

# PEG-OG-KLIK PÅ 45 MIN.

## Indhold

1	Hvad kan Peg-og-klik bruges til? .....	2
2	Lav en tabel.....	2
2.1	Start Peg-og-klik.....	2
2.2	Vælg datasæt og population .....	2
2.3	Tabeller for flere år? .....	3
2.4	Vælg variabler der skal summeres.....	3
2.5	Vælg akser og aggregeringer .....	4
2.6	Konfidentialitetsindstillinger.....	4
2.7	Testkørsel eller fuld population .....	5
2.8	Hvor ligger output?.....	5
2.9	Hvordan åbnes outputfilen? .....	5
2.10	Intet output – noget er gået galt .....	6
3	Lav en aggregering.....	6
3.1	Kopier!.....	6
4	Hvilke variabler findes der?.....	7
5	Nogle gængse variabler.....	8
6	Opgaver .....	9
6.1	Opgave 1: Virksomheders omsætning.....	9
6.2	Opgave 2: Lav et numerisk format .....	9
6.3	Opgave 3: Andel i årene 2010-2015 .....	9
6.4	Opgave 4: Lav et format, der udvælger 3 kommuner.....	9
6.5	Opgave 5: Beskæftigelsesstigning .....	9
6.6	Opgave 6: Flytninger .....	9
6.7	Opgave 7: Indtægtsgevinst ved pendling .....	9
6.8	Opgave 8: Pendling i 2 regioner .....	9
6.9	Opgave 9: Pendling med nord/syd-opdeling af Aalborg .....	9
7	Løsninger .....	10
7.1	Opgave 1: Virksomheders omsætning.....	10
7.2	Opgave 2: Lav et numerisk format .....	10
7.3	Opgave 3: Andel i årene 2010-2015 .....	11
7.4	Opgave 4: Lav et format, der udvælger 3 kommuner.....	11
7.5	Opgave 5: Stigning i antal ansatte .....	12
7.6	Opgave 6: Flytninger .....	12
7.7	Opgave 7: Indtægtsgevinst ved pendling .....	13

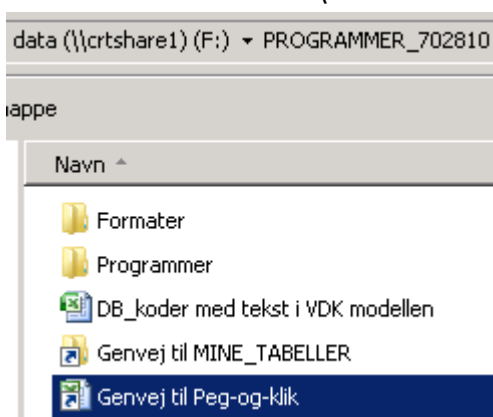
## 1 Hvad kan Peg-og-klik bruges til?

Peg-og-klik er et alternativ til DSTs Statistikbank, hvor Overvågningsmodellens brugere kan udtrække meget fleksible tabeller. Hvis den ønskede tabel findes i Statistikbanken er det nemmest at tage den der fra. Men hvis der er brug for mere specifikke eller særlige oplysninger, kan Overvågningsmodellens database af person- og virksomhedsdata tilgås vha. Peg-og-klik.

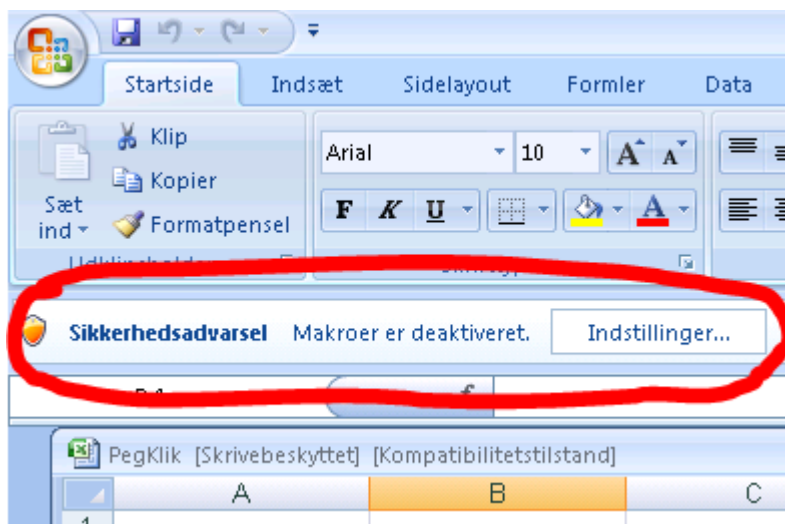
## 2 Lav en tabel

### 2.1 Start Peg-og-klik

På F-drevet under F:\PROGRAMMER\_702810 ligger "Genvej til Peg-og-klik".



