Zurich University of applied sciences
School of Life Sciences and facility management
Institute of computational life sciences

Lessons from the Past: Visualizing Switzerland's Pandemic History to prepare for the future.

Project Report | Module PODSV

By
Xenia Zbinden, Elisabeth Hermann,
Penélope Plos and Nicole Cieplinski

Submission date: 06.06.2025 | Last Update: 29.05.2025

Academic reviewer:
Prof. Dr. Wibke Weber
ZHAW Angewandte Linguistik
Theaterstrasse 15c, 8400 Winterthur

Prof. Dr. phil. Kaspar Staub
UZH Institute of Evolutionary Medicine
Winterthurerstrasse 190, 8075 Zürich

Table of Contents

Overview of the Raw Data:

Dataset Name	Source	Storage location
Dataset 1 History of Pandemics in Switzerland	Kaspar Staub and Wiebke Weber from leaddata.ch	Uploaded Excel file: 1_History_Pandemics.xlsx in GitHub Repo: Project_PODSV/Data at main · plospen1/Project_PODSV
Dataset 2	Kaspar Staub and Wiebke Weber from leaddata.ch	Uploaded Excel files: 2_All_cantons_1953-1958_Mortality.xlsx, 2_Data_cantons_incidence_weekly_56_58_NEW.xlsx and 2_Population_cantons.xlsx in GitHub Repo: Project_PODSV/Data at main · plospen1/Project_PODSV
Dataset 3 3_Todesursachen Schweiz ohne Alter 1876-2002	Kaspar Staub and Wiebke Weber from leaddata.ch	Uplooaded Excel file: 3_Todesursachen Schweiz ohne Alter 1876-2002 in GitHub Rep: Project_PODSV/Data at main · plospen1/Project_PODSV
full_data	covid-19-data/public/data/cases_deaths at master · owid/covid-19-data	Uplooaded Excel file: full_data: Project_PODSV/Data at main · plospen1/Project_PODSV

Table 1: Summary Table of the Raw Datasets

Details Dataset 1

Description:

- Annual mortality statistics in Switzerland from 1880 to 2022, including pandemic-related deaths and excess mortality estimates.

Source Details:

- Historical data from Switzerland
- COVID-19 data from Switzerland

Data Acquisition:

- File: 1_History_Pandemics.xlsx, contact Person Wiebke Weber

Legal Aspects:

- The data is publicly available and free to use for educational and analytical purposes.

Data Governance:

- Category: Public
- No personal data is included.

Access Information:

- Provided as an Excel file. On GitHub, linked at Speicherort.
- Can be accessed using common tools (Excel, Python pandas, etc.)

Data Catalog Dataset 1

Index	Column Name	Data Type	Values / Validation	Description

1	Jahr	Integer	1880–2022	Year of observation
2	Todesfälle_Grippe_100000	Float	≥ 0	Influenza deaths per 100,000 population
3	Todesfälle_Covid_100000	Float	≥ 0	COVID-19 deaths per 100,000 population
4	TodesfälleGrippe	Float	≥ 0	Total number of influenza deaths
5	TodesfälleCOVID	Float	≥ 0	Total number of COVID-19 deaths
6	Population	Integer	≥ 0	Total population of Switzerland
7	Überasterblichkeit_Alles	Float	Any value (%)	Total excess mortality percentage
8	ÜberasterblichkeitPlus	Float	≥ 0	Positive excess mortality percentage
9	ÜberasterblichkeitMinus	Float	≤ 0	Negative excess mortality percentage

Table 2: Data description for the 1_History_Pandemics table

Dataquality Dataset 1

Methodology

- The analysis was performed using automated data profiling tools such as ydata.
- Descriptive statistics and a Pearson correlation matrix were used.
- The dataset includes 143 annual observations from 1880 to 2022.
- There are 9 variables total: 7 numeric and 2 categorical.

General Statistics

- Number of variables: 9
- Number of observations: 143
- Time span: 1880–2022
- Missing cells: 33.3% of all values
- No duplicate rows or invalid entries
- Mean year: 1951
- Mean population: approximately 5.2 million

Key Variables and Values

- TodesfälleGrippe (Influenza deaths): 3 to 21,491 per year, mean = 767, 2.1% missing.

