# Socioeconomic position in BHLL data

This document gives an overview of some suggested variables to use when describing socioeconomic position (SEP) in the Better Health in Late Life (BHLL) data.

* These variables will be easy (for the statistician) to access. They will (hopefully) be used in multiple BHLL-studies and therefore support comparability across BHLL-studies.
* This document gives documentation on how they are calculated.

Nevertheless, you are of course also welcome to define other SEP-covariates, if these are more relevant for your study.

This overview has been done by Morten Madsen and Cathrine Hjorth

For a more thorough description of the SEP-variables, and the use within epidemiological studies, please consult SEPLINE by CF Hjorth et.al. – in press, [reference?]

The dataset with SEP-variables for BHLL (and for the background population in same age-groups as well) can be found at: ***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\BHLLprop\_data\BHLL\_SEP.sas7bdat***

The BHLL-questionnaires were answered in the fall 2021. The SEP-covariates are from December 31st - 2020, if nothing else is noted.

**Bold** text references names in SAS-datasets (or formats). Either in the original DST data or variables in our BHLL-SEP-dataset. Names in ***italic-bold*** marks names of tables at statistics Denmark.

In the appendix you can find the name of the categories, as well as the number of people within each category from the background population (all Danes aged 50-65 in the fall 2021).

## Summary:

I expect that in most studies you will only be interested in: **educationgr**, **Emp\_grup**, **Income\_quintile** and **Civst\_**. And maybe also: **Familie\_type\_**, **Urbanicity** and **Opr\_land\_DK**. The extra levels of the SEP-variables are included in the dataset in case you find a strong confounding effect of one of the SEP-variables and want to explore the risk of residual confounding by looking at more detailed data.

## Education:

* The highest attained educational level is stored in **educationgr** (4 levels) and **edu\_detailed** (12 levels) education level.
  + Based on the variable **HFAUDD** from the table ***UDDA\_\****. Based on the formats: **$AUDD2023\_HOVED\_L1L2\_KT** and **$AUDD2023\_HOVED\_L1L5\_KT.**
    - Corrections:
      * SEPLINE suggests categorizing For citizens born prior to 1958: 8., 9. And 10th grade counts as secondary, to account for the shift from a 7–year to a 9–year compulsory school system in 1972. At the time, this (realskolen) was more prestigious and provided wider job opportunities than completing only the compulsory years of schooling.
      * A variable edu\_imputed indicating which of the education registrations that are imputed (usually from immigrants)
  + **educationgr** can take the four values: {*low …, secondary… and higher education* and *missing}*.
* **edu\_detailed** (based on the format **$AUDD2023\_HOVED\_L1L5\_K.**), 12 different levels. (Including an additional splitting of primary school into {realskole, 8-10th grade, other primary school}
* The format **edu\_grp** can translate **edu\_detailed** into text.
* Also available is the variable **HFAUDD** itself.

## Labor market status/occupation:

* **Emp\_grup** [4 levels] (and **emp\_detailed**[13 levels]) Based on **PROC\_STATUS\_KODE** from dataset **idap\*** for year 2020 (or 2019 or 2018, in case of no registration in 2020).
  + The translation from **PSOC\_STATUS\_CODE** to **emp\_detailed** and further on to **Emp\_grup** including the numbers in each category can be found in the appendix

## Income /Wealth:

* **Income\_quintile** Income-quintiles … based om the variable **AEKVIVADISP\_13** from the table ***ind\****.
  + Average of the most recent three years (2018, 2019 and 2020) Equivalized disposable household income (in case of a missing registration – then average of the non-missing registrations).
    - Negative incomes are set to missing.
  + Quintiles among background population (within same sex and 5-yr age groups)
* Also available:
  + **Income\_fam**: the average yearly income used for determining quintiles.
  + **Wealth\_quintile** (and **famformrest\_ny05\_2020**): Household wealth (based on **famformrest\_ny05** from the ***faik\_2020*** dataset) in quintiles among background population (within same sex and 5-yr age groups)

## Cohabitation / Marital status /housing type:

Marital status and cohabitation (aggregated so marriage and same-sex-marriage are treated equally).

* **Civst\_**: Marital status at interview-date (identified based on the dataset ***civ2022***)
  + Same sex marriage has been classified together with marriage.
* **Familie\_type\_:** Family type (based on **familie\_type** from ***BEF***-dataset per 31dec2020)
* **Hustype\_**: Housing (based on **hustype** from ***BEF***-dataset per 31dec2020) (Be aware that many (13%) are classified as living in households with multiple families – maybe parents living with grown up kids?)

## Ethnicity

* Three variables:
  + **Opr\_land\_DK**: Denmark is land of origin (based on **opr\_land** in ***BEF***-dataset) (89% from DK).
  + **Statsb\_DK**: Danish citized (based on **statsb** in ***BEF***-dataset). (94% citizen in DK).

