

Index

Symbols and Numerics

& (and) logical operator, 614
* (asterisk), 586
: (colon), 641
|| (concatenation) operator, 503–504
{ } (curly braces), 612, 616
/ (divide) operator, 614
\$ (dollar sign), 646
[] (double brackets), 610
== (double equal sign), 614, 648
= (equal) operator, 303–304, 508, 517
> (greater than) operator, 517, 614
>= (greater than or equal to) operator, 517, 614
(hash symbol), 640, 719
<- (is assigned, gets) operator, 614
-> (is assigned to) operator, 614
< (less than) operator, 517, 614
<= (less than or equal to) operator, 517, 614
- (minus) operator, 614
%% (modulo) operator, 614, 650
! (not) logical operator, 614
!= (not equals) operator, 614
<> (not equal) predicate, 517
| (or) logical operator, 614
+ (plus) operator, 614
^ (power) operator, 614
[] (single brackets), 610
~ (tilde) operator, 667
* (times) operator, 614
1:1 (one-to-one) relationship, 171
1NF (Codd's First Normal Form), 472–473

A

ABS function, 499–500
Absolute mode, for macros, 77–78
Access category (Tableau), 292

access management, in Power BI, 253–254
accessing
 macros, 77
 reports, 133–136
 Tableau Bridge Client, 364
ACOS function, 500
actual cardinality, 506
adaptive models, 128
add_edge() function, 758, 760
add_feature() method, 756
adding
 annotations in Matplotlib, 733–735
 color to dashboards in Tableau, 377–378
 content to dashboards, 242–244
 context
 about, 104–105
 with annotations, 105, 106
 with data, 105
 with graphical elements, 105, 106
 data
 to imported, DirectQuery, and composite models, 196–197
 in tables, 195–196
 fonts to dashboards in Tableau, 377–378
 grids in Matplotlib, 727–728
 labels in Matplotlib, 733–734
 legends in Matplotlib, 733–734, 735–737
 markers in Matplotlib, 731–733
 metadata to worksheets in Tableau Desktop, 357
 objects to dashboards in Tableau, 376–383
 tables, 191–192
 text to dashboards in Tableau, 377–378
add_node() function, 758, 760
ADLS (Microsoft Azure Data Lake Storage), 43, 45
aes() function, 672
aesthetic mapping, 672–673
aggregate functions, 533

- Aggregate option, 325
- aggregating
 - data using GROUP statement, 69
 - in Tableau, 273
 - tables, 179
- AI (artificial intelligence)
 - Q&A feature and, 231–232
 - as a role of big data, 26
- Airtable, 57
- Alerts and Notifications feature (Tableau Cloud), 428
- aligning report pages, 235
- ALL clause, 523–525
- all-inclusive, 351
- alternative hypotheses, 628–629
- Amazon Simple Storage Service (S3), 43, 45
- Amazon Web Services (AWS), 30, 44
- Analysis menu (Tableau Desktop), 359–360
- analytic data, operational data compared with, 47–48
- analyzing
 - big data with enterprise business intelligence practices, 39–40
 - data
 - in Power BI Desktop, 129–130
 - for trends, 12
 - networks with statnet and igraph, 621–622
 - spatial point patterns with spatstat, 622
- AND connective, 531–532
- and (&) logical operator, 614
- Animation option (Format menu of Tableau Desktop), 361, 362
- annotate() function, 735
- annotations
 - adding context with, 105, 106
 - adding in Matplotlib, 733–735
- anomalies, detecting, 162–163
- ANY clause, 523–525
- Apache Flume, 29
- Apache Hadoop, 156
- Apache Kafka, 29
- Apache Sqoop, 29
- append() function, 643
- appending queries, 175–176
- application file format, 53
- applying
 - conditional formatting, 233
 - constraints
 - about, 479
 - assertions, 484–485
 - column constraints, 480–481
 - foreign key constraints, 483–484
 - table constraints, 481–483
 - data science to subject areas, 55–56
 - interactive elements in Tableau Desktop, 356
 - mathematical modeling to data science tasks, 54
- approximate numerics, 464–466
- apps, accessing reports from, 134–135
- area boundary polygon, 102
- area charts, 91, 92, 213–214, 408–410
- ARIMA (AutoRegressive Integrated Moving Average) time series forecasting, 619
- arranging data, in data models
 - about, 199
 - hiding data, 200–201
 - sorting by/grouping by, 200
- ARRAY type, 474
- array value, 506
- artificial intelligence (AI)
 - Q&A feature and, 231–232
 - as a role of big data, 26
- “as a service” databases, 127
- ASIN function, 500
- as-is statement, 176
- assertions, 484–485
- asterisk (*), 586
- ATAN function, 500
- attributes, 617, 711
- attributes() function, 619
- automatic relationships, creating, 197–198
- automating
 - data flows, 334
 - Excel tasks with Macros, 77–79
- AutoRegressive Integrated Moving Average (ARIMA) time series forecasting, 619

AVG function, 493

axes

 formatting in Matplotlib, 726–727

 representing time on, 749–750

Azure Streams, 242

Azure Synapse, 30

B

background, for report pages, 235

Background Images option (Map menu in Tableau Desktop), 361

Background Layers option (Map menu in Tableau Desktop), 361

Background Maps option (Map menu in Tableau Desktop), 361

bar graphs/charts/plots

 about, 92, 93

 clustered, 211–212

 combining with line charts, 215

 in ggplot2, 675–676, 679–683

 in Matplotlib, 740–741

 in Power BI, 209–213, 215

 in R, 653–654, 660–661, 663–666

 side-by-side, 400

 stacked, 209–211, 212–213, 400

 standard, horizontal, 400

 in Tableau, 400–401

barplot() function, 660

Beginning Programming with Python For Dummies, 3rd Edition (Mueller), 692, 697

Best Practice icon, 4

best practices, for dashboards in Tableau, 382–383

BETWEEN clause, 517–519

big data

 about, 19

 analyzing with enterprise business intelligence practices, 39–40

 connecting with business intelligence, 39

 importance of data, 30–35

 overhyping, 28

 roles of

 about, 20

 artificial intelligence (AI), 26

 data reuse, 26–27

 in data science and engineering, 36–38

 decision-making, 22

 insight management, 24–25

 measuring, 23

 monitoring, 23–24

 operations, 20–21

 problem solving, 26

 reporting, 25–26

 strategy, 21–22

 sources of, 35

 the three Vs, 27, 28–30

BIGINT type, 463–464

BINARY LARGE OBJECT (BLOB) type, 469

binary large objects (BLOBs), 46

binary strings, 468–469

BINARY type, 468

BINARY VARYING type, 468

binning, in Tableau Prep Builder, 343–344

blending data, from multiple sources, 307–308

BLOBs (binary large objects), 46

blue pill, in Tableau, 272

Boolean values, 469, 506

box and whisker plots

 in ggplot2, 686–688

 in Matplotlib, 743–744

 in R, 656–657, 669–670

 in Tableau, 415–416

boxplot() function, 670, 743

Brainstorm step, in designing for target audience, 85–86

Breakdown (Visualizations pane), 218

bubble plots, 94, 95

bubbles, in Tableau, 402–406

buckets, 501

built systems, 38

built-in functions, in Python, 709–711

bullet charts, in Tableau, 418–419

business intelligence (BI)

 about, 12–13

 connecting big data with, 39

buttons
 integrating in dashboards in Tableau, 380–381
 story workspace, 385, 386
Buxton, Stephen (author)
 Querying XML, 472
byte, 498

C

C++, 38
c() function, 638–639, 641
Calculated Fields section (Analysis menu of Tableau Desktop), 360
call to action, in stories in Tableau, 384–385
capacities, in Power BI, 113
captions, resizing in stories, 389
cardinality, 499, 506
CARDINALITY function, 499, 506
cards
 as indicators, 223–224
 in Tableau Desktop, 352
Cartesian product joins, 592–593
Cartopy, 753–756
CASE expression, 507–509
CAST expression, 510–511
cast() function, 680
CAST function, 510–511
catalogs
 in relational database hierarchy, 452
 relational databases, 452
Categories (Visualizations pane), 218
categorizing fields, in Tableau Prep Builder, 340–342
cbind() function, 648
CEIL(ING) function, 501
changing
 data in tables, 195–196
 data to imported, DirectQuery, and composite models, 196–197
 root tables to data relationships, 300–301
 sensitivity labels, 263
CHARACTER LARGE OBJECT (CLOB) type, 467
character strings, 464–468
CHARACTER type, 466
CHARACTER VARYING type, 467
character vector, 641
CHARACTER_LENGTH function, 498
charts. *See also* graphs
 about, 25
 in Excel, 71, 74–75
Cheat Sheet (website), 4
CHECK constraint, 481
checking data structures and column properties, 163–164
Choose most functional type for purpose step, in designing for target audience, 86–87
choosing
 data graphic type
 about, 90–91
 comparative graphics, 94–95
 spatial plots and maps, 101–103
 standard chart graphics, 91–94
 statistical plots, 95–99
 topology structures, 100–101
 data visualizations, 205–206
 design style, 87–90
 graphs in Matplotlib, 740–745
 Tableau versions, 281
 tools for data science strategies, 57–58
circles, in Tableau, 402–406
clarity, visualizing for data, 12
classes, in Python, 711–713
Clean Step option, 325
Cleanliness category (Tableau), 293
cleansing data, 162–165
Clear option (Format menu of Tableau Desktop), 361, 362
Clear tool, on toolbar in Tableau Desktop, 367
clinical information scientists, 56
clipboard data, working with, 309–311
cloropleth, 101–102
cloud providers, data storage pricing for, 44
clustered bar charts, 211–212
clustered column charts, 211–212
clustered index scan, 542

- clustering indexes, 603–604
- COALESCE expression, 509–510
- coastlines() method, 756
- Codd's First Normal Form (1NF), 472–473
- code listings
 - aggregating data using GROUP statement, 69
 - assigning objects and concatenating in R, 610
 - defining classes in Python, 712
 - inner JOIN() function, 68
 - narrowing results using HAVING clause, 69
 - objects in R, 617–618
 - outer JOIN() function, 68
 - querying data using SELECT(), WHERE(), and DATE(), 66–67
- coding, 55
- collaborating and sharing
 - in Power BI
 - about, 247
 - access management, 253–254
 - configuring workspaces, 250–253
 - creating workspaces, 250–253
 - data lineage, 258–262
 - data protection, 262–264
 - dataflows, 261–262
 - datasets, 261–262
 - managing data, 254–258
 - Microsoft Excel, 256
 - paginated reports, 258
 - Quick Insights, 256–257
 - settings and storage, 254
 - Usage Metrics reports, 257
 - workspaces, 247–254
 - in Power BI Services, 137–138
 - results, 133
 - in Tableau Cloud
 - about, 435–436
 - collections, 437–439
 - Explore feature, 439–440
 - recommendations engine, 440–441
 - shared with me, 436–437
 - sharing content, 435–436
- collections
 - of data, 12, 53–54
 - in Tableau Cloud, 437–439
 - types for, 473–474
- colnames() function, 676
- colon (:), 641
- color coding, in Tableau, 304
- colors
 - adding to dashboards in Tableau, 377–378
 - using in Matplotlib, 730–731
- column charts
 - clustered, 211–212
 - stacked, 209–211, 212–213
- column constraints, 480–481
- column data
 - evaluating and transforming types
 - about, 171
 - combining queries, 174–180
 - finding and creating keys for joins, 171–173
 - shaping column data to meet Power Query requirements, 173–174
 - shaping to meet Power Query requirements, 173–174
- column indexes, 64–65
- column properties, checking, 163–164
- column values, 488
- column-name joins, 597–598
- columns
 - deleting in tables, 193–195
 - in relational database hierarchy, 452
 - renaming in tables, 193–195
 - in Tableau Prep Builder, 339–340
- combining
 - data sources in Tableau, 294–296
 - queries, 174–180
 - SQL with host languages, 457
- comma-separated values (CSV) file format, 53
- comments, in R, 640
- comparative graphics, 94–95
- comparison operators, 560–563, 568–570
- comparison predicates, 516, 517

