

An empirical urban–rural typology of Swedish municipalities

Correspondence analysis and hierarchical clustering on the 2022 sampling

2026-05-22

Table of contents

1	What the analysis does	1
2	Inputs and preparation	1
2.1	Two adjustments to the previous CA	2
2.2	A note on the demography block	2
3	The space of municipalities	3
4	A first attempt at an empirical typology: four clusters	4
4.1	Cluster profiles	4
4.1.1	Cluster 1: Industrial-rural towns	5
4.1.2	Cluster 2: Peri-rural commuter belt	5
4.1.3	Cluster 3: Regional service centres	5
4.1.4	Cluster 4: Metropolitan & knowledge hubs	5
4.2	What each cluster looks like	6
4.3	Geography of the clusters	8
5	Takeaways	8
6	An optional finer cut: six clusters	9
6.1	What each new sub-cluster looks like	10
6.1.1	Cluster 1: Northern & remote	10
6.1.2	Cluster 2: Central industrial towns	11
6.1.3	Cluster 3: Peri-rural commuter belt	11
6.1.4	Cluster 4: Regional service centres	11
6.1.5	Cluster 5: Affluent suburbs & university satellites	11
6.1.6	Cluster 6: Inner Stockholm core	12
6.2	Reading the two splits	12

1 What the analysis does

Sweden’s 290 municipalities are routinely sorted into urban and rural by administrative or population-size rules. These categories are convenient but external to the data. The work reported here builds a typology *from* the data: we look at how municipalities differ on a wide set of structural variables (education, employment, housing, mobility, demography) and let those differences group municipalities together.

There are two steps. First, a correspondence analysis (CA) places every municipality in a low-dimensional space where distance reflects how similar the municipalities are on the active variables. Second, a hierarchical clustering on that space (Ward’s method) cuts the cloud into a small number of groups. The (preliminary) result is a four-cluster typology that runs from industrial and rural municipalities at one end to wealthy metropolitan suburbs and knowledge hubs at the other.

2 Inputs and preparation

The data come from Statistics Sweden via the project’s 2022 sampling. Six variable blocks are used as **active**:

- **Education:** Four levels of attainment (primary/lower secondary through post-graduate)
- **Employment:** 16 activity sectors (agriculture through arts & recreation)
- **Housing:** Rented, tenant-owned, and owner-occupied dwellings
- **Workplace:** Commuters in, commuters out, working & living in the same municipality
- **Migration:** Inmigrations and outmigrations
- **Demography:** Retirees and number of localities

Two further blocks are projected as **supplementary**:

- **Educational provision:** Counts of preschool, primary, secondary, adult and higher-education units by ownership (total, public, private)
- **Opinion:** Survey-based satisfaction on preschool, elementary and high school

2.1 Two adjustments to the previous CA

A plain CA on the raw counts puts Stockholm, Göteborg and Malmö far out on the first axis simply because they are large. The axis becomes a size axis and tells us little about *how* municipalities differ. Two adjustments fix this:

1. **Block normalisation.** Within each variable block, every municipality is rescaled so that its block-total is the same constant. After this step every municipality contributes the same row mass to the CA, and every block contributes the same total weight. The axes are no longer dominated by absolute counts.
2. **Provision as a per-capita rate.** The supplementary provision counts are divided by a population proxy (the row sum of the education block) so that “five universities in Stockholm” and “one university in Skellefteå” are compared on the rate scale rather than the raw count scale. Without this step rare types like *university* or *Sami school* are tugged to extreme positions and clutter the biplot (they still are, but not as much).

Table 1: Eigenvalues of the first five CA dimensions. Dimensions 1 and 2 together account for two-thirds of the variability.

Dimension	Eigenvalue	% inertia	Cumulative %
Dim 1	0.0368	43.7	43.7
Dim 2	0.0195	23.1	66.9
Dim 3	0.0077	9.1	76.0
Dim 4	0.0037	4.4	80.4
Dim 5	0.0035	4.1	84.5

2.2 A note on the demography block

Two of the variables we wanted to use, *Number of retirees* and *Number of localities*, do not belong to any of the natural topics described above. They are stand-alone counts that, on their own, would simply scale with municipality size.

