pandas

pandas:powerfulPythondataanalysis  
toolkit  
Release0.23.0  
WesMcKinney&PyDataDevelopmentTeam  
May15,2018

CONTENTS  
1What'sNew  
3  
1.1v0.23.0(May15,2017)  
.........................................  
3  
1.1.1Newfeatures  
..........................................  
5  
1.1.1.1JSONread/writeround-trippablewith  
orient=˜table˜  
.............  
5  
1.1.1.2  
.assign()  
acceptsdependentarguments  
......................  
6  
1.1.1.3Mergingonacombinationofcolumnsandindexlevels  
...............  
7  
1.1.1.4Sortingbyacombinationofcolumnsandindexlevels  
................  
8  
1.1.1.5ExtendingPandaswithCustomTypes(Experimental)  
................  
8  
1.1.1.6New  
observed  
keywordforexcludingunobservedcategoriesin  
groupby  
....  
9  
1.1.1.7Rolling/Expanding.apply()accepts  
raw=False  
topassa  
Series  
tothefunction  
11  
1.1.1.8  
DataFrame.interpolate  
hasgainedthe  
limit\_area  
kwarg  
........  
12  
1.1.1.9  
get\_dummies  
nowsupports  
dtype  
argument  
...................  
13  
1.1.1.10Timedeltamodmethod  
................................  
14  
1.1.1.11  
.rank()  
handles  
inf  
valueswhen  
NaN  
arepresent  
................  
14  
1.1.1.12  
Series.str.cat  
hasgainedthe  
join  
kwarg  
..................  
15  
1.1.1.13  
DataFrame.astype  
performscolumn-wiseconversionto  
Categorical  
...  
16  
1.1.1.14OtherEnhancements  
..................................  
16  
1.1.2BackwardsincompatibleAPIchanges  
.............................  
18  
1.1.2.1Dependencieshaveincreasedminimumversions  
...................  
18  
1.1.2.2Instantiationfromdictspreservesdictinsertionorderforpython3.6+  
........  
19  
1.1.2.3DeprecatePanel  
....................................  
20  
1.1.2.4pandas.core.commonremovals  
............................  
21  
1.1.2.5Changestomakeoutputof  
DataFrame.apply  
consistent  
............  
21  
1.1.2.6Concatenationwillnolongersort  
...........................  
23  
1.1.2.7BuildChanges  
.....................................  
23  
1.1.2.8IndexDivisionByZeroFillsCorrectly  
........................  
23  
1.1.2.9Extractionofmatchingpatternsfromstrings  
.....................  
24  
1.1.2.10Defaultvalueforthe  
ordered  
parameterof  
CategoricalDtype  
........  
25  
1.1.2.11Betterpretty-printingofDataFramesinaterminal  
..................  
26  
1.1.2.12DatetimelikeAPIChanges  
..............................  
27  
1.1.2.13OtherAPIChanges  
..................................  
28  
1.1.3Deprecations  
..........................................  
30  
1.1.4Removalofpriorversiondeprecations/changes  
........................  
31  
1.1.5PerformanceImprovements  
...................................  
32  
1.1.6DocumentationChanges  
....................................  
33  
1.1.7BugFixes  
............................................  
34  
1.1.7.1Categorical  
.......................................  
34  
1.1.7.2Datetimelike  
......................................  
34  
1.1.7.3Timedelta  
.......................................  
35  
1.1.7.4Timezones  
.......................................  
36  
i

1.1.7.5Offsets  
.........................................  
37  
1.1.7.6Numeric  
........................................  
37  
1.1.7.7Strings  
.........................................  
37  
1.1.7.8Indexing  
........................................  
37  
1.1.7.9MultiIndex  
.......................................  
38  
1.1.7.10I/O  
...........................................  
39  
1.1.7.11Plotting  
........................................  
40  
1.1.7.12Groupby/Resample/Rolling  
..............................  
40  
1.1.7.13Sparse  
.........................................  
41  
1.1.7.14Reshaping  
.......................................  
41  
1.1.7.15Other  
..........................................  
42  
1.2v0.22.0(December29,2017)  
......................................  
43  
1.2.1BackwardsincompatibleAPIchanges  
.............................  
43  
1.2.1.1ArithmeticOperations  
.................................  
43  
1.2.1.2GroupingbyaCategorical  
...............................  
44  
1.2.1.3Resample  
.......................................  
44  
1.2.1.4RollingandExpanding  
................................  
46  
1.2.2Compatibility  
..........................................  
46  
1.3v0.21.1(December12,2017)  
......................................  
47  
1.3.1RestoreMatplotlibdatetimeConverterRegistration  
......................  
47  
1.3.2Newfeatures  
..........................................  
48  
1.3.2.1ImprovementstotheParquetIOfunctionality  
....................  
48  
1.3.2.2OtherEnhancements  
..................................  
48  
1.3.3Deprecations  
..........................................  
48  
1.3.4PerformanceImprovements  
...................................  
48  
1.3.5BugFixes  
............................................  
48  
1.3.5.1Conversion  
.......................................  
48  
1.3.5.2Indexing  
........................................  
49  
1.3.5.3I/O  
...........................................  
49  
1.3.5.4Plotting  
........................................  
49  
1.3.5.5Groupby/Resample/Rolling  
..............................  
49  
1.3.5.6Reshaping  
.......................................  
50  
1.3.5.7Numeric  
........................................  
50  
1.3.5.8Categorical  
.......................................  
50  
1.3.5.9String  
.........................................  
50  
1.4v0.21.0(October27,2017)  
........................................  
50  
1.4.1Newfeatures  
..........................................  
52  
1.4.1.1IntegrationwithApacheParquetformat  
.....................  
52  
1.4.1.2  
infer\_objects  
typeconversion  
..........................  
52  
1.4.1.3Improvedwarningswhenattemptingtocreatecolumns  
...............  
53  
1.4.1.4  
drop  
nowalsoacceptsindex/columnskeywords  
...................  
53  
1.4.1.5  
rename  
,  
reindex  
nowalsoacceptaxiskeyword  
.................  
54  
1.4.1.6  
CategoricalDtype  
forspecifyingcategoricals  
.................  
55  
1.4.1.7  
GroupBy  
objectsnowhavea  
pipe  
method  
.....................  
56  
1.4.1.8  
Categorical.rename\_categories  
acceptsadict-like  
............  
57  
1.4.1.9OtherEnhancements  
..................................  
57  
1.4.2BackwardsincompatibleAPIchanges  
.............................  
59  
1.4.2.1Dependencieshaveincreasedminimumversions  
...................  
59  
1.4.2.2Sum/Prodofall-NaNoremptySeries/DataFramesisnowconsistentlyNaN  
....  
59  
1.4.2.3IndexingwithalistwithmissinglabelsisDeprecated  
................  
60  
1.4.2.4NAnamingChanges  
..................................  
61  
1.4.2.5IterationofSeries/IndexwillnowreturnPythonscalars  
...............  
61  
1.4.2.6IndexingwithaBooleanIndex  
............................  
62  
1.4.2.7  
PeriodIndex  
resampling  
..............................  
63  
ii

1.4.2.8Improvederrorhandlingduringitemassignmentinpd.eval  
.............  
64  
1.4.2.9DtypeConversions  
...................................  
65  
1.4.2.10MultiIndexConstructorwithaSingleLevel  
.....................  
66  
1.4.2.11UTCLocalizationwithSeries  
.............................  
66  
1.4.2.12ConsistencyofRangeFunctions  
...........................  
67  
1.4.2.13NoAutomaticMatplotlibConverters  
.........................  
68  
1.4.2.14OtherAPIChanges  
..................................  
68  
1.4.3Deprecations  
..........................................  
69  
1.4.3.1Series.selectandDataFrame.select  
..........................  
70  
1.4.3.2Series.argmaxandSeries.argmin  
...........................  
70  
1.4.4Removalofpriorversiondeprecations/changes  
........................  
70  
1.4.5PerformanceImprovements  
...................................  
71  
1.4.6DocumentationChanges  
....................................  
71  
1.4.7BugFixes  
............................................  
71  
1.4.7.1Conversion  
.......................................  
71  
1.4.7.2Indexing  
........................................  
72  
1.4.7.3I/O  
...........................................  
72  
1.4.7.4Plotting  
........................................  
73  
1.4.7.5Groupby/Resample/Rolling  
..............................  
73  
1.4.7.6Sparse  
.........................................  
74  
1.4.7.7Reshaping  
.......................................  
74  
1.4.7.8Numeric  
........................................  
75  
1.4.7.9Categorical  
.......................................  
75  
1.4.7.10PyPy  
..........................................  
75  
1.4.7.11Other  
..........................................  
75  
1.5v0.20.3(July7,2017)  
..........................................  
76  
1.5.1BugFixes  
............................................  
76  
1.5.1.1Conversion  
.......................................  
76  
1.5.1.2Indexing  
........................................  
76  
1.5.1.3I/O  
...........................................  
76  
1.5.1.4Plotting  
........................................  
77  
1.5.1.5Reshaping  
.......................................  
77  
1.5.1.6Categorical  
.......................................  
77  
1.6v0.20.2(June4,2017)  
..........................................  
77  
1.6.1Enhancements  
..........................................  
78  
1.6.2PerformanceImprovements  
...................................  
78  
1.6.3BugFixes  
............................................  
78  
1.6.3.1Conversion  
.......................................  
78  
1.6.3.2Indexing  
........................................  
79  
1.6.3.3I/O  
...........................................  
79  
1.6.3.4Plotting  
........................................  
79  
1.6.3.5Groupby/Resample/Rolling  
..............................  
79  
1.6.3.6Sparse  
.........................................  
79  
1.6.3.7Reshaping  
.......................................  
80  
1.6.3.8Numeric  
........................................  
80  
1.6.3.9Categorical  
.......................................  
80  
1.6.3.10Other  
..........................................  
80  
1.7v0.20.1(May5,2017)  
..........................................  
80  
1.7.1Newfeatures  
..........................................  
82  
1.7.1.1  
agg  
APIforDataFrame/Series  
............................  
82  
1.7.1.2  
dtype  
keywordfordataIO  
..............................  
84  
1.7.1.3  
.to\_datetime()  
hasgainedan  
origin  
parameter  
...............  
85  
1.7.1.4GroupbyEnhancements  
................................  
85  
1.7.1.5BettersupportforcompressedURLsin  
read\_csv  
.................  
86  
iii

1.7.1.6PickleI/Onowsupportscompression  
.......................  
86  
1.7.1.7UInt64SupportImproved  
...............................  
87  
1.7.1.8GroupByonCategoricals  
...............................  
88  
1.7.1.9TableSchemaOutput  
.................................  
89  
1.7.1.10SciPysparsematrixfrom/toSparseDataFrame  
....................  
89  
1.7.1.11ExceloutputforstyledDataFrames  
..........................  
90  
1.7.1.12IntervalIndex  
......................................  
91  
1.7.1.13OtherEnhancements  
..................................  
93  
1.7.2BackwardsincompatibleAPIchanges  
.............................  
95  
1.7.2.1PossibleincompatibilityforHDF5formatscreatedwithpandas<0.13.0  
......  
95  
1.7.2.2MaponIndextypesnowreturnotherIndextypes  
..................  
96  
1.7.2.3AccessingdatetimeofIndexnowreturnIndex  
.................  
97  
1.7.2.4pd.uniquewillnowbeconsistentwithextensiontypes  
................  
97  
1.7.2.5S3FileHandling  
....................................  
99  
1.7.2.6PartialStringIndexingChanges  
............................  
99  
1.7.2.7Concatofdifferentdtypeswillnotautomaticallyupcast  
............  
100  
1.7.2.8PandasGoogleBigQuerysupporthasmoved  
.....................  
100  
1.7.2.9MemoryUsageforIndexismoreAccurate  
......................  
100  
1.7.2.10DataFrame.sort\_indexchanges  
............................  
101  
1.7.2.11GroupbyDescribeFormatting  
.............................  
103  
1.7.2.12WindowBinaryCorr/CovoperationsreturnaMultiIndexDataFrame  
........  
104  
1.7.2.13HDFStorewherestringcomparison  
..........................  
105  
1.7.2.14Index.intersectionandinnerjoinnowpreservetheorderoftheleftIndex  
......  
105  
1.7.2.15PivotTablealwaysreturnsaDataFrame  
.......................  
106  
1.7.2.16OtherAPIChanges  
..................................  
107  
1.7.3Reorganizationofthelibrary:PrivacyChanges  
........................  
108  
1.7.3.1ModulesPrivacyHasChanged  
............................  
108  
1.7.3.2  
pandas.errors  
..................................  
109  
1.7.3.3  
pandas.testing  
.................................  
110  
1.7.3.4  
pandas.plotting  
.................................  
110  
1.7.3.5OtherDevelopmentChanges  
.............................  
110  
1.7.4Deprecations  
..........................................  
110  
1.7.4.1Deprecate  
.ix  
.....................................  
110  
1.7.4.2DeprecatePanel  
....................................  
111  
1.7.4.3Deprecategroupby.agg()withadictionarywhenrenaming  
.............  
112  
1.7.4.4Deprecate.plotting  
..................................  
114  
1.7.4.5OtherDeprecations  
..................................  
114  
1.7.5Removalofpriorversiondeprecations/changes  
........................  
115  
1.7.6PerformanceImprovements  
...................................  
116  
1.7.7BugFixes  
............................................  
116  
1.7.7.1Conversion  
.......................................  
116  
1.7.7.2Indexing  
........................................  
117  
1.7.7.3I/O  
...........................................  
118  
1.7.7.4Plotting  
........................................  
119  
1.7.7.5Groupby/Resample/Rolling  
..............................  
120  
1.7.7.6Sparse  
.........................................  
120  
1.7.7.7Reshaping  
.......................................  
120  
1.7.7.8Numeric  
........................................  
121  
1.7.7.9Other  
..........................................  
121  
1.8v0.19.2(December24,2016)  
......................................  
121  
1.8.1Enhancements  
..........................................  
122  
1.8.2PerformanceImprovements  
...................................  
122  
1.8.3BugFixes  
............................................  
122  
1.9v0.19.1(November3,2016)  
.......................................  
123  
iv

1.9.1PerformanceImprovements  
...................................  
124  
1.9.2BugFixes  
............................................  
124  
1.10v0.19.0(October2,2016)  
........................................  
125  
1.10.1Newfeatures  
..........................................  
127  
1.10.1.1  
merge\_asof  
forasof-styletime-seriesjoining  
...................  
127  
1.10.1.2  
.rolling()  
isnowtime-seriesaware  
.......................  
129  
1.10.1.3  
read\_csv  
hasimprovedsupportforduplicatecolumnnames  
...........  
131  
1.10.1.4  
read\_csv  
supportsparsing  
Categorical  
directly  
................  
132  
1.10.1.5CategoricalConcatenation  
...............................  
133  
1.10.1.6Semi-MonthOffsets  
..................................  
134  
1.10.1.7NewIndexmethods  
..................................  
134  
1.10.1.8GoogleBigQueryEnhancements  
...........................  
135  
1.10.1.9Fine-grainednumpyerrstate  
..............................  
136  
1.10.1.10  
get\_dummies  
nowreturnsintegerdtypes  
.....................  
136  
1.10.1.11Downcastvaluestosmallestpossibledtypein  
to\_numeric  
............  
136  
1.10.1.12pandasdevelopmentAPI  
...............................  
137  
1.10.1.13Otherenhancements  
..................................  
138  
1.10.2APIchanges  
...........................................  
140  
1.10.2.1  
Series.tolist()  
willnowreturnPythontypes  
.................  
140  
1.10.2.2  
Series  
operatorsfordifferentindexes  
.......................  
141  
1.10.2.3  
Series  
typepromotiononassignment  
.......................  
144  
1.10.2.4  
.to\_datetime()  
changes  
.............................  
145  
1.10.2.5Mergingchanges  
....................................  
145  
1.10.2.6  
.describe()  
changes  
...............................  
146  
1.10.2.7  
Period  
changes  
...................................  
148  
1.10.2.8Index  
+  
/  
-  
nolongerusedforsetoperations  
.....................  
149  
1.10.2.9  
Index.difference  
and  
.symmetric\_difference  
changes  
........  
150  
1.10.2.10  
Index.unique  
consistentlyreturns  
Index  
....................  
151  
1.10.2.11  
MultiIndex  
constructors,  
groupby  
and  
set\_index  
preservecategoricaldtypes  
151  
1.10.2.12  
read\_csv  
willprogressivelyenumeratechunks  
..................  
153  
1.10.2.13SparseChanges  
....................................  
153  
1.10.2.14Indexerdtypechanges  
.................................  
156  
1.10.2.15OtherAPIChanges  
..................................  
157  
1.10.3Deprecations  
..........................................  
158  
1.10.4Removalofpriorversiondeprecations/changes  
........................  
158  
1.10.5PerformanceImprovements  
...................................  
159  
1.10.6BugFixes  
............................................  
160  
1.11v0.18.1(May3,2016)  
..........................................  
165  
1.11.1Newfeatures  
..........................................  
166  
1.11.1.1CustomBusinessHour  
................................  
166  
1.11.1.2  
.groupby(..)  
syntaxwithwindowandresampleoperations  
...........  
167  
1.11.1.3Methodchaininngimprovements  
...........................  
169  
1.11.1.4Partialstringindexingon  
DateTimeIndex  
whenpartofa  
MultiIndex  
....  
171  
1.11.1.5AssemblingDatetimes  
.................................  
172  
1.11.1.6OtherEnhancements  
..................................  
173  
1.11.2Sparsechanges  
.........................................  
174  
1.11.3APIchanges  
...........................................  
175  
1.11.3.1  
.groupby(..).nth()  
changes  
..........................  
175  
1.11.3.2numpyfunctioncompatibility  
.............................  
177  
1.11.3.3Using  
.apply  
ongroupbyresampling  
........................  
177  
1.11.3.4Changesin  
read\_csv  
exceptions  
..........................  
178  
1.11.3.5  
to\_datetime  
errorchanges  
............................  
179  
1.11.3.6OtherAPIchanges  
...................................  
180  
1.11.3.7Deprecations  
......................................  
180  
v

1.11.4PerformanceImprovements  
...................................  
180  
1.11.5BugFixes  
............................................  
181  
1.12v0.18.0(March13,2016)  
........................................  
183  
1.12.1Newfeatures  
..........................................  
185  
1.12.1.1Windowfunctionsarenowmethods  
.........................  
185  
1.12.1.2Changestorename  
..................................  
187  
1.12.1.3RangeIndex  
......................................  
187  
1.12.1.4Changestostr.extract  
.................................  
188  
1.12.1.5Additionofstr.extractall  
................................  
189  
1.12.1.6Changestostr.cat  
...................................  
190  
1.12.1.7Datetimelikerounding  
.................................  
190  
1.12.1.8FormattingofIntegersinFloatIndex  
.........................  
192  
1.12.1.9Changestodtypeassignmentbehaviors  
........................  
193  
1.12.1.10to\_xarray  
........................................  
194  
1.12.1.11LatexRepresentation  
.................................  
195  
1.12.1.12  
pd.read\_sas()  
changes  
..............................  
195  
1.12.1.13Otherenhancements  
..................................  
195  
1.12.2BackwardsincompatibleAPIchanges  
.............................  
196  
1.12.2.1NaTandTimedeltaoperations  
............................  
196  
1.12.2.2Changestomsgpack  
..................................  
197  
1.12.2.3Signaturechangefor.rank  
...............................  
198  
1.12.2.4BuginQuarterBeginwithn=0  
............................  
199  
1.12.2.5ResampleAPI  
.....................................  
199  
1.12.2.6Changestoeval  
....................................  
204  
1.12.2.7OtherAPIChanges  
..................................  
206  
1.12.2.8Deprecations  
......................................  
207  
1.12.2.9Removalofdeprecatedindexers  
.........................  
207  
1.12.2.10Removalofpriorversiondeprecations/changes  
....................  
210  
1.12.3PerformanceImprovements  
...................................  
210  
1.12.4BugFixes  
............................................  
210  
1.13v0.17.1(November21,2015)  
......................................  
213  
1.13.1Newfeatures  
..........................................  
214  
1.13.1.1ConditionalHTMLFormatting  
............................  
214  
1.13.2Enhancements  
..........................................  
214  
1.13.3APIchanges  
...........................................  
216  
1.13.3.1Deprecations  
......................................  
216  
1.13.4PerformanceImprovements  
...................................  
216  
1.13.5BugFixes  
............................................  
217  
1.14v0.17.0(October9,2015)  
........................................  
218  
1.14.1Newfeatures  
..........................................  
220  
1.14.1.1DatetimewithTZ  
...................................  
220  
1.14.1.2ReleasingtheGIL  
...................................  
222  
1.14.1.3Plotsubmethods  
....................................  
222  
1.14.1.4Additionalmethodsfor  
dt  
accessor  
.........................  
223  
1.14.1.5PeriodFrequencyEnhancement  
............................  
225  
1.14.1.6SupportforSASXPORT  
.............................  
225  
1.14.1.7SupportforMathFunctionsin.eval()  
.........................  
226  
1.14.1.8ChangestoExcelwith  
MultiIndex  
........................  
226  
1.14.1.9GoogleBigQueryEnhancements  
...........................  
227  
1.14.1.10DisplayAlignmentwithUnicodeEastAsianWidth  
.................  
228  
1.14.1.11Otherenhancements  
..................................  
228  
1.14.2BackwardsincompatibleAPIchanges  
.............................  
232  
1.14.2.1ChangestosortingAPI  
................................  
232  
1.14.2.2Changestoto\_datetimeandto\_timedelta  
.......................  
233  
vi

1.14.2.3ChangestoIndexComparisons  
............................  
234  
1.14.2.4ChangestoBooleanComparisonsvs.None  
.....................  
235  
1.14.2.5HDFStoredropnabehavior  
..............................  
236  
1.14.2.6Changesto  
display.precision  
option  
.....................  
237  
1.14.2.7Changesto  
Categorical.unique  
........................  
237  
1.14.2.8Changesto  
bool  
passedas  
header  
inParsers  
...................  
238  
1.14.2.9OtherAPIChanges  
..................................  
238  
1.14.2.10Deprecations  
......................................  
239  
1.14.2.11Removalofpriorversiondeprecations/changes  
....................  
240  
1.14.3PerformanceImprovements  
...................................  
241  
1.14.4BugFixes  
............................................  
242  
1.15v0.16.2(June12,2015)  
.........................................  
245  
1.15.1Newfeatures  
..........................................  
246  
1.15.1.1Pipe  
..........................................  
246  
1.15.1.2OtherEnhancements  
..................................  
247  
1.15.2APIChanges  
..........................................  
247  
1.15.3PerformanceImprovements  
...................................  
247  
1.15.4BugFixes  
............................................  
248  
1.16v0.16.1(May11,2015)  
.........................................  
249  
1.16.1Enhancements  
..........................................  
250  
1.16.1.1CategoricalIndex  
....................................  
250  
1.16.1.2Sample  
.........................................  
252  
1.16.1.3StringMethodsEnhancements  
............................  
253  
1.16.1.4OtherEnhancements  
..................................  
254  
1.16.2APIchanges  
...........................................  
256  
1.16.2.1Deprecations  
......................................  
256  
1.16.3IndexRepresentation  
......................................  
256  
1.16.4PerformanceImprovements  
...................................  
258  
1.16.5BugFixes  
............................................  
258  
1.17v0.16.0(March22,2015)  
........................................  
260  
1.17.1Newfeatures  
..........................................  
261  
1.17.1.1DataFrameAssign  
...................................  
261  
1.17.1.2Interactionwithscipy.sparse  
..............................  
262  
1.17.1.3StringMethodsEnhancements  
............................  
264  
1.17.1.4Otherenhancements  
..................................  
265  
1.17.2BackwardsincompatibleAPIchanges  
.............................  
266  
1.17.2.1ChangesinTimedelta  
.................................  
266  
1.17.2.2IndexingChanges  
...................................  
267  
1.17.2.3CategoricalChanges  
..................................  
269  
1.17.2.4OtherAPIChanges  
..................................  
271  
1.17.2.5Deprecations  
......................................  
273  
1.17.2.6Removalofpriorversiondeprecations/changes  
....................  
273  
1.17.3PerformanceImprovements  
...................................  
273  
1.17.4BugFixes  
............................................  
274  
1.18v0.15.2(December12,2014)  
......................................  
276  
1.18.1APIchanges  
...........................................  
277  
1.18.2Enhancements  
..........................................  
279  
1.18.3Performance  
...........................................  
280  
1.18.4BugFixes  
............................................  
281  
1.19v0.15.1(November9,2014)  
.......................................  
282  
1.19.1APIchanges  
...........................................  
282  
1.19.2Enhancements  
..........................................  
286  
1.19.3BugFixes  
............................................  
287  
1.20v0.15.0(October18,2014)  
........................................  
288  
vii

1.20.1Newfeatures  
..........................................  
289  
1.20.1.1CategoricalsinSeries/DataFrame  
...........................  
289  
1.20.1.2TimedeltaIndex/Scalar  
.................................  
291  
1.20.1.3MemoryUsage  
....................................  
293  
1.20.1.4.dtaccessor  
......................................  
294  
1.20.1.5Timezonehandlingimprovements  
..........................  
297  
1.20.1.6Rolling/ExpandingMomentsimprovements  
.....................  
298  
1.20.1.7Improvementsinthesqliomodule  
..........................  
302  
1.20.2BackwardsincompatibleAPIchanges  
.............................  
302  
1.20.2.1Breakingchanges  
...................................  
302  
1.20.2.2InternalRefactoring  
..................................  
308  
1.20.2.3Deprecations  
......................................  
308  
1.20.2.4Removalofpriorversiondeprecations/changes  
....................  
308  
1.20.3Enhancements  
..........................................  
309  
1.20.4Performance  
...........................................  
313  
1.20.5BugFixes  
............................................  
313  
1.21v0.14.1(July11,2014)  
..........................................  
317  
1.21.1APIchanges  
...........................................  
317  
1.21.2Enhancements  
..........................................  
318  
1.21.3Performance  
...........................................  
319  
1.21.4Experimental  
..........................................  
319  
1.21.5BugFixes  
............................................  
320  
1.22v0.14.0(May31,2014)  
.........................................  
322  
1.22.1APIchanges  
...........................................  
323  
1.22.2DisplayChanges  
........................................  
327  
1.22.3TextParsingAPIChanges  
...................................  
329  
1.22.4GroupbyAPIChanges  
.....................................  
329  
1.22.5SQL  
...............................................  
332  
1.22.6MultiIndexingUsingSlicers  
..................................  
333  
1.22.7Plotting  
.............................................  
338  
1.22.8PriorVersionDeprecations/Changes  
..............................  
339  
1.22.9Deprecations  
..........................................  
339  
1.22.10KnownIssues  
..........................................  
340  
1.22.11Enhancements  
..........................................  
340  
1.22.12Performance  
...........................................  
345  
1.22.13Experimental  
..........................................  
345  
1.22.14BugFixes  
............................................  
345  
1.23v0.13.1(February3,2014)  
........................................  
350  
1.23.1OutputFormattingEnhancements  
...............................  
351  
1.23.2APIchanges  
...........................................  
352  
1.23.3PriorVersionDeprecations/Changes  
..............................  
354  
1.23.4Deprecations  
..........................................  
354  
1.23.5Enhancements  
..........................................  
354  
1.23.6Performance  
...........................................  
357  
1.23.7Experimental  
..........................................  
358  
1.23.8BugFixes  
............................................  
358  
1.24v0.13.0(January3,2014)  
........................................  
358  
1.24.1APIchanges  
...........................................  
358  
1.24.2PriorVersionDeprecations/Changes  
..............................  
361  
1.24.3Deprecations  
..........................................  
361  
1.24.4IndexingAPIChanges  
.....................................  
361  
1.24.5Float64IndexAPIChange  
...................................  
363  
1.24.6HDFStoreAPIChanges  
....................................  
365  
1.24.7DataFramereprChanges  
....................................  
367  
viii

1.24.8Enhancements  
..........................................  
368  
1.24.9Experimental  
..........................................  
375  
1.24.10InternalRefactoring  
.......................................  
379  
1.24.11BugFixes  
............................................  
381  
1.25v0.12.0(July24,2013)  
..........................................  
381  
1.25.1APIchanges  
...........................................  
381  
1.25.2I/OEnhancements  
.......................................  
385  
1.25.3OtherEnhancements  
......................................  
387  
1.25.4ExperimentalFeatures  
.....................................  
389  
1.25.5BugFixes  
............................................  
390  
1.26v0.11.0(April22,2013)  
.........................................  
392  
1.26.1SelectionChoices  
........................................  
393  
1.26.2SelectionDeprecations  
.....................................  
393  
1.26.3Dtypes  
..............................................  
394  
1.26.4DtypeConversion  
........................................  
395  
1.26.5DtypeGotchas  
.........................................  
396  
1.26.6DatetimesConversion  
......................................  
398  
1.26.7APIchanges  
...........................................  
400  
1.26.8Enhancements  
..........................................  
400  
1.27v0.10.1(January22,2013)  
........................................  
402  
1.27.1APIchanges  
...........................................  
402  
1.27.2Newfeatures  
..........................................  
403  
1.27.3HDFStore  
............................................  
403  
1.28v0.10.0(December17,2012)  
......................................  
408  
1.28.1Fileparsingnewfeatures  
....................................  
408  
1.28.2APIchanges  
...........................................  
409  
1.28.3Newfeatures  
..........................................  
414  
1.28.4WideDataFramePrinting  
....................................  
414  
1.28.5UpdatedPyTablesSupport  
...................................  
416  
1.28.6NDimensionalPanels(Experimental)  
.............................  
419  
1.29v0.9.1(November14,2012)  
.......................................  
420  
1.29.1Newfeatures  
..........................................  
420  
1.29.2APIchanges  
...........................................  
423  
1.30v0.9.0(October7,2012)  
.........................................  
424  
1.30.1Newfeatures  
..........................................  
424  
1.30.2APIchanges  
...........................................  
425  
1.31v0.8.1(July22,2012)  
..........................................  
426  
1.31.1Newfeatures  
..........................................  
426  
1.31.2Performanceimprovements  
...................................  
426  
1.32v0.8.0(June29,2012)  
..........................................  
426  
1.32.1Supportfornon-uniqueindexes  
................................  
427  
1.32.2NumPydatetime64dtypeand1.6dependency  
.........................  
427  
1.32.3Timeserieschangesandimprovements  
............................  
427  
1.32.4Othernewfeatures  
.......................................  
428  
1.32.5Newplottingmethods  
......................................  
429  
1.32.6OtherAPIchanges  
.......................................  
429  
1.32.7Potentialportingissuesforpandas<=0.7.3users  
.......................  
429  
1.33v.0.7.3(April12,2012)  
.........................................  
431  
1.33.1Newfeatures  
..........................................  
432  
1.33.2NABooleanComparisonAPIChange  
.............................  
433  
1.33.3OtherAPIChanges  
.......................................  
434  
1.34v.0.7.2(March16,2012)  
.........................................  
435  
1.34.1Newfeatures  
..........................................  
435  
1.34.2Performanceimprovements  
...................................  
435  
ix

1.35v.0.7.1(February29,2012)  
.......................................  
435  
1.35.1Newfeatures  
..........................................  
436  
1.35.2Performanceimprovements  
...................................  
436  
1.36v.0.7.0(February9,2012)  
........................................  
436  
1.36.1Newfeatures  
..........................................  
436  
1.36.2APIChangestointegerindexing  
................................  
437  
1.36.3APItweaksregardinglabel-basedslicing  
...........................  
439  
1.36.4ChangestoSeries  
[]  
operator  
.................................  
440  
1.36.5OtherAPIChanges  
.......................................  
441  
1.36.6Performanceimprovements  
...................................  
441  
1.37v.0.6.1(December13,2011)  
.......................................  
442  
1.37.1Newfeatures  
..........................................  
442  
1.37.2Performanceimprovements  
...................................  
442  
1.38v.0.6.0(November25,2011)  
.......................................  
442  
1.38.1NewFeatures  
..........................................  
442  
1.38.2PerformanceEnhancements  
..................................  
444  
1.39v.0.5.0(October24,2011)  
........................................  
444  
1.39.1NewFeatures  
..........................................  
444  
1.39.2PerformanceEnhancements  
..................................  
445  
1.40v.0.4.3throughv0.4.1(September25-October9,2011)  
........................  
445  
1.40.1NewFeatures  
..........................................  
445  
1.40.2PerformanceEnhancements  
..................................  
446  
2Installation  
447  
2.1PlanfordroppingPython2.7  
.......................................  
447  
2.2Pythonversionsupport  
..........................................  
447  
2.3Installingpandas  
.............................................  
447  
2.3.1InstallingwithAnaconda  
....................................  
447  
2.3.2InstallingwithMiniconda  
....................................  
448  
2.3.3InstallingfromPyPI  
......................................  
449  
2.3.4InstallingwithActivePython  
..................................  
449  
2.3.5InstallingusingyourLinuxdistribution'spackagemanager.  
..................  
449  
2.3.6Installingfromsource  
......................................  
449  
2.4Runningthetestsuite  
..........................................  
449  
2.5Dependencies  
...............................................  
450  
2.5.1RecommendedDependencies  
..................................  
450  
2.5.2OptionalDependencies  
.....................................  
450  
3Contributingtopandas  
453  
3.1Wheretostart?  
..............................................  
454  
3.2Bugreportsandenhancementrequests  
.................................  
454  
3.3Workingwiththecode  
..........................................  
455  
3.3.1Versioncontrol,Git,andGitHub  
................................  
455  
3.3.2GettingstartedwithGit  
.....................................  
455  
3.3.3Forking  
.............................................  
455  
3.3.4Creatingadevelopmentenvironment  
..............................  
455  
3.3.4.1InstallingaCCompiler  
................................  
456  
3.3.4.2CreatingaPythonEnvironment  
............................  
456  
3.3.4.3CreatingaPythonEnvironment(pip)  
.........................  
457  
3.3.5Creatingabranch  
........................................  
457  
3.4Contributingtothedocumentation  
....................................  
458  
3.4.1Aboutthe  
pandas  
documentation  
................................  
458  
3.4.1.1pandasdocstringguide  
................................  
458  
3.4.2Howtobuildthe  
pandas  
documentation  
............................  
474  
x

3.4.2.1Requirements  
.....................................  
474  
3.4.2.2Buildingthedocumentation  
..............................  
474  
3.4.2.3Buildingmasterbranchdocumentation  
........................  
475  
3.5Contributingtothecodebase  
......................................  
475  
3.5.1Codestandards  
.........................................  
476  
3.5.1.1C(cpplint)  
.......................................  
476  
3.5.1.2Python(PEP8)  
.....................................  
477  
3.5.1.3BackwardsCompatibility  
...............................  
477  
3.5.2TestingWithContinuousIntegration  
..............................  
478  
3.5.3Test-drivendevelopment/codewriting  
.............................  
478  
3.5.3.1Writingtests  
......................................  
479  
3.5.3.2Transitioningto  
pytest  
...............................  
479  
3.5.3.3Using  
pytest  
....................................  
480  
3.5.4Runningthetestsuite  
......................................  
481  
3.5.5Runningtheperformancetestsuite  
...............................  
482  
3.5.6Documentingyourcode  
....................................  
483  
3.6Contributingyourchangesto  
pandas  
..................................  
483  
3.6.1Committingyourcode  
.....................................  
483  
3.6.2Pushingyourchanges  
......................................  
484  
3.6.3Reviewyourcode  
........................................  
484  
3.6.4Finally,makethepullrequest  
..................................  
485  
3.6.5Updatingyourpullrequest  
...................................  
485  
3.6.6Deleteyourmergedbranch(optional)  
.............................  
485  
4Packageoverview  
487  
4.1DataStructures  
..............................................  
487  
4.1.1Whymorethanonedatastructure?  
...............................  
487  
4.2Mutabilityandcopyingofdata  
......................................  
488  
4.3GettingSupport  
.............................................  
488  
4.4Community  
................................................  
488  
4.5ProjectGovernance  
............................................  
488  
4.6DevelopmentTeam  
............................................  
488  
4.7InstitutionalPartners  
...........................................  
489  
4.8License  
..................................................  
489  
510Minutestopandas  
491  
5.1ObjectCreation  
..............................................  
491  
5.2ViewingData  
...............................................  
493  
5.3Selection  
.................................................  
494  
5.3.1Getting  
.............................................  
494  
5.3.2SelectionbyLabel  
.......................................  
495  
5.3.3SelectionbyPosition  
......................................  
496  
5.3.4BooleanIndexing  
........................................  
497  
5.3.5Setting  
..............................................  
498  
5.4MissingData  
...............................................  
499  
5.5Operations  
................................................  
500  
5.5.1Stats  
...............................................  
500  
5.5.2Apply  
..............................................  
501  
5.5.3Histogramming  
.........................................  
501  
5.5.4StringMethods  
.........................................  
502  
5.6Merge  
...................................................  
502  
5.6.1Concat  
..............................................  
502  
5.6.2Join  
...............................................  
503  
5.6.3Append  
.............................................  
504  
xi