- complex data visualization, 228–231
- complex format, in Python, 706
- components, 45, 641
- composite data models, adding and modifying data to, 196–197
- composite storage models, 150–151
- Comprehensive R Archive Network (CRAN), 619, 630–631
- concatenate() function, 610–611, 743
- concatenation (| |) operator, 503–504
- concurrency control, 48
- condition joins, 596–597
- conditional formatting
 - applying, 233
 - in Excel, 71, 72–74
- conditional value, 507–510
- configuring
 - dashboards
 - in Power BI, 240
 - in Tableau, 372–374
 - data flows, 317–318
 - queries for data loading, 180–181
 - report pages, 233–235
 - report visualizations, 232–239
 - sensitivity labels, 263
 - workspaces, 250–253
- connecting
 - big data with business intelligence, 39
 - to online sources, 160
 - relational databases, 452–453
- constraints
 - applying
 - about, 479
 - assertions, 484–485
 - column constraints, 480–481
 - foreign key constraints, 483–484
 - table constraints, 481–483
 - designing, 63–64
- consuming data, 53–54
- CONTAINS filter, 330
- content, adding to dashboards, 242–244
- context, adding
 - about, 104–105
 - with annotations, 105, 106
 - with data, 105
 - with graphical elements, 105, 106
- continuation prompt, 640
- continuous compared with discrete data, in Tableau, 272
- Continuous field type, 341
- continuous line chart, 406
- CONVERT function, 496
- converting
 - data types with CAST expression, 510–511
 - data visualizations to cross tabulation tables in Tableau, 419–422
- copying data, 48–49
- Copy/Paste Formatting option (Format menu of Tableau Desktop), 361, 362
- core-based licensing model (Tableau), 283
- cores consumed, 284
- correlated subqueries, 566–571, 579–584
- correlations
 - defined, 218
 - displaying in Matplotlib, 747–748
- COS function, 500
- COSH function, 500
- COUNT function, 492–493
- CRAN (Comprehensive R Archive Network), 619, 630–631
- creating
 - automatic relationships, 197–198
 - bar charts in Tableau, 400
 - bullet charts in Tableau, 418
 - circles in Tableau, 403–404
 - continuous line charts in Tableau, 407–408
 - dashboards, 240–241
 - data lakes, 42
 - data models
 - about, 17, 183–184
 - adding and modifying data to imported, DirectQuery, and composite models, 196–197
 - Data view, 184–186
 - defining data types, 188–189
 - importing queries, 186–188

- managing formatting and data type properties, 189–191
- managing tables, 191–196
- Modeling view, 184–186
- in Power BI Desktop, 128–129
- data relationships, 297–298
- data science pipeline, 692–693
- data visualizations, 204–205
- directed graphs, 759–760
- dual combination charts in Tableau, 411
- Gantt charts in Tableau, 417
- heat maps in Tableau, 394–395
- highlight tables in Tableau, 395–396
- items in Power BI Desktop, 130–131
- keys for joins, 171–173
- manual relationships, 198–199
- pie charts in Tableau, 398–399
- PivotTables, 76
- published data sources, 334–336
- scatter plots in Tableau, 413–414
- stacked area charts in Tableau, 410
- stories in Tableau
 - about, 383, 386–389
 - evaluating story workspace, 385–386
 - formatting, 389–390
 - planning, 384–385
 - synthesizing data through Tableau story, 383–384
- template apps, 138
- treemaps in Tableau, 402
- undirected graphs, 757–758
- worksheets in Tableau Desktop, 355, 356
- workspaces, 250–253
- Creator bundle (Tableau), 281, 284
- CRM (customer relationship management), 153
- cross joins, 592–593
- CSV (comma-separated values) file format, 53
- .csv files, 321
- CURRENT_TIMESTAMP function, 502
- curly braces {}, 612, 616
- Current Layout section (Dashboard menu of Tableau Desktop), 358

- CURRENT_DATE function, 502
- CURRENT_TIME function, 502
- CURRENT_USER, 491
- customer relationship management (CRM), 153
- Customization category (Tableau), 293
- customizing dashboards, in Tableau, 374–376
- Cycle Fields/and Swap Rows and Columns section (Analysis menu of Tableau Desktop), 360

D

- Dashboard menu (Tableau Desktop), 357–358
- Dashboard pane, in Tableau, 373–374
- dashboards
 - accessing reports from, 136
 - adding content, 242–244
 - best practices for, in Tableau, 382–383
 - configuring, 240
 - creating, 240–241
 - data visualization and
 - about, 239–240
 - adding content, 242–244
 - configuring, 240
 - creating new, 240–241
 - pinning reports, 245–246
 - embedding reports on, 138
 - extending in Tableau, 381–382
 - Power BI, 117
 - resizing in stories, 390
 - in Tableau
 - about, 274, 371–372
 - adding objects to, 376–383
 - configuring, 372–374
 - customizing, 374–376
 - in Tableau Desktop, 369–370
 - working with in Power BI Services, 136–137
- data
 - about, 9–10, 167
 - adding
 - context with, 105
 - to imported, DirectQuery, and composite models, 196–197
 - aggregating using GROUP statement, 69

data (*continued*)

- analyzing in Power BI Desktop, 129–130
- arranging in data models
 - about, 199
 - hiding data, 200–201
 - sorting by/grouping by, 200
- business intelligence, 12–13
- cleansing, 162–165
- collecting, 53–54
- configuring queries for data loading, 180–181
- consuming, 53–54
- copying, 48–49
- data analysis, 16–17
- data analytics, 13–14
- data management, 15
- defined, 127
- evaluating and transforming column data types
 - about, 171
 - combining queries, 174–180
 - finding and creating keys for joins, 171–173
 - shaping column data to meet Power Query requirements, 173–174
- filtering, 207–208
- getting
 - to know using Excel, 71–75
 - from Microsoft-based file systems, 151–153
 - from sources, 142–146
- hiding, 200–201
- importance of, 30–35
- information compared with, 33–34
- ingesting in Power BI Desktop, 125–126
- joining input steps and, 323–324
- latency of, 49
- lifecycle of, 167–168
- loading
 - about, 162–165
 - in Python, 695
- modifying to imported, DirectQuery, and composite models, 196–197
- normalizing, 348–349
- pivoting with, 345–348
- preparing for data visualization

- about, 141
- checking data structures and column properties, 163–164
- cleansing data, 162–165
- data statistics, 164–165
- detecting anomalies and inconsistencies, 162–163
- getting data from Microsoft-based file systems, 151–153
- getting data from sources, 142–146
- loading data, 162–165
- managing data source settings, 146–147
- relational data sources, 153–162
- shared vs. local datasets, 147–150
- storage and connection modes, 150–151
- transforming data, 162–165
- protecting when sharing and collaborating, 262–264
- qualitative, 23
- quantitative, 23
- querying
 - about, 53–54
 - using `SELECT()`, `WHERE()`, and `DATE()`, 66–67
- raw, 16
- reformatting using PivotTables, 75–77
- refreshing
 - in input step, 318
 - in Power BI Services, 138–139
- resolving errors
 - during data import, 182
 - during importing of, 182
- resolving inconsistencies
 - about, 168
 - finding root cause, 170
 - removing rows using Power Query, 170
 - replacing values, 168–170
- semi-structured, 11
- storing, 30, 31–32
- structured, 10–11
- summarizing using PivotTables, 75–77
- transforming, 162–165
- types of, 45–46

- unstructured, 11
- visualizing, 17
- data access, Python for, 697–698
- data analysis, 16–17
- data analytics, 13–14
- data art, 84
- data cleansing
 - about, 13
 - options for, 332–333
 - in Tableau, 279
- data connections (Tableau)
 - about, 285, 288–289
- data relationships
 - about, 296
 - advantages of, 296–297
 - changing root table of, 300–301
 - creating, 297–298
 - disadvantages of, 297
 - editing, 299
 - moving tables to create different, 299–300
 - removing tables from, 301–302
- data sources
 - combining, 294–296
 - options for, 286–288
 - setting up, 292–293
- joining data
 - about, 302
 - blending data from multiple sources, 307–308
 - clipboard data, 309–311
 - creating joins, 304–306
 - join types, 302–303
 - joining fields that contain null values, 306
 - setting up join clauses, 303–304
- requesting new connectors, 291
- Tableau Desktop, 289–290
- Tableau Online, 290–291
- Tableau Prep, 289–290
- Tableau Server, 290–291
- data display, Python for, 698
- data engineering, 36–38
- data engineers, 52
- data fields, in Tableau, 270
- data flows
 - about, 314–316
 - automating, 334
 - configuring, 317–318
 - filtering, 329–333
 - grouping, 327–328
 - input step, 318–324
 - nurturing, 324–327
- data frames
 - about, 610
 - in R, 646–649
- data governance, 15
- data graphics
 - selecting types
 - about, 90–91
 - comparative graphics, 94–95
 - spatial plots and maps, 101–103
 - standard chart graphics, 91–94
 - statistical plots, 95–99
 - topology structures, 100–101
 - testing, 103–104
- data insights
 - defined, 52
 - deriving from statistical methods, 55
- data integrity, 62
- data journalists, 56
- data lakes
 - about, 30, 41–42
 - building, 42
 - components of, 45
 - copying data, 48–49
 - data services for, 43
 - data types, 45–46
 - enterprise-data warehouse, 49–50
 - expanding, 43–45
 - operational vs. analytic data, 47–48
- data life cycle, Tableau versions by, 282–283
- data lineage, for sharing and collaborating, 258–262
- data loading, configuring queries for, 180–181
- data management, 15
- data mart, 30

- Data menu (Tableau Desktop), 354–355
- data mining, 13
- data models
 - about, 183
 - arranging data
 - about, 199
 - hiding data, 200–201
 - sorting by/grouping by, 200
 - creating
 - about, 17, 183–184
 - adding and modifying data to imported, DirectQuery, and composite models, 196–197
 - Data view, 184–186
 - defining data types, 188–189
 - importing queries, 186–188
 - managing formatting and data type properties, 189–191
 - managing tables, 191–196
 - Modeling view, 184–186
 - in Power BI Desktop, 128–129
 - defined, 129
 - managing relationships
 - about, 197
 - creating automatic relationships, 197–198
 - creating manual relationships, 198–199
 - deleting relationships, 199
 - publishing, 201–202
- data presentation, 13
- data processing. *See* information technology (IT)
- data product managers, 56
- data relationships
 - about, 296
 - advantages of, 296–297
 - changing root tables of, 300–301
 - creating, 297–298
 - disadvantages of, 297
 - editing, 299
 - moving tables to create different, 299–300
 - removing tables from, 301–302
- data reuse, as a role of big data, 26–27
- data schema, 11
- data science
 - about, 36–37, 51–52
 - data insights
 - about, 52–53
 - applying data science to subject areas, 55–56
 - applying mathematical modeling to data science tasks, 54
 - coding, 55
 - collecting data, 53–54
 - consuming data, 53–54
 - deriving from statistical methods, 55
 - querying data, 53–54
 - database design
 - about, 62–63
 - defining data types, 63
 - designing constraints, 63–64
 - normalizing databases, 64–66
 - Excel
 - about, 70
 - automating tasks with macros, 77–79
 - getting to know data, 71–75
 - reformatting with PivotTables, 75–77
 - summarizing with PivotTables, 75–77
 - Python and
 - about, 691–692
 - creating data science pipeline, 692–693
 - learning Python, 695–696
 - role of Python in data science, 693–695
 - using Python ecosystem for data science, 699–702
 - working with Python, 697–699
 - relational databases
 - about, 58–60
 - keys, 61–62
 - role of big data in, 36–38
 - selecting tools for, 57–58
 - SQL databases
 - about, 58–60
 - keys, 61–62
 - SQL functions, 66–69
 - using Python for, 703–705

Data Science For Dummies, 3rd Edition (Pierson), 54, 55

- data scientists, 693–694
- data services, for data lakes, 43
- data showcasing, 83
- Data Source tab, in Tableau Desktop, 353
- data sources
 - blending data from multiple, 307–308
 - creating published, 334–336
 - getting data from, 142–146
 - managing settings for, 146–147
 - publishing using Tableau Cloud, 336
 - relational, 153–162
 - in Tableau
 - about, 268–269, 286–288
 - combining, 294–296
 - setting up, 292–293
- data statistics, 164–165
- data storytelling, 82–83
- data structures
 - about, 337–338
 - categorizing fields, 340–342
 - checking, 163–164
 - columns, 339–340
 - for data visualization
 - about, 342
 - binning, 343–344
 - distributions, 344–345
 - histograms, 343–344
 - outliers, 344–345
 - pivoting with data, 345–348
 - fields, 339–340
 - in Python, 707
 - records, 338–339
 - rows, 338–339
- data types
 - converting with CAST expression, 510–511
 - for data lakes, 45–46
 - defining, 63, 188–189
 - managing properties for, 189–191
 - in Python
 - about, 705–706
 - dictionaries, 708
 - lists, 707
 - numbers, 706
 - sets, 708
 - strings, 706–707
 - tuples, 707–708
 - in R, 626
 - in SQL
 - about, 461–462
 - approximate numerics, 464–466
 - ARRAY type, 474
 - BIGINT, 463–464
 - BINARY LARGE OBJECT (BLOB) type, 469
 - binary strings, 468–469
 - BINARY type, 468
 - BINARY VARYING type, 468
 - Booleans, 469
 - CHARACTER LARGE OBJECT (CLOB) type, 467
 - character strings, 466–468
 - CHARACTER type, 466
 - CHARACTER VARYING type, 467
 - collection types, 473–474
 - DATE type, 469
 - datetimes, 469–471
 - DECFLOAT type, 464
 - DECIMAL type, 464
 - distinct, 476
 - DOUBLE PRECISION type, 465–466
 - exact numerics, 462
 - FLOAT type, 466
 - INTEGER, 462–463
 - intervals, 471
 - NATIONAL CHARACTER LARGE OBJECT type, 467–468
 - NATIONAL CHARACTER type, 467–468
 - NATIONAL CHARACTER VARYING type, 467–468
 - NUMERIC type, 464
 - REAL type, 465
 - REF types, 475
 - ROW type, 472–473
 - SMALLINT, 463
 - structured, 477–478

data types (*continued*)

