



Datenbeschreibung

Bildungsstand – Geodaten

Ausgabe 2025

Herausgeber: Bundesamt für Statistik (BFS)
Auskunft: geostat@bfs.admin.ch
Redaktion: BILD-S, GEO, POP
Themenbereich: 00 Statistische Grundlagen, 02 Bevölkerung

Konzept Layout: Sektion PUB
Download: www.statistik.ch
Copyright: BFS, Neuchâtel 2025
Wiedergabe unter Angabe der Quelle für nicht-kommerzielle Nutzung gestattet

Inhaltsverzeichnis

1.	Einführung	3
2.	Räumliche Gliederung	3
3.	Methodik	3
4.	Datenqualität und Datenschutz	4
5.	Datenstruktur	5
6.	Variablenbeschreibung	5
7.	Anwendungen	7
8.	Kontakt	7
9.	Referenzen	7

1. Einführung

Das Produkt «Bildungsstand – Geodaten» liefert schweizweit feinräumige Informationen über den höchsten Bildungsabschluss der ständigen Wohnbevölkerung im Alter von 25 bis 64 Jahren. Die Datengrundlage bilden mehrere statistische Erhebungen der Bundesverwaltung und werden durch ein zweistufiges Imputationsmodell ergänzt, das fehlende Angaben schätzt (OFS, 2025f).

Das Bildungsniveau wird in den Geodaten nach einer dreistufigen Nomenklatur (BFS, 2025a) ausgewiesen:

- Obligatorische Schule;
- Sekundarstufe II;
- Tertiärstufe.

Die Daten stehen für die Referenzjahre 1990 und 2000 sowie jährlich ab 2017 zur Verfügung. Für jedes Referenzjahr werden sowohl der aktuelle Stand als auch eine Zeitreihe, die alle verfügbaren Jahre umfasst, publiziert. Der aktuelle Referenzstand liegt jeweils mindestens zwei Jahre zurück, die Aktualisierung erfolgt jährlich. Die Datensätze werden als CSV-Dateien mit Koordinaten sowie als GeoPackage im Schweizer Referenzsystem LV 95 (EPSG:2056) bereitgestellt und lassen sich direkt in GIS-Systeme einlesen und weiterverarbeiten.

2. Räumliche Gliederung

2.1 Abgedecktes Gebiet

Das Produkt deckt das gesamte Territorium der Schweiz ab. Kleinräumige Analysen des Bildungsniveaus werden durch die Verwendung der Hektarauflösung ermöglicht; gleichzeitig sind aggregierte Informationen auf der Ebene von USPAT1-Regionen und Gemeinden verfügbar.

2.2 Räumliche Entitäten

Basierend auf den Daten zum Bildungsniveau in der feinsten räumlichen Auflösung – nach metrischen Gebäudekoordinaten (vgl. Kapitel 3.3 Geocodierung) – werden räumliche Aggregationen nach Hektar sowie nach USPAT1-Regionen und Gemeinden für die Datendiffusion aufbereitet.

Hektaren

Jeder Hektar wird durch die Koordinaten seines südwestlichen Eckpunkts («E_KOORD»/»N_KOORD» nach LV95) identifiziert. Die Daten werden als 100×100-Meter-Raster publiziert. Die ausgewiesenen Werte beziehen sich auf das Vorkommen innerhalb der jeweiligen Fläche und stellen Aggregationen der Einzelrecords dar, die einer Rasterzelle beziehungsweise einem Hektar zugeordnet sind.

USPAT1-Regionen

Für die gesamte Schweiz wurden 3607 statistische Grundeinheiten der ersten Stufe (USPAT1 für «unités spatiales statistiques de base de premier niveau») festgelegt (BFS, 2025b). Die Aggregation der Informationen zum Bildungsniveau in diesen räumlichen Einheiten ermöglichen detailliertere regionale Analysen als dies auf Gemeindeebene möglich ist.

Gemeinden

Die Aggregation der Daten auf Gemeindeebene erfolgt anhand der Gemeindegrenzen des Produkts swissBOUNDARIES3D (swisstopo, 2025). Da sich Gemeindegrenzen über die Zeit verändern können, wird jeweils der am 1. Januar des Publikationsjahres gültige Zustand verwendet.