5.7Grouping  
.................................................  
505  
5.8Reshaping  
................................................  
505  
5.8.1Stack  
..............................................  
506  
5.8.2PivotTables  
...........................................  
507  
5.9TimeSeries  
................................................  
508  
5.10Categoricals  
...............................................  
510  
5.11Plotting  
..................................................  
511  
5.12GettingDataIn/Out  
...........................................  
512  
5.12.1CSV  
...............................................  
512  
5.12.2HDF5  
..............................................  
513  
5.12.3Excel  
..............................................  
513  
5.13Gotchas  
..................................................  
514  
6Tutorials  
515  
6.1InternalGuides  
..............................................  
515  
6.2pandasCookbook  
............................................  
515  
6.3Lessonsfornewpandasusers  
......................................  
516  
6.4PracticaldataanalysiswithPython  
...................................  
516  
6.5Exercisesfornewusers  
.........................................  
516  
6.6Modernpandas  
..............................................  
517  
6.7Excelchartswithpandas,vincentandxlsxwriter  
............................  
517  
6.8VideoTutorials  
..............................................  
517  
6.9VariousTutorials  
.............................................  
517  
7Cookbook  
519  
7.1Idioms  
..................................................  
519  
7.1.1if-then...  
............................................  
519  
7.1.2Splitting  
.............................................  
520  
7.1.3BuildingCriteria  
........................................  
521  
7.2Selection  
.................................................  
523  
7.2.1DataFrames  
...........................................  
523  
7.2.2Panels  
..............................................  
525  
7.2.3NewColumns  
..........................................  
525  
7.3MultiIndexing  
..............................................  
526  
7.3.1Arithmetic  
............................................  
527  
7.3.2Slicing  
..............................................  
528  
7.3.3Sorting  
.............................................  
530  
7.3.4Levels  
..............................................  
530  
7.4MissingData  
...............................................  
530  
7.4.1Replace  
.............................................  
531  
7.5Grouping  
.................................................  
531  
7.5.1ExpandingData  
.........................................  
536  
7.5.2Splitting  
.............................................  
536  
7.5.3Pivot  
...............................................  
537  
7.5.4Apply  
..............................................  
539  
7.6Timeseries  
................................................  
541  
7.6.1Resampling  
...........................................  
541  
7.7Merge  
...................................................  
542  
7.8Plotting  
..................................................  
543  
7.9DataIn/Out  
................................................  
544  
7.9.1CSV  
...............................................  
544  
7.9.1.1ReadingmultipletocreateasingleDataFrame  
.................  
545  
7.9.1.2Parsingdatecomponentsinmulti-columns  
......................  
545  
7.9.1.3Skiprowbetweenheaderanddata  
..........................  
546  
xii

7.9.2SQL  
...............................................  
547  
7.9.3Excel  
..............................................  
547  
7.9.4HTML  
..............................................  
547  
7.9.5HDFStore  
............................................  
547  
7.9.6BinaryFiles  
...........................................  
548  
7.10Computation  
...............................................  
549  
7.11Timedeltas  
................................................  
549  
7.12AliasingAxisNames  
...........................................  
551  
7.13CreatingExampleData  
..........................................  
552  
8IntrotoDataStructures  
553  
8.1Series  
...................................................  
553  
8.1.1Seriesisndarray-like  
......................................  
555  
8.1.2Seriesisdict-like  
........................................  
556  
8.1.3VectorizedoperationsandlabelalignmentwithSeries  
.....................  
557  
8.1.4Nameattribute  
.........................................  
558  
8.2DataFrame  
................................................  
558  
8.2.1FromdictofSeriesordicts  
...................................  
559  
8.2.2Fromdictofndarrays/lists  
...................................  
560  
8.2.3Fromstructuredorrecordarray  
.................................  
560  
8.2.4Fromalistofdicts  
.......................................  
561  
8.2.5Fromadictoftuples  
......................................  
561  
8.2.6FromaSeries  
..........................................  
561  
8.2.7AlternateConstructors  
.....................................  
561  
8.2.8Columnselection,addition,deletion  
..............................  
562  
8.2.9AssigningNewColumnsinMethodChains  
..........................  
564  
8.2.10Indexing/Selection  
.......................................  
566  
8.2.11Dataalignmentandarithmetic  
.................................  
567  
8.2.12Transposing  
...........................................  
570  
8.2.13DataFrameinteroperabilitywithNumPyfunctions  
......................  
570  
8.2.14Consoledisplay  
.........................................  
571  
8.2.15DataFramecolumnattributeaccessandIPythoncompletion  
.................  
574  
8.3Panel  
...................................................  
574  
8.3.1From3Dndarraywithoptionalaxislabels  
...........................  
575  
8.3.2FromdictofDataFrameobjects  
................................  
575  
8.3.3FromDataFrameusing  
to\_panel  
method  
..........................  
576  
8.3.4Itemselection/addition/deletion  
...............................  
577  
8.3.5Transposing  
...........................................  
577  
8.3.6Indexing/Selection  
.......................................  
577  
8.3.7Squeezing  
............................................  
578  
8.3.8ConversiontoDataFrame  
....................................  
578  
8.4DeprecatePanel  
.............................................  
579  
9EssentialBasicFunctionality  
583  
9.1HeadandTail  
...............................................  
583  
9.2Attributesandtherawndarray(s)  
....................................  
584  
9.3Acceleratedoperations  
..........................................  
585  
9.4Flexiblebinaryoperations  
........................................  
585  
9.4.1Matching/broadcastingbehavior  
...............................  
586  
9.4.2Missingdata/operationswithvalues  
............................  
589  
9.4.3FlexibleComparisons  
......................................  
590  
9.4.4BooleanReductions  
.......................................  
590  
9.4.5Comparingifobjectsareequivalent  
..............................  
592  
9.4.6Comparingarray-likeobjects  
..................................  
592  
xiii

9.4.7Combiningoverlappingdatasets  
................................  
593  
9.4.8GeneralDataFrameCombine  
..................................  
594  
9.5Descriptivestatistics  
...........................................  
595  
9.5.1Summarizingdata:describe  
..................................  
597  
9.5.2IndexofMin/MaxValues  
....................................  
600  
9.5.3Valuecounts(histogramming)/Mode  
.............................  
601  
9.5.4Discretizationandquantiling  
..................................  
602  
9.6Functionapplication  
...........................................  
603  
9.6.1TablewiseFunctionApplication  
................................  
603  
9.6.2RoworColumn-wiseFunctionApplication  
..........................  
605  
9.6.3AggregationAPI  
........................................  
607  
9.6.3.1Aggregatingwithmultiplefunctions  
.........................  
608  
9.6.3.2Aggregatingwithadict  
................................  
609  
9.6.3.3MixedDtypes  
.....................................  
609  
9.6.3.4Customdescribe  
....................................  
610  
9.6.4TransformAPI  
.........................................  
610  
9.6.4.1Transformwithmultiplefunctions  
..........................  
612  
9.6.4.2Transformingwithadict  
...............................  
613  
9.6.5ApplyingElementwiseFunctions  
................................  
613  
9.6.6ApplyingwithaPanel  
.....................................  
615  
9.7Reindexingandalteringlabels  
......................................  
617  
9.7.1Reindexingtoalignwithanotherobject  
............................  
619  
9.7.2Aligningobjectswitheachotherwith  
align  
.........................  
620  
9.7.3Fillingwhilereindexing  
....................................  
622  
9.7.4Limitsonwhilereindexing  
...............................  
623  
9.7.5Droppinglabelsfromanaxis  
..................................  
624  
9.7.6Renaming/mappinglabels  
...................................  
625  
9.8Iteration  
..................................................  
626  
9.8.1iteritems  
.............................................  
627  
9.8.2iterrows  
.............................................  
628  
9.8.3itertuples  
............................................  
629  
9.9.dtaccessor  
................................................  
630  
9.10Vectorizedstringmethods  
........................................  
633  
9.11Sorting  
..................................................  
634  
9.11.1ByIndex  
............................................  
634  
9.11.2ByValues  
............................................  
635  
9.11.3ByIndexesandValues  
.....................................  
636  
9.11.4searchsorted  
...........................................  
637  
9.11.5smallest/largestvalues  
.....................................  
637  
9.11.6Sortingbyamulti-indexcolumn  
................................  
639  
9.12Copying  
.................................................  
639  
9.13dtypes  
...................................................  
639  
9.13.1defaults  
.............................................  
642  
9.13.2upcasting  
............................................  
642  
9.13.3astype  
..............................................  
643  
9.13.4objectconversion  
........................................  
645  
9.13.5gotchas  
.............................................  
648  
9.14Selectingcolumnsbasedon  
dtype  
...................................  
649  
10WorkingwithTextData  
653  
10.1SplittingandReplacingStrings  
.....................................  
655  
10.2Concatenation  
..............................................  
658  
10.2.1ConcatenatingasingleSeriesintoastring  
...........................  
658  
10.2.2ConcatenatingaSeriesandsomethinglist-likeintoaSeries  
..................  
659  
xiv

10.2.3ConcatenatingaSeriesandsomethingarray-likeintoaSeries  
.................  
659  
10.2.4ConcatenatingaSeriesandanindexedobjectintoaSeries,withalignment  
..........  
660  
10.2.5ConcatenatingaSeriesandmanyobjectsintoaSeries  
.....................  
662  
10.3Indexingwith  
.str  
...........................................  
664  
10.4ExtractingSubstrings  
...........................................  
664  
10.4.1Extractmatchineachsubject(extract)  
..........................  
664  
10.4.2Extractallmatchesineachsubject(extractall)  
.........................  
666  
10.5TestingforStringsthatMatchorContainaPattern  
...........................  
668  
10.6CreatingIndicatorVariables  
.......................................  
669  
10.7MethodSummary  
............................................  
669  
11OptionsandSettings  
673  
11.1Overview  
.................................................  
673  
11.2GettingandSettingOptions  
.......................................  
674  
11.3SettingStartupOptionsinpython/ipythonEnvironment  
........................  
675  
11.4FrequentlyUsedOptions  
.........................................  
675  
11.5AvailableOptions  
............................................  
681  
11.6NumberFormatting  
...........................................  
682  
11.7UnicodeFormatting  
...........................................  
683  
11.8TableSchemaDisplay  
..........................................  
684  
12IndexingandSelectingData  
685  
12.1DifferentChoicesforIndexing  
......................................  
685  
12.2Basics  
...................................................  
686  
12.3AttributeAccess  
.............................................  
689  
12.4Slicingranges  
..............................................  
691  
12.5SelectionByLabel  
............................................  
692  
12.5.1Slicingwithlabels  
.......................................  
695  
12.6SelectionByPosition  
..........................................  
696  
12.7SelectionByCallable  
..........................................  
700  
12.8IXIndexerisDeprecated  
.........................................  
702  
12.9IndexingwithlistwithmissinglabelsisDeprecated  
..........................  
703  
12.9.1Reindexing  
...........................................  
704  
12.10SelectingRandomSamples  
.......................................  
705  
12.11SettingWithEnlargement  
........................................  
707  
12.12Fastscalarvaluegettingandsetting  
...................................  
708  
12.13Booleanindexing  
.............................................  
709  
12.14Indexingwithisin  
............................................  
710  
12.15The  
where()  
MethodandMasking  
..................................  
713  
12.15.1Mask  
..............................................  
716  
12.16The  
query()  
Method  
..........................................  
717  
12.16.1  
MultiIndexquery()  
Syntax  
...............................  
719  
12.16.2  
query()  
UseCases  
......................................  
720  
12.16.3  
query()  
PythonversuspandasSyntaxComparison  
.....................  
721  
12.16.4The  
in  
and  
notin  
operators  
.................................  
722  
12.16.5Specialuseofthe  
==  
operatorwith  
list  
objects  
.......................  
724  
12.16.6BooleanOperators  
.......................................  
726  
12.16.7Performanceof  
query()  
...................................  
727  
12.17DuplicateData  
..............................................  
727  
12.18Dictionary-like  
get()  
method  
.....................................  
730  
12.19The  
lookup()  
Method  
.........................................  
731  
12.20Indexobjects  
...............................................  
731  
12.20.1Settingmetadata  
........................................  
732  
12.20.2SetoperationsonIndexobjects  
.................................  
733  
xv

12.20.3Missingvalues  
.........................................  
733  
12.21Set/ResetIndex  
.............................................  
734  
12.21.1Setanindex  
...........................................  
734  
12.21.2Resettheindex  
.........................................  
735  
12.21.3Addinganadhocindex  
.....................................  
736  
12.22Returningaviewversusacopy  
.....................................  
736  
12.22.1Whydoesassignmentfailwhenusingchainedindexing?  
...................  
737  
12.22.2Evaluationordermatters  
....................................  
738  
13MultiIndex/AdvancedIndexing  
741  
13.1Hierarchicalindexing(MultiIndex)  
...................................  
741  
13.1.1CreatingaMultiIndex(hierarchicalindex)object  
.......................  
741  
13.1.2Reconstructingthelevellabels  
.................................  
744  
13.1.3BasicindexingonaxiswithMultiIndex  
............................  
744  
13.1.4Levels  
.........................................  
745  
13.1.5Dataalignmentandusing  
reindex  
..............................  
746  
13.2Advancedindexingwithhierarchicalindex  
...............................  
747  
13.2.1Usingslicers  
..........................................  
749  
13.2.2Cross-section  
..........................................  
753  
13.2.3Advancedreindexingandalignment  
..............................  
755  
13.2.4Swappinglevelswith  
swaplevel()  
.............................  
756  
13.2.5Reorderinglevelswith  
reorder\_levels()  
........................  
756  
13.3Sortinga  
MultiIndex  
.........................................  
757  
13.4TakeMethods  
..............................................  
760  
13.5IndexTypes  
................................................  
761  
13.5.1CategoricalIndex  
........................................  
761  
13.5.2Int64IndexandRangeIndex  
...................................  
764  
13.5.3Float64Index  
..........................................  
764  
13.5.4IntervalIndex  
..........................................  
767  
13.6MiscellaneousindexingFAQ  
.......................................  
768  
13.6.1Integerindexing  
.........................................  
768  
13.6.2Non-monotonicindexesrequireexactmatches  
.........................  
769  
13.6.3Endpointsareinclusive  
.....................................  
770  
13.6.4IndexingpotentiallychangesunderlyingSeriesdtype  
.....................  
771  
14Computationaltools  
773  
14.1StatisticalFunctions  
...........................................  
773  
14.1.1PercentChange  
.........................................  
773  
14.1.2Covariance  
...........................................  
773  
14.1.3Correlation  
...........................................  
775  
14.1.4Dataranking  
..........................................  
776  
14.2WindowFunctions  
............................................  
777  
14.2.1MethodSummary  
........................................  
780  
14.2.2RollingWindows  
........................................  
782  
14.2.3Time-awareRolling  
.......................................  
784  
14.2.4RollingWindowEndpoints  
...................................  
786  
14.2.5Time-awareRollingvs.Resampling  
..............................  
787  
14.2.6CenteringWindows  
.......................................  
787  
14.2.7BinaryWindowFunctions  
...................................  
788  
14.2.8Computingrollingpairwisecovariancesandcorrelations  
...................  
789  
14.3Aggregation  
...............................................  
791  
14.3.1Applyingmultiplefunctions  
..................................  
793  
14.3.2ApplyingdifferentfunctionstoDataFramecolumns  
......................  
794  
14.4ExpandingWindows  
...........................................  
795  
xvi

14.4.1MethodSummary  
........................................  
796  
14.5ExponentiallyWeightedWindows  
....................................  
798  
15Workingwithmissingdata  
803  
15.1Missingdatabasics  
............................................  
803  
15.1.1When/whydoesdatabecomemissing?  
............................  
803  
15.1.2Valuesconsideredﬁmissingﬂ  
..................................  
804  
15.2Datetimes  
.................................................  
805  
15.3Insertingmissingdata  
..........................................  
806  
15.4Calculationswithmissingdata  
......................................  
807  
15.4.1Sum/ProdofEmpties/Nans  
...................................  
808  
15.4.2NAvaluesinGroupBy  
.....................................  
809  
15.5Cleaning/missingdata  
......................................  
809  
15.5.1Fillingmissingvalues:  
..................................  
809  
15.5.2FillingwithaPandasObject  
...................................  
811  
15.5.3Droppingaxislabelswithmissingdata:dropna  
........................  
812  
15.5.4Interpolation  
..........................................  
813  
15.5.4.1InterpolationLimits  
..................................  
819  
15.5.5ReplacingGenericValues  
....................................  
821  
15.5.6String/RegularExpressionReplacement  
............................  
822  
15.5.7NumericReplacement  
.....................................  
824  
15.6Missingdatacastingrulesandindexing  
.................................  
826  
16GroupBy:split-apply-combine  
829  
16.1Splittinganobjectintogroups  
......................................  
830  
16.1.1GroupBysorting  
........................................  
832  
16.1.2GroupByobjectattributes  
....................................  
832  
16.1.3GroupBywithMultiIndex  
...................................  
833  
16.1.4GroupingDataFramewithIndexLevelsandColumns  
.....................  
835  
16.1.5DataFramecolumnselectioninGroupBy  
...........................  
836  
16.2Iteratingthroughgroups  
.........................................  
837  
16.3Selectingagroup  
.............................................  
838  
16.4Aggregation  
...............................................  
838  
16.4.1Applyingmultiplefunctionsatonce  
..............................  
840  
16.4.2ApplyingdifferentfunctionstoDataFramecolumns  
......................  
841  
16.4.3Cython-optimizedaggregationfunctions  
............................  
842  
16.5Transformation  
..............................................  
843  
16.5.1Newsyntaxtowindowandresampleoperations  
........................  
848  
16.6Filtration  
.................................................  
850  
16.7Dispatchingtoinstancemethods  
.....................................  
852  
16.8Flexible  
apply  
.............................................  
853  
16.9Otherusefulfeatures  
...........................................  
856  
16.9.1Automaticexclusionofﬁnuisanceﬂcolumns  
..........................  
856  
16.9.2Handlingof(un)observedCategoricalvalues  
.........................  
856  
16.9.3NAandNaTgrouphandling  
..................................  
857  
16.9.4Groupingwithorderedfactors  
.................................  
857  
16.9.5GroupingwithaGrouper  
.............................  
857  
16.9.6Takingtherowsofeachgroup  
...............................  
859  
16.9.7Takingthenthrowofeachgroup  
................................  
859  
16.9.8Enumerategroupitems  
.....................................  
861  
16.9.9Enumerategroups  
........................................  
862  
16.9.10Plotting  
.............................................  
863  
16.9.11Pipingfunctioncalls  
......................................  
864  
16.10Examples  
.................................................  
865  
xvii

16.10.1Regroupingbyfactor  
......................................  
865  
16.10.2Multi-columnfactorization  
...................................  
865  
16.10.3GroupbybyIndexerto`resample'data  
.............................  
866  
16.10.4ReturningaSeriestopropagatenames  
.............................  
866  
17Merge,join,andconcatenate  
869  
17.1Concatenatingobjects  
..........................................  
869  
17.1.1Setlogicontheotheraxes  
...................................  
872  
17.1.2Concatenatingusing  
append  
.................................  
873  
17.1.3Ignoringindexesontheconcatenationaxis  
...........................  
875  
17.1.4Concatenatingwithmixedndims  
................................  
875  
17.1.5Moreconcatenatingwithgroupkeys  
..............................  
877  
17.1.6AppendingrowstoaDataFrame  
................................  
880  
17.2Database-styleDataFramejoining/merging  
...............................  
881  
17.2.1Briefprimeronmergemethods(relationalalgebra)  
......................  
883  
17.2.2Checkingforduplicatekeys  
..................................  
886  
17.2.3Themergeindicator  
.......................................  
886  
17.2.4MergeDtypes  
..........................................  
887  
17.2.5Joiningonindex  
........................................  
890  
17.2.6Joiningkeycolumnsonanindex  
................................  
891  
17.2.7JoiningasingleIndextoaMulti-index  
.............................  
893  
17.2.8Joiningwithtwomulti-indexes  
.................................  
894  
17.2.9Mergingonacombinationofcolumnsandindexlevels  
....................  
894  
17.2.10Overlappingvaluecolumns  
...................................  
895  
17.2.11JoiningmultipleDataFrameorPanelobjects  
..........................  
896  
17.2.12MergingtogethervalueswithinSeriesorDataFramecolumns  
.................  
897  
17.3Timeseriesfriendlymerging  
.......................................  
897  
17.3.1MergingOrderedData  
.....................................  
897  
17.3.2MergingAsOf  
..........................................  
898  
18ReshapingandPivotTables  
901  
18.1ReshapingbypivotingDataFrameobjects  
................................  
901  
18.2Reshapingbystackingandunstacking  
..................................  
903  
18.2.1MultipleLevels  
.........................................  
907  
18.2.2MissingData  
..........................................  
908  
18.2.3WithaMultiIndex  
........................................  
910  
18.3ReshapingbyMelt  
............................................  
911  
18.4CombiningwithstatsandGroupBy  
...................................  
913  
18.5Pivottables  
................................................  
914  
18.5.1Addingmargins  
.........................................  
917  
18.6Crosstabulations  
.............................................  
917  
18.6.1Normalization  
..........................................  
918  
18.6.2AddingMargins  
.........................................  
919  
18.7Tiling  
...................................................  
919  
18.8Computingindicator/dummyvariables  
.................................  
920  
18.9Factorizingvalues  
............................................  
923  
19TimeSeries/Datefunctionality  
925  
19.1Overview  
.................................................  
926  
19.2Timestampsvs.TimeSpans  
.......................................  
926  
19.3ConvertingtoTimestamps  
........................................  
928  
19.3.1ProvidingaFormatArgument  
.................................  
928  
19.3.2AssemblingDatetimefromMultipleDataFrameColumns  
...................  
929  
19.3.3InvalidData  
...........................................  
929  
xviii

19.3.4EpochTimestamps  
.......................................  
929  
19.3.5FromTimestampstoEpoch  
...................................  
930  
19.3.6Usingthe  
origin  
Parameter  
.................................  
931  
19.4GeneratingRangesofTimestamps  
....................................  
931  
19.4.1CustomFrequencyRanges  
...................................  
933  
19.5TimestampLimitations  
..........................................  
934  
19.6Indexing  
.................................................  
934  
19.6.1PartialStringIndexing  
.....................................  
935  
19.6.2Slicevs.ExactMatch  
......................................  
940  
19.6.3ExactIndexing  
.........................................  
941  
19.6.4Truncating&FancyIndexing  
..................................  
942  
19.7Time/DateComponents  
.........................................  
943  
19.8DateOffsetObjects  
............................................  
944  
19.8.1ParametricOffsets  
.......................................  
946  
19.8.2UsingOffsetswith  
Series  
/  
DatetimeIndex  
.......................  
947  
19.8.3CustomBusinessDays  
.....................................  
948  
19.8.4BusinessHour  
..........................................  
949  
19.8.5CustomBusinessHour  
.....................................  
951  
19.8.6OffsetAliases  
..........................................  
952  
19.8.7CombiningAliases  
.......................................  
953  
19.8.8AnchoredOffsets  
........................................  
954  
19.8.9AnchoredOffsetSemantics  
...................................  
955  
19.8.10Holidays/HolidayCalendars  
..................................  
956  
19.9TimeSeries-RelatedInstanceMethods  
.................................  
959  
19.9.1Shifting/Lagging  
........................................  
959  
19.9.2FrequencyConversion  
.....................................  
960  
19.9.3FillingForward/Backward  
...................................  
960  
19.9.4ConvertingtoPythonDatetimes  
................................  
960  
19.10Resampling  
................................................  
961  
19.10.1Basics  
..............................................  
961  
19.10.2Upsampling  
...........................................  
963  
19.10.3SparseResampling  
.......................................  
964  
19.10.4Aggregation  
...........................................  
965  
19.11TimeSpanRepresentation  
........................................  
968  
19.11.1Period  
..............................................  
968  
19.11.2PeriodIndexandperiod\_range  
.................................  
969  
19.11.3PeriodDtypes  
..........................................  
971  
19.11.4PeriodIndexPartialStringIndexing  
..............................  
972  
19.11.5FrequencyConversionandResamplingwithPeriodIndex  
...................  
974  
19.12ConvertingBetweenRepresentations  
..................................  
975  
19.13RepresentingOut-of-BoundsSpans  
...................................  
976  
19.14TimeZoneHandling  
...........................................  
977  
19.14.1WorkingwithTimeZones  
...................................  
977  
19.14.2AmbiguousTimeswhenLocalizing  
..............................  
982  
19.14.3TZAwareDtypes  
........................................  
983  
20TimeDeltas  
987  
20.1Parsing  
..................................................  
987  
20.1.1to\_timedelta  
...........................................  
988  
20.1.2Timedeltalimitations  
......................................  
989  
20.2Operations  
................................................  
989  
20.3Reductions  
................................................  
993  
20.4FrequencyConversion  
..........................................  
994  
20.5Attributes  
.................................................  
996  
xix

20.6TimedeltaIndex  
..............................................  
997  
20.6.1UsingtheTimedeltaIndex  
....................................  
998  
20.6.2Operations  
............................................  
1000  
20.6.3Conversions  
...........................................  
1000  
20.7Resampling  
................................................  
1001  
21CategoricalData  
1003  
21.1ObjectCreation  
..............................................  
1003  
21.1.1SeriesCreation  
.........................................  
1003  
21.1.2DataFrameCreation  
.......................................  
1005  
21.1.3ControllingBehavior  
......................................  
1006  
21.1.4RegainingOriginalData  
....................................  
1008  
21.2CategoricalDtype  
.............................................  
1008  
21.2.1EqualitySemantics  
.......................................  
1009  
21.3Description  
................................................  
1010  
21.4Workingwithcategories  
.........................................  
1010  
21.4.1Renamingcategories  
......................................  
1011  
21.4.2Appendingnewcategories  
...................................  
1013  
21.4.3Removingcategories  
......................................  
1013  
21.4.4Removingunusedcategories  
..................................  
1013  
21.4.5Settingcategories  
........................................  
1014  
21.5SortingandOrder  
............................................  
1015  
21.5.1Reordering  
...........................................  
1016  
21.5.2MultiColumnSorting  
......................................  
1017  
21.6Comparisons  
...............................................  
1018  
21.7Operations  
................................................  
1020  
21.8Datamunging  
..............................................  
1021  
21.8.1Getting  
.............................................  
1021  
21.8.2Stringanddatetimeaccessors  
..................................  
1023  
21.8.3Setting  
..............................................  
1024  
21.8.4Merging  
.............................................  
1026  
21.8.5Unioning  
............................................  
1026  
21.8.6Concatenation  
..........................................  
1028  
21.9GettingDataIn/Out  
...........................................  
1029  
21.10MissingData  
...............................................  
1031  
21.11DifferencestoR's  
factor  
.........................................  
1032  
21.12Gotchas  
..................................................  
1032  
21.12.1MemoryUsage  
.........................................  
1032  
21.12.2  
Categorical  
isnota  
numpy  
array  
................................  
1033  
21.12.3dtypeinapply  
..........................................  
1034  
21.12.4CategoricalIndex  
........................................  
1034  
21.12.5SideEffects  
...........................................  
1035  
22Visualization  
1037  
22.1BasicPlotting:  
plot  
...........................................  
1037  
22.2OtherPlots  
................................................  
1040  
22.2.1Barplots  
.............................................  
1042  
22.2.2Histograms  
...........................................  
1045  
22.2.3BoxPlots  
............................................  
1051  
22.2.4AreaPlot  
............................................  
1059  
22.2.5ScatterPlot  
...........................................  
1061  
22.2.6HexagonalBinPlot  
.......................................  
1065  
22.2.7Pieplot  
.............................................  
1067  
22.3PlottingwithMissingData  
........................................  
1071  
xx

22.4PlottingTools  
...............................................  
1072  
22.4.1ScatterMatrixPlot  
.......................................  
1072  
22.4.2DensityPlot  
...........................................  
1073  
22.4.3AndrewsCurves  
........................................  
1074  
22.4.4ParallelCoordinates  
.......................................  
1075  
22.4.5LagPlot  
.............................................  
1076  
22.4.6AutocorrelationPlot  
......................................  
1077  
22.4.7BootstrapPlot  
..........................................  
1078  
22.4.8RadViz  
.............................................  
1079  
22.5PlotFormatting  
..............................................  
1080  
22.5.1Settingtheplotstyle  
......................................  
1080  
22.5.2Generalplotstylearguments  
..................................  
1080  
22.5.3ControllingtheLegend  
.....................................  
1081  
22.5.4Scales  
..............................................  
1082  
22.5.5PlottingonaSecondaryY-axis  
.................................  
1083  
22.5.6SuppressingTickResolutionAdjustment  
...........................  
1086  
22.5.7AutomaticDateTickAdjustment  
................................  
1089  
22.5.8Subplots  
.............................................  
1089  
22.5.9UsingLayoutandTargetingMultipleAxes  
..........................  
1090  
22.5.10PlottingWithErrorBars  
....................................  
1093  
22.5.11PlottingTables  
.........................................  
1095  
22.5.12Colormaps  
............................................  
1098  
22.6Plottingdirectlywithmatplotlib  
.....................................  
1103  
22.7Trellisplottinginterface  
.........................................  
1104  
23Styling  
1105  
23.1BuildingStyles  
..............................................  
1105  
23.1.1BuildingStylesSummary  
....................................  
1107  
23.2FinerControl:Slicing  
..........................................  
1108  
23.3FinerControl:DisplayValues  
......................................  
1108  
23.4BuiltinStyles  
...............................................  
1109  
23.4.1Barcharts  
............................................  
1109  
23.5SharingStyles  
..............................................  
1111  
23.6OtherOptions  
..............................................  
1111  
23.6.1Precision  
............................................  
1111  
23.6.2Captions  
.............................................  
1112  
23.6.3TableStyles  
...........................................  
1112  
23.6.4HidingtheIndexorColumns  
..................................  
1112  
23.6.5CSSClasses  
...........................................  
1113  
23.6.6Limitations  
...........................................  
1113  
23.6.7Terms  
..............................................  
1113  
23.7Funstuff  
.................................................  
1113  
23.8ExporttoExcel  
..............................................  
1114  
23.9Extensibility  
...............................................  
1115  
23.9.1Subclassing  
...........................................  
1115  
24IOTools(Text,CSV,HDF5,...)  
1117  
24.1CSV&Text  
.............................................  
1117  
24.1.1Parsingoptions  
.........................................  
1117  
24.1.1.1Basic  
..........................................  
1118  
24.1.1.2ColumnandIndexLocationsandNames  
.......................  
1118  
24.1.1.3GeneralParsing  
............................  
1119  
24.1.1.4NAandMissingDataHandling  
............................  
1120  
24.1.1.5DatetimeHandling  
...................................  
1120  
xxi

24.1.1.6Iteration  
........................................  
1121  
24.1.1.7Quoting,Compression,andFileFormat  
.......................  
1121  
24.1.1.8ErrorHandling  
.....................................  
1122  
24.1.2Specifyingcolumndatatypes  
..................................  
1122  
24.1.3SpecifyingCategoricaldtype  
..................................  
1124  
24.1.4NamingandUsingColumns  
..................................  
1126  
24.1.4.1Handlingcolumnnames  
................................  
1126  
24.1.5Duplicatenamesparsing  
....................................  
1127  
24.1.5.1Filteringcolumns(  
usecols  
)  
............................  
1128  
24.1.6CommentsandEmptyLines  
..................................  
1128  
24.1.6.1Ignoringlinecommentsandemptylines  
.......................  
1128  
24.1.6.2Comments  
.......................................  
1130  
24.1.7DealingwithUnicodeData  
...................................  
1131  
24.1.8Indexcolumnsandtrailingdelimiters  
.............................  
1131  
24.1.9DateHandling  
..........................................  
1132  
24.1.9.1SpecifyingDateColumns  
...............................  
1132  
24.1.9.2DateParsingFunctions  
................................  
1135  
24.1.9.3InferringDatetimeFormat  
...............................  
1135  
24.1.9.4InternationalDateFormats  
..............................  
1136  
24.1.10Specifyingmethodforconversion  
........................  
1137  
24.1.11ThousandSeparators  
......................................  
1137  
24.1.12NAValues  
............................................  
1138  
24.1.13  
.............................................  
1138  
24.1.14ReturningSeries  
........................................  
1138  
24.1.15Booleanvalues  
.........................................  
1139  
24.1.16Handlingﬁbadﬂlines  
......................................  
1139  
24.1.17Dialect  
..............................................  
1140  
24.1.18QuotingandEscapeCharacters  
.................................  
1141  
24.1.19FileswithFixedWidthColumns  
................................  
1141  
24.1.20Indexes  
.............................................  
1143  
24.1.20.1Fileswithanﬁimplicitﬂindexcolumn  
.........................  
1143  
24.1.20.2Readinganindexwitha  
MultiIndex  
.......................  
1143  
24.1.20.3Readingcolumnswitha  
MultiIndex  
.......................  
1144  
24.1.21Automaticallyﬁsnifthedelimiter  
.............................  
1145  
24.1.22ReadingmultipletocreateasingleDataFrame  
......................  
1146  
24.1.23Iteratingthroughchunkbychunk  
.............................  
1146  
24.1.24Specifyingtheparserengine  
..................................  
1148  
24.1.25Readingremote  
......................................  
1148  
24.1.26WritingoutData  
........................................  
1148  
24.1.26.1WritingtoCSVformat  
................................  
1148  
24.1.26.2Writingaformattedstring  
...............................  
1149  
24.2JSON  
...................................................  
1149  
24.2.1WritingJSON  
..........................................  
1149  
24.2.1.1OrientOptions  
.....................................  
1150  
24.2.1.2DateHandling  
.....................................  
1152  
24.2.1.3FallbackBehavior  
...................................  
1153  
24.2.2ReadingJSON  
.........................................  
1153  
24.2.2.1DataConversion  
....................................  
1154  
24.2.2.2TheNumpyParameter  
.................................  
1157  
24.2.3Normalization  
..........................................  
1158  
24.2.4Linedelimitedjson  
.......................................  
1159  
24.2.5TableSchema  
..........................................  
1160  
24.3HTML  
..................................................  
1164  
24.3.1ReadingHTMLContent  
....................................  
1164  
xxii

24.3.2WritingtoHTML  
.....................................  
1168  
24.3.3HTMLTableParsingGotchas  
.................................  
1171  
24.4Excel  
................................................  
1172  
24.4.1ReadingExcelFiles  
.......................................  
1172  
24.4.1.1  
ExcelFile  
class  
...................................  
1173  
24.4.1.2SpecifyingSheets  
...................................  
1173  
24.4.1.3Readinga  
MultiIndex  
...............................  
1174  
24.4.1.4ParsingColumns  
...............................  
1175  
24.4.1.5ParsingDates  
.....................................  
1175  
24.4.1.6CellConverters  
....................................  
1176  
24.4.1.7dtype  
..................................  
1176  
24.4.2WritingExcelFiles  
.......................................  
1176  
24.4.2.1WritingExcelFilestoDisk  
..............................  
1176  
24.4.2.2WritingExcelFilestoMemory  
............................  
1177  
24.4.3Excelwriterengines  
......................................  
1177  
24.4.4StyleandFormatting  
......................................  
1178  
24.5Clipboard  
.................................................  
1178  
24.6Pickling  
..................................................  
1179  
24.6.1Compressedpickle  
.....................................  
1180  
24.7msgpack  
.................................................  
1182  
24.7.1Read/WriteAPI  
.........................................  
1184  
24.8HDF5(PyTables)  
.............................................  
1185  
24.8.1Read/WriteAPI  
.........................................  
1187  
24.8.2FixedFormat  
..........................................  
1188  
24.8.3TableFormat  
..........................................  
1189  
24.8.4HierarchicalKeys  
........................................  
1189  
24.8.5StoringTypes  
..........................................  
1191  
24.8.5.1StoringMixedTypesinaTable  
............................  
1191  
24.8.5.2StoringMulti-IndexDataFrames  
...........................  
1192  
24.8.6Querying  
............................................  
1193  
24.8.6.1QueryingaTable  
...................................  
1193  
24.8.6.2Usingtimedelta64[ns]  
.................................  
1196  
24.8.6.3Indexing  
........................................  
1197  
24.8.6.4QueryviaDataColumns  
...............................  
1198  
24.8.6.5Iterator  
.........................................  
1200  
24.8.6.6AdvancedQueries  
...................................  
1201  
24.8.6.7MultipleTableQueries  
................................  
1203  
24.8.7DeletefromaTable  
.......................................  
1205  
24.8.8Notes&Caveats  
........................................  
1205  
24.8.8.1Compression  
......................................  
1205  
24.8.8.2ptrepack  
........................................  
1206  
24.8.8.3Caveats  
.........................................  
1207  
24.8.9DataTypes  
............................................  
1207  
24.8.9.1CategoricalData  
....................................  
1207  
24.8.9.2StringColumns  
....................................  
1208  
24.8.10ExternalCompatibility  
.....................................  
1210  
24.8.11Performance  
...........................................  
1211  
24.9Feather  
..................................................  
1212  
24.10Parquet  
..................................................  
1213  
24.11SQLQueries  
...............................................  
1215  
24.11.1pandas.read\_sql\_table  
......................................  
1216  
24.11.2pandas.read\_sql\_query  
.....................................  
1217  
24.11.3pandas.read\_sql  
.........................................  
1218  
24.11.4pandas.DataFrame.to\_sql  
....................................  
1219  
xxiii

