United Nations ECE/CES/2016/21

Economic and Social Council

Distr.: General
12 February 2016

Original: English

Economic Commission for Europe

Conference of European Statisticians

Sixty-fourth plenary session
Paris, 27-29 April 2016
Item 6 of the provisional agenda
Geospatial information services based on official statistics

Statistical and Geospatial Information – an Australian perspective on challenges and opportunities

Note by the Australian Bureau of Statistics

Summary

This paper examines the use of location enabled information for place based decision making and the subsequent need for integrated statistical and geospatial information. It comments on this in the context of the needs of users to analyse and visualise this data for 'place based decision making', the international cooperation occurring on this topic and the need to further these efforts.

The paper is presented to the Conference of European Statisticians seminar on "Geospatial information services based on official statistics" for discussion.

GE.16-02013(E)

I. Introduction

1. Over the past 10 years, the demand for statistical information that is linked to a location has accelerated rapidly, with demand coming from users in all sectors: government, commercial, not-for-profit, academia, and citizens. Growth in demand for this location information, including geospatially enabled statistics, has occurred because people and organisations want to gain insights into their populations or topic of interest in increasingly greater detail. An important facet of this insight relates to the location of the populations under consideration; in many instances the insights are specifically required at the local level. Moreover, rapid increases in the mobility of people, products and services are driving a need to develop greater insight into the complex dynamics within and between regions. Past assumptions about where people live, work, play or access services no longer hold in the modern world.

2. More recently, insights derived from the flows of people and business transactions between locations, as well as more traditional regional demographic and business statistics are being used in a growing trend towards "place based decision making". In order to gain the insights required for these place based decisions, the data being used must be integrated across a variety of disparate information sources, including information from statistical, administrative, geospatial and big data sources. Therefore, users have not only increased their demands for location information, they are also demanding simpler integration of data across the various data sources used in their analysis.

3. The targets and global indicator framework for the Sustainable Development Goals (SDGs) also include the requirement for disaggregation of many indicators by 'geographic location'. The need for a new era in location information within national and international statistical systems has therefore gained even greater importance. Through the SDG process, the statistical community is beginning to understand more clearly that the full range of geospatial information and earth observations are able to provide new and consistent data sources and methodologies to provide information for the SDG indicators and to enrich official statistics.

4. A final factor that is enabling the growing use of location information and place based decision making has been developments in geospatial technologies. These developments now mean that geospatial visualisation (i.e. mapping data) and geospatial analysis have moved, from being tools that were the exclusive domain of experts, to being tools that are accessible to a wider range of users, in a wider range of contexts. From a user perspective:

(a) Geospatial Visualisation is a very powerful tool that makes information more accessible by placing it a real world context. These visualisations can also provide new insights by showing spatial relationships in the data.

(b) Geospatial Analysis provides access to additional insights when combined with standard statistical analysis methods; statistical functions can be combined with any number of geospatial operations that incorporate concentration, proximity and distance, and connection and obstruction.

5. Geospatially enabled statistical and administrative data is one form of location information, and its growing use to inform place based and broader decision making has attracted strong interest from National Statistical Offices and the International Statistical Community. As a result it is now readily accepted that integrating statistical and geospatial information is critical for:

(a) local, national and global decision making processes;

(b) measuring and monitoring sustainable development goals;

(c) supporting data sharing between institutions;

(d) promoting investment and capability building in geospatial and statistical information;

(e) building institutional collaboration between geospatial and statistical communities; and

(f) for unlocking new insights that would never have been possible by looking at socio-economic or geospatial data in isolation.

6. To progress these goals the United Nations Statistical Commission (UNSC) and the United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM) has formed the United Nations Expert Group on the Integration of Statistical and Geospatial Information. Within the Expert Group, members of the global and national statistical and geospatial information communities are working together to ensure greater collaboration on these goals. The Expert Group has proposed a Global Statistical Geospatial Framework to support consistent geospatial enablement of socio-economic data, which also allows this data to be integrated with geospatial information infrastructure[1].

II. Place based decision making - data and geospatial tools

A. Place based decision making

7. The various sectors within countries – government, commercial, not-for-profit, academia, and citizens – require information for important place based decisions. These place based decisions will also have a strong role in ensuring countries' progress towards achieving the Sustainable Development Goals (SDGs), while the information required for this decision making will also be an important input to reporting on the SDG targets and indicators.

8. Engagement by the Australian Bureau of Statistics with a wide range of regional data users has found that place based decisions cover a wide variety of domains, including:

(a) Local area economic development - planning for local jobs and economic growth, or managing decline or change in the structure of local economies.

(b) Service planning for locations - planning infrastructure, transport, housing, health, education, recreation, culture and many other services.

(c) Business location planning - most businesses use small area data to understand their consumer markets, labour markets and decide business locations.

(d) Addressing locational disadvantage - policy interventions or service delivery dealing with disadvantage concentrated in locations.

(e) Liveability of places - developing the quality of life and liveability of places for residents, workers and visitors; as well as to attract investment in infrastructure and economic activity.

(f) Local community resilience – supporting the ability of local residents and businesses to adapt to environmental disaster, economic shock or downturn, or other change.

[1] http://unstats.un.org/unsd/statcom/doc15/2015-37-GeoInfo-E.pdf

(g) Local community wellbeing - promoting healthy and socially active communities.

(h) Local social cohesion – supporting positive community interaction and civic participation across diverse and changing communities, while preventing social disorder.

B. Disparate data sources

9. The information required to make informed decisions across these domains requires information on a wide variety of topics and from a disparate range of sources. Within National Statistical Offices (NSOs), and across other information agencies, each of these specific topics has often been viewed as separate fields of statistics or information. In many cases, the data required by users across these topics does not use common classifications or frameworks, including the use of consistent regional geographies. However, because of the importance of location for place based decision making, regional geographies are often the point at which much of this integration must occur and so this is a central focus of concern by users.

C. Common place based questions

10. Despite the variety of applications, these place based decisions usually relate to a few basic questions. By examining these questions it is easy to identify and understand the main problems faced by users in integrating data and how these users employ geospatial tools to enhance their decision making. The current and potential roles of the international and statistical communities can also be identified.

D. Who is here?

11. The size of the population and the characteristics of the people who live or work in a region is a fundamental information component required to gain insights to a place and begin to formulate decisions. This information has been a very traditional domain of the NSO and has been a strong element of the traditional census; however, increasingly more detailed data on these topics can be derived from administrative datasets (when this data is adequately geospatially enabled).

12. From a data integration perspective, it is important that the data is geospatially enabled in a consistent and accurate manner and that this information be made available for regions that have a common basis, preferably for consistent small area regions. This allows information to be compared on a consistent geographic basis and to be flexibly aggregated to meet user needs. Consistent and accurate geospatial enablement of data is best supported by a consistent framework, such as Australia's Statistical Spatial Framework[2], which is underpinned by nationally consistent data and supported by common geocoding methods. Use of a common regionalised geography for dissemination may take the form of a hierarchy of nested regional geographies, such as those contained in the Australian

[2] http://www.nss.gov.au/nss/home.nsf/pages/Statistical%20Spatial%20Framework%20Homepage (see also Appendix F, this paper)

Statistical Geography Standard[3]; or by using nested grids, such as those being pursued in Europe[4] and Australia[5].