- TodesfälleCOVID: 3 values only, 97.9% missing.
- Todesfälle_Grippe_100000: 0.0385 to 554 per 100k, mean = 18.2, minimal missing.
- Todesfälle_Covid_100000: 3 values, 97.9% missing.
- ÜbersterblichkeitMinus: -13.8% to 0%, mean = -2.87%, 52.4% missing.
- ÜbersterblichkeitPlus: 0% to 49.3%, mean = 3.74%, 46.2% missing.
- Übersterblichkeit_Alles: nearly complete, -13.8% to 49.3%, mean = 0.66%.

Interpretation and Insights

- The dataset is ideal for long-term influenza mortality analysis.
- COVID-19 data is too incomplete for reliable modeling.
- Excess mortality is a valuable measure for both direct and indirect pandemic impact.
- Population normalization (per 100k) allows comparison across time.
- Negative excess mortality years may reflect successful health measures or mild seasons.

Data Quality Assessment

- Completeness: moderate (mainly missing values in COVID fields).
- Consistency: high (no duplicates or major outliers).
- Usability: excellent for flu and excess mortality; limited for COVID.

Conclusion

- Long-term flu mortality trends
- Excess mortality during pandemics
- Data-driven storytelling

Details Dataset 2

Description:
This dataset contains detailed demographic and epidemiological records for Switzerland and its cantons from 1953 to 1958, including:

- 2_Population_cantons.xlsx : Population per canton
- 2_Data_cantons_incidence_weekly_56_58_NEW.xlsx: Influenza incidence (weekly & monthly)
- 2_All_cantons_1953-1958_Mortality.xlsx: Total and influenza-specific mortality per canton

Source Details:

- Data provided via leaddata.ch
- For more information please contact Wiebke Weber or Kaspar Staub

Data Acquisition:
Provided as three Excel files:

- 2_Population_cantons.xlsx
- 2_Data_cantons_incidence_weekly_56_58_NEW.xlsx
- 2_All_cantons_1953-1958_Mortality.xlsxLegal

Legal Aspects:

- Data is publicly available for educational and research purposes.
- No personal data; anonymized at the canton level

Data Governance:

- Category: Public
- No personal or sensitive information

Access Information:

- Provided as an Excel file. On GitHub, linked at Speicherort.

- Can be imported using Excel, Python (pandas), or R
- Cleaned datasets used in Jupyter Notebook for full reproducibility

Data catalog Dataset 2

Index	Column Name	Data Type	Values / Validation	Description
1	Month	String	January – December	Observation month
2	Year	Integer	1953 – 1958	Observation year
3	Parameter	String	Deaths Influenza	Parameter for the values
4	ZH	Integer	≥ 0	Deaths for specific canton per month
5	BE	Integer	≥ 0	Deaths for specific canton per month
6	LU	Integer	≥ 0	Deaths for specific canton per month
7	UR	Integer	≥ 0	Deaths for specific canton per month
8	SZ	Integer	≥ 0	Deaths for specific canton per month
9	OW	Integer	≥ 0	Deaths for specific canton per month
10	NW	Integer	≥ 0	Deaths for specific canton per month
11	GL	Integer	≥ 0	Deaths for specific canton per month
12	ZG	Integer	≥ 0	Deaths for specific canton per month
13	FR	Integer	≥ 0	Deaths for specific canton per month
14	SO	Integer	≥ 0	Deaths for specific canton per month
15	BS	Integer	≥ 0	Deaths for specific canton per month
16	BL	Integer	≥ 0	Deaths for specific canton per month
17	SH	Integer	≥ 0	Deaths for specific canton per month
18	AR	Integer	≥ 0	Deaths for specific canton per month

19	AI	Integer	≥ 0	Deaths for specific canton per month
20	SG	Integer	≥ 0	Deaths for specific canton per month
21	GR	Integer	≥ 0	Deaths for specific canton per month
22	AG	Integer	≥ 0	Deaths for specific canton per month
23	TG	Integer	≥ 0	Deaths for specific canton per month
24	TI	Integer	≥ 0	Deaths for specific canton per month
25	VD	Integer	≥ 0	Deaths for specific canton per month
26	VS	Integer	≥ 0	Deaths for specific canton per month
27	NE	Integer	≥ 0	Deaths for specific canton per month
28	GE	Integer	≥ 0	Deaths for specific canton per month
29	CH	Integer	≥ 0	Deaths for the whole of Switzerland per month

Table 3: Data description for the 2_all_cantons_1953-1958_Mortality table

Index	**Column Name**	**Data Type**	**Values / Validation**	**Description**
1	StartReportingPeriod	Datetime	YYYY-MM-DD	Starting date of observation period per week
2	EndReportingPeriod	Datetime	YYYY-MM-DD	End date of observation period per week
3	Month	String	January – December	Observation month
4	Parameter	String	Cases Influenza	Parameter for the values
5	ZH	Integer	≥ 0	Cases for specific canton per week
6	BE	Integer	≥ 0	Cases for specific canton per week
7	LU	Integer	≥ 0	Cases for specific canton per week