2.3 Koordinatensystem

Alle Koordinaten werden im Schweizer Koordinatensystem LV 95 (EPSG:2056) angegeben. Die zur Aggregation benutzten Geometrien der USPAT1-Regionen und Gemeinden liegen ebenfalls in diesem Bezugsrahmen vor.

3. Methodik

Die vorliegenden Geodaten entstehen in einem mehrstufigen Prozess. Er umfasst die Zusammenführung verschiedener Datenquellen, die Behandlung fehlender Werte (Imputation) sowie die Geokodierung über das Gebäude- und Wohnungsregister.

3.1 Datenquellen

Das Bildungsniveau der ständigen Wohnbevölkerung wird jährlich aus einer Vielzahl von Quellen ermittelt (OFS, 2025f). Die verschiedenen Quellen werden plausibilisiert und zusammengeführt, um ein möglichst vollständiges Bild der Bildungsniveaus zu erhalten. Siehe OFS 2025f für mehr Informationen.

3.2 Imputation und fehlende Werte (Missings)

Das Bildungsniveau von Personen ohne direkt zugeordnete Angabe wird auf Basis der Strukturhebung sowie zusätzlicher Variablen aus verschiedenen Registern imputiert. Insgesamt sind rund 10 Prozent der Angaben zur höchsten abgeschlossenen Ausbildung imputiert (OFS, 2025f).

Seit dem Referenzjahr 2017 liegt für jede Person entweder über eine Quelle (ca. 90 Prozent) oder durch Imputation (ca. 10 Prozent) ein Bildungsstand vor. Für die Referenzjahre 1990 und 2000 bestehen hingegen einige fehlende Werte

oder Missings, die nicht imputiert werden können. Die Veröffentlichung der Daten 1990 oder 2000 erfolgt pro räumliche Entität nur dann, wenn für mehr als 50 Prozent der Personen ein Bildungsniveau ermittelt werden kann.

3.3 Geokodierung

Die Angaben zum Bildungsniveau sind auf der Basis des eidgenössischen Gebäude- und Wohnungsregisters (GWR) geokodiert. Das GWR enthält neben den Gebäudekoordinaten auch den eidgenössischen Gebäudeidentifikator (EGID) (BFS, 2025c), der ebenfalls in den Personenregistern geführt wird. Dadurch können Personen – und damit auch ihr Bildungsniveau – räumlich verortet werden.

Der Koordinatengrundbestand des GWR setzt sich aus Koordinaten aus den vollumfänglich geokodierten Volkszählungen 1990 und 2000 sowie aus den Koordinaten der Gebäude der Betriebszählungen des sekundären und tertiären Sektors zwischen 1995 und 2008 beziehungsweise der Gebäude der in den Betriebszählungen zwischen 1996 und 2008 erfassten Landwirtschaftsbetriebe zusammen. Seit 2010 werden die GWR-Koordinaten laufend erhoben, ergänzt sowie bei Bedarf zusammen mit der amtlichen Vermessung revidiert und optimiert.

Optimierung durch aktuelle GWR-Daten

Da sich die Geokodierung im GWR laufend verbessert, werden die Gebäudekoordinaten des Publikationsjahres verwendet. Voraussetzung ist, dass die Bildungsdaten einem Gebäude (EGID) zugeordnet werden können und dass die aktuelle GWR-Koordinate (Publikationsjahr) nicht mehr als 200m von der zum Erhebungsjahr der Bildungsdaten gültigen Koordinate abweicht. Liegt die Differenz über 200m, bleibt die ursprüngliche Koordinate des Erhebungsjahres bestehen, da in solchen Fällen nicht nur Verbesserungen der Geokodierung, sondern auch Korrekturen vorliegen können.

Für das Erhebungsjahr 1990 wurde die Zuordnung der Personen zu Gebäuden über den VZ-CODE (Gemeindennummer + Zählkreis + Gebäudenummer) hergestellt und mittels Mapping-Table dem aktuellen EGID zugeordnet, da damals der EGID noch nicht existierte.

