Index

A
“abnormal” periods  86
Abrahams, C.  15, 247, 258
adjudication, strategy development and  307
adverse codes  243–244
“adverse” impact  31
“aggregated up” default events  123
Akaikes Information Criterion (AIC)  253
analytical methods  96–99
applicants
changes in profiles of  282–283
type of, as segmentation area  106
applications, scoring, Basel II and  128–129
area under the receiver operating characteristic curve (AUC)  155–156, 249
assessment, validation and  325, 337–340
asymmetry  342
attributes  11
AUC (area under the receiver operating characteristic curve)  155–156, 249
augmentation
fuzzy  231–232
simple  226–227
augmentation 2 method  227
authorizations, strategy development and  307
availability
of data  74–76, 105–106
future  136–137

B
backward elimination  206–207
Bad definition  80–89, 93, 95–99, 141–142
bankruptcy scorecards  85
bankruptcy/charge-off, versus mild delinquency  297
banks, big data and  161–163
Basel Committee on Banking Supervision (BCBS)
regulation  239 16
Basel II Accord  2–3, 16, 47, 81, 98, 104, 128–129
See also Default definition
Bayes Rule  126–127, 257
BCBS (Basel Committee on Banking Supervision) regulation  239 16
Bean, R.  154
behavior scorecards  85, 202–203
benchmark  339
Bender B.  15
Benjamin N.  94
Bholat, D.  154
“Big Brother” issues, big data and  168–170
big data
about  150
“Big Brother” issues and  168–170
collection process  158
credit scoring and  158, 159–164
ethical considerations  164–170
value of  154–158
variety of  152–153
velocity of  151–152
volume of  150–151
black box  20, 52, 71
budgets, for continuous education  69
bureau data-based method  223–225
bureau scores
versus custom in-house behavior scores  296–297
migration of  228
business considerations, for scorecard development  197–199
business intelligence, application of  7–8
business plans  46–56
business reasoning 136
business reports, common 398–399
See also specific reports
business sources, as segmentation area 105
business-based approach 196–197
“buy-in” 155

C
canceled accounts 100
case-based reasoning 233–234
Cathcart, A. 94
causality, social media data and 139
CCAF (Comprehensive Credit Assessment Framework) 299
CCAR (Comprehensive Capital Analysis and Review), 16, 47
CECL (Current Expected Credit Loss) 16
Characteristic Analysis report 273–274, 284–287
characteristics
initial analysis of 179–200
selection of 133–140
charge-off scorecards 85
checkpoints, model risk management and 364
chi-square 88, 180, 185, 186, 207, 356n23
classifier models 334
closed accounts 100
clustering 109–111, 233
coding
error in 284
maintaining continuity of 70
Cohort Analysis 390–394
collections, strategy development and 307
communication, strategy development and 306–309
competitive environment, changes in 137
compliance manager, as a role in scorecard development 31
Comprehensive Capital Analysis and Review (CCAR) 16, 47
Comprehensive Credit Assessment Framework (CCAF) 299
confidence gap, for in-house scorecard development 54
consensus method 96
controls, model risk management and 365–366
core elements 332
corporate decision making, changing 47
corporate risk
management staff, as a role in scorecard development 30
strategy development and 308
correlation 178–179
cost
increasing savings 47
of in-house scorecard development 52
of outsourced scorecard development 50
cost ratio 257
Cox proportional hazard 200
credibility of developer, of outsourced scorecard development 50–51
credit application analysis reports 366–377
credit crisis, of 2008 17
credit limit-granting strategy 311–313
credit scorecard
See modeling scorecard
credit scoring
big data and 159–164
data collection process and 158
future of 16
increasing knowledge of 16–17
in insurance sector 16
credit scoring manager, as a role in scorecard development 23–24
Crone, S.F. 74–75
cross selling 309–310
c-statistic 112–113, 144, 254–255, 356n23
Current Expected Credit Loss (CECL) 16
custom in-house behavior scores, versus bureau scores 296–297
“custom” scorecards 14
customer service, strategy development and 308
cutoffs
analyzing 300–302
setting 297–306
cycle time
of in-house scorecard development 53–54
of outsourced scorecard development 51
INDEX

