**Data Quality Statement**

Historical Business Listings from Sands & McDougall directories





[deeca.vic.gov.au](file:///Users/fionadurante/Downloads/deeca.vic.gov.au)

## Summary information

|  |  |
| --- | --- |
| Data asset name: | Historical Business Listings from Sands & McDougall directories |
| Date: | August 2019 |
| Data asset owner: | Environment Protection Branch, Land and Environment Policy Division, Department of Energy, Environment and Climate Action  Data custodian - Manager Contaminated Land Policy and Planning |
| Description of data asset: | This dataset contains the business listings from selected years of the Sands & McDougall directories, a pre-cursor to the Yellow Pages. The directories are kept in hard copy at State Library Victoria and have been digitised and used for Victoria Unearthed.  The dataset was created by digitising the business directories that were published at approximately 10-year intervals between 1896 and 1974. The extracted information includes the business name, address and business type as listed in the Sands & McDougall directories. The extracted address information was subsequently geocoded to provide a point location that can be mapped using GIS software.  Additional information has been provided to allow the accuracy of the results to be evaluated with each record containing the data used at each stage of the data capture process; from the extraction from the directory to the final QA checks.  The dataset also categorises business types using the 2006 Australian and New Zealand Standard Industrial Classification (ANZSIC) business classifications. A relevant ANZSIC ‘subdivision’ is included for each business listing. This information has been included to assist with data analysis and is not part of the original Sands & McDougall directories. |
| Legislation and authority: | The dataset was created in response to Recommendation 14.1 from the Independent Inquiry into the Environment Protection Authority (2016) |
| **Scope and coverage:** | Statewide |
| **Reference period:** | The data includes mapping of business locations as listed in the trades and business section from Sands & McDougall directories from the years: 1896, 1905, 1915, 1925, 1935, 1945, 1955, 1965, 1974. |
| **Frequency and timing:** | The data has not been updated since its initial release in August 2019. |
| **Formats available:** | The data is viewable in [Victoria Unearthed](https://mapshare.vic.gov.au/victoriaunearthed/) where data can be downloaded as shape file, CSV and XLSX formats  The data can be accessed in multiple spatial data formats at Datashare <https://datashare.maps.vic.gov.au> and via DataVic at <https://discover.data.vic.gov.au/dataset/historical-business-listings-from-sands-mcdougall-directories2>  The data can be accessed as a Web Map Service (WMS) and Web Feature Service (WFS) via opendata.maps.vic.gov.au |
| **Notes** | Any information contained in Historical Business Listings from Sands & McDougall is intended to be used only as a guide and should not be relied upon as being either complete or accurate.  Historical business information can help Victorians understand more about businesses that were located at a particular place, at a point in time. Historical land use data can provide information to assist determining if a site may have been contaminated. For most business types there is no increased risk of contamination from a business activity. However, some business activities have a higher likelihood of causing contamination.  The Historical Business Listings and are not evidence of contamination. The information may, however, help you decide to make your own further enquiries about land and/or groundwater at a particular location.  The Historical Business Listings do not provide definitive information as to the actual activity that occurred on that site. For example, a manufacturing business listed at an address may not be the factory floor, but a head office for that business. A business name is not necessarily evidence that a particular activity occurred at that location. |

## Data quality description

### Accuracy

The Sands & McDougall directories have presented particular challenges regarding data accuracy as Victoria's street names and numbers may have changed drastically over the past century, and so has the way we record information.

Over half a million historical business records have been geocoded by an automated process for Victoria Unearthed. Many of these records have been manually checked and verified, to ensure the correct location on the map. However, many still remain unverified and may be located incorrectly in Victoria Unearthed.

There may also be some errors in the data that occurred in the transfer from hardcopy to digital, and also in geocoding the data to display in the Victoria Unearthed map.

The location of data points differ in precision. In many regional and suburban areas, the business address listed in the Sands & McDougall directories only indicates a town or suburb and does not include a street name or street number.

