

Analysis Report

Prepared for: James Wilson, Advertising Manager, PLO Corporation

Prepared by: Eleutheria Kanavou, Data Engineer, datamine.it

November 15th, 2008

Report number: 000-0002

Executive Summary

Objective

The hereby report summarizes the results of the extended data mining analysis performed for a well-known manufacturer of electrical fittings and switches. The initial data provided regards a market research, which served as the input for a bunch of advanced methodologies and algorithms run to reveal underlying structure and patterns that reside as latent across the data. The paragraphs to follow include, among others, a careful selection of the most significant out of these results, in terms of relevance, consistency and accuracy. The results are presented in a comprehensible and easily digestible format, ready to support decision making processes.

Goals

The analysis performed served a single goal: To extensively study the given data set in order to search for and find out the most important of the rules and patterns hidden within the data. The study, eventually, contributes the shaping of these patterns into usable knowledge, while putting focus on the given variables of specific interest.

Means

The tools and approaches used for extracting the underlying patterns out of the available data set lie in the conjunction of Artificial Intelligence / Machine Learning and Statistics, an area commonly called Data Mining. The datamine.it team leverages on extended research experience on the topic to utilize state-of-the-art tools and techniques and provide you with the most insightful of the results, while yet in an absolutely familiar way.

Outcomes

Among the vast number of results occurred and the most significant out of them to be appeared throughout the report, a sneak peek of the insights gained is provided here:

- Consumers that find endurance as the most important characteristic and have discussed their potential choices with an electrician will tend to select the product.
- People who have a high degree of knowledge about the brand name brand, they appear to prefer it above others.
- Those who don't care about the price and have seen the product in an exhibition are persuaded to buy it.
- Consumers that find endurance as the most important characteristic and have read about the brand in a magazine before tend to choose it.

The totality of contents of this report consist a work and property of datamine.it ltd.

Table of Contents

The context	4
Data, in general	4
Data Mining, in general	4
Data Mine.it, in specific	4
The content	5
Analysis of the data set	5
The analysis	8
Introduction	8
Best rules discovered	8
General outcomes	12
Appendix I: Data set attributes	13
Description of data set attributes	13
Appendix II: Rules discovered	24
List of significant rules discovered	24
Contact Information	26

The context

Data, in general

Data stand as the least biased input to decision making, a pure source of insights and knowledge. And data today is generated, stored and, literally, used in an unprecedented rate. However, time spent on consuming these data remains constant and, what's more, the typical tools to serve this task turn out to be incapable; the resulting 'data gap' is today an omnipresent reality. In this context, common and widely used techniques and approaches, like surveys and the way they are analyzed, or statistical reports, clearly cannot respond efficiently to the hurdles data volume and its in depth analysis pose. If all these leave much to be desired, datamine.it and the on-hand report comes to the rescue, at least for the data in focus.

Data Mining, in general

Where classical approaches prove to be ineffective of the scale, speed and simplicity needed, artificial intelligence comes to join statistics and provide the much needed solution. Data mining that is, and you can visualize it as the way and process of searching for secrets bared in the sand, or drilling for gold in a mine -thus 'mining'-, but in a truly systematic and efficient way. In our case, stone stands for data and gold for the insights and knowledge hidden within the data set, while the single purpose of this report is to provide you with evidence on the existence and the description of this very treasure.

That said, a miner with a mattock in his hand is a very rough way to conceptualize the complexity and state-of-the-art of the processes executed. A diverse and extended set of exploration and filtering algorithms, next to a variety of learning and meta-learning techniques, were utilized, optimized and evaluated, while the problem is a computationally intensive one and demands a highly customized approach.

Data Mine.it, in specific

The paragraphs to follow aim at providing insight on the patterns that emerge from the extended -in both width and depth- data mining analysis of the given data set. A bunch of sophisticated machine learning algorithms were run and fine-tuned by one or more datamine.it engineers to end up on extracting outcomes and patterns that make perfect sense for your dataset and really provide you with insights you never imagined before, or never thought them as being well proven; we like to call it "a tale of discovery, from your data to the report on hand". What's more, rest assured we've worked really hard to separate the wheat from the chaff, all the peculiar terminology included. And if you were used to concern a pie chart or a histogram as the most insightful thing you could expect from a data analysis, get ready to be astonished on the pages to follow.

The content

Analysis of the data set

The initial dataset consisted of 37 attributes (you may visualize it as the number of 'questions performed') and 319 instances (the number of 'samples collected'). The analytical description of attributes is provided in the Appendix I, while Table 1 that follows gives a very sneak peek.

Description	Quantity
attributes	253
nominal	252
numeric	1
target	1
instances	309
missing	0
uniques (on average)	0

Table 1: Data set at a glance

Let's take a deeper view. Table 2 provides the titles of all attributes, which consist the data set. These are referred here to provide you with a broader view of the data in focus that are potentially utilized in the results of the following pages. Again, you may find a more detailed description of the submitted attributes in Appendix I.

#	Name	#	Name	#	Name
1	Age	22	Install dimmer	43	Mostimpchar_quality
2	Gender	23	Install shades	44	Mostimpchar_price
3	Building ownership	24	Install sounds	45	Impchar_security
4	Building type	25	Install homecinema	46	Impchar_reliability
5	Building age	26	Install notistal	47	Impchar_endurance
6	Decision making electrician	27	Reasonnotistal_unnecessary	48	Impchar_functions
7	See products magazines	28	Reasonnotistal_expensive	49	Impchar_design
8	See products living exhibition	29	Howimportantbuy	50	Impchar_quality
9	See products compexhibition	30	Reliability_brand	51	Impchar_colours
10	Knownbrand_brand	31	Reasonbought_brand	52	Impchar_varietydesigns
11	See products samples	32	Rebuy	53	Impchar_price
12	See products nowhere	33	Reasonbuy_brand	54	Impchar_matchhome
13	Discussion electrician	34	Security_brand	55	Impchar_brand
14	Install dimmer	35	Mostimpchar_security	56	Qualitypriceratio_brand