D

data
See also big data
availability of 74–76, 105–106
collection process for 77–78, 142–144, 158
construction of 142–144
exploring 175
insufficient 58
missing 134
mistakes in capture of 284
nonavailability of 58
nonpredictive 58
normalizing 88
quality of 58, 74–76
rise of alternate sources of 16
typical items of 77
validation and 325
weak 58
Data Change Log 143
data integration 157
data mining 6
data quirks 143–144
data review
about 74–76
data gathering 77–78
implementation plan 117
methodology 116–117
project parameters
defined 78–103
segmentation 103–116
data scientists 22–23, 156
data silos 154–155
Davenport, T. 157
decision trees 111–112, 200
decisions, anticipating impacts of 8–9
Default definition
about 120–121
application scoring 128–129
Default event 121–124
prediction horizon 124–125
validation of Default rate and recalibration 126–128
Default event 120, 121–124
Default rate 120, 124–128
delays, in accessing data 58
delinquency
comparing 98–99
to cross-sell propensity 297
defined 121
mild, versus bankruptcy/charge-off 297
delinquency maturity curve 82–83
Delinquency Migration report 394–396
delinquency rate 82
delinquency report 386–390
delinquency scorecards 85
demographics, as segmentation areas 105
deployment, validation and 340–341
development
about 140–141
cost of 115
development database
about 132
adjusting for prior probabilities 144–147
development data collection and construction 142–144
sample specification 132–140
sampling 140–142
DFAST (Dodd-Frank Act Stress Test) 31, 47
disclosure
for in-house scorecard development 54
vendor 345–347
discriminant analysis 200, 321
“disparate” impact 31
Dodd-Frank Act Stress Test (DFAST) 31, 47
“downstream” effect 165
“due diligence” policies 13–14

E
EAD (Exposure at Default) 3
ease in collection 135
education, strategy development and 308–309
effects, isolated 379–380
enterprise risk management staff, as a role in scorecard development 30
Equal Credit Opportunity Act 31
estimation methodology, for models 331–337
ethical considerations, with big data 164–170
ETL (“point-and-click”-based Extract-Transform-Load) software 6
on-going quality assurance of vendors 351–352
validation assessment 337–340
vendor management considerations 323–326
vendor model implementation and deployment 340–341
vendor model purpose 326–331
Gini Coefficient 75, 88, 144, 180, 185, 207, 249, 254–255, 356n23
GL-44 Guidelines on Internal Governance 258, 259
Good account 99–103, 141–142
Google flu algorithm 16
“go-to-market” strategy 67
Graves, P. 151
Groenfeldt, T. 162

H
Hadoop 157
Hand, D. 333
Hanson, S.G. 94
hard lowside cutoff 317
HELOCs (home equity lines of credit) 289
heuristic (experience-based) segmentation 105–109
high-risk applicants, strategies for 13
holistic model risk 321–322
home equity lines of credit (HELOCs), 289
Hosmer-Lemeshow 88
human intervention 135
human resource risk 65–68

I
IFRS (International Financial Reporting Standards) 4, 47, 77
illusion of progress 333–334
impact, gauging 303–306
implementation about 276
cost of 115
plan for 117
pre-implementation validation 276–279
strategy development 291–317
validation and 326, 340–341
improvement, comparing 112–114
inactive accounts 100
| INDEX | 433 |

