Page references followed by f indicate an illustrated figure; followed by t indicate a table; followed by b indicate a box; followed by e indicate an example.

A

AAPL, 67, 149
  average daily trading volume, 160
  exchange auction, 183
  options, leverage, 154f
  out-of-sample accuracy, 127
Accern, 5
AcquireMedia, NewsEdge database, 5
ACWV, 22
Adverseselection, 170
  reduction, 174–177
Aggressor tag, 189, 207
  method, 190t, 193
  strategy, 194–195
  trading algorithm, 192
Akaike information criterion (AIC), 87
  criteria, 53
Algorithmic trading, 1
  flowchart, 2f
  impact, 215
Algoseek, 184
AlgoTracker, 9

Alpha, 28
  models, 29
API, usage/provision, 1, 3
AR(10) model, coefficients (application), 62t
AR(10) trading strategy, application, 62f
AR(16) trading strategy, application, 204f
Arbitrage, 124
ARIMA, 59
  forecast function, 68
Arima function, 65
ARMA (p, q), 63–67, 203
ARMA(3,7) trading strategy, application, 205f
ARMA (2,5) model coefficients, application, 65t
ARMA models, 71
ARMA (2,5) trading strategy, application, 66f
AR models, 71
AR(p), 60–63, 203
ARQ_BVPS, usage, 39–40
ARQ_EPS, division, 39–40
ARQ_PB reciprocal, 40
Array processing, 7
Artificial intelligence (AI)  
  basis, 85–86  
  techniques, 83–84, 206–207  
Astec database, 5  
At-the-money (ATM)  
  call implied volatilities, 46–47  
  options, usage, 153  
At-the-money (ATM) call/put options, 41–42  
At-the-money (ATM) options, 133–134  
AUD.USD, 60  
  AR(10) model coefficients, 62t  
  AR(10) trading strategy, usage, 62f  
  ARMA (2,5) model coefficients, 65t  
  ARMA (2,5) trading strategy, 66f  
Automated execution platforms, 9  
  Autoregressive coefficients, 68t  
Auto-regressive process, 60  

B  
Backtesting, 154  
  execution platforms, 9  
  intraday strategies, 177–185  
  profit, 208  
Backtesting platforms, 6–9  
Backwardation, 124  
Bagging, 85, 93–95  
  examples, 94f, 95f  
Bankhaus Herstatt, 11  
Base currency, 202  
BATS  
  BZX exchange, fees charge reasons, 168b  
  Hide Not Slide Order, 164  
Bayesian information criterion (BIC), 61, 87  
  criterion/criteria, 53  
  usage, 128–129  
  minimization, 63  
Bayes Net Toolbox, 84, 103  
Best bid and offer (BBO), 3, 164  
  construction, ITCH messages (usage), 178e–181e  
  displayed quotes, 164  
  prices, 3  
  quotes, 192  
  size, capture, 196  
  usage, examples, 185t  
BIC. See Bayesian information criterion  
Bid-ask bounce, 206  
  reduction, 60  
Bid-ask quotes, 211  
Bid-ask spread, 186  
Big-market-capitalization stocks, 31  
Bitcoins, 59, 201  
  facts, 201–203  
Bitfinex, 211  
BitStamp, 202–203  
  BTC.USD feed, 207  
Black-box, 85–86  
Black Monday, SPY return, 13  
Bloomberg, 5  
  terminal, advantage, 3–4  
Bollinger band trading strategy, 205f, 206f  
  usage, 220  
Book-to-market ratio (BM)  
  factor model, fitting, 39e–41e  
  log, 37  
Book-to-market stocks, 29  
Boosting, 96–98  
  effect, 97f  
Brightwater Associates All Weather fund, 22  
Brokers, 10–12  
  API, usage, 6, 9, 10  
BTC China, 202  
BTC.CNY, 207  
BTC.USD, 202  
  AR(16) trading strategy, 204f  
  ARMA(3,7) trading strategy, 205f  
  data, 206  
  equity curve, Bollinger band trading strategy, 206f
Buffett, Warren, 27
Bulk volume classification (BVC), usage, 187b–188b, 194
Buy and sell signals, generation, 205–206
buyOrderBook, 179–180