#	Name	#	Name	#	Name
15	Install shades	36	Mostimpchar_reliability	57	Brandknowledge_brand
16	Install sounds	37	Mostimpchar_endurance	58	Quality_brand2
17	Install homecinema	38	Mostimpchar_design	59	Price_brand
18	design_brand	39	Mostimpchar_colour	60	Advertising_magazines
19	Reasonnotistal_unnecessary	40	Functions_brand	61	Advertising_leaflet
20	Reasonnotistal_expensive	41	bestdesign_brand	62	Buyer_electrician
21	Howimportantbuy	42	Endurance_brand	63	Quality_brand

Table 2: Titles of attributes in use

As the target for the analysis performed served the single attribute of 'brandbought_brand' (#253). In other words, the analysis performed attempt to extract relationships and insights of all other attributes in regard to this one. Table 3 provides more details on this attribute, next to the distribution of its values in the given data set in Figure 1. Figures of all the attributes are given in the Appendix I.

#	Name	Type	Values	Missing	Distinct	Unique
253	Brandbought_brand	nominal	0 , 1	0(0%)	2	0(0%)

Table 3: Description of the target attribute

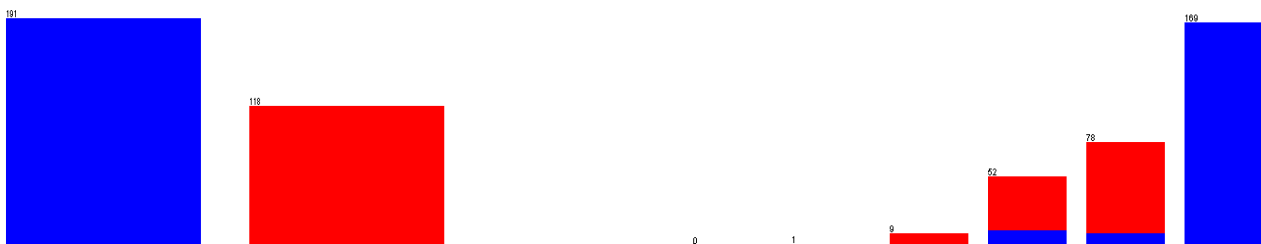


Figure 1: a) Distribution of the target attribute, b) Distribution of attribute 'reliability_brand', in regard to the target attribute

Due to the sample's complexity and size, various advanced filtering techniques were repeatedly utilized to firstly rank these attributes according to their correlation and informational value in regards to the analysis' target, and then put focus on the ones that matter the most. Table 4 presents the 10 most valuable out of these, as occurred by such a process, while Table 5 contributes the ones of least informational value.

#	Name
1	qualitypriceratio_brand
2	price_brand
3	design_brand
4	knownbrand_brand
5	reliability_brand
6	functions_brand
7	bestdesign_brand
8	endurance_brand
9	quality_brand2
10	security_brand

Table 4: Attributes of most informational value

#	Name
1	brandmostinterest_brand3
2	brandmostinterest_brand4
3	impchar_price
4	reasonnotinstal_unnecessary
5	otherchar_endurance
6	seeproducts_livingexhibition
7	brandknowledge_brand5
8	brandbought_nobrand
9	brandbought_brand6
10	brandbought_brand7

Table 5: Attributes of low informational value

Given the rough description of the submitted data set and the analysis framework deployed before, the next paragraph stands as the core of this report, moving to the actual results of the knowledge discovery process.

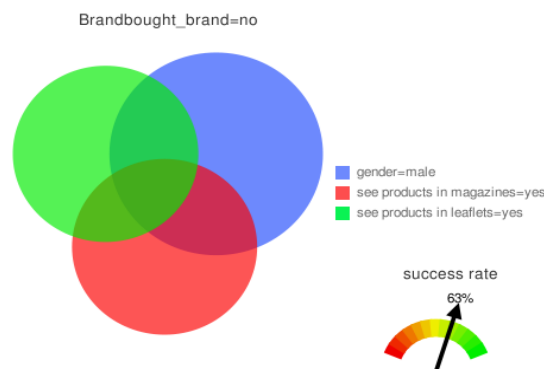
The analysis

Introduction

As referred above, the analysis performed utilized an extended variety of advanced data mining techniques and machine learning algorithms, next to the outcomes of the data set's analysis, to finally extract the best and brightest of its latent patterns. Significant effort was also put into transforming these patterns and analysis results into some direct, tangible and easily comprehensible outcomes.

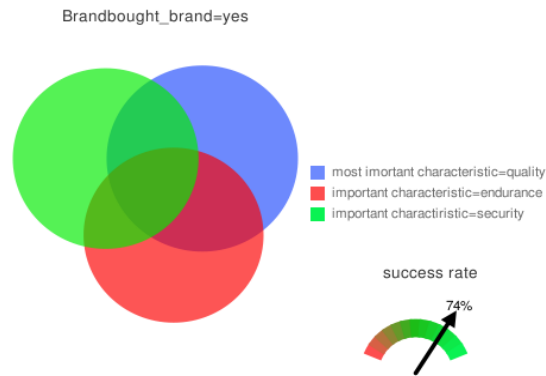
Best rules discovered

The pages to follow describe in words and figures the most significant out of the rules discovered, in other words the most distinguishable of the patterns emerged out of the extensive mining processes performed. Each pattern is also described by the number of cases that validates it across the data set, as well as its success rate. Apart from the rules presented here, Appendix II provides an extended list of (less or more) significant rules discovered, essentially contributing to the formation and understanding of the latent knowledge in the given data set.