INCLUDE option, LOGISTIC procedure 214  
independent validation, model risk management and 365  
Indeterminate account 99–103  
industry trends/changes, for local credit bureaus 70  
infering indeterminates 234–235  
information, usability of 72  
information technology (IT), strategy development and 306  
Information Value (IV) 180, 186–197, 369  
infrasctructure design about 36–39  
creating modeling data sets 41  
data gathering/organization 39–40  
data mining 41–42  
for in-house scorecard development 53  
model implementation 43–44  
reporting and analytics 44  
validation/backtesting 43  
“inherited” default events 123  
in-house scorecard development about 64, 223–225  
benefits and drawbacks of 52–54  
human resource risk 65–68  
technology and knowledge stagnation risk 68–72  
insights of in-house scorecard development 53  
of outsourced scorecard development 51  
insurance accounts 100  
insurance sector, credit scoring in 16  
integrated infrastructure, model risk management and 365  
inization, importance of 70  
intelligent scorecard development 31  
Interactive Grouping node 193  
interactive regression 213–214  
intercept adjustment 356n25  
internal development, versus external development 49–56  
International Financial Reporting Standards (IFRS) 4, 47, 77  
interpretability 135  
iterative reclassification 232–233  
IT/IS managers, as a role in scorecard development 29–30  
IV (Information Value) 180, 186–197, 369  
J  
Jackson, L. 94  
judgmental models 56, 95, 96  
K  
Kabbage 153  
KGB (known good/bad) line 232  
K-means clustering 109–111  
knowledge base, building 69  
knowledge gain of in-house scorecard development 53  
of outsourced scorecard development 51  
knowledge gap 71–72  
known good/bad (KGB) line 232  
Kolmogorov-Smirnov (KS) 88, 112–113, 126, 144, 202, 249, 251, 253–254, 356n23  
Kullback’s divergence statistic 356n23  
kurtosis 342  
L  
Law of Large Numbers 92  
LDP (low-default portfolio) 91–95  
legal issues 135–136, 309  
legal staff, as a role in scorecard development 31  
LGD (Loss Given Default) 3  
Lift/Concentration curve 257  
line of credit 310  
loan to value ratio (LTV) 55, 75  
local presence, vendors with 67  
localisation, for in-house scorecard development 54  
logical trend 186–197  
LOGISTIC procedure 146, 214  
logistic regression 200, 204–207  
logit function 146, 204  
long-run average 127  
Lorenz curves 201–202, 255  
Loss Given Default (LGD) 3  
low-default portfolio (LDP) 91–95  
LTV (loan to value ratio) 55, 75
machine learning 17
management direction, model risk management and 364–365
management reports about 266–267
characteristic reports 273–274
Gains table 267–273
Maner, C. 15
marginal event rate 270
marketing campaigns for 283
strategy development and 306
materiality, of outstanding amount 122
matrix approach 295–297
matrix-sequential hybrid approach 297
MC (multicollinearity) 178
mean squared error (MSE) 252
median 356n24
memory-based reasoning 233–234
methodology for data review 116–117
model risk management and 365 validation and 325
misclassification 247–250, 257
missing data 134
missing values 175–178
model 356n24
model developer, as a participant in model risk 322–323
Model Ordering option 214
model owner, as a participant in model risk 322, 356n16
model risk management, as a participant in model risk 323
model user, as a participant in model risk 323
model validation/vetting staff, as a role in scorecard development 28–29
MODEL y= statement 214
modeling scorecard 4 models estimation methodology 331–337
cjudgmental 56, 95, 96
risk management for 363–366
monitoring in conjunction with other scores 347–351
validation and 326, 341–351
monthly cohort 81–82
MSE (mean squared error), 252
multicollinearity (MC) 178
multiple regression 213–214
nearest neighbor (clustering) 233
net present value (NPV) 100
neural networks 200, 321
niche competition 283
non-scorecard model 93–94
nonsegmented data set 143
non-validated scorecards 287–291
normalizing data 88
NPV (net present value) 100
“offer declined” accounts 100
offset method 145–146
OLS (ordinary least squares) regression 178
one-time events 377–378
operational areas, risk scoring in 15
operational considerations, for scorecard development 197–199
operational efficiency, increasing 47
operational manager(s), as a role in scorecard development 25–28
operational use, behavior scorecards for 85
options, evaluating 293
ordinary least squares (OLS) regression 178
organizational objectives, identifying 46–49
out of range assumptions 332
outliers 175–178
out-of-time stability validation 276
out-of-time testing 334
outsourced scorecard development, benefits and drawbacks of 50–52
override report 383–386
overrides 315–317, 344–345
oversampling 144
parcelling 228–231
partial least squares 200
PD (Probability of Default) 3, 81
performance windows 80–86, 92
planning
about 46
creating business plans 46–56
creating project plans 57–60
scorecard format 60–61
Pluto, K. 94
Point in Time models 127–128
“point of no return” 97, 121
“point-and-click”-based Extract-Transform-Load (ETL)
software 6
points allocation 245–246
policy rules 293, 313–315
“pooled data” scorecards 15
population, versus sample 163–164
Population Stability Index (PSI) 88, 343, 357n29
Portfolio Monitoring
reports 360–377
portfolio risk manager, as a role in scorecard development 23–24
post-implementation
about 399–401
portfolio monitoring reports 360–377
reacting to change 377–399
scorecard monitoring reports 360–377
prediction horizon 120, 124–125
predictive modeling 107, 108
predictive power 47, 133
predictive variables 71–72
pre-implementation validation 276–279
privacy laws, social media data and 138
probabilities, adjusting for prior 144–147
Probability of Default (PD) 3, 81
“probation” period 125
process default 121–122
processing, as a segment 115
product manager(s), as a role in scorecard development 25
product owned, as segmentation area 106
product risk manager, as a role in scorecard development 23–24
product type, as segmentation area 105
products
upgrades of 309
widely used 67
profitability, increasing 47
profitability analysis 96
Pro-Forma Variable Distribution Analysis 87, 288
project manager, as a role in scorecard development 29
project parameters 78–103
about 74–76
data gathering 77–78
defined 78–103
implementation plan 117
methodology 116–117
segmentation 103–116
project plans, creating 57–60
project risks, identifying 57–58
project team, identifying 59–60
proprietary objective function 355n3
proxies 95
PSI (population stability index), 88, 343, 357n29
public opinion, social media data and 138
purpose, validation and 325
p-values 207
Q
quality assurance, validation and 326, 351–352
quality risk/gap 71
quantiles 357n29
“quarantine” period 125
quarterly cohort 81–82
R
random allocation 228
random applicants 142
ratios, creating 136
recalibration 126–128, 343
Recent Bad 93
regression
interactive 213–214
logistic 204–207
multiple 213–214
single 212–213
regulatory compliance
about 47
of in-house scorecard development 53
regulatory compliance (Continued) of outsourced scorecard development 51
INDEX ▶ 437

