

Features

„items.csv“

Column name	Description	Value range	Missing values
pid	Product number	Natural number	No
manufacturer	Manufacturer	Natural number	No
group	Product group	String consisting of capital letters and numbers	No
content	Package content	Positive floating-point number or string in the format "number X number", e.g. "5X10"	No
unit	Unit	String of capital letters	No

pharmForm	Dosage form	Three-digit string of letters	Yes
genericProduct	Generic flag	{0,1}	No
salesIndex	Dispensing regulation code	Natural number	No
category	Main shop category	Natural number	Yes
campaignIndex	Action label	{A, B, C}	Yes
rrp	Reference price	Positive floating-point number	No

„train.csv“, „class.csv“

Column name	Description	Value range	Missing values
lineID	Key for unique identification of user action	Natural number	No
day	Day in the observed time period	Natural number	No
pid	Product number	Natural number	No
adFlag	Advertising flag indicating if the product is part of an advertising campaign	{1, 0}	No
availability	Availability status	{1, 2, 3, 4}	No
competitorPrice	Lowest competitor's price	Positive floating-point number	Yes

click*	Click flag	{0,1}	No
basket*	Shopping basket flag	{0,1}	No
order*	Purchase flag	{0,1}	No
price	Product price	Positive floating-point number	No
revenue*	Revenue	Positive floating-point number or "0"	No

*) Column missing in the "class.csv" file