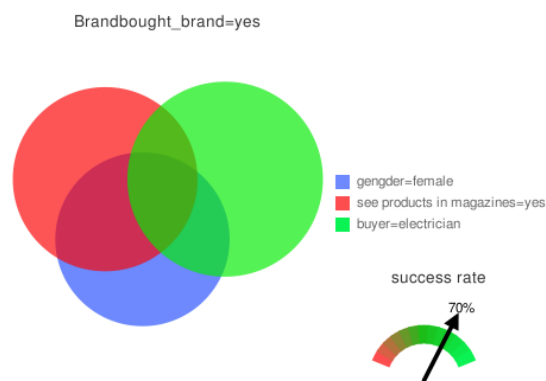
Rule 1: if gender=male & see products in magazines=yes & see products in leaflets=yes then brandbought_brand=no (63% success)

Rule 1 indicates that males that have seen the product in magazines or leaflets don't buy the product with a certainty 63%.



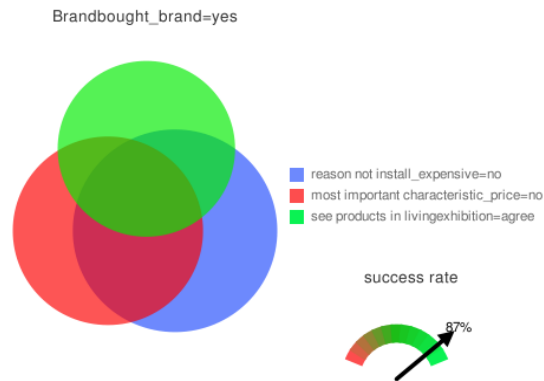
Rule 2: if most important char_quality=yes & important char_endurance=yes & important char_security=yes then brandbought_brand= yes (74% success)

Rule 2 suggests, with a certainty of 74%, that those who find quality, endurance and security as important criteria for their decision will prefer the specific brand.



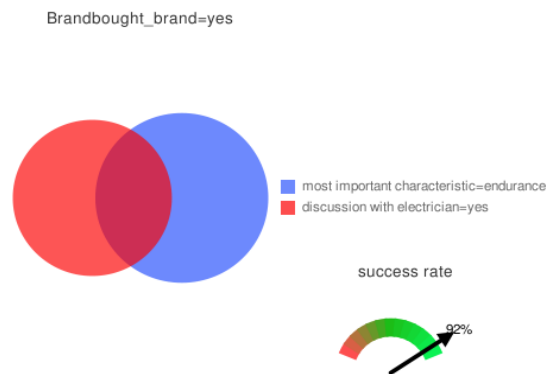
Rule 3: if gender= female & see products in magazine=yes & buyer electrician=yes then brandbought_brand=yes (70% success)

This rule provides the insight that a female consumer who is a relative reader of magazines and assigns the purchase decision to an electrician, is expected to buy the brand.



Rule 4: if reason not install_expensive =no & most important char_price=np & see product living exhibition =agree then brandbought_brand = yes (87% success).

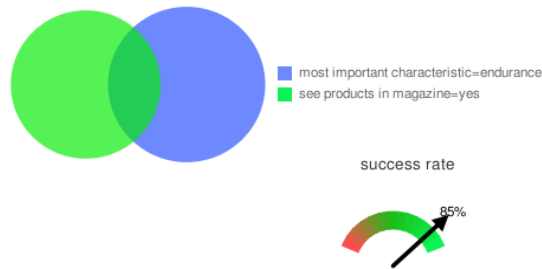
The pattern emerging from this rule indicates that if cost is not a reason to avert an electrical fitting of this brand, while price remains not an important characteristic and live exhibitions are a medium for choice, then the brand will be preferred, and this is something that comes with an 87% rate of success.



Rule 5: If most important char_endurance= yes & discussion electrician = yes then brandbought_brand= yes (92% success)

Rule 5 reveals that an individual who happens to find endurance extremely considerable and resides on electrician's opinion appears to buy the brand. The rule is supported by the given data set at a 92% rate of success.

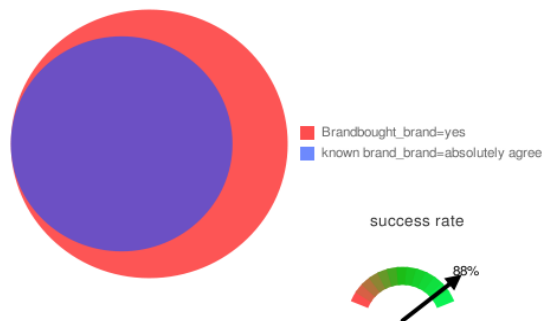
Brandbought_brand=yes



Rule 6: if most important char_ endurance= yes & advertising in magazines = yes then brandbought_brand= yes (85% success)

This rule indicates that a consumer who happens to find endurance as the most important characteristic and reads advertisements in magazines will choose the specific brand, with a certainty of 85%.

Brandbought_brand=yes



Rule 7: if known brand_brand= absolutely agree then brandbought_brand=yes (88% success)

This rule suggests that people, who are absolutely informed about the specific brand name, will tend to prefer it ad this stands as a core outcome of this study.

Again, the rules demonstrated here are a small part from the best of the rules found, while a much more extended set of them can be found at Appendix II.

General outcomes

The extended analysis performed and the numbers of results presented in the previous pages, as long as in the Appendix II, clearly shaped out a number of outcomes, the most significant of which are also deployed hereby:

- If quality, endurance and security are the decision's criteria, the consumer will select the brand
- Male consumers, who consider endurance as important and reside on an electrician's view, are expected to select the brand. Same for female consumers who read magazines.
- Whoever is informed about the brand, will tend to select it when it comes to purchase such equipment. The brand's value is strong.

While the results found are presented at full extent in the Appendixes below (including the attributes analytical description and plots, most valuable -information wise- attributes and a really big list of rules extracted), it is by now clear that the on hand analysis has contributed deep insights, yet simple descriptions, on the patterns and knowledge that were lying unveiled through the submitted data set. This tale of discovery, from your data to the report on hand, seemed to reach its end, at least on the part of maximizing the value of your data input. We do believe you'll come to validate this, while we continuously remain at your request for shaping the next episode of your data tales.

Appendix I: Data set attributes

Description of data set attributes

The list of attributes of the given data set is provided here.

