
title: "LLM Manual: Tolkning av Segregationsbarometerns Data"

author: "Boverket - Segregationsbarometern"

date: "2024"

Manual för LLM-tolkning av Segregationsbarometerns Socioekonomiska Data

Innehållsförteckning

1. [Introduktion](#1-introduktion)
2. [Datastruktur](#2-datastruktur)
3. [Tolkningsramverk](#3-tolkningsramverk)
4. [Indikatorsystem](#4-indikatorsystem)
5. [Kvalitetssäkring](#5-kvalitetssäkring)
6. [Appendix](#6-appendix)

1. Introduktion

1.1 Syfte

Denna manual är utformad för att vägleda språkmodeller (LLM) i tolkningen av socioekonomiska data från Segregationsbarometern.

1.2 Omfattning

Manualen täcker samtliga socioekonomiska indikatorer i Segregationsbarometern på RegSO-nivå.

2. Datastruktur

2.1 Kodningssystem

[indicator_id][indicator_abbreviation][background_variable_id]

Exempel:

- t10boa019 = Bostadsarea, åldersgrupp 0-19 år
- t24sysstot = Total förvärvsfrekvens

2.2 Geografisk Indelning

- **RegSO** (regionalt statistikområde)
- Definierat av SCB
- Hierarkisk områdesindelning

3. Tolkningsramverk

3.1 Indikator typer

Procentbaserade indikatorer

- Valdeltagande (t26, t27, t28)
- Sysselsättning (t24)
- Utbildningsresultat (t16, t19)

Absoluta värden

- Disponibel inkomst (t1)
- Bostadsarea (t10)
- Meritvärden (t17)
- Ohälsodagar (t30)

Kvotmått

- Flyttningskvot (t12)

3.2 Bakgrundsvariabler

Kod	Betydelse
----	-----
019	0-19 år
2025	20-25 år
2644	26-44 år
4564	45-64 år
65	65+ år
kv	Kvinnor
ma	Män
in	Inrikes född
ut	Utrikes född
utb1-4	Utbildningsnivåer

4. Indikatorsystem

4.1 Ekonomiska förhållanden

Kod	Indikator	Mätmetod
----	-----	-----
t1	Disponibel inkomst	Median per konsumtionsenhet
t2	Ekonomisk standard	Låg/hög, percentilbaserad
t4	Försörjningsstöd	Andel mottagare
t6	Barnfattigdom	Låg inkomststandard

4.2 Boende och mobilitet

Kod	Indikator	Mätmetod
----	-----	-----
t9	Upplåtelseform	Fördelning %
t10	Bostadsarea	Kvadratmeter/person
t12	Flyttmönster	In-/utflyttningsskvot

4.3 Utbildning och arbete

| Kod | Indikator | Mätmetod |

|-----|-----|-----|

| t15-t19 | Utbildningsresultat | Olika utfallsmått |

| t21 | UVAS | Andel % |

| t23 | Arbetslöshet | Längd >6 mån |

| t24 | Sysselsättning | Förvärvsfrekvens |

5. Kvalitetssäkring

5.1 Valideringsregler

1. Procentvärden: 0-100

2. Kvotmått: Kan överstiga 1.0

3. Totalvärden: Måste vara konsistenta

5.2 Felhantering

- Hantering av saknade värden (us)

- Identifiering av extremvärden

- Logiska kontroller mellan delgrupper

6. Appendix

6.1 Fullständig indikatorlista

...

t1: Disponibel inkomst

t2: Ekonomisk standard

t4: Försörjningsstöd

t6: Barnfattigdom

t8: Utbildningsnivå

t9: Upplåtelseform

t10: Bostadsarea

t12: Flyttkvot

t15: Förskola

t16: Gymnasiebehörighet

t17: Meritvärden

t19: Gymnasieexamen

t21: UVAS

t23: Arbetslöshet

t24: Sysselsättning

t26: Riksdagsval

t27: Landstingsval

t28: Kommunval

t30: Ohälsodagar

...

6.2 Standardnotationer

- tot = totalvärde
- us = uppgift saknas
- Åldersgrupper anges i intervall
- Utbildningsnivåer numreras 1-4

[Previous sections remain the same until section 6...]

