Index

Page numbers in italics refer to tables and figures.

acquisition models  6, 306–7
“adults–classic social users” segment, mobile telephony  319–20
affinity models (see association models)
“affluent mass” segment, retail banking  269, 289
“affluent” segment, retail banking  269
agglomerative models  42, 44–5, 61
“aggregate/group by” database function  149
aggregated information  51, 160
aggregation levels  205
anomaly detection  59, 84, 100
antecedent events  51
ARPU (average revenue per unit)  217, 296
association models  4, 7, 14, 38, 50–56, 64
   comparison with classification models  56
“average shoppers” segment, retailing  347

banking (see credit cards, retail banking)
Bayesian networks  25
behavioral fields  68, 154
behavioral segmentation  4–5, 14, 135, 203–13
   business understanding phase  203–5
consumer markets  192, 194–5
credit cards  258–64
   data preparation phase  205
data understanding phase  205
fixed telephony  296
mobile telephony  294, 295, 316
retailing  334
binning (IBM SPSS Modeler)  301, 302
“borrowers–credit card users” segment, retail banking  283, 284, 285, 287, 288, 289
brand management  214
bubble plots  242–3
business markets  200–2
“business mass” segment, retail banking  269, 289
business meanings  73
business objectives  203–4
retailing  283–8
business understanding phase  11, 203–5, 216–18
“business users” segment, mobile telephony  324–6

C&RT (classification and regression trees) algorithm  123, 125
C5.0 algorithm  124, 125
call detail records (CDRs)  163–4, 292
“cash advance users” segment, credit cards  261
categorical fields  18, 21
categorization levels  227
CDRs (call detail records)  162–3, 292
CHAID (chi-square automatic interaction detector) algorithm 125
IBM SPSS Modeler 126, 127, 128, 129–30
channel management 213
chi-square test 125
churn models 6, 14, 33–6, 195
classification and regression trees (C&RT) algorithm 123, 125
classification models 3, 14, 19–32, 61 behavioral segmentation 135
comparison with association models 55
decision trees 110, 111
evaluation 25–31
marketing applications 32–3
scoring 32
cluster centers/centroids 101, 102, 103, 238
cluster profiling 44, 100–8, 110, 130, 209, 223
credit cards 237–9
decision tree models 119–20, 209
mobile telephony 308
IBM SPSS Modeler 102–5
retail banking 276–83
clustering fields 101–2
clustering models 4, 5, 14, 39–46, 64, 82–94, 194, 195, 208, 223
K-means 46, 83, 85–8
Kohonen networks 46, 83, 91–4
profiling phase 45, 100–8
TwoStep clustering models 46, 83, 88–91
clustering solutions 96–100, 109, 127–31, 209
cohesion, clusters 97–8, 99, 100
communalities (components) 78–9
component scores 79–80, 237
confidence, decision tree rules 121
confidence measures 55
confidence scores 18, 21, 31
consequent events 51
consumer markets 191–200
continuous outcomes 37
core segments 211
mobile telephony 292–4
retail banking 264, 268
correlation 38, 48, 66
credit cards 225–64
CRISP-DM (cross-industry standard process for data mining) process model 11
CRM (customer relationship management) 1–2, 4–8, 14
cross-selling models 6, 14, 221
current tables 163, 179
mobile telephony 163–7
retail banking 138–40
retailing 179–80
customer categorization 141
customer information table mobile telephony 164–5, 166