- TIME WITH TIME ZONE type, 470
- TIME WITHOUT TIME ZONE type, 470
- TIMESTAMP WITH TIME ZONE type, 471
- TIMESTAMP WITHOUT TIME ZONE type, 470–471
- user-defined types, 475–478
- XML type, 471–472

in Tableau, 269–270

data values, entering

- about, 487–488
- column values, 488
- literal values, 488–489
- row values, 488
- variables, 490–491

Data view, 184–186

data visualizations

- about, 81–82, 227
- adding context
 - about, 104–105
 - with annotations, 105, 106
 - with data, 105
 - with graphical elements, 105, 106
- artificial intelligence (AI) and Q&A feature, 231–232

complex, 228–231

dashboards

- about, 239–240
- adding content, 242–244
- configuring, 240
- creating new, 240–241
- pinning reports, 245–246

data structures for

- about, 342
- binning, 343–344
- categorizing fields, 340–342
- distributions, 344–345
- histograms, 343–344
- outliers, 344–345

designing for target audience

- about, 84–85
- Brainstorm step, 85–86

Choose most functional type for purpose step, 86–87

Define the purpose step, 86

in Matplotlib

- about, 739
- bar charts, 740–741
- boxplots, 743–744
- depicting groups, 746–747
- displaying correlations, 747–748
- histograms, 741–743
- plotting geographical data, 752–756
- plotting time series, 748–752
- scatterplots, 744–745, 746–748
- selecting graphs, 740–745
- visualizing graphs, 757–760

in Power BI

- about, 203
- area charts, 213–214
- bar charts, 209–213, 215
- column charts, 209–213
- creating visualizations, 204–205
- Donut charts, 219, 220
- filtering data, 207–208
- Funnel charts, 217–218
- indicators, 222–225
- line charts, 213–214, 215
- maps, 221–222
- Pie charts, 219, 220
- reports, 203–208
- Ribbon charts, 216
- Scatter charts, 218–219
- selecting visualizations, 205–206
- treemaps, 219–221
- types, 209–225
- Waterfall charts, 216–217

preparing data for

- about, 141
- checking data structures and column properties, 163–164
- cleansing data, 162–165
- data statistics, 164–165

- detecting anomalies and inconsistencies, 162–163
- getting data from Microsoft-based file systems, 151–153
- getting data from sources, 142–146
- loading data, 162–165
- managing data source settings, 146–147
- relational data sources, 153–162
- shared vs. local datasets, 147–150
- storage and connection modes, 150–151
- transforming data, 162–165
- for reports
 - about, 232–233
 - applying conditional formatting, 233
 - configuring report page, 233–235
 - exporting reports, 235–236
 - preparing reports for distribution, 236–239
- selecting data graphic type
 - about, 90–91
 - comparative graphics, 94–95
 - spatial plots and maps, 101–103
 - standard chart graphics, 91–94
 - statistical plots, 95–99
 - topology structures, 100–101
- selecting design style, 87–90
- in Tableau
 - about, 391–392
 - area chart, 408–410
 - bar chart, 400–401
 - box and whisker plot, 415–416
 - bubbles, 402–406
 - bullet chart, 418–419
 - circles, 402–406
 - converting to cross tabulation tables, 419–422
 - dual combination chart, 411–412
 - Gantt chart, 416–418
 - heat map, 394–395
 - highlight table, 395–396
 - histogram, 414–415
 - line chart, 406–408
 - maps with/without symbols, 396–398
 - pie chart, 398–399
 - publishing, 422–424
 - scatter plot, 412–414
 - text table, 392–394
 - treemap, 401–402
- table-based, 228–231
- testing data graphics, 103–104
- types of
 - about, 82
 - data art, 84
 - data showcasing, 83
 - data storytelling, 82–83
- data warehouse, 30, 49
- databases
 - “as a service,” 127
 - creating unions among, 320–324
 - designing, 62–66
 - nonrelational, 156–158
 - normalizing, 64
 - in Power BI Desktop, 126–128
 - relational
 - about, 58–62, 155–156, 445
 - catalogs, 452
 - connections, 452–453
 - functional dependencies, 447–448
 - hierarchy of, 452
 - keys, 61–62, 448–450
 - multisets, 446–447
 - object-extended, 46
 - paths, 454
 - privileges, 451
 - relational database hierarchy, 452
 - relational model and, 445–446
 - relations, 446–447
 - routines, 453
 - schemas, 451
 - sessions, 452–453
 - sets, 446–447
 - tables, 446–447
 - transactions, 452–453
 - users, 450
 - views, 450
- SQL, 58–62

- dataflows
 - about, 114–115
 - for sharing and collaborating, 261–262
- DataFrame objects, 718
- datasets
 - about, 115
 - shared vs. local, 147–150
 - for sharing and collaborating, 261–262
 - types for data, 138–139
- Dataverse button (Home Ribbon), 185
- date data type, 63
- DATE () function, querying data using, 66–67
- DATE type, 469
- datetime value functions, 502
- datetimes, 469–471, 504–505
- datum, 30
- DECFLOAT type, 464
- DECIMAL type, 464
- decision engines, 607
- decision-making, 13, 22
- decomposition trees, 230
- dedicated capacity, 113
- Define the purpose step, in designing for target audience, 86
- defining
 - data types, 63, 188–189
 - line appearance in Matplotlib
 - about, 729
 - adding markers, 731–733
 - line styles, 729–730
 - using colors, 730–731
 - plots in Matplotlib, 722–723
- DELETE statement, using subqueries in, 571–574
- deleting
 - columns in tables, 193–195
 - relationships, 199
 - sheets in Tableau Desktop, 370
 - stories, 390
 - tables, 193
- dependencies, multiple, 64
- dependent variables, in R, 625
- depicting groups, in Matplotlib, 746–747
- descriptive data analytics, 13
- descriptive statistics, 624
- design style, selecting, 87–90
- designing
 - databases, 62–66
 - for target audience
 - about, 84–85
 - Brainstorm step, 85–86
 - Choose most functional type for purpose step, 86–87
 - Define the purpose step, 86
- detecting anomalies and inconsistencies, 162–163
- determinant, 447
- Developer tab (Excel), 77
- developing
 - automatic relationships, 197–198
 - bar charts in Tableau, 400
 - bullet charts in Tableau, 418
 - circles in Tableau, 403–404
 - continuous line charts in Tableau, 407–408
 - dashboards, 240–241
 - data lakes, 42
 - data models
 - about, 17, 183–184
 - adding and modifying data to imported, DirectQuery, and composite models, 196–197
 - Data view, 184–186
 - defining data types, 188–189
 - importing queries, 186–188
 - managing formatting and data type properties, 189–191
 - managing tables, 191–196
 - Modeling view, 184–186
 - in Power BI Desktop, 128–129
- Device Layouts section (Dashboard menu of Tableau Desktop), 357, 358
- diagnostic data analytics, 14
- dictionaries, in Python, 708
- DiGraph () function, 759–760
- dim () function, 644
- Dimension field type, 341
- dimensions, in Tableau, 270–272
- directed graphs, creating, 759–760

- directors of data science, 56
- DirectQuery
 - about, 138–139, 150–151
 - adding and modifying data to data models, 196–197
- Discrete field type, 341
- discrete line chart, 407
- displaying correlations, in Matplotlib, 747–748
- DISTINCT clause, 526
- distinct number, 164
- distinct types, 476
- distributions
 - graphing in R, 652–653
 - preparing reports for, 236–239
 - in Tableau Prep Builder, 344–345
- divide (/) operator, 614
- dollar sign (\$), 646
- domain-specific, 36
- Donut charts, 219, 220
- dot charts
 - in ggplot2, 676–679
 - in R, 662–663
- dot notation, 713
- double brackets ([[]]), 610
- double equal sign (==), 614, 648
- DOUBLE PRECISION type, 465–466
- downloading R, 608
- drawing multiple lines and plots, in Matplotlib, 723–724
- draw_networks() function, 760
- dual combination chart, in Tableau, 411–412
- dual mode, 151
- dual-line chart, 407
- Duplicate Field option, 329
- Duplicate tool, on toolbar in Tableau Desktop, 367
- DurusWorks, 698

E

- edges, 757
- Edit Locations option (Map menu in Tableau Desktop), 361
- editing
 - data relationships, 299
 - reports in Power BI Services, 132–136
- Editing view (Power BI), 116
- ELT (extraction, transformation, and loading), 48–49
- embedded SQL, 457–460
- embedding reports and dashboards, on websites, 138
- enabling updates, in Tableau Desktop, 357
- Encapsulated Postscript (ES), 724
- Enter Data button (Home Ribbon), 185
- entering data values
 - about, 487–488
 - column values, 488
 - literal values, 488–489
 - row values, 488
 - variables, 490–491
- enterprise business intelligence practices, analyzing big data with, 39–40
- enterprise resource planning (ERP), 153
- enterprise-data warehouse, 49–50
- environments, Notebook, 753–754
- equal (=) operator, 303–304, 508, 517
- equi-joins, 594–596
- ERP (enterprise resource planning), 153
- errors
 - resolving during data import, 182
 - types of, 629–630
- ES (Encapsulated Postscript), 724
- ETL (Extract, Transport, Load) tool, 49, 313, 314
- evaluating and transforming column data types
 - about, 171
 - combining queries, 174–180
 - finding and creating keys for joins, 171–173
 - shaping column data to meet Power Query requirements, 173–174
- evaluating story workspace, in Tableau, 385–386
- exact numerics, 462
- Excel Connector, connecting using, 141–143
- Excel files, 321
- Excel Workbook button (Home Ribbon), 185
- EXCEPT operation, 590

- executing SQL statements
 - about, 455–456
 - combining with host languages, 457
 - embedded SQL, 457–460
 - interactive SQL, 456
 - module language, 460–461
- existence, testing for, 567
- EXISTS clause, 525–526
- EXP function, 500
- expanding
 - data lakes, 43–45
 - tables, 179
- explicit join, 171
- Explore feature (Tableau Cloud), 439–440
- Explorer bundle (Tableau), 281, 283–284
- exporting
 - dashboards as graphics, 358
 - reports, 235–236
- expressions (SQL)
 - array value, 506
 - Boolean value, 506
 - CAST, 510–511
 - conditional value, 507–510
 - datetime value, 504–505
 - interval value, 505
 - numeric value, 503
 - row value, 511–512
 - string value, 503–504
- extending dashboards, in Tableau, 381–382
- external routine, 453
- externally invoked routine, 453
- Extract, Transport, Load (ETL) tool, 49, 313, 314
- EXTRACT function, 498
- extracting input step updates, 319–320
- extraction, transformation, and loading (ELT), 48–49

F

- factor, 653
- factor() function, 647
- Favorites (Tableau Cloud), 432–434

- features
 - of graphs in R, 658–660
 - Tableau Cloud
 - Favorites, 432–434
 - Personal Space, 430–431
 - Recents, 434–435
- fields
 - categorizing in Tableau Prep Builder, 340–342
 - in Tableau Prep Builder, 339–340
- File menu (Tableau Desktop), 353–354
- files
 - creating unions among, 320–324
 - in Power BI Desktop, 126–128
- filter cards, for report pages, 235
- Filter pane, 235
- filtering
 - data, 207–208
 - data flows, 329–333
- filters
 - in Excel, 71, 72
 - in Tableau, 273
- finding
 - keys for joins, 171–173
 - patterns for graphs
 - about, 651–652
 - bar graphs, 653–654
 - box and whisker plot, 656–657
 - graphing distributions, 652–653
 - pie graph, 654
 - scatter plot, 654–655
 - root cause of data inconsistencies, 170
- First Normal Form (1NF), 472–473
- Fit tool, on toolbar in Tableau Desktop, 368
- Fix Axes tool, on toolbar in Tableau Desktop, 368
- float format, in Python, 706
- FLOAT type, 466
- floating-point number, 465
- FLOOR function, 501
- fonts, adding to dashboards in Tableau, 377–378
- for loops, in R, 649–650
- forecast package, 619