Hvis brugerfladen ikke kommer frem er det sandsynligvis fordi du har makroer deaktiveret i Excel. Det varierer mellem forskellige versioner af Excel, hvordan makroer aktiveres. I Excel 2007 kommer et bånd frem med sikkerhedsadvarsel og en knap til aktivere makroer.

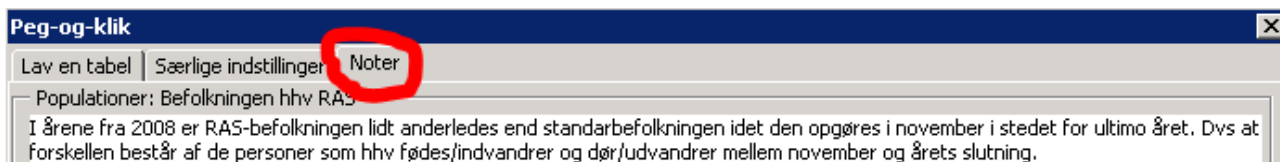


### 2.2 Vælg datasæt og population

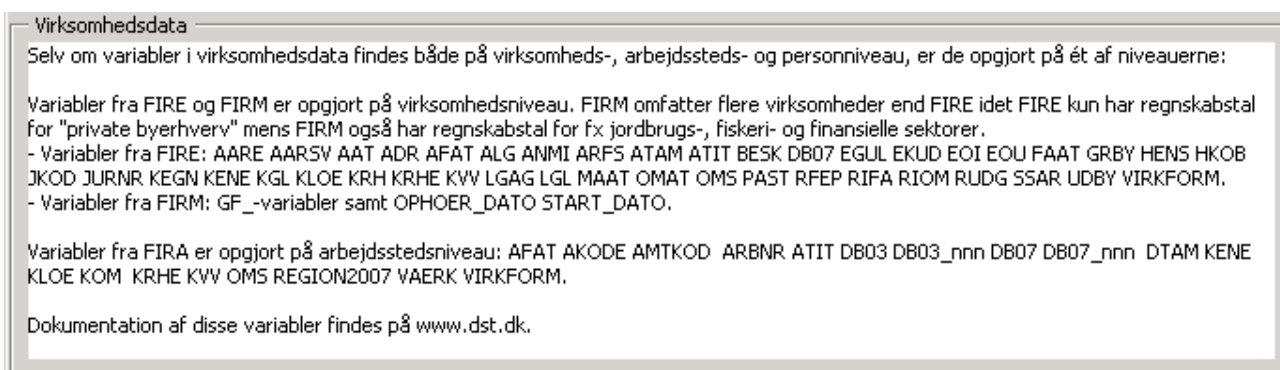
Tabeller kan baseres på enten person- eller virksomhedsdata.

I persondata er enheden altid enkeltpersoner, som kan vælges fra forskellige populationer: Enten standatdbefolkningen (hele befolkningen pr ultimo året) evt.

afgrænset efter alder, eller RAS-befolkningen, som opgøres i november, og evt kan afgrænses til de som er beskæftigede på det tidspunkt. I de fleste tilfælde vil standardbefolkningen være at foretrække, men hvis der kun ønskes arbejdsmarkedsrelaterede oplysninger, kan RAS-befolkningen være mere relevant. I RAS-befolkningen dækker kun variabler fra RAS-registret alle personer: [Link til indhold af RAS](#). Dette er også beskrevet under fanen "Noter":



I virksomhedsdata kan enheden enten være virksomheder, arbejdssteder, eller ansatte. Oplysninger på virksomheds- og arbejdsstedsniveau findes som udgangspunkt kun for private byerhverv (dvs. ingen primære eller finansielle erhverv) men nogle oplysninger findes for alle virksomheder. Dette er også beskrevet under "Noter":



### 2.3 Tabeller for flere år?

Som udgangspunkt trækkes tabeller fra

Serie fra år

ét år:

men det er også muligt at trække en serie af tabeller på én gang:

Serie fra år  Serie til

Hvis der udtrækkes en serie, skal sumvariabler og alle akser komme fra samme dataår og output er tilpasset så det viser udvikling over tid.