24.11.5WritingDataFrames  
......................................  
1221  
24.11.5.1SQLdatatypes  
....................................  
1222  
24.11.6ReadingTables  
.........................................  
1222  
24.11.7Schemasupport  
.........................................  
1223  
24.11.8Querying  
............................................  
1223  
24.11.9Engineconnectionexamples  
..................................  
1224  
24.11.10AdvancedSQLAlchemyqueries  
................................  
1225  
24.11.11Sqlitefallback  
..........................................  
1226  
24.12GoogleBigQuery  
.............................................  
1226  
24.13StataFormat  
...............................................  
1226  
24.13.1WritingtoStataformat  
.....................................  
1226  
24.13.2ReadingfromStataformat  
...................................  
1227  
24.13.2.1CategoricalData  
....................................  
1228  
24.14SASFormats  
...............................................  
1228  
24.15Otherformats  
.............................................  
1229  
24.15.1netCDF  
.............................................  
1229  
24.16PerformanceConsiderations  
.......................................  
1229  
25EnhancingPerformance  
1233  
25.1Cython(WritingCextensionsforpandas)  
................................  
1233  
25.1.1Purepython  
...........................................  
1233  
25.1.2PlainCython  
..........................................  
1234  
25.1.3Addingtype  
...........................................  
1235  
25.1.4Usingndarray  
..........................................  
1236  
25.1.5Moreadvancedtechniques  
...................................  
1237  
25.2UsingNumba  
...............................................  
1238  
25.2.1Jit  
................................................  
1238  
25.2.2Vectorize  
............................................  
1239  
25.2.3Caveats  
.............................................  
1239  
25.3ExpressionEvaluationvia  
eval()  
...................................  
1240  
25.3.1SupportedSyntax  
........................................  
1240  
25.3.2  
eval()  
Examples  
.......................................  
1241  
25.3.3The  
DataFrame.eval  
method  
...............................  
1242  
25.3.4LocalVariables  
.........................................  
1244  
25.3.5  
pandas.eval()  
Parsers  
...................................  
1246  
25.3.6  
pandas.eval()  
Backends  
.................................  
1246  
25.3.7  
pandas.eval()  
Performance  
................................  
1247  
25.3.8TechnicalMinutiaRegardingExpressionEvaluation  
.....................  
1248  
26Sparsedatastructures  
1249  
26.1SparseArray  
...............................................  
1251  
26.2SparseIndexobjects  
...........................................  
1251  
26.3SparseDtypes  
..............................................  
1251  
26.4SparseCalculation  
............................................  
1254  
26.5Interactionwithscipy.sparse  
.......................................  
1254  
26.5.1SparseDataFrame  
........................................  
1254  
26.5.2SparseSeries  
...........................................  
1255  
27FrequentlyAskedQuestions(FAQ)  
1259  
27.1DataFramememoryusage  
........................................  
1259  
27.2UsingIf/TruthStatementswithpandas  
.................................  
1261  
27.2.1Bitwiseboolean  
.........................................  
1261  
27.2.2Usingthe  
in  
operator  
......................................  
1262  
27.3  
NaN  
,Integer  
NA  
valuesand  
NA  
typepromotions  
............................  
1262  
xxiv

27.3.1Choiceof  
NA  
representation  
..................................  
1262  
27.3.2Supportforinteger  
NA  
.....................................  
1263  
27.3.3  
NA  
typepromotions  
.......................................  
1263  
27.3.4WhynotmakeNumPylikeR?  
.................................  
1264  
27.4DifferenceswithNumPy  
.........................................  
1264  
27.5Thread-safety  
...............................................  
1264  
27.6Byte-OrderingIssues  
...........................................  
1264  
28rpy2/Rinterface  
1267  
28.1TransferringRdatasetsintoPython  
...................................  
1267  
28.2ConvertingDataFramesintoRobjects  
..................................  
1268  
29pandasEcosystem  
1269  
29.1StatisticsandMachineLearning  
.....................................  
1269  
29.1.1Statsmodels  
...........................................  
1269  
29.1.2sklearn-pandas  
.........................................  
1269  
29.2Visualization  
...............................................  
1269  
29.2.1Bokeh  
..............................................  
1269  
29.2.2seaborn  
.............................................  
1269  
29.2.3yhat/ggplot  
...........................................  
1270  
29.2.4Vincent  
.............................................  
1270  
29.2.5IPythonVega  
..........................................  
1270  
29.2.6Plotly  
..............................................  
1270  
29.2.7QtPandas  
............................................  
1270  
29.3IDE  
....................................................  
1270  
29.3.1IPython  
.............................................  
1270  
29.3.2quantopian/qgrid  
........................................  
1271  
29.3.3Spyder  
..............................................  
1271  
29.4API  
....................................................  
1271  
29.4.1pandas-datareader  
........................................  
1271  
29.4.2quandl/Python  
..........................................  
1271  
29.4.3pydatastream  
..........................................  
1271  
29.4.4pandaSDMX  
..........................................  
1271  
29.4.5fredapi  
..............................................  
1272  
29.5Domain  
.............................................  
1272  
29.5.1Geopandas  
...........................................  
1272  
29.5.2xarray  
..............................................  
1272  
29.6Out-of-core  
................................................  
1272  
29.6.1Dask  
...............................................  
1272  
29.6.2Dask-ML  
............................................  
1272  
29.6.3Blaze  
..............................................  
1272  
29.6.4Odo  
...............................................  
1272  
29.7Datavalidation  
..............................................  
1273  
29.7.1Engarde  
.............................................  
1273  
29.8ExtensionDataTypes  
..........................................  
1273  
29.8.1cyberpandas  
...........................................  
1273  
29.9Accessors  
.................................................  
1273  
30ComparisonwithR/Rlibraries  
1275  
30.1QuickReference  
.............................................  
1275  
30.1.1Querying,Filtering,Sampling  
.................................  
1275  
30.1.2Sorting  
.............................................  
1276  
30.1.3Transforming  
..........................................  
1276  
30.1.4GroupingandSummarizing  
..................................  
1276  
xxv

30.2BaseR  
..................................................  
1276  
30.2.1SlicingwithR's  
c  
........................................  
1276  
30.2.2  
aggregate  
..........................................  
1278  
30.2.3  
match  
/  
%in%  
.........................................  
1279  
30.2.4  
tapply  
.............................................  
1279  
30.2.5  
subset  
.............................................  
1280  
30.2.6  
with  
..............................................  
1280  
30.3plyr  
....................................................  
1281  
30.3.1  
ddply  
.............................................  
1282  
30.4reshape/reshape2  
............................................  
1282  
30.4.1  
melt.array  
.........................................  
1282  
30.4.2  
melt.list  
..........................................  
1283  
30.4.3  
melt.data.frame  
.....................................  
1283  
30.4.4  
cast  
..............................................  
1284  
30.4.5  
factor  
.............................................  
1286  
31ComparisonwithSQL  
1287  
31.1SELECT  
.................................................  
1287  
31.2WHERE  
.................................................  
1288  
31.3GROUPBY  
...............................................  
1290  
31.4JOIN  
...................................................  
1292  
31.4.1INNERJOIN  
..........................................  
1292  
31.4.2LEFTOUTERJOIN  
......................................  
1293  
31.4.3RIGHTJOIN  
..........................................  
1293  
31.4.4FULLJOIN  
...........................................  
1293  
31.5UNION  
..................................................  
1294  
31.6PandasequivalentsforsomeSQLanalyticandaggregatefunctions  
..................  
1295  
31.6.1TopNrowswithoffset  
.....................................  
1295  
31.6.2TopNrowspergroup  
......................................  
1295  
31.7UPDATE  
.................................................  
1297  
31.8DELETE  
.................................................  
1297  
32ComparisonwithSAS  
1299  
32.1DataStructures  
..............................................  
1299  
32.1.1GeneralTerminologyTranslation  
................................  
1299  
32.1.2  
DataFrame  
/  
Series  
....................................  
1299  
32.1.3  
Index  
.............................................  
1300  
32.2DataInput/Output  
............................................  
1300  
32.2.1ConstructingaDataFramefromValues  
............................  
1300  
32.2.2ReadingExternalData  
.....................................  
1300  
32.2.3ExportingData  
.........................................  
1301  
32.3DataOperations  
.............................................  
1301  
32.3.1OperationsonColumns  
.....................................  
1301  
32.3.2Filtering  
.............................................  
1302  
32.3.3If/ThenLogic  
..........................................  
1302  
32.3.4DateFunctionality  
.......................................  
1303  
32.3.5SelectionofColumns  
......................................  
1304  
32.3.6SortingbyValues  
........................................  
1305  
32.4StringProcessing  
.............................................  
1305  
32.4.1Length  
..............................................  
1305  
32.4.2Find  
...............................................  
1306  
32.4.3Substring  
............................................  
1306  
32.4.4Scan  
...............................................  
1306  
32.4.5Upcase,Lowcase,andPropcase  
................................  
1307  
xxvi

32.5Merging  
.................................................  
1308  
32.6MissingData  
...............................................  
1309  
32.7GroupBy  
.................................................  
1311  
32.7.1Aggregation  
...........................................  
1311  
32.7.2Transformation  
.........................................  
1311  
32.7.3ByGroupProcessing  
......................................  
1312  
32.8OtherConsiderations  
...........................................  
1313  
32.8.1DiskvsMemory  
........................................  
1313  
32.8.2DataInterop  
...........................................  
1313  
33ComparisonwithStata  
1315  
33.1DataStructures  
..............................................  
1315  
33.1.1GeneralTerminologyTranslation  
................................  
1315  
33.1.2  
DataFrame  
/  
Series  
....................................  
1315  
33.1.3  
Index  
.............................................  
1316  
33.2DataInput/Output  
............................................  
1316  
33.2.1ConstructingaDataFramefromValues  
............................  
1316  
33.2.2ReadingExternalData  
.....................................  
1316  
33.2.3ExportingData  
.........................................  
1317  
33.3DataOperations  
.............................................  
1317  
33.3.1OperationsonColumns  
.....................................  
1317  
33.3.2Filtering  
.............................................  
1318  
33.3.3If/ThenLogic  
..........................................  
1318  
33.3.4DateFunctionality  
.......................................  
1319  
33.3.5SelectionofColumns  
......................................  
1319  
33.3.6SortingbyValues  
........................................  
1320  
33.4StringProcessing  
.............................................  
1321  
33.4.1FindingLengthofString  
....................................  
1321  
33.4.2FindingPositionofSubstring  
..................................  
1321  
33.4.3ExtractingSubstringbyPosition  
................................  
1322  
33.4.4ExtractingnthWord  
.......................................  
1322  
33.4.5ChangingCase  
.........................................  
1323  
33.5Merging  
.................................................  
1323  
33.6MissingData  
...............................................  
1325  
33.7GroupBy  
.................................................  
1327  
33.7.1Aggregation  
...........................................  
1327  
33.7.2Transformation  
.........................................  
1327  
33.7.3ByGroupProcessing  
......................................  
1328  
33.8OtherConsiderations  
...........................................  
1328  
33.8.1DiskvsMemory  
........................................  
1328  
34APIReference  
1329  
34.1Input/Output  
...............................................  
1329  
34.1.1Pickling  
.............................................  
1329  
34.1.1.1pandas.read\_pickle  
..................................  
1329  
34.1.2FlatFile  
.............................................  
1330  
34.1.2.1pandas.read\_table  
...................................  
1331  
34.1.2.2pandas.read\_csv  
....................................  
1336  
34.1.2.3pandas.read\_fwf  
....................................  
1341  
34.1.2.4pandas.read\_msgpack  
.................................  
1347  
34.1.3Clipboard  
............................................  
1347  
34.1.3.1pandas.read\_clipboard  
.................................  
1347  
34.1.4Excel  
..............................................  
1347  
34.1.4.1pandas.read\_excel  
...................................  
1348  
xxvii

34.1.4.2pandas.ExcelFile.parse  
................................  
1351  
34.1.5JSON  
..............................................  
1351  
34.1.5.1pandas.read\_json  
...................................  
1352  
34.1.5.2pandas.io.json.json\_normalize  
.............................  
1355  
34.1.5.3pandas.io.json.build\_table\_schema  
..........................  
1356  
34.1.6HTML  
..............................................  
1357  
34.1.6.1pandas.read\_html  
...................................  
1357  
34.1.7HDFStore:PyTables(HDF5)  
..................................  
1359  
34.1.7.1pandas.read\_hdf  
....................................  
1360  
34.1.7.2pandas.HDFStore.put  
.................................  
1361  
34.1.7.3pandas.HDFStore.append  
...............................  
1361  
34.1.7.4pandas.HDFStore.get  
.................................  
1362  
34.1.7.5pandas.HDFStore.select  
................................  
1362  
34.1.7.6pandas.HDFStore.info  
.................................  
1363  
34.1.7.7pandas.HDFStore.keys  
................................  
1363  
34.1.8Feather  
.............................................  
1363  
34.1.8.1pandas.read\_feather  
..................................  
1363  
34.1.9Parquet  
.............................................  
1363  
34.1.9.1pandas.read\_parquet  
..................................  
1364  
34.1.10SAS  
...............................................  
1364  
34.1.10.1pandas.read\_sas  
....................................  
1364  
34.1.11SQL  
...............................................  
1365  
34.1.12GoogleBigQuery  
........................................  
1365  
34.1.12.1pandas.read\_gbq  
....................................  
1365  
34.1.13STATA  
.............................................  
1366  
34.1.13.1pandas.read\_stata  
...................................  
1366  
34.1.13.2pandas.io.stata.StataReader.data  
............................  
1368  
34.1.13.3pandas.io.stata.StataReader.data\_label  
........................  
1368  
34.1.13.4pandas.io.stata.StataReader.value\_labels  
.......................  
1368  
34.1.13.5pandas.io.stata.StataReader.variable\_labels  
......................  
1368  
34.1.13.6pandas.io.stata.StataWriter  
.........................  
1369  
34.2Generalfunctions  
.............................................  
1369  
34.2.1Datamanipulations  
.......................................  
1369  
34.2.1.1pandas.melt  
......................................  
1369  
34.2.1.2pandas.pivot  
......................................  
1371  
34.2.1.3pandas.pivot\_table  
...................................  
1371  
34.2.1.4pandas.crosstab  
....................................  
1373  
34.2.1.5pandas.cut  
.......................................  
1375  
34.2.1.6pandas.qcut  
......................................  
1378  
34.2.1.7pandas.merge  
.....................................  
1379  
34.2.1.8pandas.merge\_ordered  
.................................  
1381  
34.2.1.9pandas.merge\_asof  
..................................  
1382  
34.2.1.10pandas.concat  
.....................................  
1386  
34.2.1.11pandas.get\_dummies  
.................................  
1389  
34.2.1.12pandas.factorize  
....................................  
1391  
34.2.1.13pandas.unique  
.....................................  
1393  
34.2.1.14pandas.wide\_to\_long  
.................................  
1394  
34.2.2Top-levelmissingdata  
.....................................  
1398  
34.2.2.1pandas.isna  
.......................................  
1399  
34.2.2.2pandas.isnull  
......................................  
1400  
34.2.2.3pandas.notna  
......................................  
1401  
34.2.2.4pandas.notnull  
.....................................  
1402  
34.2.3Top-levelconversions  
......................................  
1404  
34.2.3.1pandas.to\_numeric  
...................................  
1404  
xxviii

34.2.4Top-leveldealingwithdatetimelike  
..............................  
1405  
34.2.4.1pandas.to\_datetime  
..................................  
1406  
34.2.4.2pandas.to\_timedelta  
..................................  
1409  
34.2.4.3pandas.date\_range  
...................................  
1410  
34.2.4.4pandas.bdate\_range  
..................................  
1412  
34.2.4.5pandas.period\_range  
..................................  
1413  
34.2.4.6pandas.timedelta\_range  
................................  
1414  
34.2.4.7pandas.infer\_freq  
...................................  
1415  
34.2.5Top-leveldealingwithintervals  
................................  
1415  
34.2.5.1pandas.interval\_range  
.................................  
1416  
34.2.6Top-levelevaluation  
.......................................  
1417  
34.2.6.1pandas.eval  
.......................................  
1417  
34.2.7Testing  
.............................................  
1419  
34.2.7.1pandas.test  
.......................................  
1419  
34.3Series  
...................................................  
1419  
34.3.1Constructor  
...........................................  
1419  
34.3.1.1pandas.Series  
.....................................  
1419  
34.3.2Attributes  
............................................  
1621  
34.3.2.1pandas.Series.empty  
..................................  
1622  
34.3.2.2pandas.Series.is\_copy  
.................................  
1622  
34.3.2.3pandas.Series.name  
..................................  
1622  
34.3.3Conversion  
...........................................  
1622  
34.3.4Indexing,iteration  
........................................  
1622  
34.3.4.1pandas.Series.\_\_iter\_\_  
.................................  
1623  
34.3.5Binaryoperatorfunctions  
....................................  
1623  
34.3.6Functionapplication,GroupBy&Window  
...........................  
1624  
34.3.7Computations/DescriptiveStats  
................................  
1625  
34.3.8Reindexing/Selection/Labelmanipulation  
..........................  
1626  
34.3.9Missingdatahandling  
......................................  
1627  
34.3.10Reshaping,sorting  
.......................................  
1627  
34.3.11Combining/joining/merging  
.................................  
1628  
34.3.12Timeseries-related  
.......................................  
1628  
34.3.13DatetimelikeProperties  
.....................................  
1628  
34.3.13.1pandas.Series.dt.date  
.................................  
1629  
34.3.13.2pandas.Series.dt.time  
.................................  
1629  
34.3.13.3pandas.Series.dt.year  
.................................  
1629  
34.3.13.4pandas.Series.dt.month  
................................  
1629  
34.3.13.5pandas.Series.dt.day  
..................................  
1629  
34.3.13.6pandas.Series.dt.hour  
.................................  
1630  
34.3.13.7pandas.Series.dt.minute  
................................  
1630  
34.3.13.8pandas.Series.dt.second  
................................  
1630  
34.3.13.9pandas.Series.dt.microsecond  
.............................  
1630  
34.3.13.10pandas.Series.dt.nanosecond  
.............................  
1630  
34.3.13.11pandas.Series.dt.week  
.................................  
1630  
34.3.13.12pandas.Series.dt.weekofyear  
.............................  
1630  
34.3.13.13pandas.Series.dt.dayofweek  
..............................  
1630  
34.3.13.14pandas.Series.dt.weekday  
...............................  
1630  
34.3.13.15pandas.Series.dt.dayofyear  
..............................  
1631  
34.3.13.16pandas.Series.dt.quarter  
................................  
1631  
34.3.13.17pandas.Series.dt.is\_month\_start  
............................  
1631  
34.3.13.18pandas.Series.dt.is\_month\_end  
............................  
1631  
34.3.13.19pandas.Series.dt.is\_quarter\_start  
...........................  
1632  
34.3.13.20pandas.Series.dt.is\_quarter\_end  
............................  
1632  
34.3.13.21pandas.Series.dt.is\_year\_start  
.............................  
1633  
xxix

34.3.13.22pandas.Series.dt.is\_year\_end  
.............................  
1634  
34.3.13.23pandas.Series.dt.is\_leap\_year  
.............................  
1634  
34.3.13.24pandas.Series.dt.daysinmonth  
.............................  
1635  
34.3.13.25pandas.Series.dt.days\_in\_month  
...........................  
1635  
34.3.13.26pandas.Series.dt.tz  
...................................  
1635  
34.3.13.27pandas.Series.dt.freq  
..................................  
1635  
34.3.13.28pandas.Series.dt.to\_period  
...............................  
1636  
34.3.13.29pandas.Series.dt.to\_pydatetime  
............................  
1637  
34.3.13.30pandas.Series.dt.tz\_localize  
..............................  
1637  
34.3.13.31pandas.Series.dt.tz\_convert  
..............................  
1639  
34.3.13.32pandas.Series.dt.normalize  
..............................  
1640  
34.3.13.33pandas.Series.dt.strftime  
................................  
1640  
34.3.13.34pandas.Series.dt.round  
.................................  
1641  
34.3.13.35  
.................................  
1642  
34.3.13.36pandas.Series.dt.ceil  
..................................  
1643  
34.3.13.37pandas.Series.dt.month\_name  
.............................  
1643  
34.3.13.38pandas.Series.dt.day\_name  
..............................  
1644  
34.3.13.39pandas.Series.dt.days  
.................................  
1644  
34.3.13.40pandas.Series.dt.seconds  
...............................  
1644  
34.3.13.41pandas.Series.dt.microseconds  
............................  
1644  
34.3.13.42pandas.Series.dt.nanoseconds  
.............................  
1644  
34.3.13.43pandas.Series.dt.components  
.............................  
1644  
34.3.13.44pandas.Series.dt.to\_pytimedelta  
............................  
1645  
34.3.13.45pandas.Series.dt.total\_seconds  
............................  
1645  
34.3.14Stringhandling  
.........................................  
1646  
34.3.14.1pandas.Series.str.capitalize  
..............................  
1648  
34.3.14.2pandas.Series.str.cat  
..................................  
1650  
34.3.14.3pandas.Series.str.center  
................................  
1652  
34.3.14.4pandas.Series.str.contains  
...............................  
1652  
34.3.14.5pandas.Series.str.count  
.................................  
1654  
34.3.14.6pandas.Series.str.decode  
................................  
1655  
34.3.14.7pandas.Series.str.encode  
................................  
1656  
34.3.14.8pandas.Series.str.endswith  
...............................  
1656  
34.3.14.9pandas.Series.str.extract  
................................  
1657  
34.3.14.10pandas.Series.str.extractall  
...............................  
1658  
34.3.14.11pandas.Series.str  
.................................  
1660  
34.3.14.12pandas.Series.str  
................................  
1660  
34.3.14.13pandas.Series.str.get  
..................................  
1661  
34.3.14.14pandas.Series.str.index  
.................................  
1662  
34.3.14.15pandas.Series.str.join  
.................................  
1663  
34.3.14.16pandas.Series.str.len  
..................................  
1664  
34.3.14.17pandas.Series.str.ljust  
.................................  
1664  
34.3.14.18pandas.Series.str.lower  
................................  
1664  
34.3.14.19pandas.Series.str.lstrip  
.................................  
1665  
34.3.14.20pandas.Series.str.match  
................................  
1666  
34.3.14.21pandas.Series.str.normalize  
..............................  
1666  
34.3.14.22pandas.Series.str.pad  
..................................  
1666  
34.3.14.23pandas.Series.str.partition  
...............................  
1667  
34.3.14.24pandas.Series.str.repeat  
................................  
1667  
34.3.14.25pandas.Series.str.replace  
................................  
1668  
34.3.14.26pandas.Series.str  
.................................  
1670  
34.3.14.27pandas.Series.str.rindex  
................................  
1670  
34.3.14.28pandas.Series.str.rjust  
.................................  
1670  
34.3.14.29pandas.Series.str.rpartition  
...............................  
1671  
xxx

34.3.14.30pandas.Series.str.rstrip  
.................................  
1672  
34.3.14.31pandas.Series.str.slice  
.................................  
1672  
34.3.14.32pandas.Series.str.slice\_replace  
.............................  
1672  
34.3.14.33pandas.Series.str.split  
.................................  
1673  
34.3.14.34pandas.Series.str.rsplit  
.................................  
1675  
34.3.14.35pandas.Series.str.startswith  
..............................  
1675  
34.3.14.36pandas.Series.str.strip  
.................................  
1676  
34.3.14.37pandas.Series.str.swapcase  
..............................  
1677  
34.3.14.38pandas.Series.str.title  
.................................  
1678  
34.3.14.39pandas.Series.str.translate  
...............................  
1679  
34.3.14.40pandas.Series.str.upper  
................................  
1679  
34.3.14.41pandas.Series.str.wrap  
.................................  
1681  
34.3.14.42pandas.Series.str  
.................................  
1682  
34.3.14.43pandas.Series.str.isalnum  
...............................  
1682  
34.3.14.44pandas.Series.str.isalpha  
................................  
1682  
34.3.14.45pandas.Series.str.isdigit  
................................  
1682  
34.3.14.46pandas.Series.str.isspace  
................................  
1682  
34.3.14.47pandas.Series.str.islower  
................................  
1683  
34.3.14.48pandas.Series.str.isupper  
................................  
1683  
34.3.14.49pandas.Series.str.istitle  
.................................  
1683  
34.3.14.50pandas.Series.str.isnumeric  
..............................  
1683  
34.3.14.51pandas.Series.str.isdecimal  
..............................  
1683  
34.3.14.52pandas.Series.str.get\_dummies  
............................  
1683  
34.3.15Categorical  
...........................................  
1684  
34.3.15.1pandas.api.types.CategoricalDtype  
..........................  
1684  
34.3.15.2pandas.Categorical  
...................................  
1685  
34.3.15.3pandas.Series.cat.categories  
..............................  
1689  
34.3.15.4pandas.Series.cat.ordered  
...............................  
1689  
34.3.15.5pandas.Series.cat.codes  
................................  
1689  
34.3.15.6pandas.Series.cat.rename\_categories  
.........................  
1690  
34.3.15.7pandas.Series.cat.reorder\_categories  
.........................  
1691  
34.3.15.8pandas.Series.cat.add\_categories  
...........................  
1691  
34.3.15.9pandas.Series.cat.remove\_categories  
.........................  
1692  
34.3.15.10pandas.Series.cat.remove\_unused\_categories  
.....................  
1692  
34.3.15.11pandas.Series.cat.set\_categories  
............................  
1693  
34.3.15.12pandas.Series.cat.as\_ordered  
.............................  
1693  
34.3.15.13pandas.Series.cat.as\_unordered  
............................  
1694  
34.3.16Plotting  
.............................................  
1694  
34.3.16.1pandas.Series.plot.area  
................................  
1694  
34.3.16.2pandas.Series.plot.bar  
.................................  
1694  
34.3.16.3pandas.Series.plot.barh  
................................  
1695  
34.3.16.4pandas.Series.plot.box  
.................................  
1695  
34.3.16.5pandas.Series.plot.density  
...............................  
1695  
34.3.16.6pandas.Series.plot.hist  
.................................  
1696  
34.3.16.7pandas.Series.plot.kde  
.................................  
1696  
34.3.16.8pandas.Series.plot.line  
.................................  
1697  
34.3.16.9pandas.Series.plot.pie  
.................................  
1697  
34.3.17Serialization/IO/Conversion  
.................................  
1698  
34.3.18Sparse  
..............................................  
1698  
34.3.18.1pandas.SparseSeries.to\_coo  
..............................  
1698  
34.3.18.2pandas.SparseSeries.from\_coo  
............................  
1699  
34.4DataFrame  
................................................  
1700  
34.4.1Constructor  
...........................................  
1700  
34.4.1.1pandas.DataFrame  
...................................  
1700  
xxxi

34.4.2Attributesandunderlyingdata  
.................................  
1932  
34.4.2.1pandas.DataFrame.is\_copy  
..............................  
1933  
34.4.3Conversion  
...........................................  
1933  
34.4.4Indexing,iteration  
........................................  
1933  
34.4.4.1pandas.DataFrame.\_\_iter\_\_  
..............................  
1934  
34.4.5Binaryoperatorfunctions  
....................................  
1934  
34.4.6Functionapplication,GroupBy&Window  
...........................  
1935  
34.4.7Computations/DescriptiveStats  
................................  
1935  
34.4.8Reindexing/Selection/Labelmanipulation  
..........................  
1937  
34.4.9Missingdatahandling  
......................................  
1938  
34.4.10Reshaping,sorting,transposing  
.................................  
1938  
34.4.11Combining/joining/merging  
.................................  
1939  
34.4.12Timeseries-related  
.......................................  
1939  
34.4.13Plotting  
.............................................  
1939  
34.4.13.1pandas.DataFrame.plot.area  
..............................  
1940  
34.4.13.2pandas.DataFrame.plot.bar  
..............................  
1940  
34.4.13.3pandas.DataFrame.plot.barh  
..............................  
1941  
34.4.13.4pandas.DataFrame.plot.box  
..............................  
1943  
34.4.13.5pandas.DataFrame.plot.density  
............................  
1943  
34.4.13.6pandas.DataFrame.plot.hexbin  
............................  
1944  
34.4.13.7pandas.DataFrame.plot.hist  
..............................  
1946  
34.4.13.8pandas.DataFrame.plot.kde  
..............................  
1946  
34.4.13.9pandas.DataFrame.plot.line  
..............................  
1947  
34.4.13.10pandas.DataFrame.plot.pie  
..............................  
1948  
34.4.13.11pandas.DataFrame.plot.scatter  
.............................  
1949  
34.4.14Serialization/IO/Conversion  
.................................  
1950  
34.4.15Sparse  
..............................................  
1951  
34.4.15.1pandas.SparseDataFrame.to\_coo  
...........................  
1951  
34.5Panel  
...................................................  
1952  
34.5.1Constructor  
...........................................  
1952  
34.5.1.1pandas.Panel  
......................................  
1952  
34.5.2Attributesandunderlyingdata  
.................................  
2085  
34.5.3Conversion  
...........................................  
2085  
34.5.4Gettingandsetting  
.......................................  
2085  
34.5.5Indexing,iteration,slicing  
...................................  
2086  
34.5.5.1pandas.Panel.\_\_iter\_\_  
.................................  
2086  
34.5.6Binaryoperatorfunctions  
....................................  
2086  
34.5.7Functionapplication,GroupBy  
.................................  
2087  
34.5.8Computations/DescriptiveStats  
................................  
2087  
34.5.9Reindexing/Selection/Labelmanipulation  
..........................  
2088  
34.5.9.1pandas.Panel.drop  
...................................  
2088  
34.5.10Missingdatahandling  
......................................  
2088  
34.5.11Reshaping,sorting,transposing  
.................................  
2089  
34.5.12Combining/joining/merging  
.................................  
2089  
34.5.13Timeseries-related  
.......................................  
2089  
34.5.14Serialization/IO/Conversion  
.................................  
2089  
34.6Index  
...................................................  
2089  
34.6.1pandas.Index  
..........................................  
2090  
34.6.1.1pandas.Index.T  
....................................  
2091  
34.6.1.2pandas.Index.base  
...................................  
2091  
34.6.1.3pandas.Index.data  
...................................  
2091  
34.6.1.4pandas.Index.dtype  
..................................  
2091  
34.6.1.5pandas.Index.dtype\_str  
................................  
2091  
34.6.1.6pandas.Inde  
...................................  
2092  
xxxii

34.6.1.7pandas.Index.hasnans  
.................................  
2092  
34.6.1.8pandas.Index.inferred\_type  
..............................  
2092  
34.6.1.9pandas.Index.is\_monotonic  
..............................  
2092  
34.6.1.10pandas.Index.is\_monotonic\_decreasing  
........................  
2092  
34.6.1.11pandas.Index.is\_monotonic\_increasing  
........................  
2092  
34.6.1.12pandas.Index.is\_unique  
................................  
2093  
34.6.1.13pandas.Index.itemsize  
.................................  
2093  
34.6.1.14pandas.Index.nbytes  
..................................  
2093  
34.6.1.15pandas.Index.ndim  
...................................  
2093  
34.6.1.16pandas.Index.shape  
..................................  
2093  
34.6.1.17pandas.Index.size  
...................................  
2093  
34.6.1.18pandas.Index.strides  
..................................  
2093  
34.6.1.19pandas.Index.values  
..................................  
2093  
34.6.1.20pandas.Index.all  
....................................  
2096  
34.6.1.21pandas.Index.any  
...................................  
2097  
34.6.1.22pandas.Index.append  
.................................  
2097  
34.6.1.23pandas.Index.argmax  
.................................  
2098  
34.6.1.24pandas.Index.argmin  
..................................  
2098  
34.6.1.25pandas.Index.argsort  
..................................  
2098  
34.6.1.26pandas.Index.asof  
...................................  
2099  
34.6.1.27pandas.Index.asof\_locs  
................................  
2099  
34.6.1.28pandas.Index.astype  
..................................  
2099  
34.6.1.29pandas.Index.contains  
.................................  
2099  
34.6.1.30pandas.Index.copy  
...................................  
2099  
34.6.1.31pandas.Index.delete  
..................................  
2100  
34.6.1.32pandas.Index.difference  
................................  
2100  
34.6.1.33pandas.Index.drop  
...................................  
2100  
34.6.1.34pandas.Index.drop\_duplicates  
.............................  
2101  
34.6.1.35pandas.Index.dropna  
..................................  
2101  
34.6.1.36pandas.Index.duplicated  
................................  
2102  
34.6.1.37pandas.Index.equals  
..................................  
2103  
34.6.1.38pandas.Index.factorize  
.................................  
2103  
34.6.1.39pandas.Inde  
...................................  
2104  
34.6.1.40pandas.Index.format  
..................................  
2105  
34.6.1.41pandas.Index.get\_duplicates  
..............................  
2105  
34.6.1.42pandas.Index.get\_indexer  
...............................  
2106  
34.6.1.43pandas.Index.get\_indexer\_for  
.............................  
2106  
34.6.1.44pandas.Index.get\_indexer\_non\_unique  
........................  
2107  
34.6.1.45pandas.Index.get\_level\_values  
............................  
2107  
34.6.1.46pandas.Index.get\_loc  
.................................  
2107  
34.6.1.47pandas.Index.get\_slice\_bound  
.............................  
2108  
34.6.1.48pandas.Index.get\_value  
................................  
2108  
34.6.1.49pandas.Index.get\_values  
................................  
2108  
34.6.1.50pandas.Index.groupby  
.................................  
2109  
34.6.1.51pandas.Index.identical  
.................................  
2109  
34.6.1.52pandas.Index.insert  
..................................  
2110  
34.6.1.53pandas.Index.intersection  
...............................  
2110  
34.6.1.54pandas.Index.is\_  
....................................  
2110  
34.6.1.55pandas.Index.is\_categorical  
..............................  
2110  
34.6.1.56pandas.Index.isin  
...................................  
2111  
34.6.1.57pandas.Index.isna  
...................................  
2112  
34.6.1.58pandas.Index.isnull  
..................................  
2113  
34.6.1.59pandas.Index.item  
...................................  
2114  
34.6.1.60pandas.Index.join  
...................................  
2114  
xxxiii