8	UR	Integer	≥ 0	Cases for specific canton per week
	SZ	Integer	≥ 0	Cases for specific canton per week
9	OW	Integer	≥ 0	Cases for specific canton per week
10	NW	Integer	≥ 0	Cases for specific canton per week
11	GL	Integer	≥ 0	Cases for specific canton per week
12	ZG	Integer	≥ 0	Cases for specific canton per week
13	FR	Integer	≥ 0	Cases for specific canton per week
14	SO	Integer	≥ 0	Cases for specific canton per week
15	BS	Integer	≥ 0	Cases for specific canton per week
16	BL	Integer	≥ 0	Cases for specific canton per week
17	SH	Integer	≥ 0	Cases for specific canton per week
18	AR	Integer	≥ 0	Cases for specific canton per week
19	AI	Integer	≥ 0	Cases for specific canton per week
20	SG	Integer	≥ 0	Cases for specific canton per week
21	GR	Integer	≥ 0	Cases for specific canton per week
22	AG	Integer	≥ 0	Cases for specific canton per week
23	TG	Integer	≥ 0	Cases for specific canton per week
24	TI	Integer	≥ 0	Cases for specific canton per week
25	VD	Integer	≥ 0	Cases for specific canton per week
26	VS	Integer	≥ 0	Cases for specific canton per week
27	NE	Integer	≥ 0	Cases for specific canton per week
28	GE	Integer	≥ 0	Cases for specific canton per week
29	CH	Integer	≥ 0	Cases for the whole of

				Switzerland per week

Table 4: Data description for the 2_Data_cantons_incidence_weekly_56_58_NEW file

Index	Column Name	Data Type	Values / Validation	Description
1	Year	Integer	1950 – 1970	Observation Year
2	ZH	Integer	≥ 0	Population per Canton per Year
3	BE	Integer	≥ 0	Population per Canton per Year
4	LU	Integer	≥ 0	Population per Canton per Year
5	UR	Integer	≥ 0	Population per Canton per Year
6	SZ	Integer	≥ 0	Population per Canton per Year
7	OW	Integer	≥ 0	Population per Canton per Year
8	NW	Integer	≥ 0	Population per Canton per Year
9	GL	Integer	≥ 0	Population per Canton per Year
10	ZG	Integer	≥ 0	Population per Canton per Year
11	FR	Integer	≥ 0	Population per Canton per Year
12	SO	Integer	≥ 0	Population per Canton per Year
13	BS	Integer	≥ 0	Population per Canton per Year
14	BL	Integer	≥ 0	Population per Canton per Year
15	SH	Integer	≥ 0	Population per Canton per Year
16	AR	Integer	≥ 0	Population per Canton per Year
17	AI	Integer	≥ 0	Population per Canton per Year
18	SG	Integer	≥ 0	Population per Canton per Year

19	GR	Integer	≥ 0	Population per Canton per Year
20	AG	Integer	≥ 0	Population per Canton per Year
21	TG	Integer	≥ 0	Population per Canton per Year
22	TI	Integer	≥ 0	Population per Canton per Year
23	VD	Integer	≥ 0	Population per Canton per Year
24	VS	Integer	≥ 0	Population per Canton per Year
25	NE	Integer	≥ 0	Population per Canton per Year
26	GE	Integer	≥ 0	Population per Canton per Year
27	CH	Integer	≥ 0	Population per Canton per Year

Table 5: Data description for the 2_Population_cantons file

Dataquality Dataset 2

Dataquality for data_set2_mortality

Methodology

- The analysis was performed using automated data profiling tools such as ydata

General Statistics

- Number of variables: 29
- Variable types: 2 Categorical & 27 Numeric
- Number of observations: 144
- Time span: 1953 – 1958
- Missing cells: 2.5% of all values
- No duplicate rows or invalid entries
- Mean year: 1955.5
- Mean mortality for the whole of Switzerland per Month: 2069.9143

Key Variables and Values

- Month: categorical, January – December, missing = 0
- Year: numerical, 1953 – 1958, mean = 1955.5, missing = 0
- Parameter: categorical, 3 Distinct values, missing = 0
- All cantons (26 different columns): numerical, missing 4 (2.8%)
- CH: numerical, missing 4 (2.8%), mean 2069.9143

