**Schweizer Viehbestand**

Victoria Hummel, Leonard Kieffer, Gianmaria Pallecchi

Data Report

# Raw data

## Overview table of raw data sets

|  |  |  |
| --- | --- | --- |
| **Name data set** | **Source** | **Storage location** |
| cattle-map-canton.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/cattle-map-canton.html |
| goats-map-canton.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/goats-map-canton.html |
| sheep-map-canton.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/sheep-map-canton.html |
| cattle-map-commune.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/cattle-map-commune.html |
| goats-map-commune.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/goats-map-commune.html |
| sheep-map-commune.csv | Identitas AG: Develops and operates applications for the identification, registration, and monitoring of livestock and pets. | https://tierstatistik.identitas.ch/de/sheep-map-commune.html |
| slaughterhouse.csv (if downloaded directly: BLV-Bewilligungsliste\_YYYY-MM-DD.csv [date of download]). | Federal Food Safety and Veterinary Office BLV: Among other things, grants permits for food businesses. | https://www.blv.admin.ch/blv/de/home/lebensmittel-und-ernaehrung/rechts-und-vollzugsgrundlagen/bewilligung-und-meldung/listen-bewilligter-betriebe.html  (filtered by results from "Sektion I Fleisch von als Haustiere gehaltenen Huftieren - Suchresultate: 529") |
| swissBOUNDARIES3D\_1\_5\_  TLM\_KANTONSGEBIET.shp | Federal Office of Topography) swisstopo: Surveys and documents landscape/subsoil of Switzerland and provides spatial geodata. | https://www.swisstopo.admin.ch/de/landschaftsmodell-swissboundaries3d#swissBOUNDARIES3D---Download  (retrieved from folder: swissboundaries3d\_2024-01\_2056\_5728.shp.zip) |

## Details Dataset 1 (cattle-map-canton.csv)

* Description of the dataset: Registered (living) cattle per canton Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | canton | string | name of canton of all Swiss cantons | name of canton data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of cattle per canton |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of cattle per km2 in canton |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of cattle per 100 inhabitants in canton |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in canton, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in canton, with count |  |

## Data Quality Dataset 1

* The dataset is cleaned (converted numeric to floats, handled missing data if present, sorted) and saved as cattle-cleaned-canton.csv.
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.

## Details Dataset 2 (goats-map-canton.csv)

* Description of the dataset: Registered (living) goats per canton: Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | canton | string | name of canton of all Swiss cantons | name of canton data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of goats per canton |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of goats per km2 in canton |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of goats per 100 inhabitants in canton |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in canton, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in canton, with count |  |

### Data Quality Dataset 2

* The dataset is cleaned (converted numeric to floats, handled missing data if present, sorted) and saved as goats-cleaned-canton.csv.
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.

## Details Dataset 3 (sheep-map-canton.csv)

* Description of the dataset: Registered (living) sheep per canton: Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | canton | string | name of canton of all Swiss cantons | name of canton data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of sheep per canton |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of sheep per km2 in canton |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of sheep per 100 inhabitants in canton |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in canton, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in canton, with count |  |

### Data Quality Dataset 3

* The dataset is cleaned (converted numeric to floats, handled missing data if present, sorted) and saved as sheep-cleaned-canton.csv.
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.

## Details Dataset 4 (cattle-map-commune.csv)

* Description of the dataset: Registered (living) cattle per commune: Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | commune | string | name of commune of all Swiss communes | name of commune data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of cattle per commune |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of cattle per km2 in commune |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of cattle per 100 inhabitants in commune |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in commune, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in commune, with count |  |

### Data Quality Dataset 4

* Names of communes are standardized (lowercase, removed parentheses, unified abbreviations [e.g. st., st, sankt -> saint]).
* Dataset was combined with Geodata for visualizing on the map
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.
* However, there were conflicts:
  + Some communes in the dataset merged together
  + Some names of communes don't match the communes in Geodata
  + Multiple communes sharing the same name (in different cantons)
  + Lakes listed as communes

## Details Dataset 5 (goats-map-commune.csv)

* Description of the dataset: Registered (living) goats per commune: Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 5

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | commune | string | name of commune of all Swiss communes | name of commune data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of cattle per commune |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of sheep per km2 in commune |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of sheep per 100 inhabitants in commune |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in commune, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in commune, with count |  |

### Data Quality Dataset 5

* Names of communes are standardized (lowercase, removed parentheses, unified abbreviations [e.g. st., st, sankt -> saint]).
* Dataset was combined with Geodata for visualizing on the map
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.
* However, there were conflicts:
  + Some communes in the dataset merged together
  + Some names of communes don't match the communes in Geodata
  + Multiple communes sharing the same name (in different cantons)
  + Lakes listed as communes

## Details Dataset 6 (sheep-map-commune.csv)

* Description of the dataset: Registered (living) sheep per commune: Count absolute, count per surface area (km2) without waters, count per 100 inhabitants, count and rank of top 5 breeds and names.
* Data source: Identitas AG – Develops and operates applications for the identification, registration, and monitoring of livestock and pets.
* Data acquisition: Data set downloaded as a csv-file, directly from the data source Identitas AG. Identitas AG used data from Tierverkehrsdatenbank (TVD), which is in fact operated by Identitas AG. TVD marks cloven-hoofed-animals (e.g. cattle, goats and sheep) with an ear tag and builds the basis for traceability of animals.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: The data is publicly available.
* Data access: The data is available as direct download on the Identitas AG website.

### Data Catalog Dataset 6

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | commune | string | name of commune of all Swiss communes | name of commune data belongs to | Primary Key |
| 2 | count | integer | positive integers | total count of sheep per commune |  |
| 3 | countPerSurfacekm2 | float | positive decimals | count of sheep per km2 in commune |  |
| 4 | countPer100Inhabitants | float | positive decimals | count of sheep per 100 inhabitants in commune |  |
| 5 | top5breeds | string | list of 5 pairs (breed, count) separated by comma | 5 most common breeds in commune, with count |  |
| 6 | top5names | string | list of 5 pairs (name, count) separated by comma | 5 most common names in commune, with count |  |

### Data Quality Dataset 6

* Names of communes are standardized (lowercase, removed parentheses, unified abbreviations [e.g. st., st, sankt -> saint]).
* Dataset was combined with Geodata for visualizing on the map
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.
* However, there were conflicts:
  + Some communes in the dataset merged together
  + Some names of communes don't match the communes in Geodata
  + Multiple communes sharing the same name (in different cantons)
  + Lakes listed as communes

## Details Dataset 7 (slaughterhouse.csv)

* Description of the dataset: Dataset containing businesses that process meat from hoofed animals kept as pets. (Unique) permit number of business, name of business, address, postal code, category (slaughterhouse and/or processing plant), further activities performed by business (e.g. additional types of animals processed), type of animal processed.
* Data source: Federal Food Safety and Veterinary Office BLV: Among other things, collects lists of approved food businesses handling animal products.
* Data acquisition: The cantonal law enforcement agencies grant the mandatory permits for businesses handling animal products. The BLV then collects the data and creates lists of all businesses handling animal products.
* Legal aspects: Data replication needs written approval of originator (BLV).
* Data governance: Data is publicly available.
* Data access: The data is available from the BLV website as a download. The resulting csv can be filtered by type of product, e.g. by (various types of) meats, honey, collagen etc. In our case, the results were filtered by "Section I: meat from hoofed animals kept as pets".

### Data Catalog Dataset 7

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Column index | Column name | Data type | Values | Short description | Primary/ Foreign Key |
| 1 | Bew.-Nr | integer | positive integers | (unique)permit number of slaughterhouse | Primary Key |
| 2 | Firmenname | string | names of slaughterhouses | names of slaughterhouses |  |
| 3 | Adresse | string | address of slaughterhouse | name of street with number |  |
| 2 | PLZ | integer | positive integers | postal code |  |
| 2 | Ort/Region | string | area and region | municipality and abbreviated canton |  |
| 2 | Kategorie | string | SH, CP | category:  SH = slaughterhouse, CP = processing plant |  |
| 2 | Weitere Aktivitäten | string | various abbreviations | further activities performed by slaughterhouse (e.g. slaughtering caged game) |  |
| 2 | Tierart | string | B, C, O, P, S | type of animal:  B = bovine genus,  C = goat-like,  O = sheep,  P = swine  S = solid-hoofed animal |  |

### Data Quality Dataset 7

* Missing data in the dataset was removed.
* Dataset was combined with OpenCage Geocoding APi for visualizing on the map and stored (with latitude/longitude) as slaughterhouse\_with\_coordinates.csv.
* The process is to be seen in the script *Slaugherhouse\_html.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.

