

## OECD WORKSHOP

### *Modernising statistical systems for better data on regions and cities*

November 5<sup>th</sup>, 2018  
OECD Conference Centre  
CC13

#### Concept note

The availability and quality of usable statistical information is increasingly important for the success of individuals, firms and places. In the contemporary and rapidly evolving data ecosystem, the emergence of new available information opens important opportunities to foster knowledge on regions and cities, as well as for individuals, firms and stakeholders operating there. New sources of information, such as big data, administrative data and new processing technologies based on open-source platforms can bring an unprecedented improvement in the capacity to produce comparable statistics at different geographic levels and flexible spatial detail, going beyond the shortcomings of traditional data sources.

In this context, many National Statistical Offices (NSO) are in the process of modernising their statistical systems and infrastructures. Modernisation, however, has many dimensions. In addition to new data sources and processing technologies, it involves new partnerships and collaborations with data providers, new organisational forms, new visualisation and dissemination approaches, etc. This process of modernisation can potentially create important opportunities for our capacity to measure different types of phenomena, ranging from individual behaviour and socio-economic conditions to environmental trends, at a very detailed geographical scale, potentially helping regions and cities to grow and prosper sustainably.

The data ecosystem has brought about new challenges and opportunities for NSOs and other organisations responsible for the production of statistical information. Big data are mostly unstructured and very heterogeneous in terms of quality, coverage and reliability. No systematic architecture is available universally to make such data coherent and usable for both policy making and policy analysis. In this respect, crucial aspects on which it is important for NSOs and data experts to focus are definitions, methods and quality standards, to ensure new data can be coherently used and integrated with existing statistical information.

The integration of different data sources should also be considered in the framework of a new data landscape where the number and type of organisations (and individuals) that are both users and producers of data has increased, including subnational and local governments. The latter produce a number of administrative records that, especially if developed, coordinated and integrated with existing information and in partnership with National Statistical Offices, can improve the geographical granularity, timeliness and completeness of territorial information on a wide range of topics.

The OECD Working Party on Territorial Indicators is organising a Workshop to discuss, among NSOs and other country and institution representatives interested in regional development, how the modernisation of statistical systems can improve territorial information. It will focus on four pathfinder themes of relevance within the field of regional development policy, addressing different features that characterise the process of modernisation of national statistical systems for better data and information on regions and cities.

The workshop will be organised around an opening session and four thematic sessions, as follows:

**Opening session:** Modernising statistical systems: opportunities for regions and cities

- 1) Housing indicators in regions and cities. The role of administrative data, crowdsourcing and web-scraping
- 2) Defining functional geographies for policy making and analysis: methods, standards, data
- 3) Measuring accessibility to services in all places. Geospatial information, open platforms
- 4) Business clusters: mapping exercise and visualisation tools for regional development policy

**Concluding remarks** on lessons learned and next steps

Thanks to the discussions in the different thematic sessions, the workshop will aim to address a few overarching questions, which are the following.

- a. What does “modernisation of statistical systems” mean? What are the pillars of such a modernisation?
- b. How can innovation in the data ecosystem create new opportunities for data on regions and cities?
- c. What is the contribution that the community of experts on regions and cities can bring to the modernisation of statistical systems?

## Expected outcomes

The workshop will contribute to shape future research initiatives discussed within the WPTI and support the programme of work of the Regional Development Policy Committee. Based on the documents and presentations by the participants, workshop proceedings will be prepared in a consolidated document in early 2019, with a synthesis and conclusion for each topic.

# Agenda

## 9h00-9h30 **OPENING SESSION: Modernising statistical systems: opportunities for regions and cities**

**Welcome:** Alessandro Alasia, Statistics Canada and Rudiger Ahrend, OECD.

**Opening presentation:** *The importance of data innovation for analysis on cities* - Peter Elmlund, Ax:son Jonson Foundation.

## 9h30-11h00 **SESSION 1: Housing indicators in regions and cities: affordability, quality, characteristics**

- Speakers:**
- *Measuring local rent prices using linked big data* - Nigel Henretty, ONS, United Kingdom.
  - *Housing policy issues and relevance of Tax Authority data to produce house prices statistics in Portugal* - Francisco Vala, INE, Portugal & Jorge Malheiros, CEG-IGOT-University of Lisbon.
  - *Open data on buildings with satellite imagery processing* - Jubal Harpster, Bing Maps Team, Microsoft.
  - *Visualisation and mapping of building open data* - Mikel Maron, Mapbox.

### Abstract

Housing is a policy domain where cities and local governments can have an important role. The characteristics and quality of housing as well as its affordability are crucial elements of the material conditions and well-being of people. Therefore, housing characteristics should be measured with appropriate indicators in order to ensure more effective urban and regional policy. However, while national averages are available practically everywhere and often in long time series, international comparisons at detailed and policy-relevant geographical scales (i.e. regions, cities and urban agglomerations) are hardly available.

Data innovation can help in this respect. Building registers can be updated and integrated through geo-spatial information and crowdsourcing methods. Similarly, housing prices are increasingly monitored at granular geographies through administrative registers of transactions. Web scraping can represent another relevant source for timely information on housing prices.

This session will bring together experiences throughout OECD countries in producing housing statistics at a granular geographic scale and will illustrate how data innovation could help us move forward in this data domain.

### Questions to be addressed:

- What is the potential of administrative data and web-scraping to measure housing prices at the local level?
- How to visualise and assess the characteristics of the housing stock in a geographically granular way?
- How to best integrate building registers with more unconventional data (i.e. crowdsourcing)?

