Codebook for data repository
Building age map, Vienna, around 1920

Ulrich Kral, Ferdinand Reimer

Contents

Preface

This Codebook

- **is part** of the GitHub repository building.map_1920. Check out the readme file to learn more about the corresponding datasets and the use and re-use the underlying R Markdown file.

- **refers** to the geospatial datasets "Building age map around 1920" and "Scope of analog building age map 1920", which can be retrieved from Zenodo.

- **specifies** the formats of the datasets and the data fields in the attribute tables of the datasets.

- **is for users** of the datasets who want to use and re-use the datasets.

1 General dataset specifications

1.1 Location and geographical coverage

The spatial coverage is the city of Vienna, which is the capital of Austria (Latitude: 48° 12' 30.56" N, 16° 22' 19.49" E).

1.2 Dataset format

Geospatial data

- File extension: *.shp
- Coordinate reference system: EPSG-Code 31256.
- Software readability: Examples: QGIS, ArcGIS, R and FME Data Integration Platform from Safe Software

Attribute tables

- File extension: *.csv
- CSV specifications: The data records are separated by semicolons (";"). The decimal separator is a decimal point ".". The Universal Coded Character Set (UCS) is UTF-8.
- Software readability: Files are readable with the aforementioned software tools.

2 Dataset "Building age map around 1920"

2.1 Introduction

This chapter specifies the geospatial dataset Building age map around 1920, briefly called BSM_1920.

The dataset includes 80640 data entries (rows) and 28 data fields (columns). The dataset can be retrieved from the Zenodo repository, whereas the dataset can be combined from the following files

- "BSM_1920.shp", which includes the polygon-features and an unique identifier in the data field "ID".
- "BSM_1920_attribute_table_final.csv", which is the attribute table without geometry information. It also includes the data field "ID" and therefore can be linked with the "Building stock and age 1920.shp" file.

2.2 Attribute table

The dataset includes 28 data fields.

Field	Name
ID	Identifier polygon-feature
ID.bs	Identifier building schematic
Area	Area
UD.1920	Urban district number 1920
UD.2020	Urban district number 2020
Location.city_limit_1920	Location city limit
Location.abam_1920	Location ABAM scope

Field	Name
Polygon.source_id	Identifier polygon source
Polygon.source_name	Name of polygon source
Editor	Editor's name
CD.abam	Construction date ABAM_1920
CD.bs	Construction date BS
TP_pub.date	Temporal presence given by date
TP_pub.date_year	Temporal presence given by year
TP_pub.date_period	Temporal presence given by periods
PoC	Period of construction
PoC.source_name	Source name of period of construction
ID.Address.2019	Identifier for address point
Address.2019_reloc	Address point re-location
Address.2019	Address in 2019
Address.1920	Address in 1920
Address_reviewed	Review status of address
Address_revised	Revised address
Polygon.source.Scale	Scale
Location.usm	Spatial presence built-up area 1918
Location.lsm	Spatial presence built-up area 1912
Presence.conf	Spatial and temporal presence
CD.bs_fit	Construction date fitting

2.3 Comments

This section comments on the individual data fields by giving background information and details for using the data.

2.3.1 ID

The field includes a unique character for each of the 80640 polygon-features. The character the combination of "UD" for urban district", the district number "UD.2020", an underline "_" and an integer. Example: "UD1_1543".

It is noted that the integer at the end of the "ID" is unique per urban district and doesn't follow a specific logical order. It is also noted the series of integers has gaps, which is due to the removal and merging of polygon-features during the clean-up process and the integration of all building footprints from more than 100 different maps sheets into a single map.

2.3.2 ID.bs

The field includes a unique character for 32569 polygon-features. It corresponds with the data field "ID" of the Building schematic of Vienna in the late 1920s dataset and therefore allows to assign data records "year of construction", "year of purchase", "cadastral community" and "floors".

It is noted that the "ID.bs" records are based on address matching between the BSM_1920 and the Building schematic of Vienna in the late 1920s dataset. The addresses in the BSM_1920 dataset are not completely verified, because we were not able to reconstruct the address data of 1920 based on address labels in historical building stock maps.

2.3.3   Area

The field specifies the area of the polygon-feature in squared meters. It is noted that polygon-features represent built-up land but not necessary buildings in a definitional sense.

2.3.4   UD.1920

The field specifies the location of the polygon-feature with respect to the urban district distribution in 1920. At this time, the city was divided into 21 districts.

The field covers the following unique records: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, NA. The integers 1-21 stand for the number of the urban districts. "NA" records are for polygon-features beyond the city boundary 1920.

2.3.5   UD.2020

The field specifies the location of the polygon-feature with respect to the urban district distribution in 2020.

