
Contents

	Acknowledgments	xvii
	Introduction	xix
Chapter 1	Cases Queries	1
	Examining Source Data	2
	Flattened Nested Case Table	3
	Specific Source Columns	4
	Examining Training Data	5
	Examining Specific Cases	6
	Examining Test Cases	7
	Examining Model Cases Only	8
	Examining Another Model	9
	Expanding the Nested Table	10
	Sorting Cases	11
	Model and Structure Columns	12
	Specific Model Columns	13
	Distinct Column Values 1/2	13
	Distinct Column Values 2/2	14
	Cases by Cluster 1/4	15
	Cases by Cluster 2/4	16
	Cases by Cluster 3/4	17
	Cases by Cluster 4/4	18
	Content Query	18
	Decision Tree Cases	19
	Decision Tree Content	20
	Time Series Cases	21
	Sequence Clustering Cases 1/2	22
	Sequence Clustering Cases 2/2	23
	Neural Network and Naïve Bayes Cases	24
	Order By with Top	25
	Sequence Clustering Nodes 1/2	26
	Sequence Clustering Nodes 2/2	27

Chapter 2	Content Queries	29
	Content Query	30
	Updating Cluster Captions	31
	Content with New Caption	31
	Changing Caption Back	32
	Content Columns	33
	Node Type	34
	Flattened Content	34
	Flattened Content with Subquery	35
	Subquery Columns	36
	Subquery Column Aliases	37
	Subquery Where Clause	38
	Individual Cluster Analysis	39
	Demographic Analysis	40
	Renaming Clusters	41
	Querying Renamed Clusters	42
	Clusters with Predictable Columns	43
	Narrowing Down Content	43
	Flattening Content Again	44
	Some Tidying Up	45
	More Tidying Up	46
	Looking at Bike Buyers	47
	Who Are the Best Customers?	48
	How Did All Customers Do?	49
	Decision Tree Content	49
	Decision Tree Node Types	50
	Decision Tree Content Columns	51
	Flattened Column	52
	Honing the Result	53
	Just the Bike Buyers	54
	Tidying Up	54
	VBA in DMX	55
	Association Content	56
	Market Basket Analysis	57
	Naïve Bayes Content	58
	Naïve Bayes Node Type	59

	Flattening Naïve Bayes Content	60
	Naïve Bayes Content Subquery 1/2	61
	Naïve Bayes Content Subquery 2/2	62
Chapter 3	Prediction Queries with Decision Trees	65
	Select on Mining Model 1/6	66
	Select on Mining Model 2/6	67
	Select on Mining Model 3/6	67
	Select on Mining Model 4/6	68
	Select on Mining Model 5/6	68
	Select on Mining Model 6/6	69
	Prediction Query	70
	Aliases and Formatting	72
	Natural Prediction Join	73
	More Demographics	74
	Natural Prediction Join Broken	76
	Natural Prediction Join Fixed	77
	Nonmodel Columns	78
	Ranking Probabilities	79
	Predicted Versus Actual	80
	Bike Buyers Only	81
	More Demographics	82
	Choosing Inputs 1/3	84
	Choosing Inputs 2/3	84
	Choosing Inputs 3/3	85
	All Inputs and All Customers	86
	Singletons 1/6	87
	Singletons 2/6	88
	Singletons 3/6	88
	Singletons 4/6	89
	Singletons 5/6	90
	Singletons 6/6	91
	New Customers	92
	New Bike-Buying Customers	93
	A Cosmetic Touch	94
	PredictHistogram() 1/2	95
	PredictHistogram() 2/2	96

Chapter 4	Prediction Queries with Time Series	99
	Analyzing All Existing Sales	100
	Analyzing Existing Sales by Category	101
	Analyzing Existing Sales by Specific Periods—Lag() 1/3	102
	Analyzing Existing Sales by Specific Periods—Lag() 2/3	103
	Analyzing Existing Sales by Specific Periods—Lag() 3/3	103
	PredictTimeSeries() 1/11	104
	PredictTimeSeries() 2/11	105
	PredictTimeSeries() 3/11	106
	PredictTimeSeries() 4/11	106
	PredictTimeSeries() 5/11	107
	PredictTimeSeries() 6/11	108
	PredictTimeSeries() 7/11	108
	PredictTimeSeries() 8/11	109
	PredictTimeSeries() 9/11	110
	PredictTimeSeries() 10/11	110
	PredictTimeSeries() 11/11	111
	PredictStDev()	112
	What-If 1/3	113
	What-If 2/3	114
	What-If 3/3	115
Chapter 5	Prediction and Cluster Queries with Clustering	117
	Cluster Membership 1/3	118
	Cluster Membership 2/3	119
	Cluster Membership 3/3	119
	ClusterProbability() 1/2	120
	ClusterProbability() 2/2	121
	Clustering Parameters	121
	Another ClusterProbability	122
	Cluster Content 1/2	123
	Cluster Content 2/2	123
	PredictCaseLikelihood() 1/3	124
	PredictCaseLikelihood() 2/3	125
	PredictCaseLikelihood() 3/3	125
	Anomaly Detection	126
	Cluster with Predictable Column 1/3	127
	Cluster with Predictable Column 2/3	127