Resten af beskrivelsen her er for "Ej serie".

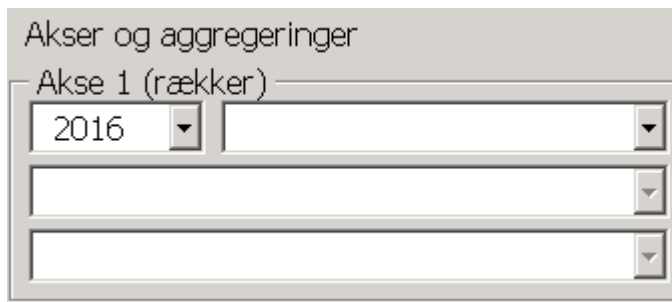
### 2.4 Vælg variabler der skal summeres

En tabel vil altid indeholde antal enheder i cellerne (personer/virksomheder/arbejdssteder) og kan derudover summere op til 4 numeriske variabler, der kan vælges frit fra alle dataår:

Sum i år  Sum af variabel  Sum af variabel

## 2.5 Vælg akser og aggregeringer

Tabellen kan frit defineres i op til 6 dimensioner blandt alle variabler. De 2 første akser definerer den enkelte tabel (hvad skal der være i rækker og søljer) f.eks. ALDER \* KOEN. De øvrige 4 akser bruges til at opdele populationen, f.eks. efter bopælskommune KOM og højst fuldførte uddannelse HFAUDD. Dette eksempel vil give en meget detaljeret tabel for alder-/kønsfordeling for hver uddannelse i hver kommune. I de fleste tilfælde vil det være praktisk at aggregere aksevariabler med passende formater, f.eks. 10-års intervaller for alder, gruppering af uddannelser, eller region i stedet for kommune.



År: Akserne kan baseres på samme eller forskellige år. Det 'almindelige' vil være at lave tabeller med akser fra samme dataår, men vha. akser fra forskellige dataår kan man f.eks. trække overgangstabeller (f.eks. Akse1 fra 2015 og Akse 2 fra 2016) eller betinge på bopæl, beskæftigelse, eller andet fra tidligere år.

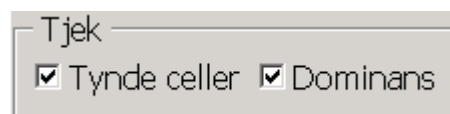
Aksevariabel: Den variabel, akse skal baseres på.

Aggregering: Alle aggregeringer kan vælges, men udtrækket vil ikke fungere hvis formatet ikke passer til aksevariablen. Aggregeringer med samme navn som aksevariablen står øverst på listen sammen med brugerens egne aggregeringer (mere om dette nedenfor).

Andel: Hvis tabellen laves for virksomhedsdata på virksomheds- eller arbejdsstedeniveau, kan andelen af ansatte, der opfylder et givet kriterie, opgøres. Se f.eks. filen "andel\_ALDER\_over60.sas", som definerer 0/1-indikator for at være mindst 60 år.

## 2.6 Konfidentialitetsindstillinger

Som udgangspunkt tjekkes for tynde celler ( $N < 3$ ) og dominans (de 2 største værdier udgør tilsammen 80%) og disse celler blændes. Hvis både tynde og dominerede celler tillades, føjes en "Alle"-kategori til alle akser.



DST kræver generelt at tynde celler ikke rapporteres, mens dominanskriteriet kun gælder for virksomhedsdata. De generelle konfidentialitetskrav kan findes [her](#) og [her](#) (især "Retningslinjer for hjemsendelse af analyseresultater").

Bemærk: Det, at Peg-og-klik blænder tynde celler og celler, der ikke opfylder dominanskriterium, garanterer ikke at tabellen opfylder DSTs konfidentialitetskrav og CRT påtager sig intet ansvar for, hvad der hjemsendes.