## Urbanicity:

* **Urbanicity** (5 levels) are based on municipality and the DST-format **$KOMGRP\_V4\_k.**
  + {hovedstadskommuner(25%), storbykommuner(11%), provinsbykommuner(24%), oplandskommuner(18%), landkommuner(23%)}
* Based on municipality of residence 31dec2020.

## SAS-dataset with SEP-data:

The SAS-dataset with SEP-data (for the full Danish population aged 50-65 in the autumn 2021) can be found at:

***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\BHLLprop\_data\SEP.sas7bdat***

***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\BHLLprop\_data\SEP\_17jun2025.sas7bdat***

The relevant formats can be created with: ***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\01a SEP Formats.sas***

The SAS-programs I have used for preparing the dataset:

***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\09a – prep SEP.sas***

***E:\ProjektDB\KEA\Workdata\708194\BHLL\Studybase\_prep\09b SEP\_covar.sas***

We expect that the overall categories within each component will be sufficient, but we have added the other categories, to also make it easy to dig deeper into SEP.

Future work: In the representativeness study Morten and Cathrine are working on exploring the impact of SEP and developing weights to make the BHLL-cohort representative of the total population.

Morten Madsen 17/6-2025

## Appendix: Hierarchies of education and employment, and count for marital status etc.:

### Education:

| **Table of edu\_detailed by educationgr** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **edu\_detailed** | **educationgr** | | | | |
| **Frequency** | **Low education (primary school)** | **Secondary education or vocationale training** | **Higher education (short/medium/long)** | **missing** | **Total** |
| **10: Grundskole** | 24832 | 0 | 0 | 0 | 24832 |
| **11: 8-10 klasse** | 183662 | 0 | 0 | 0 | 183662 |
| **12: Realskolekode** | 24413 | 0 | 0 | 0 | 24413 |
| **20: Gymnasiale uddannelser** | 0 | 60333 | 0 | 0 | 60333 |
| **30: Erhvervsfaglige uddannelser** | 0 | 473041 | 0 | 0 | 473041 |
| **35: Adgangsgivende uddannelsesfor** | 0 | 464 | 0 | 0 | 464 |
| **40: Korte videregående uddannelse** | 0 | 0 | 66683 | 0 | 66683 |
| **50: Mellemlange videregående udda** | 0 | 0 | 203302 | 0 | 203302 |
| **60: Bacheloruddannelser, BACH** | 0 | 0 | 12325 | 0 | 12325 |
| **70: Lange videregående uddannelse** | 0 | 0 | 101098 | 0 | 101098 |
| **80: Ph.d. og forskeruddannelser** | 0 | 0 | 10836 | 0 | 10836 |
| **90: Uoplyst** | 0 | 0 | 0 | 58063 | 58063 |
| **Total** | 232907 | 533838 | 394244 | 58063 | 1219052 |
|  |  |  |  |  |  |