13. Data showing 'who is here' is commonly used in Geospatial Visualisation (i.e. mapping), usually in the form of choropleth (shaded area) maps. This data is also increasingly being used in far more elaborate, often interactive, visualisations. Some examples are shown below in Appendix A.

E. What is here?

14. A region is often defined by what is located within it – physical infrastructure, urbanisation and land use, the natural environment, businesses and services. However, obtaining data to give accurate, up to date information can be complex. The data that describes these physical features is principally the domain of the geospatial community and the base of this information is often supplied by National Mapping Agencies (NMAs). Because of the different data models and standards used with geospatial data, accessing and integrating this geospatial data with statistical information can often be a time consuming, technical task.

15. Data on the built and natural environment is often integrated in standard mapping products by National Mapping Agencies, and are often produced as part of National Spatial Data Infrastructures (NSDIs). However, the capacity for data integration is not necessarily a key feature of all NSDIs. National Mapping Agencies are now beginning to prioritise data integration within NSDI's, such as Australia's Foundation Spatial Data Framework[6], and the geospatial standards community are implementing a range of technical solutions. Many NSDIs also incorporate statistical and administrative geographies; when these geographies are integrated within an NSDI, this can greatly simplify integration efforts.

16. Much of the information on the built and natural environment is captured by remotely sensed or other Earth observation techniques. Earth observation information ranges from local to global coverage, and is very large in volume, resolution and frequency. The geospatial community also often makes use of large and often complex grids (known as raster grids) to disseminate this data. These grids are typically based on a square array, such as a 1 kilometre by 1 kilometre grid, or a 25 metre by 25 metre grid. The variety, coverage, frequency and scale of this data mean that Earth observation data can be used to enrich official statistics and place based decision making.

17. The United Nations Expert Group on the Integration of Statistical and Geospatial Information is encouraging communication and collaboration between NMAs and NSOs. This collaboration will encourage greater integration of data and prioritisation of mapping and Earth observation datasets that support the work of NSOs, such as mapping required for Census collection activities, as well as the needs of statistical and administrative data users.

18. Maps of the built and natural environment can assist in understanding the nature and extent of geographic regions, and provide important context for the socio-economic data that is associated with these areas (see first map example in Appendix B). That said, the greatest benefit that can be derived from this data on the built and natural environment is in

[3] http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Statistical+Geography+Standard+(ASGS)

[4] http://ec.europa.eu/eurostat/web/gisco/geostat-project

[5] http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/1270.0.55.007main+features12011

[6] http://www.anzlic.gov.au/foundation_spatial_data_framework

Geospatial Analysis, where this data is used in a Geographic Information System (GIS) to apply geospatial and statistical functions to derive location or regional insights from the data. The second map below in Appendix B contains a visualisation of data obtained from Geospatial Analysis showing where land development is occurring through identification of the subdivision of property boundaries for urban development. This, and more complex analysis, is where the true power of geospatial technologies and information exists.

G. How do people, businesses and services inter-relate across regions?

19. People, businesses and services participate in a complex interplay of local and global connections. Information is needed to describe the important local connections within or between regions or cities within a country, as well as those that occur more broadly. These connections can include: population flows and mobility, employment and consumption patterns, movement and portability of goods and services, and business supply chains. Transactional data may be used to provide insight to these topics, and big data sources are increasingly also being assessed for use.

20. Due to the nature of transactional and other big data information sources, they are often unstructured or use classifications or frameworks that are not useful for data integration or analysis. This often makes the use of this data with other sources a significant challenge. Allocating this unstructured data to a location is often also problematic; as the locational elements are often derived and, therefore, imprecise or incomplete (e.g. mobile phone tower data). The United Nations Expert Group on the Integration of Statistical and Geospatial Information has acknowledged that little has been done to systematically assess the locational aspects of big data and have identified this as a topic to be examined in the coming years.

21. Analysis of inter-relationships across regions, requires input of information about the ‘who’ and ‘what’ of regions, and also requires this data to be analysed and understood in the context of the relative locations and connections between these various elements. These insights can only be obtained using both statistical and geospatial analysis methods, most commonly using Geographic Information Systems (GIS). As an example, the first map in Appendix C below shows how use of motor vehicle transport (derived from the Australian Census) interrelates with public transport availability. Further analysis of this data would be possible using GIS to assess how distance from public transport infrastructure affects usage.

H. What value or benefit do people or businesses derive from places?

22. Many people and businesses move to regions that best fit their needs or wants, and the value they derive from those areas affects how they might flourish and develop. This is driven by a range of complex factors which only detailed analysis can seek to gain insights about. Many disparate aspects influence this topic for people and businesses. For population, the following aspects are thought to be important; social and family support and cohesion, employment and educational opportunities, service availability, and amenity and housing affordability. For business, the following are thought to be influential: access to labour and markets, availability of complementary services and production, and increasingly the ability to innovate and adapt.

23. The ability to understand how these factors influence social or economic development is strongly associated with the inter-relationships identified in the previous question, but requires the addition of some element of ‘value’ to be input to these relationships. Geospatially enabled data can input some of this value information into

analysis through the location or regional attributes in the data. These data can include: land value and amenity information, consumer and labour market information, indicators of community cohesion, or transport time or cost values derived from transport networks. Again the ability to integrate data from different sources is fundamental to undertaking this type of analysis. The map in Appendix D shows a grid of land value information that could be a key input to this type of analysis.

I. What will this place be like in future?

24. Understanding how a region or place might develop is often fundamental to the people and organisations located there, as well as the organisations that provide services to that region. Governments, businesses and communities cannot simply plan for the 'now'; for planning they need well framed projections that can support different scenarios and parameters. These include demographic, economic and spatial data and models that can produce useful outputs at the regional level.

25. Through the use of common standard and frameworks, these models can integrate data from a range of sources, including information from other countries to further test and refine the models to improve their predictive power. The United Nations Expert Group on the Integration of Statistical and Geospatial Information has identified the use of grid based statistics as a means of enhancing international comparability of information from a range of sources. Currently a model for a common international grid is under consideration internationally; the Discrete Global Grid System standard[7] under development within the Open Geospatial Consortium.

26. The maps in Appendix E show gridded population information in relation to transport networks in two cities – Melbourne, Australia and Paris, France. Models developed and derived from this and other related information from cities around the world could be used to evaluate the potential impact of different forms of transport infrastructure on the population density and characteristics of developing cities.