## 2.7 Testkørsel eller fuld population

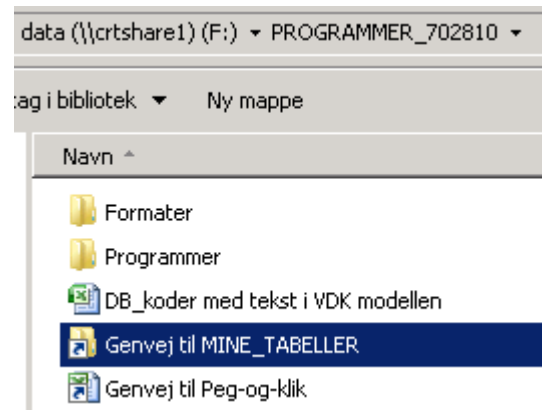
Før du trækker tabellen for fuld population, kan du tjekke tabellens opbygning/indhold med en hurtig kørsel for 10.000 observationer:



HUSK at lukke testtabellen før du laver fuld kørsel, ellers kan testfilen ikke overskrives.

## 2.8 Hvor ligger output?

Tabellen placeres nummereret .xls-fil i mappen MINE\_TABELLER på D-drevet på den maskine du er logget på:

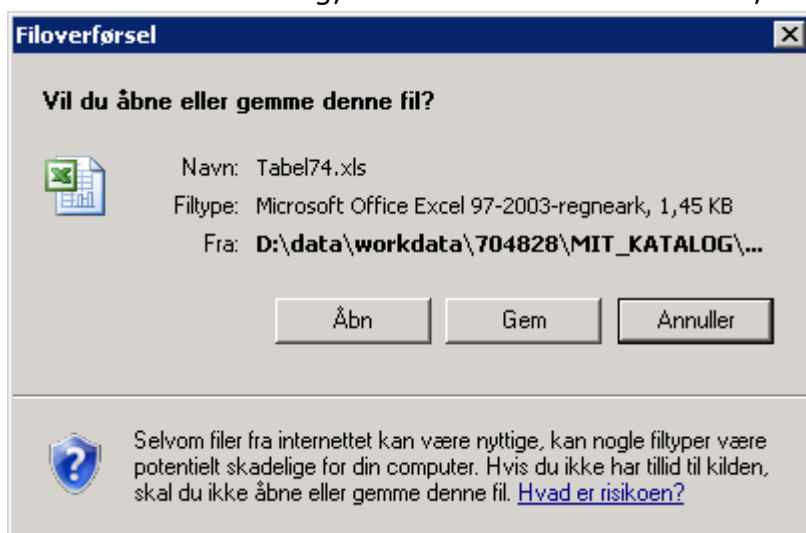


D:\data\workdata\...\MIT\_KATALOG\MINE\_TABELLER

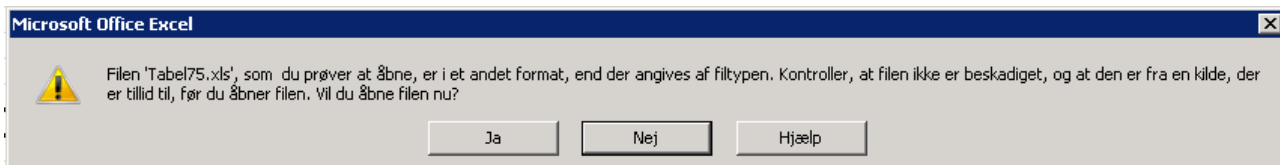
Der findes en genvej til denne mappe samme sted som genvejen til Peg-og-klik:

## 2.9 Hvordan åbnes outputfilen?

Når tabellen er færdig, kommer denne meddelelse, hvor du kan klikke "Åbn":



Selv om output ligger i .xls-fil, som åbner i Excel, er indholdet egentlig html og Excel vil derfor spørge:



Hvis du laver beregninger i regnearket og vil gemme dem, skal du gemme i et "rigtigt" Excel-format som .xlsx eller lignende.

## 2.10 Intet output – noget er gået galt

Log-filer placeres i LOG-undermappen til MINE\_TABELLER hvor tabeloutput placeres og kan give en ide om, hvad der er gået galt. Problemer skyldes oftest

- En aggregering/format fungerer ikke eller passer ikke med aksevariablen: Se nedenfor.
- Du har prøvet at lave en større tabel, eller flere store tabeller, end SAS/Excel kan håndtere. Dette sker hvis du på én gang bruger for mange aksevariable med for mange niveauer, og kan afhjælpe ved at tilpasse aggregeringer så du kun får detaljerede oplysninger hvor det er relevant. F.eks. kan det være smart at aggregere KOM eller POSTNR udenfor den relevante region.