C

C (programming language), 7
C# (programming language), 9
C++ (programming language), 7, 162
CAGR. See Compound annual growth rate
Calendar spread quotes data, 184e–185e
Call implied volatility (CIV), 42
Call-put implied volatilities, 45–46
Calmar ratio, 13–14, 123, 129–130, 139
basis, 143
elevation, 134–135
Capital allocation, tangency portfolio
(correspondence), 20
Capital asset pricing model (CAPM), 42
Center for Research of Security Prices
(CRSP)
data, 3, 43
midprices, usage, 67
survivorship-bias-free database, 31–32
Chattopadhyay two-factor model, 50
Child node, 89
CIV. See Call implied volatility
Classification tree, 98–99
CLS bank, establishment, 11
CME equity index futures, 124
CME Globex, crude oil futures
expiration, 135
CNY, 202
Cognitive dissonance, creation, 108–109
Coinsetter messages, 179
Cointegration, Johansen test, 70
Commodity Futures Trading Commission
(CFTC), 224
Commodity Pool, 224, 225, 226
(endnote)
Commodity Pool Operator (CPO), 222
Commodity Trading Advisor (CTA),
222, 224

Compound annual growth rate (CAGR),
12–14
calculation, 33
generation, 39, 48
in-sample CAGR, increase, 88, 92
maximization, 14
obtaining, 91
out-of-sample CAGR, increase, 38
percentage value, 87, 104, 143
training set level, 37
unlevered CAGR
expectation, 45–46
obtaining, 47
Compustat data, 4–5
Computer hardware stocks
Kalman filter trading strategy,
application, 75f
VAR(1) model, application, 68t
VAR(1) trading strategy,
application, 70f
VEC(0) model, application, 70t
Constant offsets, 68t
Constant Proportion Portfolio
Insurance (CPPI), 221
Contracts trading dates, 136t
Correlation trading, 142
Covariance
example, 68t
matrix, 16–17, 72
CQG, 5, 6
CQG Data Factory, 182
Cross-border cross-exchange
arbitrage, 211
Cross-exchange arbitrage, 211–212
Cross-sectional factors, 33–37
loading, 33
model, fitting, 33e–37e
Cross-sectional mean reversion,
152–155
Cross-sectional two-factor model, 37–38
Cross-validated accuracy, 92–93
Cross-validated regression tree, 110, 111f
Cross validation, 92–93
Cross validation data set, usage, 107
Cross-validation test, training set (subset omission), 92f
Cross-validation, usage, 101f
Crowd-sourcing earnings estimates, availability, 4–5
CRSP. See Center for Research of Security Prices
Crude oil futures
  gamma scalping (usage), straddle (combination), 140c–141e
  options straddles, shorting, 135e–137e
CSI, usage, 2–3
cTree.m, 99
Cumulative distribution function (CDF), 188, 188e
Cumulative returns
  RV(t+1)–RV(t) strategy, cumulative returns, 130f
  VX-ES roll returns strategy, cumulative returns, 124f
  XIV-SPY roll return strategy, cumulative returns, 126f
Currency conversion counters, 218
Currency markets, 175–176
Currency returns, 59
Currency traders, 11
CUSIP, 145
Cython compilers, usage, 162

D
Daily data, futures traders
  usage, 4
Daily options data, collection, 4
Dark pools, routing, 171–174
Data
  aggregation, 109–111
  mining, 83–84
  normalization, 109
  structure, 190t, 192t
  vendors, 6
Days to cover (DTC), 48
Debugging productivity, 7
Deep learning, 108–109
Delta, 138
constraints, 145
  defining, 119
Delta-hedging, 144
Delta-neutral, meaning, 119
Deltix, 9, 162
Direct Market Access (DMA), 10, 164
Discrete (categorical) response variables, 98
Dispersion strategy, 144
Dispersion trading, 120, 142–152
  example, 143f
  usage, 144e–152e
Diversification, 225
DMA. See Direct Market Access
Dow Jones, 5
Drawdowns, 122
Dumb beta, 28–29