III. Issues and challenges

27. There are many challenges that will need to be overcome to ensure we can make effective use of geospatially enabled statistics and other location information for place based decision making, and to help measure and monitor the Sustainable Development Goals. Geospatial standards and frameworks have been highlighted as a key issue in this paper as these are vital to ensuring efficient access to and use of this information. The Australian Government, with Australian Bureau of Statistics (ABS) as a key partner, are developing access to relevant public sector data[8], and ABS will collaborate to develop new statistical solutions for future output of regional statistics and applied geospatial analysis. While researchers and statisticians are currently working on the challenges involved in generating meaningful regional statistics using the data sets they have immediately available to them.

[7] http://www.opengeospatial.org/projects/groups/dggsswg
[8] https://www.dpmc.gov.au/pmc/publication/public-sector-data-management-report

A. Global Statistical Geospatial Framework

28. The importance of integrating geospatial information with statistics has been acknowledged by the United Nations Statistical Commission and the United Nations Committee of Experts on Global Geospatial Information Management. This was reflected in the joint decision to establish the United Nations Expert Group on the Integration of Statistical and Geospatial Information with a primary mandate to develop a statistical geospatial framework as a global standard for the integration of statistical and geospatial information.

29. The ABS recognized some time ago the need for and challenge of better integration of geospatial and statistical information, and responded by developing the Statistical Spatial Framework. This framework provides Australia with a common approach to connecting socio-economic information to a location, and improves the accessibility and usability of geospatially-enabled information.

30. The Statistical Spatial Framework consists of five principles that are considered essential for integrating geospatial and socio-economic information. These are: authoritative geospatial infrastructure and geocoding; data management; common geography; statistical and spatial metadata interoperability; and accessible and useable geospatially enabled information. Attachment 1 details the vision, goals and outcomes of these principles, as well as the Australian implementation of these generic principles.

31. The Expert Group concluded that the Statistical Spatial Framework developed by the ABS is a practical, principle-based framework that is flexible enough to be customized to suit national contexts. Moreover, the framework is an effective mechanism to improve the geospatial enablement of statistical information, so that it can be effectively integrated with data from national and international geospatial and statistical information infrastructures to support analysis and decision making. The framework is also recognized as a tool for encouraging thinking and discussion between geospatial and statistical communities, especially as institutions modernise and transform their infrastructure and technology.

32. The Statistical Spatial Framework and its global parent, the Global Statistical Geospatial Framework, provide a bridge between the official statistical and geospatial information. Statistical units are geocoded using a spatial addressing layer. Data is then released for common geographic boundaries in the boundary layer. This supports integration of data across diverse socio-economic and geospatial information resources. As more data is released on a common geography, it creates an increasingly valuable network of integrated information.

B. Partnership between National Statistical Offices and National Mapping Agencies

33. When establishing the United Nations Expert Group on the Integration of Statistical and Geospatial Information, the United Nations Statistical Commission and the United Nations Committee of Experts on Global Geospatial Information Management identified the need to reach out and developing best practices by bringing together both statistical and geospatial professional communities. Through the Expert Group's work the two communities have started the process of uniting together to integrate statistical and geospatial information. This will empower countries to better measure the progress of their societies and economies, to unleash the power of statistics and location information in place based decision making, and to help measure and monitor the Sustainable Development Goals.

34. Key to this journey of collaboration has been a willingness and desire to cooperate. For example, through the development of the Sustainable Development Goals indicator framework the statistical community is beginning to understand more clearly the value of geospatial information and earth observation data. Particularly, that these data types are able to provide new and consistent data sources and methodologies to integrate multiple 'location-based' variables to support, inform and enrich official statistics, place based decision making and the SDGs indicators. They are able to fill data gaps and/or improve the temporal and geospatial resolutions of data, by incorporating information from various sources, particularly those related to the environment.

35. An example, of this developing relationship is between the ABS and Geoscience Australia. The ABS recently released a new strategic Engagement Plan based on three engagement principles: nurture, understand and collaborate. These principles target three outcomes:

(a) shared benefit through open relationships that recognize others play a significant role in Australia's statistical system;

(b) shared vision by committing ABS to understanding the current and future policy landscape, and the pressures, priorities and pain points of stakeholders; and

(c) shared solutions by committing ABS to work together with partners in purposeful ways to inform discussions and decisions, and improve Australia's statistical system.

36. The relationship between Australia's official statisticians and geoscience experts has adopted these three principles and is producing productive results through a growing partnership. The joint representation on UN-GGIM is nurturing professional relationships. We also listen and seek to understand the pressures, priorities and pain points of our two communities and institutions. For instance, the Chief Executive Officer of Geoscience Australia was the guest speaker at an inaugural ABS Transformation Showcase event. Geoscience Australia's CEO shared their experiences of institutional transformation, informing ABS as it embarks on its own transformation. And finally, we are sharing solutions through many collaborative efforts underway, especially those aimed at integrating statistical and geospatial information.

37. One example of this collaborative work is a joint project that is exploring the use of Earth Observation data in the Australian Geoscience Data Cube. This project will develop new and useful statistical products for agricultural, environmental and urban statistical applications. The organisations are combining their expertise, tools and experiences in geospatial information, Earth observations data and the compilation, analysis and dissemination of official statistics. The aim is for ABS and Geoscience Australia geospatial experts, working with the more traditional official statisticians, to bridge the methodological gap between the two professions. This approach has proved successful in recent projects producing the Population Grid and the Land Accounts, and the ABS is hoping to replicate it with this new, large, high-value data source: the Australian Geoscience Data Cube.

IV. Conclusions and recommendations

38. This paper concludes with the following recommendations:

(a) NSOs support users undertaking analysis of location information and making place based decisions.

(b) NSOs support the production of geospatially enabled data, and implement and encourage the use more broadly of geospatial standards and frameworks for this data.

(c) NSOs engage in the international harmonisation of location information, and associated standards and frameworks.

(d) NSOs seek to make fuller use of geospatial information and earth observation data to enrich official statistics.

(e) NSOs support the work of the United Nations Expert Group on the Integration of Statistical and Geospatial Information to:

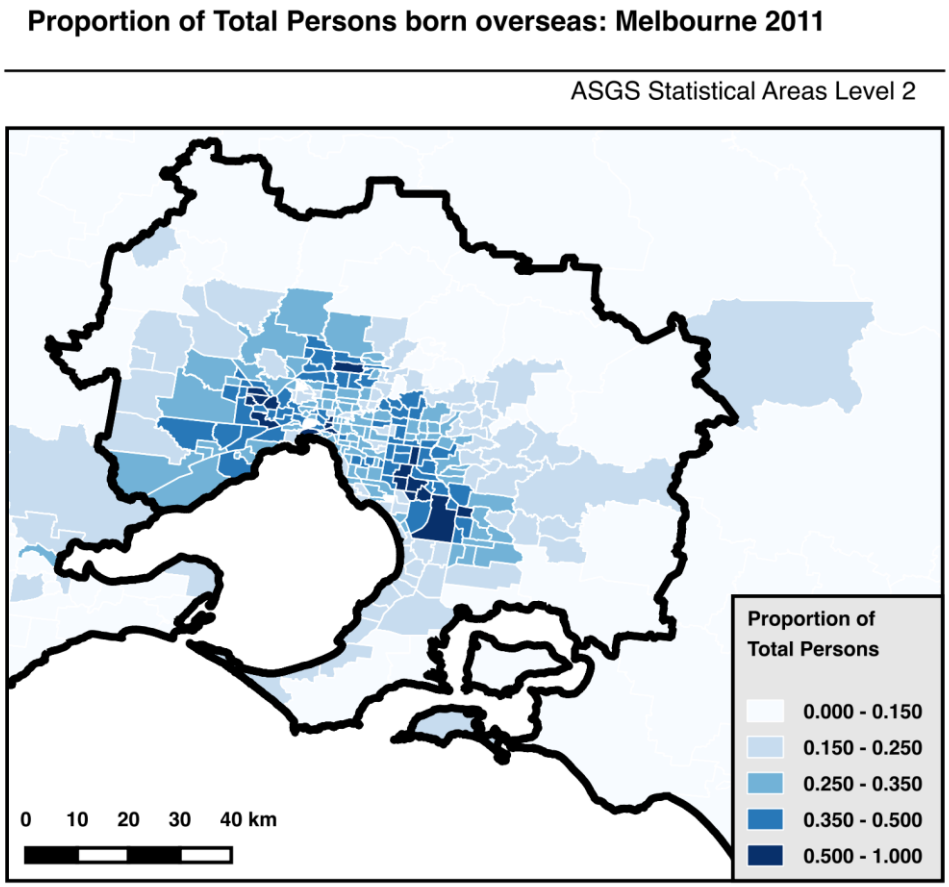
(i) Adopt a Global Statistical Geospatial Framework, and

(ii) Encourage the collaboration and partnership between of NSOs and NMAs.

Appendix A

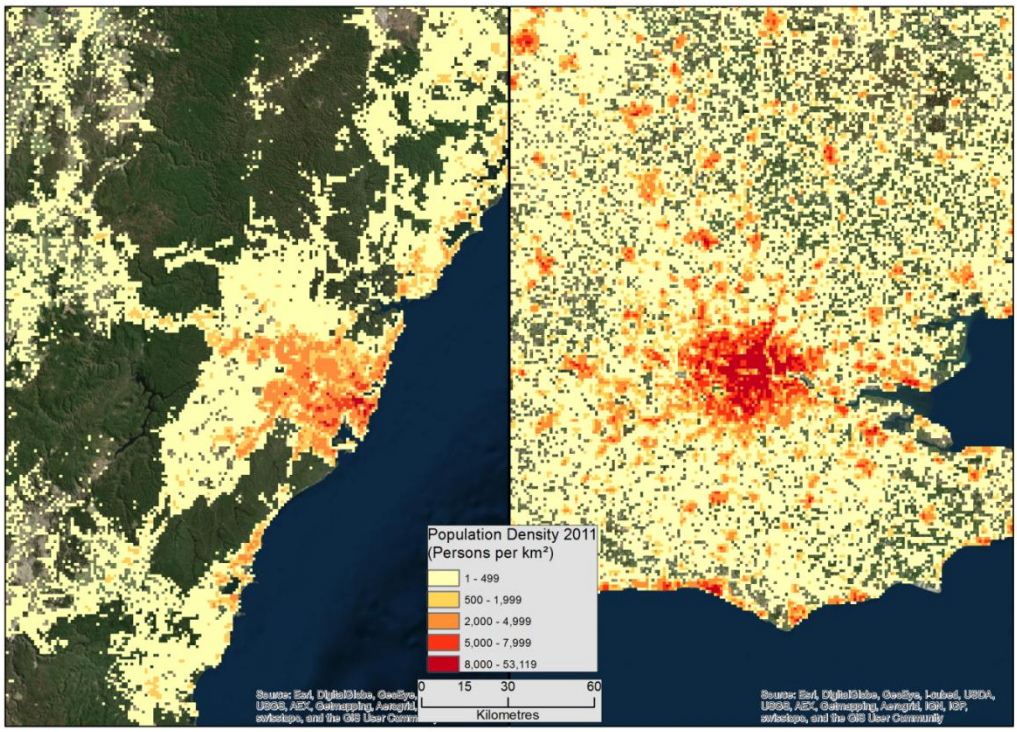
Maps - Who is here?

Map 1
Proportion of total persons born overseas: Melbourne - 2011



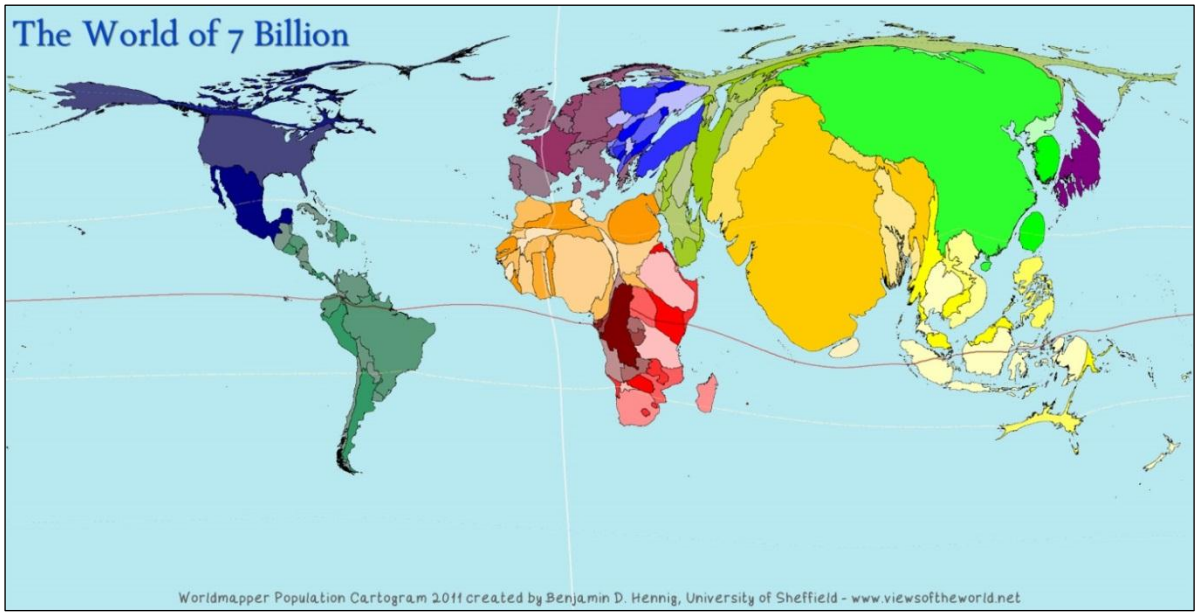
Source: ABS, 2011 Census of Population and Housing

Map 2
Population density, 1km^2 grid: Sydney, Australia and London, England – 2011



Source: ABS, 2011 Census of Population and Housing, Australian Population Grid and Eurostat GEOSTAT 2011 grid dataset