The normalisation step we use needs at least **two variables** in a block. With only one variable, the rescaling “divide each row by its own block total” turns every value into the same constant, and the variable carries no information into the CA. So a one-variable block disappears from the analysis.

We had three options for retirees and localities:

- **Drop them.** They would not influence the result at all. We would lose information that is meaningful to the urban–rural question (an older population, a denser locality structure).
- **Attach each to an existing topic.** There is no natural home: retirees are not “employment” and localities are not “housing”.
- **Group the two together into their own block** (“Demography”). Each municipality’s retirees and localities are rescaled so that the two of them add up to the same constant, just like every other block.

I went with the third option for now. The price is that the resulting share “retirees share of (retirees + localities)” mixes units (people and places), so it is not a tidy proportion in the way “share of housing that is rented” is. The benefit is that the *contrast* it captures is meaningful: a municipality with many retirees per locality (an ageing population concentrated in a few settlements) sits on one side, and one with many localities per resident retiree (a population spread thinly across many small settlements) on the other. That contrast lines up with the urban–rural gradient we care about, and so it’s kept (for now).

3 The space of municipalities

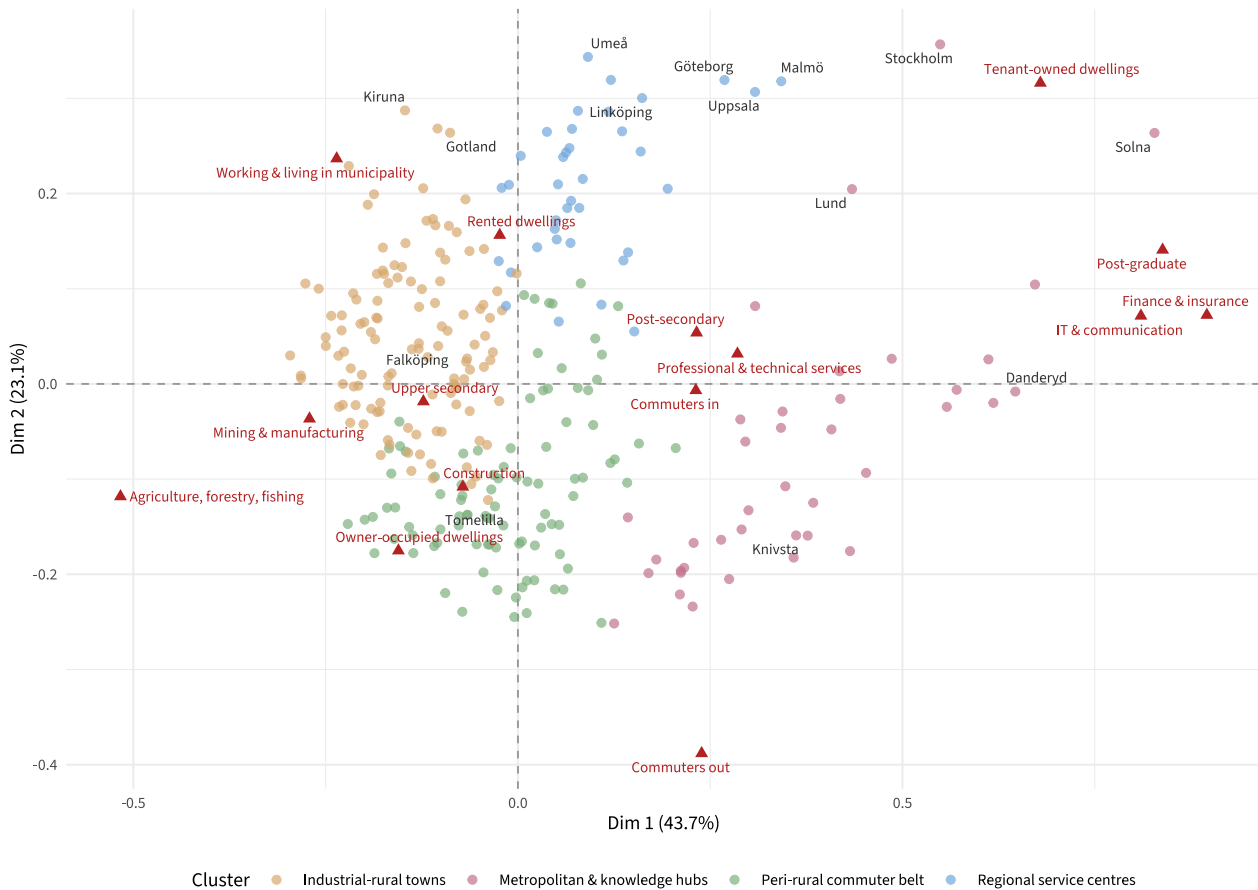


Figure 1: CA biplot, first two dimensions. Municipalities (dots) are coloured by their cluster (see next section). The most contributive active variables are shown in red.

Dim 1 (43.7%) is a rural-to-urban axis. To the left sit modalities tied to extractive and traditional employment (agriculture, mining & manufacturing), owner-occupied housing, and primary or upper-secondary education; the municipalities there are northern, industrial or sparsely populated (e.g. Kiruna, Gotland, Falköping). To the right sit modalities tied to the knowledge economy (IT & communication, finance, professional services), post-secondary and post-graduate attainment, and tenant-owned (apartment) housing; the right pole is anchored by Solna, Stockholm, Sundbyberg, Lund and Uppsala.