## 11h15-13h00 **SESSION 2: Functional areas in all territories: data sources and methods**

- Speakers:**
- *Multisource data for Labour Market Areas* - Luisa Franconi, Istat.
  - *Defining functional areas in Canada through open source packages* - Peter Murphy, Statistics Canada.
  - *Using mobile phone data to understand functional geographies* - Johan van der Valk, Statistics Netherlands.
  - *The functional urban area of Prague: testing consistency with mobile phone data* - Jiří Čtyroký, Prague Institute of Planning and Development, Czech Republic.

### Abstract

Regional development policy must rely on sound and reliable statistics. One crucial aspect of such statistics is the geographic unit of analysis. The need of granular and meaningful geographies for analysis and policy underlies the creation of several concepts, such as metropolitan areas, labour market areas, daily urban systems or, more generally, "functional areas". These concepts have been used extensively in OECD countries with the purpose of overcoming the limitations of administrative boundaries to better understand local economies. Nevertheless, the potential of using new sources of data has been only partially addressed. In addition, many of the existing definitions of functional areas tend focus on cities and their surrounding commuting zones. The functional organisation of space in less urbanised regions is something that still needs to be studied in depth from an international comparative perspective.

In recent years, some OECD countries have implemented functional areas for the entire national territory. The concept is mature enough to deserve some discussion on pros and cons of existing methods and data and on possible recommendations for increased comparability across countries. A standard definition of functional areas can significantly enhance our understanding of rural-urban linkages, improve the quality of labour market statistics for small jurisdictions, and facilitate collaboration of small municipalities outside the areas of influence of cities. For this type of analysis, in addition to the information available through Census, survey data, and administrative records, new sources of data, such as mobile phone and grid-level data, can add insights on how the movement of people in space determines new geographies of labour market, consumption or leisure.

This session, through country experiences focuses on comparing functional areas across countries. Elements of discussions will include possible harmonisation of methodologies, open-source tools to foster harmonisation and differences observed between functional areas based on conventional and less conventional data sources.

### Questions to be addressed:

- What methods would be most appropriate to compare functional areas throughout OECD countries?
- What are the data sources and methods that best allow defining functional areas throughout the entire national territories?
- To what extent does our understanding of functional areas change with respect to the type of data sources used (i.e. conventional vs. unconventional sources) or the type of linkages considered?

- Speakers:**
- *Access to services by settlement size in Europe* - Lewis Dijkstra, European Commission.
  - *Using Google data to measure access to amenities in cities* - Talia Kaufmann, OECD.
  - *Measuring walkable neighbourhoods using new open data* - Claudia Baranzelli, EC Joint Research Centre.
  - *Measuring access to services in French regions* - David Levy, INSEE.
  - *Access to services, recent innovations in the Netherlands* - Wim Vosselman, Statistics Netherlands.

### Abstract

Social and economic inclusion has become a pillar of the economic policy in several member countries. The extent to which many types of public services are physically accessible to individuals located in different places is a crucial aspect of people's well-being. Citizens and policy makers are increasingly calling for indicators at different territorial levels on the extent to which services are accessible.

In this domain, elements of modernisation relate to new data source and new open source platforms that could be used to generate a standard set of accessibility measures. The increasing availability of geo-referenced information on services (i.e. schools, hospitals, public transport infrastructure, amenities, etc.) as well as the availability of open information on the location of people at regular and detailed geographies (i.e. population grid data) and on transport infrastructure provision (i.e. OpenStreetMap and Google Maps) make it possible to measure access to services in a consistent way across countries. At the same time, all these tools provide an opportunity to overcome the rigidity of administrative geographies and to adapt any measure of physical accessibility to a customised unit of analysis.

### Questions to be addressed:

- a) What characteristics should accessibility indicators have in order to best support policy decisions?
- b) What sources of data allow the highest reliability of international comparisons and accuracy of the measures?

- Speakers:**
- *Regional competitiveness and clusters* - Christian Ketels, BCG Henderson Institute.
  - *Business Clusters: the experience of Mexico* - Oscar Gasca Brito, INEGI, Mexico.
  - *Clusters and regional development: a practitioner perspective* - Ksenia Shadrina, US Department of Commerce.

### Abstract

Firms are often concentrated in space and, including through self-organisation and policy support, they are connected through local linkages and spillovers. OECD member countries have a long tradition of declared or undeclared industrial policies that support businesses that are geographically concentrated in space. Similarly, long academic tradition has elaborated on concepts of industrial districts, and industrial clusters (in particular of SMEs), among others.

From a data perspective, measuring clusters and understanding their performance presents many challenges. Nevertheless, in recent years, some OECD member countries have been carrying out initiatives on cluster mapping portals. Data on geographic clusters and their internal firm dynamics can help understand how economic activities play out in places and inform policy action for regional development.

Mapping business clusters and being able to measure their performance with a regional perspective requires granular data both in terms of geography and sectors. Well-developed, harmonised and accessible business registers would help a robust cluster mapping as well as sound international comparisons of clusters, their performance and business dynamism.

### Questions to be addressed:

- a) What are the most appropriate methodologies and data sources to map business clusters at subnational level consistently across OECD countries?
- b) How to overcome limitations in the access to business registers and firm micro-databases?
- c) How to measure the performance and dynamism of business clusters?
- d) What is the best way to disseminate/visualise the local nature and performance of business clusters?

## CLOSING REMARKS

The closing remarks will recap the main outcomes of the previous session and possibly highlight future steps to be taken in for each topic.