Dokumentation af interviewundersøgelser

# Kvalitetsdeklaration

|  |  |
| --- | --- |
| Opgave | AAU JOB 2018 |
| Kunde | Aalborg Universitet AAU |
| *Delopgaver udført af:*  Population  Stikprøve  Dataindsamling | Danmarks Statistik  Danmarks Statistik  Danmarks Statistik |
| Population | 2.076.617 personer |
| Bruttostikprøve | 3.162 personer |
| Nettostikprøve | 3.117 personer |
| Antal svar (fulde besvarelser) | **1.244 personer** |
| Delvise | 145 personer |
| Antal svar med både delvise og fulde besvarelser | 1.389 personer |
| Svarprocent i forhold til nettostikprøve (fulde besvarelser) | **39,9 pct.** |
| Indsamlingsmetode | Web og telefon |
| Svar fordelt på indsamlingsmetode (telefon-web) | Web 814 svar / 65,4 pct.  Telefon 430 svar / 34,6 pct. |
| Bortfaldskategorier og delvise | 2 Delvist gennemført 145 pers.  3 Nægter 291  4 Øvrigt bortfald 13  5 Sprogvanskeligheder 17  7 Ikke kontakt på tlf.nr. 104  8 Ikke fundet tlf.nr. 534  9 IP ikke kontaktet 45  10 Ikke truffet 769 |
| Undersøgelsesperiode (web-telefon-post) | Web 14. marts 2019 - 25. maj 2019  Telefon 19. marts 2019 - 14. april 2019 |
| Udsendelse og rykkere | Udsendelse af digital post (brev til dem der ikke er tilmeldt digitalpost) den 14. marts.  Reminder med digitalpost (brev til dem der ikke er tilmeldt digitalpost) den 26. marts.  Rykker digital post 12. april  Rykker digitalpost til ubesvarede og delvise besvarelser 03. maj |
| Sprog og oversættelse af brev og skema | Nej |
| Anvendt tolkning på andre sprog | Nej |
| Pilotundersøgelse | Ja |
| Instruktion på undersøgelse | Ja – Danmarks Statistik |
| Opregning og vægtning af data | Ja |

# Særligt for denne undersøgelse

Population

Populationen dannes på baggrund af E-Indkomst basisregister (EIR). Datarens er foregået i nedenstående steps:

1. Data er udtrukket for november 2018 (4.645.733 observationer).
2. Hvis en person er registreret med flere end ét job, så er jobbet med de højeste antal timer beholdt. Hvis der er flere jobs med de samme antal timer, så er dét job med den seneste slutdato beholdt (2.805.645 observationer tilbage). I nedenstående eksempler bliver observationen med 16 timer beholdt for person A. For person B er er alle observationer ens, så beholdes den første observation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pers | BJO\_JOB\_SLUT\_DATO | BJO\_LOENTIMER | BJO\_SMALT\_LOENBELOEB | PRODNR |
| A | 30. november 2018 | 6,08 | 1195,7 | 11 |
| A | 30. november 2018 | 6,08 | 1195,7 | 11 |
| A | 30. november 2018 | 16,00 | 3150,5 | 22 |
| A | 30. november 2018 | 2,85 | 560 | 33 |
| A | 30. november 2018 | 2,85 | 560 | 33 |
| A | 30. november 2018 | 12,80 | 2525,3 | 44 |
| A | 30. november 2018 | 12,80 | 2525,3 | 55 |
| B | 30. november 2018 | 55,14 | 12.309 | 66 |
| B | 30. november 2018 | 55,14 | 12.309 | 66 |
| B | 30. november 2018 | 55,14 | 12.309 | 66 |
| B | 30. november 2018 | 55,14 | 12.309 | 66 |