Map 3
World population cartogram, national population density – 2011



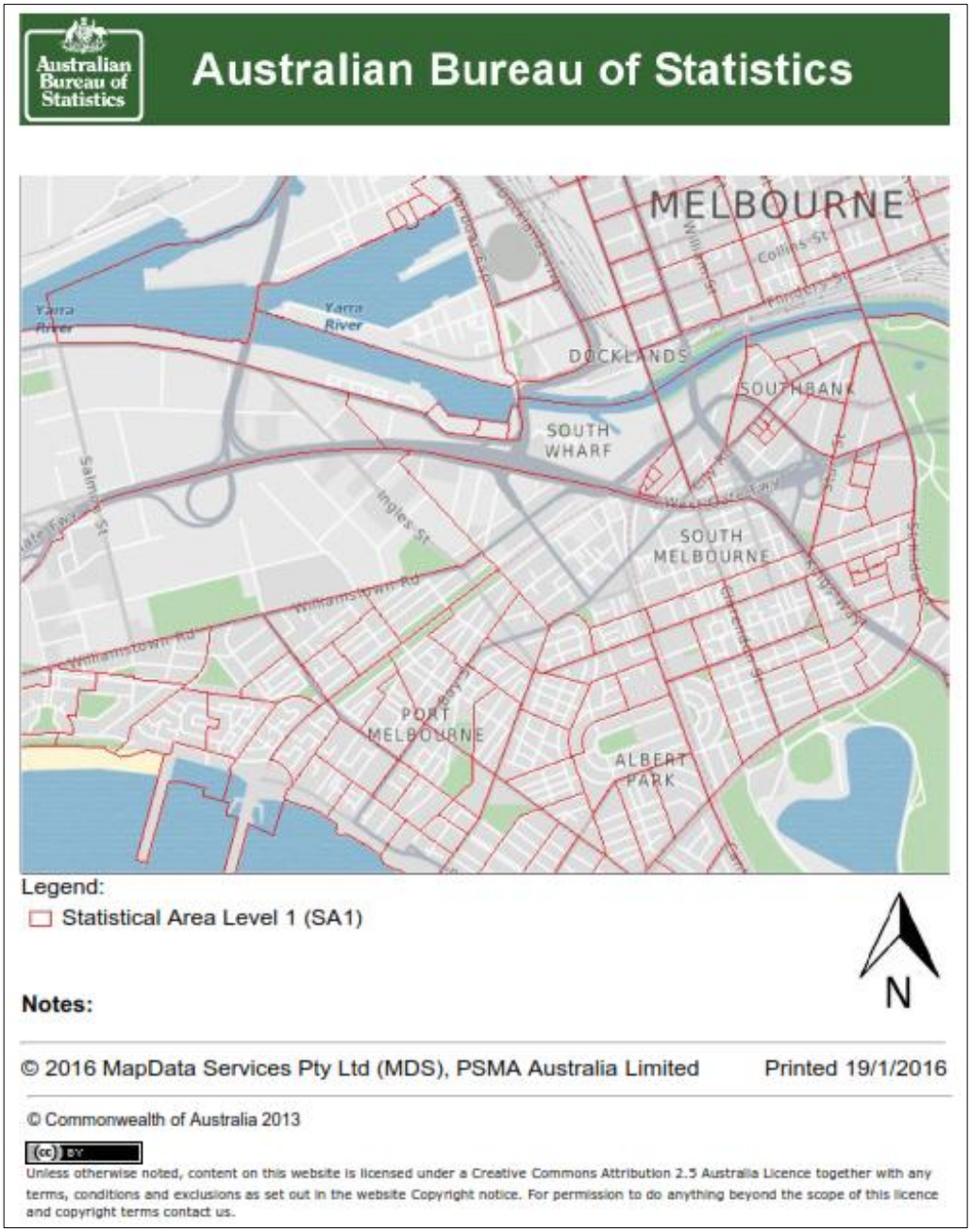
Source: Benjamin Hennig, Senior Research Fellow in the School of Geography and the Environment at the University of Oxford < http://www.viewsoftheworld.net/?p=1889>

Full visualisation link: <http://www.viewsoftheworld.net/wp-content/uploads/2011/10/WorldPopulationAnimation.gif>

Appendix B

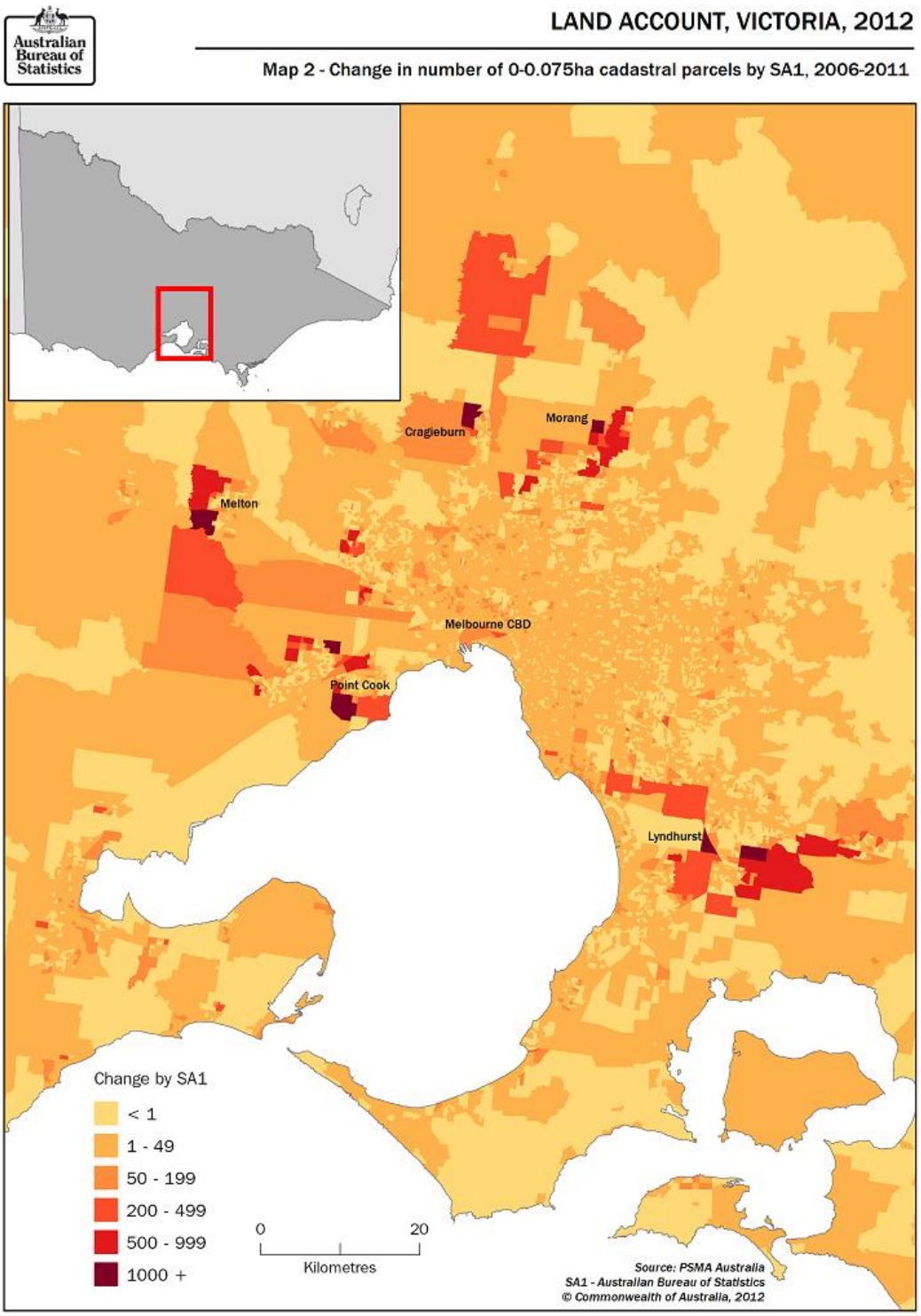
Maps - What is here?