34.6.1.61pandas.Index.map  
...................................  
2115  
34.6.1.62pandas.Index.max  
...................................  
2115  
34.6.1.63pandas.Index.memory\_usage  
.............................  
2116  
34.6.1.64pandas.Index.min  
...................................  
2116  
34.6.1.65pandas.Index.notna  
..................................  
2117  
34.6.1.66pandas.Index.notnull  
..................................  
2117  
34.6.1.67pandas.Index.nunique  
.................................  
2118  
34.6.1.68pandas.Index.putmask  
.................................  
2118  
34.6.1.69pandas.Index.ravel  
...................................  
2119  
34.6.1.70pandas.Index.reindex  
.................................  
2119  
34.6.1.71pandas.Index.rename  
.................................  
2119  
34.6.1.72pandas.Index.repeat  
..................................  
2119  
34.6.1.73pandas.Index.searchsorted  
...............................  
2120  
34.6.1.74pandas.Index.set\_names  
................................  
2121  
34.6.1.75pandas.Index.set\_value  
................................  
2122  
34.6.1.76pandas.Index.shift  
...................................  
2122  
34.6.1.77pandas.Index.slice\_indexer  
..............................  
2123  
34.6.1.78pandas.Index.slice\_locs  
................................  
2124  
34.6.1.79pandas.Index.sort\_values  
...............................  
2125  
34.6.1.80pandas.Index.sortlevel  
.................................  
2125  
34.6.1.81pandas.Index.str  
....................................  
2126  
34.6.1.82pandas.Index.summary  
................................  
2126  
34.6.1.83pandas.Index.symmetric\_difference  
..........................  
2126  
34.6.1.84pandas.Index.take  
...................................  
2127  
34.6.1.85pandas.Index.to\_frame  
.................................  
2127  
34.6.1.86pandas.Index.to\_native\_types  
.............................  
2128  
34.6.1.87pandas.Index.to\_series  
.................................  
2128  
34.6.1.88pandas.Index.tolist  
...................................  
2128  
34.6.1.89pandas.Index.transpose  
................................  
2129  
34.6.1.90pandas.Index.union  
..................................  
2129  
34.6.1.91pandas.Index.unique  
..................................  
2129  
34.6.1.92pandas.Index.value\_counts  
..............................  
2129  
34.6.1.93pandas.Index.where  
..................................  
2130  
34.6.2Attributes  
............................................  
2130  
34.6.2.1pandas.Index.has\_duplicates  
.............................  
2131  
34.6.2.2pandas.Index.is\_all\_dates  
...............................  
2131  
34.6.2.3pandas.Index.name  
..................................  
2131  
34.6.2.4pandas.Index.names  
..................................  
2131  
34.6.2.5pandas.Index.empty  
..................................  
2131  
34.6.3ModifyingandComputations  
..................................  
2131  
34.6.3.1pandas.Index.is\_boolean  
................................  
2132  
34.6.3.2pandas.Inde  
................................  
2132  
34.6.3.3pandas.Index.is\_integer  
................................  
2132  
34.6.3.4pandas.Index.is\_interval  
................................  
2133  
34.6.3.5pandas.Index.is\_lexsorted\_for\_tuple  
.........................  
2133  
34.6.3.6pandas.Index.is\_mixed  
................................  
2133  
34.6.3.7pandas.Index.is\_numeric  
...............................  
2133  
34.6.3.8pandas.Index.is\_object  
................................  
2133  
34.6.4MissingValues  
.........................................  
2133  
34.6.5Conversion  
...........................................  
2133  
34.6.5.1pandas.Index.view  
...................................  
2133  
34.6.6Sorting  
.............................................  
2134  
34.6.7Toperations  
....................................  
2134  
34.6.8Combining/joining/setoperations  
..............................  
2134  
xxxiv

34.6.9Selecting  
............................................  
2134  
34.7NumericIndex  
..............................................  
2135  
34.7.1pandas.RangeIndex  
.......................................  
2135  
34.7.1.1pandas.RangeIndex.from\_range  
............................  
2136  
34.7.2pandas.Int64Index  
.......................................  
2136  
34.7.3pandas.UInt64Index  
.......................................  
2137  
34.7.4pandas.Float64Index  
......................................  
2137  
34.8CategoricalIndex  
.............................................  
2138  
34.8.1pandas.CategoricalIndex  
....................................  
2138  
34.8.1.1pandas.CategoricalIndex.rename\_categories  
.....................  
2139  
34.8.1.2pandas.CategoricalIndex.reorder\_categories  
.....................  
2140  
34.8.1.3pandas.CategoricalIndex.add\_categories  
.......................  
2141  
34.8.1.4pandas.CategoricalIndex.remove\_categories  
.....................  
2141  
34.8.1.5pandas.CategoricalIndex.remove\_unused\_categories  
.................  
2142  
34.8.1.6pandas.CategoricalIndex.set\_categories  
........................  
2142  
34.8.1.7pandas.CategoricalIndex.as\_ordered  
.........................  
2143  
34.8.1.8pandas.CategoricalIndex.as\_unordered  
........................  
2143  
34.8.1.9pandas.CategoricalIndex.map  
.............................  
2143  
34.8.2CategoricalComponents  
....................................  
2144  
34.8.2.1pandas.CategoricalIndex.codes  
............................  
2145  
34.8.2.2pandas.CategoricalIndex.categories  
..........................  
2145  
34.8.2.3pandas.CategoricalIndex.ordered  
...........................  
2145  
34.9IntervalIndex  
...............................................  
2145  
34.9.1pandas.IntervalIndex  
......................................  
2145  
34.9.1.1pandas.IntervalIndex.closed  
..............................  
2146  
34.9.1.2pandas.IntervalIndex.is\_non\_overlapping\_monotonic  
................  
2147  
34.9.1.3pandas.IntervalIndex.left  
...............................  
2147  
34.9.1.4pandas.IntervalIndex.length  
..............................  
2147  
34.9.1.5pandas.IntervalIndex.mid  
...............................  
2147  
34.9.1.6pandas.IntervalIndex.right  
...............................  
2147  
34.9.1.7pandas.IntervalIndex.values  
..............................  
2147  
34.9.1.8pandas.IntervalIndex.contains  
.............................  
2147  
34.9.1.9pandas.IntervalIndex.from\_arrays  
...........................  
2148  
34.9.1.10pandas.IntervalIndex.from\_breaks  
..........................  
2149  
34.9.1.11pandas.IntervalIndex.from\_tuples  
...........................  
2150  
34.9.1.12pandas.IntervalIndex.get\_indexer  
...........................  
2151  
34.9.1.13pandas.IntervalIndex.get\_loc  
.............................  
2151  
34.9.2IntervalIndexComponents  
...................................  
2152  
34.10MultiIndex  
................................................  
2153  
34.10.1pandas.MultiIndex  
.......................................  
2153  
34.10.1.1pandas.MultiIndex.names  
...............................  
2154  
34.10.1.2pandas.MultiIndex.nlevels  
...............................  
2154  
34.10.1.3pandas.MultiIndex.levshape  
..............................  
2154  
34.10.1.4pandas.MultiIndex.from\_arrays  
............................  
2155  
34.10.1.5pandas.MultiIndex.from\_tuples  
............................  
2155  
34.10.1.6pandas.MultiIndex.from\_product  
...........................  
2156  
34.10.1.7pandas.MultiIndex.set\_levels  
.............................  
2156  
34.10.1.8pandas.MultiIndex.set\_labels  
.............................  
2157  
34.10.1.9pandas.MultiIndex.to\_hierarchical  
..........................  
2158  
34.10.1.10pandas.MultiIndex.to\_frame  
..............................  
2159  
34.10.1.11pandas.MultiIndex.is\_lexsorted  
............................  
2159  
34.10.1.12pandas.MultiIndex.sortlevel  
..............................  
2159  
34.10.1.13pandas.MultiIndex.droplevel  
.............................  
2159  
34.10.1.14pandas.MultiIndex.swaplevel  
.............................  
2160  
xxxv

34.10.1.15pandas.MultiIndex.reorder\_levels  
...........................  
2160  
34.10.1.16pandas.MultiIndex.remove\_unused\_levels  
......................  
2161  
34.10.2pandas.IndexSlice  
........................................  
2161  
34.10.3MultiIndexConstructors  
....................................  
2162  
34.10.4MultiIndexAttributes  
......................................  
2162  
34.10.4.1pandas.MultiIndex.levels  
...............................  
2162  
34.10.4.2pandas.MultiIndex.labels  
...............................  
2162  
34.10.5MultiIndexComponents  
....................................  
2162  
34.10.5.1pandas.MultiIndex.unique  
...............................  
2163  
34.10.6MultiIndexSelecting  
......................................  
2163  
34.10.6.1pandas.MultiIndex.get\_loc  
..............................  
2163  
34.10.6.2pandas.MultiIndex.get\_indexer  
............................  
2164  
34.10.6.3pandas.MultiIndex.get\_level\_values  
.........................  
2165  
34.11DatetimeIndex  
..............................................  
2165  
34.11.1pandas.DatetimeIndex  
.....................................  
2165  
34.11.1.1pandas.DatetimeIndex.year  
..............................  
2167  
34.11.1.2pandas.DatetimeIndex.month  
.............................  
2168  
34.11.1.3pandas.DatetimeIndex.day  
...............................  
2168  
34.11.1.4pandas.DatetimeIndex.hour  
..............................  
2168  
34.11.1.5pandas.DatetimeIndex.minute  
.............................  
2168  
34.11.1.6pandas.DatetimeIndex.second  
.............................  
2168  
34.11.1.7pandas.DatetimeIndex.microsecond  
..........................  
2168  
34.11.1.8pandas.DatetimeIndex.nanosecond  
..........................  
2168  
34.11.1.9pandas.DatetimeIndex.date  
..............................  
2168  
34.11.1.10pandas.DatetimeIndex.time  
..............................  
2168  
34.11.1.11pandas.DatetimeIndex.dayofyear  
...........................  
2169  
34.11.1.12pandas.DatetimeIndex.weekofyear  
..........................  
2169  
34.11.1.13pandas.DatetimeIndex.week  
..............................  
2169  
34.11.1.14pandas.DatetimeIndex.dayofweek  
...........................  
2169  
34.11.1.15pandas.DatetimeIndex.weekday  
............................  
2169  
34.11.1.16pandas.DatetimeIndex.quarter  
.............................  
2169  
34.11.1.17pandas.DatetimeIndex.freq  
..............................  
2169  
34.11.1.18pandas.DatetimeIndex.freqstr  
.............................  
2169  
34.11.1.19pandas.DatetimeIndex.is\_month\_start  
........................  
2169  
34.11.1.20pandas.DatetimeIndex.is\_month\_end  
.........................  
2170  
34.11.1.21pandas.DatetimeIndex.is\_quarter\_start  
........................  
2170  
34.11.1.22pandas.DatetimeIndex.is\_quarter\_end  
.........................  
2171  
34.11.1.23pandas.DatetimeIndex.is\_year\_start  
..........................  
2172  
34.11.1.24pandas.DatetimeIndex.is\_year\_end  
..........................  
2172  
34.11.1.25pandas.DatetimeIndex.is\_leap\_year  
..........................  
2173  
34.11.1.26pandas.DatetimeIndex.inferred\_freq  
.........................  
2174  
34.11.1.27pandas.DatetimeIndex.normalize  
...........................  
2175  
34.11.1.28pandas.DatetimeIndex.strftime  
............................  
2175  
34.11.1.29pandas.DatetimeIndex.snap  
..............................  
2176  
34.11.1.30pandas.DatetimeIndex.tz\_convert  
...........................  
2176  
34.11.1.31pandas.DatetimeIndex.tz\_localize  
...........................  
2177  
34.11.1.32pandas.DatetimeIndex.round  
.............................  
2178  
34.11.1.33pandas.DatetimeInde  
..............................  
2179  
34.11.1.34pandas.DatetimeIndex.ceil  
...............................  
2180  
34.11.1.35pandas.DatetimeIndex.to\_period  
...........................  
2181  
34.11.1.36pandas.DatetimeIndex.to\_perioddelta  
.........................  
2181  
34.11.1.37pandas.DatetimeIndex.to\_pydatetime  
.........................  
2182  
34.11.1.38pandas.DatetimeIndex.to\_series  
............................  
2182  
34.11.1.39pandas.DatetimeIndex.to\_frame  
............................  
2182  
xxxvi

34.11.1.40pandas.DatetimeIndex.month\_name  
.........................  
2183  
34.11.1.41pandas.DatetimeIndex.day\_name  
...........................  
2183  
34.11.2Time/DateComponents  
.....................................  
2183  
34.11.2.1pandas.DatetimeIndex.tz  
...............................  
2184  
34.11.3Selecting  
............................................  
2184  
34.11.3.1pandas.DatetimeIndex.indexer\_at\_time  
........................  
2184  
34.11.3.2pandas.DatetimeIndex.indexer\_between\_time  
....................  
2185  
34.11.4Toperations  
....................................  
2185  
34.11.5Conversion  
...........................................  
2185  
34.12TimedeltaIndex  
..............................................  
2186  
34.12.1pandas.TimedeltaIndex  
.....................................  
2186  
34.12.1.1pandas.TimedeltaIndex.days  
..............................  
2187  
34.12.1.2pandas.TimedeltaIndex.seconds  
............................  
2187  
34.12.1.3pandas.TimedeltaIndex.microseconds  
.........................  
2187  
34.12.1.4pandas.TimedeltaIndex.nanoseconds  
.........................  
2187  
34.12.1.5pandas.TimedeltaIndex.components  
..........................  
2188  
34.12.1.6pandas.TimedeltaIndex.inferred\_freq  
.........................  
2188  
34.12.1.7pandas.TimedeltaIndex.to\_pytimedelta  
........................  
2188  
34.12.1.8pandas.TimedeltaIndex.to\_series  
...........................  
2188  
34.12.1.9pandas.TimedeltaIndex.round  
.............................  
2189  
34.12.1.10pandas.TimedeltaInde  
.............................  
2189  
34.12.1.11pandas.TimedeltaIndex.ceil  
..............................  
2190  
34.12.1.12pandas.TimedeltaIndex.to\_frame  
...........................  
2191  
34.12.2Components  
...........................................  
2191  
34.12.3Conversion  
...........................................  
2192  
34.13PeriodIndex  
................................................  
2192  
34.13.1pandas.PeriodIndex  
.......................................  
2192  
34.13.1.1pandas.PeriodIndex.day  
................................  
2194  
34.13.1.2pandas.PeriodIndex.dayofweek  
............................  
2194  
34.13.1.3pandas.PeriodIndex.dayofyear  
.............................  
2194  
34.13.1.4pandas.PeriodIndex.days\_in\_month  
..........................  
2194  
34.13.1.5pandas.PeriodIndex.daysinmonth  
...........................  
2194  
34.13.1.6pandas.PeriodIndex.freq  
................................  
2194  
34.13.1.7pandas.PeriodIndex.freqstr  
..............................  
2194  
34.13.1.8pandas.PeriodIndex.hour  
...............................  
2195  
34.13.1.9pandas.PeriodIndex.is\_leap\_year  
...........................  
2195  
34.13.1.10pandas.PeriodIndex.minute  
..............................  
2195  
34.13.1.11pandas.PeriodIndex.month  
..............................  
2195  
34.13.1.12pandas.PeriodIndex.quarter  
..............................  
2195  
34.13.1.13pandas.PeriodIndex.second  
..............................  
2195  
34.13.1.14pandas.PeriodIndex.week  
...............................  
2195  
34.13.1.15pandas.PeriodIndex.weekday  
.............................  
2195  
34.13.1.16pandas.PeriodIndex.weekofyear  
............................  
2195  
34.13.1.17pandas.PeriodIndex.year  
................................  
2196  
34.13.1.18pandas.PeriodIndex.asfreq  
...............................  
2196  
34.13.1.19pandas.PeriodIndex.strftime  
..............................  
2197  
34.13.1.20pandas.PeriodIndex.to\_timestamp  
...........................  
2197  
34.13.1.21pandas.PeriodIndex.tz\_convert  
............................  
2198  
34.13.1.22pandas.PeriodIndex.tz\_localize  
............................  
2198  
34.13.2Attributes  
............................................  
2198  
34.13.2.1pandas.PeriodIndex.end\_time  
.............................  
2199  
34.13.2.2pandas.PeriodIndex.qyear  
...............................  
2199  
34.13.2.3pandas.PeriodIndex.start\_time  
............................  
2199  
34.13.3Methods  
.............................................  
2199  
xxxvii

34.14Scalars  
..................................................  
2199  
34.14.1Period  
..............................................  
2199  
34.14.1.1pandas.Period  
.....................................  
2199  
34.14.2Attributes  
............................................  
2208  
34.14.2.1pandas.Period.end\_time  
................................  
2209  
34.14.2.2pandas.Period.freq  
...................................  
2209  
34.14.2.3pandas.Period.freqstr  
.................................  
2209  
34.14.2.4pandas.Period.is\_leap\_year  
..............................  
2209  
34.14.2.5pandas.Period.month  
.................................  
2209  
34.14.2.6pandas.Period.ordinal  
.................................  
2209  
34.14.2.7pandas.Period.quarter  
.................................  
2209  
34.14.2.8pandas.Period.qyear  
..................................  
2209  
34.14.2.9pandas.Period.weekday  
................................  
2209  
34.14.2.10pandas.Period.weekofyear  
...............................  
2209  
34.14.2.11pandas.Period.year  
...................................  
2210  
34.14.3Methods  
.............................................  
2210  
34.14.3.1pandas.Period.now  
...................................  
2210  
34.14.4Timestamp  
...........................................  
2210  
34.14.4.1pandas.Timestamp  
...................................  
2210  
34.14.5Properties  
............................................  
2220  
34.14.5.1pandas.Timestamp.asm8  
................................  
2220  
34.14.5.2pandas.Timestamp.day  
................................  
2220  
34.14.5.3pandas.Timestamp.dayofweek  
.............................  
2220  
34.14.5.4pandas.Timestamp.dayofyear  
.............................  
2220  
34.14.5.5pandas.Timestamp.days\_in\_month  
..........................  
2221  
34.14.5.6pandas.Timestamp.daysinmonth  
...........................  
2221  
34.14.5.7pandas.Timestamp.fold  
................................  
2221  
34.14.5.8pandas.Timestamp.hour  
................................  
2221  
34.14.5.9pandas.Timestamp.is\_leap\_year  
............................  
2221  
34.14.5.10pandas.Timestamp.is\_month\_end  
...........................  
2221  
34.14.5.11pandas.Timestamp.is\_month\_start  
..........................  
2221  
34.14.5.12pandas.Timestamp.is\_quarter\_end  
..........................  
2221  
34.14.5.13pandas.Timestamp.is\_quarter\_start  
..........................  
2221  
34.14.5.14pandas.Timestamp.is\_year\_end  
............................  
2221  
34.14.5.15pandas.Timestamp.is\_year\_start  
............................  
2221  
34.14.5.16pandas.Timestamp.max  
................................  
2222  
34.14.5.17pandas.Timestamp.microsecond  
............................  
2222  
34.14.5.18pandas.Timestamp.min  
................................  
2222  
34.14.5.19pandas.Timestamp.minute  
...............................  
2222  
34.14.5.20pandas.Timestamp.month  
...............................  
2222  
34.14.5.21pandas.Timestamp.nanosecond  
............................  
2222  
34.14.5.22pandas.Timestamp.quarter  
...............................  
2222  
34.14.5.23pandas.Timestamp.resolution  
.............................  
2222  
34.14.5.24pandas.Timestamp.second  
...............................  
2222  
34.14.5.25pandas.Timestamp.tzinfo  
...............................  
2222  
34.14.5.26pandas.Timestamp.value  
................................  
2222  
34.14.5.27pandas.Timestamp.week  
................................  
2223  
34.14.5.28pandas.Timestamp.weekofyear  
............................  
2223  
34.14.5.29pandas.Timestamp.year  
................................  
2223  
34.14.6Methods  
.............................................  
2223  
34.14.6.1pandas.Timestamp.freq  
................................  
2224  
34.14.6.2pandas.Timestamp.freqstr  
...............................  
2224  
34.14.6.3pandas.Timestamp.isoformat  
.............................  
2224  
34.14.7Interval  
.............................................  
2224  
xxxviii

34.14.7.1pandas.Interval  
.....................................  
2224  
34.14.8Properties  
............................................  
2228  
34.14.9Timedelta  
............................................  
2228  
34.14.9.1pandas.Timedelta  
...................................  
2228  
34.14.10Properties  
............................................  
2232  
34.14.10.1pandas.Timedelta.freq  
.................................  
2232  
34.14.10.2pandas.Timedelta.is\_populated  
............................  
2232  
34.14.10.3pandas.Timedelta.max  
.................................  
2233  
34.14.10.4pandas.Timedelta.min  
.................................  
2233  
34.14.10.5pandas.Timedelta.value  
................................  
2233  
34.14.11Methods  
.............................................  
2233  
34.15Frequencies  
................................................  
2233  
34.15.1pandas.tseries.frequencies.to\_offset  
..............................  
2233  
34.16Window  
..................................................  
2234  
34.16.1Standardmovingwindowfunctions  
..............................  
2234  
34.16.1.1pandas.core.window.Rolling.count  
..........................  
2235  
34.16.1.2pandas.core.window.Rolling.sum  
...........................  
2235  
34.16.1.3pandas.core.window.Rolling.mean  
..........................  
2237  
34.16.1.4pandas.core.window.Rolling.median  
.........................  
2237  
34.16.1.5pandas.core.window.Rolling.var  
............................  
2238  
34.16.1.6pandas.core.window.Rolling.std  
............................  
2239  
34.16.1.7pandas.core.window.Rolling.min  
...........................  
2240  
34.16.1.8pandas.core.window.Rolling.max  
...........................  
2241  
34.16.1.9pandas.core.window.Rolling.corr  
...........................  
2241  
34.16.1.10pandas.core.window.Rolling.cov  
...........................  
2241  
34.16.1.11pandas.core.window.Rolling.skew  
..........................  
2242  
34.16.1.12pandas.core.window.Rolling.kurt  
...........................  
2242  
34.16.1.13pandas.core.window.Rolling.apply  
..........................  
2243  
34.16.1.14pandas.core.window.Rolling.aggregate  
........................  
2243  
34.16.1.15pandas.core.window.Rolling.quantile  
.........................  
2245  
34.16.1.16pandas.core.window.Window.mean  
..........................  
2246  
34.16.1.17pandas.core.window.Window.sum  
...........................  
2247  
34.16.2Standardexpandingwindowfunctions  
.............................  
2248  
34.16.2.1pandas.core.window.Expanding.count  
........................  
2248  
34.16.2.2pandas.core.window.Expanding.sum  
.........................  
2249  
34.16.2.3pandas.core.window.Expanding.mean  
.........................  
2251  
34.16.2.4pandas.core.window.Expanding.median  
.......................  
2251  
34.16.2.5pandas.core.window.Expanding.var  
..........................  
2252  
34.16.2.6pandas.core.window.Expanding.std  
..........................  
2253  
34.16.2.7pandas.core.window.Expanding.min  
.........................  
2254  
34.16.2.8pandas.core.window.Expanding.max  
.........................  
2255  
34.16.2.9pandas.core.window.Expanding.corr  
.........................  
2255  
34.16.2.10pandas.core.window.Expanding.cov  
..........................  
2255  
34.16.2.11pandas.core.window.Expanding.skew  
.........................  
2256  
34.16.2.12pandas.core.window.Expanding.kurt  
.........................  
2256  
34.16.2.13pandas.core.window.Expanding.apply  
........................  
2257  
34.16.2.14pandas.core.window.Expanding.aggregate  
......................  
2257  
34.16.2.15pandas.core.window.Expanding.quantile  
.......................  
2259  
34.16.3Exponentially-weightedmovingwindowfunctions  
......................  
2260  
34.16.3.1pandas.core.window.EWM.mean  
...........................  
2260  
34.16.3.2pandas.core.window.EWM.std  
............................  
2260  
34.16.3.3pandas.core.window.EWM.var  
............................  
2260  
34.16.3.4pandas.core.window.EWM.corr  
............................  
2261  
34.16.3.5pandas.core.window.EWM.cov  
............................  
2261  
xxxix

34.17GroupBy  
.................................................  
2262  
34.17.1Indexing,iteration  
........................................  
2262  
34.17.1.1pandas.core.groupby.GroupBy.\_\_iter\_\_  
........................  
2262  
34.17.1.2pandas.core.groupby.GroupBy.groups  
.........................  
2262  
34.17.1.3pandas.core.groupby.GroupBy.indices  
........................  
2262  
34.17.1.4pandas.core.groupby.GroupBy.get\_group  
.......................  
2262  
34.17.1.5pandas.Grouper  
....................................  
2263  
34.17.2Functionapplication  
......................................  
2264  
34.17.2.1pandas.core.groupby.GroupBy.apply  
.........................  
2264  
34.17.2.2pandas.core.groupby.GroupBy.aggregate  
.......................  
2265  
34.17.2.3pandas.core.groupby.GroupBy.transform  
.......................  
2265  
34.17.2.4pandas.core.groupby.GroupBy.pipe  
..........................  
2265  
34.17.3Computations/DescriptiveStats  
................................  
2267  
34.17.3.1pandas.core.groupby.GroupBy.all  
...........................  
2267  
34.17.3.2pandas.core.groupby.GroupBy.any  
..........................  
2268  
34.17.3.3pandas.core.groupby.GroupBy  
..........................  
2268  
34.17.3.4pandas.core.groupby.GroupBy.count  
.........................  
2268  
34.17.3.5pandas.core.groupby.GroupBy.cumcount  
.......................  
2268  
34.17.3.6pandas.core.groupby.GroupBy.f  
..........................  
2269  
34.17.3.7pandas.core.groupby.GroupBy  
..........................  
2269  
34.17.3.8pandas.core.groupby.GroupBy.head  
..........................  
2269  
34.17.3.9pandas.core.groupby.GroupBy.last  
..........................  
2270  
34.17.3.10pandas.core.groupby.GroupBy.max  
..........................  
2270  
34.17.3.11pandas.core.groupby.GroupBy.mean  
.........................  
2270  
34.17.3.12pandas.core.groupby.GroupBy.median  
........................  
2270  
34.17.3.13pandas.core.groupby.GroupBy.min  
..........................  
2271  
34.17.3.14pandas.core.groupby.GroupBy.ngroup  
........................  
2271  
34.17.3.15pandas.core.groupby.GroupBy.nth  
..........................  
2272  
34.17.3.16pandas.core.groupby.GroupBy.ohlc  
..........................  
2273  
34.17.3.17pandas.core.groupby.GroupBy.prod  
..........................  
2273  
34.17.3.18pandas.core.groupby.GroupBy.rank  
..........................  
2273  
34.17.3.19pandas.core.groupby.GroupBy.pct\_change  
......................  
2274  
34.17.3.20pandas.core.groupby.GroupBy.size  
..........................  
2274  
34.17.3.21pandas.core.groupby.GroupBy.sem  
..........................  
2274  
34.17.3.22pandas.core.groupby.GroupBy.std  
...........................  
2275  
34.17.3.23pandas.core.groupby.GroupBy.sum  
..........................  
2275  
34.17.3.24pandas.core.groupby.GroupBy.var  
..........................  
2275  
34.17.3.25pandas.core.groupby.GroupBy.tail  
..........................  
2275  
34.17.3.26pandas.core.groupby.DataFrameGroupBy.agg  
....................  
2277  
34.17.3.27pandas.core.groupby.DataFrameGroupBy.all  
.....................  
2278  
34.17.3.28pandas.core.groupby.DataFrameGroupBy.any  
....................  
2279  
34.17.3.29pandas.core.groupby.DataFrameGroupBy  
....................  
2279  
34.17.3.30pandas.core.groupby.DataFrameGroupBy.corr  
....................  
2279  
34.17.3.31pandas.core.groupby.DataFrameGroupBy.count  
...................  
2279  
34.17.3.32pandas.core.groupby.DataFrameGroupBy.cov  
....................  
2280  
34.17.3.33pandas.core.groupby.DataFrameGroupBy.cummax  
..................  
2281  
34.17.3.34pandas.core.groupby.DataFrameGroupBy.cummin  
..................  
2281  
34.17.3.35pandas.core.groupby.DataFrameGroupBy.cumprod  
.................  
2281  
34.17.3.36pandas.core.groupby.DataFrameGroupBy.cumsum  
..................  
2281  
34.17.3.37pandas.core.groupby.DataFrameGroupBy.describe  
..................  
2282  
34.17.3.38pandas.core.groupby.DataFrameGroupBy.diff  
....................  
2285  
34.17.3.39pandas.core.groupby.DataFrameGroupBy.f  
....................  
2287  
34.17.3.40pandas.core.groupby.DataFrameGroupBy  
...................  
2287  
34.17.3.41pandas.core.groupby.DataFrameGroupBy  
....................  
2289  
xl

34.17.3.42pandas.core.groupby.DataFrameGroupBy.hist  
....................  
2289  
34.17.3.43pandas.core.groupby.DataFrameGroupBy.idxmax  
..................  
2291  
34.17.3.44pandas.core.groupby.DataFrameGroupBy.idxmin  
..................  
2291  
34.17.3.45pandas.core.groupby.DataFrameGroupBy.mad  
....................  
2292  
34.17.3.46pandas.core.groupby.DataFrameGroupBy.pct\_change  
................  
2292  
34.17.3.47pandas.core.groupby.DataFrameGroupBy.plot  
....................  
2292  
34.17.3.48pandas.core.groupby.DataFrameGroupBy.quantile  
..................  
2292  
34.17.3.49pandas.core.groupby.DataFrameGroupBy.rank  
....................  
2293  
34.17.3.50pandas.core.groupby.DataFrameGroupBy.resample  
.................  
2294  
34.17.3.51pandas.core.groupby.DataFrameGroupBy.shift  
....................  
2294  
34.17.3.52pandas.core.groupby.DataFrameGroupBy.size  
....................  
2295  
34.17.3.53pandas.core.groupby.DataFrameGroupBy.skew  
....................  
2295  
34.17.3.54pandas.core.groupby.DataFrameGroupBy.take  
....................  
2295  
34.17.3.55pandas.core.groupby.DataFrameGroupBy.tshift  
...................  
2297  
34.17.3.56pandas.core.groupby.SeriesGroupBy.nlargest  
.....................  
2297  
34.17.3.57pandas.core.groupby.SeriesGroupBy.nsmallest  
....................  
2298  
34.17.3.58pandas.core.groupby.SeriesGroupBy.nunique  
.....................  
2299  
34.17.3.59pandas.core.groupby.SeriesGroupBy.unique  
.....................  
2299  
34.17.3.60pandas.core.groupby.SeriesGroupBy.value\_counts  
..................  
2300  
34.17.3.61pandas.core.groupby.SeriesGroupBy.is\_monotonic\_increasing  
...........  
2300  
34.17.3.62pandas.core.groupby.SeriesGroupBy.is\_monotonic\_decreasing  
...........  
2300  
34.17.3.63pandas.core.groupby.DataFrameGroupBy.corrwith  
..................  
2300  
34.17.3.64pandas.core.groupby.DataFrameGroupBy.boxplot  
..................  
2301  
34.18Resampling  
................................................  
2302  
34.18.1Indexing,iteration  
........................................  
2302  
34.18.1.1pandas.core.resample.Resampler.\_\_iter\_\_  
......................  
2302  
34.18.1.2pandas.core.resample.Resampler.groups  
.......................  
2302  
34.18.1.3pandas.core.resample.Resampler.indices  
.......................  
2302  
34.18.1.4pandas.core.resample.Resampler.get\_group  
......................  
2302  
34.18.2Functionapplication  
......................................  
2303  
34.18.2.1pandas.core.resample.Resampler.apply  
........................  
2303  
34.18.2.2pandas.core.resample.Resampler.aggregate  
......................  
2304  
34.18.2.3pandas.core.resample.Resampler.transform  
......................  
2306  
34.18.2.4pandas.core.resample.Resampler.pipe  
.........................  
2306  
34.18.3Upsampling  
...........................................  
2307  
34.18.3.1pandas.core.resample.Resampler.f  
.........................  
2308  
34.18.3.2pandas.core.resample.Resampler  
.......................  
2308  
34.18.3.3pandas.core.resample.Resampler  
.........................  
2310  
34.18.3.4pandas.core.resample.Resampler.pad  
.........................  
2311  
34.18.3.5pandas.core.resample.Resampler.nearest  
.......................  
2312  
34.18.3.6pandas.core.resample.Resampler  
........................  
2312  
34.18.3.7pandas.core.resample.Resampler.asfreq  
........................  
2315  
34.18.3.8pandas.core.resample.Resampler.interpolate  
.....................  
2315  
34.18.4Computations/DescriptiveStats  
................................  
2317  
34.18.4.1pandas.core.resample.Resampler.count  
........................  
2317  
34.18.4.2pandas.core.resample.Resampler.nunique  
.......................  
2317  
34.18.4.3pandas.core.resample.Resampler  
.........................  
2317  
34.18.4.4pandas.core.resample.Resampler.last  
.........................  
2318  
34.18.4.5pandas.core.resample.Resampler.max  
.........................  
2318  
34.18.4.6pandas.core.resample.Resampler.mean  
........................  
2318  
34.18.4.7pandas.core.resample.Resampler.median  
.......................  
2318  
34.18.4.8pandas.core.resample.Resampler.min  
.........................  
2318  
34.18.4.9pandas.core.resample.Resampler.ohlc  
.........................  
2318  
34.18.4.10pandas.core.resample.Resampler.prod  
........................  
2319  
xli

34.18.4.11pandas.core.resample.Resampler.size  
.........................  
2319  
34.18.4.12pandas.core.resample.Resampler.sem  
.........................  
2319  
34.18.4.13pandas.core.resample.Resampler.std  
.........................  
2319  
34.18.4.14pandas.core.resample.Resampler.sum  
.........................  
2319  
34.18.4.15pandas.core.resample.Resampler.var  
.........................  
2320  
34.19Style  
...................................................  
2320  
34.19.1StylerConstructor  
........................................  
2320  
34.19.1.1pandas.io.formats.style.Styler  
.............................  
2320  
34.19.2StylerAttributes  
.........................................  
2332  
34.19.2.1pandas.io.formats.style.Styler.env  
...........................  
2332  
34.19.2.2pandas.io.formats.style.Styler.template  
........................  
2332  
34.19.2.3pandas.io.formats.style.Styler.loader  
.........................  
2332  
34.19.3StyleApplication  
........................................  
2332  
34.19.4BuiltinStyles  
..........................................  
2333  
34.19.5StyleExportandImport  
....................................  
2333  
34.20Plotting  
..................................................  
2333  
34.20.1pandas.plotting.andrews\_curves  
................................  
2333  
34.20.2pandas.plotting.bootstrap\_plot  
.................................  
2334  
34.20.3pandas.plotting.deregister\_matplotlib\_converters  
.......................  
2335  
34.20.4pandas.plotting.lag\_plot  
....................................  
2335  
34.20.5pandas.plotting.parallel\_coordinates  
..............................  
2335  
34.20.6pandas.plotting.radviz  
......................................  
2336  
34.20.7pandas.plotting.register\_matplotlib\_converters  
.........................  
2337  
34.20.8pandas.plotting.scatter\_matrix  
.................................  
2338  
34.21Generalutilityfunctions  
.........................................  
2339  
34.21.1Workingwithoptions  
......................................  
2339  
34.21.1.1pandas.describe\_option  
................................  
2339  
34.21.1.2pandas.reset\_option  
..................................  
2343  
34.21.1.3pandas.get\_option  
...................................  
2346  
34.21.1.4pandas.set\_option  
...................................  
2350  
34.21.1.5pandas.option\_context  
.................................  
2354  
34.21.2Testingfunctions  
........................................  
2354  
34.21.2.1pandas.testing.assert\_frame\_equal  
..........................  
2354  
34.21.2.2pandas.testing.assert\_series\_equal  
..........................  
2355  
34.21.2.3pandas.testing.assert\_index\_equal  
...........................  
2356  
34.21.3Exceptionsandwarnings  
....................................  
2356  
34.21.3.1pandas.errors.DtypeWarning  
.............................  
2357  
34.21.3.2pandas.errors.EmptyDataError  
............................  
2358  
34.21.3.3pandas.errors.OutOfBoundsDatetime  
.........................  
2358  
34.21.3.4pandas.errors.ParserError  
...............................  
2358  
34.21.3.5pandas.errors.ParserWarning  
.............................  
2358  
34.21.3.6pandas.errors.PerformanceWarning  
..........................  
2359  
34.21.3.7pandas.errors.UnsortedIndexError  
...........................  
2359  
34.21.3.8pandas.errors.UnsupportedFunctionCall  
.......................  
2359  
34.21.4Datatypesrelatedfunctionality  
.................................  
2359  
34.21.4.1pandas.api.types.union\_categoricals  
..........................  
2360  
34.21.4.2pandas.api.types.infer\_dtype  
.............................  
2361  
34.21.4.3pandas.api.types.pandas\_dtype  
............................  
2364  
34.21.4.4pandas.api.types.is\_bool\_dtype  
............................  
2365  
34.21.4.5pandas.api.types.is\_categorical\_dtype  
.........................  
2365  
34.21.4.6pandas.api.types.is\_complex\_dtype  
..........................  
2366  
34.21.4.7pandas.api.types.is\_datetime64\_any\_dtype  
......................  
2366  
34.21.4.8pandas.api.types.is\_datetime64\_dtype  
........................  
2367  
34.21.4.9pandas.api.types.is\_datetime64\_ns\_dtype  
.......................  
2367  
xlii