1. Observationer med løntimer (BJO\_LOENTIMER) mindre end 17,09 (laveste decil) eller lønbeløb (BJO\_SMALT\_LOENBELOEB) mindre end 2.727 kr. (laveste decil) om måneden er slettet (2.651.297 observationer tilbage).
2. Data er kørt sammen med Erhvervsregistret (ESR) pr. 3. kvartal 2018, som er nyeste version med beskæftigelsesoplysninger på nuværende tidspunkt. Nøglen mellem de to registre er arbejdsstedsnummeret (ARB\_NR). Personer fra arbejdssteder i EIR, hvor arbejdsstedet ikke har match i ESR er sorteret fra. Yderligere er personer fra arbejdssteder, som er registreret med en gennemsnitlig årsværk på mindre end 5 årsværk gennem de seneste 4 kvartaler, plus alle arbejdssteder som har ”O Offentlig administration, forsvar og politi” i branche 21-grupperingen sorteret fra. Efter dette step er der 2.173.960 observationer tilbage.
3. Efterfølgende er populationen yderligere afgrænset til personer i befolkningen pr. 31. december 2018 i alderen 18+ (Danmarks Statistiks befolkningsstatistik pr. 1. kvartal 2019). Alder er beregnet pr. 1. marts 2019. Udvalgspopulationen er således på 2.076.617.

Bruttostikprøve

Populationen er inddelt i 10 strata ved hjælp af arbejdsstedets beliggenhedsregion og størrelse i årsværk i to grupper (5-49 og 50+årsværk). Der er først udvalgt 3.162 personer til bruttostikprøve til undersøgelsen blandt populationen på de 2.076.617. Personer i arbejdssteder med 5-49 årsværk udgør 1/3-del af stikprøven og personer i arbejdssteder med 50+årsværk udgør 2/3-del.

Antal personer udvalgt i stratum med arbejdssteder med 5-49 årsværk er beregnet i forhold til det pågældende stratums andel i forhold til populationen af personer i alle arbejdssteder med 5-49 årsværk (982.858) samlet. Samme princip er brugt for personer i arbejdssteder med 50+ årsværk (1.093.759).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Strata | Antal udvalgt | | Antal svar | Strata-størrelse | Delpopulation  5-49 / 50+ årsværk | Population i alt |
| 81 Nordjylland - 5-49 årsværk | 109 | 41 | | 102.050 | 982.858 | 2.076.617 |
| 81 Nordjylland - 50+ årsværk | 182 | 72 | | 94.251 | 1.093.759 | 2.076.617 |
| 82 Midtjylland - 5-49 årsværk | 252 | 85 | | 235.423 | 982.858 | 2.076.617 |
| 82 Midtjylland - 50+ årsværk | 464 | 199 | | 240.295 | 1.093.759 | 2.076.617 |
| 83 Syddanmark - 5-49 årsværk | 229 | 90 | | 214.346 | 982.858 | 2.076.617 |
| 83 Syddanmark - 50+ årsværk | 392 | 169 | | 203.070 | 1.093.759 | 2.076.617 |
| 84 Hovedstaden - 5+ årsværk | 329 | 112 | | 307.591 | 982.858 | 2.076.617 |
| 84 Hovedstaden - 50+ årsværk | 883 | 340 | | 457.535 | 1.093.759 | 2.076.617 |
| 85 Sjælland - 5-49 årsværk | 132 | 50 | | 123.448 | 982.858 | 2.076.617 |
| 85 Sjælland - 50+ årsværk | 190 | 86 | | 98.608 | 1.093.759 | 2.076.617 |

Nettostikprøve

3.117 personer er kontaktet og de udgør nettostikprøven. 45 er ikke blevet kontaktet (har adressebeskyttelse, ukendt adresse, er udvandrede eller døde).

Opregning

Vi har tildelt vægte til alle, som har deltaget i undersøgelsen, 1.244 personer.

Følgende opregningsmodel er anvendt:

koen\*strata

aldergrp\*strata

loengrp2\*strata

udgrp\*strata

Følgende baggrundsvariable indgår i opregningen.