Gemeinde- und USPAT-Zuordnung

Die Angaben zu Gemeinde und USPAT stammen ebenfalls grösstenteils aus dem GWR (Stand des Publikationsjahres). In Fällen, in denen eine direkte Verknüpfung mit aktuellen GWR-Daten via EGID nicht möglich ist (insbesondere in den Referenzjahren 1990 und 2000), werden diese Informationen zusätzlich durch einen räumlichen Verschnitt mit den im Publikationsjahr gültigen Geometrien der Gemeinden (gemäss swissBOUNDARIES3D) und USPAT1-Regionen ergänzt.

Sammelhektaren

Personen ohne Adresse in der Gemeinde bzw. Personen, für die kein EGID vorhanden ist, können nicht lokalisiert und somit auch nicht geokodiert werden. Um die Vollständigkeit der

Erhebung zu gewährleisten, wurden diese Personen gemeindefeise der sogenannten Zentrumsordinate der Gemeinde (Sammelhektare) zugewiesen. Die Daten dieser Sammelhektaren können in kleinräumigen Analysen unerwünschte Einflüsse haben oder Ergebnisse verfälschen. Deshalb wurden sie aus dem Hektarraster und den USPAT1-Regionen entfernt. In der Gemeindeaggregation sind sie hingegen enthalten.

4. Datenqualität und Datenschutz

4.1 Datenschutz

Zum Schutz von Informationen über Individuen gelten für die Veröffentlichung der Angaben zum höchsten Bildungsstand (drei Niveaus: obligatorische Schule, Sekundarstufe II, Tertiärstufe) folgende Regeln:

Mindestens 10 Personen pro Entität

Angaben werden unterdrückt, wenn in einer räumlichen Entität (Hektar, Gemeinde oder USPAT1) weniger als 10 Personen der Wohnbevölkerung 25–64 Jahre leben.

Klassenbildung

Die Anzahl der Personen pro Bildungsniveau wird nicht in absoluten Werten, sondern als prozentuale Anteile der Gesamtzahl der Personen der Wohnbevölkerung im Alter von 25 bis 64 Jahren (pro räumlicher Einheit) angegeben. Diese prozentualen Anteile der drei Bildungsniveaus werden im Hektarraster in Klassen ausgewiesen. Die fünf Klassen sind:

- 1 = 0–<20 %;
- 2 = 20–<40 %;
- 3 = 40–<60 %;
- 4 = 60–<80 %;
- 5 = 80–100 %.

Auch die Anzahl Personen der Wohnbevölkerung 25–64 Jahre wird, in Hektarauflösung, in Klassen ausgewiesen:

- 1 = 10–19 Personen;
- 2 = 20–29 Personen;
- 3 = 30–49 Personen;
- 4 = 50–119 Personen;
- 5 = mehr als 120 Personen.

Prozentwerte für USPAT1 und Gemeinden ab 100 Personen

Prozentwerte zu den drei Bildungsniveaus werden nur publiziert, wenn in einer USPAT1-Region oder Gemeinde mindestens 100 Personen («BBTOT25_64») leben und nicht alle denselben Bildungsabschluss aufweisen. Andernfalls werden ausschliesslich Angaben nach Klassen publiziert.

4.2 Datenqualität

Die folgenden Punkte beschreiben die zentralen Qualitätsaspekte der Daten.

- **Räumliche und zeitliche Stabilität**

Die Verwendung von Hektarrastern und USPAT1-Regionen gewährleistet eine eindeutige und über die Zeit stabile Objektidentifikation.

- **Feinräumige Abdeckung (urbaner Fokus)**

In ländlichen Gebieten werden aufgrund geringerer Bevölkerungszahlen und der Datenschutzmassnahmen viele Hektaren unterdrückt; Publikationen auf Hektarebene betreffen daher vor allem urbane Gebiete.