E
Earning alpha, 28
EEMV, 22
Efficient frontier, 16, 17f
  calculation, quadratic programming (usage), 17b–20b
  determination, 17
Eigenvalues, descending magnitude, 50
Eigenvectors, 51
EM algorithm, 103
EMC, 67
Emissions, 102, 104
  probabilities, 102
Enron, large-cap stocks, 28
Ensemble method, 95–96
  entryThreshold, selection, 191–192, 210
  entryZscore, 206
Equity curve, 91, 129–130
Error correction matrix, 70t
eSignal, 6
Estimate function, 72–73
  application, 76
Estimize data, 4–5
Eurodollar futures (ED) outright market, BBO (usage), 185t
European options, put-call parity, 45–46
EURUSD, out-of-sample accuracy, 127
Event-driven strategies, 14, 133–137
EWA-EWC, linear regression relationship, 76
EWA, shorting, 78
EWC-EWA
Kalman filter trading strategy, application, 79f
offset, Kalman filter estimate (usage), 78f
out-of-sample EWC-EWA, Kalman filter trading strategy (application), 79f
EWC treatment, 76
Exchanges, order book (examples), 165t, 167t
Exchange-traded futures (ETFs), 2–3 components, 17
EWA, 76
expected return, 18
portfolio, efficient frontier (determination), 17
SPLV, 22
SPY next-day return, 85
Exchange-traded notes (ETNs), 121
exitThreshold, 210
Exponential moving average, 72

Face-to-face meetings, impact, 223
Factor
BM factor model, fitting, 39e–41e
cross-sectional factors, 33–37
exposure, 29
fundamental cross-sectional factor model, 37f
input factors loadings, size-independent characteristics, 34t
loadings, 29, 35
nonlinear model, application, 54–55
model, 27
implementation, 29
predictive statistical factors (discovery), Principal Component Analysis (usage), 51e–53e
risks, 27
ROE factor model, fitting, 39e–41e
selection, 113t
statistical factors, 49–53
time-series factors, 29–31, 50
two-factor model, 37–40
FactorLoadings matrix, 52
Fama-French factors, usage, 31e–32e
Fama-French three-factor model, 50
FCM. See Futures Commission Merchants
feedforwardnet function, 105
Filter function, usage, 74
Filtering, meaning, 74
fitctree function, 98
fitensemble function, 96–97
fitlm function, 32, 35, 52, 86, 113
fitrtree function, 94, 97, 98
Fivefold cross validation, imposition, 98
Flash crashes, 171
Forecast functions, 65
Function filter, application, 77
Fundamental factors, usage, 114f
Fundamental financial data, 4–5
Futures
daily historical data, 2–3
traders, daily data (usage), 4
Futures Commission Merchants (FCM), 10–11
FX brokers, money (earning), 10
FX market, 177