- foreign key, 61–62
- foreign key constraints, 483–484
- Format menu (Tableau Desktop), 362–363
- Format option
 - Format menu of Tableau Desktop, 361, 362
 - Story menu of Tableau Desktop, 359
- formatting
 - axes in Matplotlib, 726–727
 - managing (and data type properties), 189–191
 - report visualizations, 232–239
 - stories in Tableau, 389–390
 - worksheets in Tableau Desktop, 356–357
- Formatting and Images section (Dashboard menu of Tableau Desktop), 358
- FROM clauses, 514–515
- FTP server, 711
- full outer join, 176, 302–303, 602–603
- function call, 613
- functional coding style, of Python, 694
- functional dependencies, 447–448
- functions. *See also specific functions*
 - aggregate, 533
 - Python, 709–711
 - R, 612–615, 638–639
 - set, 533
 - SQL, 66–69, 491–502
 - user-defined (R), 639–640
- Funnel charts, 217–218
- fuzzy matching, 176–177

G

- Gantt charts
 - about, 94–95, 96
 - in Tableau, 416–418
- gauges, 223
- General Architecture for Text Engineering (GATE), 69
- generating
 - automatic relationships, 197–198
 - bar charts in Tableau, 400
 - bullet charts in Tableau, 418
 - circles in Tableau, 403–404
 - continuous line charts in Tableau, 407–408
 - dashboards, 240–241
 - data lakes, 42
 - data models
 - about, 17, 183–184
 - adding and modifying data to imported, DirectQuery, and composite models, 196–197
 - Data view, 184–186
 - defining data types, 188–189
 - importing queries, 186–188
 - managing formatting and data type properties, 189–191
 - managing tables, 191–196
 - Modeling view, 184–186
 - in Power BI Desktop, 128–129
 - generic vectors, 609
 - Geocoding option (Map menu in Tableau Desktop), 361
 - geographical data, plotting in Matplotlib, 752–756
 - geom function, 672
 - geom_bar() function, 675–676
 - geom_boxplot() function, 686
 - geom_dotplot() function, 677
 - geom_histogram() function, 672–673
 - geom_jitter() function, 688
 - Get Data button (Home Ribbon), 185
 - getting
 - axes in Matplotlib, 725–726
 - data
 - from Microsoft-based file systems, 151–153
 - from sources, 142–146
 - ggpairs() function, 685
 - ggplot() function, 672–673, 676, 679–680, 681–682, 683, 686
 - ggplot2
 - about, 671–672
 - bar plots, 675–676, 679–683
 - box plots, 686–688
 - dot charts, 676–679
 - histograms, 672–675
 - scatter plot matrix, 683–686
 - scatter plots, 683
 - visualizing R statistics with, 620–621

- Google Colab, 699
- Google Forms, 58
- Google Sheets, 58, 70
- granularity, in maps, 221
- graph models, 100, 101
- graphical elements, adding context with, 105, 106
- graphing distributions, in R, 652–653
- graphs
 - about, 651
 - directed, 759–760
 - exporting dashboards as, 358
 - finding patterns
 - about, 651–652
 - bar graphs, 653–654
 - box and whisker plot, 656–657
 - graphing distributions, 652–653
 - pie graph, 654
 - scatter plot, 654–655
- Matplotlib
 - about, 722
 - adding grids, 727–728
 - defining plots, 722–723
 - drawing multiple lines and plots, 723–724
 - formatting axes, 726–727
 - getting axes, 725–726
 - saving, 724–725
- R
 - about, 657
 - bar plots, 660–661, 663–666
 - box plots, 669–670
 - dot charts, 662–663
 - features, 658–660
 - histograms, 657–658
 - pie graphs, 662
 - scatter plots, 666–669
 - saving dashboards as, 358
 - selecting in Matplotlib, 740–745
 - undirected, 757–758
 - visualizing in Matplotlib, 757–760
- greater than (>) operator, 517, 614
- greater than or equal to (>=) operator, 517, 614
- green pill, in Tableau, 272

- grid() function, 728
- grids, adding in Matplotlib, 727–728
- Grids section (Dashboard menu of Tableau Desktop), 357, 358
- GROUP BY clauses, 533–535
- Group Members tool, on toolbar in Tableau Desktop, 367, 368
- GROUP statement, aggregating data using, 69
- grouping data flows, 327–328
- groups, depicting in Matplotlib, 746–747

H

- Hadoop Distributed File System (HDFS), 30, 43
- Hadoop MapReduce, 38
- handle, 726
- hash symbol (#), 640, 719
- HAVING clause
 - about, 535–536, 570–571
 - narrowing results using, 69
 - SQL queries and, 551–554
- HDFS (Hadoop Distributed File System), 30, 43
- head() function, 669
- heat map, in Tableau, 394–395
- Help feature (Tableau Cloud), 427, 428
- Help menu (Tableau Desktop), 365
- helper queries, 180
- Hide Field option, 329
- hiding data, 200–201
- hierarchical classification, 100, 101
- highlight table, in Tableau, 395–396
- Highlight tool, on toolbar in Tableau Desktop, 367, 368
- histograms
 - about, 95, 98
 - in ggplot2, 672–675
 - in Matplotlib, 741–743
 - in R, 657–658
 - in Tableau, 414–415
 - in Tableau Prep Builder, 343–344
- Home Ribbon (Power BI model view), 185–186
- host language, 461
- HubSpot, 37

- human resources (HR), role of big data in, 20
- Hyman, Jack A. (author)
 - Tableau For Dummies*, 2nd Edition, 275
- hypotenuse, 639–640
- hypotenuse() function, 639–640
- hypotheses
 - alternative, 628–629
 - null, 628–629
 - testing, 628

I

- icons, explained, 3–4
- IDE (Integrated Development Environment), 699
- if statements, in R, 649–650
- igraph, analyzing networks with, 621–622
- images
 - adding to dashboards, 242
 - integrating in dashboards in Tableau, 378–379
- impedance mismatch, 479
- imperative coding style, of Python, 694
- implicit join, 171
- Import mode, 138
- imported data models, adding and modifying data to, 196–197
- importing queries, 186–188
- IN clause, 519–520, 567–568
- inconsistencies
 - detecting, 162–163
 - resolving for data
 - about, 168
 - finding root cause, 170
 - removing rows using Power Query, 170
 - replacing values, 168–170
- indentation, need for, 698–699
- independent variables, in R, 625
- index
 - column, 64–65
 - defined, 59
- indicators
 - cards, 223–224
 - gauges, 223
 - key performance indicators (KPIs), 224–225
 - multi-cards, 223–224
- inferential statistics, 624, 628
- information, data compared with, 33–34
- information technology (IT), 31
- ingesting data, in Power BI Desktop, 125–126
- init function, 712
- inner join, 176, 302–303
- inner JOIN() function, 68
- inner joins, 598–599
- Insert Flow option, 326
- INSERT statement, using subqueries in, 571–574
- insight management, as a role of big data, 24–25
- INTEGER type, 462–463
- integers, in Python, 706
- Integrated Development Environment (IDE), 699
- integrating
 - buttons in dashboards in Tableau, 380–381
 - images in dashboards in Tableau, 378–379
 - logos in dashboards in Tableau, 378–379
 - web pages in dashboards in Tableau, 379–380
- interactive, 608
- interactive elements, applying in Tableau Desktop, 356
- interactive SQL, 456
- Internet of Things (IoT), 35
- Internet resources
 - Cartopy, 753
 - Cheat Sheet, 4
 - Comprehensive R Archive Network (CRAN), 619, 630–631
 - DurusWorks, 698
 - embedding reports on, 138
 - General Architecture for Text Engineering (GATE), 69
 - Keras, 701
 - MATLAB, 721
 - Matplotlib library, 701, 721, 741
 - MySQLdb, 697
 - Natural Language Toolkit (NLTK), 69
 - NetworkX, 702
 - NumPy library, 700

Internet resources (*continued*)

pandas library, 700

Power BI Desktop, 120

Quixote, 698

R, 620

RStudio, 631

Scikit-learn library, 700

SciPy library, 699

Tableau, 291

TensorFlow, 701

INTERSECT operation, 588–589

interval data, in R, 626

intervals, 471, 505

is assigned, gets (<-) operator, 614

is assigned to (->) operator, 614

items

creating in Power BI Desktop, 130–131

publishing in Power BI Desktop, 130–131

iterating, in R, 615–616

J

Java, 38

JavaScript library, 55

JavaScript Object Notation (JSON) files, 159–160

JOIN() function

inner, 68

outer, 68

Join option, 325

joins

data

about, 302

blending data from multiple sources, 307–308

clipboard data, 309–311

creating, 304–306

input steps and, 323–324

joining fields that contain null values, 306

setting up clauses, 303–304

types of, 302–303

defined, 172

finding and creating keys for, 171–173

keys for, 172

SQL

about, 591–592

Cartesian product, 592–593

ON clause, 603

clustering indexes, 603–604

column-name, 597–598

condition, 596–597

conditions for, 603–604

cross, 592–593

equi-join, 594–596

full outer, 602–603

inner, 598–599

left outer, 599–601

natural, 596

outer, 599–603

right outer, 602

WHERE clause, 603

JSON (JavaScript Object Notation) files, 159–160

Jupyter Notebook, 699

K

Keep Only Field option, 329

Keras, 701

key influencers, 230–231

key performance indicators (KPIs),
224–225

keys

about, 61–62, 64

finding and creating for joins, 171–173

relational databases, 448–450

knowledge management, relationship with insight
management, 25

L

labels, adding in Matplotlib, 733–734

Labels/Measures/Marks/Data section (Analysis
menu of Tableau Desktop), 359, 360

labs() function, 674, 678

latency, 29, 49

- Layout pane, in story workspace, 386
- learning Python, 695–696
- left anti join, 176
- left join, 302–303
- left outer join, 176, 599–601
- legend() function, 736
- legends, adding in Matplotlib, 733–734, 735–737
- Legends, Highlights, Filters, Parameters section (Analysis menu of Tableau Desktop), 360
- less than (<) operator, 517, 614
- less than or equal to (<=) operator, 517, 614
- libraries (Python)
 - about, 713
 - Matplotlib, 716–718
 - NumPy, 714–715
 - Pandas, 718–719
 - Scikit-learn, 719
 - SciPy, 716
 - support for, 697
- licensing
 - about, 118
 - Power BI, 119
 - for Tableau, 280
- lifecycle, of data, 167–168
- LIKE clause, 520–521
- line charts
 - about, 92, 93
 - combining with bar charts, 215
 - continuous, 406
 - discrete, 407
 - dual-, 407
 - in Power BI, 213–214, 215
 - in Tableau, 406–408
- linear topological structures, 100
- lines
 - defining appearance of in Matplotlib
 - about, 729
 - adding markers, 731–733
 - line styles, 729–730
 - using colors, 730–731
 - drawing multiple in Matplotlib, 723–724
 - styles in Matplotlib, 729–730
- Lines, labels, and captions option (Format menu of Tableau Desktop), 361, 362
- lines() function, 659–660
- list() function, 645
- LISTAGG function, 494
- lists
 - about, 609
 - in Python, 707
 - in R, 645–646
- literal values, 488–489
- Live Connection, 138–139, 150–151
- LN function, 500
- loading
 - data, 162–165
 - data in Python, 695
- local storage, 150–151
- Location category (Tableau), 292
- LOG() function, 500
- LOG10 function, 500
- logarithmic functions, 500
- logical connectives, 531–532
- logical vector, 641
- logically centralized, data lakes as, 43
- logos, integrating in dashboards in Tableau, 378–379
- long format, in Python, 706
- loops, in Python, 708–709
- LOWER function, 496
- low-latency data feed, 49
- ls() function, 638

M

- machine learning, 37
- machine learning engineers, 37, 56
- macros, automating Excel tasks using, 77–79
- Manage Relationships button (Home Ribbon), 186
- Manage Roles button (Home Ribbon), 186
- management, as a benefit of data models, 128