#	Name	Type	Values	Missing	Distinct	Unique
1	questionID	numeric	-	0	-	309
2	town	nominal	{capital, suburbs, other_town, area4, area5, 0}	0	6	0
3	age	nominal	{age#1, age#2, age#3, age#4, age#5, 0}	0	6	0
4	gender	nominal	{male, female}	0	3	0
5	who_chose	nominal	{same, other_house_member, no_house_member, 0}	0	4	0
6	start_search	nominal	{less_than_1_year, 1_year_ago, 2_years_ago, 3_years_ago, 0}	0	5	0
7	building_ownership	nominal	{privately_owned, rented, DA, 0}	0	4	1
8	building_type	nominal	{permanent, rural, for_rent, second_house, store, for_sale, 0}	0	7	1
9	building_age	nominal	{new, renovation, 0}	0	3	0
10	building_charact	nominal	{apartment, 2_floors, 3+floors, private, duplex_house, maisonette, store, 0}	0	8	0
11	decisionmaking_partner	nominal	{0,1}	0	2	0
12	decisionmaking_kids	nominal	{0,1}	0	2	0
13	decisionmaking_parents	nominal	{0,1}	0	2	0
14	decisionmaking_relatives_friends	nominal	{0,1}	0	2	0
15	decisionmaking_electrician	nominal	{0,1}	0	2	0
16	decisionmaking_none	nominal	{0,1}	0	2	0
17	decisionmakingsup_engineer	nominal	{0,1}	0	2	0
18	decisionmakingsup_electrician	nominal	{0,1}	0	2	0
19	decisionmakingsup_elstore	nominal	{0,1}	0	2	0
20	decisionmakingsup_friends	nominal	{0,1}	0	2	0
21	seeproducts_magazines	nominal	{0,1}	0	2	0
22	seeproducts_livingexhibition	nominal	{0,1}	0	2	0
23	seeproducts_compexhibition	nominal	{0,1}	0	2	0
24	seeproducts_elstore	nominal	{0,1}	0	2	0
25	seeproducts_samples	nominal	{0,1}	0	2	0
26	seeproducts_nowhere	nominal	{0,1}	0	2	0
27	buyproducts_atonce	nominal	{0,1}	0	2	0
28	buyproducts_home	nominal	{0,1}	0	2	0
29	buyproducts_moreresearch	nominal	{0,1}	0	2	0
30	buyproducts_noproducts	nominal	{0,1}	0	2	0

#	Name	Type	Values	Missing	Distinct	Unique
31	whatstage_startsearch	nominal	{sta_arxika_sxedia, sthn_arxikh_pr osfo- ra, htan_ka8orismena, den_asxolh8 hka, allo, DA, 0}	0	7	0
32	discussion_engineer	nominal	{0,1}	0	2	0
33	discussion_electrician	nominal	{0,1}	0	2	0
34	discussion_friends	nominal	{0,1}	0	2	0
35	discussion_27	nominal	{0,1}	0	2	0
36	placedisc_engineer	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	2
37	placedisc_architect	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	1
38	placedisc_builder	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	0
39	placedisc_contractor	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	0
40	placedisc_electrician	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	0
41	placedisc_friend	nominal	{at_his_office, at_my_place, building, at_work, DA, 0}	0	6	0
42	mecaknowledge_dimmer	nominal	{0,1}	0	2	0
43	mecaknowledge_shades	nominal	{0,1}	0	2	0
44	mecaknowledge_sound	nominal	{0,1}	0	2	0
45	mecaknowledge_homecinema	nominal	{0,1}	0	2	0
46	mecaknowledge_notinstal	nominal	{0,1}	0	2	0
47	infomecas_engineer	nominal	{0,1}	0	2	0
48	infomecas_electrician	nominal	{0,1}	0	2	0
49	infomecas_elstore	nominal	{0,1}	0	2	0
50	infomecas_friends	nominal	{0,1}	0	2	0
51	infomecas_magazines	nominal	{0,1}	0	2	0
52	infomecas_livingexhibition	nominal	{0,1}	0	2	0
53	infomecas_31	nominal	{0,1}	0	2	0
54	instal_dimmer	nominal	{0,1}	0	2	0
55	instal_shades	nominal	{0,1}	0	2	0
56	instal_sound	nominal	{0,1}	0	2	0
57	instal_homecinema	nominal	{0,1}	0	2	0
58	instal_notinstal	nominal	{0,1}	0	2	0
59	reasonnotinstal_unnecessary	nominal	{0,1}	0	2	0
60	reasonnotinstal_expensive	nominal	{0,1}	0	2	0
61	howimportantbuy	nominal	{not_important, not_much_important, fair_important, enough_important, very_important, DA, 0}	0	7	0
62	mostimpchar_security	nominal	{0,1}	0	2	0
63	mostimpchar_reliability	nominal	{0,1}	0	2	0
64	mostimpchar_endurance	nominal	{0,1}	0	2	0
65	mostimpchar_design	nominal	{0,1}	0	2	0
66	mostimpchar_colour	nominal	{0,1}	0	2	0
67	mostimpchar_quality	nominal	{0,1}	0	2	0
68	mostimpchar_price	nominal	{0,1}	0	2	0
69	impchar_security	nominal	{0,1}	0	2	0