Gamma scalping, 120, 137–141
path-dependent profit, 138
usage, example, 140e–141e
Gaussian, CDF, 188e
function, 193
Gaussian distribution, 13–15
Gaussian fluctuations, 217–218
Gaussian noise, 60
assumption, 73
GBP.USD, 211
GDI. See Gini’s Diversity Index
GDX
  out-of-sample spread, 220f
  pair trading, 219–220
Generalized Autoregressive Conditional
  Heteroskedasticity (GARCH), 126, 129
  basis, 128
  prediction, 129–130
  usage, 130
Gini’s Diversity Index (GDI), 98
GLD
  out-of-sample accuracy, 127
  out-of-sample spread, 220f
  pair trading, 219–220
Goldman Sachs, Sigma X, 171–172
Greeks, 120, 145
Greenblatt, Joel, 39
Gross margin, 33
Gross market value, computation, 150
H
  Head-of-queue priority, obtaining, 166
  Hedge portfolio, 29
  Hedge ratio
    display, 125
    estimates, 76
  Herd-like behavior, 83
  Herstatt risk, 11
  HFT, 168–169
Hidden Markov Model (HMM), 101–105
  hidden states transition
    probabilities, 103f
  supervised training, usage, 85
  unsupervised training, usage, 85
  Hidden variable, 71–72
  Hide-and-light, 164
  order, usage, 178
  Hide Not Slide Order, 164
  High-leverage point, 53–54
  Historical data, retrieval, 136–137
  Historical market data, 2–5
  Historical stock prices, adjustment, 2–3
  HML, 29, 30
    portfolio return, 31
time-series factor, 33
  value factor, 31
HMM. See Hidden Markov Model
  hmmdecode function, 104
  hmmviterbi function, 105
  Hodrick-Prescott filter, 74
  HotspotFX, 178, 186
    user acceptance testing (UAT) account, 176–177
HPQ, 67
Hyperplane, usage, 100
I
  IB. See Interactive Brokers
  IBES database, usage, 4–5
  IBM, trading, 211
  IBPy, usage, 7
  Identity matrix, 72
  IEX, 171–174
  Immediate-or-cancel (IOC)
    limit IOC order, 172
    modifier, usage, 170–171
    orders, 177
  Implied market volatility, daily change, 47
  Implied moments, 41–43
    high/low implied moments, 43e–45e
  Implied volatilities, 133
    call-put implied volatilities, 45–46
    defining, 41–42
    factor, 54
    historical realized volatility, ratio, 139–140
    monthly change, 45
    OTM put–ATM call implied volatilities, 46–47
    storage, 43
  Implied volatility, cross-sectional mean
    reversion, 120, 152–155
    impVolCrossSectionalMeanReversion.m, 153
  Index straddle/SPX component straddle
    (contrast), dispersion trading
    (usage), 144e–152e
  Initial guesses, dependence (reduction), 107–108
Innovations, 60
Input factors
  loadings, size-independent characteristics, 34t
  size-independent characteristic, 112t
In-sample CAGR
  increase, 88, 92
  obtaining, 107
In-sample computer hardware stocks,
  Kalman filter trading strategy (application), 75f
Integrated development environment (IDE), 7
Interactive Brokers (IB), 224
  latency, 162
  Point of Presence (POP), 161
Interactive Data, 6
Intermarket sweep order (ISO), 166–167
  book walking process, 169e–170e
  sending sequence, impact, 170
In-the-money (ITM), 138
Intraday futures strategy, 182
Intraday trading, 4, 159
Investment Advisor, 224
Investment decisions, face-to-face meetings, 223
Investment management services, offer, 218–219
Investors, benefits, 223
IQFeed, 6
ITCH data, 178
ITCH feed, 161, 178, 186
ITCH messages, usage, 178e–181e, 186
ITG, 173
  POSIT, 171–172

J
Java (programming language), 8
Johansen test, 70

K
Kalman filter, 59, 71, 104
  application, 76
  estimate, examples, 77f, 78f
  trading strategy, application, 75f
  usage, 74, 125f
  $k$-dimensional Gaussian white noise, 71
Kernel function, 100–101
Kernel transformation, 101
Kelly, 12, 13, 19, 20, 21, 22, 24, 121, 123, 159, 199, 224
KF_beta_XIV_SPY.m, 125
kfoldLoss function, 93
Kibot, 5
Kurt, computation, 43–44
Kurtosis, 42
  factor, 564