- **Aggregierte Hektardaten**

Auf Gemeinde- und USPAT1-Ebene weichen die aus Hektardaten aggregierten Werte von den direkten USPAT1- und Gemeindedaten ab, da unterdrückte Hektaren (Suppression) nicht enthalten sind. Die aggregierten USPAT1-Werte entsprechen den Gemeindedaten abzüglich der Sammelhektaren (sofern keine weiteren Datenschutzmassnahmen greifen).

- **Imputation**

Rund 10 % der Werte zur höchsten abgeschlossenen Ausbildung sind imputiert. Die Imputation ist räumlich nicht zufällig verteilt: Einzelne Hektaren können ausschliesslich imputierte Werte enthalten, während die Anteile auf Gemeinde- und USPAT1-Ebene bis rund 50 % reichen können.

- **Fehlende Angaben (Missings)**

Ab 2017 liegt für jede Person eine Angabe zum höchsten Bildungsstand vor, entweder aus einer Quelle oder imputiert. Für die Erhebungsjahre 1990 und 2000 bestehen Missings («BBTOT25_64_MISS»). Publiziert werden pro räumliche Entität nur Werte, wenn mindestens die Hälfte der Angaben der Wohnbevölkerung («BBTOT25_64») der Entität vorhanden ist. Dadurch werden zusätzlich zu den Datenschutzmassnahmen einige weitere Hektaren unterdrückt. Für das Referenzjahr 2000 liegen für die Gemeinde Hérémence (BFS-Nummer 6084) sowie für einzelne USPAT1-Regionen daher keine Angaben zum Bildungsstand vor.

Enthält die Variablen mit Hektarkoordinaten (100m Rasterweite) im LV95. Das Koordinatenpaar «E_KOORD»/«N_KOORD» repräsentiert bei der Hektarauflösung die südwestliche Ecke einer Zelle. Für jede Zelle beziehen sich die ausgewiesenen Werte auf die gesamte 100×100m-Fläche.

- **GeoPackage (GPKG)**

Enthält zusätzlich zu den Variablen auch die vollständigen Geometrien: Polygone für Hektarzellen, USPAT1-Regionen und Gemeinden, wobei für die Gemeinden die generalisierten Geometrien K4 (BFS, 2025d) verwendet werden. Die Variablen im GPKG der Zeitreihendatei sind nach Erhebungsjahren (z. B. 1990, 2000, 2017+) aufgespalten. Jede Variable liegt somit in der Form VARIABLEN-NAME_JAHR vor, z. B. COMPULSORY_2017, SECONDARYII_2019 oder IMPUTATION_2000.

Dateinamen

Die Dateinamen sind so gestaltet, dass sie möglichst selbst erklärend sind.

Beispiele für das jeweils aktuelle Referenzjahr:

- *bildungsstand_YYYY_AGGR_2056.csv*
- *bildungsstand_YYYY_2056.gpkg* (enthält Layer für Hektaren, USPAT1-Regionen und Gemeinden)

wobei:

- **«YYYY»:** Referenzjahr «ERHJAHR»;
- **«AGGR»:** Räumliches Aggregationsniveau – *ha* für Hektaren, *gmde* für Gemeinden und *uspat* für USPAT1-Regionen;
- **«_2056»:** EPSG-Code des Koordinatensystems (LV 95).

Beispiel für eine Zeitreihendatei:

- *bildungsstand_AGGR_2056.csv*
- *bildungsstand_2056.gpkg* (enthält Layer für Hektaren, USPAT1-Regionen und Gemeinden)

wobei:

- **«AGGR»:** Räumliches Aggregationsniveau – *ha* für Hektaren, *gmde* für Gemeinden und *uspat* für USPAT1-Regionen;
- **«_2056»:** EPSG-Code des Koordinatensystems (LV 95).

5. Datenstruktur

5.1 Dateien und Dateinamen

Für jedes räumliche Aggregationsniveau – Hektaren, Gemeinden und USPAT1-Regionen – werden sowohl die jeweils aktuellen Geodaten als auch Zeitreihen-Dateien mit allen verfügbaren Referenzjahren bereitgestellt.

