

Product Data: Request Forms

The product files, known as 'All Articles', contain the product's eancode and all information relevant to that particular product across the years from 2006 - 2020.

These files include information on the product's description, manufacturer details, quantity, private-label information, units, and more. The files are most useful when matched using the eancode to the sales data, where they can give more detailed descriptions of individual product lines. However, it should be noted that the files use real-world data which is messy by nature. Therefore, it can be expected that the files you receive will need to be cleaned to fit your own specific research question.

Additional observations:

1. Some of the variables record data for specific years or only where relevant to the product. These include the alphanumerical "Type" and "text" columns (e.g. IType and Itext).
2. Alphanumeric variables can be converted between filetypes or when loading into programming languages such as R without being called to do so. For this reason, the user should take extra care to ensure they run checks on alphanumeric columns before analysing them. If using in R, it is recommended to load the file using "read.csv()" as opposed to "read_csv()".

The following 30 variables are included within these files:

Column	Column Name (DK)	Description	Datatype
1	EanCodeText	The product's individual eancode: this can be used to match to the 'eancode' column of the Sales files.	Alphanumeric factor
2	ProductGroupCode	The numeric product group code, as described in "Consumer Product Categories.docx". There are 127 product categories in total.	Numeric
3	PgText	The product group name, as given by GfK.	Alphanumeric factor

		These names are in Danish, with the product group code in parenthesis beside (e.g., Pasta- og Risretter (056))	
4	ArticleText	<p>The products description, which varies in level of detail between products.</p> <p>Most descriptions include the products name, brand, and unit quantity (e.g., Betty Crocker Cake Mix 500 gr)</p>	Alphanumeric factor
5	Cmrk	Check with GFK	Numeric
6	ManufacturerName	<p>The higher-level manufacturer name of the product or brand. This can be the same as the BrandText column for independent brands or may vary if the brand belongs to a bigger chain such as with private label brands.</p> <p>(e.g., for product of eancode 5711040000000, BrandText = Elmelund, ManufacturerName = Dansk Supermarked)</p>	Alphanumeric factor
7	BrandText	The product's brand name. This can vary from the manufacturer name, particularly with private labels or product lines which belong to bigger product chains.	Alphanumeric factor
8	DType	<p>Check with GFK</p> <p>Numbers between 0 to 99, where many of the data entried are blank or NAs</p>	Numeric Column
9	Dtext	<p>A supporting variable which gives more information on the products size or sub-category.</p> <p>Check with GFK</p> <p>Looks to be more specific product information which ranges from stating the number of items in the product (e.g., 10 stk) to more specific categories such as 'Breakfast Tea') – a good one to check with GFK</p>	Alphanumeric factor
10	FType	<p>Check with GFK</p> <p>Numbers between 0 and 20, where most of the datafields are blank or NAs</p>	Numeric

11	Ftext	<p>A supporting variable which gives more information on the products size or sub-category.</p> <p>Similar to Dtext – seems to give more information on the units or specific category – not intuitive how it's been coded</p>	Alphanumeric factor
12	GType	Numbers between 0 and 25, where most of the datafields are blank or NA, or 99	Numeric
13	Gtext	<p>A supporting variable which gives more information on the product's nature, whether by specifying the unit, packaging, or product contents.</p> <p><i>(e.g., for the category 'Dishwashing', it specifies whether the product is a tablet or liquid)</i></p>	Alphanumeric factor
14	IType	Numbers between 0 and 20, where most of the entries are blank or NAs	Numeric
15	Itext	A supporting variable which gives more information on the product's pre-packaging and cooking preparation (e.g. oven preparation, microwave, pre-packaged etc)	Alphanumeric factor
16	JType	Numbers between 1 and 9,, where most of the entries are blank or NAs	Numeric
17	Jtext	A supporting variable which gives information on the product's physical properties (e.g., is it a liquid or powder), the material it was made from (e.g., pap/paper) and also sometimes indicates whether the product is organic or not (e.g., Økologisk)	Alphanumeric factor
18	KType	Numbers between 0 and 34, where most of the entries are blank or NAs	Numeric
19	Ktext	<p>A supporting variable which contains more information on the quantity within the product, as well as extra information on drinks categories, food categories, some body products.</p> <p>This column also indicates some products which have been produced in Denmark or not.</p>	Alphanumeric factor
20	LType	Numbers between 1 and 9, where most of the entries are blank or NAs	Numeric
21	Ltext	A supporting variable which contains information on some food products preparation such as grilled or	Alphanumeric factor

		marinated, frosted or not frosted, sliced, natural, and breaded.	
22	MType	Numbers of 1, 2, 8, and 9, where most of the entries are blank or NAs	Numerical
23	Mtext	A supporting variable which indicates whether the product is Danish or not. This column is not well recorded with many NAs so is likely to require extra processing by the user.	Alphanumeric factor
24	BrandTextB	This variable indicates whether the product is a Private Label, Brand, Unknown, or NA.	Alphanumeric factor
25	EcologyText	This variable indicates whether the product is organic or not (categories: Organic, Not Organic, Unknown, NA)	Factor
26	UnitText	This variable indicates the units of the product (e.g., ml, g, pk etc)	Alphanumeric factor
27	EanCodeNABS	A secondary eancode Check with GFK	Alphanumeric factor
28	PL_Brand	This variable indicates which private label brand the product belongs to (e.g., Grøn Balance)	Alphanumeric factor
29	PL_SuperM_Chain	This variable indicates which chain the private label brand belongs to The 6 key Danish supermarket chains include Aldi, Coop, Dagrofa, Lidl, Rema, Salling. There is also some extra information on smaller chains or foreign chains such as Edeka (Swedish) and Rewe (German).	String factor
30	PL_Hierarchy	This variable indicates which private label level the product's brand belongs to. Readers are encouraged to check the private label encoding descriptions if planning to use this variable to know which brands are included in each label. Labels = Premium, Mainstream, Budget	String factor

Data request notes:

When requesting data, please clearly state the product group and product group code for the categories you are interested in. Please also include any extra details you feel may be necessary. The best way to get your data in a timely matter is to be clear about what you would like, including stating the timeframe you are interested in.

The reader is highly encouraged to run data checks on the columns to ensure the datatypes match when loading, in particular ensuring the EanCodeText column has no NAs.