6. Few-Shot Learning Examples

6.1 Grundläggande Tolkning

****Input Data:****

region,code,year,regsokod,regso,t2lekstot,t2hekstot
regso,2983,2022,2983,RegSO-Annelund-Lönngården,30.55,1.25

****Exempel 1 - Basal Tolkning:****

Q: Vad visar värdena t2lekstot och t2hekstot för RegSO-Annelund-Lönngården?

A: I området RegSO-Annelund-Lönngården har:

- 30.55% av befolkningen låg ekonomisk standard (t2lekstot)
- 1.25% av befolkningen hög ekonomisk standard (t2hekstot)

Detta indikerar en betydande ekonomisk utsatthet i området.

****Exempel 2 - Sammansatt Tolkning:****

Q: Analysera den ekonomiska polariseringen i området.

A: Området uppvisar:

1. Låg ekonomisk polarisering då:
 - Andelen med låg ekonomisk standard (30.55%) är markant högre än
 - Andelen med hög ekonomisk standard (1.25%)
2. Detta ger en polariseringskvot på 0.041 (1.25/30.55)
3. Området karakteriseras av ekonomisk homogenitet i den lägre delen av inkomstskalan

6.2 Demografisk Analys

****Input Data:****

region,t2leks019,t2leks2025,t2leks2644,t2leks4564,t2leks65
RegSO-Annelund-Lönngården,43.89,29.11,24.74,30.58,40.54

^^^

****Exempel 3 - Åldersfördelad Analys:****

^^^

Q: Vilka åldersgrupper är mest ekonomiskt utsatta i området?

A: Analysen visar att:

1. Högst andel låg ekonomisk standard finns bland:
 - Barn/unga 0-19 år: 43.89%
 - Äldre 65+ år: 40.54%
2. Lägst andel finns i åldersgruppen:
 - 26-44 år: 24.74%
3. Detta indikerar särskild ekonomisk utsatthet i områdets ytterkanter av åldersspannet

^^^

6.3 Intersektionell Analys

****Input Data:****

^^^

region,t2leksma,t2lekskv,t2leksin,t2leksut
RegSO-Annelund-Lönngården,29.95,31.22,27.23,35.62

^^^

****Exempel 4 - Gruppjämförelse:****

^^^

Q: Analysera skillnader i ekonomisk utsatthet mellan olika demografiska grupper.

A: Data visar följande mönster:

1. Könsskillnader:
 - Kvinnor: 31.22% låg ekonomisk standard
 - Män: 29.95% låg ekonomisk standard
 - Könsgap: 1.27 procentenheter
2. Födelsebakgrund:
 - Inrikes födda: 27.23% låg ekonomisk standard
 - Utrikes födda: 35.62% låg ekonomisk standard
 - Gap: 8.39 procentenheter
3. Huvudsaklig observation: Födelsebakgrund uppvisar större skillnader än kön

^^^

6.4 Utbildningsnivå och Ekonomi

****Input Data:****

^^^

region,t2leksutb1,t2leksutb2,t2leksutb3,t2leksutb4
RegSO-Annelund-Lönngården,50.00,22.83,29.17,18.20

^^^

****Exempel 5 - Utbildningsanalys:****

Q: Hur relaterar utbildningsnivå till ekonomisk utsatthet?

A: Analysen visar ett tydligt mönster:

1. Högst andel låg ekonomisk standard:
 - Förgymnasial utbildning (utb1): 50.00%
2. Lägst andel:
 - Längre eftergymnasial utbildning (utb4): 18.20%
3. Observationer:
 - Linjärt samband bryts för eftergymnasialt övrigt (utb3)
 - Skillnad mellan högst/lägst: 31.80 procentenheter

6.5 Komplex Mönsteridentifiering

Exempel 6 - Holistisk Analys:

Q: Identifiera huvudsakliga socioekonomiska mönster i området.

A: Baserat på alla tillgängliga indikatorer:

1. Särskilt utsatta grupper:
 - Barn/unga (43.89% låg ekonomisk standard)
 - Personer med förgymnasial utbildning (50.00%)
 - Utrikes födda (35.62%)
2. Skyddande faktorer:
 - Högre utbildning (18.20% för eftergymnasialt ≥ 2 år)
 - Ålder 26-44 år (24.74%)
3. Strukturella observationer:
 - Tydlig utbildningsgradient
 - Åldersbetingad sårbarhet
 - Betydande integration/segregationsmönster

7. Appendix

[Previous appendix content remains the same...]

Slut på manual