Dim 2 (23.1%) distinguishes municipalities by how their residents relate to the labour market. The top of the axis is anchored by “Working & living in the municipality”; self-contained labour markets, often regional centres like Göteborg, Malmö, Umeå and Uppsala. The bottom is dominated by “Commuters out”; residential municipalities whose population works elsewhere, typically the Stockholm-region suburbs (Knivsta, Salem, Staffanstorp).

Table 2: Cluster sizes and labels.

Cluster	Label	Municipalities	Share (%)
1	Industrial-rural towns	119	41.0
2	Peri-rural commuter belt	95	32.8
3	Regional service centres	37	12.8
4	Metropolitan & knowledge hubs	39	13.4

4 A first attempt at an empirical typology: four clusters

Hierarchical clustering (Ward) on the CA row coordinates yields four clusters; the gap between four and three clusters in the inertia profile is large, while five clusters add only a small refinement.



Figure 2: Ward dendrogram on the CA coordinates. The cut into four clusters falls where the merge cost rises sharply.

4.1 Cluster profiles

For every cluster we list the variables whose mean inside the cluster departs most strongly from the national mean (v . t test statistic from HCPC). A high positive value means the variable is *over*-represented in the cluster.

4.1.1 Cluster 1: Industrial-rural towns

Typical municipalities (closest to the cluster centre): *Säffle, Falköping, Åmål, Lindesberg, Mora.*

Table 3: Top over-represented variables, cluster 1.

Variable	Share in cluster (%)	Share overall (%)	v.test
Working & living in municipality	8.66	6.94	Inf
Mining & manufacturing	2.52	2.06	Inf
Agriculture, forestry, fishing	0.69	0.50	31.1
Upper secondary	7.51	6.93	27.7
Public Komvux	0.37	0.26	24.5
Komvux	0.37	0.26	24.5

4.1.2 Cluster 2: Peri-rural commuter belt

Typical municipalities (closest to the cluster centre): *Tomelilla, Osby, Sölvesborg, Tibro, Klippan.*

Table 4: Top over-represented variables, cluster 2.

Variable	Share in cluster (%)	Share overall (%)	v.test
Commuters out	5.40	4.19	Inf
Owner-occupied dwellings	8.38	7.44	35.9
Bad opinion on highschool	0.85	0.62	29.0
Good opinion on preschool	2.65	2.31	22.5
Good opinion on elementary school	1.64	1.46	15.0
Bad opinion on elementary school	0.54	0.45	14.2

4.1.3 Cluster 3: Regional service centres

Typical municipalities (closest to the cluster centre): *Kalmar, Borås, Nyköping, Växjö, Gävle.*

Table 5: Top over-represented variables, cluster 3.