- managing
 - cases, 507–508
 - data source settings, 146–147
 - data when sharing and collaborating, 254–258
 - formatting and data type properties, 189–191
 - null values, 478–479
 - relationships in data models
 - about, 197
 - creating automatic relationships, 197–198
 - creating manual relationships, 198–199
 - deleting relationships, 199
 - tables
 - adding data in, 195–196
 - adding tables, 191–192
 - deleting columns, 193–195
 - deleting tables, 193
 - modifying data in, 195–196
 - renaming columns, 193–195
 - renaming tables, 192–193
 - worksheets in Tableau Desktop, 356
- manual relationships, creating, 198–199
- many-to-many relationships, 100, 171
- Map Legend option (Map menu in Tableau Desktop), 361
- Map menu (Tableau Desktop), 361–362
- Map Options option (Map menu in Tableau Desktop), 361
- mapping spatial point patterns, with spatstat, 622
- maps, 221–222
- maps with/without symbols, in Tableau, 396–398
- mark type, 392
- markers, adding in Matplotlib, 731–733
- marketing data scientists, 56
- massively parallel processing (MPP) platform, 38
- MATCH clause, 527–532
- mathematical modeling, applying to data science tasks, 54
- mathematics, 54
- MATLAB application. *See* Matplotlib
- MATLAB For Dummies* (Mueller), 721

- Matplotlib library
 - about, 701–702, 716–718, 721
 - adding
 - annotations, 733–735
 - labels, 733–734
 - legends, 733–734, 735–737
 - data visualizations
 - about, 739
 - bar charts, 740–741
 - boxplots, 743–744
 - depicting groups, 746–747
 - displaying correlations, 747–748
 - histograms, 741–743
 - plotting geographical data, 752–756
 - plotting time series, 748–752
 - scatterplots, 744–745, 746–748
 - selecting graphs, 740–745
 - visualizing graphs, 757–760
 - defining line appearance
 - about, 729
 - adding markers, 731–733
 - line styles, 729–730
 - using colors, 730–731
 - graphs
 - about, 722
 - adding grids, 727–728
 - defining plots, 722–723
 - drawing multiple lines and plots, 723–724
 - formatting axes, 726–727
 - getting axes, 725–726
 - saving, 724–725
 - website, 741
- matrices
 - about, 229, 609
 - in R, 643–645
 - scatter plot, 667–669
- `matrix()` function, 644
- MAX function, 493
- maximum cardinality, 506
- `mean()` function, 612

- Measure field type, 341
- measures, in Tableau, 270–272
- measuring, as a role of big data, 23
- median, 656
- `melt()` function, 680
- Melton, Jim (author)
 - Querying XML*, 472
- menus (Tableau Desktop)
 - about, 353
 - Analysis menu, 359–360
 - Dashboard menu, 357–358
 - Data menu, 354–355
 - File menu, 353–354
 - Format menu, 362–363
 - Help menu, 365
 - Map menu, 361–362
 - Server menu, 363–364
 - Story menu, 358–359
 - Window menu, 364–365
 - Worksheet menu, 355–357
- merging queries, 176–180
- metadata, adding to worksheets in Tableau Desktop, 357
- Microsoft 365 sensitivity labels, 262
- Microsoft AI engine, 231
- Microsoft Azure, 44
- Microsoft Azure Data Lake Storage (ADLS), 43, 45
- Microsoft Cosmos DB, 156–157
- Microsoft Excel
 - about, 70
 - automating tasks with macros, 77–79
 - getting to know your data using, 71–75
 - reformatting with PivotTables, 75–77
 - for sharing and collaborating, 256
 - summarizing with PivotTables, 75–77
- Microsoft Power Apps, 58
- Microsoft Teams, collaborating inside Power BI Services using, 137
- Microsoft-based file systems, getting data from, 151–153
- MIN function, 493
- minus (-) operator, 614
- MOD function, 500
- Modeling view, 184–186
- models, training in Python, 695–696
- modifying
 - data in tables, 195–196
 - data to imported, DirectQuery, and composite models, 196–197
 - root tables to data relationships, 300–301
 - sensitivity labels, 263
- modifying clauses (SQL)
 - about, 514
 - ALL clause, 523–525
 - ANY clause, 523–525
 - BETWEEN clause, 517–519
 - IN clause, 519–520
 - FROM clauses, 514–515
 - DISTINCT clause, 526
 - EXISTS clause, 525–526
 - GROUP BY clauses, 533–535
 - HAVING clauses, 535–536
 - LIKE clause, 520–521
 - MATCH clause, 527–532
 - NOT IN clause, 519–520
 - NOT LIKE clause, 520–521
 - NULL clause, 521–522
 - ORDER BY clauses, 536–538
 - OVERLAPS clause, 527
 - SOME clause, 523–525
 - UNIQUE clause, 526
 - WHERE clauses, 515–517
- modular arithmetic, 650
- module language (SQL), 460–461
- modulo (%%) operator, 614, 650
- monitoring, as a role of big data, 23–24
- moving tables, to create different data relationships, 299–300
- MPP (massively parallel processing) platform, 38
- Mueller, John (author)
 - Beginning Programming with Python For Dummies*, 3rd Edition, 692, 697
 - MATLAB For Dummies*, 721
- multi-cards, as indicators, 223–224

- multinomial logit model, 619
- multiple dependencies, 64
- multisets, 446–447, 474
- My Workspace, 114
- MySQL, 60
- MySQLdb, 697

N

- narratives, in stories in Tableau, 385
- NATIONAL CHARACTER LARGE OBJECT type, 467–468
- NATIONAL CHARACTER type, 467–468
- NATIONAL CHARACTER VARYING type, 467–468
- natural joins, 596
- Natural Language Toolkit (NLTK), 69
- Navigation pane (Power BI), 117–118
- Navigator, in story workspace, 386, 389
- n-dimensional arrays, 714
- nested queries, 574–579
- networks, analyzing with statnet and igraph, 621–622
- NetworkX, 702
- New Column button (Home Ribbon), 186
- New Dashboard section (Dashboard menu of Tableau Desktop), 357, 358
- New Data Source tool, on toolbar in Tableau Desktop, 366
- New Measure button (Home Ribbon), 186
- New Rows option, 325
- New Story option (Story menu of Tableau Desktop), 359
- New Table button (Home Ribbon), 186
- NLTK (Natural Language Toolkit), 69
- nodes, 757
- nominal data, in R, 626
- nonexistence, testing for, 567
- non-interactive, 608
- nonredundancy, of columns, 64
- nonrelational data systems, semi-structured data and, 11
- nonrelational (NoSQL) databases, 11, 156–158
- normalizing
 - data, 348–349
 - databases, 64
- NoSQL databases, 11, 156–158
- NOT connective, 532
- not equal (<>) predicate, 517
- not equals (!=) operator, 614
- NOT IN clause, 519–520
- NOT LIKE clause, 520–521
- not (!) logical operator, 614
- NOT NULL constraint, 480
- Notebook environments, 753–754
- NULL clause, 521–522
- null hypotheses, 628–629
- null values
 - joining fields that contain, 306
 - managing, 478–479
- Null Values Filter, 331
- NULL IF expression, 509
- numbers, in Python, 706
- NUMERIC type, 464
- numeric value
 - about, 503
 - functions
 - about, 497
 - ABS, 499–500
 - ACOS, 500
 - ASIN, 500
 - ATAN, 500
 - CARDINALITY, 499
 - CEIL (ING), 501
 - CHARACTER_LENGTH, 498
 - COS, 500
 - COSH, 500
 - EXP, 500
 - EXTRACT, 498
 - FLOOR, 501
 - LN, 500
 - LOG (), 500

- LOG10, 500
- logarithmic, 500
- MOD, 500
- OCTET_LENGTH, 498–499
- POSITION, 498
- POWER, 500
- SIN, 500
- SINH, 500
- SQRT, 501
- TAN, 500
- TANH, 500
- trigonometric, 500
- WIDTH_BUCKET, 501
- numerical data type, 63
- numerical vectors, in R, 642–643
- numerics
 - approximate, 464–466
 - exact, 462
- NumPy library, 700, 714–715

O

- object-extended relational databases, 46
- object-oriented coding style, of Python, 694
- object-oriented language, 609
- objects, in R, 617–619
- OCTET_LENGTH function, 498–499
- octothorpe, 640
- ON clause, joins and, 603
- OneDrive, getting data from, 151, 152
- one-to-one (1:1) relationship, 171
- online services, for data sources, 127
- online sources, 160–162
- on-premises, 44
- opening
 - Recents page in Tableau Cloud, 434
 - shared reports, 134
- operational data, analytic data compared with, 47–48
- operations, as a role of big data, 20–21
- operators, in R, 612–615
- OR connective, 532, 555

- or (|) logical operator, 614
- ORDER BY clauses, 536–538, 547–551
- ordinal data, in R, 626
- outer JOIN() function, 68
- outer joins, 599–603
- outliers, in Tableau Prep Builder, 344–345
- Output option, 326
- overhyping big data, 28
- OVERLAPS clause, 527
- OVERLAY function, 497

P

- packed circle diagrams, 94, 96
- paginated reports, 236, 258
- Paint Roller icon, 234
- pairs() function, 669, 683, 685
- Pandas library, 700, 718–719
- Panes option (Format menu of Tableau Desktop), 361, 362
- parallel processing, Python for, 697
- paths, 454
- patterns, finding for graphs
 - about, 651–652
 - bar graphs, 653–654
 - box and whisker plot, 656–657
 - graphing distributions, 652–653
 - pie graph, 654
 - scatter plot, 654–655
- Pause Auto Update tool, on toolbar in Tableau Desktop, 366
- PDF (Portable Document Format), 724
- Percentage section (Analysis menu of Tableau Desktop), 360
- Performance category (Tableau), 293
- performing what-if analysis, 17
- Personal Space (Tableau Cloud), 430–431
- pie graphs/charts
 - about, 92, 94
 - in Power BI, 219, 220
 - in R, 654, 662
 - in Tableau, 398–399

- Pierson, Lillian (author)
 - Data Science For Dummies*, 3rd Edition, 54, 55
- pinning reports, 245–246
- Pivot option, 325
- PivotCharts, 75, 76
- pivoting, with data, 345–348
- PivotTables
 - creating, 76
 - reformatting and summarizing data using, 75–77
- planning stories, in Tableau, 384–385
- `plot()` function, 666, 670, 729, 731, 732, 748, 750, 752
- `plotpairs()` function, 684
- `plot.show()` function, 722
- plots/plotting
 - defining in Matplotlib, 722–723
 - drawing multiple in Matplotlib, 723–724
 - geographical data in Matplotlib, 752–756
 - time series in Matplotlib, 748–752
- `plt.axes()` function, 725–726
- `plt.figure()` function, 726
- `plt.ioff()` function, 725
- `plt.plot()` function, 722, 723
- `plt.savefig()` function, 724
- plus (+) operator, 614
- PNG (Portable Network Graphic) format, 724
- point map, 102, 103
- `polyfit()` function, 748
- `polyid()` function, 748
- polymorphic table functions, 502
- populations, in R, 624–625
- Portable Document Format (PDF), 724
- Portable Network Graphic (PNG) format, 724
- POSITION function, 498
- positional mapping, 638
- PostgreSQL, 60
- Postscript (PS), 724
- Power BI
 - about, 109–110, 123
 - big data in, 39–40
 - data visualizations in
 - about, 203
 - area charts, 213–214
 - bar charts, 209–213, 215
 - column charts, 209–213
 - creating visualizations, 204–205
 - Donut charts, 219, 220
 - filtering data, 207–208
 - Funnel charts, 217–218
 - indicators, 222–225
 - line charts, 213–214, 215
 - maps, 221–222
 - Pie charts, 219, 220
 - reports, 203–208
 - Ribbon charts, 216
 - Scatter charts, 218–219
 - selecting visualizations, 205–206
 - treemaps, 219–221
 - types, 209–225
 - Waterfall charts, 216–217
- Power BI Desktop
 - about, 110, 112, 124–125
 - analyzing data, 129–130
 - building data models, 128–129
 - creating items, 130–131
 - files and databases, 126–128
 - ingesting data, 125–126
 - publishing items, 130–131
- Power BI Services
 - about, 110, 112–113, 132
 - collaborating inside, 137–138
 - editing reports, 132–136
 - refreshing data, 138–139
 - viewing reports, 132–136
 - working with dashboards, 136–137
- Power Map, 110, 111–112
- Power Pivot, 110, 111
- Power Q&A, 110, 112
- Power Query, 110–111
- Power View, 110, 111
- products
 - about, 118
 - Desktop vs. Services options, 120–121
 - license options, 119

