

Scanner data in the CPI/HICP

Which scanner data?

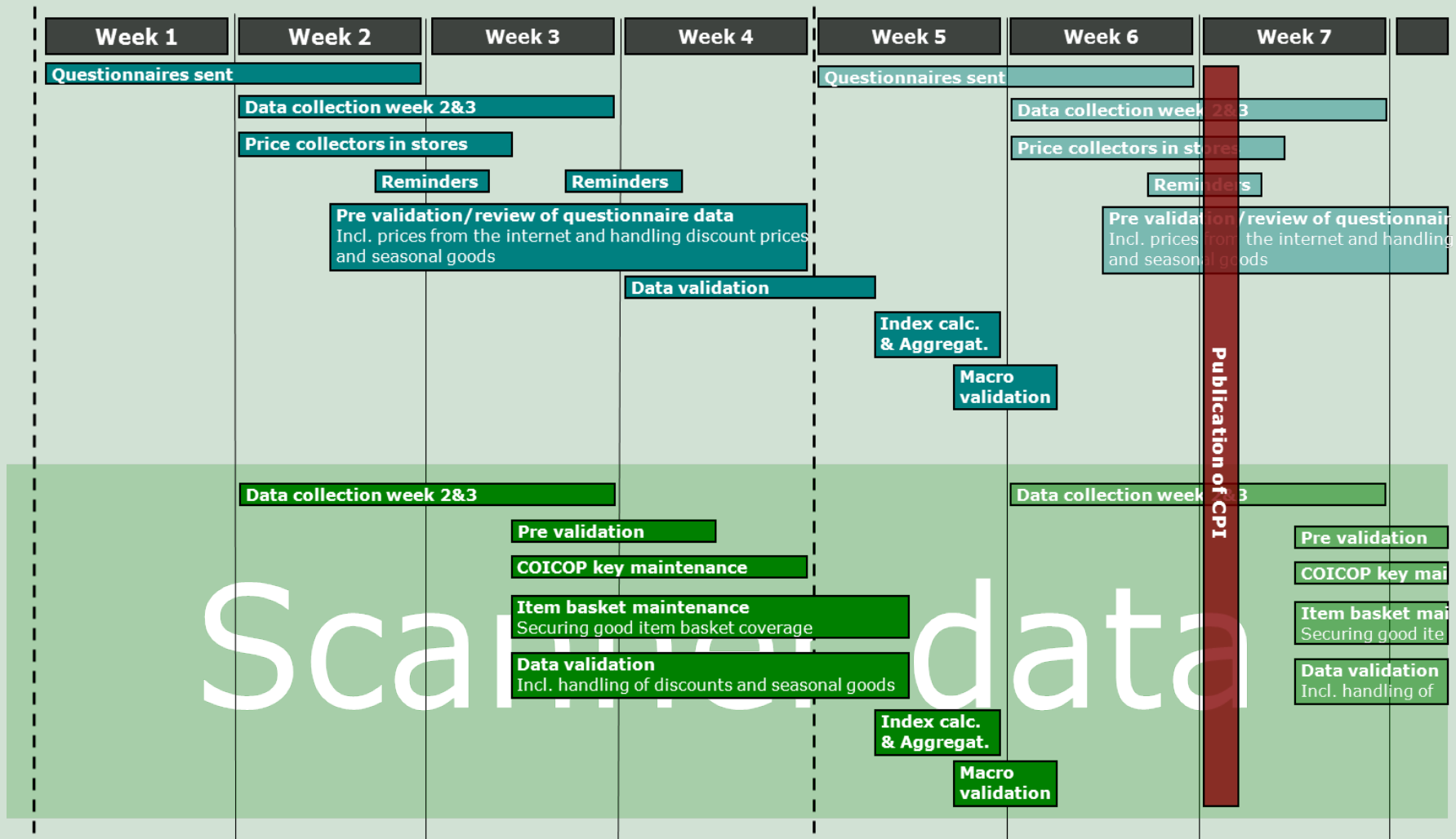
Date	EAN number	Turnover	Volume	Unit	Amount per unit	Product number	Product description
1104	2920080800007	3402,70	211	Gram	300	910076003	Bacon 2x150 G.
1104	2920080800007	2119,65	163	Gram	300	910076003	Bacon 2x150 G.
1104	2920080800007	1516,05	108	Gram	300	910076003	Bacon 2x150 G.
1104	2920080800007	1478,13	105	Gram	300	910076003	Bacon 2x150 G.
1104	2921056000005	302,50	14	Gram	200	911056001	Chicken filet
1104	2921056000005	102,50	5	Gram	200	911056001	Chicken filet



Which chains?