* **Koen (køn)**

1 = '1. Mænd'

2 = '2. Kvinder'

* **Aldergrp (Alder i 3 grupper)**

1 = '1. 18-39 år'

2 = '2. 40-59 år'

3= ’3. 60 år og derover’

* **Loengrp2 (Lønindkomst fra E-Indkomst i 2 grupper)**

1 = '-32.950 (middelværdi)’

2 = '32.950’

* **Udgrp (Højeste fuldførte uddannelse i 3 grupper)**

1 = '1. Grundskole eller uoplyst'

2 = '2. Gymnasial / erhvervsfaglig (EUD) / KVU'

3= ’3. MVU/Bachelor/LVU/PHD’

# Generelt om population og stikprøve

Skal have bopæl i Danmark

Ud over den eventuelle afgrænsning af populationen, som kan være lavet ud fra Danmarks Statistiks registre eller kundens egne registre, så afgrænses personer i populationen af den seneste befolkningsstatus. Den dannes hvert kvartal 30 dage efter kvartalsafslutning dvs. at 1. januar er klar primo februar osv.

For at tælles med i den danske befolkning, skal man have bopæl i Danmark og have tildelt et CPR-nummer.

Kvalitet af cpr

CPR’s befolkningsstatus er ret præcis. Usikkerheden på den samlede befolkning er få promille. Kun ved undersøgelser, hvor personer har flyttet ind og ud af Danmark, kan der være en større usikkerhed.

Personer som flytter til udlandet

Personer, som flytter ud af Danmark, er ikke tvunget til at lade det registrerer nogen steder. Der er derfor en mindre gruppe personer, som stadig er registret som en del af den danske befolkning, på trods af at de ikke længere har adresse i Danmark.

Udenlandske studerende

Uddannelsessøgende skal have et CPR-nummer for at blive optaget på en dansk uddannelse. De skal dog også være optaget på studiet før de kan få opholdstilladelse, og der kan derfor være et mindre antal personer, som ikke tælles med.

Adressebeskyttelse

Omkring 0,5 pct. af den danske befolkning har adressebeskyttelse, hvilket betyder, at de ikke deltager i interviewundersøgelser fra Danmarks Statistik. Desuden er der personer i Danmark som ikke kan kontaktes pga. at der ikke er oplysninger om deres adresse.

Sikring af repræsentativitet i undersøgelser

Danmarks Statistik danner populationer med udgangspunkt i Befolkningsregisteret (CPR) samt øvrige registeroplysninger. I denne proces anvender vi et unikt person-ID for alle personer (en anonym CPR-nøgle). Vi tester og fejlsøger data for at sikre at hver person kun optræder én gang i vores population.

Efter at vi har lavet stikprøvedesignet, danner vi stikprøven ved hjælp af SAS Proc SurveySelect. Her anvender vi SRS-metode (Simple Random Sample). Denne metode giver samme tilfældige udvalgschance (uden tilbagelægning) for alle deltagere. Metoden danner en ny tilfældig seedning hver gang og dermed en sikring af, at det ikke er de samme personer, som bliver udvalgt fra populationen gentagene gange.

Du kan læse mere om [PROC SURVEYSELECT funktionen her](http://support.sas.com/rnd/app/da/new/dasurvey.html#The SURVEYSELECT Procedure)

SAS-koden

Kodeeksempel:

**proc** **surveyselect**

data=pop\_select sampsize=**10050**

method=srs out=stikip

OUTSIZE STATS;

**run**;

Udvalgssandsynligheden

Vi kan også udvælge en stratificeret stikprøve, hvor der er forskellige udvalgssandsynligheder. Derfor gemmer vi altid udvalgssandsynligheden og stikprøvevægten for alle deltagere i undersøgelsen.