L
Lagrange multiplier method, usage, 20
Lagrange multiplier, zero method, 21
Latency
  arbitrage, 172–173
  minimization, 161–162
  reduction, 160–163, 175
  types, 160–161
learn_params_dbn_em function, 103
Lehman Brothers, large-cap stocks, 28
Levered compounded returns,
  comparison, 122e–123e
Lime Brokerage, Strategy Studio, 9
Limit-on-Close (LOC) order, 183
Limit-on-Open (LOO) order, 183
Linear regression
  relationship, 76
  usage, 50
Liquidation problem, 22
Liquidity, 48–49, 54, 168
  addition, 164–168
    fees charge, BATS BZX exchange reasons, 168b
    elevation, 119–120
    factor, 49f
    liquidity-sorted stocks, 49
    taking, 168–171
Liquidity provider (LP), 176
Live market data, 5–6
Lloyd’s of London, 10
LMAX, 177
Log prices, SSM application, 73
Log return
calculation, 14
mean, net return mean (contrast), 15b–16b, 16t
standard deviation, 15
Long-short portfolio, 29
construction, 42, 48
Long SPY, short VX contrast, 121f
performance comparison, 122t
Long stocks, high/low implied moments, 43e–45e
Low frequency data, 182e–184e
LSBoost, specification, 97
MATLAB, 162
array, storage, 145
backtesting platform, 6
binary files, 144–145
code fragment, 69, 76
Data Structures package, 179
datenum, 209
estimate function, 72–73
Optimization Toolbox, 18
serial date number, 190
Statistics and Machine Learning Toolbox, 32, 50–51, 84, 101
usage, 3, 7–9
Matrices, estimated values, 73t
Maximum drawdown, 13–14, 22
Maximum likelihood estimation method, 72
Maximum variance portfolio, construction, 22
MDP Market Data, 186
Mean log return, calculation, 15
Mean reversion strategy, 137–138, 204–206
Mean-reversion trading models, 168
Mean-reverting strategy, 14, 178
Mean-reverting technique, 102–103
Mean-reverting trading rules, 91
Mean squared error (MSE), 89
Mex files, usage, 162
MF Global bankruptcy, 10
Mind/health, 215–218
Minimum variance portfolio, 18–19
MinLeafSize, selection, 90
Momentum strategy, 14, 163
Money, management, 222–225
Moneyness, 47
Moving averages, 59
model, 72
usage, 72
MSE. See Mean squared error
Murphy, Kevin, 84
MXN.USD, 203

M
Machine learners, ambiguity, 102
Machine learning, 109
techniques, 84
Macroeconomic factors, 31
Managed accounts, 224, 225
Market
capitalization, 112
index, returns, 30
inefficiency, 83
microstructure, 159
orders, usage, 189, 195
risk, 27
Marketcetera, 9
Market data
historical market data, 2–5
latency, 161–162
live market data, 5–6
Market makers, 176
Market-neutral portfolio, out-of-sample CAGR, 39
Market-neutral results, 39
Market-on-Close (MOC) order, 183
Market-on-Open (MOO) order, 183
MarketPsych, 5
Markowitz portfolio optimization method, 14
MAR ratio, 13–14
N

NaN, 40, 113
  returns, 51–52
  unknown parameters, 76
Nanex, 5
National Best Bid Offer (NBBO), 168–169, 172
  basis, 166
  liquidity, 177
  midprice, 173
  quote size, 167
National best offer (NBO), 165, 167
National Futures Association (NFA), 10–11
National Oilseed Processors Association (NOPA), crush report, 134
Net asset value (NAV), 224, 225
drawdown, 143
Net return calculation, 14
  mean, log return mean (contrast), 15b–16b, 16t
  standard deviation, 15
Network architectures, performance comparison
  averaging, usage, 108t
  retraining, usage, 108t
Neural network, 105–109
Neural Network Toolbox, 105
Neurons, impact, 105, 107
New calendar spread limit order, addition, 185t
News data, 5
Next-day return (prediction),
  Fama-French factors (usage), 31e–32e
Next-month return (prediction), ROE/B factor model (fitting), 39–41
Next-period return, 128
Next-quarter return (prediction),
  cross-sectional factor model (fitting), 33e–37e
NFA. See National Futures Association
Ninjatrader, 9
Nondisplayed orders, presence, 187
Nonzero cross-sectional correlations, 67
Normalization, 109–111
NTAP, 67, 70
Numba compilers, usage, 162
Numerical quadratic programming method, usage, 16–17
NumPredictorsToSample, 96