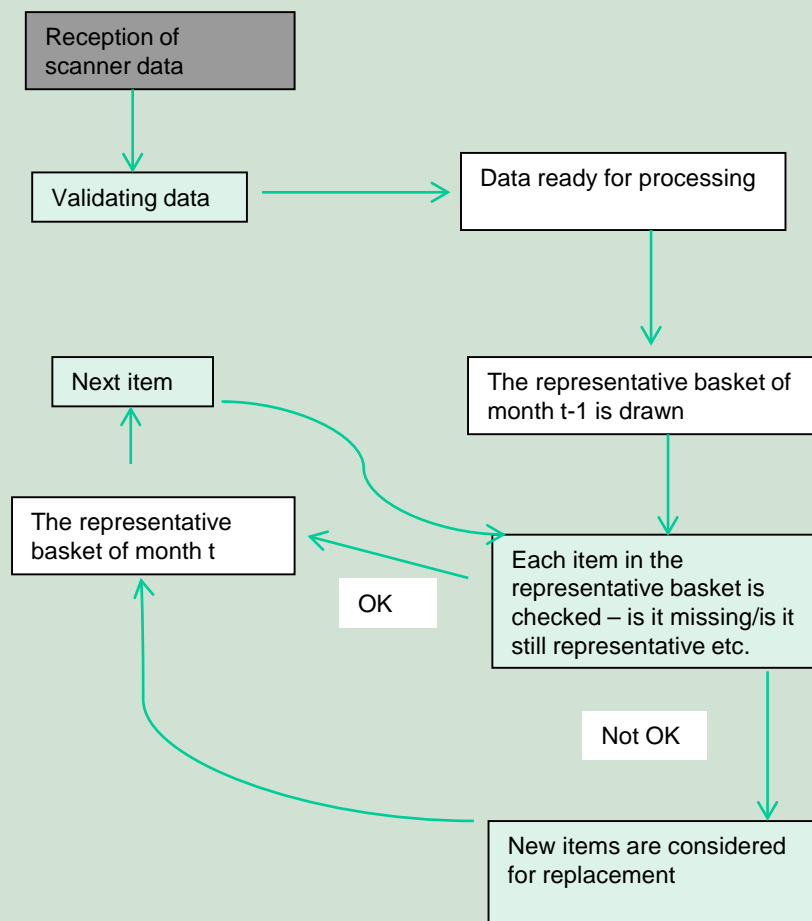
Supermarket	Weekly reception of scanner data since	Approx. sales percentage of the total Danish Food and beverage market
Dansk Supermarked	January 2011	25%
COOP	January 2011	25%
Rema 1000	January 2011	10%

Scanner data in the production



Scanner data

CPI/HICP by the use of scanner data





Data processing- 3 steps

- **First step**

Data checks:

1. It is checked whether the program date matches the given week
2. The data is checked for null-values
3. The number of records are checked
4. The numerical variables are checked for characters and vice versa




Continued

- **Second step:**
- The use of the supermarket chains own classifications manually assigned to COICOP. This manual assignment is only done once and only changed if the supermarket chain decides to change their classification.
- Where the supermarket classification is not precise enough (e.g. when covering several 6-digit COICOP categories) the classification is assigned to a special residual 4-digit COICOP group. These residual groups are monitored and handled by a search-word process.
- If a residual group has more than 5 % of the weekly turnover compared to the corresponding COICOP 4-digit group it has been decided to apply the search-word process.
- Future EANs will be processed by the actual search-word list and will therefore not show up in the residual groups anymore. Hence, the production team's work is stored properly and automatically applied in future COICOP assignments



Continued

- **Third step:**
- The third step deals with the COICOP-key quality. The different EAN-COICOP connections between the different supermarket chains should be the same and is checked for consistency.
- A manual check system for the rightful distribution of the EANs to the COICOP has been put in place. Thereby it is possible to manually alter the COICOP-group assigned to an EAN. This system is based upon Excel and SAS.



