

November 2024

SynPop 2022

Synthetische Bevölkerung der Schweiz 2022

Beschreibung

Die Synthetische Bevölkerung der Schweiz 2022, kurz SynPop, ist ein Datensatz, der aus 8,95 Millionen Personen (oder Agenten) besteht. Dieser Datensatz wurde für die Verwendung in den Verkehrsmodellen NPVM des ARE (Bundesamt für Raumentwicklung) und SIMBA MOBi der SBB (Schweizerische Bundesbahnen) entwickelt. Die Basisquelle für SynPop ist der STATPOP-Datensatz des Bundesamts für Statistik (BFS), der mit anderen Datensätzen und Quellen kombiniert wird, um verschiedene andere relevante Attribute zu modellieren (z. B. die Verfügbarkeit von Mobilitätsressourcen wie Autos oder Fahrkarten für den öffentlichen Verkehr). SynPop 2022 basiert zwar zum Teil auf amtlichen Volkszählungsdaten, ist aber kein brauchbarer Ersatz für STATPOP, da für die offene Veröffentlichung die Genauigkeit reduziert und die sensiblen Datenquellen anonymisiert wurden (siehe unten). Die meisten Personenmerkmale im Datensatz sind jedoch synthetisch (modelliert). Um zu verstehen, wie die synthetische Bevölkerung konstruiert wurde, lesen Sie bitte den entsprechenden Bericht. Dieser Datensatz wird öffentlich zugänglich gemacht, in der Hoffnung, die Zusammenarbeit zwischen Forschung und Wirtschaft in der Schweiz zu fördern und die Entscheidungsfindung im öffentlichen Bereich zu verbessern.

Anonymisierung

Die Anonymisierung der Daten besteht in der Aggregation sensibler Attribute und in der räumlichen Randomisierung. Die räumliche Randomisierung erfolgt durch den Austausch der Wohnorte von Personen innerhalb eines bestimmten Gebiets - definiert durch ein Cluster von mindestens 1500 Personen und nicht durch die ursprüngliche Gebiets-ID. In Kombination tragen beide Maßnahmen dazu bei, das Offenlegungsrisiko von Einzelpersonen zu verringern.

Quellen und Referenzen

1. Das mikroskopische Transportmodell der SBB: [SIMBA MOBi](#)
2. Das NPVM-Verkehrsmodell des ARE (Website auf Deutsch, Französisch und Italienisch): www.are.admin.ch/npvm
3. Website zu den Aktivitäten des ARE zur Landnutzungsmodellierung: www.are.admin.ch/flnm
4. FaLC-Landnutzungsmodell: <http://www.falc-sim.org/en/home.html>
5. Offizielle Volkszählungserhebung: [STATPOP](#)

Kontakt

- SBB: Davi Guggisberg, davi.guggisberg@sbb.ch
- ARE: Thomas Schatzmann, thomas.schatzmann@are.admin.ch

Variable	Type	Categories	Description	Source data	Source variable	Treatment
person_id	Integer	-	Unique person identifier	synthetic	person_id	None (identity)
sex	String	F, M	Sex	STATPOP	sex	Simple recoding
age	String	0-17,18-24,25-44,45-64,65-74,75+	Age (grouped)	STATPOP	dbirth	Converted to age and grouped
nation	String	swiss, non-swiss	Nationality (binary)	STATPOP	nation	Grouped
child_in_household	Boolean	true, false	Whether in the household someone is 18 or older and another person is under 18	STATPOP	household_id	Derived from household structure
household_size	String	1, 2, 3-4, 5+	Number of persons in household (households with 10 or more were first split).	STATPOP	household_id	Derived from household structure
zone_id	Integer	-	Unique identifier of the NPVM-Zone	STATPOP	location_id	Randomized
xcoord	Decimal	-	X Coordinate	STATPOP	X	Randomized
ycoord	Decimal	-	Y Coordinate	STATPOP	Y	Randomized
level_of_employment	String	0, 1-39, 40-79, 80-100	Level of employment, i.e., employment rate (in percentage groups)	synthetic	type 2	Grouped
position_in_bus	String	null, ceo, employee, apprentice, bus_management, management	Current business position status. Differences are made between CEO (in German: Geschäftsführer:in), business management (Geschäftsleitung), management (qualifizierte Mitarbeiter:in), employee (einfache Mitarbeiter:in) and apprentice (Lehrling). The difference between <i>einfache</i> (employee) and <i>qualifizierte</i> (management) employees is based on a definition derived from the Mobility and Transport Microcensus 2015.	synthetic	position_in_bus	Simple recoding
position_in_edu	String	pupil, apprentice, student, null	Current education status. "pupil" corresponds to "Schüler:in" in German. It includes primary school pupils (Primärschüler:in) and secondary	synthetic	position_in_edu	Simple recoding

			<p>school pupils (Sekundarschüler:in) Their age is between 6 and 24 years old.</p> <p>“student” corresponds to “Studenten” in German. Their age is between 18 and 64 years old.</p> <p>“null” corresponds to a person who is currently not undertaking formal education.</p> <p>Data source: Federal Statistical Office (Bildungsstatistiken SDL (Lernende), SBG (Berufliche Grundbildung) und SHIS (Schweizerisches Hochschulinformationssystem), BFS).</p>			
education	String	primary, secondary, tertiary	<p>Highest education level achieved by person. Data source: Federal Statistical Office (Strukturerhebung 2021).</p> <p>“primary” (without post-compulsory education) corresponds to “keine Ausbildung” in German and Code 1 in Strukturerhebung, and includes those who did not yet finish primary education.</p> <p>“secondary” (Secondary School II) corresponds to “Sekundarstufe” in German and Codes 2 and 3 in Strukturerhebung.</p> <p>“tertiary” corresponds to “Tertiärstufe” in German and Codes 4 and 5 in Strukturerhebung.</p>	synthetic	education	Simple recoding
language	String	german, french, italian, romansh, other	Language (4 classes)	synthetic	language	Simple recoding
household_income	String	0-50k, 50-95k, 95-145k, 145+	Available income for the household (in CHF per year)	synthetic	income	Derived from "Income (Person)" and grouped
driving_licence	Boolean	true, false	Whether the person owns a driving licence	synthetic	driving_licence	Simple recoding
cars_in_hh	String	0, 1, 2, 3+	Number of cars in the household	synthetic	hcar_ownership	Derived from household

cars_drivetype	String	-	Engine type and size of cars, separated by "_". Types: 'Benzin', 'Diesel', 'Elektrisch', 'Hybrid', 'HybridStecker'. Sizes: 'Klein' (up to 80kW), 'Mittel' (80-150kW), 'Gross' (over 150kW).	synthetic	hcar_drivetype_1,2,3	Derived from household
car_available	Boolean	true, false	Whether the person has access to a car. Always false if driving_licence is false or cars_in_hh > 0, otherwise probabilistic.	synthetic	car_available	Simple recoding
public_transport	String	null, HTA, GA, VA, HTA+VA	Public transport subscription. "GA" corresponds to "GA travelcard" (in German: Generalabonnement). "HTA" corresponds to "Half Fare Travelcard" (in German: Halbtax). "VA" corresponds to "regional travelcards" (in German: Verbund-Abos).	synthetic	public_transport	Simple recoding
bike_ownership	String	null, bike, eBike25, bike+eBike25, eBike45, bike+eBike45	Type of bicycle(s) owned by this person. eBike25 and eBike45 corresponds to electric bicycles licenced to up to 25km/h and 45km/h maximum speed, respectively.	synthetic	bike_ownership	Simple recoding