

Statistics based on roads and addresses

- How do we create household clusters?

For each municipality, we divide all households into clusters, which as a rule must consist of **a minimum of 400 households** (read below for how to ensure a minimum number of households in the clusters).

Before we settle on the final division into clusters, we divide households into smaller groups (by sections of road sides) depending on the type of roads they belong to. This is done in the following way:

Division of roads into 3 different types of roads

For each municipality, all roads are grouped as follows:

1. Small roads

Roads with less than 5 households. Small roads cannot be divided into road sides nor into sections of roadsides.

2. Non-divided roads

Roads with a minimum of 5 households, but where the 1-4 households are located on one side of the road. Then these roads cannot be divided into two roadsides, so each road is thus perceived as one section of a road.

3. Roads divided into roadsides or sections of roadsides

Roads with a minimum of 10 households. The roads can be divided into roadsides or sections of roads.

- a. Either a minimum of 5 households on each roadside
- b. Or 0 households on one side of the road and at least 10 households on the other side of the road.

The households are thus divided into 2 sections depending on the road side.

Clustering

Once the roads have been divided into types, we form the final clusters that form the basis of the statistics. The clusters are primarily formed on the basis of the statistical variable (e.g. personal income or educational level). Clusters are formed differently for households depending on whether they belong to type 1, 2 or 3.

For households with road type 1

The clustering of households in road type 1 takes place in one of the following two ways.

1. If there are at least 20 households with road type 1 in the municipality, the households are put together in an independent cluster (ClusterID = 0). Then we calculate the average value for this cluster on the respective

statistical variable (e.g. personal income). There must be a minimum of 20 households in the cluster.

2. If there are less than 20 households in road type 1 in the municipality, the households are considered to belong to an independent group that will be included in the cluster formation with other sections of roads. This means that these households will be included in a cluster of households with road type 2 and / or road type 3.

For households in road types 2 and 3 and road type 1 (see point 2 above)

We calculate average values on the statistical variable for each section of a road. Then we sort the road sections based on the calculated average values.

Subsequently, the road sections are grouped according to the sorting into clusters, which consists of at least 400 households. Meaning that road sections are grouped into clusters with other road sections that are most similar to each other relative to the calculated average value.

Discretionary requirements

For the clusters formed on the basis of the average values of the road sections, the requirement is that they must contain at least 400 households. For ClusterID = 0, the requirement is that there should be a minimum of 20 households, which is the general rule for calculating statistics in smaller geographical areas.

It is a further requirement that in the distribution of households into categories of the statistical variable all households must not be placed in one group. If that is the case, the cluster will be expanded by several sections until there are at least 2 different categories of households.

For example: If statistics are calculated on income, no clusters should be found, where all households are in the range of DKK 300,000-400,000.

What does the actual output look like?

The customer receives a file that contains both statistical information at the cluster level and an overview of which households belong to which clusters. The statistical information includes an average value for each cluster as well as a distribution of households at different intervals on the basis of the statistical variable.

The table below shows statistics on each cluster

Municipality	Cluster-ID	Grp 1	Grp 2	Grp 3	Grp 4	Grp 5	Number of households	Average
XX	1	117	62	122	17	81	400	404,4
XX	2	95	63	86	56	100	400	399,2
YY	1	36	270	19	45	31	400	450,2
YY	2	91	106	86	72	45	400	390,1
ZZ	1	157	2	127	95	19	400	370,9
ZZ	2	117	39	82	104	58	400	363,8
...								
...								

The table below shows which clusters the individual sections of road and households belong to.

Municipality	Cluster-ID	Postal code	Name of the road	Side of the road	House no. from	House no. to
XX	1	100	AAA	1	1	9
XX	1	100	BBB	1	1	15
XX	2	200	CCC	2	2	16
XX	2	200	DDD	2	2	44
YY	1	300	EEE	1	1	17
YY	1	300	FFF	1	1	9
YY	2	400	GGG	0	1	4
YY	2	400	HHH	0	1	5
...						