O

Object-oriented design, 8
Off-diagonal elements, 73t
One-factor model, 38f
OpenQuant, 9
Opportunity cost, suffering, 163
Optimal leverage, 121–122
calculation, 13
OptionMetrics, Ivy DB databases, 144–145
Options
  absence, 120–126
  crude oil futures options straddles, shorting, 135e–137e
  factors, 46–47
  historical data, 4
  prices, usage, 41–47
  strategies, 119
ORATS.com, daily options data, 4
Order book
  examples, 165t, 167t
  imbalance, 196
  ISO walking process, 169e–170e
Order book imbalance of a stock (ORCL), 196
Order flow, 175, 186–195, 207–208
determination, bulk volume
classification (usage), 187b–188b
  payment, 10
  strategy, examples, 189e–194e, 208e–210e
Order Protection Rule, 164, 169, 171
requirements, 170
Order status (latency type), 160
Order submission (latency type), 160
Order type, 163–174
Out-of-bag observations, 94
Out-of-sample CAGR
  increase, 38
  production, 40
  zero level, 39
Out-of-sample computer hardware
  stocks, Kalman filter trading
  strategy (application), 75f
Out-of-sample data, 100
Out-of-sample EWC-EWA, Kalman filter
  trading strategy (application), 79f
Out-of-sample performance,
  degrading, 95
Out-of-the-money (OTM), 138
  call/put options, 41–42
  call/SPX options, tenor duration,
    127–128
  options, embedded leverage level, 153
  OTM put–At-the-money (ATM) call
  implied volatilities, 46–47
Outright market, BBO (usage), 185t
OZS option, trading, 134

P

Parallel processing, 169
Parametric distributions, 104
Parent node, 89
Partial derivatives, usage, 21
Pearson correlation, 54
Perceptron (sigmoidal) function, 100–101
Performance
  attribution, 30
  measures, computation, 111
  metrics, 1, 12–14
  speed, comparison, 7
PFGBest bankruptcy, 10
Pipeline, 173
PIV. See Put implied volatility
Point of Presence (POP), 161
Polynomial function, specification,
  100–101
Portfolio
  optimization, 14–22
    Markowitz portfolio optimization
    method, 14
  Sharpe ratio, maximization, 20b–21b
    vega-neutral, 142–143
POSIT, 171–172
Post No Preference Blind Order, 164
  predict function, 36
Prediction error minimization, 107
Predictive factors, 201
Predictive power, 54
Predictive statistical factors (discovery),
  Principal Component Analysis
    (usage), 51e–53e
Predictors
  matrix, 113
  sampling, 95
Price data, 31
Price protection, 172
Price-to-earnings (P/E) ratio, 33, 35, 86
Principal Component Analysis (PCA),
  usage, 51e–53e
Profitable strategies, cessation
  (reasons), 219
Profit & Loss (P&L), 139
  computation, 149–151
  negative level, generation, 138
  responsibility, 217–218
Profit & Loss (P&L), usage, 12–13, 114
Progress Apama, 162
Put-call parity, 153
  deviation, 45–46
Put implied volatility (PIV), 42
Put-call implied volatilities, 45–46
PyCharm, 7
Python, usage, 3, 7–9, 162

Q

Quadprog (programming function),
  usage, 18
Quadratic programming, usage, 17b–20b
Qualcomm, Inc., 105
Quandl, 6, 112
  data, 4–5
  database, 34
  usage, 3
QuantConnect, 9
QuantGo, 4, 6, 182, 189
QuantHouse, 6, 9, 161
Quantiacs, 222
Quantmasscapital, 222
Quantopian, 7, 9, 222
Quant trading, programming languages (ranking), 8t

R

Random forest, 95–96
Random noise, effect, 30
Random subspace, 85, 95–96
Random walk, 60
Ravenpak, news sentiment pack, 5
Realized volatility (RV), 130
Registered investment advisor (RIA), 224
Regression coefficient, 195
Regression tree, 89–92
boosting, effect, 97f
learner, bagging (application), 94 trading model, 91f, 95f

