



MONTHLY POPULATION DATA

Privacy Impact Assessment

14 July 2022



INFO by Design

It's about them - putting your customers at the centre
of how you manage their information



Crown copyright ©

[See Copyright and terms of use](#) for our copyright, attribution, and liability statements.

Citation

Data Ventures and Stats NZ (2022). Population density: Privacy impact assessment. Retrieved from www.stats.govt.nz.

ISBN 978-1-98-858333-4 (online)

Published in July 2022 by

Data Ventures and Stats NZ Tatauranga Aotearoa
Wellington, New Zealand

Contact

Data Ventures: dataventures@stats.org.nz

www.stats.govt.nz



Contents

Introduction and Purpose	4
Product Scope.....	5
Assessment of Information Type	5
Assessment of Privacy Impacts	6
Collecting or obtaining information	6
Storage, security and retention	7
Access and correction	8
Accuracy.....	8
Use and disclosure	9
Unique identifiers	9
Conclusion and Recommendations.....	9

Introduction and Purpose

1. This Privacy Impact Assessment (PIA) examines the privacy impacts of Data Ventures' Monthly Population Data product built for Ministry of Business, Innovation and Employment (MBIE) using data acquired from telecommunications companies (data providers), namely Spark and Vodafone.
2. Data Ventures is a business unit of Stats NZ, specialising in creative data projects involving brokering of data from public and private sector organisations. Its role is to develop and demonstrate innovative ways to utilise large administrative datasets to produce more timely and relevant statistics and insights for New Zealand. These insights are then utilised to help customers make more informed, data-driven decisions.
3. The Monthly Population Data product is related to Data Ventures' existing Population Density product. Population Density contains measures of population in each hour in each suburb, split by local residents, domestic visitors, and international visitors. The Monthly Population Data product will contain measures of the number of unique local residents, domestic visitors, and international visitors in each Regional Tourism Organisation (RTO) area. Further, a subset of international visitors, 'short-term international visitors' (non-New Zealand residents, staying in NZ for less 90 days) will also be provided.
4. Data Ventures and MBIE have heard from stakeholders in the tourism industry that there are still gaps in the available data about tourism, particularly domestic tourism. The Monthly Population Data product is intended to provide an easier-to-use, publicly accessible snapshot of tourism trends that stakeholders can use together with other datasets about tourism to draw insights and make decisions. The definitions of international and domestic visitors have also been tailored to tourism industry feedback.
5. This will be achieved using datasets containing mobile location estimate data, sourced from data providers. Qrious (Spark) will provide input data that contains the number of distinct devices that appeared as local, domestic, international, and short-term international in each RTO in each month.
6. Vodafone will provide the same anonymised input data from the previous Population Density product.
7. Data Ventures will then apply Stats NZ's processes for data quality, assurance, and modelling on population density. It will do the following:
 - a. Receive the data from both providers
 - b. Process the Vodafone data to create aggregates in the same form as the Spark data
 - c. Combine the two datasets to create device counts for each segment in each area
 - d. Use supplementary datasets to weight the device counts into population estimates
8. The data series will start from January 2019. The datasets are anonymised and aggregated to provide monthly snapshots of tourism activity at given places in New Zealand. This product will be publicly accessible.
9. Anticipated use cases among the tourism industry include:
 - a. Planning marketing campaigns to spread tourism across the year
 - b. Infrastructure planning
 - c. Understanding trends in tourism during the recovery from COVID-19

Product Scope

10. In scope of this PIA is Data Ventures' use of population datasets as provided by third parties.
11. Out of scope of this PIA is an assessment of the privacy implications of third-party compilation of data acquired by Data Ventures.
12. Also out of scope is any assessment on other products which may derive from, be related to, or are extensions of, the Population Density or Monthly Population Data product.
13. A separate PIA will be produced if any such changes to the collection and use of data occur for other products.

Assessment of Information Type

14. This section considers how the information for the Monthly Population Data product compares to the definition of 'personal information' in the Privacy Act 2020 (Privacy Act).
15. Data Ventures makes the Monthly Population Data product with datasets relating to location estimates acquired from data providers.
16. There are many processes in place to protect against Data Ventures being able to identify an individual via its data. For example:
 - a. Qrious (Spark) and Data Ventures signed a Change Request to cover the provision of monthly unique data rather than hourly data.
 - b. Data it receives from Spark will be highly aggregated compared to the previous Population Density product. The data supplied to Data Ventures contains a total device count for each population segment (local/domestic/international) in a regional tourism organisation area within a monthly time range.
 - c. Data it receives from Vodafone is already anonymised. The data includes location (cell tower site), timestamp, a flag showing whether a device has an NZ sim card or international, and a hashed IMEI, which cannot be used to determine the owner of the device or to join the location data with other datasets. This data is processed by Data Ventures' technical staff into aggregated device counts similar to the Spark dataset.
 - d. Cell tower sites do not provide highly specific location information (particularly in rural locations, a single cell tower may cover a large area).
 - e. Data Ventures uses large geographies that provide greater confidentiality to the data (Regional Tourism Organisation areas). The Regional Tourism Organisation area classification used for the Monthly Population Data product can be found at the following link:
<https://datafinder.stats.govt.nz/layer/104934-regional-tourism-organisation-areas-2020-clipped/>
 - f. Data Ventures combines the data from multiple data providers to create an aggregate total view.
 - g. After combining data from multiple data providers, Data Ventures models the data to create population counts and randomly rounds those population counts to base three. This process means the output data further obscures information about individuals.
17. Section 7 of the Privacy Act states that personal information means information about an identifiable individual.
18. The information Data Ventures acquires is anonymised and/or aggregated by the data providers