Variable	Share in cluster (%)	Share overall (%)	v.test
Mid opinion on highschool	2.30	1.37	Inf
Tenant-owned dwellings	3.22	1.86	Inf
Working & living in municipality	8.65	6.94	36.0
Mid opinion on preschool	2.12	1.35	34.2
Mid opinion on elementary school	2.33	1.54	33.3
Rented dwellings	5.86	4.58	32.6

4.1.4 Cluster 4: Metropolitan & knowledge hubs

Typical municipalities (closest to the cluster centre): *Huddinge, Partille, Tyresö, Upplands-Bro, Knivsta.*

Table 6: Top over-represented variables, cluster 4.

Variable	Share in cluster (%)	Share overall (%)	v.test
Commuters out	7.00	4.19	Inf
Tenant-owned dwellings	4.06	1.86	Inf
IT & communication	0.93	0.34	Inf
Post-secondary	6.27	4.32	Inf
Finance & insurance	0.47	0.15	38.2
Commuters in	3.87	2.75	36.8

4.2 What each cluster looks like

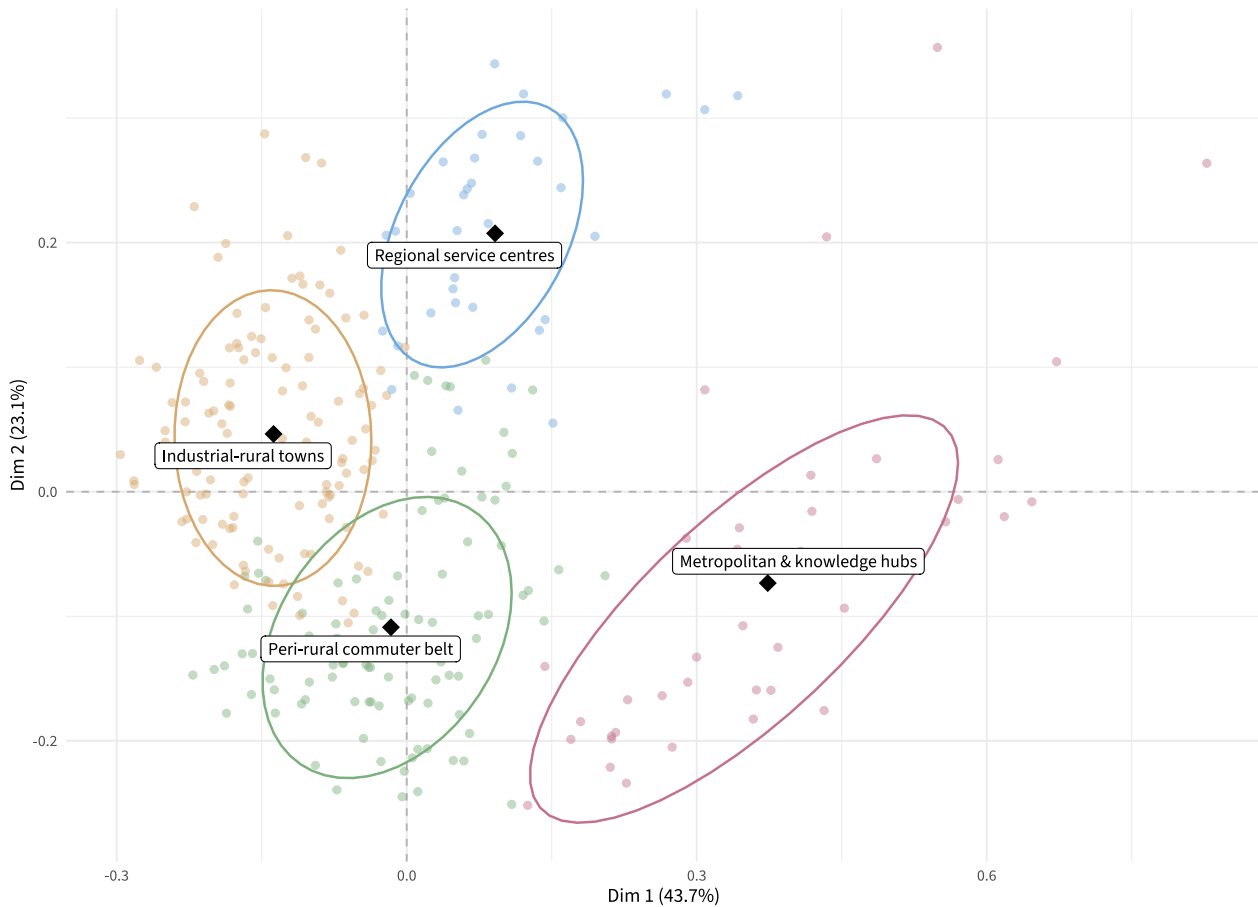


Figure 3: Cluster centroids on the CA plane. The arrows give an at-a-glance view of where each cluster sits in the rural-urban (Dim 1) and self-contained-vs-commuter (Dim 2) plane.