Drawing and maintaining a representative basket

- Statistics Denmark has decided to use a representative basket methodology for the calculation of the CPI/HICP. This is in contrast with a full population method, which is prone to drift and bias problems due to difficulties in taking proper account of seasonal goods and goods on discount leaving the sample.



Selection criteria

The system for developing a representative basket is based upon data processing in SAS. It has been decided that the different COICOP 1 and 2 groups as a general rule needs to be in the data for all 12 months of 2011 and be part of the 50% highest turnover for each shop. However individual selection criteria are also applied, so as to assure that the number of observations match the desired.

Sub- groups with too many observations

- The sub-groups' number of observations are chosen from the two-criteria-sample based on highest turnover. This means that each supermarket chain's (store level) share of the COICOP sub-group turnover is multiplied with the total number of observations desired for the sub-group, determining the number of observations per chain and store. Then the observations are chosen based on highest turnover.

Sub-groups with too few observations

- What we do is that, for each sub-group all of 2011's scanner data is collected and each item's share of the sub-groups total turnover is calculated. Then we examine which criterion we can set for the sub-group for how many months the item is available in data, i.e. the stability of the item. The stability criterion is set so that generally the best-selling items become part of the sample.

System for maintaining the representative basket

- It-staff have set up an Excel-SAS interface system which is to be monthly maintained by the HICP-section. For the different COICOP groups the section is presented with goods that have entered the scanner data, and that are to replace EANs, that are no longer present
- The EANs that are new candidates for the representative basket cover the last 4 months and they are sorted by shops and highest turnover

Continued

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	Stregkodedata - håndtering af kurv															
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																