providing it. It does not contain any information which will allow for the identification of an individual by Data Ventures. However, the information received from Vodafone may enable tracking of the movements of a device, like cookies on a website, and therefore has been considered personal information.

Assessment of Privacy Impacts

19. This section will consider the proposed approach to implementing it against the information privacy principles of the Privacy Act to ascertain whether the approach is in the spirit of the privacy principles and will operate using best practice for information management. Consideration is also given to the Office of the Privacy Commissioner's (OPC) 'Principles for the safe and effective use of data and analytics' (data and analytics principles), which can be found at the following link:

<https://privacy.org.nz/assets/Uploads/Principles-for-the-safe-and-effective-use-of-data-and-analytics-guidance.pdf>

Collecting or obtaining information

20. Principle 1 of the Privacy Act states that personal information should not be collected by any agency unless the information is collected for a lawful purpose connected with a function of the agency, and the collection is necessary for that purpose.
21. Data Ventures has been created to produce products for organisations using Stats and non-Stats NZ datasets to help them make data-informed decisions. The Monthly Population Data product will fulfil this role by providing an indicator of the population counts – including visitor counts – across New Zealand at different geographic areas and time periods using location estimates data as the base datasets. This will be modelled to populations with other supplementary datasets.
22. The location estimates data to be used in the Monthly Population Data product is an estimate of the total population count within a monthly range within a geographic area. The data attributes used for this purpose will be the:
- month
 - segment (local residents, domestic visitors, international visitors, or short-term international visitors)
 - Regional Tourism Organisation area
 - count.

[Example output to be added below be added when web development is complete]

23. The information Data Ventures is collecting is aligned with its function and necessary for it to fulfil the purpose of creating the Population Density product. Population Density also could provide clear public benefits, in line with the OPC's data and analytics principles, by allowing for better planning by intended customers.
24. Principle 2 of the Privacy Act requires that agencies shall collect personal information directly from the individual concerned, unless (among other things) the collection from a third party is authorised by the person concerned or is not reasonably practicable in the circumstances of the particular case.
25. Data Ventures will acquire datasets of location estimates from data providers. The datasets will be anonymised and/or aggregated by the data providers prior to Data Ventures receiving them and will not contain any data allowing identification of an individual. It is not reasonably practical for Data Ventures to acquire this data directly from the individual party.
26. Additionally, Data Ventures has been advised that the possibility for the provision of data from a

data provider to a third party is part of the standard terms and conditions of the service agreed by the data providers' customer, such as informing them that user statistics that do not identify the customer may be shared with third parties.

27. Data Ventures has conducted a robust process with the Privacy and Legal teams of both data providers to ensure that the provision of data by them to Data Ventures is in line with their own standard terms and conditions and the OPC's data and analytics principles.
28. Therefore, there is both authorisation for the collection of information from the data providers, as well as it not being reasonably practical to acquire the data any other way. Data Ventures has also been guided by the OPC's data and analytics principle of focusing on people by considering the methods to be used for the Monthly Population Data product and protecting the privacy of individuals by collecting information that is either aggregated or anonymised.
29. Principle 3 of the Privacy Act requires that where an agency collects personal information directly from the individual concerned, the agency shall take reasonable steps to ensure that the individual concerned is aware of the fact that the information is being collected, the purpose of collection, who will receive it, and whether the collection is required by law. Agencies can forego this communication if (among other reasons) they believe that the individual authorises collection, or that communication is not reasonably practicable in the circumstances.
30. Data Ventures is not collecting data for the Monthly Population Data product directly from individuals. It is not reasonably practicable for Data Ventures to communicate the collection directly to individuals. However, Data Ventures has several policies to maintain transparency around its data use, including providing a methodology document that MBIE can publish alongside the data, which will include any limitations of the data, and publishing its Privacy Impact Assessments. This is in line with the OPC's data and analytics principle of maintaining transparency.
31. Principle 4 requires that information is not collected by means that are unlawful, unfair, or unreasonably intrusive.
32. As set out above, the information is being collected from data providers in alignment with the terms and conditions they have with their customers, which the customer can view and agree to prior to using the service, so is being collected in a lawful manner that is in line with the terms and conditions.
33. The datasets being provided to Data Ventures will be anonymised and/or aggregated and therefore will not be unreasonably intrusive.