Interpretation and Insights

- The data gives insight on how much the flu raised the mortality rates compared to other flu seasons and over the whole time
- Gives a good overall picture to how bad the Asian flu was and how much the population was impacted

Data Quality Assessment

- Completeness: moderate (data missing form September – December 1958)
- Consistency: high
- Usability: good for the analysis of the impact of the asian flu on the mortality

Conclusion

- Mortality Trends specific for the period before and during the Asian flu
- Data-Driven storytelling

Dataquality for data_set2_population

Methodology

- The analysis was performed using automated data profiling tools such as ydata

General Statistics

- Number of variables: 30
- Variable types: 2 DateTime, 27 Numeric, 1 Categorical
- Number of observations: 314
- Time span: 14.12.1952 – 14.12.1958
- Missing cells: 0.0% of all values
- No duplicate rows or invalid entries

Key Variables and Values

- StartReportingPeriod: DateTime, 14.12.1952 – 14.12.1958, missing = 0
- EndReportingPeriod: DateTime, 20.12.1952 – 20.12.1958, missing = 0
- Month: numerical, 1-12, missing = 0
- Parameter: categorical, cases Influenza
- All cantons (26 different columns): numerical, missing 0
- CH: numerical, missing 0, mean 1143.5

Interpretation and Insights

- The data gives insight on how much the flu impacted Switzerland as a whole and canton specific
- Gives a good overall picture to how bad the Asian flu was and how much the population was impacted

Data Quality Assessment

- Completeness: high
- Consistency: high
- Usability: good for the analysis if the Asian flu made a big impact on the overall population.

Conclusion

- Analysis on the Population of Switzerland before and during the Asian flu
- Data-Driven storytelling

Dataquality for data_set2_incidence_weekly

Methodology

- The analysis was performed using automated data profiling tools such as ydata

General Statistics

- Number of variables: 27
- Variable types: 27 Numeric
- Number of observations: 21
- Time span: 1950 – 1970
- Missing cells: 0.0% of all values
- No duplicate rows or invalid entries
- Mean year: 1955.5

Key Variables and Values

- Year: numerical, 1950 – 1970, mean = 1960, missing = 0
- All cantons (26 different columns): numerical, missing 0
- CH: numerical, missing 0, mean 5452523.8

Interpretation and Insights

- The data is really good for comparing the difference between the Asian flu and other flu seasons
- Gives a good overall picture to how bad the Asian flu was to a normal flu

Data Quality Assessment

- Completeness: high
- Consistency: high
- Usability: good to compare the Asian flu cases to other flu seasons

Conclusion

- Analysis of the Influenza cases before, during and after the Asian flu
- Data-Driven storytelling

Details Dataset 3

Description:

- Annual mortality statistics in Switzerland from 1880 to 2022, including pandemic-related deaths and excess mortality estimates.

Source Details:

- For more information please contact Wiebke Weber or Kaspar Staub

Data Acquisition:

- File: 3_Todesursachen Schweiz ohne Alter 1876-2002.xlsx, contact Wiebke Weber

Legal Aspects:

- The data is publicly available and free to use for educational and analytical purposes.

Data Governance:

- Category: Public
- No personal data is included.

Access Information:

- Provided as an Excel file. On GitHub, linked at Speicherort.
- Can be accessed using common tools (Excel, Python pandas, etc.)

Data catalog Dataset 3

Index	Column Name	Data Type	Values / Validation	Description
1	Jahr	Float	1876 - 2004	Durchschnittliches Beobachtungsjahr
2	Todesfälle_Infektionen_Total	Float	≥ 0	Gesamt: Infektions- und parasitäre Krankheiten
3	Pocken	Float	≥ 0	Pocken-Todesfälle
4	Scharlach	Float	≥ 0	Scharlach-Todesfälle
5	Masern	Float	≥ 0	Masern-Todesfälle
6	Typhus_Paratyphus	Float	≥ 0	Typhus und Paratyphus
7	Diphtherie	Float	≥ 0	Diphtherie-Todesfälle
8	Keuchhusten	Float	≥ 0	Keuchhusten-Todesfälle
9	Tuberkulose_Total	Float	≥ 0	Alle Tuberkulosefälle
10	Tuberkulose_Lungen	Float	≥ 0	Tuberkulose der Lungen
11	Tuberkulose_Andere_Organe	Float	≥ 0	Tuberkulose anderer Organe