### Occupation hierarchies:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Aggregating employment data into two new variables: emp\_grup (4 levels) and emp\_detailed (12 levels) based on PSOC\_STATUS\_kode*** | | | | | | |
| ***Counts/percentages are among the 1,200,000 Danes aged 50-65 autumn 2021*** | | | | | | |
| **emp\_grup** | **emp\_detailed** | **PSOC\_STATUS\_KODE** | **Frequency** | **Percent** |
| **.** | **.** |  | 353 | 0.03 |
| **Employed** | **Self-employed** | **110 Selvstændige (primær status ult. nov.)** | 51337 | 4.21 |
| **Employed** | **Self-employed** | **112 Arbejdsgivere (kun relevant for IDA-databasen)** | 18173 | 1.49 |
| **Employed** | **Self-employed** | **120 Medarbejdende ægtefæller** | 1593 | 0.13 |
| **Employed** | **Employee - high level + manager** | **131 Lønmodtagere med ledelsesarbejde** | 54727 | 4.49 |
| **Employed** | **Employee - high level + manager** | **132 Lønmodtagere på højeste niveau** | 221702 | 18.19 |
| **Employed** | **Employee - middle level** | **133 Lønmodtagere på mellemniveau** | 105359 | 8.64 |
| **Employed** | **Employee - basic level** | **134 Lønmodtagere på grundniveau** | 331533 | 27.20 |
| **Employed** | **Employee - basic level** | **135 Andre lønmodtagere** | 73470 | 6.03 |
| **Employed** | **Employee - basic level** | **136 Lønmodtagere u.n.a.** | 63955 | 5.25 |
| **Employed** | **Student** | **511 Personer under uddannelse (ordinær)** | 1237 | 0.10 |
| **Employed** | **Student** | **517 Modtagere af SU, incl. få '516 Udenlandske studeren** | 46 | 0.00 |
| **Unemployed** | **Temporarily unemployed** | **200 Arbejdsløse** | 35245 | 2.89 |
| **Unemployed** | **Temporarily unemployed** | **312 Feriedagpenge** | 23 | 0.00 |
| **Unemployed** | **Temporarily unemployed** | **316 Barselsfravær fra ledighed** | 15 | 0.00 |
| **Unemployed** | **Social security** | **311 Støttet beskæftigede uden løn** | 1629 | 0.13 |
| **Unemployed** | **Social security** | **313 Vejledning og opkvalificering** | 932 | 0.08 |
| **Unemployed** | **Social security** | **318 Kontanthjælp (passiv)** | 14607 | 1.20 |
| **Unemployed** | **Social security** | **319 Integrationsydelse** | 1513 | 0.12 |
| **Unemployed** | **Social security** | **320 Revalidering** | 47 | 0.00 |
| **Unemployed** | **Social security** | **321 Ressouceforløb** | 4775 | 0.39 |
| **Unemployed** | **Social security** | **322 Jobafklaringsforløb** | 4125 | 0.34 |
| **Unemployed** | **Social security** | **513 Øvrige uden for arbejdsstyrken** | 37379 | 3.07 |
| **Unemployed** | **Social security** | **514 Kursister** | 1222 | 0.10 |
| **Unemployed** | **Flex job** | **314 Ledighedsydelse** | 8129 | 0.67 |
| **Unemployed** | **Flex job** | **413 Fleksydelse** | 1878 | 0.15 |
| **Sick-leave / disability pension** | **Sick leave** | **317 Sygefravær fra ledighed** | 14055 | 1.15 |
| **Sick-leave / disability pension** | **Disability pension** | **411 Førtidspension** | 132644 | 10.88 |
| **Pension** | **Post-employment pension** | **412 Efterløn** | 26743 | 2.19 |
| **Pension** | **Age retirement pension** | **415 Anden pension** | 10536 | 0.86 |

| **emp\_grup** | **emp\_detailed** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| --- | --- | --- | --- | --- | --- |
| **.** | **.** | 353 | 0.03 | 353 | 0.03 |
| **Employed** | **Self-employed** | 71103 | 5.83 | 71456 | 5.86 |
| **Employed** | **Employee - high level + manager** | 276429 | 22.68 | 347885 | 28.54 |
| **Employed** | **Employee - middle level** | 105359 | 8.64 | 453244 | 37.18 |
| **Employed** | **Employee - basic level** | 468958 | 38.47 | 922202 | 75.65 |
| **Employed** | **Student** | 1283 | 0.11 | 923485 | 75.76 |
| **Unemployed** | **Temporarily unemployed** | 35283 | 2.89 | 958768 | 78.65 |
| **Unemployed** | **Social security** | 66229 | 5.43 | 1024997 | 84.09 |
| **Unemployed** | **Flex job** | 10007 | 0.82 | 1035004 | 84.91 |
| **Sick-leave / disability pension** | **Sick leave** | 14055 | 1.15 | 1049059 | 86.06 |
| **Sick-leave / disability pension** | **Disability pension** | 132644 | 10.88 | 1181703 | 96.94 |
| **Pension** | **Post-employment pension** | 26743 | 2.19 | 1208446 | 99.14 |
| **Pension** | **Age retirement pension** | 10536 | 0.86 | 1218982 | 100.00 |

### Marital status, urbanicity etc.

| **civst\_** | **Frequency** | **Percent** |
| --- | --- | --- |
| **Ugift** | 223633 | 18.34 |
| **Gift/partnerskab** | 722586 | 59.27 |
| **Skilt / ophævet partnerskab** | 241020 | 19.77 |
| **Enke/Enkemand / længstlevende p** | 31813 | 2.61 |

| **familie\_type\_** | **Frequency** | **Percent** |
| --- | --- | --- |
| **ægtepar m.m.** | 696208 | 57.11 |
| **Samlevende par** | 48543 | 3.98 |
| **samboende par** | 89907 | 7.38 |
| **Enlig** | 384394 | 31.53 |

| **hustype\_** | **Frequency** | **Percent** |
| --- | --- | --- |
| **Enlig mand** | 149503 | 12.26 |
| **Enlig kvinde** | 161313 | 13.23 |
| **Ægtepar** | 614640 | 50.42 |
| **Par i øvrigt** | 133388 | 10.94 |
| **Andre husstande bestående af flere fam** | 160208 | 13.14 |

| **urbanicity** | **Frequency** | **Percent** |
| --- | --- | --- |
| **Hovedstadskommuner** | 300173 | 24.62 |
| **Storbykommuner** | 136446 | 11.19 |
| **Provinsbykommuner** | 288595 | 23.67 |
| **Oplandskommuner** | 217088 | 17.81 |
| **Landkommuner** | 276750 | 22.70 |