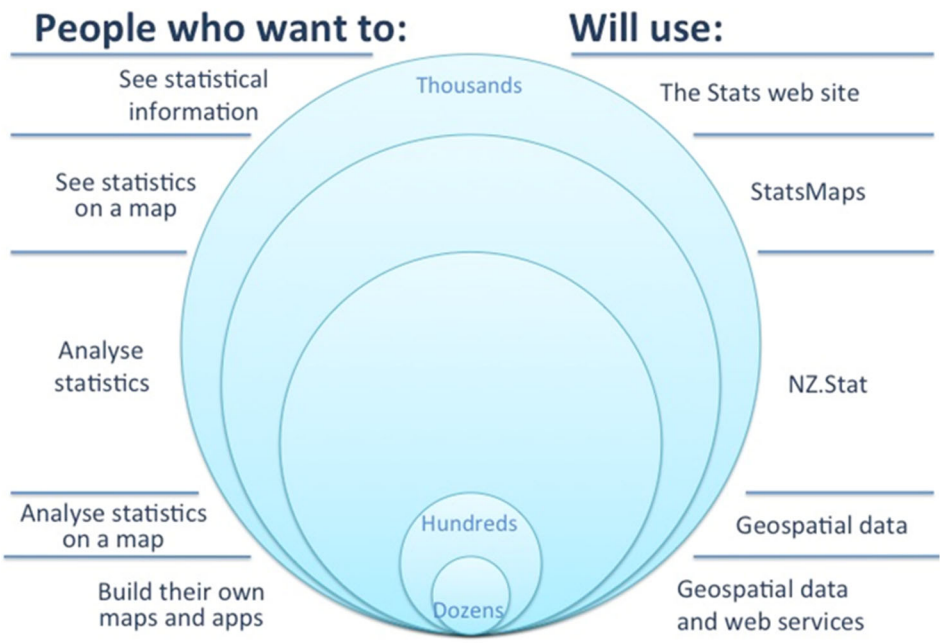

Figure: An indicative representation of the likely number of users, the things they want to do and the services they might want to use[29].

The *StatsMaps suite*[30] of products allows users to visualise geospatial data. It does not provide open data for reuse. Visualisations are useful for the public to see the data as a map rather than the traditional statistical data table. The geographic boundaries are needed to show the physical extent of the area – a dimension that can't be shown in a spreadsheet but is important when comparing values between areas to identify spatial patterns in the data.

The *DataFinder*[31] is a geospatial/statistical data service. It is the same platform that Land Information New Zealand, LINZ, uses for the LINZ Data Service[32] and the Ministry for Environment Data Service[33] for environmental information. This has been a big improvement to the way Statistics New Zealand were delivering geospatial boundary data previously – through large zip files on their web site. Now users can select from a wide array of data; select to download a specific area of interest (instead of the entire dataset as they had to previously), and choose one of many formats they wish to receive the data in.

---

[28] https://ggim.un.org/meetings/GGIM-committee/9th-Session/documents/The_GSGF.pdf

[29] Draft from Statistics New Zealand: "Statistics New Zealand's Open Geospatial Data".

[30] http://archive.stats.govt.nz/StatsMaps/Home.aspx

[31] https://datafinder.stats.govt.nz/

[32] https://data.linz.govt.nz/

[33] https://data.mfe.govt.nz/

The **Department of Statistics Malaysia** (DOSM) provides *MyLocal Stats*[34]*,* a comprehensive interactive visualisation online platform featuring statistics from national level down to local level. The statistics are downloadable and reusable. DOSM has also developed a special mobile application for MyLocal Stats which enables users to access the statistics at local level.

The STATS Geoportal[35] provides local-level statistics that allows users to search, access, display and view information by population / demographic, social and economic data at various geographical levels. The portal is expected to be able to meet the overall openness criteria in the near future.

Way Forward

There are many countries with a lot to share and show, not least in connection with actions and follow-up to the 2030 Agenda and leaving no one behind from a location perspective. The examples given in this paper only gives a hint of some sources of inspiration.

With increased use of administrative data sources, such as population registers, taxation files and various geospatial data, such as building or real estate coordinates, the possibilities to produce relevant and timely local-level statistics is expected to grow. NSOs have an important role to play in providing local-level statistics covering the whole country, taking into account the need to balance openness and privacy.

The UN FOC-FPOS, Working Group on Open Data, can play a role in providing guidance on how to apply the principles of open data when releasing local-level statistics. This paper is a first attempt to high-light factors to consider when making local-level statistics more open, especially when the local level geography differs from administrative regions which most countries have as standard geographies. Geospatial patterns rarely follow administrative boundaries, which is why showing local-level statistics on more functional areas, such as urban and rural areas, are increasingly important.

[34] https://www.dosm.gov.my/v1/index.php?r=column/cthree&menu_id=UEq1NkpJUFYzRzBKclE4V2JEb2I3QT09

[35] https://www.mycensus.gov.my/geostats/main.php