12	Tuberkulose_Gehirnhaut	Float	≥ 0	Tuberkulose der Hirnhaut
13	Tuberkulose_Knochen	Float	≥ 0	Tuberkulose der Knochen und Gelenke
14	Übrige_Krankheiten_Total	Float	≥ 0	Sonstige Infektionskrankheiten
15	Epidem_Kinderlähmung	Float	≥ 0	Kinderlähmung
16	Rotlauf	Float	≥ 0	Rotlauf-Todesfälle
17	Syphilis	Float	≥ 0	Syphilis-Todesfälle
18	Aids	Float	≥ 0	Aids-Todesfälle
19	Atmungsorgane_Total	Float	≥ 0	Todesfälle durch Erkrankungen der Atmungsorgane
20	Grippe	Float	≥ 0	Grippe-Todesfälle
21	Lungenentzündung	Float	≥ 0	Lungenentzündung
22	Bronchitis	Float	≥ 0	Bronchitis
23	Neubildungen_Krebs_Total	Float	≥ 0	Bösartige Neubildungen insgesamt
24	Krebs_Gesamt	Float	≥ 0	Gesamtanzahl Krebsfälle
25	Krebs_Magen_Darm	Float	≥ 0	Magen-Darm-Krebs
26	Krebs_Weibl_Brustdrüse	Float	≥ 0	Brustkrebs (weiblich)
27	Krebs_Bronchien_Lunge	Float	≥ 0	Lungenkrebs/Bronchien
28	Nervensystem_Kreislauf_Total	Float	≥ 0	Gesamt: Nerven- und Kreislauferkrankungen
29	Nervensystem_Total	Float	≥ 0	Nervenkrankheiten insgesamt
30	Hirnschlag	Float	≥ 0	Schlaganfall
31	Kreislauferkrankungen_Total	Float	≥ 0	Herz-Kreislauf-Erkrankungen insgesamt
32	Herzkrankheiten	Float	≥ 0	Herzkrankheiten
33	Arterienverkalkung	Float	≥ 0	Arteriosklerose
34	Blut_Stoffwechsel_Total	Float	≥ 0	Erkrankungen des Blutes und Stoffwechsels
35	Anämie_Leukämie	Float	≥ 0	Anämie & Leukämie
36	Diabetes	Float	≥ 0	Diabetes mellitus
37	Verdauungsorgane_Total	Float	≥ 0	Erkrankungen des Verdauungssystems
38	MagenDarmKatarrh_Kinder	Float	≥ 0	Magen-Darm-Katarrh bei Kindern
39	MagenDarmKatarrh_Erwachsene	Float	≥ 0	...bei Erwachsenen
40	MagenDarm_Geschwür	Float	≥ 0	Magen-Darm-Geschwüre
41	Blinddarmentzündung	Float	≥ 0	Appendizitis
42	Leberzirrhose	Float	≥ 0	Leberzirrhose
43	Harnapparat_Total	Float	≥ 0	Krankheiten des Harnapparats
44	Nierenentzündung	Float	≥ 0	Nierenentzündung

45	Gewaltsamer_Tod_Total	Float	≥ 0	Gewaltsamer Tod insgesamt
46	Selbsttötung	Float	≥ 0	Suizid
47	Unfälle_Total	Float	≥ 0	Unfalltote insgesamt
48	Verkehrsunfälle_Gesamt	Float	≥ 0	Verkehrsunfälle gesamt
49	Verkehrsunfälle_Strasse	Float	≥ 0	Straßenverkehrsunfälle
50	Verkehrsunfälle_Bahn	Float	≥ 0	Bahnverkehrsunfälle
51	Verkehrsunfälle_Flug	Float	≥ 0	Flugunfälle
52	Sturz	Float	≥ 0	Sturzunfälle
53	Ertrinken	Float	≥ 0	Ertrinken
54	Mord_Totschlag_Total	Float	≥ 0	Mord/Totschlag + unklare Ursachen
55	Mord_Totschlag	Float	≥ 0	Mord und Totschlag
56	Übrige_Krankheiten_Total	Float	≥ 0	Übrige bekannte Ursachen
57	Chronischer_Alkoholismus	Float	≥ 0	Alkoholismus
58	Schwangerschaft_Geburt	Float	≥ 0	Schwangerschaft & Geburt
59	Angeborene_Krankheiten_Säuglinge_Total	Float	≥ 0	Angeborene Säuglingskrankheiten
60	Frühgeburt	Float	≥ 0	Frühgeburt
61	Alterschwäche	Float	≥ 0	Altersschwäche
62	Todesursache_Arzt_beigezogen	Float	≥ 0	Arzt beigezogen
63	Todesursache_Arzt_nicht_beigezogen	Float	≥ 0	Arzt nicht beigezogen
64	Todesursache_Unbekannt_Total	Float	≥ 0	Todesursache unbekannt (gesamt)
65	Todesfälle_Gesamt	Float	≥ 0	Gesamttodesfälle