## 3 Lav en aggregering

Du kan helt frit lave aggregeringer, som passer til de udtræk du har brug for. Dvs. du kan tilpasse gruppering af uddannelser, sogne, aldersgrupper, indkomstintervaller, boligstørrelser eller andre aksevariable.

En aggregering er et SAS-format, som ligger i en .sas-fil enten på Share-maskinen i mappen F:\PROGRAMMER\702810\Formater, hvor det kan bruges af alle, eller på din lokale maskine i mappen

D:\data\workdata\

### 3.1 Kopier!

Langt den nemmeste måde at lave en aggregering er at kopiere en lignende eksisterende aggregeringsfil til D-drevet og tilpasse den. For at en aggregering skal virke er det...

- Vigtigt at filnavnet svarer til formatets navn så filnavnet er:  
fmt\_<formatnavn>.sas
- Bedst hvis format- og filnavn svarer til den variabel formatet skal anvendes på, så formatet står øverst på listen.
- Vigtigt at syntax'en er 100% korrekt: Dollartegn på formater til karaktervariable, semikolon og anførselstegn de rigtige steder, "run;" til sidst osv. Det er nemmest at tilpasse eksisterende format og testkøre resultatet i SAS (tjek i loggen at det virker).

Udover F:\PROGRAMMER\702810\Formater –mappen er der et par steder med variabelinformation, der kan være relevant:

- DSTs formatbibliotek inde i forskermaskinen med skrivebordsgenvejen
- [Høj kvalitetsdokumentation og TIMES](#) på dst.dk
- [Register- og variabeloversigt](#) kan hjælpe med at finde det du søger
- Nogle gange virker det at Google efter "dst.dk <variabelnavn>"



#### 4 Hvilke variabler findes der?

Projektdatabasen indeholder op i mod 400 personvariabler og 150 virksomhedsvariabler, som er tilgængelige i Peg-og-klik. Derudover er der et stort antal andre variabler, der ikke umiddelbart er tilgængelige, f.eks. sundhed/offentligt forbrug, e-indkomster, DREAM, mm.

Princip: Variablene skal hedde det samme som i de rå registre, også når navnene er mindre oplagte.

I mappen F:\PROGRAMMER\_702810 ligger lige nu tre filer, der kan hjælpe med at finde variabler:

- "Om CRTs virksomhedsregister" beskriver virksomhedsregistrets opbygning og indhold.
- "Oversigt\_grunddata\_18maj" viser hvilke variabler, der er tilgængelige i hvilke år i Peg-og-klik, for hhv person- og virksomhedsdata, samt labels, der kan hjælpe med at søge efter relevante variabler.
- "Oversigt\_grunddata\_2018maj" viser hvilke variabler, der er leveret i hvilke registre fra DST.

Navn	Ændringsdato
Formater	23-05-2018 09:48
Programmer	24-05-2018 03:05
DB_koder med tekst i VDK modellen	24-08-2015 14:57
Genvej til MINE_TABELLER	29-11-2015 22:05
Genvej til Peg-og-klik	12-10-2017 10:50
Om CRTs virksomhedsregister	25-10-2017 10:40
Oversigt_grunddata_2018maj	24-05-2018 03:03
Oversigt_leverede_2018maj	24-05-2018 03:12

Register- og variabelnavne svarer i store træk til [Register- og variabeloversigt](#).

Planen er, at opdatere disse filer, især "Om CRTs virksomhedsregister", og tilføje en "Om CRTs personregister" med beskrivelse af opbygning og indhold af personregistret. Generelt arbejdes der løbende på at forbedre tilgængelighed, brugervenlighed osv.

## 5 Nogle gængse variabler

<b>Persondata</b>	
ALDER	Alder
KOEN	Køn
KOM	Bopælskommune
POSTNR	Postnummer
SOGN	Sogn
DSTNAVN	Bynavn
DSTKODE	Bykode
STORHED	Bystørrelse
KYSTAFST1	Bopæl <1000 m fra kyst
KYSTAFST3	Bopæl <3000 m fra kyst
ARBKOM	Arbejdsstedskommune
AFSTAND	Pendlingsafstand
BESK	Beskæftiget
DB07	Hovedbranche DB07
SOC_STATUS_KODE	Social status kode
HFAUDD	Højeste fuldførte uddannelse
UDD	Igangværende uddannelse
UDDKOM	Uddannelseskommune
IELAND	Oprindelsesland
STATSB	Statsborgerskabskode
PERINDKIALT_13	Personindkomst i alt ekskl. beregnet lejeværdi af egen bolig og før fr
DISPON_13	Indkomst efter skat og renter tillagt beregnet lejeværdi af egen bolig.
EJDV	Ejendomsværdi
SOCIO13	Socioøkonomisk klassifikation version 2013
DISCO_ALLE_INDK_13	Fagklassifikation for beskæftigelsesforhold, fra 1991 TIL ÅRET