retail banking 138, 139
retailing 179–81, 182
customer profitability 217
customer relationship management (CRM) 1–2, 4–8, 14
customer segmentation (see segmentation)
customer signatures 35
data auditing 299–300
data enrichment 206–7
data mart (see mining data mart)
data mining models 3, 13
data preparation commands, IBM SPSS Modeler 162, 163, 164
data reduction 47–50, 65, 66, 207–8, 235
data understanding and preparation phase 11, 12, 205, 218
data validation 206
database functions 149
decision rules 24
association models 51, 55–6
classification models 111
decision trees 121
sequence models 57
decision tree models 118–19
decision trees 22–3, 110–27, 209, 211, 213, 254–6, 326–9
advantages 121–3
K-means 46, 83, 85–8, 131
  comparison with Kohonen networks 93
IBM SPSS Modeler 84, 88–89, 91, 94–6
key performance indicators (see KPIs)
key usage attributes 182
Kohonen networks 46, 83, 91–4, 131–2
  comparison with K-means 93
IBM SPSS Modeler 84, 97, 98
KPIs (key performance indicators) 102
  credit cards 229, 243, 244, 259
  mobile telephony 305–6
latency period (datasets) 135
life-cycle segmentation (consumer markets) 193, 198, 199
life-stage segmentation (business markets) 201, 202
life-time value 217
lift measure 28, 55
Likert scales 49
linear correlation 38, 48, 66
logistics regression 35
lookup tables 138
  mobile telephony 170–2
  retail banking 143–8
  retailing 179, 183
loyalty-based segmentation 193, 196–8
machine learning 61–2
management strategy, segmentation 213–16
market basket analysis 51
market research 209–10
marketing applications, classification models 32–3
marketing services management 214
MARPU (marginal average revenue per unit) 217, 294, 295, 296, 301, 304, 305, 316–17
“mature users” segment, credit cards 257
maximum benefit points 31
MCIFs (marketing customer information files) 148–55
  mobile telephony 172–7, 295
IBM SPSS Modeler 150, 152, 153
  retail banking 155–60, 271
  retailing 184–7
“merge/join” database function 149
mining data mart 133–87
  mobile telephony 160–77, 295
  retail banking 137–48, 155–60, 271
  retailing 177–87
mobile telephony
  customer profitability 217
  mining data mart 160–77
    current tables 163–7
    lookup tables 170–2
  MCIFs 172–7
    monthly tables 167–70
  segmentation 194–5, 197–8, 204, 291–329, 331
model evaluation phase 11, 12
modeling phase 11, 12
monthly tables
  mobile telephony 167–70
  retail banking 140–3
  retailing 150–3
multicollinearity (correlation) 50
naming, segments 210
NBA (next best activity strategies) 8–10
needs/attitudinal segmentation 193, 199–200, 333
network codes table, mobile telephony 172, 173
network types table, mobile telephony 172, 173
neural networks 24–5
next best activity (NBA strategies) 8–10
null hypotheses 68
“occasional users” segment, credit cards 253–4, 258
“older families/retired” segment, retailing 347, 348
“oldies” segment, credit cards 250–1
“oldies–basic users” segment, mobile telephony 317–19
OLSR (ordinary least square regression) 37
“one-off/occasional buyers” segment, credit cards 263
ordinary least square regression (see OLSR)
outgoing network table, mobile telephony 170
outgoing usage table, mobile telephony 168–9
outliers (see also anomaly detection) 39, 46, 60, 90, 206