Cluster 1: Industrial-rural towns.

The largest cluster (119 municipalities). These are self-contained, often Norrland or central-Sweden municipalities. Residents tend to work and live in the same place, employment is concentrated in mining, manufacturing and agriculture, housing is overwhelmingly owner-occupied, and the typical educational ceiling is upper secondary. Kiruna and Gotland sit at the cluster's far edge.

Cluster 2: Peri-rural commuter belt.

95 municipalities. Smaller southern and central rural municipalities (Tomelilla, Osby, Klippan, Sölvesborg) where many residents commute out to a nearby labour market. Owner-occupied housing dominates, construction and agriculture are visible employment sectors, and survey opinions of local high schools tend to be *worse* than the national average. Consistent with a thin local educational offer that pushes choices elsewhere.

Cluster 3: Regional service centres.

37 municipalities, including Göteborg, Malmö, Umeå, Uppsala, Linköping, Kalmar, Borås, Växjö and Gävle. They share the markers of a mid-sized regional city: substantial rented and tenant-owned housing, residents who work where they live, public administration, post-secondary attainment, and average satisfaction scores on local services. These are Sweden's "self-contained" mid-tier cities.

Cluster 4: Metropolitan and knowledge hubs.

39 municipalities, mostly Stockholm-county suburbs (Huddinge, Tyresö, Knivsta, Sundbyberg, Solna, Danderyd, Lidingö, Täby, Nacka) plus Stockholm and Lund. Outbound commuting, finance, IT, professional services, post-secondary and post-graduate attainment all surge above the national mean; tenant-owned (apartment) housing is the norm. This cluster captures both the inner-city core of the Stockholm region and the wealthy residential suburbs that depend on it.