#	Name	Type	Values	Missing	Distinct	Unique
70	impchar_reliability	nominal	{0,1}	0	2	0
71	impchar_endurance	nominal	{0,1}	0	2	0
72	impchar_functions	nominal	{0,1}	0	2	0
73	impchar_design	nominal	{0,1}	0	2	0
74	impchar_colours	nominal	{0,1}	0	2	0
75	impchar_varietydesigns	nominal	{0,1}	0	2	0
76	impchar_quality	nominal	{0,1}	0	2	0
77	impchar_price	nominal	{0,1}	0	2	0
78	impchar_matchhome	nominal	{0,1}	0	2	0
79	impchar_brand	nominal	{0,1}	0	2	0
80	impchar_qualitypriceratio	nominal	{0,1}	0	2	0
81	impchar_dontknow	nominal	{0,1}	0	2	0
82	otherchar_security	nominal	{0,1}	0	2	0
83	otherchar_reliability	nominal	{0,1}	0	2	0
84	otherchar_endurance	nominal	{0,1}	0	2	0
85	otherchar_functions	nominal	{0,1}	0	2	0
86	otherchar_design	nominal	{0,1}	0	2	0
87	otherchar_quality	nominal	{0,1}	0	2	0
88	otherchar_price	nominal	{0,1}	0	2	0
89	otherchar_dontknow	nominal	{0,1}	0	2	0
90	1stbrand_BRAND 1	nominal	{0,1}	0	2	0
91	1stbrand_BRAND	nominal	{0,1}	0	2	0
92	1stbrand_BRAND 8	nominal	{0,1}	0	2	0
93	1stbrand_othbrand	nominal	{0,1}	0	2	0
94	1stbrand_nobrand	nominal	{0,1}	0	2	0
95	1stbrand_dontknow	nominal	{0,1}	0	2	0
96	othbrand_BRAND 1	nominal	{0,1}	0	2	0
97	othbrand_BERKER	nominal	{0,1}	0	2	0
98	othbrand_BRAND 2	nominal	{0,1}	0	2	0
99	othbrand_BRAND	nominal	{0,1}	0	2	0
100	othbrand_BRAND 8	nominal	{0,1}	0	2	0
101	othbrand_othbrand	nominal	{0,1}	0	2	0
102	othbrand_none	nominal	{0,1}	0	2	0
103	othbrand_dontknow	nominal	{0,1}	0	2	0
104	brandknowledge_BRAND 3	nominal	{0,1}	0	2	0
105	brandknowledge_BRAND 4	nominal	{0,1}	0	2	0
106	brandknowledge_BRAND 2	nominal	{0,1}	0	2	0
107	brandknowledge_BRAND 6	nominal	{0,1}	0	2	0
108	brandknowledge_BRAND 7	nominal	{0,1}	0	2	0
109	brandknowledge_BRAND	nominal	{0,1}	0	2	0
110	brandknowledge_BRAND 5	nominal	{0,1}	0	2	0
111	brandknowledge_BRAND 8	nominal	{0,1}	0	2	0
112	brandknowledge_BRAND 9	nominal	{0,1}	0	2	0
113	brandknowledge_BRAND 5	nominal	{0,1}	0	2	0
114	brandknowledge_none	nominal	{0,1}	0	2	0
115	brandmostinterest_BRAND 3	nominal	{0,1}	0	2	0
116	brandmostinterest_BRAND 4	nominal	{0,1}	0	2	0
117	brandmostinterest_BRAND 2	nominal	{0,1}	0	2	0
118	brandmostinterest_BRAND	nominal	{0,1}	0	2	0
119	brandmostinterest_BRAND 8	nominal	{0,1}	0	2	0

#	Name	Type	Values	Missing	Distinct	Unique
120	brandmostinterest_BRAND 9	nominal	{0,1}	0	2	0
121	brandmostinterest_othbrand	nominal	{0,1}	0	2	0
122	brandmostinterest_none	nominal	{0,1}	0	2	0
123	brandmostinterest_dontknow	nominal	{0,1}	0	2	0
124	reasonbought_BRAND 3	nominal	{nice, modern, colours, 0}	0	4	1
125	reasonbought_BRAND 4	nominal	{high_quality, 0}	0	2	0
126	reasonbought_BRAND 2	nominal	{good_price, discount, price_vs_quality, price_vs_design, 0}	0	5	1
127	reasonbought_BRAND 6	nominal	{high_secure,0}	0	2	0
128	reasonbought_BRAND 7	nominal	{known_brand, good_brand, trusted_brand, reliable_brand, eponymous_brand, 0}	0	6	0
129	reasonbought_BRAND	nominal	{high_compact, 0}	0	2	0
130	reasonbought_BRAND 8	nominal	{good_reputation, told_to_be_better, heard_good_opinions, recommended,0}	0	5	1
131	reasonbought_BRAND 9	nominal	{very_practical, good_operation, good_technique, easy_setting, appearing_no_problems,0}	0	6	1
132	reasonbought_nobrand	nominal	{satisfied, fitted_in_the_house, same_brand_before, husband_choice, this_brand_already, replace_older, electrician_decision,0}	0	8	2
133	reasondislikebrand_BRAND 3	nominal	{not_my_taste, very_simple_design,very_raw_design, no_modern_design, old_fashioned, large_fittings,0}	0	7	3
134	reasondislikebrand_BRAND 4	nominal	{not_good_quality, easily_damaged, colour_fade,colour_darken,very_sensitive, not_reliable,0}	0	7	0
135	reasondislikebrand_BRAND 2	nominal	{very_expensive,0}	0	2	0
136	reasondislikebrand_BRAND 6	nominal	{not_resistant, low_endurance, not_safe,0}	0	4	2
137	reasondislikebrand_BRAND 5	nominal	{not_easy_used_by_children, not_steady, large_installation_place, easy_detached, wires_malfunction,0}	0	6	2
138	reasondislikebrand_nobrand	nominal	{_997,0}	0	2	0
139	brandranking	numeric	-	0	-	2
140	rebuy	nominal	{buy_again,buy_other_brand, don't_know, 0}	0	4	0
141	reasonrebuy_BRAND 3	nominal	{nice, modern, colours, 0}	0	4	2
142	reasonrebuy_BRAND 4	nominal	{high_quality, 0}	0	2	0
143	reasonrebuy_BRAND 2	nominal	{good_price, discount, price_vs_quality, price_vs_design, 0}	0	5	1
144	reasonrebuy_BRAND 6	nominal	{high_secure,0}	0	2	0