The field covers the following unique records: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23. The integers 1-23 stand for the number of the urban districts.

2.3.6   Location.city_limit_1920

This data field specifies if the polygon-feature is located inside our outside the city boundary of 1920.

The field covers the following unique records: in, out.

2.3.7   Location.abam_1920

The analog building age map 1920 covers 36% of the city area at this time. This field specifies if the polygon-feature is inside or outside of the coverage area of the analog building age map 1920.

The field covers the following unique records: in, out.

2.3.8   Polygon.source_id

The field identifies the data sources from which the building footprints were retrieved. The field covers 136 unique records. The first two positions indicate the type of data source, whereas

- "FF" stands for "Fortführungsmappe" (in German). It is an analog cadastral map sheet, which we used to vectorize building footprints. The cadastral maps sheets are available in TIFF format and can be retrieved from Federal Office of Metrology and Surveying (BEV). The nomenclatur is presented in Figure @ref(fig:nomenclatur).

- "FW" stands for "Feuerwehrplan" (in German). It is an historical fire brigade map sheet, which we used to vectorize building footprints. The cadastral maps sheets are available in TIFF format and can be retrieved from Federal Office of Metrology and Surveying (BEV).
- "MA" stands for "Magistratratsabteilung" (in German). It is the city authority - department 41, who provided the polygon-features. The polygon-features can be retrieved from Open Data Österreich.
- "BE" stands for "Bezirksplan" (in German). It covers historical urban district maps. The urban district maps are available in TIFF format and can be retrieved from Municipal and Provincial Archives of Vienna.

Figure 1: Nomenclatur for identifying historical cadastral map sheets.

- "BA" stands for "Baualterskarte" (in German). It is the analog building age map 1920. It covers historical urban district maps. The urban district maps are available in TIFF format and can be retrieved from Municipal and Provincial Archives of Vienna.
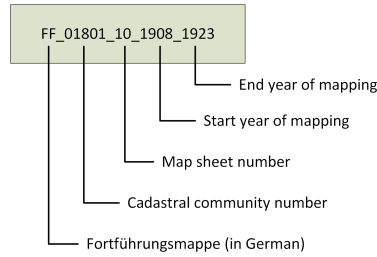
2.3.9   Polygon.source_name

This field groups the "Polygon.source_id" records by the type of the cartographic document into: City map 2018, Historic urban district map, Historic fire brigade map, Historic cadaster, Analog building age map 1920, Analog building age map 1920 (digital raw data).

2.3.10   Editor

The field covers name of the person who digitized the building footprints from analog maps or the institution that provided the polygon-features.

The field covers the following unique records: City authority (MA41), Andreas Danzinger, Mateo Petz, Hannah Wolfinger, Ingeborg Hengl, Sarah Prunner, Lupina Prospero, Klara Stuppacher, Ulrich Kral, Friedrich Hauer.

2.3.11   CD.abam

The field covers construction periods of buildings, which were retrieved from the analog building age map 1920.

The field covers the following unique records: , 1871-1880, 0-1850, 1861-1870, 1851-1860, 1901-1910, 1911-1920, 1881-1890, 1891-1900, 1819-1876, 0-1818, 1877-1890, 0-1920. Blank records indicate that no information is available.

2.3.12   CD.bs

The field covers construction years of buildings, which were retrieved from Building schematic of Vienna in the late 1920s.

It is noted that the assignment of construction years is based on address matching and the addresses in the BAM_1920 were not entirely validated. It is also noted that the BSM_1920 includes polygon-features with multiple addresses. Due to the address matching, these polygon-features potentially got multiple construction year assignments.

The field covers 735 unique records.

2.3.13   TP_pub.date

The field stands for the temporal presence of the building footprint. The temporal presence was retrieved from the publication data of the analog building stock map, from which we retrieved and digitized the building footrint. In other words, the presence is confirmed if the builing is shown on the analog buildig age map. The temporal presence is either given by years or periods. Periods are given for the cadastral map sheets (see data field "Polygon.source_id"), because the map sheets were used to record changes in a specific period.

The field covers the following unique records: 30.

2.3.14   TP_pub.date_year

The field includes either the year or the end of the period as given in "TP_pub.date". It documents the temporal presence of the building footprints by a single year.

The field covers the following unique records: 1906, 1907, 1908, 1910, 1911, 1912, 1913, 1914, 1915, 1920, 1921, 1922, 1923, 1927, 1928, 1929, 1930, 1935, 1936.

2.3.15   TP_pub.date_period

This field groups the "TP_pub.date_year" records into the following periods: 1906-1912, 1920-1923, 1927-1936.