34.21.4.10pandas.api.types.is\_datetime64tz\_dtype  
.......................  
2368  
34.21.4.11pandas.api.types.is\_extension\_type  
..........................  
2369  
34.21.4.12  
............................  
2369  
34.21.4.13pandas.api.types.is\_int64\_dtype  
............................  
2370  
34.21.4.14pandas.api.types.is\_integer\_dtype  
...........................  
2371  
34.21.4.15pandas.api.types.is\_interval\_dtype  
..........................  
2371  
34.21.4.16pandas.api.types.is\_numeric\_dtype  
..........................  
2372  
34.21.4.17pandas.api.types.is\_object\_dtype  
...........................  
2372  
34.21.4.18pandas.api.types.is\_period\_dtype  
...........................  
2373  
34.21.4.19pandas.api.types.is\_signed\_integer\_dtype  
.......................  
2373  
34.21.4.20pandas.api.types.is\_string\_dtype  
...........................  
2374  
34.21.4.21pandas.api.types.is\_timedelta64\_dtype  
........................  
2375  
34.21.4.22pandas.api.types.is\_timedelta64\_ns\_dtype  
......................  
2375  
34.21.4.23pandas.api.types.is\_unsigned\_integer\_dtype  
.....................  
2376  
34.21.4.24pandas.api.types.is\_sparse  
...............................  
2376  
34.21.4.25pandas.api.types.is\_dict\_like  
.............................  
2377  
34.21.4.26e  
..............................  
2377  
34.21.4.27pandas.api.types.is\_list\_like  
..............................  
2378  
34.21.4.28pandas.api.types.is\_named\_tuple  
...........................  
2378  
34.21.4.29pandas.api.types.is\_iterator  
..............................  
2379  
34.21.4.30pandas.api.types.is\_bool  
................................  
2380  
34.21.4.31pandas.api.types.is\_categorical  
............................  
2380  
34.21.4.32pandas.api.types.is\_complex  
.............................  
2380  
34.21.4.33pandas.api.types.is\_datetimetz  
............................  
2380  
34.21.4.34  
................................  
2381  
34.21.4.35pandas.api.types.is\_hashable  
.............................  
2381  
34.21.4.36pandas.api.types.is\_integer  
..............................  
2381  
34.21.4.37pandas.api.types.is\_interval  
..............................  
2381  
34.21.4.38pandas.api.types.is\_number  
..............................  
2381  
34.21.4.39pandas.api.types.is\_period  
...............................  
2382  
34.21.4.40pandas.api.types.is\_re  
.................................  
2382  
34.21.4.41pandas.api.types.is\_re\_compilable  
..........................  
2383  
34.21.4.42pandas.api.types.is\_scalar  
...............................  
2383  
34.22Extensions  
................................................  
2383  
34.22.1pandas.api.extensions.register\_dataframe\_accessor  
......................  
2384  
34.22.2pandas.api.extensions.register\_series\_accessor  
.........................  
2385  
34.22.3pandas.api.extensions.register\_index\_accessor  
.........................  
2386  
34.22.4pandas.api.extensions.ExtensionDtype  
.............................  
2387  
34.22.4.1pandas.api.extensions.ExtensionDtype.kind  
.....................  
2388  
34.22.4.2pandas.api.extensions.ExtensionDtype.name  
.....................  
2388  
34.22.4.3pandas.api.extensions.ExtensionDtype.names  
....................  
2388  
34.22.4.4pandas.api.extensions.ExtensionDtype.type  
......................  
2388  
34.22.4.5pandas.api.extensions.ExtensionDtype.construct\_from\_string  
............  
2388  
34.22.4.6pandas.api.extensions.ExtensionDtype.is\_dtype  
...................  
2389  
34.22.5pandas.api.extensions.ExtensionArray  
.............................  
2389  
34.22.5.1pandas.api.extensions.ExtensionArray.dtype  
.....................  
2391  
34.22.5.2pandas.api.extensions.ExtensionArray.nbytes  
.....................  
2391  
34.22.5.3pandas.api.extensions.ExtensionArray.ndim  
.....................  
2391  
34.22.5.4pandas.api.extensions.ExtensionArray.shape  
.....................  
2391  
34.22.5.5pandas.api.extensions.ExtensionArray.argsort  
....................  
2391  
34.22.5.6pandas.api.extensions.ExtensionArray.astype  
.....................  
2392  
34.22.5.7pandas.api.extensions.ExtensionArray.copy  
.....................  
2392  
34.22.5.8pandas.api.extensions.ExtensionArray.factorize  
...................  
2392  
34.22.5.9pandas.api.extensions.ExtensionArray  
.....................  
2393  
xliii

34.22.5.10pandas.api.extensions.ExtensionArray.isna  
......................  
2393  
34.22.5.11pandas.api.extensions.ExtensionArray.take  
......................  
2393  
34.22.5.12pandas.api.extensions.ExtensionArray.unique  
....................  
2395  
34.22.6pandas.Index.asi8  
........................................  
2395  
34.22.7pandas.Index.holds\_integer  
...................................  
2395  
34.22.8pandas.Index.is\_type\_compatible  
................................  
2395  
34.22.9pandas.Index.nlevels  
......................................  
2395  
34.22.10pandas.Index.sort  
........................................  
2395  
34.22.11pandas.Panel.agg  
........................................  
2395  
34.22.12pandas.Panel.aggregate  
.....................................  
2395  
34.22.13pandas.Panel.is\_copy  
......................................  
2395  
34.22.14pandas.Series.imag  
.......................................  
2396  
34.22.15pandas.Series.real  
........................................  
2396  
35Developer  
2397  
35.1StoringpandasDataFrameobjectsinApacheParquetformat  
......................  
2397  
36Internals  
2401  
36.1Indexing  
.................................................  
2401  
36.1.1MultiIndex  
...........................................  
2402  
36.1.2Values  
..............................................  
2402  
36.2SubclassingpandasDataStructures  
...................................  
2403  
37ExtendingPandas  
2405  
37.1RegisteringCustomAccessors  
......................................  
2405  
37.2ExtensionTypes  
.............................................  
2406  
37.2.1  
ExtensionDtype  
......................................  
2406  
37.2.2  
ExtensionArray  
......................................  
2406  
37.3SubclassingpandasDataStructures  
...................................  
2407  
37.3.1OverrideConstructorProperties  
................................  
2407  
37.3.2OriginalProperties  
...................................  
2409  
38ReleaseNotes  
2411  
38.1pandas0.23.0  
...............................................  
2411  
38.1.1Thanks  
.............................................  
2412  
38.2pandas0.22.0  
...............................................  
2421  
38.3pandas0.21.1  
...............................................  
2421  
38.3.1Thanks  
.............................................  
2421  
38.3.1.1Contributors  
......................................  
2422  
38.4pandas0.21.0  
...............................................  
2423  
38.4.1Thanks  
.............................................  
2423  
38.4.1.1Contributors  
......................................  
2424  
38.5pandas0.20.0/0.20.1  
..........................................  
2429  
38.5.1Thanks  
.............................................  
2430  
38.6pandas0.19.2  
...............................................  
2436  
38.6.1Thanks  
.............................................  
2436  
38.7pandas0.19.1  
...............................................  
2437  
38.7.1Thanks  
.............................................  
2437  
38.8pandas0.19.0  
...............................................  
2438  
38.8.1Thanks  
.............................................  
2439  
38.9pandas0.18.1  
...............................................  
2442  
38.9.1Thanks  
.............................................  
2442  
38.10pandas0.18.0  
...............................................  
2444  
38.10.1Thanks  
.............................................  
2444  
38.11pandas0.17.1  
...............................................  
2447  
xliv

38.11.1Thanks  
.............................................  
2448  
38.12pandas0.17.0  
...............................................  
2449  
38.12.1Thanks  
.............................................  
2450  
38.13pandas0.16.2  
...............................................  
2453  
38.13.1Thanks  
.............................................  
2453  
38.14pandas0.16.1  
...............................................  
2454  
38.14.1Thanks  
.............................................  
2455  
38.15pandas0.16.0  
...............................................  
2456  
38.15.1Thanks  
.............................................  
2457  
38.16pandas0.15.2  
...............................................  
2458  
38.16.1Thanks  
.............................................  
2459  
38.17pandas0.15.1  
...............................................  
2460  
38.17.1Thanks  
.............................................  
2460  
38.18pandas0.15.0  
...............................................  
2461  
38.18.1Thanks  
.............................................  
2461  
38.19pandas0.14.1  
...............................................  
2464  
38.19.1Thanks  
.............................................  
2464  
38.20pandas0.14.0  
...............................................  
2465  
38.20.1Thanks  
.............................................  
2466  
38.21pandas0.13.1  
...............................................  
2469  
38.21.1NewFeatures  
..........................................  
2469  
38.21.2APIChanges  
..........................................  
2469  
38.21.3ExperimentalFeatures  
.....................................  
2469  
38.21.4Improvementstoexistingfeatures  
...............................  
2469  
38.21.5BugFixes  
............................................  
2470  
38.22pandas0.13.0  
...............................................  
2472  
38.22.1NewFeatures  
..........................................  
2472  
38.22.2ExperimentalFeatures  
.....................................  
2472  
38.22.3Improvementstoexistingfeatures  
...............................  
2472  
38.22.4APIChanges  
..........................................  
2475  
38.22.5InternalRefactoring  
.......................................  
2478  
38.22.6BugFixes  
............................................  
2480  
38.23pandas0.12.0  
...............................................  
2485  
38.23.1NewFeatures  
..........................................  
2486  
38.23.2Improvementstoexistingfeatures  
...............................  
2486  
38.23.3APIChanges  
..........................................  
2487  
38.23.4ExperimentalFeatures  
.....................................  
2489  
38.23.5BugFixes  
............................................  
2489  
38.24pandas0.11.0  
...............................................  
2492  
38.24.1NewFeatures  
..........................................  
2492  
38.24.2Improvementstoexistingfeatures  
...............................  
2493  
38.24.3APIChanges  
..........................................  
2495  
38.24.4BugFixes  
............................................  
2496  
38.25pandas0.10.1  
...............................................  
2498  
38.25.1NewFeatures  
..........................................  
2498  
38.25.2APIChanges  
..........................................  
2499  
38.25.3Improvementstoexistingfeatures  
...............................  
2499  
38.25.4BugFixes  
............................................  
2500  
38.26pandas0.10.0  
...............................................  
2501  
38.26.1NewFeatures  
..........................................  
2501  
38.26.2ExperimentalFeatures  
.....................................  
2502  
38.26.3APIChanges  
..........................................  
2502  
38.26.4Improvementstoexistingfeatures  
...............................  
2503  
38.26.5BugFixes  
............................................  
2504  
xlv

38.27pandas0.9.1  
...............................................  
2505  
38.27.1NewFeatures  
..........................................  
2506  
38.27.2APIChanges  
..........................................  
2506  
38.27.3Improvementstoexistingfeatures  
...............................  
2506  
38.27.4BugFixes  
............................................  
2506  
38.28pandas0.9.0  
...............................................  
2508  
38.28.1NewFeatures  
..........................................  
2508  
38.28.2Improvementstoexistingfeatures  
...............................  
2508  
38.28.3APIChanges  
..........................................  
2509  
38.28.4BugFixes  
............................................  
2509  
38.29pandas0.8.1  
...............................................  
2513  
38.29.1NewFeatures  
..........................................  
2513  
38.29.2Improvementstoexistingfeatures  
...............................  
2513  
38.29.3BugFixes  
............................................  
2514  
38.30pandas0.8.0  
...............................................  
2515  
38.30.1NewFeatures  
..........................................  
2515  
38.30.2Improvementstoexistingfeatures  
...............................  
2516  
38.30.3APIChanges  
..........................................  
2517  
38.30.4BugFixes  
............................................  
2518  
38.31pandas0.7.3  
...............................................  
2519  
38.31.1NewFeatures  
..........................................  
2519  
38.31.2APIChanges  
..........................................  
2520  
38.31.3BugFixes  
............................................  
2520  
38.32pandas0.7.2  
...............................................  
2521  
38.32.1NewFeatures  
..........................................  
2521  
38.32.2APIChanges  
..........................................  
2521  
38.32.3Improvementstoexistingfeatures  
...............................  
2521  
38.32.4BugFixes  
............................................  
2521  
38.33pandas0.7.1  
...............................................  
2522  
38.33.1NewFeatures  
..........................................  
2522  
38.33.2Improvementstoexistingfeatures  
...............................  
2522  
38.33.3BugFixes  
............................................  
2523  
38.34pandas0.7.0  
...............................................  
2523  
38.34.1NewFeatures  
..........................................  
2523  
38.34.2APIChanges  
..........................................  
2525  
38.34.3Improvementstoexistingfeatures  
...............................  
2525  
38.34.4BugFixes  
............................................  
2526  
38.34.5Thanks  
.............................................  
2529  
38.35pandas0.6.1  
...............................................  
2530  
38.35.1APIChanges  
..........................................  
2530  
38.35.2NewFeatures  
..........................................  
2530  
38.35.3Improvementstoexistingfeatures  
...............................  
2530  
38.35.4BugFixes  
............................................  
2531  
38.35.5Thanks  
.............................................  
2532  
38.36pandas0.6.0  
...............................................  
2532  
38.36.1APIChanges  
..........................................  
2532  
38.36.2NewFeatures  
..........................................  
2532  
38.36.3Improvementstoexistingfeatures  
...............................  
2533  
38.36.4BugFixes  
............................................  
2534  
38.36.5Thanks  
.............................................  
2535  
38.37pandas0.5.0  
...............................................  
2536  
38.37.1APIChanges  
..........................................  
2536  
38.37.2DeprecationsRemoved  
.....................................  
2537  
38.37.3NewFeatures  
..........................................  
2537  
xlvi

38.37.4Improvementstoexistingfeatures  
...............................  
2538  
38.37.5BugFixes  
............................................  
2539  
38.37.6Thanks  
.............................................  
2540  
38.38pandas0.4.3  
...............................................  
2540  
38.38.1NewFeatures  
..........................................  
2540  
38.38.2Improvementstoexistingfeatures  
...............................  
2540  
38.38.3APIChanges  
..........................................  
2540  
38.38.4BugFixes  
............................................  
2540  
38.38.5Thanks  
.............................................  
2541  
38.39pandas0.4.2  
...............................................  
2541  
38.39.1NewFeatures  
..........................................  
2541  
38.39.2Improvementstoexistingfeatures  
...............................  
2541  
38.39.3APIChanges  
..........................................  
2542  
38.39.4BugFixes  
............................................  
2542  
38.39.5Thanks  
.............................................  
2542  
38.40pandas0.4.1  
...............................................  
2542  
38.40.1NewFeatures  
..........................................  
2542  
38.40.2Improvementstoexistingfeatures  
...............................  
2543  
38.40.3APIChanges  
..........................................  
2543  
38.40.4BugFixes  
............................................  
2543  
38.40.5Thanks  
.............................................  
2543  
38.41pandas0.4.0  
...............................................  
2544  
38.41.1NewFeatures  
..........................................  
2544  
38.41.2Improvementstoexistingfeatures  
...............................  
2545  
38.41.3APIChanges  
..........................................  
2546  
38.41.4BugFixes  
............................................  
2547  
38.41.5Thanks  
.............................................  
2548  
38.42pandas0.3.0  
...............................................  
2548  
38.42.1Newfeatures  
..........................................  
2549  
38.42.2Improvementstoexistingfeatures  
...............................  
2549  
38.42.3APIChanges  
..........................................  
2549  
38.42.4BugFixes  
............................................  
2549  
Bibliography  
2551  
PythonModuleIndex  
2553  
xlvii

xlviii

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
PDFVersion  
ZippedHTML  
Date  
:May15,2018  
Version  
:0.23.0  
BinaryInstallers:  
https://pypi.org/project/pandas  
SourceRepository:  
http://github.com/pandas-dev/pandas  
Issues&Ideas:  
https://github.com/pandas-dev/pandas/issues  
Q&ASupport:  
http://stackovw.com/questions/tagged/pandas  
DeveloperMailingList:  
http://groups.google.com/group/pydata  
pandas  
isa  
Python  
packageprovidingfast,xible,andexpressivedatastructuresdesignedtomakeworkingwith  
ﬁrelationalﬂorﬁlabeledﬂdatabotheasyandintuitive.Itaimstobethefundamentalhigh-levelbuildingblockfordoing  
practical,  
realworld  
dataanalysisinPython.Additionally,ithasthebroadergoalofbecoming  
themostpowerful  
andopensourcedataanalysis/manipulationtoolavailableinanylanguage  
.Itisalreadywellonitsway  
towardthisgoal.  
pandasiswellsuitedformanydifferentkindsofdata:  
Ł  
Tabulardatawithheterogeneously-typedcolumns,asinanSQLtableorExcelspreadsheet  
Ł  
Orderedandunordered(notnecessarilyed-frequency)timeseriesdata.  
Ł  
Arbitrarymatrixdata(homogeneouslytypedorheterogeneous)withrowandcolumnlabels  
Ł  
Anyotherformofobservational/statisticaldatasets.Thedataactuallyneednotbelabeledatalltobeplaced  
intoapandasdatastructure  
Thetwoprimarydatastructuresofpandas,  
Series  
(1-dimensional)and  
DataFrame  
(2-dimensional),handlethe  
vastmajorityoftypicalusecasesinstatistics,socialscience,andmanyareasofengineering.ForRusers,  
DataFrame  
provideseverythingthatR's  
data.frame  
providesandmuchmore.pandasisbuiltontopof  
NumPy  
andisintendedtointegratewellwithinacomputingenvironmentwithmanyother3rdpartylibraries.  
Herearejustafewofthethingsthatpandasdoeswell:  
Ł  
Easyhandlingof  
missingdata  
(representedasNaN)inpointaswellaspointdata  
Ł  
Sizemutability:columnscanbe  
insertedanddeleted  
fromDataFrameandhigherdimensionalobjects  
Ł  
Automaticandexplicit  
dataalignment  
:objectscanbeexplicitlyalignedtoasetoflabels,ortheusercan  
simplyignorethelabelsandlet  
Series  
,  
DataFrame  
,etc.automaticallyalignthedataforyouincomputations  
Ł  
Powerful,xible  
groupby  
functionalitytoperformsplit-apply-combineoperationsondatasets,forbothag-  
gregatingandtransformingdata  
Ł  
Makeit  
easytoconvert  
ragged,differently-indexeddatainotherPythonandNumPydatastructuresinto  
DataFrameobjects  
Ł  
Intelligentlabel-based  
slicing  
,  
fancyindexing  
,and  
subsetting  
oflargedatasets  
Ł  
Intuitive  
merging  
and  
joining  
datasets  
Ł  
Flexible  
reshaping  
andpivotingofdatasets  
Ł  
Hierarchical  
labelingofaxes(possibletohavemultiplelabelspertick)  
Ł  
RobustIOtoolsforloadingdatafrom  
  
(CSVanddelimited),Exceldatabases,andsaving/loading  
datafromtheultrafast  
HDF5format  
Ł  
Timeseries  
functionality:daterangegenerationandfrequencyconversion,movingwindowstatistics,  
movingwindowlinearregressions,dateshiftingandlagging,etc.  
CONTENTS  
1

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Manyoftheseprinciplesareheretoaddresstheshortcomingsfrequentlyexperiencedusingotherlanguages/  
researchenvironments.Fordatascientists,workingwithdataistypicallydividedintomultiplestages:mungingand  
cleaningdata,analyzing/modelingit,thenorganizingtheresultsoftheanalysisintoaformsuitableforplottingor  
tabulardisplay.pandasistheidealtoolforallofthesetasks.  
Someothernotes  
Ł  
pandasis  
fast  
.Manyofthelow-levelalgorithmicbitshavebeenextensivelytweakedin  
Cython  
code.However,  
aswithanythingelsegeneralizationusuallyperformance.Soifyoufocusononefeatureforyour  
applicationyoumaybeabletocreateafasterspecializedtool.  
Ł  
pandasisadependencyof  
statsmodels  
,makingitanimportantpartofthestatisticalcomputingecosystemin  
Python.  
Ł  
pandashasbeenusedextensivelyinproductioninapplications.  
Note:  
ThisdocumentationassumesgeneralfamiliaritywithNumPy.Ifyouhaven'tusedNumPymuchoratall,do  
investsometimein  
learningaboutNumPy  
  
Seethepackageoverviewformoredetailaboutwhat'sinthelibrary.  
2  
CONTENTS

CHAPTER  
ONE  
WHAT'SNEW  
Thesearenewfeaturesandimprovementsofnoteineachrelease.  
1.1  
v0.23.0(May15,2017)  
Thisisamajorreleasefrom0.22.0andincludesanumberofAPIchanges,deprecations,newfeatures,enhancements,  
andperformanceimprovementsalongwithalargenumberofbuges.Werecommendthatallusersupgradetothis  
version.  
Highlightsinclude:  
Ł  
Round-trippableJSONformatwith`table'orient  
.  
Ł  
InstantiationfromdictsrespectsorderforPython3.6+  
.  
Ł  
Dependentcolumnargumentsforassign  
.  
Ł  
Merging/sortingonacombinationofcolumnsandindexlevels  
.  
Ł  
ExtendingPandaswithcustomtypes  
.  
Ł  
Excludingunobservedcategoriesfromgroupby  
.  
Ł  
ChangestomakeoutputshapeofDataFrame.applyconsistent  
.  
Checkthe  
APIChanges  
and  
deprecations  
beforeupdating.  
Warning:  
StartingJanuary1,2019,pandasfeaturereleaseswillsupportPython3only.See  
Planfordropping  
Python2.7  
formore.  
What'snewinv0.23.0  
Ł  
Newfeatures  
Œ  
JSONread/writeround-trippablewith  
orient=˜table˜  
Œ  
.assign()  
acceptsdependentarguments  
Œ  
Mergingonacombinationofcolumnsandindexlevels  
Œ  
Sortingbyacombinationofcolumnsandindexlevels  
Œ  
ExtendingPandaswithCustomTypes(Experimental)  
Œ  
New  
observed  
keywordforexcludingunobservedcategoriesin  
groupby  
3

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Œ  
Rolling/Expanding.apply()accepts  
raw=False  
topassa  
Series  
tothefunction  
Œ  
DataFrame.interpolate  
hasgainedthe  
limit\_area  
kwarg  
Œ  
get\_dummies  
nowsupports  
dtype  
argument  
Œ  
Timedeltamodmethod  
Œ  
.rank()  
handles  
inf  
valueswhen  
NaN  
arepresent  
Œ  
Series.str.cat  
hasgainedthe  
join  
kwarg  
Œ  
DataFrame.astype  
performscolumn-wiseconversionto  
Categorical  
Œ  
OtherEnhancements  
Ł  
BackwardsincompatibleAPIchanges  
Œ  
Dependencieshaveincreasedminimumversions  
Œ  
Instantiationfromdictspreservesdictinsertionorderforpython3.6+  
Œ  
DeprecatePanel  
Œ  
pandas.core.commonremovals  
Œ  
Changestomakeoutputof  
DataFrame.apply  
consistent  
Œ  
Concatenationwillnolongersort  
Œ  
BuildChanges  
Œ  
IndexDivisionByZeroFillsCorrectly  
Œ  
Extractionofmatchingpatternsfromstrings  
Œ  
Defaultvalueforthe  
ordered  
parameterof  
CategoricalDtype  
Œ  
Betterpretty-printingofDataFramesinaterminal  
Œ  
DatetimelikeAPIChanges  
Œ  
OtherAPIChanges  
Ł  
Deprecations  
Ł  
Removalofpriorversiondeprecations/changes  
Ł  
PerformanceImprovements  
Ł  
DocumentationChanges  
Ł  
BugFixes  
Œ  
Categorical  
Œ  
Datetimelike  
Œ  
Timedelta  
Œ  
Timezones  
Œ  
Offsets  
Œ  
Numeric  
Œ  
Strings  
Œ  
Indexing  
4  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Œ  
MultiIndex  
Œ  
I/O  
Œ  
Plotting  
Œ  
Groupby/Resample/Rolling  
Œ  
Sparse  
Œ  
Reshaping  
Œ  
Other  
1.1.1  
Newfeatures  
1.1.1.1  
JSONread/writeround-trippablewith  
orient=˜table˜  
A  
DataFrame  
cannowbewrittentoandsubsequentlyreadbackviaJSONwhilepreservingmetadatathroughusage  
ofthe  
orient=˜table˜  
argument(see  
GH18912  
and  
GH9146  
).Previously,noneoftheavailable  
orient  
values  
guaranteedthepreservationofdtypesandindexnames,amongstothermetadata.  
In[1]:  
df  
=  
pd  
.  
DataFrame({  
˜  
foo  
˜  
:[  
1  
,  
2  
,  
3  
,  
4  
],  
...:  
˜  
bar  
˜  
:[  
˜  
a  
˜  
,  
˜  
b  
˜  
,  
˜  
c  
˜  
,  
˜  
d  
˜  
],  
...:  
˜  
baz  
˜  
:pd  
.  
date\_range(  
˜  
2018-01-01  
˜  
,freq  
=  
˜  
d  
˜  
,periods  
=  
4  
),  
...:  
˜  
qux  
˜  
:pd  
.  
Categorical([  
˜  
a  
˜  
,  
˜  
b  
˜  
,  
˜  
c  
˜  
,  
˜  
c  
˜  
])  
...:  
},index  
=  
pd  
.  
Index(  
range  
(  
4  
),name  
=  
˜  
idx  
˜  
))  
...:  
In[2]:  
df  
Out[2]:  
foobarbazqux  
idx  
01a2018-01-01a  
12b2018-01-02b  
23c2018-01-03c  
34d2018-01-04c  
In[3]:  
df  
.  
dtypes  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[3]:  
,  
!  
fooint64  
barobject  
bazdatetime64[ns]  
quxcategory  
dtype:object  
In[4]:  
df  
.  
to\_json(  
˜  
test.json  
˜  
,orient  
=  
˜  
table  
˜  
)  
In[5]:  
new\_df  
=  
pd  
.  
read\_json(  
˜  
test.json  
˜  
,orient  
=  
˜  
table  
˜  
)  
In[6]:  
new\_df  
Out[6]:  
foobarbazqux  
idx  
01a2018-01-01a  
12b2018-01-02b  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
5

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
23c2018-01-03c  
34d2018-01-04c  
In[7]:  
new\_df  
.  
dtypes  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[7]:  
,  
!  
fooint64  
barobject  
bazdatetime64[ns]  
quxcategory  
dtype:object  
Pleasenotethatthestring  
index  
isnotsupportedwiththeroundtripformat,asitisusedbydefaultin  
write\_json  
toindicateamissingindexname.  
In[8]:  
df  
.  
index  
.  
name  
=  
˜  
index  
˜  
In[9]:  
df  
.  
to\_json(  
˜  
test.json  
˜  
,orient  
=  
˜  
table  
˜  
)  
In[10]:  
new\_df  
=  
pd  
.  
read\_json(  
˜  
test.json  
˜  
,orient  
=  
˜  
table  
˜  
)  
In[11]:  
new\_df  
Out[11]:  
foobarbazqux  
01a2018-01-01a  
12b2018-01-02b  
23c2018-01-03c  
34d2018-01-04c  
In[12]:  
new\_df  
.  
dtypes  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[12]:  
,  
!  
fooint64  
barobject  
bazdatetime64[ns]  
quxcategory  
dtype:object  
1.1.1.2  
.assign()  
acceptsdependentarguments  
The  
DataFrame.assign()  
nowacceptsdependentkeywordargumentsforpythonversionlaterthan3.6(seealso  
PEP468  
).Laterkeywordargumentsmaynowrefertoearlieronesiftheargumentisacallable.Seethe  
documentation  
here  
(  
GH14207  
)  
In[13]:  
df  
=  
pd  
.  
DataFrame({  
˜  
A  
˜  
:[  
1  
,  
2  
,  
3  
]})  
In[14]:  
df  
Out[14]:  
A  
01  
12  
23  
In[15]:  
df  
.  
assign(B  
=  
df  
.  
A,C  
=  
lambda  
x:x[  
˜  
A  
˜  
]  
+  
x[  
˜  
B  
˜  
])  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[15]:  
(continuesonnextpage)  
6  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
ABC  
0112  
1224  
2336  
Warning:  
Thismaysubtlychangethebehaviorofyourcodewhenyou'reusing  
.assign()  
toupdatean  
existingcolumn.Previously,callablesreferringtoothervariablesbeingupdatedwouldgettheﬁoldﬂvalues  
PreviousBehavior:  
In[2]:  
df  
=  
pd  
.  
DataFrame({  
"  
A  
"  
:[  
1  
,  
2  
,  
3  
]})  
In[3]:  
df  
.  
assign(A  
=  
lambda  
df:df  
.  
A  
+  
1  
,C  
=  
lambda  
df:df  
.  
A  
\*  
-  
1  
)  
Out[3]:  
AC  
02-1  
13-2  
24-3  
NewBehavior:  
In[16]:  
df  
.  
assign(A  
=  
df  
.  
A  
+  
1  
,C  
=  
lambda  
df:df  
.  
A  
\*  
-  
1  
)  
Out[16]:  
AC  
02-2  
13-3  
24-4  
1.1.1.3  
Mergingonacombinationofcolumnsandindexlevels  
Stringspassedto  
DataFrame.merge()  
asthe  
on  
,  
left\_on  
,and  
right\_on  
parametersmaynowrefertoeither  
columnnamesorindexlevelnames.Thisenablesmerging  
DataFrame  
instancesonacombinationofindexlevels  
andcolumnswithoutresettingindexes.Seethe  
Mergeoncolumnsandlevels  
documentationsection.(  
GH14355  
)  
In[17]:  
left\_index  
=  
pd  
.  
Index([  
˜  
K0  
˜  
,  
˜  
K0  
˜  
,  
˜  
K1  
˜  
,  
˜  
K2  
˜  
],name  
=  
˜  
key1  
˜  
)  
In[18]:  
left  
=  
pd  
.  
DataFrame({  
˜  
A  
˜  
:[  
˜  
A0  
˜  
,  
˜  
A1  
˜  
,  
˜  
A2  
˜  
,  
˜  
A3  
˜  
],  
....:  
˜  
B  
˜  
:[  
˜  
B0  
˜  
,  
˜  
B1  
˜  
,  
˜  
B2  
˜  
,  
˜  
B3  
˜  
],  
....:  
˜  
key2  
˜  
:[  
˜  
K0  
˜  
,  
˜  
K1  
˜  
,  
˜  
K0  
˜  
,  
˜  
K1  
˜  
]},  
....:  
index  
=  
left\_index)  
....:  
In[19]:  
right\_index  
=  
pd  
.  
Index([  
˜  
K0  
˜  
,  
˜  
K1  
˜  
,  
˜  
K2  
˜  
,  
˜  
K2  
˜  
],name  
=  
˜  
key1  
˜  
)  
In[20]:  
right  
=  
pd  
.  
DataFrame({  
˜  
C  
˜  
:[  
˜  
C0  
˜  
,  
˜  
C1  
˜  
,  
˜  
C2  
˜  
,  
˜  
C3  
˜  
],  
....:  
˜  
D  
˜  
:[  
˜  
D0  
˜  
,  
˜  
D1  
˜  
,  
˜  
D2  
˜  
,  
˜  
D3  
˜  
],  
....:  
˜  
key2  
˜  
:[  
˜  
K0  
˜  
,  
˜  
K0  
˜  
,  
˜  
K0  
˜  
,  
˜  
K1  
˜  
]},  
....:  
index  
=  
right\_index)  
....:  
In[21]:  
left  
.  
merge(right,on  
=  
[  
˜  
key1  
˜  
,  
˜  
key2  
˜  
])  
Out[21]:  
ABkey2CD  
key1  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
7

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
K0A0B0K0C0D0  
K1A2B2K0C1D1  
K2A3B3K1C3D3  
1.1.1.4  
Sortingbyacombinationofcolumnsandindexlevels  
Stringspassedto  
DataFrame.sort\_values()  
asthe  
by  
parametermaynowrefertoeithercolumnnamesor  
indexlevelnames.Thisenablessorting  
DataFrame  
instancesbyacombinationofindexlevelsandcolumnswithout  
resettingindexes.Seethe  
SortingbyIndexesandValues  
documentationsection.(  
GH14353  
)  
#BuildMultiIndex  
In[22]:  
idx  
=  
pd  
.  
MultiIndex  
.  
from\_tuples([(  
˜  
a  
˜  
,  
1  
),(  
˜  
a  
˜  
,  
2  
),(  
˜  
a  
˜  
,  
2  
),  
....:  
(  
˜  
b  
˜  
,  
2  
),(  
˜  
b  
˜  
,  
1  
),(  
˜  
b  
˜  
,  
1  
)])  
....:  
In[23]:  
idx  
.  
names  
=  
[  
˜  
first  
˜  
,  
˜  
second  
˜  
]  
#BuildDataFrame  
In[24]:  
df\_multi  
=  
pd  
.  
DataFrame({  
˜  
A  
˜  
:np  
.  
arange(  
6  
,  
0  
,  
-  
1  
)},  
....:  
index  
=  
idx)  
....:  
In[25]:  
df\_multi  
Out[25]:  
A  
firstsecond  
a16  
25  
24  
b23  
12  
11  
#Sortby˜second˜(index)and˜A˜(column)  
In[26]:  
df\_multi  
.  
sort\_values(by  
=  
[  
˜  
second  
˜  
,  
˜  
A  
˜  
])  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[26]:  
,  
!  
A  
firstsecond  
b11  
12  
a16  
b23  
a24  
25  
1.1.1.5  
ExtendingPandaswithCustomTypes(Experimental)  
Pandasnowsupportsstoringarray-likeobjectsthataren'tnecessarily1-DNumPyarraysascolumnsinaDataFrameor  
valuesinaSeries.Thisallowsthird-partylibrariestoimplementextensionstoNumPy'stypes,similartohowpandas  
implementedcategoricals,datetimeswithtimezones,periods,andintervals.  
Asademonstration,we'lluse  
cyberpandas  
,whichprovidesan  
IPArray  
typeforstoringipaddresses.  
8  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
In[1]:  
from  
cyberpandas  
import  
IPArray  
In[2]:  
values  
=  
IPArray([  
...:  
0  
,  
...:  
3232235777  
,  
...:  
42540766452641154071740215577757643572  
...:  
])  
...:  
...:  
IPArray  
isn'tanormal1-DNumPyarray,butbecauseit'sapandas~pandas.api.extension.ExtensionArray,itcanbe  
storedproperlyinsidepandas'containers.  
In[3]:  
ser  
=  
pd  
.  
Series(values)  
In[4]:  
ser  
Out[4]:  
00.0.0.0  
1192.168.1.1  
22001:db8:85a3::8a2e:370:7334  
dtype:ip  
Noticethatthedtypeis  
ip  
.Themissingvaluesemanticsoftheunderlyingarrayarerespected:  
In[5]:  
ser  
.  
isna()  
Out[5]:  
0True  
1False  
2False  
dtype:bool  
Formore,seethe  
extensiontypes  
documentation.Ifyoubuildanextensionarray,publicizeitonour  
ecosystempage  
.  
1.1.1.6  
New  
observed  
keywordforexcludingunobservedcategoriesin  
groupby  
Groupingbyacategoricalincludestheunobservedcategoriesintheoutput.Whengroupingbymultiplecategorical  
columns,thismeansyougetthecartesianproductofallthecategories,includingcombinationswherethereareno  
observations,whichcanresultinalargenumberofgroups.Wehaveaddedakeyword  
observed  
tocontrolthis  
behavior,itdefaultsto  
observed=False  
forbackward-compatiblity.(  
GH14942  
,  
GH8138  
,  
GH15217  
,  
GH17594  
,  
GH8669  
,  
GH20583  
,  
GH20902  
)  
In[27]:  
cat1  
=  
pd  
.  
Categorical([  
"  
a  
"  
,  
"  
a  
"  
,  
"  
b  
"  
,  
"  
b  
"  
],  
....:  
categories  
=  
[  
"  
a  
"  
,  
"  
b  
"  
,  
"  
z  
"  
],ordered  
=  
True  
)  
....:  
In[28]:  
cat2  
=  
pd  
.  
Categorical([  
"  
c  
"  
,  
"  
d  
"  
,  
"  
c  
"  
,  
"  
d  
"  
],  
....:  
categories  
=  
[  
"  
c  
"  
,  
"  
d  
"  
,  
"  
y  
"  
],ordered  
=  
True  
)  
....:  
In[29]:  
df  
=  
pd  
.  
DataFrame({  
"  
A  
"  
:cat1,  
"  
B  
"  
:cat2,  
"  
values  
"  
:[  
1  
,  
2  
,  
3  
,  
4  
]})  
In[30]:  
df[  
˜  
C  
˜  
]  
=  
[  
˜  
foo  
˜  
,  
˜  
bar  
˜  
]  
\*  
2  
In[31]:  
df  
Out[31]:  
ABvaluesC  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
9

