

Variable	Data type	Range	Description
ID	numerical	Integer	Shows a unique identifier of a customer.
Day	numerical	Integer	Day when the customer has visited the store
Incidence	categorical	{0,1}	Purchase Incidence The customer has not purchased an item from the category of 0interest The customer has purchased an item from the category of 1interest
Brand	categorical	{0,1,2,3,4,5}	Shows which brand the customer has purchased No brand was 0purchased 1,2,3,4,5Brand ID
Quantity	numerical	integer	Number of items bought by the customer from the product category of interest
Last_Inc_Brand	categorical	{0,1,2,3,4,5}	Shows which brand the customer has purchased on their previous store visit No brand was 0purchased 1,2,3,4,5Brand ID
Last_Inc_Quantity	numerical	integer	Number of items bought by the customer from the product category of interest during their previous store visit
Price_1	numerical	real	Price of an item from Brand 1 on a particular day
Price_2	numerical	real	Price of an item from Brand 2 on a particular day
Price_3	numerical	real	Price of an item from Brand 3 on a particular day
Price_4	numerical	real	Price of an item from Brand 4 on a particular day
Price_5	numerical	real	Price of an item from Brand 5 on a particular day

Promotion_1	categorical	{0,1}	Indicator whether Brand 1 was on promotion or not on a particular day There is no 0promotion There is 1promotion
Promotion_2	categorical	{0,1}	Indicator of whether Brand 2 was on promotion or not on a particular day There is no 0promotion There is 1promotion
Promotion_3	categorical	{0,1}	Indicator of whether Brand 3 was on promotion or not on a particular day There is no 0promotion There is 1promotion
Promotion_4	categorical	{0,1}	Indicator of whether Brand 4 was on promotion or not on a particular day There is no 0promotion There is 1promotion
Promotion_5	categorical	{0,1}	Indicator of whether Brand 5 was on promotion or not on a particular day There is no 0promotion There is 1promotion
Sex	categorical	{0,1}	Biological sex (gender) of a customer. In this dataset there are only 2 different options. 0male 1female
Marital status	categorical	{0,1}	Marital status of a customer. 0single non-single (divorced / separated / married / widowed) 1
Age	numerical	Integer	The age of the customer in years, calculated as current year minus the year of birth of the customer at the time of creation of the dataset Min value (the lowest age observed in the dataset) 18 Max value (the highest age observed in the dataset) 75

Education	categorical	{0,1,2,3}	Level of education of the customer other / 0unknown 1high school 2university graduate 3school
Income	numerical	real	Self-reported annual income in US dollars of the customer. Min value (the lowest income observed in the dataset) 38247 Max value (the highest income observed in the dataset) 309364
Occupation	categorical	{0,1,2}	Category of occupation of the customer. unemployed / 0unskilled skilled employee / 1official management / self-employed / highly qualified employee / 2officer
Settlement size	categorical	{0,1,2}	The size of the city that the customer lives in. 0small city 1mid-sized city 2big city