<b>Virksomhedsdata</b>	
AARE	Årets resultat (efter selskabsskat)
AARSV	Antal ansatte (i årsværk) i regnskabsperioden.
ATIT	Investeringer, årets tilgang
DB07	DB07 branche
KOM	Kommune
OMS	Omsætning
CRT_AKG	Afkastningsgrad = $ARFS/AT$
CRT_EKF	Egenkapitalforrentning = $AARE/(EGUL-KEGN/2)$



## 6 Opgaver

### 6.1 Opgave 1: Virksomheders omsætning

Hvor mange virksomheder er der i hver region, og hvad er deres samlede omsætning?

### 6.2 Opgave 2: Lav et numerisk format

Lav et format, der kan opdele en numerisk variabel, som f.eks. OMS, i intervallerne "op til 1 mio", "1-10 mio", og "over 10 mio".

### 6.3 Opgave 3: Andel i årene 2010-2015

Hvilken andel af virksomhederne i hver region havde i årene 2010-2015 en omsætning på over 10 mio?

### 6.4 Opgave 4: Lav et format, der udvælger 3 kommuner

Lav et format, der udvælger de to kommuner Skive og Viborg, og samler resten som "Øvrige"

### 6.5 Opgave 5: Beskæftigelsesstigning

Hvilke brancher har, på 27-niveau, i hver af de to kommuner haft størst stigning i ansbeskæftigelsen fra 2010 til 2015?

### 6.6 Opgave 6: Flytninger

Hvad er flyttemønstret mellem regionerne i perioden 2010-2016?

### 6.7 Opgave 7: Indtægtsgevinst ved pendling

Havde kvinder eller mænd størst indtægtsmæssig gevinst ved at pendle i 2015?

### 6.8 Opgave 8: Pendling i 2 regioner

Træk pendlingsmønster på sogneniveau for Region Nord og Midt, mens andre regioner samles i "Ikke Nord/Midt".

### 6.9 Opgave 9: Pendling med nord/syd-opdeling af Aalborg

Træk pendlingsmønster på kommuneniveau for Region Nord og Midt, mens andre regioner samles i "Ikke Nord/Midt". Aalborg kommune skal opdeles i nord og syd (nord er sogne 8375, 8376, 8377, 8378, 8381, 8382, 8383, 8384, 8385, 8386, 8387, 9136, 9225. Syd er sogne 8276, 8318, 8319, 8320, 8321, 8322, 8323, 8324, 8325, 8326, 8327, 8328, 8329, 8330, 8331, 8332, 8333, 8342, 8343, 8344, 8345, 8356, 8357, 8362, 8363, 8364, 8365, 8366, 8367, 8368, 8369, 8370, 8371, 8372, 8373, 8374, 8388, 8389, 8390, 9064, 9076, 9107, 9110, 9157, 9187, 9197, 9227, 9232, 9237).

## 7 Løsninger

### 7.1 Opgave 1: Virksomheders omsætning

Hvor mange virksomheder er der i hver region, og hvad er deres samlede omsætning?

Peg-og-klik

Lav en tabel | Særlige indstillinger | Noter

Datasæt og population  
Virksomhedsdata | 1a. Virksomheder (FIRM)

Serie fra år: Ej serie | Sum i år: 2015 | Sum af variabel: OMS | Sum af variabel: [ ]

Check:  Tynde celler  Dominans | Sum af variabel: [ ] | Sum af variabel: [ ]

Akser og aggregeringer  
Akse 1 (rækker): 2015 | KOM | Akse 2 (søjler): 2015 | [ ]  
KOM\_5\_regioner | [ ]

### 7.2 Opgave 2: Lav et numerisk format

Lav et format, der kan opdele en numerisk variabel, som f.eks. OMS, i intervallerne "op til 1 mio", "1-10 mio", og "over 10 mio".

```
proc format;
  value OMS_3a
    low-<1000 = '1. Under 1 mio'
    1000-<10000 = '2. 1-10 mio'
    10000-high = '3. Over 10 mio';
run;
```

### 7.3 Opgave 3: Andel i årene 2010-2015

Hvilken andel af virksomhederne i hver region havde i årene 2010-2015 en omsætning på over 10 mio?