#	Name	Type	Values	Missing	Distinct	Unique
145	reasonrebuy_BRAND 7	nominal	{known_brand, good_brand, trusted_brand, reliable_brand, eponymous_brand, 0}	0	6	2
146	reasonrebuy_BRAND	nominal	{high_compact, 0}	0	2	0
147	reasonrebuy_BRAND 9	nominal	{very_practical, good_operation, good_technique, easy_setting, appearing_no_problems, 0}	0	6	2
148	reasonrebuy_nobrand	nominal	{satisfied, fitted_in_the_house, same_brand_before, husband_choice, this_brand_already, replace_older, electrician_decision, 0}	0	8	0
149	reasonrebuy_othbrand	nominal	{{not_my_taste, very_simple_design, very_raw_design, no_modern_design, old_fashioned, large_fittings, 0}}	0	7	1
150	reasonrebuy_none	nominal	{not_good_quality, easily_damaged, colour_fade, colour_darken, very_sensitive, not_reliable, 0}	0	7	2
151	reasonrebuy_18	nominal	{not_easy_used_by_children, not_steady, large_installation_place, easy_detached, wires_malfunction, 0}	0	6	0
152	reasonrebuy_21	nominal	{no_dimar, cant_find_the_brand, for_testing, too_new, 0}	0	5	2
153	security_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
154	functions_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
155	design_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
156	endurance_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
157	reliability_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
158	bestdesign_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
159	knownbrand_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
160	qualitypriceratio_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
161	quality_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
162	price_BRAND 3	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
163	security_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
164	functions_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
165	design_BRAND 4	nominal	{strongly disagree, disagree, don't	0	6	1

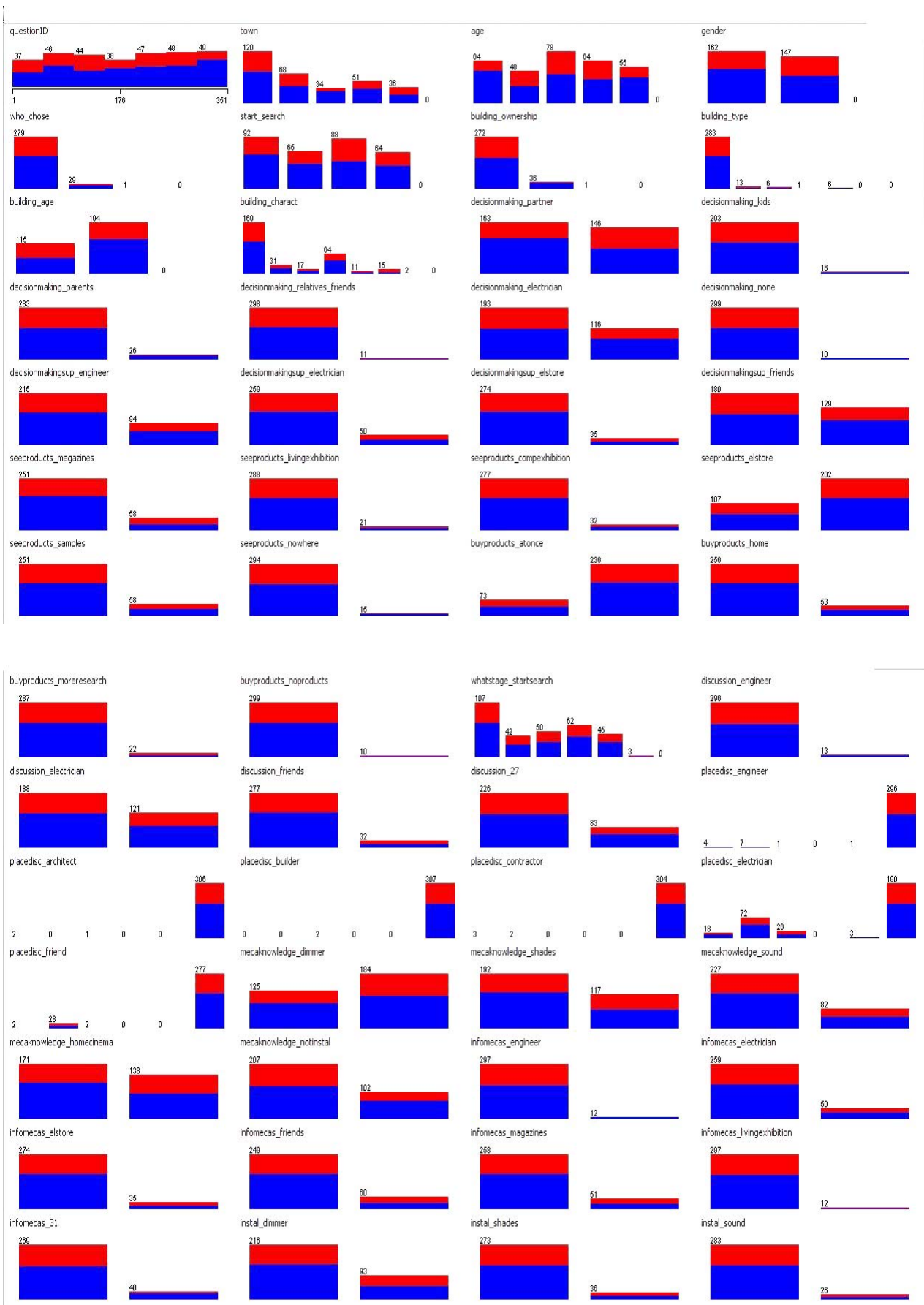
#	Name	Type	Values	Missing	Distinct	Unique
			know, agree, strongly agree, 0}			
166	endurance_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
167	reliability_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
168	bestdesign_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
169	knownbrand_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
170	qualitypriceratio_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
171	quality_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
172	price_BRAND 4	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
173	security_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	3
174	functions_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
175	design_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
176	endurance_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
177	reliability_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
178	bestdesign_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
179	knownbrand_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
180	qualitypriceratio_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
181	quality_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
182	price_BRAND 2	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
183	security_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
184	functions_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
185	design_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	3
186	endurance_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	3
187	reliability_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
188	bestdesign_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
189	knownbrand_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
190	qualitypriceratio_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
191	quality_BRAND 2	nominal	{strongly disagree, disagree, don't	0	6	0