- sharing and collaborating
 - about, 247
 - access management, 253–254
 - configuring workspaces, 250–253
 - creating workspaces, 250–253
 - data lineage, 258–262
 - data protection, 262–264
 - dataflows, 261–262
 - datasets, 261–262
 - managing data, 254–258
 - Microsoft Excel, 256
 - paginated reports, 258
 - Quick Insights, 256–257
 - settings and storage, 254
 - Usage Metrics reports, 257
 - workspaces, 247–254
- structured data and, 11
- terminology
 - about, 113
 - capacities, 113
 - dashboards, 117
 - Navigation pane, 117–118
 - reports, 115–116
 - workspaces, 114–115
- unstructured data and, 11
- Power BI Dataset button (Home Ribbon), 185
- Power BI Desktop
 - about, 110, 112, 124–125
 - analyzing data, 129–130
 - building data models, 128–129
 - creating items, 130–131
 - files and databases, 126–128
 - ingesting data, 125–126
 - license, 119
 - Power BI Services compared with, 120–121
 - publishing items, 130–131
- Power BI Embedded license, 119
- Power BI Free license, 119
- Power BI Mobile license, 119
- Power BI Premium license, 119
- Power BI Pro license, 119
- Power BI Report Builder, 237
- Power BI Report Server license, 119
- Power BI Services
 - about, 110, 112–113, 132
 - collaborating inside, 137–138
 - creating data0based paginated reports using, 237–238
 - editing reports, 132–136
 - navigation menu, 248, 249, 250
 - Power BI Desktop compared with, 120–121
 - refreshing data, 138–139
 - viewing reports, 132–136
 - working with dashboards, 136–137
- POWER function, 500
- Power Map, 110, 111–112
- power (^) operator, 614
- Power Pivot, 110, 111
- Power Q&A, 110, 112
- Power Query
 - about, 110–111
 - removing rows using, 170
 - shaping column data to meet requirements for, 173–174
- Power Query Editor, 186, 188
- Power View, 110, 111
- predicates, 516, 517
- Prediction option, 325
- predictive data analytics, 14
- preparing data, for data visualization
 - about, 141
 - checking data structures and column properties, 163–164
 - cleansing data, 162–165
 - data statistics, 164–165
 - detecting anomalies and inconsistencies, 162–163
 - getting data from Microsoft-based file systems, 151–153
 - getting data from sources, 142–146
 - loading data, 162–165
 - managing data source settings, 146–147
 - relational data sources, 153–162
 - shared vs. local datasets, 147–150
 - storage and connection modes, 150–151
 - transforming data, 162–165

- prescriptive data analytics, 14
- Presentation Mode tool, on toolbar in Tableau Desktop, 368
- primary key, 61–62, 64
- PRIMARY KEY constraint, 481–483
- print() function, 609, 710
- printing reports, 138
- privileges, 451
- probability, in R, 626–627, 659
- problem solving, as a role of big data, 26
- procedural coding style, of Python, 695
- products (Power BI)
 - about, 118
 - Desktop vs. Services options, 120–121
 - license options, 119
- Profile feature (Tableau Cloud), 428
- proportional symbol maps, 396–397
- protecting data, when sharing and collaborating, 262–264
- PS (Postscript), 724
- Publish button (Home Ribbon), 186
- publishing
 - data models, 201–202
 - data sources using Tableau Cloud, 336
 - data visualizations in Tableau, 422–424
 - items in Power BI Desktop, 130–131
- PubNub, 242
- Push mode, 139
- Python
 - about, 38
 - built-in functions, 709–711
 - classes, 711–713
 - data science and
 - about, 691–692
 - creating data science pipeline, 692–693
 - learning Python, 695–696
 - role of Python in data science, 693–695
 - using Python ecosystem for data science, 699–702
 - working with Python, 697–699
 - data types
 - about, 705–706
 - dictionaries, 708
 - lists, 707
 - numbers, 706
 - sets, 708
 - strings, 706–707
 - tuples, 707–708
 - ecosystem for data science, 699–702
 - functions, 709–711
 - learning, 695–696
 - libraries
 - about, 713
 - Matplotlib, 716–718
 - NumPy, 714–715
 - Pandas, 718–719
 - Scikit-learn, 719
 - SciPy, 716
 - loops, 708–709
 - role of in data science, 693–695
 - uses for, 55
 - using for data science, 703–705
 - working with, 697–699

Q

- Q&A feature, artificial intelligence (AI) and, 231–232
- qualitative data, 23
- quantified subqueries, 563–566
- quantitative data, 23
- queries
 - appending, 175–176
 - combining, 174–180
 - configuring for data loading, 180–181
 - data
 - about, 53–54
 - using SELECT(), WHERE(), and DATE(), 66–67
 - importing, 186–188
 - merging, 176–180
- SQL
 - about, 539–540, 557
 - correlated subqueries, 579–584
 - EXCEPT operation, 590
 - HAVING clause, 551–554
 - INTERSECT operation, 588–589

- nested queries, 574–579
- OR connective, 555
- ORDER BY clause, 547–551
- SELECT DISTINCT, 540–542
- subqueries, 557–574
- temporary tables, 542–547
- UNION operation, 584–588
- Querying XML* (Melton and Buxton), 472
- Quick Insights, for sharing and collaborating, 256–257
- Quick Measure button (Home Ribbon), 186
- quit() function, 698
- Quixote, 698

R

- R. *See also* ggplot2
 - about, 623
 - alternative hypotheses, 628–629
 - analyzing networks with statnet and igraph, 621–622
 - comments, 640
 - data types, 626
 - downloading, 608
 - error types, 629–630
 - functions, 612–615, 638–639
 - getting
 - about, 630–631
 - RStudio, 631–633
 - started with, 635–638
 - graphs
 - about, 657
 - bar plots, 660–661, 663–666
 - box plots, 669–670
 - dot charts, 662–663
 - features, 658–660
 - histograms, 657–658
 - pie graphs, 662
 - scatter plots, 666–669
 - if statements, 649–650
 - inferential statistics, 628
 - iterating in, 615–616
 - for loops, 649–650
 - mapping and analyzing spatial point patterns
 - with spatstat, 622
 - null hypotheses, 628–629
 - objects, 617–619
 - operators, 612–615
 - populations, 624–625
 - probability, 626–627
 - samples, 624–625
 - statistical analysis packages, 619–620
 - structures
 - about, 641
 - data frames, 646–649
 - lists, 645–646
 - matrices, 643–645
 - numerical vectors, 642–643
 - vectors, 641
 - user-defined functions, 639–640
 - uses for, 55
 - variables, 625
 - visualizing statistics with ggplot2, 620–621
 - vocabulary for, 608–611
 - website, 620
 - working directory, 634–635
- range() function, 740
- raster surface map, 102, 103
- ratio data, in R, 626
- raw data, 16
- RDBMSs, 38
- Reading view (Power BI), 116
- REAL type, 465
- Recent Sources button (Home Ribbon), 185
- Recents (Tableau Cloud), 434–435
- recommendations engine (Tableau Cloud), 440–441
- records, in Tableau Prep Builder, 338–339
- Redo tool, on toolbar in Tableau Desktop, 366
- REF types, 475
- referential integrity, 527, 529–531
- reformatting data, using PivotTables, 75–77
- Refresh button (Home Ribbon), 186

- refreshing
 - data in input step, 318
 - data in Power BI Services, 138–139
- relational data sources, 153–162
- relational databases
 - about, 58–60, 155–156, 445
 - catalogs, 452
 - connections, 452–453
 - functional dependencies, 447–448
 - hierarchy of, 452
 - keys, 61–62, 448–450
 - multisets, 446–447
 - object-extended, 46
 - paths, 454
 - privileges, 451
 - relational database hierarchy, 452
 - relational model and, 445–446
 - relations, 446–447
 - routines, 453
 - schemas, 451
 - sessions, 452–453
 - sets, 446–447
 - tables, 446–447
 - transactions, 452–453
 - users, 450
 - views, 450
- relational models, SQL and, 445–446
- relations, 446–447
- relationships
 - defined, 172
 - deleting, 199
 - keys for, 172
 - managing in data models
 - about, 197
 - creating automatic relationships, 197–198
 - creating manual relationships, 198–199
 - deleting relationships, 199
- Relative mode, for macros, 77–78
- Remember icon, 4
- Remove option, 329
- removing
 - rows using Power Query, 170
 - tables from data relationships, 301–302
- Rename Field option, 329
- renaming
 - columns in tables, 193–195
 - sheets in Tableau Desktop, 370
 - tables, 192–193
- rep() function, 642
- replacing values, 168–170
- Replay Animation tool, on toolbar in Tableau Desktop, 366
- reporting, as a role of big data, 25–26
- reports
 - accessing, 133–136
 - data visualization for
 - about, 232–233
 - applying conditional formatting, 233
 - configuring report page, 233–235
 - exporting reports, 235–236
 - preparing reports for distribution, 236–239
 - editing in Power BI Services, 132–136
 - embedding on dashboards and websites, 138
 - Power BI, 115–116
 - printing, 138
 - value of, 133
 - viewing in Power BI Services, 132–136
- reserved words, in SQL, 461
- resizing
 - captions in stories, 389
 - dashboards in stories, 390
- resolving
 - data inconsistencies
 - about, 168
 - finding root cause, 170
 - removing rows using Power Query, 170
 - replacing values, 168–170
 - errors during data import, 182
- resources, Internet
 - Cartopy, 753

- Cheat Sheet, 4
- Comprehensive R Archive Network (CRAN), 619, 630–631
- DurusWorks, 698
- embedding reports on, 138
- General Architecture for Text Engineering (GATE), 69
- Keras, 701
- MATLAB, 721
- Matplotlib library, 701, 721, 741
- MySQLdb, 697
- Natural Language Toolkit (NLTK), 69
- NetworkX, 702
- NumPy library, 700
- pandas library, 700
- Power BI Desktop, 120
- Quixote, 698
- R, 620
- RStudio, 631
- Scikit-learn library, 700
- SciPy library, 699
- Tableau, 291
- TensorFlow, 701
- results
 - sharing, 133
 - viewing in Python, 696
- reusability, as a benefit of data models, 128
- Ribbon charts, 216
- right anti join, 176
- right join, 302–303
- right outer join, 176, 602
- roles of big data
 - about, 20
 - artificial intelligence (AI), 26
 - data reuse, 26–27
 - in data science and engineering, 36–38
 - decision-making, 22
 - insight management, 24–25
 - measuring, 23
 - monitoring, 23–24
 - operations, 20–21
 - problem solving, 26
 - reporting, 25–26
 - strategy, 21–22
- root tables, changing of data relationships, 300–301
- routines, 453
- ROW type, 472–473
- row values, 488, 511–512
- rows
 - removing using Power Query, 170
 - in Tableau Prep Builder, 338–339
- RStudio, 631–633
- Run Auto Update tool, on toolbar in Tableau Desktop, 366

S

- SaaS (Software as a Service), 37, 120
- Salesforce, 37
- samples, in R, 624–625
- Save tool, on toolbar in Tableau Desktop, 366
- saving
 - dashboards as graphics, 358
 - in Matplotlib, 724–725
 - Tableau Prep Builder data
 - about, 333
 - automating flows, 334
 - creating published data sources, 334–336
- Scala, 38
- Scalable Vector Graphics (SVG), 724
- Scale category (Tableau), 293
- scatter plots
 - about, 98, 99, 218–219
 - in ggplot2, 683
 - in Matplotlib, 744–745, 746–748
 - in R, 654–655, 666–669
 - in Tableau, 412–414
- scatterplot matrix, 98, 99, 683–686
- schemas
 - in relational database hierarchy, 452
 - relational databases, 451
- Scikit-learn library, 700–701, 719
- SciPy library, 699–700, 716
- SCM (supply chain management), 153