Storage, security and retention

34. Principle 5 of the Privacy Act requires that an agency that holds personal information shall ensure that the information is protected by security safeguards as are reasonable to protect against loss or unauthorised access, use, modification or disclosure. These security safeguards are required when information is given to a third party.
35. Data Ventures has significant security safeguards as set out below.
36. The data sourced from the data providers will be stored in a secure cloud hosted environment. Access to this environment will be password protected with credentials only being supplied to each data provider for their own storage area. After the data from the providers is aggregated by Data Ventures it is stored in a separate, protected storage area accessible only by Data Ventures and Stats NZ for validation and methodology improvement. The aggregated device counts which are used to create population counts are not shared outside of Stats NZ. The derived population

count is stored separately and will be provided to MBIE via secure API.

37. To prevent the data being misused, access to the Vodafone input data is limited to Data Ventures technical staff. Access logging is in place for the Vodafone data stored in Data Ventures' system to ensure all access is for legitimate and necessary purposes. The Vodafone input data is used solely to create aggregated device counts and is never shared outside of Data Ventures technical staff.
38. Data received from Vodafone contains records at an individual IMEI level. To minimise the risk of this being able to be linked to an identifiable individual, the IMEI is hashed using an SHA-256 one way hashing algorithm by Vodafone prior to being provided to Data Ventures.
39. Principle 9 of the Privacy Act states that an agency that holds personal information shall not keep that information for longer than is required for the purposes for which it may lawfully be used.
40. The data sourced from each data provider, and the aggregated combined data from the providers, will be kept in long term archive storage for the express purposes of using the historic data to improve the methodology and audit any customer queries around the accuracy or validity of the data. As methodology is improved over time, access to the archived data will be required to improve the accuracy of the population estimates.
41. The open-ended, long-term storage of data for the Monthly Population Data product is not of concern because the information is de-identified or aggregated and it is being retained for the purpose of improving the product over time.

Access and correction

42. People have a right, under Principle 6 of the Privacy Act, to ask if an agency holds their personal information and to have access to that information. Principle 7 gives people the right to request correction of that information.
43. Data Ventures is not collecting information that can readily be linked to an identifiable individual. Consequently, Data Ventures will not hold personal information capable of being accessed or corrected.

Accuracy

44. Principle 8 of the Privacy Act states that agencies should ensure that personal information is accurate, up to date and relevant before it is used or disclosed.
45. Data Ventures is not collecting data that can be linked to identifiable individuals. The data collected from the data providers is based on source information and in an agreed format so can be considered accurate and up to date.
46. The information is not being used or disclosed in a manner that could have a direct or adverse impact on an individual.
47. Data Ventures will provide MBIE with a publishable version of the methodology and data definitions, as well as limitations of the data and any time periods where the data quality is lower.
48. Given these limitations, the data being used will be as accurate, up to date and relevant as possible considering its use for overall population density counts.
49. Data Ventures has heard from stakeholders across the tourism industry that the data from the different segments currently isn't available elsewhere in a timely or consistent form, particularly for domestic visitors. Regardless, any accuracy concerns would relate to the overall integrity of the product, rather than impact on individual privacy.

50. Data Ventures has followed the OPC's data and analytics principles to understand the limitations of the data, as set out above, and to ensure the data is fit for purpose by applying Stats NZ processes to it to ensure the data is compiled in such a way that meaningful information can be derived from it.

Use and disclosure

51. Principle 10 of the Privacy Act requires that information be used for the same purpose for which it was collected unless an exception applies. Principle 11 of the Privacy Act states that an agency shall not disclose personal information to another agency unless it is believed on reasonable grounds that the disclosure is for one of the purposes connected with the purpose the information was obtained. Principle 12 places restrictions on the disclosure of personal information outside New Zealand.

52. Data Ventures is acquiring information external to Stats NZ for specific products (in this case the Monthly Population Data product), applying Stats NZ processes to it and then providing it to other organisations.

53. As long as the datasets collected for the Monthly Population Data product are only used for that product, the use will be for the same purpose for which it was collected.

54. The data is not being disclosed outside of New Zealand; therefore Principle 12 does not apply to the Monthly Population Data product.

Unique identifiers

55. Under Principle 13, agencies shall not assign unique identifiers to individuals unless it is necessary for that agency to carry out its functions. Agencies shall not require individuals to disclose any unique identifier unless that disclosure is related to the purpose it was assigned.

56. Data Ventures will not receive any data allowing identification of individuals. All data received for the Monthly Population Data product will have been anonymised or aggregated by the data provider that provided it.

Conclusion and Recommendations

57. The Monthly Population Data product does not pose any material privacy risks to individuals. Data Ventures is only using aggregated or anonymised information and outputs of the product cannot be linked back to identifiable individuals in any way. This analysis of Data Ventures' processes show it is following best practice information management and the OPC's data and analytics principles.