2.3.16   PoC

This field covers the building's period of construction. The periods were retrieved from "CD.abam", "CD.bs" and "TP_pub.date_year" and harmonized as followed: 0-1818, 0-1850, 0-1920, 0-1930, 0-1940, 1851-1860, 1861-1870, 1871-1880, 1881-1890, 1891-1900, 1901-1910, 1911-1920, 1921-1930.

It is noted that the periods include two distinctive types. Firs, periods with a timespan of 10 years. Second, periods that indicate that the building has been constructed before a given year (e.g. 0-1850).

2.3.17   PoC.source_name

This field specifies the data source from which the "PoC" records were retrieved. The field covers the following unique records: TP_pub.date_year, ABSM, CD.abam_1920, ABAM_1920, CD.bs, BS_1920s.

2.3.18   ID.Address.2019

The "Address.2019" records were retrieved on 6 July 2019 from the address database Adressen Standorte Wien. It's data field "OBJECTID" is identical with the data field "ID.Address.2019".

2.3.19   Address.2019_reloc

The "Address.2019" records were retrieved from the address database Adressen Standorte Wien as point data. Points that were located outside of polygon-features were re-located to the inside of polygon-features close by. This enables the assignment of address data to polygon-features. This data field specifies if the address point was manually re-located.

The field covers the following unique records: , no, yes.

2.3.20 Address.2019

The field specifies the address in 2019, in particular the name of the traffic area and the orientation number. This field corresponds with "NAME" in the address database Adressen Standorte Wien.

The field covers 37212 unique records.

2.3.21 Address.1920

The field specifies the address in 1920, in particular the name of the traffic area and the orientation number. It is noted that the address data are not entirely validated.

The field covers 37749 unique records.

2.3.22 Address_reviewed

We reviewed some "Address.2019" records to verify if the address was valid in 1920. This data field specifies if the "Address.2019" record was reviewed or not.

The field covers the following unique records: no, yes.

2.3.23 Address_revised

Addresses under review (see "Address_reviewed") were revised based on the address labels in historical building stock maps (see "Address_reviewed"). This data field includes all addresses, as shown on historical buildings stock maps with address labels, that refer to the polygon-features. This data fields includes the "Address.2019" record if it is valid for 1920 too, as well as additional addresses. It is noted that multiple addresses (e.g. for corner buildings) are separated by a comma (";").

The field covers 24859 unique records.

2.3.24 Location.usm

The analog urban sprawl map Die räumliche Ausdehnung Wien von 1857 bis 1957 documents the built-up area in 1918. This field specifies if the polygon-feature is located within the built-up area 1918. The field covers the following unique records: yes, no.

2.3.25 Location.lsm

The land use map documents the built up area 1912. The map is available in SHP format and can be retrieved from S. Hohensinner based on personal request. This field specifies if the polygon-feature is located within the built-up area 1918. The field covers the following unique records: .

2.3.26 Presence.conf

The data field combines the information from "Location.usm" and "Location.lum" and confirms if the building footprint can be found in one of the two types of built-up areas. Building footprints outside of the built-up areas are considered as long as the temporal presence "TP_pub.date_year" equal or before 1921.

The field covers the following unique records: inside built-up areas of LSM 1912 and USM 1918, inside built-up area of USM 1918, inside built-up area of LSM 1912, outside built-up areas of LSM 1912.

2.3.27 CD.bs_fit

This data field shows if the construction year "CD.bs" fits within the period of construction "PoC". The field covers the following unique records: No 'CD.bs' record, included, before, ambiguous, after. The data record are interpreted as followed:

- "included": The construction year "CD.bs" is within the period of construction "PoC".
- "before": The construction year "CD.bs" is before the start year of the period of construction "PoC".
- "after": The construction year "CD.bs" is after the end year of the period of construction "PoC".
- "ambiguous": The "CD.bs" records include multiple years, which are within two distinctive periods of construction "PoC". So, the "CD.bs" record cannot be assigned to a single period.
- "No 'CD.bs' record": The polygon-feature has no "CD.bs" record.

3 Dataset "Scope of analog building age map 1920"

3.1 Introduction

This chapter refers to the geospatial dataset Scope of analog building age map 1920, briefly called SABAM_1920. The dataset includes 1 polygon-feature.

3.2 Attribute table

The dataset includes 2 data fields.

Field	Name
ID	Identifier of polygon-feature
Area	Area of polygon-feature

3.3 Comments

3.3.1 ID

This field includes a unique identifier for the polygon-feature.

3.3.2 Area

This field specifies the area of the polygon-feature in squared meter.

Competing interests

The authors declare no competing interests.