Map 4:
2011 Australian Statistical Geography Standard Statistical Area 1 regions and base map



Source: ABS, MapData Services Pty Ltd, PSMA Australia Limited
<http://betaworks.abs.gov.au/betaworks/betaworks.nsf/dx/asgs-boundaries-online.htm>

Map 5
Change in cadastral parcels, Land account, Victoria – 2012

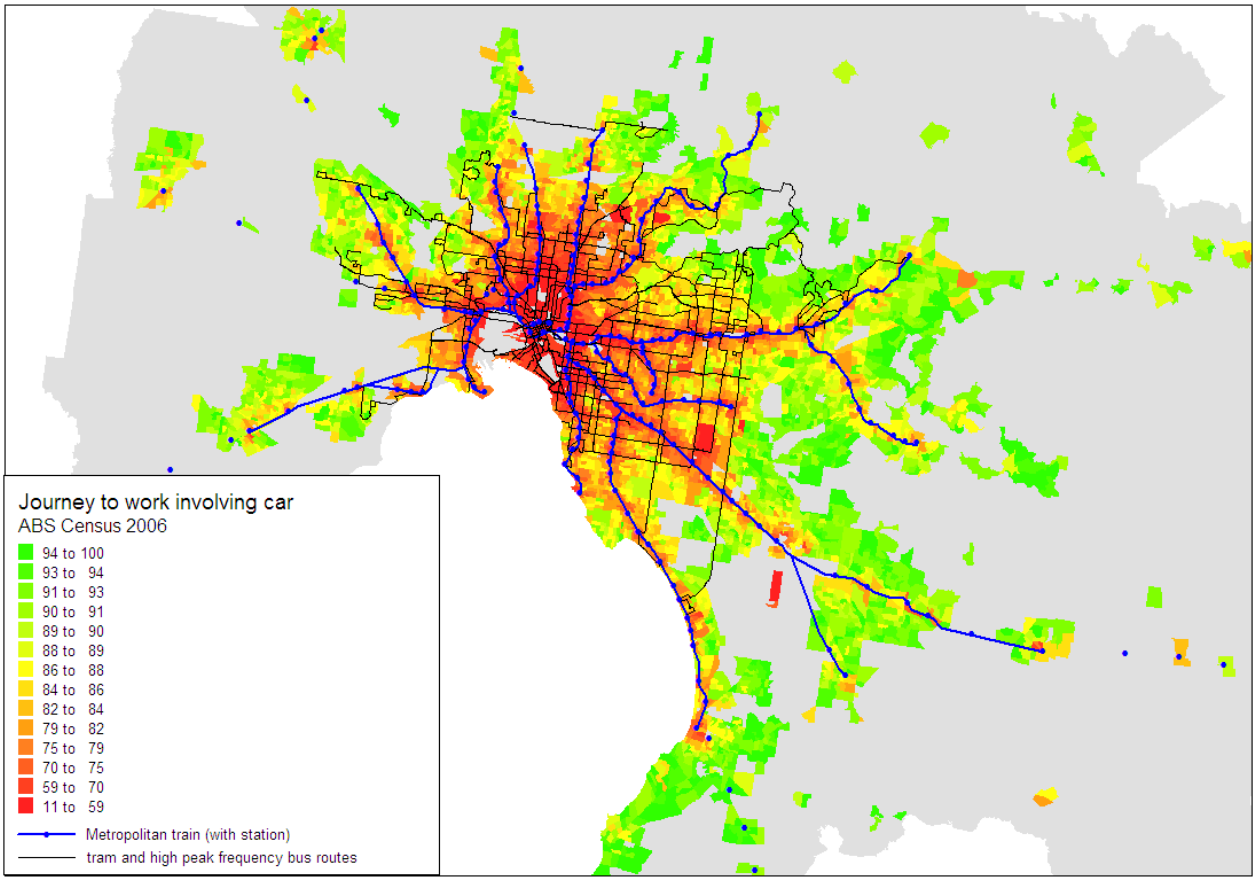


Source: ABS, Land Account, Victoria, 2012

Appendix C

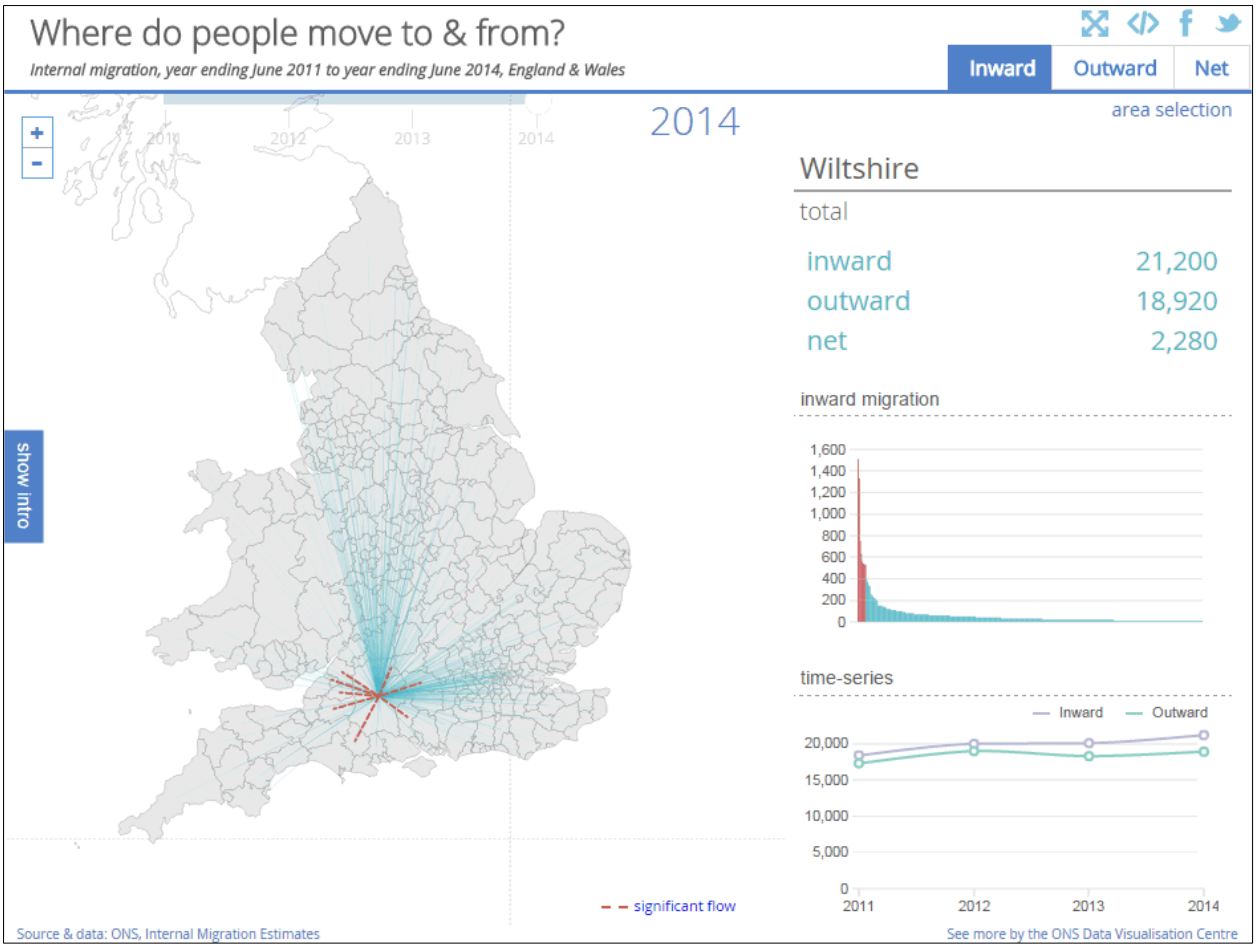
How do people, businesses and services inter-relate across regions?

Map 6