Business listings are represented as point locations, and these points represent the best known location of the business (down to address point level at best) – the data does not represent the exact location of the business on a property, nor does it represent the real extent of the historic land use.

For data that has not been manually inspected and verified, there may be errors due to incorrect OCR, incorrect text parsing or abbreviations, non-unique address listings, insufficient address data provided, changes in street naming and numbering, incorrect geocoding and other possible issues.

The original records are old and may themselves have been inaccurate, incomplete or both. For example, not all businesses would have had their information published in Sands & McDougall directories.

A business record is not necessarily evidence of business activity having historically occurred at that location. Historical business information in the directories can, for example, indicate the head office of a manufacturing business rather than the address of the factory floor.

### Completeness

The Historical Business Listings maps the locations of businesses listed in the trades and business section of the Sands & McDougall directories. The mapping does not include the residential sections of those directories.

The nine directories contained 557,445 records, of which 553,287 (99%) were able to be geocoded. 92,541 (17%) of these records were manually inspected and verified against other sources, such as historic maps and other historic datasets. Not all records could be successfully geocoded to street address level – approximately 64% of records were represented at address level, 14% at street level and 22% at locality level.

### Representative

The Historical Business Listings maps the locations of businesses listed in the trades and business section of the Sands & McDougall directories.

### Timeliness/Currency

The data includes point locations of business listings between 1896 and 1974.

### Collection

The data was compiled between March and July 2018 by GHD Pty Ltd in partnership with Lotsearch Pty Ltd. under contract with the Department of Environment, Land, Water and Planning (DELWP).

The process for creating the dataset involved digitisation and geocoding of hard copy Sands & McDougall directories held at State Library Victoria. The methodology followed the following steps:

* Text extraction and structuring: extracting text via optical character recognition (OCR) using high quality digital scans provided by State Library Victoria
* Land use categorisation: to assist with prioritising data capture and manual verification, to ensure business categories with a higher likelihood of causing contamination were prioritised for verification.
* Data cleansing and geocoding: to improve the quality of the geocoding outputs, including manual verification of records against other sources, such as historic maps and other historic datasets.
* Final clean-up and QA
* Data delivery and publishing

### Consistency

A consistent process was used to geocode the business locations. However, due to limitation in the original Sands & McDougall directories, data precision is poor for most listings outside the Melbourne metropolitan area. Most regional listings only include a town name, with no street address provided.

Data accuracy is also likely to be lower for business types unlikely to have caused contamination, due to errors in the automated geocoding process, as fewer records were manually checked against other sources.

### Fit for purpose

The data is suitable for its intended purpose: to identify the location of business listings from Sands and McDougall Directories to enable further investigations to inform land developments and planning decisions.

Historical business information can help Victorians understand more about businesses that were located at a particular place, at a point in time. Historical land use data can provide information to assist determining if a site may have been contaminated. For most business types there is no increased risk of contamination from a business activity. However some business activities have a higher likelihood of causing contamination.

The Historical Business Listings and are not evidence of contamination. The information may, however, help you decide to make your own further enquiries about land and/or groundwater at a particular location.

### The Historical Business Listings do not provide definitive information as to the actual activity that occurred on that site. For example, a manufacturing business listed at an address may not be the factory floor, but a head office for that business. A business name is not necessarily evidence that a particular activity occurred at that location.

## Disclaimer

This data asset is provided “as is”, without warranty to the suitability of the data for unspecified use.

Any information contained in Historical Business Listings from Sands & McDougall is intended to be used only as a guide and should not be relied upon as being either complete or accurate. DEECA does not verify the accuracy of all information contained in this dataset. Anyone with an interest in a particular piece of land, location, site, or property should make his or her own enquiries with relevant authorities.

## For further information

For further information regarding this data quality statement please contact the data custodian at [victoria.unearthed@delwp.vic.gov.au](mailto:victoria.unearthed@delwp.vic.gov.au)

## Document control

## Version history

|  |  |  |
| --- | --- | --- |
| Version | Date | Comments |
| 1.0 | 30 June 2023 | Approved by data custodian |