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
0ac1foo  
1ad2bar  
2bc3foo  
3bd4bar  
Toshowallvalues,thepreviousbehavior:  
In[32]:  
df  
.  
groupby([  
˜  
A  
˜  
,  
˜  
B  
˜  
,  
˜  
C  
˜  
],observed  
=  
False  
)  
.  
count()  
Out[32]:  
values  
ABC  
acbarNaN  
foo1.0  
dbar1.0  
fooNaN  
ybarNaN  
fooNaN  
bcbarNaN  
......  
yfooNaN  
zcbarNaN  
fooNaN  
dbarNaN  
fooNaN  
ybarNaN  
fooNaN  
[18rowsx1columns]  
Toshowonlyobservedvalues:  
In[33]:  
df  
.  
groupby([  
˜  
A  
˜  
,  
˜  
B  
˜  
,  
˜  
C  
˜  
],observed  
=  
True  
)  
.  
count()  
Out[33]:  
values  
ABC  
acfoo1  
dbar1  
bcfoo1  
dbar1  
Forpivottingoperations,thisbehavioris  
already  
controlledbythe  
dropna  
keyword:  
In[34]:  
cat1  
=  
pd  
.  
Categorical([  
"  
a  
"  
,  
"  
a  
"  
,  
"  
b  
"  
,  
"  
b  
"  
],  
....:  
categories  
=  
[  
"  
a  
"  
,  
"  
b  
"  
,  
"  
z  
"  
],ordered  
=  
True  
)  
....:  
In[35]:  
cat2  
=  
pd  
.  
Categorical([  
"  
c  
"  
,  
"  
d  
"  
,  
"  
c  
"  
,  
"  
d  
"  
],  
....:  
categories  
=  
[  
"  
c  
"  
,  
"  
d  
"  
,  
"  
y  
"  
],ordered  
=  
True  
)  
....:  
In[36]:  
df  
=  
DataFrame({  
"  
A  
"  
:cat1,  
"  
B  
"  
:cat2,  
"  
values  
"  
:[  
1  
,  
2  
,  
3  
,  
4  
]})  
In[37]:  
df  
Out[37]:  
ABvalues  
0ac1  
1ad2  
(continuesonnextpage)  
10  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
2bc3  
3bd4  
In[38]:  
pd  
.  
pivot\_table(df,values  
=  
˜  
values  
˜  
,index  
=  
[  
˜  
A  
˜  
,  
˜  
B  
˜  
],  
....:  
dropna  
=  
True  
)  
....:  
Out[38]:  
values  
AB  
ac1  
d2  
bc3  
d4  
In[39]:  
pd  
.  
pivot\_table(df,values  
=  
˜  
values  
˜  
,index  
=  
[  
˜  
A  
˜  
,  
˜  
B  
˜  
],  
....:  
dropna  
=  
False  
)  
....:  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[39]:  
,  
!  
values  
AB  
ac1.0  
d2.0  
yNaN  
bc3.0  
d4.0  
yNaN  
zcNaN  
dNaN  
yNaN  
1.1.1.7  
Rolling/Expanding.apply()accepts  
raw=False  
topassa  
Series  
tothefunction  
Series.rolling().apply()  
,  
DataFrame.rolling().apply()  
,  
Series.expanding().  
apply()  
,and  
DataFrame.expanding().apply()  
havegaineda  
raw=None  
parameter.Thisissimilarto  
DataFame.apply()  
.Thisparameter,if  
True  
allowsonetosenda  
np.ndarray  
totheappliedfunction.If  
False  
a  
Series  
willbepassed.Thedefaultis  
None  
,whichpreservesbackwardcompatibility,sothiswilldefault  
to  
True  
,sendingan  
np.ndarray  
.Inafutureversionthedefaultwillbechangedto  
False  
,sendinga  
Series  
.  
(  
GH5071  
,  
GH20584  
)  
In[40]:  
s  
=  
pd  
.  
Series(np  
.  
arange(  
5  
),np  
.  
arange(  
5  
)  
+  
1  
)  
In[41]:  
s  
Out[41]:  
10  
21  
32  
43  
54  
dtype:int64  
Passa  
Series  
:  
1.1.v0.23.0(May15,2017)  
11

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
In[42]:  
s  
.  
rolling(  
2  
,min\_periods  
=  
1  
)  
.  
apply(  
lambda  
x:x  
.  
iloc[  
-  
1  
],raw  
=  
False  
)  
Out[42]:  
10.0  
21.0  
32.0  
43.0  
54.0  
dtype:float64  
Mimictheoriginalbehaviorofpassingandarray:  
In[43]:  
s  
.  
rolling(  
2  
,min\_periods  
=  
1  
)  
.  
apply(  
lambda  
x:x[  
-  
1  
],raw  
=  
True  
)  
Out[43]:  
10.0  
21.0  
32.0  
43.0  
54.0  
dtype:float64  
1.1.1.8  
DataFrame.interpolate  
hasgainedthe  
limit\_area  
kwarg  
DataFrame.interpolate()  
hasgaineda  
limit\_area  
parametertoallowfurthercontrolofwhich  
NaN  
sarereplaced.Use  
limit\_area=˜inside˜  
toonlyNaNssurroundedbyvalidvaluesoruse  
limit\_area=˜outside˜  
toonly  
NaN  
soutsidetheexistingvalidvalueswhilepreservingthoseinside.  
(  
GH16284  
)Seethe  
fulldocumentationhere  
.  
In[44]:  
ser  
=  
pd  
.  
Series([np  
.  
nan,np  
.  
nan,  
5  
,np  
.  
nan,np  
.  
nan,np  
.  
nan,  
13  
,np  
.  
nan,np  
.  
,  
!  
nan])  
In[45]:  
ser  
Out[45]:  
0NaN  
1NaN  
25.0  
3NaN  
4NaN  
5NaN  
613.0  
7NaN  
8NaN  
dtype:float64  
Filloneconsecutiveinsidevalueinbothdirections  
In[46]:  
ser  
.  
interpolate(limit\_direction  
=  
˜  
both  
˜  
,limit\_area  
=  
˜  
inside  
˜  
,limit  
=  
1  
)  
Out[46]:  
0NaN  
1NaN  
25.0  
37.0  
4NaN  
511.0  
613.0  
7NaN  
(continuesonnextpage)  
12  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
8NaN  
dtype:float64  
Fillallconsecutiveoutsidevaluesbackward  
In[47]:  
ser  
.  
interpolate(limit\_direction  
=  
˜  
backward  
˜  
,limit\_area  
=  
˜  
outside  
˜  
)  
Out[47]:  
05.0  
15.0  
25.0  
3NaN  
4NaN  
5NaN  
613.0  
7NaN  
8NaN  
dtype:float64  
Fillallconsecutiveoutsidevaluesinbothdirections  
In[48]:  
ser  
.  
interpolate(limit\_direction  
=  
˜  
both  
˜  
,limit\_area  
=  
˜  
outside  
˜  
)  
Out[48]:  
05.0  
15.0  
25.0  
3NaN  
4NaN  
5NaN  
613.0  
713.0  
813.0  
dtype:float64  
1.1.1.9  
get\_dummies  
nowsupports  
dtype  
argument  
The  
get\_dummies()  
nowacceptsa  
dtype  
argument,whichadtypeforthenewcolumns.Thedefault  
remainsuint8.(  
GH18330  
)  
In[49]:  
df  
=  
pd  
.  
DataFrame({  
˜  
a  
˜  
:[  
1  
,  
2  
],  
˜  
b  
˜  
:[  
3  
,  
4  
],  
˜  
c  
˜  
:[  
5  
,  
6  
]})  
In[50]:  
pd  
.  
get\_dummies(df,columns  
=  
[  
˜  
c  
˜  
])  
.  
dtypes  
Out[50]:  
aint64  
bint64  
c\_5uint8  
c\_6uint8  
dtype:object  
In[51]:  
pd  
.  
get\_dummies(df,columns  
=  
[  
˜  
c  
˜  
],dtype  
=  
bool  
)  
.  
dtypes  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[51]:  
aint64  
bint64  
c\_5bool  
c\_6bool  
dtype:object  
1.1.v0.23.0(May15,2017)  
13

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.1.1.10  
Timedeltamodmethod  
mod  
(%)and  
divmod  
operationsarenowon  
Timedelta  
objectswhenoperatingwitheithertimedelta-like  
orwithnumericarguments.Seethe  
documentationhere  
.(  
GH19365  
)  
In[52]:  
td  
=  
pd  
.  
Timedelta(hours  
=  
37  
)  
In[53]:  
td  
%  
pd  
.  
Timedelta(minutes  
=  
45  
)  
Out[53]:  
Timedelta(˜0days00:15:00˜)  
1.1.1.11  
.rank()  
handles  
inf  
valueswhen  
NaN  
arepresent  
Inpreviousversions,  
.rank()  
wouldassign  
inf  
elements  
NaN  
astheirranks.Nowranksarecalculatedproperly.  
(  
GH6945  
)  
In[54]:  
s  
=  
pd  
.  
Series([  
-  
np  
.  
inf,  
0  
,  
1  
,np  
.  
nan,np  
.  
inf])  
In[55]:  
s  
Out[55]:  
0-inf  
10.000000  
21.000000  
3NaN  
4inf  
dtype:float64  
PreviousBehavior:  
In[11]:  
s  
.  
rank()  
Out[11]:  
01.0  
12.0  
23.0  
3NaN  
4NaN  
dtype:float64  
CurrentBehavior:  
In[56]:  
s  
.  
rank()  
Out[56]:  
01.0  
12.0  
23.0  
3NaN  
44.0  
dtype:float64  
Furthermore,previouslyifyourank  
inf  
or  
-inf  
valuestogetherwith  
NaN  
values,thecalculationwon'tdistinguish  
NaN  
fromwhenusing`top'or`bottom'argument.  
In[57]:  
s  
=  
pd  
.  
Series([np  
.  
nan,np  
.  
nan,  
-  
np  
.  
inf,  
-  
np  
.  
inf])  
In[58]:  
s  
Out[58]:  
0NaN  
(continuesonnextpage)  
14  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
1NaN  
2-inf  
3-inf  
dtype:float64  
PreviousBehavior:  
In[15]:  
s  
.  
rank(na\_option  
=  
˜  
top  
˜  
)  
Out[15]:  
02.5  
12.5  
22.5  
32.5  
dtype:float64  
CurrentBehavior:  
In[59]:  
s  
.  
rank(na\_option  
=  
˜  
top  
˜  
)  
Out[59]:  
01.5  
11.5  
23.5  
33.5  
dtype:float64  
Thesebugsweresquashed:  
Ł  
Bugin  
DataFrame.rank()  
and  
Series.rank()  
when  
method=˜dense˜  
and  
pct=True  
inwhich  
percentilerankswerenotbeingusedwiththenumberofdistinctobservations(  
GH15630  
)  
Ł  
Bugin  
Series.rank()  
and  
DataFrame.rank()  
when  
ascending=˜False˜  
failedtoreturncorrect  
ranksforif  
NaN  
werepresent(  
GH19538  
)  
Ł  
Bugin  
DataFrameGroupBy.rank()  
whererankswereincorrectwhenbothand  
NaN  
werepresent  
(  
GH20561  
)  
1.1.1.12  
Series.str.cat  
hasgainedthe  
join  
kwarg  
Previously,  
Series.str.cat()  
didnotŒincontrasttomostof  
pandas  
Œalign  
Series  
ontheirindexbefore  
concatenation(see  
GH18657  
).Themethodhasnowgainedakeyword  
join  
tocontrolthemannerofalignment,see  
examplesbelowand  
here  
.  
Inv.0.23  
join  
willdefaulttoNone(meaningnoalignment),butthisdefaultwillchangeto  
˜left˜  
inafutureversion  
ofpandas.  
In[60]:  
s  
=  
pd  
.  
Series([  
˜  
a  
˜  
,  
˜  
b  
˜  
,  
˜  
c  
˜  
,  
˜  
d  
˜  
])  
In[61]:  
t  
=  
pd  
.  
Series([  
˜  
b  
˜  
,  
˜  
d  
˜  
,  
˜  
e  
˜  
,  
˜  
c  
˜  
],index  
=  
[  
1  
,  
3  
,  
4  
,  
2  
])  
In[62]:  
s  
.  
str  
.  
cat(t)  
Out[62]:  
0ab  
1bd  
2ce  
3dc  
dtype:object  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
15

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
In[63]:  
s  
.  
str  
.  
cat(t,join  
=  
˜  
left  
˜  
,na\_rep  
=  
˜  
-  
˜  
)  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[63]:  
0a-  
1bb  
2cc  
3dd  
dtype:object  
Furthermore,  
Series.str.cat()  
nowworksfor  
CategoricalIndex  
aswell(previouslyraiseda  
ValueError  
;see  
GH20842  
).  
1.1.1.13  
DataFrame.astype  
performscolumn-wiseconversionto  
Categorical  
DataFrame.astype()  
cannowperformcolumn-wiseconversionto  
Categorical  
bysupplyingthestring  
˜category˜  
ora  
CategoricalDtype  
.Previously,attemptingthiswouldraisea  
NotImplementedError  
.  
Seethe  
ObjectCreation  
sectionofthedocumentationformoredetailsandexamples.(  
GH12860  
,  
GH18099  
)  
Supplyingthestring  
˜category˜  
performscolumn-wiseconversion,withonlylabelsappearinginagivencolumn  
setascategories:  
In[64]:  
df  
=  
pd  
.  
DataFrame({  
˜  
A  
˜  
:  
list  
(  
˜  
abca  
˜  
),  
˜  
B  
˜  
:  
list  
(  
˜  
bccd  
˜  
)})  
In[65]:  
df  
=  
df  
.  
astype(  
˜  
category  
˜  
)  
In[66]:  
df[  
˜  
A  
˜  
]  
.  
dtype  
Out[66]:  
CategoricalDtype(categories=[˜a˜,˜b˜,˜c˜],ordered=False)  
In[67]:  
df[  
˜  
B  
˜  
]  
.  
dtype  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[67]:  
,  
!  
CategoricalDtype(categories=[˜b˜,˜c˜,˜d˜],ordered=False)  
Supplyinga  
CategoricalDtype  
willmakethecategoriesineachcolumnconsistentwiththesupplieddtype:  
In[68]:  
from  
pandas.api.types  
import  
CategoricalDtype  
In[69]:  
df  
=  
pd  
.  
DataFrame({  
˜  
A  
˜  
:  
list  
(  
˜  
abca  
˜  
),  
˜  
B  
˜  
:  
list  
(  
˜  
bccd  
˜  
)})  
In[70]:  
cdt  
=  
CategoricalDtype(categories  
=  
list  
(  
˜  
abcd  
˜  
),ordered  
=  
True  
)  
In[71]:  
df  
=  
df  
.  
astype(cdt)  
In[72]:  
df[  
˜  
A  
˜  
]  
.  
dtype  
Out[72]:  
CategoricalDtype(categories=[˜a˜,˜b˜,˜c˜,˜d˜],ordered=True)  
In[73]:  
df[  
˜  
B  
˜  
]  
.  
dtype  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[73]:  
,  
!  
CategoricalDtype(categories=[˜a˜,˜b˜,˜c˜,˜d˜],ordered=True)  
1.1.1.14  
OtherEnhancements  
Ł  
Unary  
+  
nowpermittedfor  
Series  
and  
DataFrame  
asnumericoperator(  
GH16073  
)  
Ł  
Bettersupportfor  
to\_excel()  
outputwiththe  
xlsxwriter  
engine.(  
GH16149  
)  
16  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
pandas.tseries.frequencies.to\_offset()  
nowacceptsleading`+'signse.g.`+1h'.(  
GH18171  
)  
Ł  
MultiIndex.unique()  
nowsupportsthe  
level=  
argument,togetuniquevaluesfromaindex  
level(  
GH17896  
)  
Ł  
pandas.io.formats.style.Styler  
nowhasmethod  
hide\_index()  
todeterminewhethertheindex  
willberenderedinoutput(  
GH14194  
)  
Ł  
pandas.io.formats.style.Styler  
nowhasmethod  
hide\_columns()  
todeterminewhether  
columnswillbehiddeninoutput(  
GH14194  
)  
Ł  
Improvedwordingof  
ValueError  
raisedin  
to\_datetime()  
when  
unit=  
ispassedwithanon-convertible  
value(  
GH14350  
)  
Ł  
Series.fillna()  
nowacceptsaSeriesoradictasa  
value  
foracategoricaldtype(  
GH17033  
)  
Ł  
pandas.read\_clipboard()  
updatedtouseqtpy,fallingbacktoPyQt5andthenPyQt4,addingcompati-  
bilitywithPython3andmultiplepython-qtbindings(  
GH17722  
)  
Ł  
Improvedwordingof  
ValueError  
raisedin  
read\_csv()  
whenthe  
usecols  
argumentcannotmatchall  
columns.(  
GH17301  
)  
Ł  
DataFrame.corrwith()  
nowsilentlydropsnon-numericcolumnswhenpassedaSeries.Before,anex-  
ceptionwasraised(  
GH18570  
).  
Ł  
IntervalIndex  
nowsupportstimezoneaware  
Interval  
objects(  
GH18537  
,  
GH18538  
)  
Ł  
Series()  
/  
DataFrame()  
tabcompletionalsoreturnsinthelevelofa  
MultiIndex()  
.  
(  
GH16326  
)  
Ł  
read\_excel()  
hasgainedthe  
nrows  
parameter(  
GH16645  
)  
Ł  
DataFrame.append()  
cannowinmorecasespreservethetypeofthecallingdataframe'scolumns(e.g.if  
bothare  
CategoricalIndex  
)(  
GH18359  
)  
Ł  
DataFrame.to\_json()  
and  
Series.to\_json()  
nowacceptan  
index  
argumentwhichallowsthe  
usertoexcludetheindexfromtheJSONoutput(  
GH17394  
)  
Ł  
IntervalIndex.to\_tuples()  
hasgainedthe  
na\_tuple  
parametertocontrolwhetherNAisreturned  
asatupleofNA,orNAitself(  
GH18756  
)  
Ł  
Categorical.rename\_categories  
,  
CategoricalIndex.rename\_categories  
and  
Series.  
cat.rename\_categories  
cannowtakeacallableastheirargument(  
GH18862  
)  
Ł  
Interval  
and  
IntervalIndex  
havegaineda  
length  
attribute(  
GH18789  
)  
Ł  
Resampler  
objectsnowhaveafunctioning  
pipe  
method.Previously,callsto  
pipe  
weredivertedtothe  
mean  
method(  
GH17905  
).  
Ł  
is\_scalar()  
nowreturns  
True  
for  
DateOffset  
objects(  
GH18943  
).  
Ł  
DataFrame.pivot()  
nowacceptsalistforthe  
values=  
kwarg(  
GH17160  
).  
Ł  
Added  
pandas.api.extensions.register\_dataframe\_accessor()  
,  
pandas.  
api.extensions.register\_series\_accessor()  
,and  
pandas.api.extensions.  
register\_index\_accessor()  
,accessorforlibrariesdownstreamofpandastoregistercustom  
accessorslike  
.cat  
onpandasobjects.See  
RegisteringCustomAccessors  
formore(  
GH14781  
).  
Ł  
IntervalIndex.astype  
nowsupportsconversionsbetweensubtypeswhenpassedan  
IntervalDtype  
(  
GH19197  
)  
Ł  
IntervalIndex  
anditsassociatedconstructormethods(  
from\_arrays  
,  
from\_breaks  
,  
from\_tuples  
)havegaineda  
dtype  
parameter(  
GH19262  
)  
1.1.v0.23.0(May15,2017)  
17

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Added  
pandas.core.groupby.SeriesGroupBy.is\_monotonic\_increasing()  
and  
pandas.  
core.groupby.SeriesGroupBy.is\_monotonic\_decreasing()  
(  
GH17015  
)  
Ł  
Forsubclassed  
DataFrames  
,  
DataFrame.apply()  
willnowpreservethe  
Series  
subclass(if  
whenpassingthedatatotheappliedfunction(  
GH19822  
)  
Ł  
DataFrame.from\_dict()  
nowacceptsa  
columns  
argumentthatcanbeusedtospecifythecolumnnames  
when  
orient=˜index˜  
isused(  
GH18529  
)  
Ł  
Addedoption  
display.html.use\_mathjax  
so  
MathJax  
canbedisabledwhenrenderingtablesin  
Jupyter  
notebooks(  
GH19856  
,  
GH19824  
)  
Ł  
DataFrame.replace()  
nowsupportsthe  
method  
parameter,whichcanbeusedtospecifythereplacement  
methodwhen  
to\_replace  
isascalar,listortupleand  
value  
is  
None  
(  
GH19632  
)  
Ł  
Timestamp.month\_name()  
,  
DatetimeIndex.month\_name()  
,and  
Series.dt.  
month\_name()  
arenowavailable(  
GH12805  
)  
Ł  
Timestamp.day\_name()  
and  
DatetimeIndex.day\_name()  
arenowavailabletoreturndaynames  
withalocale(  
GH12806  
)  
Ł  
DataFrame.to\_sql()  
nowperformsamultivalueinsertiftheunderlyingconnectionsupportsitk  
ratherthaninsertingrowbyrow.  
SQLAlchemy  
dialectssupportingmultivalueinsertsinclude:  
mysql  
,  
postgresql  
,  
sqlite  
andanydialectwith  
supports\_multivalues\_insert  
.(  
GH14315  
,  
GH8953  
)  
Ł  
read\_html()  
nowacceptsa  
displayed\_only  
keywordargumenttocontrolswhetherornothiddenele-  
mentsareparsed(  
True  
bydefault)(  
GH20027  
)  
Ł  
read\_html()  
nowreadsall  
<tbody>  
elementsina  
<table>  
,notjustthe(  
GH20690  
)  
Ł  
quantile()  
and  
quantile()  
nowacceptthe  
interpolation  
keyword,  
linear  
bydefault  
(  
GH20497  
)  
Ł  
zipcompressionissupportedvia  
compression=zip  
in  
DataFrame.to\_pickle()  
,  
Series.  
to\_pickle()  
,  
DataFrame.to\_csv()  
,  
Series.to\_csv()  
,  
DataFrame.to\_json()  
,  
Series.  
to\_json()  
.(  
GH17778  
)  
Ł  
WeekOfMonth  
constructornowsupports  
n=0  
(  
GH20517  
).  
Ł  
DataFrame  
and  
Series  
nowsupportmatrixmultiplication(  
@  
)operator(  
GH10259  
)forPython>=3.5  
Ł  
Updated  
DataFrame.to\_gbq()  
and  
pandas.read\_gbq()  
signatureanddocumentationto  
changesfromthePandas-GBQlibraryversion0.4.0.AddsintersphinxmappingtoPandas-GBQlibrary.  
(  
GH20564  
)  
Ł  
AddednewwriterforexportingStatadtainversion117,  
StataWriter117  
.Thisformatsupports  
exportingstringswithlengthsupto2,000,000characters(  
GH16450  
)  
Ł  
to\_hdf()  
and  
read\_hdf()  
nowacceptan  
errors  
keywordargumenttocontrolencodingerrorhandling  
(  
GH20835  
)  
Ł  
cut()  
hasgainedthe  
duplicates=˜raise˜|˜drop˜  
optiontocontrolwhethertoraiseonduplicated  
edges(  
GH20947  
)  
Ł  
date\_range()  
,  
timedelta\_range()  
,and  
interval\_range()  
nowreturnalinearlyspacedindexif  
start  
,  
stop  
,and  
periods  
arebut  
freq  
isnot.(  
GH20808  
,  
GH20983  
,  
GH20976  
)  
1.1.2  
BackwardsincompatibleAPIchanges  
1.1.2.1  
Dependencieshaveincreasedminimumversions  
Wehaveupdatedourminimumsupportedversionsofdependencies(  
GH15184  
).Ifinstalled,wenowrequire:  
18  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Package  
MinimumVersion  
Required  
Issue  
python-dateutil  
2.5.0  
X  
GH15184  
openpyxl  
2.4.0  
GH15184  
beautifulsoup4  
4.2.1  
GH20082  
setuptools  
24.2.0  
GH20698  
1.1.2.2  
Instantiationfromdictspreservesdictinsertionorderforpython3.6+  
UntilPython3.6,dictsinPythonhadnoformallyordering.ForPythonversion3.6andlater,dictsareordered  
byinsertionorder,see  
PEP468  
.Pandaswillusethedict'sinsertionorder,whencreatinga  
Series  
or  
DataFrame  
fromadictandyou'reusingPythonversion3.6orhigher.(  
GH19884  
)  
PreviousBehavior(andcurrentbehaviorifonPython<3.6):  
pd  
.  
Series({  
˜  
Income  
˜  
:  
2000  
,  
˜  
Expenses  
˜  
:  
-  
1500  
,  
˜  
Taxes  
˜  
:  
-  
200  
,  
˜  
Netresult  
˜  
:  
300  
})  
Expenses  
-  
1500  
Income  
2000  
Netresult  
300  
Taxes  
-  
200  
dtype:int64  
NotetheSeriesaboveisorderedalphabeticallybytheindexvalues.  
NewBehavior(forPython>=3.6):  
In[74]:  
pd  
.  
Series({  
˜  
Income  
˜  
:  
2000  
,  
....:  
˜  
Expenses  
˜  
:  
-  
1500  
,  
....:  
˜  
Taxes  
˜  
:  
-  
200  
,  
....:  
˜  
Netresult  
˜  
:  
300  
})  
....:  
Out[74]:  
Income2000  
Expenses-1500  
Taxes-200  
Netresult300  
dtype:int64  
NoticethattheSeriesisnoworderedbyinsertionorder.Thisnewbehaviorisusedforallrelevantpandastypes  
(  
Series  
,  
DataFrame  
,  
SparseSeries  
and  
SparseDataFrame  
).  
IfyouwishtoretaintheoldbehaviorwhileusingPython>=3.6,youcanuse  
.sort\_index()  
:  
In[75]:  
pd  
.  
Series({  
˜  
Income  
˜  
:  
2000  
,  
....:  
˜  
Expenses  
˜  
:  
-  
1500  
,  
....:  
˜  
Taxes  
˜  
:  
-  
200  
,  
....:  
˜  
Netresult  
˜  
:  
300  
})  
.  
sort\_index()  
....:  
Out[75]:  
Expenses-1500  
Income2000  
Netresult300  
Taxes-200  
dtype:int64  
1.1.v0.23.0(May15,2017)  
19

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.1.2.3  
DeprecatePanel  
Panel  
wasdeprecatedinthe0.20.xrelease,showingasa  
DeprecationWarning  
.Using  
Panel  
willnowshowa  
FutureWarning  
.Therecommendedwaytorepresent3-Ddataarewitha  
MultiIndex  
ona  
DataFrame  
viathe  
to\_frame()  
orwiththe  
xarraypackage  
.Pandasprovidesa  
to\_xarray()  
methodtoautomatethisconversion.  
Formoredetailssee  
DeprecatePanel  
documentation.(  
GH13563  
,  
GH18324  
).  
In[76]:  
p  
=  
tm  
.  
makePanel()  
In[77]:  
p  
Out[77]:  
<class˜pandas.core.panel.Panel˜>  
Dimensions:3(items)x3(major\_axis)x4(minor\_axis)  
Itemsaxis:ItemAtoItemC  
Major\_axisaxis:2000-01-0300:00:00to2000-01-0500:00:00  
Minor\_axisaxis:AtoD  
ConverttoaMultiIndexDataFrame  
In[78]:  
p  
.  
to\_frame()  
Out[78]:  
ItemAItemBItemC  
majorminor  
2000-01-03A1.474071-0.964980-1.197071  
B0.7818361.846883-0.858447  
C2.353925-1.7176930.384316  
D-0.7444710.9018050.476720  
2000-01-04A-0.064034-0.845696-1.066969  
B-1.071357-1.3288650.306996  
C0.5837870.8887821.574159  
D0.7585271.1712160.473424  
2000-01-05A-1.282782-1.340896-0.303421  
B0.4411531.682706-0.028665  
C0.2214710.2284401.588931  
D1.7296890.520260-0.242861  
ConverttoanxarrayDataArray  
In[79]:  
p  
.  
to\_xarray()  
Out[79]:  
<xarray.DataArray(items:3,major\_axis:3,minor\_axis:4)>  
array([[[1.474071,0.781836,2.353925,-0.744471],  
[-0.064034,-1.071357,0.583787,0.758527],  
[-1.282782,0.441153,0.221471,1.729689]],  
[[-0.96498,1.846883,-1.717693,0.901805],  
[-0.845696,-1.328865,0.888782,1.171216],  
[-1.340896,1.682706,0.22844,0.52026]],  
[[-1.197071,-0.858447,0.384316,0.47672],  
[-1.066969,0.306996,1.574159,0.473424],  
[-0.303421,-0.028665,1.588931,-0.242861]]])  
Coordinates:  
\*  
items(items)object˜ItemA˜˜ItemB˜˜ItemC˜  
\*  
major\_axis(major\_axis)datetime64[ns]2000-01-032000-01-042000-01-05  
\*  
minor\_axis(minor\_axis)object˜A˜˜B˜˜C˜˜D˜  
20  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.1.2.4  
pandas.core.commonremovals  
Thefollowingerror&warningmessagesareremovedfrom  
pandas.core.common  
(  
GH13634  
,  
GH19769  
):  
Ł  
PerformanceWarning  
Ł  
UnsupportedFunctionCall  
Ł  
UnsortedIndexError  
Ł  
AbstractMethodError  
Theseareavailablefromimportfrom  
pandas.errors  
(since0.19.0).  
1.1.2.5  
Changestomakeoutputof  
DataFrame.apply  
consistent  
DataFrame.apply()  
wasinconsistentwhenapplyinganarbitraryuserthatreturnedalist-like  
with  
axis=1  
.Severalbugsandinconsistenciesareresolved.IftheappliedfunctionreturnsaSeries,thenpandaswill  
returnaDataFrame;otherwiseaSerieswillbereturned,thisincludesthecasewherealist-like(e.g.  
tuple  
or  
list  
isreturned)(  
GH16353  
,  
GH17437  
,  
GH17970  
,  
GH17348  
,  
GH17892  
,  
GH18573  
,  
GH17602  
,  
GH18775  
,  
GH18901  
,  
GH18919  
).  
In[80]:  
df  
=  
pd  
.  
DataFrame(np  
.  
tile(np  
.  
arange(  
3  
),  
6  
)  
.  
reshape(  
6  
,  
-  
1  
)  
+  
1  
,columns  
=  
[  
˜  
A  
˜  
,  
,  
!  
˜  
B  
˜  
,  
˜  
C  
˜  
])  
In[81]:  
df  
Out[81]:  
ABC  
0123  
1123  
2123  
3123  
4123  
5123  
PreviousBehavior:ifthereturnedshapehappenedtomatchthelengthoforiginalcolumns,thiswouldreturna  
DataFrame  
.Ifthereturnshapedidnotmatch,a  
Series  
withlistswasreturned.  
In[  
3  
]:df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
,  
3  
],axis  
=  
1  
)  
Out[  
3  
]:  
ABC  
0  
1  
2  
3  
1  
1  
2  
3  
2  
1  
2  
3  
3  
1  
2  
3  
4  
1  
2  
3  
5  
1  
2  
3  
In[  
4  
]:df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
],axis  
=  
1  
)  
Out[  
4  
]:  
0  
[  
1  
,  
2  
]  
1  
[  
1  
,  
2  
]  
2  
[  
1  
,  
2  
]  
3  
[  
1  
,  
2  
]  
4  
[  
1  
,  
2  
]  
5  
[  
1  
,  
2  
]  
dtype:  
object  
NewBehavior:Whentheappliedfunctionreturnsalist-like,thiswillnow  
always  
returna  
Series  
.  
1.1.v0.23.0(May15,2017)  
21

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
In[82]:  
df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
,  
3  
],axis  
=  
1  
)  
Out[82]:  
0[1,2,3]  
1[1,2,3]  
2[1,2,3]  
3[1,2,3]  
4[1,2,3]  
5[1,2,3]  
dtype:object  
In[83]:  
df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
],axis  
=  
1  
)  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[83]:  
,  
!  
0[1,2]  
1[1,2]  
2[1,2]  
3[1,2]  
4[1,2]  
5[1,2]  
dtype:object  
Tohaveexpandedcolumns,youcanuse  
result\_type=˜expand˜  
In[84]:  
df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
,  
3  
],axis  
=  
1  
,result\_type  
=  
˜  
expand  
˜  
)  
Out[84]:  
012  
0123  
1123  
2123  
3123  
4123  
5123  
Tobroadcasttheresultacrosstheoriginalcolumns(theoldbehaviourforlist-likesofthecorrectlength),youcanuse  
result\_type=˜broadcast˜  
.Theshapemustmatchtheoriginalcolumns.  
In[85]:  
df  
.  
apply(  
lambda  
x:[  
1  
,  
2  
,  
3  
],axis  
=  
1  
,result\_type  
=  
˜  
broadcast  
˜  
)  
Out[85]:  
ABC  
0123  
1123  
2123  
3123  
4123  
5123  
Returninga  
Series  
allowsonetocontroltheexactreturnstructureandcolumnnames:  
In[86]:  
df  
.  
apply(  
lambda  
x:Series([  
1  
,  
2  
,  
3  
],index  
=  
[  
˜  
D  
˜  
,  
˜  
E  
˜  
,  
˜  
F  
˜  
]),axis  
=  
1  
)  
Out[86]:  
DEF  
0123  
1123  
2123  
3123  
4123  
5123  
22  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.1.2.6  
Concatenationwillnolongersort  
Inafutureversionofpandas  
pandas.concat()  
willnolongersortthenon-concatenationaxiswhenitisnot  
alreadyaligned.Thecurrentbehavioristhesameastheprevious(sorting),butnowawarningisissuedwhen  
sort  
is  
notandthenon-concatenationaxisisnotaligned(  
GH4588  
).  
In[87]:  
df1  
=  
pd  
.  
DataFrame({  
"  
a  
"  
:[  
1  
,  
2  
],  
"  
b  
"  
:[  
1  
,  
2  
]},columns  
=  
[  
˜  
b  
˜  
,  
˜  
a  
˜  
])  
In[88]:  
df2  
=  
pd  
.  
DataFrame({  
"  
a  
"  
:[  
4  
,  
5  
]})  
In[89]:  
pd  
.  
concat([df1,df2])  
Out[89]:  
ab  
011.0  
122.0  
04NaN  
15NaN  
Tokeepthepreviousbehavior(sorting)andsilencethewarning,pass  
sort=True  
In[90]:  
pd  
.  
concat([df1,df2],sort  
=  
True  
)  
Out[90]:  
ab  
011.0  
122.0  
04NaN  
15NaN  
Toacceptthefuturebehavior(nosorting),pass  
sort=False  
Notethatthischangealsoappliesto  
DataFrame.append()  
,whichhasalsoreceiveda  
sort  
keywordforcontrol-  
lingthisbehavior.  
1.1.2.7  
BuildChanges  
Ł  
Buildingpandasfordevelopmentnowrequires  
cython>=0.24  
(  
GH18613  
)  
Ł  
Buildingfromsourcenowexplicitlyrequires  
setuptools  
in  
setup.py  
(  
GH18113  
)  
Ł  
Updatedcondarecipetobeincompliancewithconda-build3.0+(  
GH18002  
)  
1.1.2.8  
IndexDivisionByZeroFillsCorrectly  
Divisionoperationson  
Index  
andsubclasseswillnowdivisionofpositivenumbersbyzerowith  
np.inf  
,division  
ofnegativenumbersbyzerowith  
-np.inf  
and  
0/0  
with  
np.nan  
.Thismatchesexisting  
Series  
behavior.  
(  
GH19322  
,  
GH19347  
)  
PreviousBehavior:  
In[6]:  
index  
=  
pd  
.  
Int64Index([  
-  
1  
,  
0  
,  
1  
])  
In[7]:  
index  
/  
0  
Out[7]:  
Int64Index([0,0,0],dtype=˜int64˜)  
#Previousbehavioryieldeddifferentresultsdependingonthetypeofzerointhe  
,  
!  
divisor  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
23