Journey to work and public transport infrastructure, Melbourne – 2006



Source: Charting Transport, 'A spatial analysis of Melbourne 2006 journey to work mode shares', <http://chartingtransport.com/2010/07/16/a-spatial-analysis-of-melbourne-2006-journey-to-work-mode-shares/> and ABS, 2006 Census of Population and Housing.

Map 7
Internal Migration, England and Wales – 2011-2014

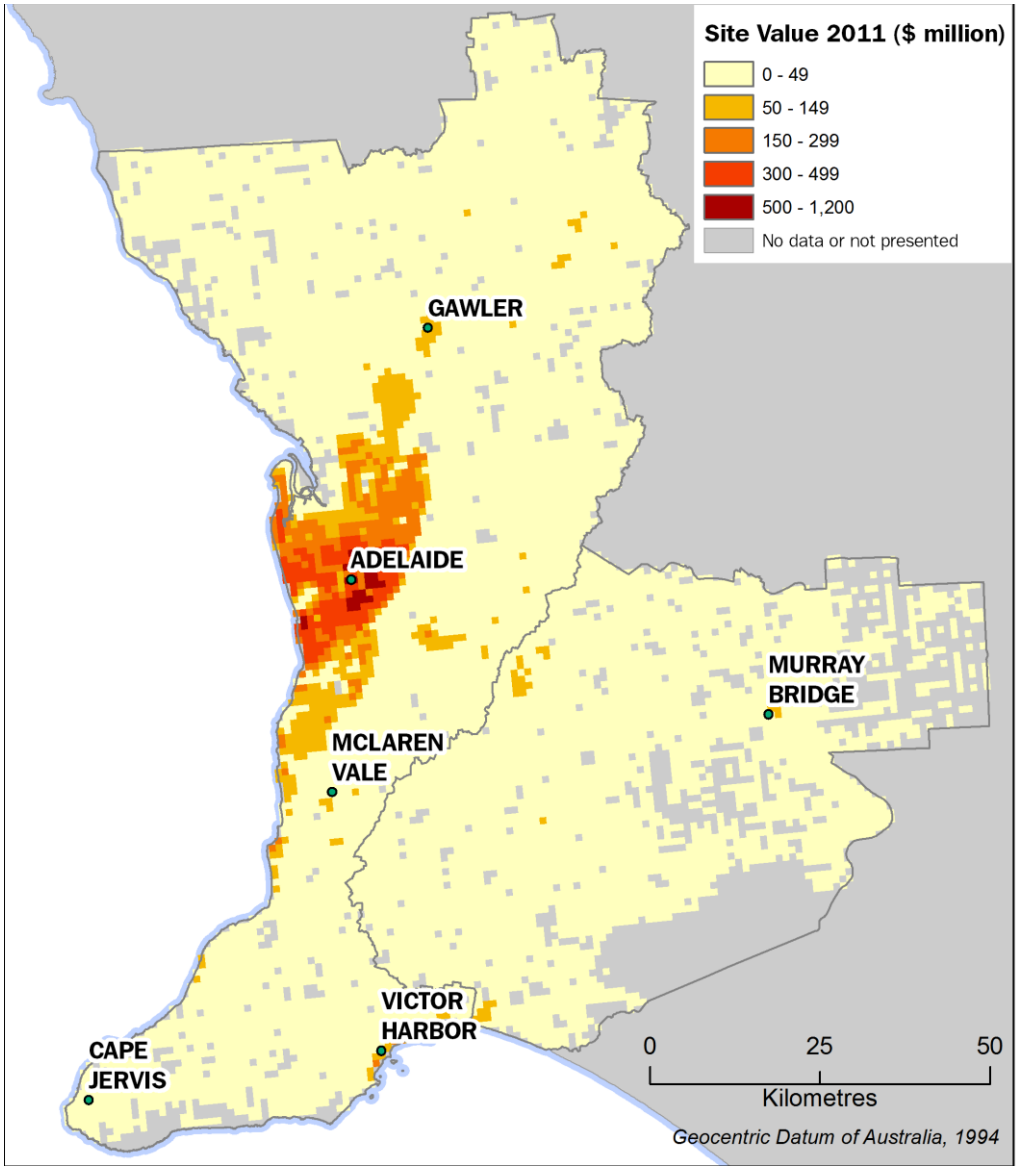


Source: ONS, Internal Migration Estimates

Appendix D

What value or benefit do people or businesses derive from places?