#	Name	Type	Values	Missing	Distinct	Unique
192	price_BRAND	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
193	security_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
194	functions_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
195	design_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
196	endurance_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
197	reliability_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
198	bestdesign_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
199	knownbrand_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
200	qualitypriceratio_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
201	quality_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
202	price_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
203	security_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
204	functions_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	3
205	design_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
206	endurance_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
207	reliability_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
208	bestdesign_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
209	knownbrand_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
210	qualitypriceratio_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
211	quality_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
212	price_XXX	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
213	security_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	3
214	functions_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
215	design_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
216	endurance_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
217	reliability_BRAND 9	nominal	{strongly disagree, disagree, don't	0	6	1

#	Name	Type	Values	Missing	Distinct	Unique
			know, agree, strongly agree, 0}			
218	bestdesign_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
219	knownbrand_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
220	qualitypriceratio_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
221	quality_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
222	price_BRAND 9	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
223	security_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	2
224	functions_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
225	design_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
226	endurance_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
227	reliability_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
228	bestdesign_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
229	knownbrand_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
230	qualitypriceratio_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	0
231	quality_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
232	price_BRAND 5	nominal	{strongly disagree, disagree, don't know, agree, strongly agree, 0}	0	6	1
233	advertising	nominal	{yes, no, 0}	0	3	1
234	advertising_magazines	nominal	{0,1}	0	2	1
235	advertising_leaflet	nominal	{0,1}	0	2	0
236	advertising_08	nominal	{0,1}	0	2	0
237	advertising_14	nominal	{0,1}	0	2	0
238	y6i_01	numeric	-	0	-	0
239	electricboard	nominal	{yes_me, yes_elec/manuf/mech, no, DA, 0}	0	5	0
240	brandbought_BRAND 3	nominal	{0,1}	0	2	0
241	brandbought_BRAND 4	nominal	{0,1}	0	2	0
242	brandbought_BRAND 2	nominal	{0,1}	0	2	1
243	brandbought_BRAND 6	nominal	{0,1}	0	2	1
244	brandbought_BRAND 7	nominal	{0,1}	0	2	0
245	brandbought_BRAND 5	nominal	{0,1}	0	2	0
246	brandbought_BRAND 8	nominal	{0,1}	0	2	1
247	brandbought_BRAND 9	nominal	{0,1}	0	2	0
248	brandbought_BRAND 5	nominal	{0,1}	0	2	0
249	brandbought_nobrand	nominal	{0,1}	0	2	0
250	brandbought_othbrand	nominal	{0,1}	0	2	1
251	brandbought_none	nominal	{0,1}	0	2	0

#	Name	Type	Values	Missing	Distinct	Unique
252	brandbought_dontknow	nominal	{0,1}	0	2	0
253	brandbought_BRAND	nominal	{0,1}	0	2	1

Table x: Analytical description of data set attributes



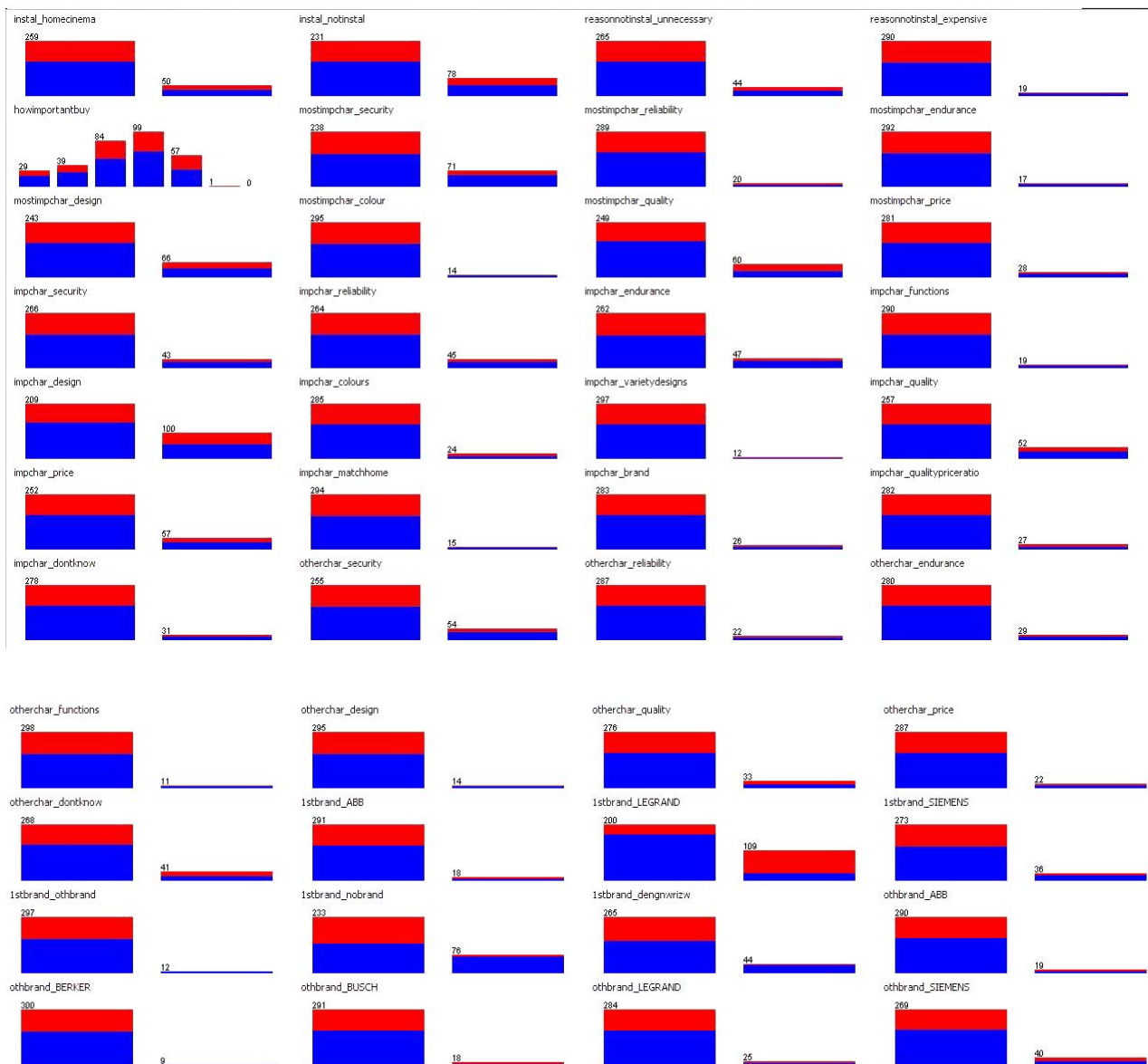


Figure x: Visualization of the data set's distribution, according to variable 'brandbought_brand'