Formate

- **CSV (Trennzeichen «;»)**

6. Variablenbeschreibung

Die mitgelieferte Variablenliste ergänzt diese Beschreibung. Sie enthält unter anderem die Datentypen der Variablen.

6.1 Zeitvariablen

- **ERHJAHR: Erhebungsjahr**
Der Stichtag entspricht dem 31.12. des Erhebungsjahrs der Daten, auch Referenzjahr genannt.
- **PUBJAHR: Publikationsjahr**
Jahr der Veröffentlichung des Datensatzes. Der für die Aggregation verwendete Stand der Gemeinden und USPAT1-Regionen ist der Stand vom 1. Januar des Veröffentlichungsjahres.

6.2 Gebietsvariablen

Diese Variablen identifizieren das Gebiet, für das die Bildungsdaten veröffentlicht werden. Aggregationen nach USPAT1-Regionen und Gemeinden basieren auf dem am 1. Januar des Publikationsjahres gültigen Gebietszustand.

- **RELI: Primärschlüssel einer Hektare**
Primärschlüssel einer Hektare. Er besteht aus den Stellen 2 bis 5 der E- und N-Koordinaten (die ersten vier Stellen von «E_KOORD» und die zweiten vier Stellen von «N_KOORD»).
- **E_KOORD: Ost-Koordinate**
- **N_KOORD: Nord-Koordinate**
Koordinatenpaar der südwestlichen Ecke einer Hektare im LV 95.
- **U1_ID: USPAT1 Code**
Der Code der USPAT1 besteht aus der Ziffer 1 gefolgt von der einstelligen Nummer der Grossregion, der zweistelligen Kantonsnummer und der dreistelligen USPAT-Nummer (BFS, 2025b).
- **U1_NAME: USPAT1 Name**
Für die Namen der USPAT1 werden die Gemeinde- oder Quartiernamen oder die Namen aus der Datenbank der Ortschaften swissNAMES3D von swisstopo übernommen. Um Verwechslungen mit den Gemeindennamen zu vermeiden, wird dem Namen «USPAT1» vorangestellt.
- **GDENR: BFS-Gemeindenummer**
Die BFS-Gemeindennummer ist für jeden ab 1960 gegebenen Gemeindegstand (Referenzdatum) eindeutig. Die gleiche Nummer kann aber für einen anderen Gemeindebestand, der zu einem anderen Zeitpunkt erstellt wurde, wiederverwendet werden (BFS, Applikation der Schweizer Gemeinden, 2025e).
- **GDEHISTID: Historisierungsnummer Gemeinde**
Die Historisierungsnummer der Gemeinde ist ein nicht sprechender Identifikator, mit dem Einträge, die in der amtlichen Gemeindegstand gestrichen oder hinzugefügt wurden, eindeutig identifiziert werden

können.

6.3 Thematische Variablen

- **BBTOT25_64**
Ständige Wohnbevölkerung im Alter von 25–64 Jahren gemäss Statistik der Bevölkerung und der Haushalte (STATPOP) für die Erhebungsjahre ab 2017. Für die Erhebungsjahre 1990 und 2000 gemäss der eidgenössischen Volkszählung (VZ).

In Hektaraufösung werden die Angaben in Klassen ausgewiesen:
 - 1 = 10–19 Personen;
 - 2 = 20–29 Personen;
 - 3 = 30–49 Personen;
 - 4 = 50–119 Personen;
 - 5 = mehr als 120 Personen.
 In den USPAT1-Regionen und den Gemeinden gibt es aus Datenschutzgründen die Klasse «0–9 Personen». Diese wird mit dem Wert 9 gekennzeichnet.
- **BBTOT25_64_MISS**
Ständige Wohnbevölkerung im Alter von 25–64 Jahren «BBTOT25_64» ohne Angabe zum Bildungsstand. Diese Variable existiert nur für die Erhebungsjahre 1990 und 2000; ab 2017 gibt es keine fehlenden Werte (Missings) mehr.
- **COMPULSORY_C** (Obligatorische Schule, klassierte Werte)
- **SECONDARYII_C** (Sekundarstufe II, klassierte Werte)
- **TERTIARY_C** (Tertiärstufe, klassierte Werte)
Klassierte Anteile der dreistufigen Nomenklatur des Bildungsniveaus. Es werden fünf Klassen unterschieden:
 - 1 = 0–<20 %;
 - 2 = 20–<40 %;
 - 3 = 40–<60 %;
 - 4 = 60–<80 %;
 - 5 = 80–100 %.
- **COMPULSORY** (Obligatorische Schule, Prozentwerte)
- **SECONDARYII** (Sekundarstufe II, Prozentwerte)
- **TERTIARY** (Tertiärstufe, Prozentwerte)
Prozentuale Anteile des höchsten Bildungsniveaus. Die Werte sind so gerundet, dass die Summe immer 100 % ergibt. Dabei werden die