Pareto principle 221, 267, 305
IBM SPSS Modeler
anomaly detection 60, 84, 100
CHAID algorithm 126, 127, 128, 129–30
cluster profiling 102–5
data auditing 299–300
data preparation commands 162, 163, 164
decision trees 116, 117, 125
K-means 84, 87–8, 91
Kohonen networks 84, 94–96, 97, 98
MCIFs 150, 152, 153
PCA 80–2, 230
RFM analysis 339–343
segmentation 230, 238, 269, 270, 301, 302, 303, 328
TwoStep clustering models 62, 91
PCA (principal component analysis) 47, 65–82, 126, 206–7
communalities 78–9
component scores 79–80
credit cards 234–6
mobile telephony 307–12
IBM SPSS Modeler 80–2, 233
retail banking 274–6, 277–8, 279–80
Pearson correlation coefficients 48, 66, 67
personalization (data) 178

“postpaid–contractual customers” segment, mobile telephony 293
potential value 217
pre-clusters 89, 90
predictive models (see also supervised models) 135
predictors 18, 117–8
“prepaid customers” segment, mobile telephony 293
principal component analysis (see PCA)
product affinities 51
product codes table, retail banking 144–5, 147
product groups 52
product groups table retail banking 146
retailing 182–3
product hierarchy table, retailing 183, 184
product management 214
product ownership and utilization table, retail banking 141, 143
product status table, retail banking 138, 140
product suggestions 51
profiling (see cluster profiling)
profiling phase 208–10
profitability curves 219, 220, 268, 269
propensity scores 3, 21, 195–6
propensity-based segmentation 192, 195
pruning, decision trees 119
“pure mass” segment, retail banking 270–1, 276, 289
rate plan codes table, mobile telephony 171
rate plan history table, mobile telephony 165, 167
ratios 154
record screening 59–61
relative error measure 38
response measures 27, 28
“restructure” database function 150
retail banking
customer profitability 218
retail banking (Continued)
  mining data mart  137–48, 155–60
    current tables  138–40
    lookup tables  143–8
  MCIFs  155–60
    monthly tables  140–3
  segmentation  141, 194, 264–90
retailing
  mining data mart  177–87
    current tables  179–80
    lookup tables  179, 183
  MCIF  184–7
    monthly tables  180–3
  segmentation  333–48
return on investment charts (ROI charts)  30
“revolvers” segment, credit cards  260
RFM (recency, frequency, monetary) analysis  168, 334–46, 348
IBM SPSS Modeler  339, 340, 341–3
RFM scores  337–8
ROI (Return On Investment) charts  30
rotation (components)  75–6
rules (see decision rules)
scoring models  211
scree plots  73–4
segment and group membership table, retail banking  141, 142
segment migration  222
segmentation  4–5, 14, 41, 65–132, 189–224
  behavioral segmentation  203–13
  business markets  200–2
  cluster models  82–94, 190
  cluster profiling  100–8, 110, 130
  clustering solutions  96–100, 108–9, 127–31
  consumer markets  191–200
  credit cards  225–63
  decision trees  110–27
  fixed telephony  329–31
  management strategy  213–16
  mobile telephony  291–329, 331
  IBM SPSS Modeler  233, 237, 270, 301, 302, 303, 328
  PCA  65–82, 126
  retail banking  264–90
  retailing  333–48
  segmentation criteria  204
  segmentation fields
    credit cards  230–3
    mobile telephony  295–300
    retail banking  272–3
  segmentation levels  204–5, 227
  segmentation populations  204, 228–30
  segmentation types  191, 192, 223
  “select” database function  149
  self-organizing maps (SOMs)  46, 92, 93–4
  separation, clusters  98–100
  sequence models  4, 8, 14, 56–9
  service types table, mobile telephony  171–2
  “share of wallet”  266
  size-based segmentation  201
  socio-demographic segmentation (see demographic segmentation)
  SOMs (self-organizing maps)  46, 92, 93–4
SQL  149, 150, 162, 163
  standard deviation  72, 95
  standardized fields  73
  statistical techniques  61–2
  sum of squares between (SSB)  99
  sum of squares error (SSE)  96
  “superstars” segment, retailing  344
  supervised models (see also classification models, estimation models, field screening models, voluntary churn models)  3, 17–39, 61–2
  cluster profiling  110
  segmentation  254–6
  supervised segmentation  110
  support, decision tree rules  121
support measure 54
SVM (support vector machine) 25

“telcos fans” segment, credit cards 247–8, 257
telecommunications (see fixed telephony, mobile telephony)
terminal nodes, decision trees 114
top-up information, mobile telephony 169–70
training phase 17, 135, 212
transaction channels table, retail banking 145, 146, 147
transaction codes table, retail banking 145, 147
transaction data 52, 53–4
retail banking 141–8
retailing 180, 181
transaction types table, retail banking 145, 148
transactions table
retail banking 142, 144
retailing 180, 182
“travelers” segment, credit cards 249–50, 257
TwoStep clustering models 46, 83, 88–90, 131
IBM SPSS Modeler 62, 91
“typical customers” segment, retailing 344
“typical users” segment, credit cards 246–7, 256

“unclassified” segment 213
unsupervised models (see also association models, clustering models, sequence models) 3–4, 39–61, 63–4
data reduction 47–50
record screening 59–61
unsupervised segmentation 127
up-selling models 6, 14, 221
“update” database function 150

validation phase 136
value-at-risk segmentation 14
value-based segmentation 14, 216–22, 223
business markets 200–1
consumer markets 191, 193–4
mobile telephony 294, 296, 300–7, 316
retail banking 267–8
retailing 333
variance 72–3
Varimax rotation (PCA) 75–6, 234
voluntary attrition models 6
voluntary churn models (see churn models)

“young –SMS users” segment, mobile telephony 321–2
“young active customers” segment, retail banking 283, 284, 285, 286, 287, 289