Table 6: Data description for the 3_Todesursachen Schweiz ohne Alter 1876-2002 table

Dataquality Dataset 3

Methodology

- The analysis was performed using automated data profiling tools such as ydata.
- Descriptive statistics and a Pearson correlation matrix were used.
- The dataset includes 250 annual observations.
- There are 82 variables total: all numerical

General Statistics

- Number of variables: 82
- Number of observations: 240
- Time span: 1876–2002
- Missing cells: 25.1% of all values
- 1 duplicate rows
- 81 unsupported variable types
- No statistical measures can be calculated

Key Variables and Values

- Jahr: Average: 1939.8

Infektions-und parasitäre Krankheiten:

- Pocken, Scharlach, Masern, Typhus, Diphtherie, Keuchhusten: Title
    - Total: Average 825
    - Pocken: Average: 31.6
    - Scharlach 1): Average: 98
    - Masern 2): Average: 134
    - Typhus Paratyphus: Average: 142
    - Diphtherie: Average: 386
    - Keuchhusten: Average: 232
- Tuberkulose
    - Total: Average: 3’163
    - Atmungsorgane
        - Total: Average: 2’417.1
        - Lungen-Tb: Average: 2’331
    - Übrige Organe
        - Total: Average: 776.5
        - Gehirnhaut, Hirnhaut: Average: 388
        - Knochen, Gelenke: Average: 273.4
- Übrige Krankheiten 4)
    - Total: Average: 607
    - Epidem. Kinderlähmung: Average: 29.5
    - Rotlauf: Average: 65
    - Syphilis: Average: 88.5
    - Aids: Average: 228
- Total: Average: 4’581

Atmungsorgane:

- Total: Average: 4’917
- Grippe: Average: 630.3
- Lungenentzünding: Average: 2’273
- Bronchitis: Average: 689.8

Jahr: Average: 1939.8

Neubildungen:

- Total Krebs (bösartige Geschwülste); Average: 7’058.1
    - Total: Average: 5’492
    - Magen, Darm (Total): Average: 2’430.6
    - Weibl. Brustdrüse: Average: 596
    - Bronchien, Lunge 7): Average: 831.6

Nervensystem und Kreislauforgane

- Total : Average: 15’009
- Nervensystem:
    - Total: Average: 2’309.6
    - Hirnschlag: Average: 1’845
- Kreislauforgane
    - Total: Average: 13’385.6
    - Herzkrankeiten: Average: 6’295
    - Arterienverkalkung: Average: 4’778.4

Blut, Stoffwechsel

- Total: Average: 1’117
- Anämie, Leukämie: Average: 278.7

- Zuckerkrankheit (Diabetes mellitus): Average: 569

Ernährungsapparat, Verdauuungsorgane

- Total: Average: 3’116.9
- Magen-Darm-Kattarrh
  - Säuglinge und Kleindkinder 9): Average: 1’155
  - Erwachsene: Average: 312.1
- Magen-Darm-geschwür: Average: 250
- Blinddarm entzündung: Average: 176.2
- Leberzirrhose: Average: 443

Harnapparat

- Total: Average: 951.7
- Nierenentzündungen: Average: 686

Jahr: Average: 1939.8

Gewaltsamer Tod:

- Total: Average: 2’989
- Selbsttötung: Average: 844.5
- Unfälle
  - Total 12): Average: 2’010
  - Verkehrsunfälle:
    - Total: Average: 577.2
    - Strassenverkehr: Average: 705
    - Bahnverkehr: Average: 70.7
    - Flugverkehr; Average: 17
  - Sturz: Average: 672.1
  - Ertrinken: Average: 183
- Mord, Totschlag, unklare Ursache
  - Total: Average: 135.2
  - Mord, Totschlag: Average: 55

Übrige bekannte Todesursachen

- Total: Average: 4’296.4
- Chronischer Alkoholismus: Average: 167
- Schwangerschaft u. Geburt: Average: 229.7
- Angeborene Säuglingskrankheiten Total: Average: 1’775
  - Frühgeburt 14): Average: 1’301.8
- Alterschwäche; Average: 965

Unbekannte Todesursachen

- Arzt beigezogen: Average: 437.7
- Arzt nicht beigezogen: Average: 1’653
- Total: Average: 1’973.5

Gesamttotal: Average: 45’535

Interpretation and Insights

- Ideal for looking at different causes of death in more or less detail

Data Quality Assessment

- Completeness: poor- moderate (lot of missing values).
- Consistency: moderate (no duplicates or major outliers).