beiden Werte mit den grössten Rundungsresten normal gerundet; der dritte Wert wird als Differenz zu 100 % bestimmt.

- **IMPUTATION**

Prozentualer Anteil imputierter Werte (OFS, 2025f) des höchsten Bildungsniveaus pro Gebietseinheit.

7. Anwendungen

Die Daten zur höchsten abgeschlossenen Ausbildung der ständigen Wohnbevölkerung sind ein zentraler Indikator für das bildungsspezifische Humankapital der Schweiz. Angesichts des demografischen Wandels, der Alterung der Bevölkerung und des steigenden Fachkräftebedarfs ist ihre Analyse besonders wichtig.

Die Geodaten in Hektarauflösung ermöglichen Analysen und Visualisierungen mit sehr hoher räumlicher Auflösung. Sie machen Unterschiede zwischen Quartieren grosser Städte sichtbar und zeigen räumliche Muster, die auf kantonaler oder kommunaler Ebene verborgen bleiben (z. B. die Lage in der Nähe einer Verkehrsachse oder einer Hochschule). Zudem erlauben sie die feinräumige Untersuchung der Beziehung zu weiteren Variablen wie Arbeitsmarkt- oder demografischen Merkmalen und unterstützen die Analyse von Entwicklungen und Trends über die Zeit.

8. Kontakt

Wir hoffen mit diesem Produkt Ihren Bedürfnissen zu entsprechen und sind offen für Bemerkungen, Verbesserungsvorschläge oder Fehlermeldungen, welche Sie gerne an geostat@bfs.admin.ch richten können. Bei thematischen Fragen zur höchsten abgeschlossenen Ausbildung können sie sich gerne an eduperspectives@bfs.admin.ch wenden.

9. Referenzen

BFS. (2025a). *Höchste abgeschlossene Ausbildung STATPOP - Produktionsprozess, Kurzbeschreibung (do-d-15.10-EXP-01)*. Neuchâtel: Bundesamt für Statistik (BFS).

BFS. (August 2025b). *Räumliche Gliederungen: USPAT – Raumbezogene statistische Grundeinheiten*. Von <https://www.bfs.admin.ch/bfs/de/home/statistiken/querschnittsthemen/raeumliche-analysen/raeumliche-gliederungen/uspat.html> abgerufen

BFS. (August 2025c). *Eid. Gebäude- und Wohnungsregister (GWR)*. Von <https://www.housing-stat.ch/de/index.html> abgerufen

BFS. (August 2025d). *Generalisierte Gemeindegrenzen*. Von <https://www.bfs.admin.ch/bfs/de/home/dienstleistungen/geostat/geodaten-bundesstatistik/raumgliederungen/generalisierte-gemeindegrenzen.html> abgerufen

BFS. (September 2025e). *Applikation der Schweizer Gemeinden*. Von <https://www.agvchapp.bfs.admin.ch> abgerufen

OFS. (2025f). *Plus haute formation achevée dans la statistique STATPOP+ – Processus et méthode de production (DO-D-15.10-EXP-01)*. Neuchâtel: Office fédéral de la statistique (OFS).

swisstopo. (August 2025). *swissBOUNDARIES3D*. Von <https://www.swisstopo.admin.ch/de/landschaftsmodell-swissboundaries3d> abgerufen