Return-on-equity (ROE) calculation, 37
factor model, fitting, 39e–41e
one-factor model, application, 39
Risk-free rate, 13
Riskless arbitrage, 28
Riskless portfolio, 16–17
Risk parity, 22
allocation method, 22
portfolio, 22
Risky assets, riskiness (comparison), 202t
Routing, 171–174
optimization, 163–174
rTree.m, 90
rTreeLSBoost.m, 97
rTree_SPX.m, 109–110
Rule 610, 211
R (language), usage, 7–8, 162
RV(t+1)–RV(t) strategy, cumulative returns, 130f
RV(t+1)–VIX(t), 131f strategy, 132f

S

SciPy packages, usage, 7
Scripting languages, 8
Securities and Exchange Commission (SEC), 224, 226 (endnote)
Securities Investor Protection Corporation (SIPC), 10–11
sellOrderBook, 179
Sequential processing, 169
Settlement risk, 11
Sharadar, Core U.S. Fundamentals database, 34
Sharpe ratio, 13–14, 36, 219
increase, 98
logarithm, maximization, 21
maximization, 19, 20b–21b
regression tree, boosting (impact), 97f value, 87, 114
Short interest, 47–48
Short options position, risk, 137
Short stocks, high/low implied moments, 43e–45e
shortStraddle_LO.m, 137, 140
Short volatility strategy, 219
Short VX, long SPY
  contrast, 121f
  performance comparison, 122t
Sigmoid functions, 105–106
Signal processing, 59
Sniffler orders, 178
SIPC. See Securities Investor Protection
  Corporation
SIP feed, 166, 183
Size-independent input factors, 112t
Skew, 46
  computation, 43–44
Skewness factor, 54
Slope, Kalman filter estimate (usage), 77f
Small-cap (SML) stocks, ROE one-factor
  model (application), 39
Smart betas, 28–29
SMB
  factor, 28, 30
  portfolio return, 31
  time-series factor, 33, 34
SNDK, 67
S&P Capital IQ (QuantHouse), 6, 161
SPLV, 22
SPX
  component straddles/index
    straddles (contrast),
    dispersion trading (usage),
    144e–152e
  dispersion trading, 143f
  index, 35
  stocks. sorting, 48
SPX component stocks
cross-validated regression
  tree, 111f
fundamental cross-sectional factor
  model, 37f
liquidity factor, 49f
one-factor model, 38f
statistical factor, 51f
stepwise regression, fundamental
  factors (usage), 114f
two-factor models, 38f

SPY
  actual return, 13
  levered compounded returns,
    comparison, 122e–123e
  long SPY, short VX
    contrast, 121f
    performance comparison, 122t
  out-of-sample accuracy, 127
  purchase, 104
  regression tree, 90f
  returns, 86
    prediction, 109
    prediction task, 105
  stepwise regression, out-of-sample
    performance, 88f
volatility, 121
  prediction, 128e–130e
XIV (hedge ratio determination),
  Kalman filter (usage), 125f
XIV-SPY roll return strategy,
  cumulative returns, 126f
SR Labs, 6, 162
State space models (SSMs), 71–80
  application, 73
  SSM_beta_EWA_EWC.m, 78
Statistical factors, 49–53
  chart, 51f
  usefulness, 50–51
Statistical significance, achievement, 88
stepwiserlm function, 87, 90
stepwiseLR.m, 90
Stepwise regression, 86–89, 106
  factors selection, 113t
  fundamental factors, usage, 114f
  out-of-sample performance, 88f
  result, 91
Stochastic indicators, 201
Stochastic oscillator, 86
Stocks
  daily historical data, 2–3
  equal dollar weight, 48
  index, buy-and-hold, 120–121
  momentum, 83
  options, portfolio (selection), 152
  returns
decomposition, 49–50
prediction, option prices (usage), 41–47
selection, application, 111–114
Stopping condition, meeting, 89
Straddles, 138
crude oil futures options straddles, shorting, 135e–137e
inception, 139
long position, 140
SPX component straddles/index straddles (contrast),
dispersion trading (usage), 144e–152e
usage, 140e–141e
Strangle, 139
shorting, 134
Strategy Studio (Lime Brokerage), 9
Strike price, 135–136
Sub-period log returns, 15
SunGard, Astec database, 5
Supervised training, usage, 85
Support vector machine (SVM), 99–101, 207
cross-validation, usage, 101f
illustration, 99f
Support vectors, set
(formation), 100
Survivorship-bias-free database, 31–32, 110
Swiss National Bank, pegging announcement, 11–12