1. Dan data

3. Opdater kurv

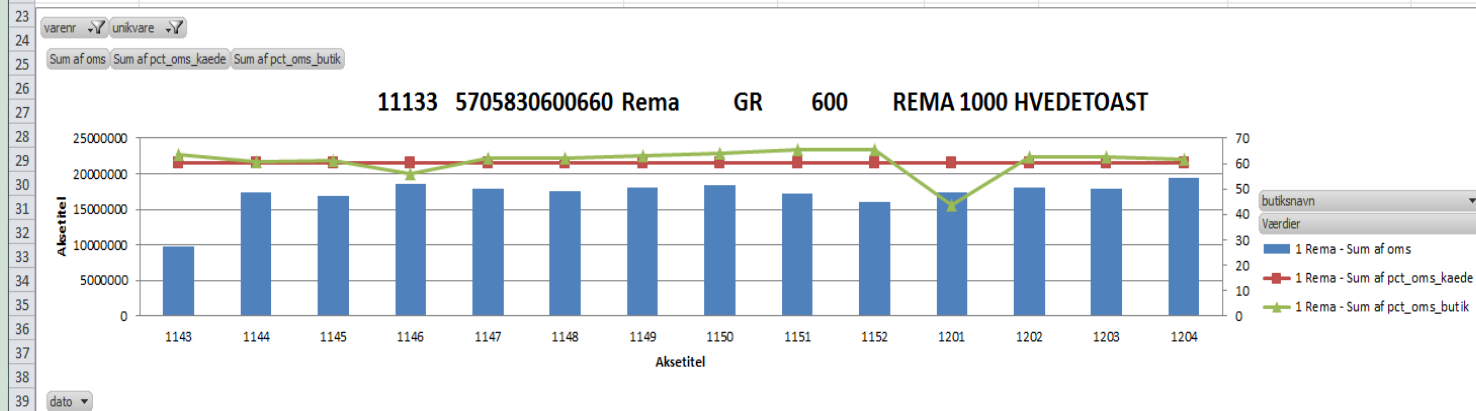
2. Hent udgåede varer

4. Dan data til Oracle

Continued

	A	B	C	D	E	F	G	H	I
1	varenr	11133							
2	unikvare	11133 5705830600660 Rema GR 600 REMA 1000 HVEDETOAST							
3									
4		Kolonnenavn							
5		1 Rema							
6	Rækkenavn	Sum af oms	Sum af pct_oms_kaede	Sum af pct_oms_butik					
7	1143	9847015	60,30938955	63,68456045	9847015	60,30938955	63,68456045		
8	1144	17294907	60,30938955	60,94545999	17294907	60,30938955	60,94545999		
9	1145	16925018	60,30938955	61,18392965	16925018	60,30938955	61,18392965		
10	1146	18643166	60,30938955	56,19646662	18643166	60,30938955	56,19646662		
11	1147	17920744	60,30938955	61,99608521	17920744	60,30938955	61,99608521		
12	1148	17568450	60,30938955	61,97021069	17568450	60,30938955	61,97021069		
13	1149	18129598	60,30938955	62,92639347	18129598	60,30938955	62,92639347		
14	1150	18396758	60,30938955	63,97433982	18396758	60,30938955	63,97433982		
15	1151	17238783	60,30938955	65,28642902	17238783	60,30938955	65,28642902		
16	1152	16012741	60,30938955	65,37598084	16012741	60,30938955	65,37598084		
17	1201	17385053	60,30938955	43,59711653	17385053	60,30938955	43,59711653		
18	1202	18036329	60,30938955	62,69056562	18036329	60,30938955	62,69056562		
19	1203	17913248	60,30938955	62,36777062	17913248	60,30938955	62,36777062		
20	1204	19371719	60,30938955	61,86375208	19371719	60,30938955	61,86375208		
21	Hovedtotal	240683529	844,3314537	854,0590606	240683529	844,3314537	854,0590606		

Dette vareløbenummer skal erstatte udgået vare i kurven



Continued

- These two systems, the representative basket (starting in January 2011) and the EXCEL-SAS interface system to maintain the representative basket, have then been implemented in our ORACLE-system for the production of the CPIs/HICPS.
- It is noted that the ORACLE-system has been implemented with semi-automatic check-systems. It is thereby possible for the HICP-section to check whether scanner data have the right amounts, taxes for the HICP-CT etc.:



Prisinddatering

Skema Stregkoder

Alle kæder

Hent alle varer

Korriger data

4943 Varer

Stregkoder

Alle butikker

Hent med mangler

Importer priser

155 Varegrupper

Gem

Vis alle varegrupper

Refresh/Fortryd

Oopdater varenr

	Ris								
1111000	BASMATI RIS COOP	1.0 kg	kg	1 kg	22.08	0	7340011352327	DAGLIGBRUGSEN	
1111000	BASMATI RIS COOP	1.0 kg	kg	1 kg	21.23	0	7340011352327	SUPERBRUGSEN	
1111000	BASMATI RIS COOP	1.0 kg	kg	1 kg	21.89	0	7340011352327	KVICKLY	
1111000	BASMATIRIS IRMA 1 KG	1.0 kg	kg	1 kg	20.16	0	5718147440010	IRMA	
1111000	BASMATIRIS 1 KG	1.0 kg	kg	1 kg	14.95	0	5701318016152	FAKTA	
1111000	GLOBAL CUISINE BASMATI RIS	1.0 kg	kg	1 kg	16.72	0	20000028121350	FØTEX/FOTEX	
1111000	GLOBAL CUISINE BASMATI RIS	1.0 kg	kg	1 kg	21.67	0	20000028121350	BILKA	
1111000	GLOBAL CUISINE JASMIN RIS	1.0 kg	kg	1 kg	16.54	0	20000028121527	FØTEX/FOTEX	
1111000	JASMINRIS 1 KG	1.0 kg	kg	1 kg	14.95	0	5701318016121	FAKTA	
1111000	ORIENTAL CLASSICS N BASMATI RIS	1.0 kg	kg	1 kg	14.95	0	20281215400032	NETTO	
1111000	ORIENTAL CLASSICS N JASMIN RIS	1.0 kg	kg	1 kg	14.95	0	20281215400063	NETTO	
1111000	PAMA GRØDRIS	0.5 kg	kg	5 kg	11.86	0	5730000421400	REMA	
	Mel								
1112100	BAGERENS HVEDEMEL 2 KG	2.0 kg	kg	2 kg	10.91	0	20000028101079	NETTO	
1112100	GLUTEN HVEDEMEL	2.0 kg	kg	2 kg	11.08	0	5701029006572	REMA	
1112100	GLUTEN HVEDEMEL	2.0 kg	kg	2 kg	20.72	0	5701029006572	FAKTA	
1112100	HVEDEMEL AMO 2 KG	2.0 kg	kg	2 kg	21.50	0	5701029006305	BILKA	
1112100	HVEDEMEL AMO 2 KG	2.0 kg	kg	2 kg	15.39	0	5701029006305	SUPERBRUGSEN	
1112100	HVEDEMEL AMO 2 KG	2.0 kg	kg	2 kg	19.47	0	5701029006305	KVICKLY	
1112100	HVEDEMEL AMO 2 KG	2.0 kg	kg	2 kg	21.55	0	5701029006305	FØTEX/FOTEX	
1112100	HVEDEMEL BALANCE 2KG	2.0 kg	kg	2 kg	21.93	0	5718147441147	IRMA	
1112100	HVEDEMEL 2 KG	2.0 kg	kg	2 kg	8.94	0	5701318001653	FAKTA	
1112100	HVEDEMEL 2 KG 1/4 PALLE NETTO	2.0 kg	kg	2 kg	9.95	0	20000028101055	NETTO	
1112100	HVEDEMEL 2 KG X-TRA	2.0 kg	kg	2 kg	9.81	0	7340011343431	DAGLIGBRUGSEN	
1112100	HVEDEMEL 2 KG X-TRA	2.0 kg	kg	2 kg	8.96	0	7340011343431	SUPERBRUGSEN	
1112100	HVEDEMEL 2 KG ÅNGLA	2.0 kg	kg	2 kg	19.92	0	7340011314974	SUPERBRUGSEN	
1112100	HVEDEMEL 2 KG ÅNGLA	2.0 kg	kg	2 kg	20.92	0	7340011314974	KVICKLY	
1112100	KORNKAMMERET HVEDEMEL ØKO	2.0 kg	kg	2 kg	21.95	0	5708731039836	FØTEX/FOTEX	
1112100	KORNKAMMERET HVEDEMEL ØKO	2.0 kg	kg	2 kg	21.95	0	5708731039836	BILKA	

Future work

Work process	Expected deadline
Finalization of IT-system	1. July 2014
Investigate when Supergros will be able to deliver data on a weekly basis	1. July 2014
Test calculations of CPI with the use of scanner data	1. July 2014-31. December 2014
Binding agreements with Supermarkets on delivering scanner data	1. August 2014
Calculation of historical CPI's back to 2011 using the scanner data	1. September 2014
Decision on suitability of CPI/HICP calculated with the use of scanner data	1. September 2014
Final decision on whether scanner data will be implemented in the CPI from January 2015	10. September 2014
Fully implemented system	1. Jan 2015-



More chains?

- The CPI-section has been in contact with Supergros with regard to the possible delivery of scanner data. Since Supergros' shops currently accounts for ca. 20 % of the turnover of food and drinks, it will improve the CPI/HICP, if they are to be included. Though they have generally been positive towards the possibility of delivering data, and have supplied a couple of weeks of test-data, no weekly delivering is in place.
- In 2014 we will try to get a final agreement with Supergros on when to start delivering weekly scanner data to us. In any case it is has been decided that we will not be able to implement the Supergros-data in the CPI from January 2015, but January 2016 could be a target.



Challenges

- Informal agreements with supermarket chains- risk of not getting data?
- Deadline for the decision on when to implement
- Maintaining of IT-systems
- Maintaining of SAS-codes