about 9–18
age of 378
bankruptcy 85
behavior 85, 202–203
building 8
characteristics of 11
charge-off 85
choosing 246–257
collaborative development of 8
“custom” 14
as decision support tools 9
delinquency 85
designing 207–215
factors driving increased use of 2
format of 60–61
generic 15, 55, 355n4
identifying roles 46–49
modeling 4
non-validated 287–291
“poled data” 15
preliminary 200–215
samples 10
scoring
accuracy for testing 276
strategy for 293–297
seasonality, effects of 86–89
segmentation
about 103–104
choosing segments 115–116
comparing improvement 112–114
experience-based (heuristic) 105–109
statistically based 109–112
segments
considerations about 292
size of 115
self-organizing maps (SOMs) 109–111
sequential approach 294–295
SEQUENTIAL option, LOGISTIC
procedure 214
shadow ratings approach 94–95
shape 342, 357n26
shift, levels of 378
short-term scorecard development 67
“showstoppers” 58
simple augmentation 226–227
single regression 212–213
60-day delinquency 100
social media data 138–139, 167
soft cutoff 225–235
Somers’ D, gamma, tau-a 257
SOMs (self-organizing maps) 109–111
Spark 157
stacked samples 92
START option, LOGISTIC
procedure 214
statistical measures 182–186
statistically based segmentation 109–112
stepwise 207–215
strategy development
communication 306–309
considerations for 291–293
overrides 315–317
policy rules 313–315
risk-adjusted actions 309–313
scoring strategy 293–297
as a segment 115
setting cutoffs 297–306
stress testing 16
structured data 158
swap set 217–219, 302
systematic errors 378
T
talent, for in-house scorecard
development 52–53
target, validation and 325, 328–329
target population, validation and 325, 327–328
Tasche, D. 94
technology
for in-house scorecards 66
issues with 379
stagnation risk 68–72
30-day delinquency 100
Through the Cycle models 127–128
through-the-door (TTD)
population 86, 356n22
time
roll rate analysis and 396–399
since changes 378–379
track records of developer,
of outsourced scorecard
development 50–51
training, availability of
regular 66–67
transparency
as a vendor management
considerations 324–326
transparency (Continued)
  of vendor model estimation 332–335
tree-based methods 321
trends, across time 377–378
TTD (through-the-door)
  population 86, 356n22

U
“unlikeliness to pay,” indications of 123

V
validation 140–141, 258–262, 276–279
  See also generic vendor scorecards
value, of big data 154–158
VARCLUS procedure 178, 210
Variable Clustering node 178
variables, transformation of 356n19
variance 342
variety, of big data 152–153
velocity, of big data 151–152
vendors
  about 67, 357n30
disclosure challenges for 345–347
management considerations 323–326
model purpose 326–331
verification 235–236
vetting, for in-house scorecard development 53
Vintage Analysis report 81–82, 390–394
volume, of big data 150–151

W
Wagner, Hendrik 3
Weight of Evidence (WOE) values 126–127, 142, 180, 186–197, 222
“what-if” analysis 292–293
Wilcoxon-Mann-Whitney test 254–255
Wilde, T. 94
WOE (Weight of Evidence) values 126–127, 142, 180, 186–197, 222

Z
ZestFinance 155
Zhang, S. 247, 258