Appendix II: Rules discovered

List of significant rules discovered

Apart from the most significant rules that were referred to in the analysis section and out of the huge bulk of rules that were found during the study of the given data set, a number of other rules are definitely worth or mentioning. These are referred to in the Table XX that follows.

#	Rule
1	if gender=male & see products in magazines=yes & see products in leaflets=yes then brandbought_brand=no (13.0/6.0)
2	if most important char_quality=yes & important char_endurance=yes & important char_security=yes then brandbought_brand= yes (14.0/2.0)
3	if gender= female & see products in magazine=yes & buyer electrician=yes then brandbought_brand=yes (17.0/3.0)
4	if reason not install_expensive =no & most important char_price=np & see product living exhibition =agree then brandbought_brand = yes (43.0/6.0).
5	If most important char_endurance= yes & discussion electrician = yes then brandbought_brand= yes (46.0/4.0)
6	if most important char_endurance= yes & advertising in magazines = yes then brandbought_brand= yes (28.0/5.0)
7	if known brand_brand= absolutely agree then brandbought_brand=yes (86.0/12.0)
8	If brandbought_dontknow = 0 AND brandbought_BRAND 8 = 0 AND qualitypriceratio_BRAND 2 = 0 AND price_BRAND 3 = 0 AND building_ownership = privately_owned:then brandbought_brand = yes (99.0/2.0)
9	If advertising_magazines = 0 AND who_chose = same AND infomecas_friends = 0 AND reasonrebuy_BRAND = 0 AND otherchar_dontknow = 0 AND brandknowledge_BRAND 3 = 0 then brandbought_brand=no (13.0)
10	If brandbought_dontknow = 0 AND knownbrand_XXX = 0 AND instal_homecinema = 0 then brandbought_brand = yes (16.0/1.0)
11	If othbrand_BRAND 8 = 0 AND infomecas_magazines = 0 then brandbought_brand = yes (5.0)
12	If brandmostinterest_BRAND = 1 then brandbought_brand=1 (140.0/22.0)
13	If reasonnotinstal_expensive = 0 AND impchar_functions = 0 AND mostimpchar_price = 0 AND reasonbought_BRAND = 0 AND instal_homecinema = 0 AND advertising = no then brandbought_brand = yes (16.0/2.0)
14	If mostimpchar_endurance = 0 AND advertising_magazines = 0 AND brandknowledge_BRAND = 0 AND gender = female then brandbought_brand = yes (28.0)
15	If reasonbought_BRAND = 0 AND discussion_electrician = 1 then brandbought_brand = no (8.0)
16	If building_ownership = privately_owned AND building_type = permanent_house AND gender = male AND building_charact = apartment then brandbought_brand = no (71.0/27.0)
17	If building_type = permanent_house AND building_charact = apartment then brandbought_brand = no (71.0/33.0)
18	If building_ownership = privately_owned AND building_charact = detached_house then brandbought_brand = no (59.0/19.0)
19	If building_type = permanent_house AND building_charact = 2_floors AND building_age = renovation then brandbought_brand = no (12.0/3.0)

#	Rule
20	If infomecas_engineer = 0 AND infomecas_31 = 1 then brandbought_brand = no (40.0/11.0)
21	If brandbought_dontknow = 0 AND brandbought_BRAND 3 = 0 AND brandmostinterest_BRAND 2 = 0 then brandbought_brand = yes (122.0/6.0)
22	If brandbought_dontknow = 0 AND brandbought_BRAND 3 = 0 AND brandbought_BRAND 8 = 0 AND brandbought_othbrand = 0 then brandbought_brand = yes (123.0/5.0)
23	If decisionmaking_parents = 0 AND decisionmaking_relatives_friends = 0 AND decisionmaking_partner = 1 then brandbought_brand=yes (32.0/12.0)
24	If qualitypriceratio_BRAND = total_agree AND gender = female then brandbought_brand=yes (30.0)
25	If advertising = yes AND building_type = permanent_house then brandbought_brand=yes (38.0/14.0)
26	If seeproducts_magazines = 1 and discussion_electrician = 0 then brandbought_BRAND=yes (33.0/15.0)
27	If building_type = permanent_house AND decisionmaking_electrician = 1 then brandbought_brand=no (61.0/21.0)
28	If advertising_leaflet = 0 AND building_type = permanent_house AND age = age#4 then brandbought_brand=no (30.0/10.0)
29	If placedisc_engineer = 0 AND mostimpchar_endurance = 0 AND brandknowledge_BRAND = 0 AND gender = female then brandbought_brand=yes (47.0)
30	If decisionmakingsup_electrician = 0 AND discussion_27 = 0 AND placedisc_friend = 0 then brandbought_brand = no (44.0/12.0)
31	If building_ownership = privately_owned AND building_age = renovation AND building_charact = detached_house then brandbought_brand = no (36.0/7.0)

Table xx: Extended list of significant rules discovered

Contact Information

This report was prepared by Eletheria Kanavou, data engineer. You may contact her directly at eletheria@datamine.it.

This report was prepared for James Wilson, Advertising Manager, PLO Corporation.

datamine.it

14 Meletiou Vasileiou Str

11 745 Athens, Greece

T +30 6937 122 065

go@datamine.it

<http://datamine.it>

This report remains the property of datamine.it and its content and format are for the exclusive use of the PLO Corporation.