Peg-og-klik

Lav en tabel | Særlige indstillinger | Noter

Datasæt og population  
Virksomhedsdata | 1a. Virksomheder (FIRM)

Serie fra år: Ej serie | Sum i år: 2015 | Sum af variabel: | Sum af variabel: |

Check  
 Tynde celler  Dominans | Sum af variabel: | Sum af variabel: |

Akser og aggregeringer  
Akse 1 (rækker): 2015 | KOM | Akse 2 (søjler): 2015 | OMS  
KOM\_5\_regioner | OMS\_3a

### 7.4 Opgave 4: Lav et format, der udvælger 3 kommuner

Lav et format, der udvælger de tre kommuner Skive, Struer og Viborg, og samler resten som "Øvrige"

Baseret på filen fmt\_KOM\_99\_landsdele.sas:

```
proc format;
  value $KOM_3b
    779 = '86 Skive'
    671 = '87 Struer'
    791 = '88 Viborg'
    other = 'XX Andet';
run;
```

## 7.5 Opgave 5: Stigning i antal ansatte

Hvilke brancher har, på 16-niveau, i hver af de tre kommuner haft størst stigning i antal ansatte fra 2010 til 2015?

Peg-og-klik

Lav en tabel | Særlige indstillinger | Noter

Datasæt og population  
Virksomhedsdata | 1a. Virksomheder (FIRM)

Serie fra år: 2010 | Serie til: 2015 | Sum af variabel: AARSV

Check  
 Tynde celler |  Dominans

Akser og aggregeringer  
Akse 1: Året | KOM | KOM\_3b  
Akse 2: Året | DB07 | DB07\_16a

## 7.6 Opgave 6: Flytninger

Hvad er flyttemønstret mellem regionerne i perioden 2010-2016? Sammenlign aldersgrupperne 0-29, 30-59, 60+.

Peg-og-klik

Lav en tabel | Særlige indstillinger | Noter

Datasæt og population  
Persondata | 1. Standardbefolkning, alle personer

Serie fra år: Ej serie | Sum i år: 2016

Check  
 Tynde celler |  Dominans

Akser og aggregeringer  
Akse 1 (rækker): 2010 | KOM | KOM\_5\_regioner  
Akse 2 (søjler): 2016 | KOM | KOM\_5\_regioner  
Akse 3 (flere tabeller): 2010 | ALDER | ALDER\_3b  
Akse 4 (flere tabeller): | |

## 7.7 Opgave 7: Indtægtsgevinst ved pendling

Havde kvinder eller mænd størst indtægtsmæssig gevinst ved at pendle i 2015?

**Peg-og-klik**

Lav en tabel | Særlige indstillinger | Noter

Datasæt og population  
 Persondata | 1. Standardbefolkning, alle personer

Serie fra år: Ej serie | Sum i år: 2016 | Sum af variabel: PERINDKIALT\_13 | Sum af variabel: [ ]

Check:  Tynde celler |  Dominans | Sum af variabel: [ ] | Sum af variabel: [ ]

Akser og aggregeringer

Akse 1 (rækker): 2015 | AFSTAND | AFSTAND\_7a | [ ]

Akse 2 (søjler): 2015 | KOEN | KOEN\_3a | [ ]

Afstand	Kvinder					Mænd				
	Antal	Andel	Gns Indk	Gevinst	Gevinst%	Antal	Andel	Gns Indk	Gevinst	Gevinst%
1. Under 5 km	486143	37%	316176			500829	35%	387945		
2. 5-10 km	250451	19%	349039	32864	10%	244186	17%	438143	50198	13%
3. 10-20 km	255955	20%	368850	19810	6%	264152	18%	462828	24685	6%
4. 20-50 km	240888	18%	387632	18782	5%	283821	20%	493820	30991	7%
5. 50-100 km	50196	4%	397001	9369	2%	93080	6%	499033	5213	1%
6. 100-200 km	16472	1%	379370	-17631	-4%	38646	3%	555252	56220	11%
7. over 200 km	7642	1%	401201	21831	6%	18935	1%	532872	-22380	-4%