- script file format, 53
- Script option, 325
- Security category (Tableau), 293
- SELECT command, 461
- SELECT DISTINCT, SQL queries and, 540–542
- SELECT() function, querying data using, 66–67
- SELECT statement, 513–514
- Selected Values Filter, 331
- selecting
 - data graphic type
 - about, 90–91
 - comparative graphics, 94–95
 - spatial plots and maps, 101–103
 - standard chart graphics, 91–94
 - statistical plots, 95–99
 - topology structures, 100–101
 - data visualizations, 205–206
 - design style, 87–90
 - graphs in Matplotlib, 740–745
 - Tableau versions, 281
 - tools for data science strategies, 57–58
- semi-structured data, 11
- sensitivity labels, 262–264
- seq() function, 639, 642
- Series objects, 718
- Server menu (Tableau Desktop), 363–364
- server side, 452
- sessions, 452–453
- SESSION_USER, 490
- set functions
 - about, 491–492, 533
 - AVG, 493
 - COUNT, 492–493
 - LISTAGG, 494
 - MAX, 493
 - MIN, 493
 - SUM, 493–494
- set_extent() method, 756
- sets
 - in Python, 708
 - relational databases, 446–447
- settings and storage, for sharing and collaborating, 254
- setup, of data sources in Tableau, 292–293
- set_xlim() function, 726
- set_xticks() function, 726
- set_ylim() function, 726
- set_yticks() function, 726
- shaping column data, to meet Power Query requirements, 173–174
- Share Workbook with Others tool, on toolbar in Tableau Desktop, 368–369
- shared capacity, 113
- shared vs. local datasets, 147–150
- shared with me (Tableau Cloud), 436–437
- SharePoint 365, getting data from, 151, 152
- sharing and collaborating
 - in Power BI
 - about, 247
 - access management, 253–254
 - configuring workspaces, 250–253
 - creating workspaces, 250–253
 - data lineage, 258–262
 - data protection, 262–264
 - dataflows, 261–262
 - datasets, 261–262
 - managing data, 254–258
 - Microsoft Excel, 256
 - paginated reports, 258
 - Quick Insights, 256–257
 - settings and storage, 254
 - Usage Metrics reports, 257
 - workspaces, 247–254
 - in Power BI Services, 137–138
 - results, 133
 - in Tableau Cloud
 - about, 435–436
 - collections, 437–439
 - Explore feature, 439–440
 - recommendations engine, 440–441
 - shared with me, 436–437
 - sharing content, 435–436
- Sheet tab, in Tableau Desktop, 353

- sheets
 - deleting in Tableau Desktop, 370
 - renaming in Tableau Desktop, 370
 - workbooks compared with, 369–370
- shelves
 - in Tableau, 274
 - in Tableau Desktop, 352
- Show Me icon, in Tableau Desktop, 353
- Show Start Page tool, on toolbar in Tableau Desktop, 366
- Show/Hide Cards tool, on toolbar in Tableau Desktop, 368
- Show/Hide Mark Labels tool, on toolbar in Tableau Desktop, 368
- Side Bar, in Tableau Desktop, 353
- side-by-side bar charts, 400
- SIN function, 500
- single brackets ([]), 610
- SINH function, 500
- sizing report pages, 235
- Slack, 37
- slicers, 228
- SMALLINT type, 463
- SME (subject matter expert), 53, 56
- Software as a Service (SaaS), 37, 120
- SOME clause, 523–525
- Sort Ascending tool, on toolbar in Tableau Desktop, 367
- Sort Descending tool, on toolbar in Tableau Desktop, 367
- sorting by/grouping by, 200
- source type, 476
- sources, of big data, 35
- Spark, 38
- spatial plots and maps, 101–103
- spatial point patterns, mapping and analyzing with spatstat, 622
- spatstat, mapping and analyzing spatial point patterns with, 622
- Split Value Filter, 333
- SQL All-in-One For Dummies*, 3rd Edition (Taylor), 461
- SQL Server button (Home Ribbon), 185
- SQL-invoked routine, 453
- SQRT function, 501
- stacked bar charts, 209–211, 212–213, 400
- stacked charts, 95, 97
- stacked column charts, 209–211, 212–213
- standard, horizontal bar charts, 400
- standard chart graphics, 91–94
- Start icon, in Tableau Desktop, 353
- statistical analysis
 - about, 13
 - R, 619–620
- statistical methods, deriving insights from, 55
- statistical plots, 95–99
- statistics, 54
- statnet, analyzing networks with, 621–622
- storage and connection modes, 138–139, 150–151
- storing data, 30, 31–32
- Story menu
 - story workspace, 386
 - Tableau Desktop, 358–359
- story pages, in Tableau Desktop, 369–370
- Story pane (story workspace), 385, 386
- story points, 388–389
- Story toolbar, in story workspace, 386
- storytelling (Tableau)
 - about, 371
 - charts as, 25
 - creating stories
 - about, 383, 386–389
 - evaluating story workspace, 385–386
 - formatting, 389–390
 - planning, 384–385
 - synthesizing data through Tableau story, 383–384
- dashboards
 - about, 371–372
 - adding objects to, 376–383
 - configuring, 372–374
 - customizing, 374–376
- strategy, as a role of big data, 21–22
- streaming data, adding to dashboards, 242
- string value, 503–504

- string value functions
 - about, 494–495
 - CONVERT, 496
 - LOWER, 496
 - OVERLAY, 497
 - SUBSTRING(), 495
 - TRANSLATE, 496
 - TRIM, 496
 - UPPER, 496
- strings, in Python, 706–707
- structured data, 10–11
- Structured Query Language (SQL)
 - about, 53
 - applying constraints
 - about, 479
 - assertions, 484–485
 - column constraints, 480–481
 - foreign key constraints, 483–484
 - table constraints, 481–483
 - data types
 - about, 461–462
 - approximate numerics, 464–466
 - ARRAY type, 474
 - BIGINT, 463–464
 - BINARY LARGE OBJECT (BLOB) type, 469
 - binary strings, 468–469
 - BINARY type, 468
 - BINARY VARYING type, 468
 - Booleans, 469
 - CHARACTER LARGE OBJECT (CLOB) type, 467
 - character strings, 466–468
 - CHARACTER type, 466
 - CHARACTER VARYING type, 467
 - collection types, 473–474
 - DATE type, 469
 - datetimes, 469–471
 - DECFLOAT type, 464
 - DECIMAL type, 464
 - distinct, 476
 - DOUBLE PRECISION type, 465–466
 - exact numerics, 462
 - FLOAT type, 466
 - INTEGER, 462–463
 - intervals, 471
 - NATIONAL CHARACTER LARGE OBJECT type, 467–468
 - NATIONAL CHARACTER type, 467–468
 - NATIONAL CHARACTER VARYING type, 467–468
 - NUMERIC type, 464
 - REAL type, 465
 - REF types, 475
 - ROW type, 472–473
 - SMALLINT, 463
 - structured, 477–478
 - TIME WITH TIME ZONE type, 470
 - TIME WITHOUT TIME ZONE type, 470
 - TIMESTAMP WITH TIME ZONE type, 471
 - TIMESTAMP WITHOUT TIME ZONE type, 470–471
 - user-defined types, 475–478
 - XML type, 471–472
 - entering data values
 - about, 487–488
 - column values, 488
 - literal values, 488–489
 - row values, 488
 - variables, 490–491
 - executing statements
 - about, 455–456
 - combining with host languages, 457
 - embedded SQL, 457–460
 - interactive SQL, 456
 - module language, 460–461
 - expressions
 - array value, 506
 - Boolean value, 506
 - CAST, 510–511
 - conditional value, 507–510
 - datetime value, 504–505
 - interval value, 505
 - numeric value, 503
 - row value, 511–512
 - string value, 503–504

- functions
 - about, 66–69, 491
 - set, 491–494
 - value, 494–502
- joins
 - about, 591–592
 - Cartesian product, 592–593
 - ON clause, 603
 - clustering indexes, 603–604
 - column-name, 597–598
 - condition, 596–597
 - conditions for, 603–604
 - cross, 592–593
 - equi-join, 594–596
 - full outer, 602–603
 - inner, 598–599
 - left outer, 599–601
 - natural, 596
 - outer, 599–603
 - right outer, 602
 - WHERE clause, 603
- managing null values, 478–479
- modifying clauses
 - about, 514
 - ALL clause, 523–525
 - ANY clause, 523–525
 - BETWEEN clause, 517–519
 - IN clause, 519–520
 - FROM clauses, 514–515
 - DISTINCT clause, 526
 - EXISTS clause, 525–526
 - GROUP BY clauses, 533–535
 - HAVING clauses, 535–536
 - LIKE clause, 520–521
 - MATCH clause, 527–532
 - NOT IN clause, 519–520
 - NOT LIKE clause, 520–521
 - NULL clause, 521–522
 - ORDER BY clauses, 536–538
 - OVERLAPS clause, 527
 - SOME clause, 523–525
 - UNIQUE clause, 526
 - WHERE clauses, 515–517
- queries
 - about, 158, 539–540, 557
 - correlated subqueries, 579–584
 - EXCEPT operation, 590
 - HAVING clause, 551–554
 - INTERSECT operation, 588–589
 - nested queries, 574–579
 - OR connective, 555
 - ORDER BY clause, 547–551
 - SELECT DISTINCT, 540–542
 - subqueries, 557–574
 - temporary tables, 542–547
 - UNION operation, 584–588
- relational databases
 - about, 445
 - catalogs, 452
 - connections, 452–453
 - functional dependencies, 447–448
 - keys, 448–450
 - multisets, 446–447
 - paths, 454
 - privileges, 451
 - relational database hierarchy, 452
 - relational model and, 445–446
 - relations, 446–447
 - routines, 453
 - schemas, 451
 - sessions, 452–453
 - sets, 446–447
 - tables, 446–447
 - transactions, 452–453
 - users, 450
 - views, 450
- routine, 453
- SELECT statement, 513–514
- uses for, 55
- using reserved words, 461
- Structured Query Language (SQL) databases
 - about, 58–60
 - keys, 61–62
- structured types, 477–478

- structures (R)
 - about, 641
 - data frames, 646–649
 - lists, 645–646
 - matrices, 643–645
 - numerical vectors, 642–643
 - vectors, 641
- subject areas, applying data science to, 55–56
- subject matter expert (SME), 53, 56
- subject-matter segregation, 65
- subqueries (SQL)
 - about, 557–558
 - correlated, 566–571
 - quantified, 563–566
 - that return multiple values, 558–560
 - that return single values, 560–563
 - using in DELETE statement, 571–574
 - using in INSERT statement, 571–574
 - using in UPDATE statement, 571–574
- subset() function, 669
- SUBSTRING() function, 495
- subtypes, 477
- sum() function, 638–639
- SUM function, 493–494
- summarization
 - about, 190–191
 - data using PivotTables, 75–77
- Summarization, Trending, Values, and Layouts section (Analysis menu of Tableau Desktop), 360
- supertypes, 477
- supply chain management (SCM), 153
- SVG (Scalable Vector Graphics), 724
- Swap Rows and Columns tool, on toolbar in Tableau Desktop, 367
- synthesizing data, through Tableau story, 383–384
- SYSTEM_USER, 491

T

- t() function, 644
- table constraints, 481–483
- table visualizations, 228–229

- Tableau
 - about, 267–268
 - big data in, 39–40
 - color coding in, 304
 - data connections
 - about, 285, 288–289
 - data relationships, 296–302
 - data sources, 292–296
 - joining data, 302–311
 - requesting new connectors, 291
 - Tableau Desktop, 289–290
 - Tableau Online, 290–291
 - Tableau Prep, 289–290
 - Tableau Server, 290–291
 - data visualizations
 - about, 391–392
 - area chart, 408–410
 - bar chart, 400–401
 - box and whisker plot, 415–416
 - bubbles, 402–406
 - bullet chart, 418–419
 - circles, 402–406
 - converting to cross tabulation tables, 419–422
 - dual combination chart, 411–412
 - Gantt chart, 416–418
 - heat map, 394–395
 - highlight table, 395–396
 - histogram, 414–415
 - line chart, 406–408
 - maps with/without symbols, 396–398
 - pie chart, 398–399
 - publishing, 422–424
 - scatter plot, 412–414
 - text table, 392–394
 - treemap, 401–402
 - structured data and, 11
 - terminology
 - aggregation, 273
 - continuous vs. discrete data, 272
 - data fields, 270
 - data source, 268–269