- Usability: poor, because of how the title are distributed (unclear for when reading the table)

Conclusion

- Insight to cause of death
- Data-driven storytelling

Details full_data

Description:

- Daily COVID-19 case and death statistics from various global locations, including cumulative and rolling aggregates (weekly, biweekly), spanning the pandemic timeline.

Source Details:

- Data sourced from a global COVID-19 tracking initiative (e.g., Our World in Data, Johns Hopkins University).

Data Acquisition:

- File: full_data.csv, in the GitHub Repo

Legal Aspects:

- The dataset is public

Data Governance:

- Category: Public
- No personal data is included.

Access Information:

- Provided as a CSV file.
- Can be accessed using common tools (Excel, Python pandas, R, etc.)

Datacatalog full_data:

Index	Column Name	Data Type	Values / Validation	Description
1	date	Date	ISO format (YYYY-MM-DD)	Date of record
2	location	String	Valid location names	Country or region name
3	new_cases	Integer	≥ 0	New reported COVID-19 cases (daily)
4	new_deaths	Integer	≥ 0	New reported COVID-19 deaths (daily)
5	total_cases	Integer	≥ 0	Total confirmed cases to date
6	total_deaths	Integer	≥ 0	Total confirmed deaths to date
7	weekly_cases	Integer	≥ 0	Aggregated cases over the past 7 days

8	weekly_deaths	Integer	≥ 0	Aggregated deaths over the past 7 days
9	biweekly_cases	Integer	≥ 0	Aggregated cases over the past 14 days
10	biweekly_deaths	Integer	≥ 0	Aggregated deaths over the past 14 days

Table 7: Data description for the full_data table

Dataquality full_data

Methodology

- Data collected and processed by a centralized public health data provider.
- Descriptive analysis done via automated profiling tools.
- Rolling aggregates (weekly, biweekly) computed from raw daily data.
- Dataset includes over 411,000 rows across 10 variables.

General Statistics

- Number of variables: 10
- Number of observations: 411,804
- Duplicate rows: 0
- Primary key: (date, location)

Key Variables and Values

- new_cases: varies widely, occasional spikes during outbreak peaks, 0.4% missing.
- new_deaths: similar pattern, 0.29% missing.
- weekly_cases, weekly_deaths: ~0.6–0.7% missing.
- biweekly_cases, biweekly_deaths: ~1% missing.
- Cumulative fields (total_cases, total_deaths) are complete.

Interpretation and Insights

- Highly suitable for time-series analysis of pandemic spread and impact.
- Useful for comparing countries/regions and identifying outbreak waves.
- Rolling aggregates provide smoothing for trend analysis.

Data Quality Assessment

- **Completeness**: High (less than 1.2% missing overall)
- **Consistency**: High (no duplicates, logical numeric values)
- **Usability**: Excellent for trend analysis, comparison, and modeling

Overview of Processed Data:

Dataset Name	Source	Storage location
data_set3_cleaned	Dataset 3	Can be found in the Github Repo under the name data_set3_cleaned: Project_PODSV/Data at main · plospen1/Project_PODSV
data_set3_infectious_diseases	Dataset 3	Can be found in the Github Repo under the name data_set3_infectious_diseases

		Project_PODSV/Data at main · plospen1/Project_PODSV

Table 8: Summary Table of the Processed Datasets

Details Processed data_set3_cleaned

Description:

- This dataset contains yearly death statistics in Switzerland (1876–2004), categorized by causes such as infectious diseases (e.g., measles, tuberculosis), respiratory diseases, cancers, cardiovascular issues, and accidents

Processing Steps:

- Headers: The original table had multi-row headers (rows 3–5). These were merged using forward fill and joined with " | " to create clear column names.
- Data Extraction: Actual data starts at row 10 (index 9). Earlier rows were removed.
- Cleanup: Index was reset and column names verified.

Access and Tools:

- Available at:
  GitHub: Project_PODSV/Data at main · plospen1/Project_PODSV
- Tools: Python and pandas
- Script: Processing done via a reproducible script using standard pandas functions.

Details Processed data_set3_infectious_diseases

Description:

- This dataset contains annual death counts in Switzerland (1876–1891) specifically related to infectious diseases. It includes categories such as smallpox, scarlet fever, measles, typhoid/paratyphoid, diphtheria, and whooping cough, alongside total infectious disease-related deaths per year.