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
In[8]:  
index  
/  
0.0  
Out[8]:  
Float64Index([-inf,nan,inf],dtype=˜float64˜)  
In[9]:  
index  
=  
pd  
.  
UInt64Index([  
0  
,  
1  
])  
In[10]:  
index  
/  
np  
.  
array([  
0  
,  
0  
],dtype  
=  
np  
.  
uint64)  
Out[10]:  
UInt64Index([0,0],dtype=˜uint64˜)  
In[11]:  
pd  
.  
RangeIndex(  
1  
,  
5  
)  
/  
0  
ZeroDivisionError:integerdivisionormodulobyzero  
CurrentBehavior:  
In[91]:  
index  
=  
pd  
.  
Int64Index([  
-  
1  
,  
0  
,  
1  
])  
#divisionbyzerogives-infinitywherenegative,+infinitywherepositive,andNaN  
,  
!  
for0/0  
In[92]:  
index  
/  
0  
Out[92]:  
Float64Index([-inf,nan,inf],dtype=˜float64˜)  
#Theresultofdivisionbyzeroshouldnotdependonwhetherthezeroisintorfloat  
In[93]:  
index  
/  
0.0  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[93]:Float64Index([-inf,  
,  
!  
nan,inf],dtype=˜float64˜)  
In[94]:  
index  
=  
pd  
.  
UInt64Index([  
0  
,  
1  
])  
In[95]:  
index  
/  
np  
.  
array([  
0  
,  
0  
],dtype  
=  
np  
.  
uint64)  
Out[95]:  
Float64Index([nan,inf],dtype=˜float64˜)  
In[96]:  
pd  
.  
RangeIndex(  
1  
,  
5  
)  
/  
0  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[96]:Float64Index([inf,inf,  
,  
!  
inf,inf],dtype=˜float64˜)  
1.1.2.9  
Extractionofmatchingpatternsfromstrings  
Bydefault,extractingmatchingpatternsfromstringswith  
str.extract()  
usedtoreturna  
Series  
ifasin-  
glegroupwasbeingextracted(a  
DataFrame  
ifmorethanonegroupwasextracted).AsofPandas0.23.0  
str.  
extract()  
alwaysreturnsa  
DataFrame  
,unless  
expand  
issetto  
False  
.Finallay,  
None  
wasanacceptedvalue  
forthe  
expand  
parameter(whichwasequivalentto  
False  
),butnowraisesa  
ValueError  
.(  
GH11386  
)  
PreviousBehavior:  
In[1]:  
s  
=  
pd  
.  
Series([  
˜  
number10  
˜  
,  
˜  
12eggs  
˜  
])  
In[2]:  
extracted  
=  
s  
.  
str  
.  
extract(  
˜  
.  
\*  
(  
\  
d  
\  
d).  
\*  
˜  
)  
In[3]:  
extracted  
Out[3]:  
010  
112  
dtype:object  
In[4]:  
type  
(extracted)  
Out[4]:  
pandas.core.series.Series  
24  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
NewBehavior:  
In[97]:  
s  
=  
pd  
.  
Series([  
˜  
number10  
˜  
,  
˜  
12eggs  
˜  
])  
In[98]:  
extracted  
=  
s  
.  
str  
.  
extract(  
˜  
.  
\*  
(  
\  
d  
\  
d).  
\*  
˜  
)  
In[99]:  
extracted  
Out[99]:  
0  
010  
112  
In[100]:  
type  
(extracted)  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[100]:pandas.core.frame.DataFrame  
Torestorepreviousbehavior,simplyset  
expand  
to  
False  
:  
In[101]:  
s  
=  
pd  
.  
Series([  
˜  
number10  
˜  
,  
˜  
12eggs  
˜  
])  
In[102]:  
extracted  
=  
s  
.  
str  
.  
extract(  
˜  
.  
\*  
(  
\  
d  
\  
d).  
\*  
˜  
,expand  
=  
False  
)  
In[103]:  
extracted  
Out[103]:  
010  
112  
dtype:object  
In[104]:  
type  
(extracted)  
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\Out[104]:pandas.core.series.Series  
1.1.2.10  
Defaultvalueforthe  
ordered  
parameterof  
CategoricalDtype  
Thedefaultvalueofthe  
ordered  
parameterfor  
CategoricalDtype  
haschangedfrom  
False  
to  
None  
toallow  
updatingof  
categories  
withoutimpacting  
ordered  
.Behaviorshouldremainconsistentfordownstreamobjects,  
suchas  
Categorical  
(  
GH18790  
)  
Inpreviousversions,thedefaultvalueforthe  
ordered  
parameterwas  
False  
.Thiscouldpotentiallylead  
tothe  
ordered  
parameterunintentionallybeingchangedfrom  
True  
to  
False  
whenusersattempttoupdate  
categories  
if  
ordered  
isnotexplicitlyasitwouldsilentlydefaultto  
False  
.Thenewbehavior  
for  
ordered=None  
istoretaintheexistingvalueof  
ordered  
.  
NewBehavior:  
In[105]:  
from  
pandas.api.types  
import  
CategoricalDtype  
In[106]:  
cat  
=  
pd  
.  
Categorical(  
list  
(  
˜  
abcaba  
˜  
),ordered  
=  
True  
,categories  
=  
list  
(  
˜  
cba  
˜  
))  
In[107]:  
cat  
Out[107]:  
[a,b,c,a,b,a]  
Categories(3,object):[c<b<a]  
In[108]:  
cdt  
=  
CategoricalDtype(categories  
=  
list  
(  
˜  
cbad  
˜  
))  
In[109]:  
cat  
.  
astype(cdt)  
Out[109]:  
(continuesonnextpage)  
1.1.v0.23.0(May15,2017)  
25

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
[a,b,c,a,b,a]  
Categories(4,object):[c<b<a<d]  
Noticeintheexampleabovethattheconverted  
Categorical  
hasretained  
ordered=True  
.Hadthedefault  
valuefor  
ordered  
remainedas  
False  
,theconverted  
Categorical  
wouldhavebecomeunordered,despite  
ordered=False  
neverbeingexplicitlyTochangethevalueof  
ordered  
,explicitlypassittothenew  
dtype,e.g.  
CategoricalDtype(categories=list(˜cbad˜),ordered=False)  
.  
Notethattheunintenionalconversionof  
ordered  
discussedabovedidnotariseinpreviousversionsduetoseparate  
bugsthatprevented  
astype  
fromdoinganytypeofcategorytocategoryconversion(  
GH10696  
,  
GH18593  
).These  
bugshavebeenedinthisrelease,andmotivatedchangingthedefaultvalueof  
ordered  
.  
1.1.2.11  
Betterpretty-printingofDataFramesinaterminal  
Previously,thedefaultvalueforthemaximumnumberofcolumnswas  
pd.options.display.  
max\_columns=20  
.Thismeantthatrelativelywidedataframeswouldnotwithintheterminalwidth,andpandas  
wouldintroducelinebreakstodisplaythese20columns.Thisresultedinanoutputthatwasrelativelydiftoread:  
IfPythonrunsinaterminal,themaximumnumberofcolumnsisnowdeterminedautomaticallysothattheprinteddata  
framewithinthecurrentterminalwidth(  
pd.options.display.max\_columns=0  
)(  
GH17023  
).IfPython  
runsasaJupyterkernel(suchastheJupyterQtConsoleoraJupyternotebook,aswellasinmanyIDEs),thisvalue  
cannotbeinferredautomaticallyandisthussetto  
20  
asinpreviousversions.Inaterminal,thisresultsinamuchnicer  
output:  
26  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Notethatifyoudon'tlikethenewdefault,youcanalwayssetthisoptionyourself.Toreverttotheoldsetting,you  
canrunthisline:  
pd  
.  
options  
.  
display  
.  
max\_columns  
=  
20  
1.1.2.12  
DatetimelikeAPIChanges  
Ł  
Thedefault  
Timedelta  
constructornowacceptsan  
ISO8601Duration  
stringasanargument  
(  
GH19040  
)  
Ł  
Subtracting  
NaT  
froma  
Series  
with  
dtype=˜datetime64[ns]˜  
returnsa  
Series  
with  
dtype=˜timedelta64[ns]˜  
insteadof  
dtype=˜datetime64[ns]˜  
(  
GH18808  
)  
Ł  
Additionorsubtractionof  
NaT  
from  
TimedeltaIndex  
willreturn  
TimedeltaIndex  
insteadof  
DatetimeIndex  
(  
GH19124  
)  
Ł  
DatetimeIndex.shift()  
and  
TimedeltaIndex.shift()  
willnowraise  
NullFrequencyError  
(whichsubclasses  
ValueError  
,whichwasraisedinolderversions)whentheindexobjectfrequencyis  
None  
(  
GH19147  
)  
Ł  
Additionandsubtractionof  
NaN  
froma  
Series  
with  
dtype=˜timedelta64[ns]˜  
willraisea  
TypeError  
insteadoftreatingthe  
NaN  
as  
NaT  
(  
GH19274  
)  
Ł  
NaT  
divisionwith  
datetime.timedelta  
willnowreturn  
NaN  
insteadofraising(  
GH17876  
)  
1.1.v0.23.0(May15,2017)  
27

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Operationsbetweena  
Series  
withdtype  
dtype=˜datetime64[ns]˜  
anda  
PeriodIndex  
willcor-  
rectlyraises  
TypeError  
(  
GH18850  
)  
Ł  
Subtractionof  
Series  
withtimezone-aware  
dtype=˜datetime64[ns]˜  
withmis-matchedtimezones  
willraise  
TypeError  
insteadof  
ValueError  
(  
GH18817  
)  
Ł  
Timestamp  
willnolongersilentlyignoreunusedorinvalid  
tz  
or  
tzinfo  
keywordarguments(  
GH17690  
)  
Ł  
Timestamp  
willnolongersilentlyignoreinvalid  
freq  
arguments(  
GH5168  
)  
Ł  
CacheableOffset  
and  
WeekDay  
arenolongeravailableinthe  
pandas.tseries.offsets  
module  
(  
GH17830  
)  
Ł  
pandas.tseries.frequencies.get\_freq\_group()  
and  
pandas.tseries.frequencies.  
DAYS  
areremovedfromthepublicAPI(  
GH18034  
)  
Ł  
Series.truncate()  
and  
DataFrame.truncate()  
willraisea  
ValueError  
iftheindexisnotsorted  
insteadofanunhelpful  
KeyError  
(  
GH17935  
)  
Ł  
Series.first  
and  
DataFrame.first  
willnowraisea  
TypeError  
ratherthan  
NotImplementedError  
whenindexisnota  
DatetimeIndex  
(  
GH20725  
).  
Ł  
Series.last  
and  
DataFrame.last  
willnowraisea  
TypeError  
ratherthan  
NotImplementedError  
whenindexisnota  
DatetimeIndex  
(  
GH20725  
).  
Ł  
Restricted  
DateOffset  
keywordarguments.Previously,  
DateOffset  
subclassesallowedarbitrarykeyword  
argumentswhichcouldleadtounexpectedbehavior.Now,onlyvalidargumentswillbeaccepted.(  
GH17176  
,  
GH18226  
).  
Ł  
pandas.merge()  
providesamoreinformativeerrormessagewhentryingtomergeontimezone-awareand  
timezone-naivecolumns(  
GH15800  
)  
Ł  
For  
DatetimeIndex  
and  
TimedeltaIndex  
with  
freq=None  
,additionorsubtractionofinteger-dtyped  
arrayor  
Index  
willraise  
NullFrequencyError  
insteadof  
TypeError  
(  
GH19895  
)  
Ł  
Timestamp  
constructornowacceptsa  
nanosecond  
keywordorpositionalargument(  
GH18898  
)  
Ł  
DatetimeIndex  
willnowraisean  
AttributeError  
whenthe  
tz  
attributeissetafterinstantiation  
(  
GH3746  
)  
Ł  
DatetimeIndex  
witha  
pytz  
timezonewillnowreturnaconsistent  
pytz  
timezone(  
GH18595  
)  
1.1.2.13  
OtherAPIChanges  
Ł  
Series.astype()  
and  
Index.astype()  
withanincompatibledtypewillnowraisea  
TypeError  
ratherthana  
ValueError  
(  
GH18231  
)  
Ł  
Series  
constructionwithan  
object  
dtypedtz-awaredatetimeand  
dtype=object  
willnow  
returnan  
object  
dtyped  
Series  
,previouslythiswouldinferthedatetimedtype(  
GH18231  
)  
Ł  
A  
Series  
of  
dtype=category  
constructedfromanempty  
dict  
willnowhavecategoriesof  
dtype=object  
ratherthan  
dtype=float64  
,consistentlywiththecaseinwhichanemptylistispassed  
(  
GH18515  
)  
Ł  
All-NaNlevelsina  
MultiIndex  
arenowassigned  
float  
ratherthan  
object  
dtype,promotingconsistency  
with  
Index  
(  
GH17929  
).  
Ł  
Levelsnamesofa  
MultiIndex  
(whennotNone)arenowrequiredtobeunique:tryingtocreatea  
MultiIndex  
withrepeatednameswillraisea  
ValueError  
(  
GH18872  
)  
Ł  
Bothconstructionandrenamingof  
Index  
/  
MultiIndex  
withnon-hashable  
name  
/  
names  
willnowraise  
TypeError  
(  
GH20527  
)  
28  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Index.map()  
cannowaccept  
Series  
anddictionaryinputobjects(  
GH12756  
,  
GH18482  
,  
GH18509  
).  
Ł  
DataFrame.unstack()  
willnowdefaulttowith  
np.nan  
for  
object  
columns.(  
GH12815  
)  
Ł  
IntervalIndex  
constructorwillraiseifthe  
closed  
parameterwithhowtheinputdataisinferred  
tobeclosed(  
GH18421  
)  
Ł  
Insertingmissingvaluesintoindexeswillworkforalltypesofindexesandautomaticallyinsertthecorrecttype  
ofmissingvalue(  
NaN  
,  
NaT  
,etc.)regardlessofthetypepassedin(  
GH18295  
)  
Ł  
Whencreatedwithduplicatelabels,  
MultiIndex  
nowraisesa  
ValueError  
.(  
GH17464  
)  
Ł  
Series.fillna()  
nowraisesa  
TypeError  
insteadofa  
ValueError  
whenpassedalist,tupleor  
DataFrameasa  
value  
(  
GH18293  
)  
Ł  
pandas.DataFrame.merge()  
nolongercastsa  
float  
columnto  
object  
whenmergingon  
int  
and  
float  
columns(  
GH16572  
)  
Ł  
pandas.merge()  
nowraisesa  
ValueError  
whentryingtomergeonincompatibledatatypes(  
GH9780  
)  
Ł  
ThedefaultNAvaluefor  
UInt64Index  
haschangedfrom0to  
NaN  
,whichimpactsmethodsthatmaskwith  
NA,suchas  
UInt64Index.where()  
(  
GH18398  
)  
Ł  
Refactored  
setup.py  
touse  
find\_packages  
insteadofexplicitlylistingoutallsubpackages(  
GH18535  
)  
Ł  
Rearrangedtheorderofkeywordargumentsin  
read\_excel()  
toalignwith  
read\_csv()  
(  
GH16672  
)  
Ł  
wide\_to\_long()  
previouslykeptnumeric-likesufesas  
object  
dtype.Nowtheyarecasttonumericif  
possible(  
GH17627  
)  
Ł  
In  
read\_excel()  
,the  
comment  
argumentisnowexposedasanamedparameter(  
GH18735  
)  
Ł  
Rearrangedtheorderofkeywordargumentsin  
read\_excel()  
toalignwith  
read\_csv()  
(  
GH16672  
)  
Ł  
Theoptions  
html.border  
and  
mode.use\_inf\_as\_null  
weredeprecatedinpriorversions,thesewill  
nowshow  
FutureWarning  
ratherthana  
DeprecationWarning  
(  
GH19003  
)  
Ł  
IntervalIndex  
and  
IntervalDtype  
nolongersupportcategorical,object,andstringsubtypes  
(  
GH19016  
)  
Ł  
IntervalDtype  
nowreturns  
True  
whencomparedagainst  
˜interval˜  
regardlessofsubtype,and  
IntervalDtype.name  
nowreturns  
˜interval˜  
regardlessofsubtype(  
GH18980  
)  
Ł  
KeyError  
nowraisesinsteadof  
ValueError  
in  
drop()  
,  
drop()  
,  
drop()  
,  
drop()  
whendroppinga  
non-existentelementinanaxiswithduplicates(  
GH19186  
)  
Ł  
Series.to\_csv()  
nowacceptsa  
compression  
argumentthatworksinthesamewayasthe  
compression  
argumentin  
DataFrame.to\_csv()  
(  
GH18958  
)  
Ł  
Setoperations(union,difference...)on  
IntervalIndex  
withincompatibleindextypeswillnowraisea  
TypeError  
ratherthana  
ValueError  
(  
GH19329  
)  
Ł  
DateOffset  
objectsrendermoresimply,e.g.  
<DateOffset:days=1>  
insteadof  
<DateOffset:  
kwds={˜days˜:1}>  
(  
GH19403  
)  
Ł  
Categorical.fillna  
nowvalidatesits  
value  
and  
method  
keywordarguments.Itnowraiseswhenboth  
ornonearematchingthebehaviorof  
Series.fillna()  
(  
GH19682  
)  
Ł  
pd.to\_datetime(˜today˜)  
nowreturnsadatetime,consistentwith  
pd.Timestamp(˜today˜)  
;pre-  
viously  
pd.to\_datetime(˜today˜)  
returneda  
.normalized()  
datetime(  
GH19935  
)  
Ł  
Series.str.replace()  
nowtakesanoptional  
regex  
keywordwhich,whensetto  
False  
,usesliteral  
stringreplacementratherthanregexreplacement(  
GH16808  
)  
Ł  
DatetimeIndex.strftime()  
and  
PeriodIndex.strftime()  
nowreturnan  
Index  
insteadofa  
numpyarraytobeconsistentwithsimilaraccessors(  
GH20127  
)  
1.1.v0.23.0(May15,2017)  
29

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
ConstructingaSeriesfromalistoflength1nolongerbroadcaststhislistwhenalongerindexis  
(  
GH19714  
,  
GH20391  
).  
Ł  
DataFrame.to\_dict()  
with  
orient=˜index˜  
nolongercastsintcolumnstoforaDataFrame  
withonlyintandcolumns(  
GH18580  
)  
Ł  
Auserthatispassedto  
Series.rolling().aggregate()  
,  
DataFrame.  
rolling().aggregate()  
,oritsexpandingcousins,willnow  
always  
bepasseda  
Series  
,rather  
thana  
np.array  
;  
.apply()  
onlyhasthe  
raw  
keyword,see  
here  
.Thisisconsistentwiththesignaturesof  
.aggregate()  
acrosspandas(  
GH20584  
)  
Ł  
RollingandExpandingtypesraise  
NotImplementedError  
uponiteration(  
GH11704  
).  
1.1.3  
Deprecations  
Ł  
Series.from\_array  
and  
SparseSeries.from\_array  
aredeprecated.Usethenormalconstructor  
Series(..)  
and  
SparseSeries(..)  
instead(  
GH18213  
).  
Ł  
DataFrame.as\_matrix  
isdeprecated.Use  
DataFrame.values  
instead(  
GH18458  
).  
Ł  
Series.asobject  
,  
DatetimeIndex.asobject  
,  
PeriodIndex.asobject  
and  
TimeDeltaIndex.asobject  
havebeendeprecated.Use  
.astype(object)  
instead(  
GH18572  
)  
Ł  
Groupingbyatupleofkeysnowemitsa  
FutureWarning  
andisdeprecated.Inthefuture,atuplepassedto  
˜by˜  
willalwaysrefertoasinglekeythatistheactualtuple,insteadoftreatingthetupleasmultiplekeys.To  
retainthepreviousbehavior,usealistinsteadofatuple(  
GH18314  
)  
Ł  
Series.valid  
isdeprecated.Use  
Series.dropna()  
instead(  
GH18800  
).  
Ł  
read\_excel()  
hasdeprecatedthe  
skip\_footer  
parameter.Use  
skipfooter  
instead(  
GH18836  
)  
Ł  
ExcelFile.parse()  
hasdeprecated  
sheetname  
infavorof  
sheet\_name  
forconsistencywith  
read\_excel()  
(  
GH20920  
).  
Ł  
The  
is\_copy  
attributeisdeprecatedandwillberemovedinafutureversion(  
GH18801  
).  
Ł  
IntervalIndex.from\_intervals  
isdeprecatedinfavorofthe  
IntervalIndex  
constructor  
(  
GH19263  
)  
Ł  
DataFrame.from\_items  
isdeprecated.Use  
DataFrame.from\_dict()  
instead,or  
DataFrame.  
from\_dict(OrderedDict())  
ifyouwishtopreservethekeyorder(  
GH17320  
,  
GH17312  
)  
Ł  
Indexinga  
MultiIndex  
ora  
FloatIndex  
withalistcontainingsomemissingkeyswillnowshowa  
FutureWarning  
,whichisconsistentwithothertypesofindexes(  
GH17758  
).  
Ł  
The  
broadcast  
parameterof  
.apply()  
isdeprecatedinfavorof  
result\_type=˜broadcast˜  
(  
GH18577  
)  
Ł  
The  
reduce  
parameterof  
.apply()  
isdeprecatedinfavorof  
result\_type=˜reduce˜  
(  
GH18577  
)  
Ł  
The  
order  
parameterof  
factorize()  
isdeprecatedandwillberemovedinafuturerelease(  
GH19727  
)  
Ł  
Timestamp.weekday\_name  
,  
DatetimeIndex.weekday\_name  
,and  
Series.dt.  
weekday\_name  
aredeprecatedinfavorof  
Timestamp.day\_name()  
,  
DatetimeIndex.  
day\_name()  
,and  
Series.dt.day\_name()  
(  
GH12806  
)  
Ł  
pandas.tseries.plotting.tsplot  
isdeprecated.Use  
Series.plot()  
instead(  
GH18627  
)  
Ł  
Index.summary()  
isdeprecatedandwillberemovedinafutureversion(  
GH18217  
)  
Ł  
NDFrame.get\_ftype\_counts()  
isdeprecatedandwillberemovedinafutureversion(  
GH18243  
)  
30  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
The  
convert\_datetime64  
parameterin  
DataFrame.to\_records()  
hasbeendeprecatedandwillbe  
removedinafutureversion.TheNumPybugmotivatingthisparameterhasbeenresolved.Thedefaultvalue  
forthisparameterhasalsochangedfrom  
True  
to  
None  
(  
GH18160  
).  
Ł  
Series.rolling().apply()  
,  
DataFrame.rolling().apply()  
,  
Series.expanding().  
apply()  
,and  
DataFrame.expanding().apply()  
havedeprecatedpassingan  
np.array  
bydefault.  
Onewillneedtopassthenew  
raw  
parametertobeexplicitaboutwhatispassed(  
GH20584  
)  
Ł  
The  
data  
,  
base  
,  
strides  
,  
flags  
and  
itemsize  
propertiesofthe  
Series  
and  
Index  
classeshavebeen  
deprecatedandwillberemovedinafutureversion(  
GH20419  
).  
Ł  
DatetimeIndex.offset  
isdeprecated.Use  
DatetimeIndex.freq  
instead(  
GH20716  
)  
Ł  
Floordivisionbetweenanintegerndarrayanda  
Timedelta  
isdeprecated.Divideby  
Timedelta.value  
instead(  
GH19761  
)  
Ł  
Setting  
PeriodIndex.freq  
(whichwasnotguaranteedtoworkcorrectly)isdeprecated.Use  
PeriodIndex.asfreq()  
instead(  
GH20678  
)  
Ł  
Index.get\_duplicates()  
isdeprecatedandwillberemovedinafutureversion(  
GH20239  
)  
Ł  
Thepreviousdefaultbehaviorofnegativeindicesin  
Categorical.take  
isdeprecated.Inafutureversion  
itwillchangefrommeaningmissingvaluestomeaningpositionalindicesfromtheright.Thefuturebehavioris  
consistentwith  
Series.take()  
(  
GH20664  
).  
Ł  
Passingmultipleaxestothe  
axis  
parameterin  
DataFrame.dropna()  
hasbeendeprecatedandwillbe  
removedinafutureversion(  
GH20987  
)  
1.1.4  
Removalofpriorversiondeprecations/changes  
Ł  
Warningsagainsttheobsoleteusage  
Categorical(codes,categories)  
,whichwereemittedforin-  
stancewhenthetwoargumentsto  
Categorical()  
haddifferentdtypes,andrecommendedtheuseof  
Categorical.from\_codes  
,havenowbeenremoved(  
GH8074  
)  
Ł  
The  
levels  
and  
labels  
attributesofa  
MultiIndex  
cannolongerbesetdirectly(  
GH4039  
).  
Ł  
pd.tseries.util.pivot\_annual  
hasbeenremoved(deprecatedsincev0.19).Use  
pivot\_table  
instead(  
GH18370  
)  
Ł  
pd.tseries.util.isleapyear  
hasbeenremoved(deprecatedsincev0.19).Use  
.is\_leap\_year  
propertyinDatetime-likesinstead(  
GH18370  
)  
Ł  
pd.ordered\_merge  
hasbeenremoved(deprecatedsincev0.19).Use  
pd.merge\_ordered  
instead  
(  
GH18459  
)  
Ł  
The  
SparseList  
classhasbeenremoved(  
GH14007  
)  
Ł  
The  
pandas.io.wb  
and  
pandas.io.data  
stubmoduleshavebeenremoved(  
GH13735  
)  
Ł  
Categorical.from\_array  
hasbeenremoved(  
GH13854  
)  
Ł  
The  
freq  
and  
how  
parametershavebeenremovedfromthe  
rolling  
/  
expanding  
/  
ewm  
methodsof  
DataFrameandSeries(deprecatedsincev0.18).Instead,resamplebeforecallingthemethods.(  
GH18601  
&  
GH18668  
)  
Ł  
DatetimeIndex.to\_datetime  
,  
Timestamp.to\_datetime  
,  
PeriodIndex.to\_datetime  
,and  
Index.to\_datetime  
havebeenremoved(  
GH8254  
,  
GH14096  
,  
GH14113  
)  
Ł  
read\_csv()  
hasdroppedthe  
skip\_footer  
parameter(  
GH13386  
)  
Ł  
read\_csv()  
hasdroppedthe  
as\_recarray  
parameter(  
GH13373  
)  
Ł  
read\_csv()  
hasdroppedthe  
buffer\_lines  
parameter(  
GH13360  
)  
1.1.v0.23.0(May15,2017)  
31

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
read\_csv()  
hasdroppedthe  
compact\_ints  
and  
use\_unsigned  
parameters(  
GH13323  
)  
Ł  
The  
Timestamp  
classhasdroppedthe  
offset  
attributeinfavorof  
freq  
(  
GH13593  
)  
Ł  
The  
Series  
,  
Categorical  
,and  
Index  
classeshavedroppedthe  
reshape  
method(  
GH13012  
)  
Ł  
pandas.tseries.frequencies.get\_standard\_freq  
hasbeenremovedinfavorof  
pandas.  
tseries.frequencies.to\_offset(freq).rule\_code  
(  
GH13874  
)  
Ł  
The  
freqstr  
keywordhasbeenremovedfrom  
pandas.tseries.frequencies.to\_offset  
infavor  
of  
freq  
(  
GH13874  
)  
Ł  
The  
Panel4D  
and  
PanelND  
classeshavebeenremoved(  
GH13776  
)  
Ł  
The  
Panel  
classhasdroppedthe  
to\_long  
and  
toLong  
methods(  
GH19077  
)  
Ł  
Theoptions  
display.line\_with  
and  
display.height  
areremovedinfavorof  
display.width  
and  
display.max\_rows  
respectively(  
GH4391  
,  
GH19107  
)  
Ł  
The  
labels  
attributeofthe  
Categorical  
classhasbeenremovedinfavorof  
Categorical.codes  
(  
GH7768  
)  
Ł  
The  
flavor  
parameterhavebeenremovedfromfunc:  
to\_sql  
method(  
GH13611  
)  
Ł  
Themodules  
pandas.tools.hashing  
and  
pandas.util.hashing  
havebeenremoved(  
GH16223  
)  
Ł  
Thetop-levelfunctions  
pd.rolling\_  
\*  
,  
pd.expanding\_  
\*  
and  
pd.ewm  
\*  
havebeenremoved(Deprecated  
sincev0.18).Instead,usetheDataFrame/Seriesmethods  
rolling  
,  
expanding  
and  
ewm  
(  
GH18723  
)  
Ł  
Importsfrom  
pandas.core.common  
forfunctionssuchas  
is\_datetime64\_dtype  
arenowremoved.  
Thesearelocatedin  
pandas.api.types  
.(  
GH13634  
,  
GH19769  
)  
Ł  
The  
infer\_dst  
keywordin  
Series.tz\_localize()  
,  
DatetimeIndex.tz\_localize()  
and  
DatetimeIndex  
havebeenremoved.  
infer\_dst=True  
isequivalentto  
ambiguous=˜infer˜  
,and  
infer\_dst=False  
to  
ambiguous=˜raise˜  
(  
GH7963  
).  
Ł  
When  
.resample()  
waschangedfromaneagertoalazyoperation,like  
.groupby()  
inv0.18.0,weput  
inplacecompatibility(witha  
FutureWarning  
),sooperationswouldcontinuetowork.Thisisnowfully  
removed,soa  
Resampler  
willnolongerforwardcompatoperations(  
GH20554  
)  
Ł  
Removelongdeprecated  
axis=None  
parameterfrom  
.replace()  
(  
GH20271  
)  
1.1.5  
PerformanceImprovements  
Ł  
Indexerson  
Series  
or  
DataFrame  
nolongercreateareferencecycle(  
GH17956  
)  
Ł  
Addedakeywordargument,  
cache  
,to  
to\_datetime()  
thatimprovedtheperformanceofconvertingdupli-  
catedatetimearguments(  
GH11665  
)  
Ł  
DateOffset  
arithmeticperformanceisimproved(  
GH18218  
)  
Ł  
Convertinga  
Series  
of  
Timedelta  
objectstodays,seconds,etc...spedupthroughvectorizationofunder-  
lyingmethods(  
GH18092  
)  
Ł  
Improvedperformanceof  
.map()  
witha  
Series/dict  
input(  
GH15081  
)  
Ł  
Theoverridden  
Timedelta  
propertiesofdays,secondsandmicrosecondshavebeenremoved,leveragingtheir  
built-inPythonversionsinstead(  
GH18242  
)  
Ł  
Series  
constructionwillreducethenumberofcopiesmadeoftheinputdataincertaincases(  
GH17449  
)  
Ł  
Improvedperformanceof  
Series.dt.date()  
and  
DatetimeIndex.date()  
(  
GH18058  
)  
Ł  
Improvedperformanceof  
Series.dt.time()  
and  
DatetimeIndex.time()  
(  
GH18461  
)  
32  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Improvedperformanceof  
IntervalIndex.symmetric\_difference()  
(  
GH18475  
)  
Ł  
Improvedperformanceof  
DatetimeIndex  
and  
Series  
arithmeticoperationswithBusiness-Monthand  
Business-Quarterfrequencies(  
GH18489  
)  
Ł  
Series()  
/  
DataFrame()  
tabcompletionlimitsto100values,forbetterperformance.(  
GH18587  
)  
Ł  
Improvedperformanceof  
DataFrame.median()  
with  
axis=1  
whenbottleneckisnotinstalled(  
GH16468  
)  
Ł  
Improvedperformanceof  
MultiIndex.get\_loc()  
forlargeindexes,atthecostofareductioninperfor-  
manceforsmallones(  
GH18519  
)  
Ł  
Improvedperformanceof  
MultiIndex.remove\_unused\_levels()  
whentherearenounusedlevels,at  
thecostofareductioninperformancewhenthereare(  
GH19289  
)  
Ł  
Improvedperformanceof  
Index.get\_loc()  
fornon-uniqueindexes(  
GH19478  
)  
Ł  
Improvedperformanceofpairwise  
.rolling()  
and  
.expanding()  
with  
.cov()  
and  
.corr()  
opera-  
tions(  
GH17917  
)  
Ł  
Improvedperformanceof  
pandas.core.groupby.GroupBy.rank()  
(  
GH15779  
)  
Ł  
Improvedperformanceofvariable  
.rolling()  
on  
.min()  
and  
.max()  
(  
GH19521  
)  
Ł  
Improvedperformanceof  
pandas.core.groupby.GroupBy.ffill()  
and  
pandas.core.  
groupby.GroupBy.bfill()  
(  
GH11296  
)  
Ł  
Improvedperformanceof  
pandas.core.groupby.GroupBy.any()  
and  
pandas.core.groupby.  
GroupBy.all()  
(  
GH15435  
)  
Ł  
Improvedperformanceof  
pandas.core.groupby.GroupBy.pct\_change()  
(  
GH19165  
)  
Ł  
Improvedperformanceof  
Series.isin()  
inthecaseofcategoricaldtypes(  
GH20003  
)  
Ł  
Improvedperformanceof  
getattr(Series,attr)  
whentheSerieshascertainindextypes.This  
estedinslowprintingoflargeSerieswitha  
DatetimeIndex  
(  
GH19764  
)  
Ł  
Fixedaperformanceregressionfor  
GroupBy.nth()  
and  
GroupBy.last()  
withsomeobjectcolumns  
(  
GH19283  
)  
Ł  
Improvedperformanceof  
pandas.core.arrays.Categorical.from\_codes()  
(  
GH18501  
)  
1.1.6  
DocumentationChanges  
ThankstoallofthecontributorswhoparticipatedinthePandasDocumentationSprint,whichtookplaceonMarch  
10th.Wehadabout500participantsfromover30locationsacrosstheworld.Youshouldnoticethatmanyofthe  
API  
docstrings  
havegreatlyimproved.  
Thereweretoomanysimultaneouscontributionstoincludeareleasenoteforeachimprovement,butthis  
GitHub  
search  
shouldgiveyouanideaofhowmanydocstringswereimproved.  
Specialthanksto  
MarcGarcia  
fororganizingthesprint.Formoreinformation,readthe  
NumFOCUSblogpost  
recap-  
pingthesprint.  
Ł  
ChangedspellingofﬁnumpyﬂtoﬁNumPyﬂ,andﬁpythonﬂtoﬁPythonﬂ.(  
GH19017  
)  
Ł  
Consistencywhenintroducingcodesamples,usingeithercolonorperiod.Rewrotesomesentencesforgreater  
clarity,addedmoredynamicreferencestofunctions,methodsandclasses.(  
GH18941  
,  
GH18948  
,  
GH18973  
,  
GH19017  
)  
Ł  
Addedareferenceto  
DataFrame.assign()  
intheconcatenatesectionofthemergingdocumentation  
(  
GH18665  
)  
1.1.v0.23.0(May15,2017)  
33