## Details Dataset 8 (swissBOUNDARIES3D\_1\_5\_TLM\_KANTONSGEBIET.shp)

* Description of the data set: Shapefile for Swiss cantons.
* Data source: (Federal Office of Topography) swisstopo: Surveys and documents landscape/subsoil of Switzerland and provides spatial geodata.
* Data acquisition: Official cadastral survey of Switzerland, the principality of Liechtenstein and exclaves of adjacent countries. 3D borders (x, y, z) of countries, cantons, districts and municipalities including names.
* Legal aspects: The data is free to use for both non-commercial and non-commercial usage; citation is mandatory.
* Data governance: Data is publicly available.
* Data access: The data is located in the folder swissboundaries3d\_2024-01\_2056\_5728.shp.zip, which is available as direct download on the swisstopo website.

### Data Quality Dataset 8

* The dataset (being a shapefile of Swiss cantons) was combined with the cattle-map-canton.csv dataset in order to visualize it on the map.
* The process is to be seen in the script *interactive\_map.ipynb.*
* The dataset is from an official source. Therefore, the quality of the data is assumed to be good. The data is assumed to be complete.
* Description of the dataset, according to the source (swisstopo):
* consistent quality and form across the entire area
* geometric accuracy ±0.5 m in position and height
* clear and stable object identification
* simple structure

# Processed Data

## Overview table of processed data

|  |  |  |
| --- | --- | --- |
| Name | Input-Data sets | Storage location |
| cattle-cleaned-canton.csv | cattle-map-canton.csv | stored locally and in Github repository |
| sheep-cleaned-canton.csv | sheep-map-canton.csv | stored locally and in Github repository |
| goats-cleaned-canton.csv | goats-map-canton.csv | stored locally and in Github repository |
| slaughterhouse\_with\_coordinates.csv | slaughterhouse.csv | stored locally and in Github repository |

## Details Processed Dataset 1 (cattle-cleaned-canton.csv)

* The original dataset cattle-map-canton.csv was slightly cleaned/modified up for processing purposes, containing the same information.
* Dataset was modified for further use:
  + Converting columns to the desired format, e.g. converting values of column 'count' from string to numeric.
  + Handle missing data (if missing data is present)
  + Sorting
* The processed dataset 1 was created in the notebook 'interactive\_map.ipynb' and is available in the repository.

## Details Processed Dataset 2 (sheep-cleaned-canton.csv)

* The original dataset sheep-map-canton.csv was slightly cleaned/modified up for processing purposes, containing the same information.
* Dataset was modified for further use:
  + Converting columns to the desired format, e.g. converting values of column 'count' from string to numeric.
  + Handle missing data (if missing data is present)
  + Sorting
* The processed dataset 2 was created in the notebook 'interactive\_map.ipynb' and is available in the repository.

## Details Processed Dataset 3 (goats-cleaned-canton.csv)

* The original dataset goats-map-canton.csv was slightly cleaned/modified up for processing purposes, containing the same information.
* Dataset was modified for further use:
  + Converting columns to the desired format, e.g. converting values of column 'count' from string to numeric.
  + Handle missing data (if missing data is present)
  + Sorting
* The processed dataset 3 was created in the notebook 'interactive\_map.ipynb' and is available in the repository.

## Details Processed Dataset 4 (slaughterhouse\_with\_coordinates.csv)

* Coordinates (latitude and longitude) added to the location of the slaughterhouses using OpenCage Geocoding API.
* Values for 3 columns of each row from the original dataset describing the location were used to find the corresponding coordinates. Then, the slaughterhouses could be displayed on the interactive map.
* The processed dataset 4 was created in the notebook 'Slaughterhouse\_html.ipynb' and is available in the repository.