Processing Steps:

- Header Extraction: Column names were taken from row 6 (index 5) of the Excel sheet. These were cleaned of empty and duplicate names, with placeholders added where needed.
- Data Selection: The actual data starts from row 8 (index 7). Only these rows were kept.
- Filtering: Only rows with valid numeric years were retained. The “Year” column was converted to int type.
- Column Selection: Only the first 8 columns (A–H) were kept, which correspond to the relevant disease categories.
- Renaming Columns: All columns were renamed to English names

Access and Tools:

- Available at: GitHub: Project_PODSV/Data ,File name: data_set3_infectious_diseases
- Tools: Python and pandas
- Script: Processing done via a reproducible script using pandas, including header cleaning, filtering valid rows, renaming, and exporting to .csv.

Bibliography:

1957-1958 Pandemic (H2N2 virus) | Pandemic Influenza (Flu) | CDC (2019). Available at: https://archive.cdc.gov/www_cdc_gov/flu/pandemic-resources/1957-1958-pandemic.html (Accessed: 29 May 2025).

Bevölkerungsdaten im Zeitvergleich, 1950-2023 - 1950-2023 | Tabelle. Available at: https://www.bfs.admin.ch/asset/de/32374798 (Accessed: 29 May 2025).

CDC (2024a) 1880-1959 Highlights in the History of Avian Influenza (Bird Flu) Timeline, Avian Influenza (Bird Flu). Available at: https://www.cdc.gov/bird-flu/avian-timeline/1880-1959.html (Accessed: 29 May 2025).

CDC (2024b) About Pandemic Influenza, Pandemic Flu. Available at: https://www.cdc.gov/pandemic-flu/basics/index.html (Accessed: 29 May 2025).

CDC (2024c) Pandemic Flu, Pandemic Flu. Available at: https://www.cdc.gov/pandemic-flu/index.html (Accessed: 29 May 2025).

GBD Results (no date) Institute for Health Metrics and Evaluation. Available at: https://vizhub.healthdata.org/gbd-results (Accessed: 29 May 2025).

Gupta, R.K., George, R.C. and Nguyen-Van-Tam, J.S. (no date) 'Bacterial Pneumonia and Pandemic Influenza Planning - Volume 14, Number 8—August 2008 - Emerging Infectious Diseases journal - CDC'. Available at: https://doi.org/10.3201/eid1408.070751.

History of influenza vaccination (no date). Available at: https://www.who.int/news-room/spotlight/history-of-vaccination/history-of-influenza-vaccination (Accessed: 29 May 2025).

Morens, D.M., Taubenberger, J.K. and Fauci, A.S. (2008) 'Predominant Role of Bacterial Pneumonia as a Cause of Death in Pandemic Influenza: Implications for Pandemic Influenza Preparedness', The Journal of infectious diseases, 198(7), pp. 962–970. Available at: https://doi.org/10.1086/591708.

Smallpox (no date). Available at: https://www.who.int/health-topics/smallpox (Accessed: 29 May 2025).

Specific causes of death (no date). Available at: https://www.bfs.admin.ch/content/bfs/en/home/statistics/health/state-health/mortality-causes-death/specific.html (Accessed: 29 May 2025).

Threats, I. of M. (US) F. on M. et al. (2005) 'The Story of Influenza', in The Threat of Pandemic Influenza: Are We Ready? Workshop Summary. National Academies Press (US). Available at: https://www.ncbi.nlm.nih.gov/books/NBK22148/ (Accessed: 29 May 2025).

Vaccine-preventable diseases (2015). Available at: https://www.ecdc.europa.eu/en/vaccine-preventable-diseases (Accessed: 29 May 2025).

Vaccines and immunization EURO (no date a). Available at: https://www.who.int/europe/health-topics/vaccines-and-immunization (Accessed: 29 May 2025).

Vaccines and immunization EURO (no date b). Available at: https://www.who.int/europe/health-topics/vaccines-and-immunization (Accessed: 29 May 2025).

Wei, Y. et al. (2019) 'Efficient synthesis of polysaccharide with high selenium content mediated by imidazole-based acidic ionic liquids', Carbohydrate Polymers, 203, pp. 157–166. Available at: https://doi.org/10.1016/j.carbpol.2018.09.047.
Zhdanov, P.V.M. (1957) 'The 1957 Influenza Pandemic in the USSR'.

Table of Figures: