

Statistical disclosure control –
Show the Swedish forest but not all of
the trees



Niklas Notstrand

Authors Michael Carlson, Ingejård Janson, Ann-
marie Flygare

Whats the problem?

It may not be possible to identify an individual (individuals, households, businesses, etc.), or reveal anything about their personal or financial circumstances via a published or extradited materials.



Whys is there any problem?

- Legal aspects
- Ethical aspects
- Confidens of the respondents



Disclosure control in i tables



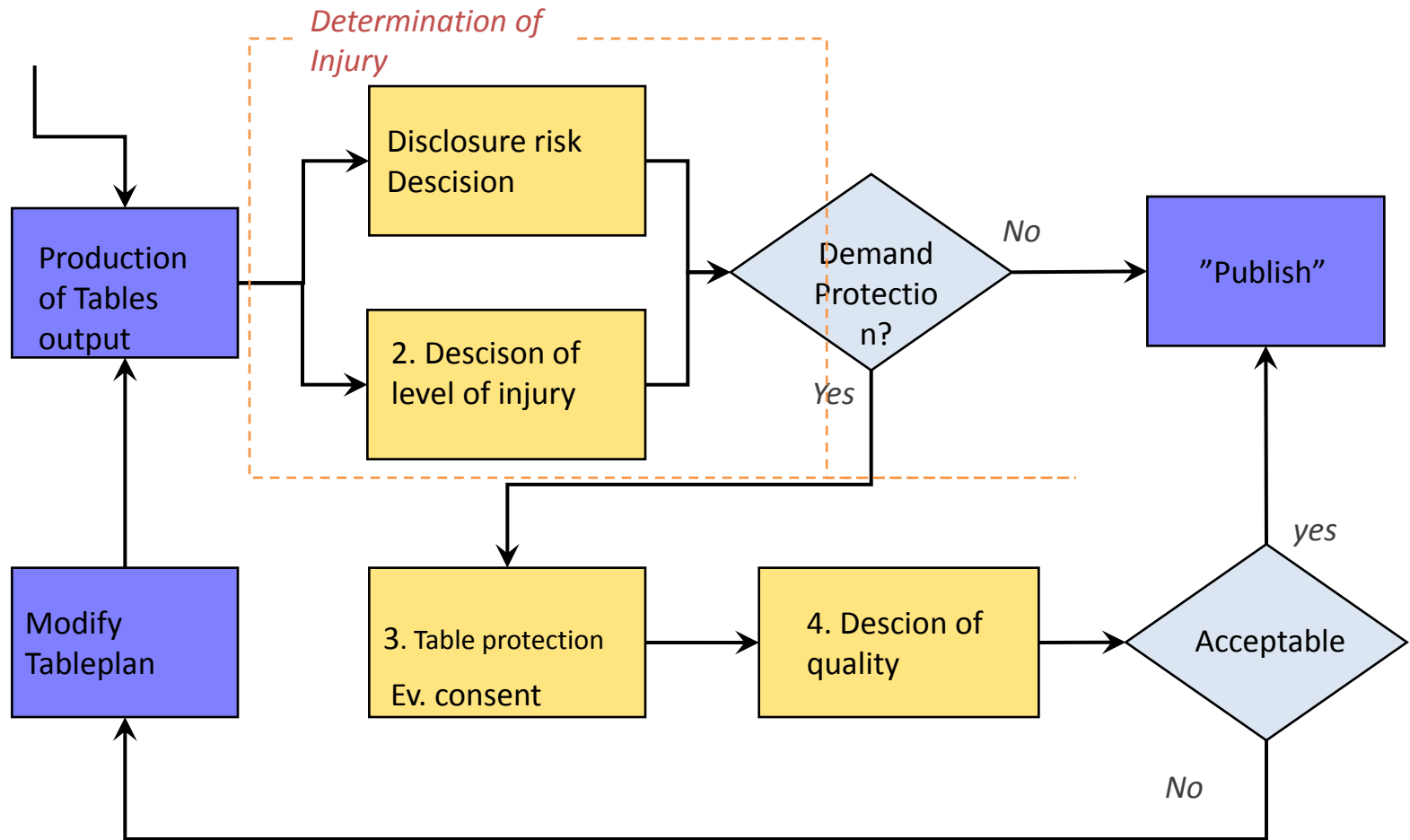
1. Risk for Disclosure
2. Need for protection of information
3. Protection against disclosure
4. Lack of missing information

Protection against disclosure

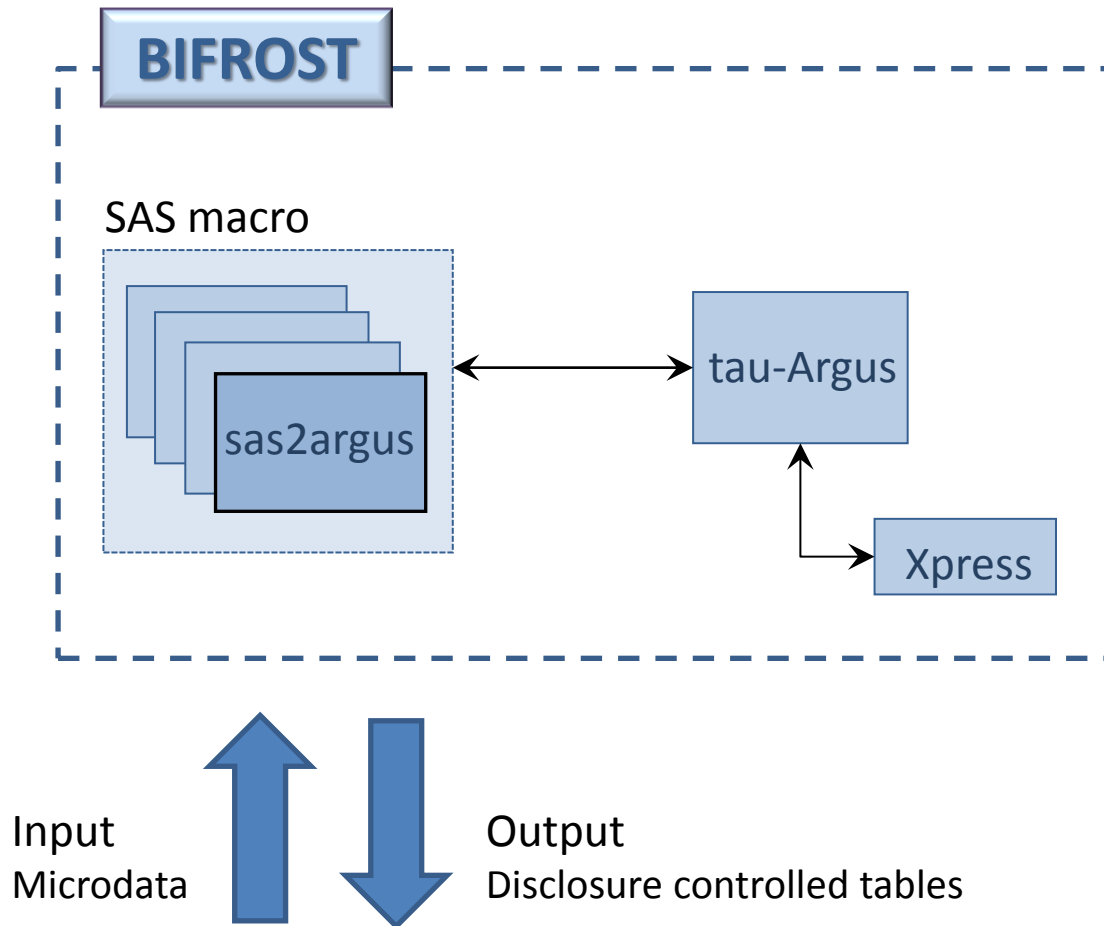
- Group strata on different levels, group columns together
- Hide sensitive cellvalues
- Round cellvalues
- Consent from respondents
- ...



Disclosure control– a part of the statistical production process



Example of IT-Tool for disclosure control of tables



Disclosure control with programming SAS-macro

- Magnitude Table with fictitious data sets where we report the use of coal, diesel and natural gas to produce electricity and heat
- Assume We've done a consent request of the three companies (consent survey)

Identity	Name of company	Consent
10	Bolag 1	1
11	Bolag 2	0
12	Bolag 3	1

Fictitious Microdata

Identit y	Code	Code name	elproduc tion	heatproduc tion
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

Result of microdata after calculations, editing and so on.