- data type, 269–270
- dimensions and measures, 270–272
- filter, 273
- workbook/worksheet, 274–275
- unstructured data and, 11
- user classifications, 283–284
- versions
 - about, 275
 - by data life cycle, 282–283
 - selecting, 281
 - Tableau Cloud, 279–280
 - Tableau Desktop, 275–278
 - Tableau Prep, 278–279
 - Tableau Server, 279–280
- website, 291
- Tableau Bridge, 293
- Tableau Bridge Client, 364
- Tableau Cloud
 - about, 279–280, 425–426
 - components of, 426–429
 - data life cycle and, 282–283
 - features
 - about, 430
 - Favorites, 432–434
 - Personal Space, 430–431
 - Recents, 434–435
 - publishing data sources using, 336
 - sharing and collaborating
 - about, 435
 - collections, 437–439
 - Explore feature, 439–440
 - recommendations engine, 440–441
 - shared with me, 436–437
 - sharing content, 435–436
- Tableau Desktop
 - about, 275–278, 351
 - data life cycle and, 282–283
 - menus
 - about, 353
 - Analysis menu, 359–360
 - Dashboard menu, 357–358
 - Data menu, 354–355
 - File menu, 353–354
 - Format menu, 362–363
 - Help menu, 365
 - Map menu, 361–362
 - Server menu, 363–364
 - Story menu, 358–359
 - Window menu, 364–365
 - Worksheet menu, 355–357
 - sheets compared with workbooks, 369–370
 - toolbar, 365–369
 - workspace, 351–353
- Tableau For Dummies*, 2nd Edition (Hyman), 275
- Tableau Prep Builder
 - about, 278–279, 313–314, 337
 - data flows
 - about, 314–316
 - configuring, 317–318
 - filtering, 329–333
 - grouping, 327–328
 - input step, 318–324
 - nurturing, 324–327
 - data life cycle and, 282–283
 - data structures
 - about, 337–338, 342
 - binning, 343–344
 - categorizing fields, 340–342
 - columns, 339–340
 - distributions, 344–345
 - fields, 339–340
 - histograms, 343–344
 - outliers, 344–345
 - pivoting with data, 345–348
 - records, 338–339
 - rows, 338–339
 - normalizing data, 348–349
 - saving data
 - about, 333
 - automating data flows, 334
 - creating published data sources, 334–336
- Tableau Server, 279–280

- table-based data visualization, 228–231
- tables
 - aggregating, 179
 - expanding, 179
 - managing
 - adding data in, 195–196
 - adding tables, 191–192
 - deleting columns, 193–195
 - deleting tables, 193
 - modifying data in, 195–196
 - renaming columns, 193–195
 - renaming tables, 192–193
 - moving to create different data relationships, 299–300
 - in relational database hierarchy, 452
 - relational databases, 446–447
 - removing from data relationships, 301–302
 - temporary, 542–547
- TAN function, 500
- TANH function, 500
- target audience, designing for
 - about, 84–85
 - Brainstorm step, 85–86
 - Choose most functional type for purpose step, 86–87
 - Define the purpose step, 86
- tasks, automating in Excel using macros, 77–79
- Taylor, Allen G. (author)
 - SQL All-in-One For Dummies*, 3rd Edition, 461
- Technical Stuff icon, 4
- template apps, creating, 138
- temporary tables, SQL queries and, 542–547
- TensorFlow, 701
- terminology
 - Power BI
 - about, 113
 - capacities, 113
 - dashboards, 117
 - Navigation pane, 117–118
 - reports, 115–116
 - workspaces, 114–115
 - Tableau
 - aggregation, 273
 - continuous vs. discrete data, 272
 - data fields, 270
 - data source, 268–269
 - data type, 269–270
 - dimensions and measures, 270–272
 - filter, 273
 - workbook/worksheet, 274–275
 - testing
 - data graphics, 103–104
 - for existence, 566–567
 - hypotheses, 628
 - for nonexistence, 567
 - text
 - adding to dashboards in Tableau, 377–378
 - as a data type, 63
 - text boxes, adding to dashboards, 242
 - text mark type, 392
 - text tables, in Tableau, 392–394
 - theme() function, 677–678
 - Themes option (Format menu of Tableau Desktop), 361, 362
 - the three Vs, of big data, 27, 28–30
 - throughput, 29
 - tilde (~) operator, 667
 - tiles, 242–244
 - time series, plotting in Matplotlib, 748–752
 - TIME WITH TIME ZONE type, 470
 - TIME WITHOUT TIME ZONE type, 470
 - times (*) operator, 614
 - TIMESTAMP WITH TIME ZONE type, 471
 - TIMESTAMP WITHOUT TIME ZONE type, 470–471
 - Tip icon, 4
 - Title and Navigator option (Story menu of Tableau Desktop), 359
 - Titles and Actions section (Dashboard menu of Tableau Desktop), 358
 - toolbar, in Tableau Desktop, 352, 365–369
 - tools, selecting for data science strategies, 57–58
 - Tooltip (Visualizations pane), 218
 - topology structures, 100–101
 - training models, in Python, 695–696

- transactions, 452–453
- Transform Data button (Home Ribbon), 186
- transforming data, 12, 162–165
- TRANSLATE function, 496
- tree network topology, 100, 101
- treemaps
 - about, 95, 97, 219–221
 - in Tableau, 401–402
- trends
 - analyzing data for, 12
 - plotting over time, 751–752
- trigonometric functions, 500
- TRIM function, 496
- tuples, 446, 707–708
- type conversion, 170

U

- UDTs (user-defined types), 475–478
- undirected graphs, creating, 757–758
- Undo tool, on toolbar in Tableau Desktop, 366
- UNION ALL operator, 175, 586–587
- UNION CORRESPONDING operator, 587–588
- union join, 302–303
- UNION operation, 584–588
- UNION operator, 584–586
- Union option, 325
- union-compatible, 584
- UNIQUE clause, 526
- UNIQUE constraint, 480–481
- unique number, 164
- unstructured data, 11
- UPDATE statement, using subqueries in, 571–574
- updates, enabling in Tableau Desktop, 357
- Updates option
 - Dashboard menu of Tableau Desktop, 358
 - Story menu of Tableau Desktop, 359
- UPPER function, 496
- Usage Metrics reports, for sharing and collaborating, 257
- use cases, in stories in Tableau, 385
- user classifications (Tableau), 283–284

- user-based licensing model (Tableau), 283
- user-defined functions, in R, 639–640
- user-defined types (UDTs), 475–478
- users, 450

V

- value functions
 - about, 494
 - datetime, 502
 - numeric, 497–501
 - polymorphic table, 502
 - string, 494–497
- values
 - of big data, 39
 - replacing, 168–170
 - of reports, 133
 - subqueries that return multiple, 558–560
 - subqueries that return single, 560–563
- Values (Visualizations pane), 218
- var () function, 638–639
- variables, 490–491, 625
- variety, of big data, 27, 29–30, 39
- VBA (Visual Basic for Applications), 77, 79
- vectors
 - about, 609
 - numerical (R), 642–643
 - in R, 641
- velocity, of big data, 27, 28–29, 39
- veracity, of big data, 39
- versions (Tableau)
 - about, 275
 - by data life cycle, 282–283
 - selecting, 281
 - Tableau Cloud, 279–280
 - Tableau Desktop, 275–278
 - Tableau Prep, 278–279
 - Tableau Server, 279–280
- video, adding to dashboards, 242
- View As button (Home Ribbon), 186
- view modes, for reports in Power BI, 116
- Viewer bundle (Tableau), 281, 283

- views/viewing
 - relational databases, 450
 - reports in Power BI Services, 132–136
 - results in Python, 696
 - in Tableau Desktop, 352
- Visual Basic for Applications (VBA), 77, 79
- Visualization enhancements option (Format menu of Tableau Desktop), 361, 362
- visualizations (data)
 - about, 81–82, 227
 - adding context
 - about, 104–105
 - with annotations, 105, 106
 - with data, 105
 - with graphical elements, 105, 106
 - artificial intelligence (AI) and Q&A feature, 231–232
 - complex, 228–231
 - dashboards
 - about, 239–240
 - adding content, 242–244
 - configuring, 240
 - creating new, 240–241
 - pinning reports, 245–246
 - data structures for
 - about, 342
 - binning, 343–344
 - categorizing fields, 340–342
 - distributions, 344–345
 - histograms, 343–344
 - outliers, 344–345
 - designing for target audience
 - about, 84–85
 - Brainstorm step, 85–86
 - Choose most functional type for purpose step, 86–87
 - Define the purpose step, 86
 - in Matplotlib
 - about, 739
 - bar charts, 740–741
 - boxplots, 743–744
 - depicting groups, 746–747
 - displaying correlations, 747–748
 - histograms, 741–743
 - plotting geographical data, 752–756
 - plotting time series, 748–752
 - scatterplots, 744–745, 746–748
 - selecting graphs, 740–745
 - visualizing graphs, 757–760
 - in Power BI
 - about, 203
 - area charts, 213–214
 - bar charts, 209–213, 215
 - column charts, 209–213
 - creating visualizations, 204–205
 - Donut charts, 219, 220
 - filtering data, 207–208
 - Funnel charts, 217–218
 - indicators, 222–225
 - line charts, 213–214, 215
 - maps, 221–222
 - Pie charts, 219, 220
 - reports, 203–208
 - Ribbon charts, 216
 - Scatter charts, 218–219
 - selecting visualizations, 205–206
 - treemaps, 219–221
 - types, 209–225
 - Waterfall charts, 216–217
 - preparing data for
 - about, 141
 - checking data structures and column properties, 163–164
 - cleansing data, 162–165
 - data statistics, 164–165
 - detecting anomalies and inconsistencies, 162–163
 - getting data from Microsoft-based file systems, 151–153
 - getting data from sources, 142–146
 - loading data, 162–165
 - managing data source settings, 146–147
 - relational data sources, 153–162
 - shared vs. local datasets, 147–150
 - storage and connection modes, 150–151
 - transforming data, 162–165

- for reports
 - about, 232–233
 - applying conditional formatting, 233
 - configuring report page, 233–235
 - exporting reports, 235–236
 - preparing reports for distribution, 236–239
- selecting data graphic type
 - about, 90–91
 - comparative graphics, 94–95
 - spatial plots and maps, 101–103
 - standard chart graphics, 91–94
 - statistical plots, 95–99
 - topology structures, 100–101
- selecting design style, 87–90
- in Tableau
 - about, 391–392
 - area chart, 408–410
 - bar chart, 400–401
 - box and whisker plot, 415–416
 - bubbles, 402–406
 - bullet chart, 418–419
 - circles, 402–406
 - converting to cross tabulation tables, 419–422
 - dual combination chart, 411–412
 - Gantt chart, 416–418
 - heat map, 394–395
 - highlight table, 395–396
 - histogram, 414–415
 - line chart, 406–408
 - maps with/without symbols, 396–398
 - pie chart, 398–399
 - publishing, 422–424
 - scatter plot, 412–414
 - text table, 392–394
 - treemap, 401–402
- table-based, 228–231
- testing data graphics, 103–104
- types of
 - about, 82
 - data art, 84
 - data showcasing, 83
 - data storytelling, 82–83

- Visualizations pane, 218
- visualizing
 - about, 130
 - data, 17
 - graphs in Matplotlib, 757–760
 - R statistics with ggplot2, 620–621
 - table, 228–229
- vocabulary, for R, 608–611
- volume, of big data, 27, 28, 39

W

- wallpaper, for report pages, 235
- Warning icon, 4
- Waterfall charts, 216–217
- web content, adding to dashboards, 242
- web pages, integrating in dashboards in Tableau, 379–380
- web programming file format, 53
- websites
 - Cartopy, 753
 - Cheat Sheet, 4
 - Comprehensive R Archive Network (CRAN), 619, 630–631
 - DurusWorks, 698
 - embedding reports on, 138
 - General Architecture for Text Engineering (GATE), 69
 - Keras, 701
 - MATLAB, 721
 - Matplotlib library, 701, 721, 741
 - MySQLdb, 697
 - Natural Language Toolkit (NLTK), 69
 - NetworkX, 702
 - NumPy library, 700
 - pandas library, 700
 - Power BI Desktop, 120
 - Quixote, 698
 - R, 620
 - RStudio, 631
 - Scikit-learn library, 700
 - SciPy library, 699
 - Tableau, 291
 - TensorFlow, 701

- what-if analysis, performing, 17
- WHERE clauses, 515–517, 603
- WHERE() function, querying data using, 66–67
- Wickham, Hadley, 671, 680
- wide-versus-tall debate, 345–348
- WIDTH_BUCKET function, 501
- Wildcard Search Filter, 331
- Wilkinson, Leland, 671
- Window menu (Tableau Desktop), 364–365
- with() function, 648
- word clouds, 95, 97
- workbooks
 - sheets compared with, 369–370
 - in Tableau, 274–275
- working directory, in R, 634–635
- Worksheet menu (Tableau Desktop), 355–357
- worksheets
 - adding metadata to in Tableau Desktop, 357
 - creating in Tableau Desktop, 355, 356
 - formatting in Tableau Desktop, 356–357
 - managing in Tableau Desktop, 356
 - in Tableau, 274–275
 - in Tableau Desktop, 369–370

- workspaces
 - collaborating inside Power BI Services using, 137
 - configuring, 250–253
 - creating, 250–253
 - Power BI, 114–115
 - for sharing and collaborating, 247–254
 - Tableau Desktop, 351–353

X

- x-coordinate, 656
- xlabel() function, 734
- .xls files, 321
- XML type, 471–472

Y

- y-coordinate, 656

Z

- zettabyte, 31–33