T
Tail-dependence, 142
Tail risk, 22, 42
Tangency portfolio, 16–17
capital allocation, correspondence, 20
determination, 19
Technical analysis, 201
Tenor, 41–42
duration, 127–128, 135
satisfaction, 145
testRatio, usage, 107
Theta, value, 138
Thomson Reuters, 5, 6, 162
Thomson Reuters News Analytics, 5
Tickdata, 5, 184
Tick rule, 186–1817
Time-series analysis, 59–60
Time-series factors, 29–31, 33, 47, 53
construction, 50
Time-series mean reversion, 152–153
Time-series techniques, 203–204
Trades and quotes (TAQ) data, 182, 184, 209
Trade secrets, leakage, 224–225
Trade tick, 192
absence, 209
Trading
Bollinger band trading strategy, 205f
contracts trading dates, 136t
correlation trading, 142
dispersion trading, 142–152
intraday trading, 159
model, regression tree basis, 91f, 95f
platforms, 6–9
criteria, 9
quant trading, programming languages (ranking), 8t
service function, 218–219
strategies, 36, 113
AI basis, 85–86
success, question, 219–221
trends, 221–222
volatility, options (absence), 120–126
Training set
CAGR, 37
predictions, 96–97
replacement, 93–94
trainRatio, indication, 107
TreeBagger function, 94–96
Trending techniques, 102–103
Trends, 221–222
Twitter feeds, usage, 221–222
Two-factor loadings, usage, 38
Two-factor model, 37–40, 38f
INDEX

U
UMD, 31
  momentum, 31
Uninformed traders, identification, 174–175
Unit variance, 188f
Univariate regression, 54
Unlevered CAGR
  expectation, 45–46
  obtaining, 47
Unsupervised training, usage, 85
U.S. Dollars (USD)
  AR(10) model coefficients,
    application, 62t
  AR(10) trading strategy,
    usage, 62f
User acceptance testing (UAT) account,
  usage, 176–177
USMV, 22
USO, out-of-sample accuracy, 127

V
valRatio, zero level, 108
Value-minus-growth factor, 29
VAR, 59
VAR(1) model, application, 68t
VAR(1) trading strategy, application, 70f
Variances, reduction (absence), 89
VAR models, usage, 71
VAR(p), 67–71
vartovec function, 69
Vector error correction (VEC), 59, 69
  models, 71
  VEC(0) model, application, 70t
Vega
  computation, 148
  value, 138
Vega-neutrality, achievement, 144
Vega-neutral weighting scheme, 143
Viterbi, Andrew, 105
Volatility (VIX)
  change, 30, 31
  implied market volatility, daily change, 47
  increase, 27, 133
  index, 47
  prediction, 126–132
    example, 128e–130e
  realized volatility (RV), 130
  risk, exposure minimum, 30
  shorting, 120
  smirk, 46
  trading volatility, options (absence), 120–126
Vol, computation, 43–44
Volume bars, data structure, 192t
VX
  levered compounded returns,
    comparison, 122e–123e
  short VX, long SPY
    contrast, 121f
    performance comparison, 122t
VX-ES
  roll returns strategy, cumulative returns, 124f
  strategy, 123
VXX
  holding, 121
  purchase/sale, 129
VXX/XIV strategy, 127

W
Wavelet filter, 74
Weak learners, 95
“Weekly Petroleum Status Report,”
  133–135
White noise, 29
  k-dimensional Guassian white noise, 71
WRDS
  data, 3
  package, 4

X
XIV/SPY (hedge ratio determination),
  Kalman filter (usage), 125f
XIV-SPY roll return strategy, cumulative returns, 126f

Z
Zero mean, 188f