4.3 Geography of the clusters

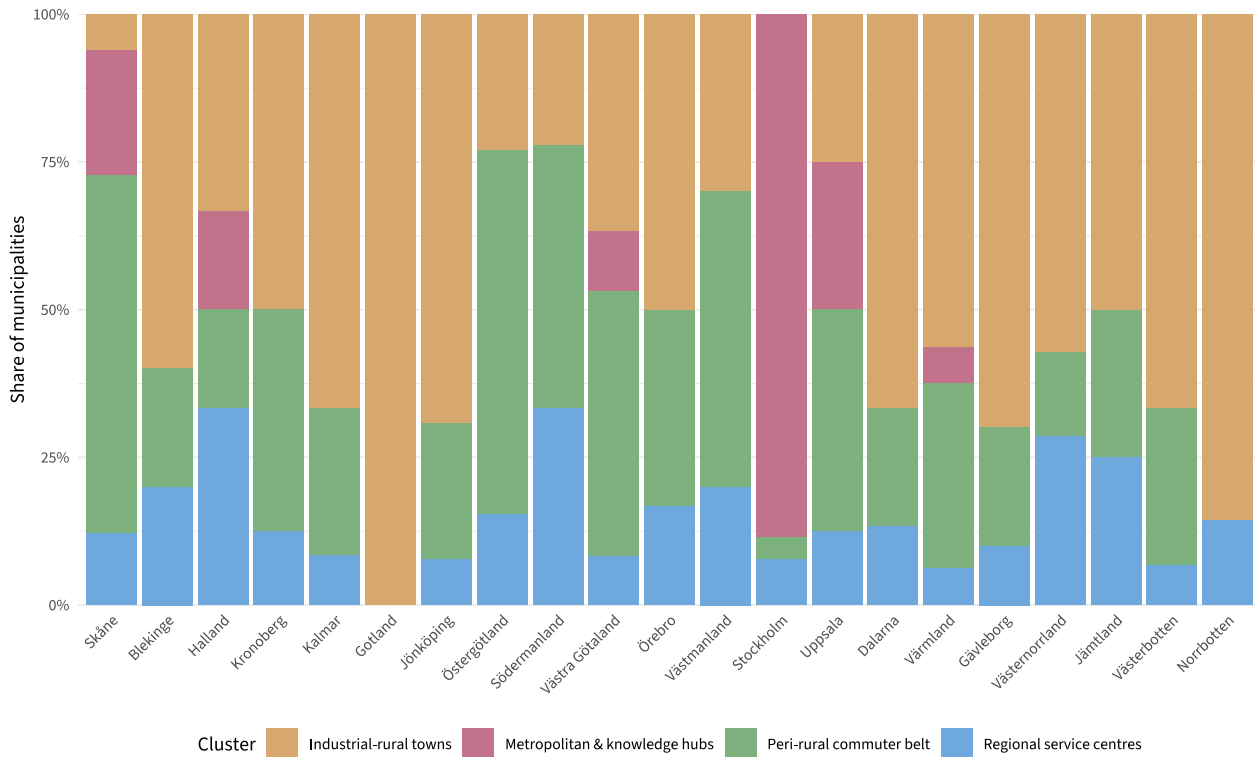


Figure 4: Cluster composition by county. Counties are ordered roughly south to north. Cluster 1 (industrial-rural) and 2 (peri-rural) dominate outside the metropolitan areas; cluster 4 (metro & knowledge hubs) is overwhelmingly Stockholm-, Uppsala- and Skåne-based; cluster 3 (regional centres) is spread thinly across most counties.

5 Takeaways

The four clusters lay out a coherent rural-to-urban gradient that the standard “urban / rural” dichotomy hides. The biggest empirical break is not between metro and non-metro but between two kinds of rural area: industrial-rural towns with their own labour market (Cluster 1) versus peri-rural municipalities that depend on outbound commuting (Cluster 2). At the urban end, the data also separate two kinds of city: regional service centres that work as self-contained labour markets (Cluster 3) and metropolitan-region municipalities tied to the knowledge economy (Cluster 4).

A second observation is that the second axis—self-contained labour market vs. outbound-commuter base—matters as much for cluster membership as the first. Sweden’s “wealthy suburb” municipalities and its mid-sized university cities sit on the same (right) side of the rural-urban axis but on opposite sides of the labour-market axis, and the clustering picks that distinction up cleanly.

The typology is a starting point: a four-cluster cut over-aggregates Stockholm, Solna and Sundbyberg with the surrounding affluent suburbs, and it pools the remote Norrland municipalities with the central industrial towns. Cutting the same tree at six clusters splits both of those merges; the next section shows that finer view.

Table 7: Six-cluster cut: cluster sizes and labels.

Cluster	Label	Municipalities	Share (%)
1	Northern & remote	42	14.5
2	Central industrial towns	77	26.6
3	Peri-rural commuter belt	95	32.8
4	Regional service centres	37	12.8
5	Affluent suburbs & university satellites	31	10.7
6	Inner Stockholm core	8	2.8

6 An optional finer cut: six clusters

Cutting the Ward tree at six clusters keeps the two middle clusters from the four-cluster cut (the peri-rural commuter belt and the regional service centres) and splits the other two:

- The **industrial-rural towns** divide into **northern/remote** municipalities (Kiruna, Gotland, Piteå, Skellefteå) and **central/southern industrial** towns (Falköping, Lindesberg, Hedemora).
- The **metropolitan & knowledge hubs** divide into the **affluent suburbs and university satellites** (Lund, Mölndal, Partille, Huddinge, Knivsta, Tyresö, Kungsbacka, ...) and a small **inner Stockholm core** (Stockholm, Solna, Sundbyberg, Danderyd, Lidingö, Täby, Nacka, Sollentuna).