	Cluster with Predictable Column 3/3	128
	Clusters and Predictions	129
Chapter 6	Prediction Queries with Association and Sequence Clustering	131
	Association Content—Item Sets	132
	Association Content—Rules	133
	Important Rules	134
	Twenty Most Important Rules	135
	Particular Product Models	136
	Another Product Model	137
	Nested Table	137
	PredictAssociation()	138
	Cross-Selling Prediction 1/7	139
	Cross-Selling Prediction 2/7	140
	Cross-Selling Prediction 3/7	140
	Cross-Selling Prediction 4/7	141
	Cross-Selling Prediction 5/7	142
	Cross-Selling Prediction 6/7	143
	Cross-Selling Prediction 7/7	143
	Sequence Clustering Prediction 1/3	144
	Sequence Clustering Prediction 2/3	145
	Sequence Clustering Prediction 3/3	146
Chapter 7	Data Definition Language (DDL) Queries	149
	Creating a Mining Structure	150
	Creating a Mining Model	152
	Training a Mining Model	153
	Structure Cases	155
	Model Cases	155
	Model Content	156
	Model Predict	157
	Specifying Structure Holdout	159
	Specifying Model Parameter	160
	Specifying Model Filter	161
	Specifying Model Drill-through	162
	Training the New Models	163
	Cases—with No Drill-through	164
	Cases—with Drill-through	164
	Structure with Holdout	165

Specifying Model Parameter, Filter, and Drill-through	166
Training New Model	167
Unprocessing a Structure	168
Model Cases with Filter and Drill-through	169
Clearing Out Cases	169
Removing Models	170
Removing Structures	170
Renaming a Model	171
Renaming a Structure	172
Making Backups	172
Removing the Backed-up Structure	173
Restoring a Backup	173
Structure with Nested Case Table	174
Model Using Nested Case Table	175
Model Training with Nested Case Table	176
Prediction Queries with Nested Cases 1/2	177
Prediction Queries with Nested Cases 2/2	178
Cube—Mining Structure	179
Cube—Mining Model	180
Cube—Model Training	181
Cube—Structure Cases	182
Cube—Model Content	183
Cube—Model Prediction	184
Chapter 8 Schema and Column Queries	187
DMSHEMA_MINING_SERVICES 1/2	188
DMSHEMA_MINING_SERVICES 2/2	189
DMSHEMA_MINING_SERVICE_PARAMETERS 1/2	189
DMSHEMA_MINING_SERVICE_PARAMETERS 2/2	190
DMSHEMA_MINING_MODELS 1/3	191
DMSHEMA_MINING_MODELS 2/3	192
DMSHEMA_MINING_MODELS 3/3	192
DMSHEMA_MINING_COLUMNS 1/3	193
DMSHEMA_MINING_COLUMNS 2/3	194
DMSHEMA_MINING_COLUMNS 3/3	194
DMSHEMA_MINING_MODEL_CONTENT 1/5	195
DMSHEMA_MINING_MODEL_CONTENT 2/5	196
DMSHEMA_MINING_MODEL_CONTENT 3/5	197

	DMSchema_Mining_Model_Content 4/5	197
	DMSchema_Mining_Model_Content 5/5	198
	DMSchema_Mining_Functions 1/3	199
	DMSchema_Mining_Functions 2/3	200
	DMSchema_Mining_Functions 3/3	201
	DMSchema_Mining_Structures 1/2	201
	DMSchema_Mining_Structures 2/2	202
	DMSchema_Mining_Structure_Columns 1/3	203
	DMSchema_Mining_Structure_Columns 2/3	204
	DMSchema_Mining_Structure_Columns 3/3	204
	DMSchema_Mining_Model_XML 1/2	205
	DMSchema_Mining_Model_Content_PMML	206
	DMSchema_Mining_Model_XML 2/2	206
	Discrete Model Columns 1/5	207
	Discrete Model Columns 2/5	207
	Discrete Model Columns 3/5	208
	Discrete Model Columns 4/5	208
	Discrete Model Columns 5/5	209
	Discretized Model Column	209
	Discretized Model Column—Minimum	210
	Discretized Model Column—Maximum	210
	Discretized Model Column—Mid Value	211
	Discretized Model Column—Range Values	211
	Discretized Model Column—Spread	212
	Continuous Model Column—Spread	213
Chapter 9	After You Finish	215
	Where to Use DMX	216
	SSRS	216
	SSIS	216
	SQL	216
	XMLA	217
	Winforms and Webforms	217
	Third-Party Software	218
	Copy and Paste	218
Appendix A	Graphical Content Queries	219
	Content Queries	220
	Graphical Content Queries in SSMS	221

	Clustering Model	222
	Time Series Model	225
	Association Rules Model	225
	Decision Trees Model	228
	Graphical Content Queries in Excel 2007	230
	Data Mining Ribbon	232
	Table Tools/Analyze Ribbon	234
	Graphical Content Queries in BIDS	236
	Opening the Adventure Works Solution	236
	Reverse-Engineering the Adventure Works Database	238
	Adventure Works Database in Connected Mode	241
	Viewing Content	242
	Tracing Generated DMX	243
	Excel Data Mining Functions	246
Appendix B	Graphical Prediction Queries	249
	Prediction Queries	250
	SSMS Prediction Queries	250
	SSRS Prediction Queries	253
	SSIS Prediction Queries	257
	Control Flow	258
	Data Flow	260
	SSAS Prediction Queries	264
	Building a Prediction Query	265
	Clustering Prediction Queries	265
	Time Series Prediction Queries	268
	Association Prediction Queries	269
	Decision Trees Prediction Queries	271
	Excel Prediction Queries	274
	Excel Data Mining Functions	277
Appendix C	Graphical DDL Queries	279
	DDL Queries	280
	SSAS in BIDS	280
	Excel 2007/2010	290
	SSIS in BIDS	295
	Index	299