

Produktbeskrivning inför implementering av röjande- kontroll

Produktnamn och enhet:

Produktkod:

1. Beskrivning av undersökningen

Undersökning: *Total- eller urvalsundersökning.*

Objektstyp: *T.ex. företag eller individ.*

Tabelltyp: *Frekvens eller magnitud (summerbar målvariabel).*

Målvariabel (variabler): *T.ex. produktion, inkomst m.m.*

Övrig information: *Hierarkiska tabeller? Länkade tabeller (finns samma variabel i flera tabeller)? Hur stora är eventuella vikter (vid urval och/eller bortfallskompensation)?*

Bortfallshantering: *Imputering? Bortfallsandel?*

Medgivande: *Kommer produkten, om det är möjligt, att satsa på medgivande?*

Tidpunkt för publicering: *Om undersökningen har preliminär och definitiv publicering, ange båda tidpunkterna.*

Publicering av tabellerna: *Var publiceras tabellerna? SM, SSD m.m.*

Samarbete med andra undersökningar: *Lämnas mikrodata eller aggregerad data till andra undersökningar på SCB? Utnyttjar produkten mikrodata eller aggregerad data från andra undersökningar på SCB?*

Uppdrag: *Gör undersökningen uppdrag som kan innebära redovisning på en finare nivå?*



2. Behovet av röjandekontroll

Se Risker och konsekvenser ... sid 4

Röjanderiskbedömning: *Finns det risk att tabellerna röjer enskilda objekt, personer eller företag?*

Skaderiskbedömning: *Bedömer ni att enskilda personer eller företag kan lida skada eller men om vi röjer deras uppgifter?*

Nuvarande röjandekontroll: *Primär- och/eller sekundärundertryckningar? På vilket sätt görs röjandekontrollen? Vilken (eventuell) regel används för primärundertryckningar (p%, dominans eller tröskelvärdesregel)? Kan betydande tidsvinster uppnås med röjandekontrollverkyget?*

3. Beskrivning av mikrodata

Kort beskrivning av produktens mikrodata: *Hämtas mikrodata till alla tabeller från samma databas? Är det dokumenterat vilka variabler som är målvariabler respektive rad- och kolumnvariabler? O.s.v. (se nedan).*

För att röjandekontroll med verktyget Bifrost ska kunna utföras är det ett krav från röjandekontrollgruppen att mikrodata är iordningställt och väldefinierat innan implementeringen börjar.

Det ska vara dokumenterat varifrån de olika tabellerna hämtar sina mikrodata, samt vilka variabler som är målvariabler och vilka variabler som definierar de olika tabellerna (t.ex. en rad- eller kolumnvariabel). Om klassificeringskoder används behöver det finnas tillhörande förklaringstext, antingen i mikrodata eller specificerat i bifogat dokument. För att kunna utföra röjandekontroll av enheter/objekt, behöver de olika enheterna/objekten som bidrar med data vara definierade i mikrodata. Om undersökningen är en urvalsundersökning behöver även de slutliga uppräkningsvikterna vara med.

Om ni har många tabeller kan det vara ett tidskrävande arbete att iordningställa mikrodata med tillhörande dokumentation och vi rekommenderar att ni börjar med det i god tid.

Exempel

Nedan visas ett exempel där mikrodata är iordningställt (fiktivt datamaterial). I det här fallet innehåller mikrodata information som beskriver hur mycket av de olika bränsleslagen som företagen har använt för att producera el respektive värme. Här utgör kod radvariabel

och el- och vämeprod kolumnvariabler. Knr (kontroll-nummer) är de olika företagens unika identifikationsnummer.

knr	kod	kodnamn	elprod	varmeprod
10	100	Stenkol	120	800
11	100	Stenkol	240	450
12	100	Stenkol	680	210
10	200	Dieselbränsle	275	1640
11	200	Dieselbränsle	860	1510
12	200	Dieselbränsle	1375	90
10	300	Naturgas	80	500
11	300	Naturgas	0	185
12	300	Naturgas	120	100

Övriga kommentarer:

Kontaktpersoner:	Lars-Erik Almberg	PMU/MFÖ-Ö
	Ingela Düring	PMU/MFS-S
	Annika Gerner	PMU/MFS-S
	Helen Lindkvist	PMU/MFÖ-Ö
	Karin Valentin Assin	PMU/MIS-S

Risker och konsekvenser vid skydd och publicering av tabeller

1. Frekvenstabeller

(redovisar antal individer/objekt i olika celler)

Generellt

- Risken för röjande är större vid totalräknade data än vid urval
- Möjligheten att röja en individ ökar när marginaler finns tillgängliga

Tabeller med små värden publiceras

Möjliga risker:

- Individer kan identifiera sig själva
- Någon annan kan identifiera en individ i tabellen
- Attributröjande kan ske om en identifierad individ också kan kopplas till ny information

Möjliga konsekvenser:

- Individen lider men av att bli identifierad av sig själva eller någon annan
- Individen lider men av att ny information röjs
- Genom att små värden syns tydligt i tabellerna kan det uppfattas som att SCB inte tar frågan om skydd på allvar

Tabeller där alla individer hamnar i samma eller ett fåtal kategorier av en variabel publiceras

Möjliga risker:

- En grupp av individer kan kopplas till ny information trots att de inte kan identifieras unikt i tabellen

Möjliga konsekvenser:

- Individer lider men av att ny information röjs

Många tabeller publiceras baserat på samma material och med samma uppsättning bakgrundsvariabler i olika kombinationer

Möjliga risker:

- Tabeller kan kombineras så att individer kan identifieras
- Ny information kan kopplas till individerna

Möjliga konsekvenser:

- Individer lider men av att bli identifierade
- Individer lider men av att ny information röjs

Tabeller eller kartor publiceras på låg geografisk nivå

Möjliga risker:

- Antalet individer i det geografiska området är så få att identifiering avsevärt underlättas
- Ny information kan kopplas till individerna

Möjliga konsekvenser:

- Individen lider men av att bli identifierad av sig själva eller någon annan
- Individen lider men av att ny information röjs
- Genom att små geografiska områden syns tydligt i tabellerna eller kartorna kan det uppfattas som att SCB inte tar frågan om skydd på allvar

2. Magnitudtabeller

(redovisar t.ex. summor av en målvariabel i olika celler)

Generellt

- Risken för röjande är större vid totalräknade data än vid urval

Tabeller publiceras där några företag/enheter/objekt är dominerande i vissa celler

Möjliga risker:

- Ett objekt kan röja ett annat objekts uppgifter

Möjliga konsekvenser:

- Det röjda objektet lider ekonomisk skada

Tabeller publiceras där vissa celler har värdet noll

Möjliga risker:

- Alla får veta att samtliga objekt som tillhör cellen har värdet noll

Möjliga konsekvenser:

- Det röjda objektet lider ekonomisk skada
- Eftersom röjandet är tydligt för många kan det allmänt uppfattas som att SCB inte tar frågan om skydd på allvar

3. Allmänt

Tabeller publiceras där summor eller cellvärden inte stämmer mellan tabeller baserade på samma material (inkonsistens)

Möjliga risker:

- Missnöjda användare

Möjliga konsekvenser:

- SCB rykte försämrats