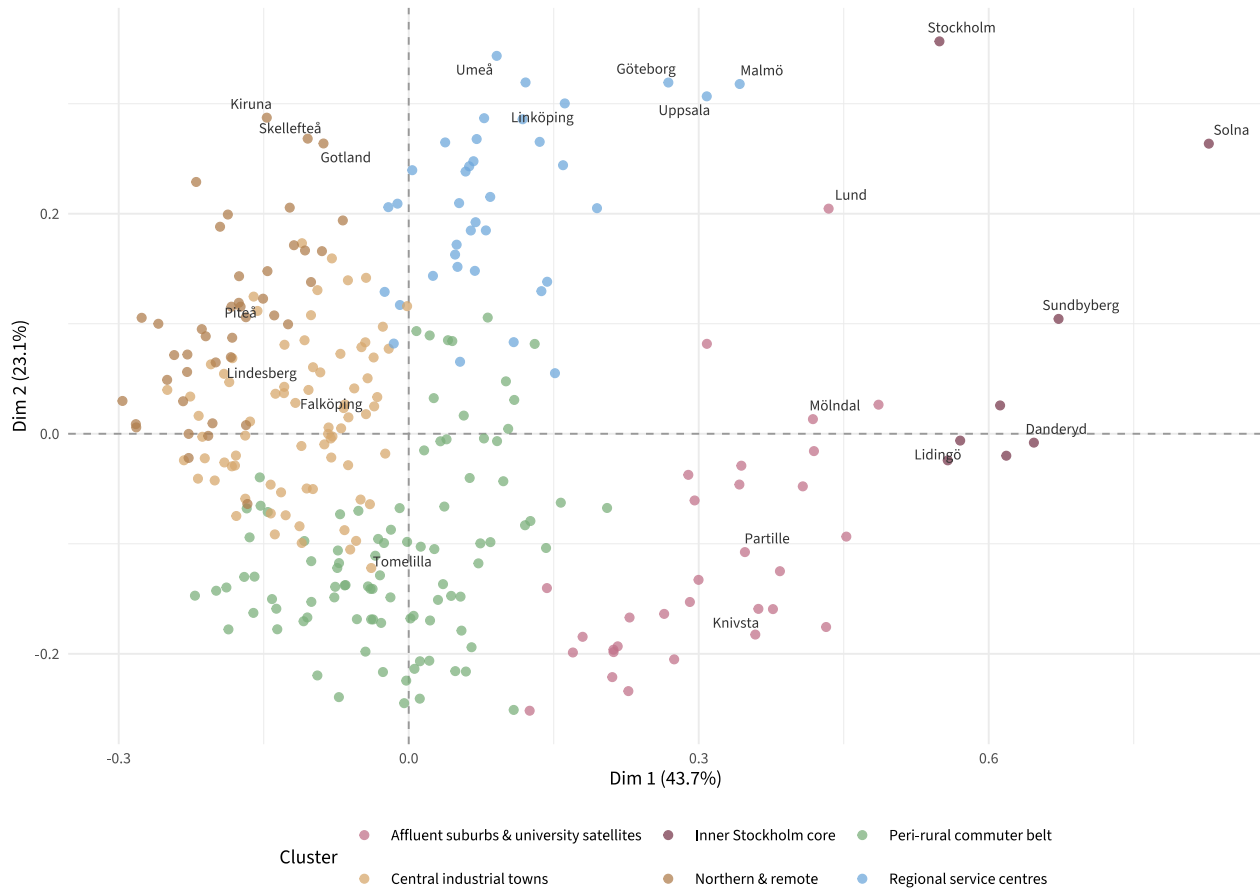


Figure 5: CA biplot, six-cluster cut. The two splits relative to the four-cluster view cuts the cloud out along Dim 1: the northern/remote group sits further left than the central industrial towns, and the inner Stockholm core sits further right than the suburban ring.

6.1 What each new sub-cluster looks like

6.1.1 Cluster 1: Northern & remote

Typical municipalities: *Piteå, Ljusdal, Ånge, Kalix, Storuman.*

Table 8: Top over-represented variables, six-cluster cut, cluster 1.

Variable	Share in cluster (%)	Share overall (%)	v.test
Working & living in municipality	10.06	6.94	Inf
Agriculture, forestry, fishing	0.81	0.50	24.0
Public Komvux	0.48	0.26	23.2
Komvux	0.48	0.26	23.2
Public preschool class	1.07	0.77	20.3
Preschool class	1.19	0.87	19.9

6.1.2 Cluster 2: Central industrial towns

Typical municipalities: *Lindesberg, Falköping, Hedemora, Åmål, Nybro.*

Table 9: Top over-represented variables, six-cluster cut, cluster 2.