Test og kvalitetssikring

Efter dannelse af stikprøven, så tester og fejlsøger vi for dubletter. Vi laver tabeloversigter på hhv. population og stikprøve for at se om sammensætningen er nogenlunde ensartet på en række baggrundsvariable (køn, alder, herkomst, uddannelse, indkomst mv.) Herudover tester vi også for om der er udvalgt flere på samme bopæl, og om bopælen er en særlig adresse (med mange personer eller ugyldig adresse).

Vi validerer herefter stikprøven op mod en aktuel CPR-befolkning (få dage gammel) for at sikre os, at vi har korrekte oplysninger på alle deltagere, samt frasortering af beskyttede, døde og udvandrede personer.

Ved at anvende denne metode og tilhørende fejlsøgningsproces sikrer vi os at vores stikprøver er repræsentative for den population de er udtrukket fra.

# Danmarks Statistiks begreber

Populationen

Vores definition af population er efter match med seneste kvartals-befolkningsstatus.

Bruttostikprøve

Den udvalgte bruttostikprøve fra populationen er uden hensyn til adressebeskyttelse og seneste udvandrede/døde.

Nettostikprøve

Nettostikprøve indeholder personer, der kan kontaktes efter seneste cpr-opdatering. Adressebeskyttede personer er ikke med. Ligeledes er personer, som står med ’Rådhusadresser’ dvs. hjemløse og personer uden fast adresse mv. heller ikke er medregnet. (vejkode 99). Forskellen på brutto og netto er i denne undersøgelse 143 personer.

# Undersøgelsesforløbet

Telefonsøgning

Vi laver et dataudtræk med interviewpersonen og op til tre øvrige personer i husstanden over 18 år. Vi anvender den seneste adressekode fra CPR. Har der været en flytning inden for de seneste måneder, så medtager vi også den tidligere adressekode. Data sendes til telefonnummerberigelse.

Brev til interviewperson

Vi sender altid et brev, e-mail eller digital post til interviewpersonen inden vi starter interview. Har vi fundet et telefonnummer, så skriver vi normalt det i brevet. Hvis personen ønsker at blive kontaktet på et andet nummer, er dette muligt ved at henvende sig til Danmarks Statistik. Ligeledes opfordrer vi i brevet til at indsende et telefonnummer på mail, hvis vi ikke har fundet noget nummer på interviewpersonen.

Web og telefonundersøgelser

Ofte har vi en kombination af web- og telefonundersøgelse evt. også postskema. I brevet til interviewpersonen sender vi adgangskode, så personen selv kan udfylde skemaet på internettet. De har normalt 8-14 dage fra udsendelse til at udfylde på nettet, før vi begynder at kontakte pr. telefon.

# Svar

Svar

Vi tæller kun fuldt gennemførte interviews med i svarprocenten.

Svarprocent

Vi udregner normalt svarprocenten ud fra nettostikprøven. Der kan dog tages højde for eventuelle *ikke relevante*.

Dataindsamlingsmetode

Det vil fremgå med hvilken metode data er indsamlet med. Telefon, web eller postskema.

# Bortfaldskategorier

Ikke truffet

Vi har fundet telefonnummer, og har måske haft kontakt (eller optaget/ikke svar), men vi har ikke kunne gennemføre interview.

Nægter

Vi har haft kontakt til interviewpersonen, der har meddelt at han/hun ikke ønsker at deltage. Denne kategori inkluderer også delvist gennemførte interviews.

Øvrigt bortfald

Det er personer, som har sygdom/handicap eller er døde, samt personer som er væk fra hjemmet i interviewperioden. Vi har normalt haft kontakt til en anden end interviewpersonen i husstanden.

Forkert telefonnummer/ophørt

Denne kategori omfatter personer, hvor vi får ’nummer i uorden’/’tre klang’, samt personer der er flyttet fra adressen uden mulighed for telefonisk kontakt.

Ikke fundet nummer

Vores telefonsøgning har ikke fundet et telefonnummer til interviewpersonen.

Ikke truffet

Forsøgt kontaktet med den valgte dataindsamlingsform, fx mindst 5 opringninger, og ikke fået kontakt eller svar.