Identity	Code	Code name	columnne	Name of columnne	value
10	300	Naturgas	el	Elproduktion	80
12	300	Naturgas	el	Elproduktion	120
10	300	Naturgas	varme	Värmeproduktion	500
11	300	Naturgas	varme	Värmeproduktion	185
12	300	Naturgas	varme	Värmeproduktion	100

SAS-program script

```
%let variabel1          =code;
%let variabel2          =columnne;
%let varde_var         =value;
%let jobname           =table_1;
%let print format      =9.;

%sas2argus (           InData      = appdat.&jobname,
                  Jobname    = &jobname,
                  Explanatory = &variabel1 &variabel2,
                  Response   = &varde_var,
                  SafetyRule = P (15,1), /*NK (1,50) | NK (2,90), */
                  /*Suppress = MOD (1,10),*/
                  Out        = inter(1),
                  RunArgus   = 1,
                  SAS         = 2,
                  Debug      = 1
                )
%save_table ( &jobname, 1 )
```

SAS-script

```
%proc_tabulate (  
    Data      = &jobname,  
    Format   = code $kod.  columnne $columnne. ,  
    Row1     = &variable1,  
    Col1     = &variable2,  
    Response = &varde_var,  
    Suppression = 0,  
    Label    = ,  
    Excel    = 0,  
    Outformat = &print format  
)
```

SAS-output result

	Electricity production	Heat production	Total
Hardcoal	1040	1460	2500
Diesel fuel	2510	3240	5750
Naturgas	200	785	985
Total	3750	5485	9235

	Electricity production	Heat production	Total
Hardcoal	1040	1460	2500
Diesel fuel	2510	P	5750
Naturgas	P	785	985
Total	3750	5485	9235

SAS-code script with result of consent

```
%consent_control ( Microdata = appdat.&jobname,  
                   Consentdata = appdat.medgiv,  
                   Jobname    = &jobname,  
                   Safety_rule = P,  
                   Safety_var  = 15,  
                   Var1        = &variabel1,  
                   Var2        = &variabel2,  
                   Response   = &varde_var  
                   )
```

SAS-kod

- **Apriori**

Kod	Kolumn	Value	Freqvar	Top1	Top2	Statusvar	Medgiv1	Medgiv2
300	el	200	2	120	80	U	1	1

- **Unsafe**

Kod	Kolumn	Value	Freqvar	Top1	Top2	Statusvar	Medgiv1	Medgiv2
200	varme	3240	3	1640	1510	U	1	0

- **Unsafe_companies**

kod	kolumn	knr	bolagsnamn	value	high	Freqvar	Top1	Top2	Medgiv1	Medgiv2
200	varme	11	Bolag 2	3240	2	3	1640	1510	1	0

SAS-code for disclosure control

```
%arb_file ( &jobname, '<SUPPRESS> MOD(1,10)' )
```

```
%sas2argus (      InData      = appdat.&jobname,  
                  Jobname     = &jobname,  
                  Explanatory = &variabel1 &variabel2,  
                  Response    = &varde_var,  
                  SafetyRule  = P (15,1),  /*NK (1,50) | NK (2,90), */  
                  /*sekundärundertryckningen går in som parameter i %arb_fil*/  
                  Out         = inter(1),  
                  RunArgus   = 2,  
                  SAS        = 2,  
  Debug          = 1  
)
```

```
%save_table ( &jobname, 2 )
```


SAS-output table

%proc_tabulate (...)

	Elproduktion	Värmeproduktion	Total
Stenkol	S	S	2500
Dieselbränsle	S	P	5750
Naturgas	200	785	985
Total	3750	5485	9235

	Elproduktion	Värmeproduktion	Total
Stenkol	X	X	2500
Dieselbränsle	X	X	5750
Naturgas	200	785	985
Total	3750	5485	9235

%proc_tab_excel_consent (...)

SAS-script

```
%let ar                =2012;  
%let man              =3;  
                        ***** Conentdatafiles editing of data *****  
  
%let variabel1        =kod;  
%let variabel2        =kolumn;  
%let varde_var       =value;  
%let jobname         =tabell_1;  
%let utskriftsformat =9.;  
  
%sas2argus ( ... )  
%save_table ( ... )  
%consent_control ( ... )  
%arb_file ( ... )  
%sas2argus ( ... )  
%save_table ( ... )  
%proc_tab_excel_consent ( ... )
```

Thank you !!!

Authors

1. PHD Michael.Carlson@stat.su.se
2. PHD Annmarie.flygare@scb.se
3. PHD Ingejard.janson@scb.se
4. PHD Helen.lindkvist@scb.se