Variable	Share in cluster (%)	Share overall (%)	v.test
Mining & manufacturing	2.97	2.06	Inf
Working & living in municipality	7.87	6.94	31.1
Primary & lower secondary	2.85	2.54	16.9
Owner-occupied dwellings	7.92	7.44	15.8
Upper secondary	7.39	6.93	15.7
Agriculture, forestry, fishing	0.62	0.50	14.1

6.1.3 Cluster 3: Peri-rural commuter belt

Typical municipalities: *Tomelilla, Osby, Sölvesborg, Tibro, Klippan.*

Table 10: Top over-represented variables, six-cluster cut, cluster 3.

Variable	Share in cluster (%)	Share overall (%)	v.test
Commuters out	5.40	4.19	Inf
Owner-occupied dwellings	8.38	7.44	35.9
Bad opinion on highschool	0.85	0.62	29.0
Good opinion on preschool	2.65	2.31	22.5
Good opinion on elementary school	1.64	1.46	15.0
Bad opinion on elementary school	0.54	0.45	14.2

6.1.4 Cluster 4: Regional service centres

Typical municipalities: *Kalmar, Borås, Nyköping, Växjö, Gävle.*

Table 11: Top over-represented variables, six-cluster cut, cluster 4.

Variable	Share in cluster (%)	Share overall (%)	v.test
Mid opinion on highschool	2.30	1.37	Inf
Tenant-owned dwellings	3.22	1.86	Inf
Working & living in municipality	8.65	6.94	36.0
Mid opinion on preschool	2.12	1.35	34.2
Mid opinion on elementary school	2.33	1.54	33.3
Rented dwellings	5.86	4.58	32.6

6.1.5 Cluster 5: Affluent suburbs & university satellites

Typical municipalities: *Hammarö, Partille, Staffanstorps, Kävlinge, Knivsta.*

Table 12: Top over-represented variables, six-cluster cut, cluster 5.

Variable	Share in cluster (%)	Share overall (%)	v.test
Commuters out	7.23	4.19	Inf
Tenant-owned dwellings	3.41	1.86	Inf
Post-secondary	5.82	4.32	35.3
IT & communication	0.77	0.34	32.4
Professional & technical services	1.96	1.33	26.2
Commuters in	3.46	2.75	21.0

6.1.6 Cluster 6: Inner Stockholm core

Typical municipalities: *Nacka, Sollentuna, Täby, Lidingö, Sundbyberg.*

Table 13: Top over-represented variables, six-cluster cut, cluster 6.

Variable	Share in cluster (%)	Share overall (%)	v.test
Public special school	0.00	0.00	Inf
Special school	0.00	0.00	Inf
Tenant-owned dwellings	6.62	1.86	Inf
Post-secondary	8.02	4.32	Inf
IT & communication	1.58	0.34	37.6
Commuters in	5.46	2.75	35.3

6.2 Reading the two splits

North vs. central industrial.

Both subgroups inherit the rural-industrial profile of the four-cluster cut (owner-occupied housing, upper-secondary education, working and living in the same municipality). The northern group is distinguished by an even stronger weight of agriculture/forestry/fishing, adult education (komvux) and preschool-class infrastructure. These are sparsely populated municipalities that have to provide a full educational service on their own at every age. The central group keeps mining & manufacturing as its single strongest difference. These are old industrial towns rather than remote rural ones.

Suburban ring vs. inner Stockholm core.

The suburban/university group (31 municipalities) is dominated by outbound commuters, post-secondary attainment and tenant-owned housing. These are residential satellites whose labour markets sit elsewhere. The inner core (8 municipalities) tightens the same profile much further: apartment housing, IT and finance employment, and inbound commuting all jump to extreme levels (several `v.test` = `Inf`). The split is meaningful because Stockholm, Solna, Sundbyberg and Lidingö are *destinations* in the commuting network, while Knivsta, Partille and Staffanstorp are *origins*. The four-cluster cut treats them as one population because they are similar relative to the rest of Sweden, but they sit at the two ends of the metropolitan labour market.