Map 8
Land value, 1km^2 grid, South Australia – 2011



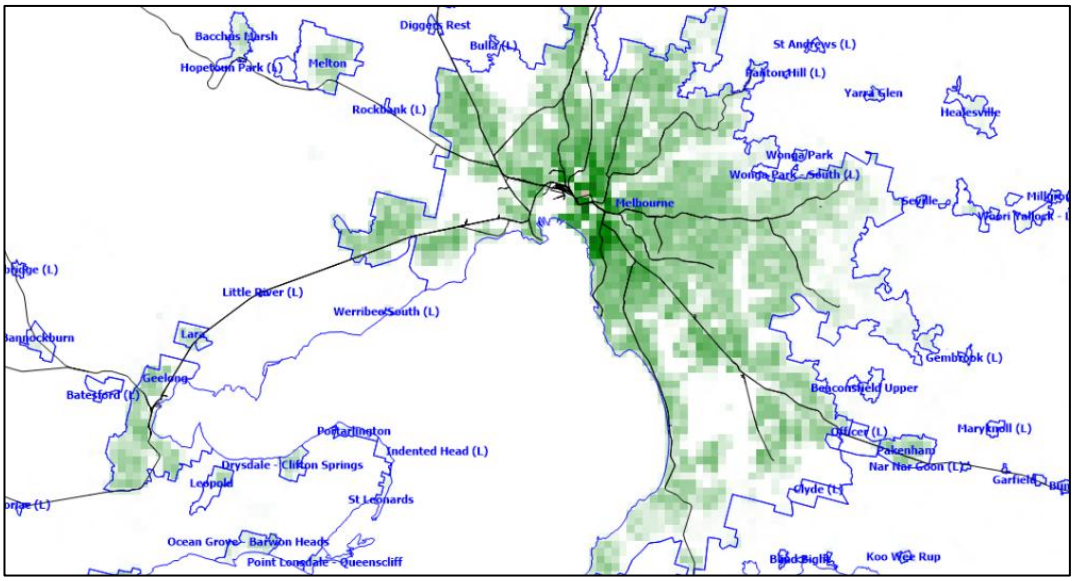
Source: ABS, Land Account, South Australia 2011 and SA Department of Planning, Transport and Infrastructure and Department of Environment.

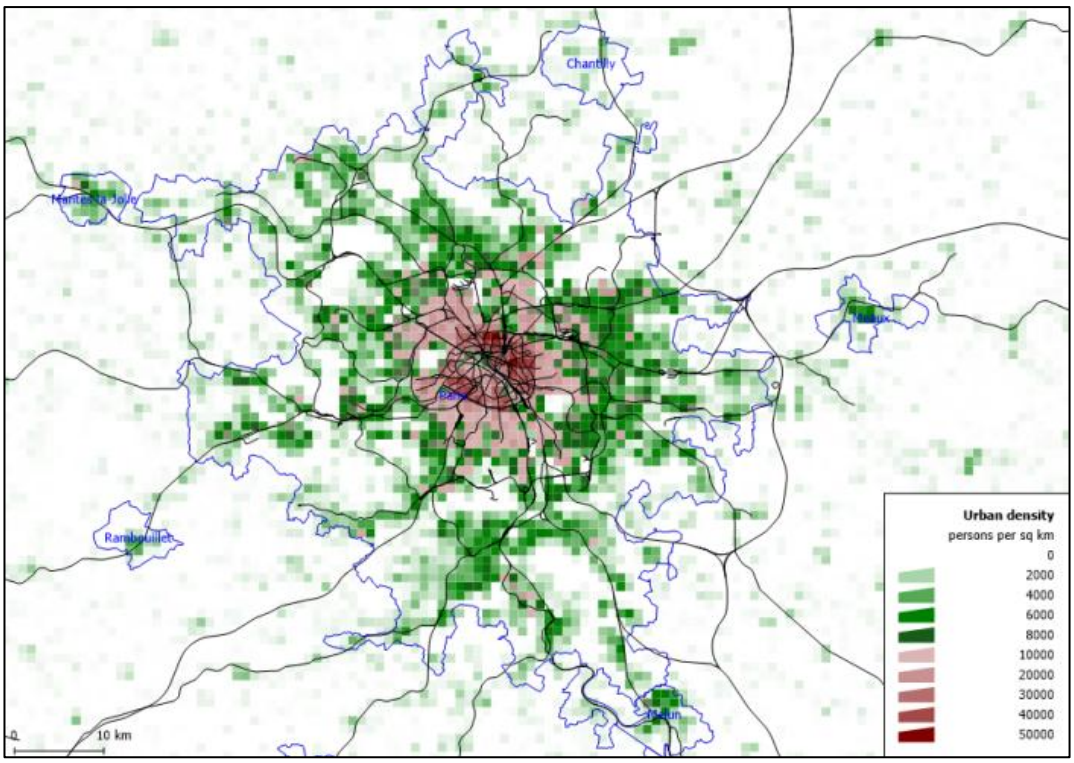
Appendix E

Maps - What will this place be like in future?

Map 9

Population Density, 1km2 grid, and Rail Infrastructure, Melbourne, Australia – 2011 and Paris, France – 2010





Source: Charting Transport, "Comparing the densities of Australian and European cities' <http://chartingtransport.com/2015/11/26/comparing-the-densities-of-australian-and-european-cities/>

Appendix F

Statistical Spatial Framework

Australia's Statistical Spatial Framework is a framework to enhance decision making by using location in a common way to allow seamless integration of administrative, statistical and spatial information resources. The following diagram outlines the vision, goals and principles embedded within the framework, and the outcomes it will achieve.

Australian Bureau of Statistics

Statistical Spatial Framework

SSF Vision	Informed decision making is enhanced by using location in a common framework to allow seamless integration of administrative, statistical and spatial information resources.				
SSF Goals	All statistical data is consistently spatially enabled		Users can discover, access, integrate, analyse and visualise statistical information seamlessly for regions of interest		
SSF Principles	Authoritative geospatial infrastructure and geocoding	Data management – geocoded unit record data	Common geographic boundaries	Interoperable Metadata	Accessible and usable geostatistics
Australian Application of SSF	Use Foundation Spatial Data and relevant National Address Management Framework (NAMF) protocols	Geocodes stored on unit records are location coordinates and ASGS Mesh Blocks. Use statistical data management frameworks.	Data is released for Australian Statistical Geography Standard (ASGS) Statistical Area regions – as a minimum	Use international statistical and geospatial metadata standards[1]	Policies, standards and guidelines support the release, access, analysis and visualisation of spatially enabled information[1]
SSF Outcomes	• Consistent address and location information • Consistent geocoding results • Consistent management of geocoding issues	• Consistent and interpretable geocode information • Simplified aggregation of data to regions • Flexibility in production of regionalised data • Effective data management, especially privacy and metadata • Clear maintenance and custodianship	• Datasets can be integrated using common geography • Use of ASGS Statistical Areas simplifies visualisation and analysis • Metadata for ASGS and other regions supports data integration and use • Conversion of data between region types is supported	• Creation, discovery, integration and use of geostatistics is supported by statistical and geospatial metadata • Use of international standards enables the application of a larger pool of technologies, and wider data access and comparisons	• Data custodians can release data with confidence • Data users can discover and access geostatistics, and undertake analysis and visualisation • Web services enable machine to machine access and dynamic linking

[1] Aspects of this component require further work.