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.1.7  
BugFixes  
1.1.7.1  
Categorical  
Warning:  
Aclassofbugswereintroducedinpandas0.21with  
CategoricalDtype  
thataffectsthecorrectness  
ofoperationslike  
merge  
,  
concat  
,andindexingwhencomparingmultipleunordered  
Categorical  
arrays  
thathavethesamecategories,butinadifferentorder.Wehighlyrecommendupgradingormanuallyaligningyour  
categoriesbeforedoingtheseoperations.  
Ł  
Bugin  
Categorical.equals  
returningthewrongresultwhencomparingtwounordered  
Categorical  
arrayswiththesamecategories,butinadifferentorder(  
GH16603  
)  
Ł  
Bugin  
pandas.api.types.union\_categoricals()  
returningthewrongresultwhenforunordered  
categoricalswiththecategoriesinadifferentorder.Thisaffected  
pandas.concat()  
withCategoricaldata  
(  
GH19096  
).  
Ł  
Bugin  
pandas.merge()  
returningthewrongresultwhenjoiningonanunordered  
Categorical  
thathad  
thesamecategoriesbutinadifferentorder(  
GH19551  
)  
Ł  
Bugin  
CategoricalIndex.get\_indexer()  
returningthewrongresultwhen  
target  
wasanun-  
ordered  
Categorical  
thathadthesamecategoriesas  
self  
butinadifferentorder(  
GH19551  
)  
Ł  
Bugin  
Index.astype()  
withacategoricaldtypewheretheresultantindexisnotconvertedtoa  
CategoricalIndex  
foralltypesofindex(  
GH18630  
)  
Ł  
Bugin  
Series.astype()  
and  
Categorical.astype()  
whereanexistingcategoricaldatadoesnotget  
updated(  
GH10696  
,  
GH18593  
)  
Ł  
Bugin  
Series.str.split()  
with  
expand=True  
incorrectlyraisinganIndexErroronemptystrings  
(  
GH20002  
).  
Ł  
Bugin  
Index  
constructorwith  
dtype=CategoricalDtype(...)  
where  
categories  
and  
ordered  
arenotmaintained(  
GH19032  
)  
Ł  
Bugin  
Series  
constructorwithscalarand  
dtype=CategoricalDtype(...)  
where  
categories  
and  
ordered  
arenotmaintained(  
GH19565  
)  
Ł  
Bugin  
Categorical.\_\_iter\_\_  
notconvertingtoPythontypes(  
GH19909  
)  
Ł  
Bugin  
pandas.factorize()  
returningtheuniquecodesforthe  
uniques  
.Thisnowreturnsa  
Categorical  
withthesamedtypeastheinput(  
GH19721  
)  
Ł  
Bugin  
pandas.factorize()  
includinganitemformissingvaluesinthe  
uniques  
returnvalue  
(  
GH19721  
)  
Ł  
Bugin  
Series.take()  
withcategoricaldatainterpreting  
-1  
in  
indices  
asmissingvaluemarkers,ratherthan  
thelastelementoftheSeries(  
GH20664  
)  
1.1.7.2  
Datetimelike  
Ł  
Bugin  
Series.\_\_sub\_\_()  
subtractinganon-nanosecond  
np.datetime64  
objectfroma  
Series  
gave  
incorrectresults(  
GH7996  
)  
Ł  
Bugin  
DatetimeIndex  
,  
TimedeltaIndex  
additionandsubtractionofzero-dimensionalintegerarrays  
gaveincorrectresults(  
GH19012  
)  
34  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
DatetimeIndex  
and  
TimedeltaIndex  
whereaddingorsubtractinganarray-likeof  
DateOffset  
objectseitherraised(  
np.array  
,  
pd.Index  
)orbroadcastincorrectly(  
pd.Series  
)  
(  
GH18849  
)  
Ł  
Bugin  
Series.\_\_add\_\_()  
addingSerieswithdtype  
timedelta64[ns]  
toatimezone-aware  
DatetimeIndex  
incorrectlydroppedtimezoneinformation(  
GH13905  
)  
Ł  
Addinga  
Period  
objecttoa  
datetime  
or  
Timestamp  
objectwillnowcorrectlyraisea  
TypeError  
(  
GH17983  
)  
Ł  
Bugin  
Timestamp  
wherecomparisonwithanarrayof  
Timestamp  
objectswouldresultina  
RecursionError  
(  
GH15183  
)  
Ł  
Bugin  
Series  
-divisionwhereoperatingonascalar  
timedelta  
raisesanexception(  
GH18846  
)  
Ł  
Bugin  
DatetimeIndex  
wherethereprwasnotshowinghigh-precisiontimevaluesattheendofaday(e.g.,  
23:59:59.999999999)(  
GH19030  
)  
Ł  
Bugin  
.astype()  
tonon-nstimedeltaunitswouldholdtheincorrectdtype(  
GH19176  
,  
GH19223  
,  
GH12425  
)  
Ł  
Buginsubtracting  
Series  
from  
NaT  
incorrectlyreturning  
NaT  
(  
GH19158  
)  
Ł  
Bugin  
Series.truncate()  
whichraises  
TypeError  
withamonotonic  
PeriodIndex  
(  
GH17717  
)  
Ł  
Bugin  
pct\_change()  
using  
periods  
and  
freq  
returneddifferentlengthoutputs(  
GH7292  
)  
Ł  
Bugincomparisonof  
DatetimeIndex  
against  
None  
or  
datetime.date  
objectsraising  
TypeError  
for  
==  
and  
!=  
comparisonsinsteadofall-  
False  
andall-  
True  
,respectively(  
GH19301  
)  
Ł  
Bugin  
Timestamp  
and  
to\_datetime()  
whereastringrepresentingabarelyout-of-boundstimestamp  
wouldbeincorrectlyroundeddowninsteadofraising  
OutOfBoundsDatetime  
(  
GH19382  
)  
Ł  
Bugin  
Timestamp.floor()  
DatetimeIndex.floor()  
wheretimestampsfarinthefutureandpast  
werenotroundedcorrectly(  
GH19206  
)  
Ł  
Bugin  
to\_datetime()  
wherepassinganout-of-boundsdatetimewith  
errors=˜coerce˜  
and  
utc=True  
wouldraise  
OutOfBoundsDatetime  
insteadofparsingto  
NaT  
(  
GH19612  
)  
Ł  
Bugin  
DatetimeIndex  
and  
TimedeltaIndex  
additionandsubtractionwherenameofthereturnedobject  
wasnotalwayssetconsistently.(  
GH19744  
)  
Ł  
Bugin  
DatetimeIndex  
and  
TimedeltaIndex  
additionandsubtractionwhereoperationswithnumpy  
arraysraised  
TypeError  
(  
GH19847  
)  
Ł  
Bugin  
DatetimeIndex  
and  
TimedeltaIndex  
wheresettingthe  
freq  
attributewasnotfullysupported  
(  
GH20678  
)  
1.1.7.3  
Timedelta  
Ł  
Bugin  
Timedelta.\_\_mul\_\_()  
wheremultiplyingby  
NaT  
returned  
NaT  
insteadofraisinga  
TypeError  
(  
GH19819  
)  
Ł  
Bugin  
Series  
with  
dtype=˜timedelta64[ns]˜  
whereadditionorsubtractionof  
TimedeltaIndex  
hadresultscastto  
dtype=˜int64˜  
(  
GH17250  
)  
Ł  
Bugin  
Series  
with  
dtype=˜timedelta64[ns]˜  
whereadditionorsubtractionof  
TimedeltaIndex  
couldreturna  
Series  
withanincorrectname(  
GH19043  
)  
Ł  
Bugin  
Timedelta.\_\_floordiv\_\_()  
and  
Timedelta.\_\_rfloordiv\_\_()  
dividingbymanyincom-  
patiblenumpyobjectswasincorrectlyallowed(  
GH18846  
)  
Ł  
Bugwheredividingascalartimedelta-likeobjectwith  
TimedeltaIndex  
performedthereciprocaloperation  
(  
GH19125  
)  
1.1.v0.23.0(May15,2017)  
35

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
TimedeltaIndex  
wheredivisionbya  
Series  
wouldreturna  
TimedeltaIndex  
insteadofa  
Series  
(  
GH19042  
)  
Ł  
Bugin  
Timedelta.\_\_add\_\_()  
,  
Timedelta.\_\_sub\_\_()  
whereaddingorsubtractinga  
np.  
timedelta64  
objectwouldreturnanother  
np.timedelta64  
insteadofa  
Timedelta  
(  
GH19738  
)  
Ł  
Bugin  
Timedelta.\_\_floordiv\_\_()  
,  
Timedelta.\_\_rfloordiv\_\_()  
whereoperatingwitha  
Tick  
objectwouldraisea  
TypeError  
insteadofreturninganumericvalue(  
GH19738  
)  
Ł  
Bugin  
Period.asfreq()  
whereperiodsnear  
datetime(1,1,1)  
couldbeconvertedincorrectly  
(  
GH19643  
,  
GH19834  
)  
Ł  
Bugin  
Timedelta.total\_seconds()  
causingprecisionerrors,forexample  
Timedelta(˜30S˜).  
total\_seconds()==30.000000000000004  
(  
GH19458  
)  
Ł  
Bugin  
Timedelta.\_\_rmod\_\_()  
whereoperatingwitha  
numpy.timedelta64  
returneda  
timedelta64  
objectinsteadofa  
Timedelta  
(  
GH19820  
)  
Ł  
Multiplicationof  
TimedeltaIndex  
by  
TimedeltaIndex  
willnowraise  
TypeError  
insteadofraising  
ValueError  
incasesoflengthmis-match(  
GH19333  
)  
Ł  
Buginindexinga  
TimedeltaIndex  
witha  
np.timedelta64  
objectwhichwasraisinga  
TypeError  
(  
GH20393  
)  
1.1.7.4  
Timezones  
Ł  
Bugincreatinga  
Series  
fromanarraythatcontainsbothtz-naiveandtz-awarevalueswillresultina  
Series  
whosedtypeistz-awareinsteadofobject(  
GH16406  
)  
Ł  
Bugincomparisonoftimezone-aware  
DatetimeIndex  
against  
NaT  
incorrectlyraising  
TypeError  
(  
GH19276  
)  
Ł  
Bugin  
DatetimeIndex.astype()  
whenconvertingbetweentimezoneawaredtypes,andconvertingfrom  
timezoneawaretonaive(  
GH18951  
)  
Ł  
Bugincomparing  
DatetimeIndex  
,whichfailedtoraise  
TypeError  
whenattemptingtocompare  
timezone-awareandtimezone-naivedatetimelikeobjects(  
GH18162  
)  
Ł  
Buginlocalizationofanaive,datetimestringina  
Series  
constructorwitha  
datetime64[ns,tz]  
dtype  
(  
GH174151  
)  
Ł  
Timestamp.replace()  
willnowhandleDaylightSavingstransitionsgracefully(  
GH18319  
)  
Ł  
Bugintz-aware  
DatetimeIndex  
whereaddition/subtractionwitha  
TimedeltaIndex  
orarraywith  
dtype=˜timedelta64[ns]˜  
wasincorrect(  
GH17558  
)  
Ł  
Bugin  
DatetimeIndex.insert()  
whereinserting  
NaT  
intoatimezone-awareindexincorrectlyraised  
(  
GH16357  
)  
Ł  
Bugin  
DataFrame  
constructor,wheretz-awareDatetimeindexandagivencolumnnamewillresultinan  
empty  
DataFrame  
(  
GH19157  
)  
Ł  
Bugin  
Timestamp.tz\_localize()  
wherelocalizingatimestampneartheminimumormaximumvalid  
valuescouldovwandreturnatimestampwithanincorrectnanosecondvalue(  
GH12677  
)  
Ł  
Bugwheniteratingover  
DatetimeIndex  
thatwaslocalizedwithedtimezoneoffsetthatroundednanosec-  
ondprecisiontomicroseconds(  
GH19603  
)  
Ł  
Bugin  
DataFrame.diff()  
thatraisedan  
IndexError  
withtz-awarevalues(  
GH18578  
)  
Ł  
Bugin  
melt()  
thatconvertedtz-awaredtypestotz-naive(  
GH15785  
)  
36  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
Dataframe.count()  
thatraisedan  
ValueError  
,if  
Dataframe.dropna()  
wascalledfora  
singlecolumnwithtimezone-awarevalues.(  
GH13407  
)  
1.1.7.5  
Offsets  
Ł  
Bugin  
WeekOfMonth  
and  
Week  
whereadditionandsubtractiondidnotrollcorrectly(  
GH18510  
,  
GH18672  
,  
GH18864  
)  
Ł  
Bugin  
WeekOfMonth  
and  
LastWeekOfMonth  
wheredefaultkeywordargumentsforconstructorraised  
ValueError  
(  
GH19142  
)  
Ł  
Bugin  
FY5253Quarter  
,  
LastWeekOfMonth  
whererollbackandrollforwardbehaviorwasinconsistent  
withadditionandsubtractionbehavior(  
GH18854  
)  
Ł  
Bugin  
FY5253  
where  
datetime  
additionandsubtractionincrementedincorrectlyfordatesontheyear-end  
butnotnormalizedtomidnight(  
GH18854  
)  
Ł  
Bugin  
FY5253  
wheredateoffsetscouldincorrectlyraisean  
AssertionError  
inarithmeticoperatons  
(  
GH14774  
)  
1.1.7.6  
Numeric  
Ł  
Bugin  
Series  
constructorwithanintorlistwherespecifying  
dtype=str  
,  
dtype=˜str˜  
or  
dtype=˜U˜  
failedtoconvertthedataelementstostrings(  
GH16605  
)  
Ł  
Bugin  
Index  
multiplicationanddivisionmethodswhereoperatingwitha  
Series  
wouldreturnan  
Index  
objectinsteadofa  
Series  
object(  
GH19042  
)  
Ł  
Buginthe  
DataFrame  
constructorinwhichdatacontainingverylargepositiveorverylargenegativenumbers  
wascausing  
OverflowError  
(  
GH18584  
)  
Ł  
Bugin  
Index  
constructorwith  
dtype=˜uint64˜  
whereint-likewerenotcoercedto  
UInt64Index  
(  
GH18400  
)  
Ł  
Bugin  
DataFrame  
xarithmetic(e.g.  
df.add(other,fill\_value=foo)  
)witha  
fill\_value  
otherthan  
None  
failedtoraise  
NotImplementedError  
incornercaseswhereeithertheframeor  
other  
haslengthzero(  
GH19522  
)  
Ł  
Multiplicationanddivisionofnumeric-dtyped  
Index  
objectswithtimedelta-likescalarsreturns  
TimedeltaIndex  
insteadofraising  
TypeError  
(  
GH19333  
)  
Ł  
Bugwhere  
NaN  
wasreturnedinsteadof0by  
Series.pct\_change()  
and  
DataFrame.pct\_change()  
when  
fill\_method  
isnot  
None  
(  
GH19873  
)  
1.1.7.7  
Strings  
Ł  
Bugin  
Series.str.get()  
withadictionaryinthevaluesandtheindexnotinthekeys,raising  
KeyError  
(  
GH20671  
)  
1.1.7.8  
Indexing  
Ł  
Bugin  
Index  
constructionfromlistofmixedtypetuples(  
GH18505  
)  
Ł  
Bugin  
Index.drop()  
whenpassingalistofbothtuplesandnon-tuples(  
GH18304  
)  
Ł  
Bugin  
DataFrame.drop()  
,  
Panel.drop()  
,  
Series.drop()  
,  
Index.drop()  
whereno  
KeyError  
israisedwhendroppinganon-existentelementfromanaxisthatcontainsduplicates(  
GH19186  
)  
1.1.v0.23.0(May15,2017)  
37

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Buginindexingadatetimelike  
Index  
thatraised  
ValueError  
insteadof  
IndexError  
(  
GH18386  
).  
Ł  
Index.to\_series()  
nowaccepts  
index  
and  
name  
kwargs(  
GH18699  
)  
Ł  
DatetimeIndex.to\_series()  
nowaccepts  
index  
and  
name  
kwargs(  
GH18699  
)  
Ł  
Buginindexingnon-scalarvaluefrom  
Series  
havingnon-unique  
Index  
willreturnvalue  
(  
GH17610  
)  
Ł  
Buginindexingwithiteratorcontainingonlymissingkeys,whichraisednoerror(  
GH20748  
)  
Ł  
Fixedinconsistencyin  
.ix  
betweenlistandscalarkeyswhentheindexhasintegerdtypeanddoesnotinclude  
thedesiredkeys(  
GH20753  
)  
Ł  
Bugin  
\_\_setitem\_\_  
whenindexinga  
DataFrame  
witha2-dbooleanndarray(  
GH18582  
)  
Ł  
Bugin  
str.extractall  
whentherewerenomatchesempty  
Index  
wasreturnedinsteadofappropriate  
MultiIndex  
(  
GH19034  
)  
Ł  
Bugin  
IntervalIndex  
whereemptyandpurelyNAdatawasconstructedinconsistentlydependingonthe  
constructionmethod(  
GH18421  
)  
Ł  
Bugin  
IntervalIndex.symmetric\_difference()  
wherethesymmetricdifferencewithanon-  
IntervalIndex  
didnotraise(  
GH18475  
)  
Ł  
Bugin  
IntervalIndex  
wheresetoperationsthatreturnedanempty  
IntervalIndex  
hadthewrongdtype  
(  
GH19101  
)  
Ł  
Bugin  
DataFrame.drop\_duplicates()  
whereno  
KeyError  
israisedwhenpassingincolumnsthat  
don'texistonthe  
DataFrame  
(  
GH19726  
)  
Ł  
Bugin  
Index  
subclassesconstructorsthatignoreunexpectedkeywordarguments(  
GH19348  
)  
Ł  
Bugin  
Index.difference()  
whentakingdifferenceofan  
Index  
withitself(  
GH20040  
)  
Ł  
Bugin  
DataFrame.first\_valid\_index()  
and  
DataFrame.last\_valid\_index()  
inpresence  
ofentirerowsofNaNsinthemiddleofvalues(  
GH20499  
).  
Ł  
Bugin  
IntervalIndex  
wheresomeindexingoperationswerenotsupportedforoverlappingornon-  
monotonic  
uint64  
data(  
GH20636  
)  
Ł  
Bugin  
Series.is\_unique  
whereextraneousoutputinstderrisshownifSeriescontainsobjectswith  
\_\_ne\_\_  
(  
GH20661  
)  
Ł  
Bugin  
.loc  
assignmentwithasingle-elementlist-likeincorrectlyassignsasalist(  
GH19474  
)  
Ł  
Buginpartialstringindexingona  
Series/DataFrame  
withamonotonicdecreasing  
DatetimeIndex  
(  
GH19362  
)  
Ł  
Buginperformingin-placeoperationsona  
DataFrame  
withaduplicate  
Index  
(  
GH17105  
)  
Ł  
Bugin  
IntervalIndex.get\_loc()  
and  
IntervalIndex.get\_indexer()  
whenusedwithan  
IntervalIndex  
containingasingleinterval(  
GH17284  
,  
GH20921  
)  
Ł  
Bugin  
.loc  
witha  
uint64  
indexer(  
GH20722  
)  
1.1.7.9  
MultiIndex  
Ł  
Bugin  
MultiIndex.\_\_contains\_\_()  
wherenon-tuplekeyswouldreturn  
True  
eveniftheyhadbeen  
dropped(  
GH19027  
)  
Ł  
Bugin  
MultiIndex.set\_labels()  
whichwouldcausecasting(andpotentiallyclipping)ofthenewlabels  
ifthe  
level  
argumentisnot0oralistlike[0,1,...](  
GH19057  
)  
38  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
MultiIndex.get\_level\_values()  
whichwouldreturnaninvalidindexonlevelofintswith  
missingvalues(  
GH17924  
)  
Ł  
Bugin  
MultiIndex.unique()  
whencalledonempty  
MultiIndex  
(  
GH20568  
)  
Ł  
Bugin  
MultiIndex.unique()  
whichwouldnotpreservelevelnames(  
GH20570  
)  
Ł  
Bugin  
MultiIndex.remove\_unused\_levels()  
whichwouldnanvalues(  
GH18417  
)  
Ł  
Bugin  
MultiIndex.from\_tuples()  
whichwouldfailtotakezippedtuplesinpython3(  
GH18434  
)  
Ł  
Bugin  
MultiIndex.get\_loc()  
whichwouldfailtoautomaticallycastvaluesbetweenandint  
(  
GH18818  
,  
GH15994  
)  
Ł  
Bugin  
MultiIndex.get\_loc()  
whichwouldcastbooleantointegerlabels(  
GH19086  
)  
Ł  
Bugin  
MultiIndex.get\_loc()  
whichwouldfailtolocatekeyscontaining  
NaN  
(  
GH18485  
)  
Ł  
Bugin  
MultiIndex.get\_loc()  
inlarge  
MultiIndex  
,wouldfailwhenlevelshaddifferentdtypes  
(  
GH18520  
)  
Ł  
Buginindexingwherenestedindexershavingonlynumpyarraysarehandledincorrectly(  
GH19686  
)  
1.1.7.10  
I/O  
Ł  
read\_html()  
nowrewindsseekableIOobjectsafterparsefailure,beforeattemptingtoparsewithanew  
parser.Ifaparsererrorsandtheobjectisnon-seekable,aninformativeerrorisraisedsuggestingtheuseofa  
differentparser(  
GH17975  
)  
Ł  
DataFrame.to\_html()  
nowhasanoptiontoaddanidtotheleading  
<table>  
tag(  
GH8496  
)  
Ł  
Bugin  
read\_msgpack()  
withanonexistentispassedinPython2(  
GH15296  
)  
Ł  
Bugin  
read\_csv()  
wherea  
MultiIndex  
withduplicatecolumnswasnotbeingmangledappropriately  
(  
GH18062  
)  
Ł  
Bugin  
read\_csv()  
wheremissingvalueswerenotbeinghandledproperlywhen  
keep\_default\_na=False  
withdictionary  
na\_values  
(  
GH19227  
)  
Ł  
Bugin  
read\_csv()  
causingheapcorruptionon32-bit,big-endianarchitectures(  
GH20785  
)  
Ł  
Bugin  
read\_sas()  
whereawith0variablesgavean  
AttributeError  
incorrectly.Nowitgivesan  
EmptyDataError  
(  
GH18184  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
wherepairsofbracesmeanttoserveasinvisibleplaceholderswerees-  
caped(  
GH18667  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
wherea  
NaN  
ina  
MultiIndex  
wouldcausean  
IndexError  
orincor-  
rectoutput(  
GH14249  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
whereanon-stringindex-levelnamewouldresultinan  
AttributeError  
(  
GH19981  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
wherethecombinationofanindexnameandthe  
index\_names=False  
optionwouldresultinincorrectoutput(  
GH18326  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
wherea  
MultiIndex  
withanemptystringasitsnamewouldresultin  
incorrectoutput(  
GH18669  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
wheremissingspacecharacterscausedwrongescapingandproduced  
non-validlatexinsomecases(  
GH20859  
)  
Ł  
Bugin  
read\_json()  
wherelargenumericvalueswerecausingan  
OverflowError  
(  
GH18842  
)  
Ł  
Bugin  
DataFrame.to\_parquet()  
whereanexceptionwasraisedifthewritedestinationisS3(  
GH19134  
)  
1.1.v0.23.0(May15,2017)  
39

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Interval  
nowsupportedin  
DataFrame.to\_excel()  
forallExceltypes(  
GH19242  
)  
Ł  
Timedelta  
nowsupportedin  
DataFrame.to\_excel()  
forallExceltypes(  
GH19242  
,  
GH9155  
,  
GH19900  
)  
Ł  
Bugin  
pandas.io.stata.StataReader.value\_labels()  
raisingan  
AttributeError  
when  
calledonveryoldNowreturnsanemptydict(  
GH19417  
)  
Ł  
Bugin  
read\_pickle()  
whenunpicklingobjectswith  
TimedeltaIndex  
or  
Float64Index  
created  
withpandaspriortoversion0.20(  
GH19939  
)  
Ł  
Bugin  
pandas.io.json.json\_normalize()  
wheresubrecordsarenotproperlynormalizedifanysub-  
recordsvaluesareNoneType(  
GH20030  
)  
Ł  
Bugin  
usecols  
parameterin  
read\_csv()  
whereerrorisnotraisedcorrectlywhenpassingastring.  
(  
GH20529  
)  
Ł  
Bugin  
HDFStore.keys()  
whenreadingawithasoftlinkcausesexception(  
GH20523  
)  
Ł  
Bugin  
HDFStore.select\_column()  
whereakeywhichisnotavalidstoreraisedan  
AttributeError  
insteadofa  
KeyError  
(  
GH17912  
)  
1.1.7.11  
Plotting  
Ł  
Bettererrormessagewhenattemptingtoplotbutmatplotlibisnotinstalled(  
GH19810  
).  
Ł  
DataFrame.plot()  
nowraisesa  
ValueError  
whenthe  
x  
or  
y  
argumentisimproperlyformed(  
GH18671  
)  
Ł  
Bugin  
DataFrame.plot()  
when  
x  
and  
y  
argumentsgivenaspositionscausedincorrectreferencedcolumns  
forline,barandareaplots(  
GH20056  
)  
Ł  
Buginformattingticklabelswith  
datetime.time()  
andfractionalseconds(  
GH18478  
).  
Ł  
Series.plot.kde()  
hasexposedtheargs  
ind  
and  
bw\_method  
inthedocstring(  
GH18461  
).Theargu-  
ment  
ind  
maynowalsobeaninteger(numberofsamplepoints).  
Ł  
DataFrame.plot()  
nowsupportsmultiplecolumnstothe  
y  
argument(  
GH19699  
)  
1.1.7.12  
Groupby/Resample/Rolling  
Ł  
Bugwhengroupingbyasinglecolumnandaggregatingwithaclasslike  
list  
or  
tuple  
(  
GH18079  
)  
Ł  
Fixedregressionin  
DataFrame.groupby()  
whichwouldnotemitanerrorwhencalledwithatuplekeynot  
intheindex(  
GH18798  
)  
Ł  
Bugin  
DataFrame.resample()  
whichsilentlyignoredunsupported(ormistyped)optionsfor  
label  
,  
closed  
and  
convention  
(  
GH19303  
)  
Ł  
Bugin  
DataFrame.groupby()  
wheretupleswereinterpretedaslistsofkeysratherthanaskeys(  
GH17979  
,  
GH18249  
)  
Ł  
Bugin  
DataFrame.groupby()  
whereaggregationby  
first  
/  
last  
/  
min  
/  
max  
wascausingtimestampsto  
loseprecision(  
GH19526  
)  
Ł  
Bugin  
DataFrame.transform()  
whereparticularaggregationfunctionswerebeingincorrectlycastto  
matchthedtype(s)ofthegroupeddata(  
GH19200  
)  
Ł  
Bugin  
DataFrame.groupby()  
passingthe  
on=  
kwarg,andsubsequentlyusing  
.apply()  
(  
GH17813  
)  
Ł  
Bugin  
DataFrame.resample().aggregate  
notraisinga  
KeyError  
whenaggregatinganon-existent  
column(  
GH16766  
,  
GH19566  
)  
40  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
DataFrameGroupBy.cumsum()  
and  
DataFrameGroupBy.cumprod()  
when  
skipna  
was  
passed(  
GH19806  
)  
Ł  
Bugin  
DataFrame.resample()  
thatdroppedtimezoneinformation(  
GH13238  
)  
Ł  
Bugin  
DataFrame.groupby()  
wheretransformationsusing  
np.all  
and  
np.any  
wereraisinga  
ValueError  
(  
GH20653  
)  
Ł  
Bugin  
DataFrame.resample()  
where  
ffill  
,  
bfill  
,  
pad  
,  
backfill  
,  
fillna  
,  
interpolate  
,  
and  
asfreq  
wereignoring  
loffset  
.(  
GH20744  
)  
Ł  
Bugin  
DataFrame.groupby()  
whenapplyingafunctionthathasmixeddatatypesandtheusersupplied  
functioncanfailonthegroupingcolumn(  
GH20949  
)  
Ł  
Bugin  
DataFrameGroupBy.rolling().apply()  
whereoperationsperformedagainsttheassociated  
DataFrameGroupBy  
objectcouldimpacttheinclusionofthegroupeditem(s)intheresult(  
GH14013  
)  
1.1.7.13  
Sparse  
Ł  
Buginwhichcreatinga  
SparseDataFrame  
fromadense  
Series  
oranunsupportedtyperaisedanuncon-  
trolledexception(  
GH19374  
)  
Ł  
Bugin  
SparseDataFrame.to\_csv  
causingexception(  
GH19384  
)  
Ł  
Bugin  
SparseSeries.memory\_usage  
whichcausedsegfaultbyaccessingnonsparseelements  
(  
GH19368  
)  
Ł  
Buginconstructinga  
SparseArray  
:if  
data  
isascalarand  
index  
isitwillcoerceto  
float64  
regardlessofscalar'sdtype.(  
GH19163  
)  
1.1.7.14  
Reshaping  
Ł  
Bugin  
DataFrame.merge()  
wherereferencinga  
CategoricalIndex  
byname,wherethe  
by  
kwarg  
would  
KeyError  
(  
GH20777  
)  
Ł  
Bugin  
DataFrame.stack()  
whichfailstryingtosortmixedtypelevelsunderPython3(  
GH18310  
)  
Ł  
Bugin  
DataFrame.unstack()  
whichcastsinttoif  
columns  
isa  
MultiIndex  
withunusedlevels  
(  
GH17845  
)  
Ł  
Bugin  
DataFrame.unstack()  
whichraisesanerrorif  
index  
isa  
MultiIndex  
withunusedlabelson  
theunstackedlevel(  
GH18562  
)  
Ł  
Fixedconstructionofa  
Series  
froma  
dict  
containing  
NaN  
askey(  
GH18480  
)  
Ł  
Fixedconstructionofa  
DataFrame  
froma  
dict  
containing  
NaN  
askey(  
GH18455  
)  
Ł  
Disabledconstructionofa  
Series  
wherelen(index)>len(data)=1,whichpreviouslywouldbroadcastthe  
dataitem,andnowraisesa  
ValueError  
(  
GH18819  
)  
Ł  
Suppressederrorintheconstructionofa  
DataFrame  
froma  
dict  
containingscalarvalueswhenthecorre-  
spondingkeysarenotincludedinthepassedindex(  
GH18600  
)  
Ł  
Fixed(changedfrom  
object  
to  
float64  
)dtypeof  
DataFrame  
initializedwithaxes,nodata,and  
dtype=int  
(  
GH19646  
)  
Ł  
Bugin  
Series.rank()  
where  
Series  
containing  
NaT  
the  
Series  
inplace(  
GH18521  
)  
Ł  
Bugin  
cut()  
whichfailswhenusingreadonlyarrays(  
GH18773  
)  
Ł  
Bugin  
DataFrame.pivot\_table()  
whichfailswhenthe  
aggfunc  
argisoftypestring.Thebehavioris  
nowconsistentwithothermethodslike  
agg  
and  
apply  
(  
GH18713  
)  
1.1.v0.23.0(May15,2017)  
41

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Ł  
Bugin  
DataFrame.merge()  
inwhichmergingusing  
Index  
objectsasvectorsraisedanException  
(  
GH19038  
)  
Ł  
Bugin  
DataFrame.stack()  
,  
DataFrame.unstack()  
,  
Series.unstack()  
whichwerenotreturn-  
ingsubclasses(  
GH15563  
)  
Ł  
Bugintimezonecomparisons,manifestingasaconversionoftheindextoUTCin  
.concat()  
(  
GH18523  
)  
Ł  
Bugin  
concat()  
whenconcattingsparseanddenseseriesitreturnsonlya  
SparseDataFrame  
.Shouldbe  
a  
DataFrame  
.(  
GH18914  
,  
GH18686  
,and  
GH16874  
)  
Ł  
Improvederrormessagefor  
DataFrame.merge()  
whenthereisnocommonmergekey(  
GH19427  
)  
Ł  
Bugin  
DataFrame.join()  
whichdoesan  
outer  
insteadofa  
left  
joinwhenbeingcalledwithmultiple  
DataFramesandsomehavenon-uniqueindices(  
GH19624  
)  
Ł  
Series.rename()  
nowaccepts  
axis  
asakwarg(  
GH18589  
)  
Ł  
Bugin  
rename()  
whereanIndexofsame-lengthtupleswasconvertedtoaMultiIndex(  
GH19497  
)  
Ł  
Comparisonsbetween  
Series  
and  
Index  
wouldreturna  
Series  
withanincorrectname,ignoringthe  
Index  
'snameattribute(  
GH19582  
)  
Ł  
Bugin  
qcut()  
wheredatetimeandtimedeltadatawith  
NaT  
presentraiseda  
ValueError  
(  
GH19768  
)  
Ł  
Bugin  
DataFrame.iterrows()  
,whichwouldinfersstringsnotcompliantto  
ISO8601  
todatetimes  
(  
GH19671  
)  
Ł  
Bugin  
Series  
constructorwith  
Categorical  
wherea  
ValueError  
isnotraisedwhenanindexofdif-  
ferentlengthisgiven(  
GH19342  
)  
Ł  
Bugin  
DataFrame.astype()  
wherecolumnmetadataislostwhenconvertingtocategoricaloradictionary  
ofdtypes(  
GH19920  
)  
Ł  
Bugin  
cut()  
and  
qcut()  
wheretimezoneinformationwasdropped(  
GH19872  
)  
Ł  
Bugin  
Series  
constructorwitha  
dtype=str  
,previouslyraisedinsomecases(  
GH19853  
)  
Ł  
Bugin  
get\_dummies()  
,and  
select\_dtypes()  
,whereduplicatecolumnnamescausedincorrectbehav-  
ior(  
GH20848  
)  
Ł  
Bugin  
isna()  
,whichcannothandleambiguoustypedlists(  
GH20675  
)  
Ł  
Bugin  
concat()  
whichraisesanerrorwhenconcatenatingTZ-awaredataframesandall-NaTdataframes  
(  
GH12396  
)  
Ł  
Bugin  
concat()  
whichraisesanerrorwhenconcatenatingemptyTZ-awareseries(  
GH18447  
)  
1.1.7.15  
Other  
Ł  
ImprovederrormessagewhenattemptingtouseaPythonkeywordasaniina  
numexpr  
backedquery  
(  
GH18221  
)  
Ł  
Buginaccessinga  
pandas.get\_option()  
,whichraised  
KeyError  
ratherthan  
OptionError  
when  
lookingupanon-existantoptionkeyinsomecases(  
GH19789  
)  
Ł  
Bugin  
testing.assert\_series\_equal()  
and  
testing.assert\_frame\_equal()  
forSeriesor  
DataFrameswithdifferingunicodedata(  
GH20503  
)  
42  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.2  
v0.22.0(December29,2017)  
Thisisamajorreleasefrom0.21.1andincludesasingle,API-breakingchange.Werecommendthatallusersupgrade  
tothisversionaftercarefullyreadingthereleasenote(singular!).  
1.2.1  
BackwardsincompatibleAPIchanges  
Pandas0.22.0changesthehandlingofemptyandall-  
NA  
sumsandproducts.Thesummaryisthat  
Ł  
Thesumofanemptyorall-  
NA  
Series  
isnow  
0  
Ł  
Theproductofanemptyorall-  
NA  
Series  
isnow  
1  
Ł  
We'veaddeda  
min\_count  
parameterto  
.sum()  
and  
.prod()  
controllingtheminimumnumberofvalid  
valuesfortheresulttobevalid.Iffewerthan  
min\_count  
non-  
NA  
valuesarepresent,theresultis  
NA  
.The  
defaultis  
0  
.Toreturn  
NaN  
,the0.21behavior,use  
min\_count=1  
.  
Somebackground:Inpandas0.21,weedalong-standinginconsistencyinthereturnvalueofall-  
NA  
seriesde-  
pendingonwhetherornotbottleneckwasinstalled.See  
Sum/Prodofall-NaNoremptySeries/DataFramesisnow  
consistentlyNaN  
.Atthesametime,wechangedthesumandprodofanempty  
Series  
toalsobe  
NaN  
.  
Basedonfeedback,we'vepartiallyrevertedthosechanges.  
1.2.1.1  
ArithmeticOperations  
Thedefaultsumforemptyorall-  
NA  
Series  
isnow  
0  
.  
pandas0.21.x  
In[1]:  
pd  
.  
Series([])  
.  
sum()  
Out[1]:  
nan  
In[2]:  
pd  
.  
Series([np  
.  
nan])  
.  
sum()  
Out[2]:  
nan  
pandas0.22.0  
In[1]:  
pd  
.  
Series([])  
.  
sum()  
Out[1]:  
0.0  
In[2]:  
pd  
.  
Series([np  
.  
nan])  
.  
sum()  
\\\\\\\\\\\\Out[2]:0.0  
Thedefaultbehavioristhesameaspandas0.20.3withbottleneckinstalled.ItalsomatchesthebehaviorofNumPy's  
np.nansum  
onemptyandall-  
NA  
arrays.  
Tohavethesumofanemptyseriesreturn  
NaN  
(thedefaultbehaviorofpandas0.20.3withoutbottleneck,orpandas  
0.21.x),usethe  
min\_count  
keyword.  
In[3]:  
pd  
.  
Series([])  
.  
sum(min\_count  
=  
1  
)  
Out[3]:  
nan  
Thankstothe  
skipna  
parameter,the  
.sum  
onanall-  
NA  
seriesisconceptuallythesameasthe  
.sum  
ofanempty  
onewith  
skipna=True  
(thedefault).  
In[4]:  
pd  
.  
Series([np  
.  
nan])  
.  
sum(min\_count  
=  
1  
)  
#skipna=Truebydefault  
Out[4]:  
nan  
1.2.v0.22.0(December29,2017)43

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
The  
min\_count  
parameterreferstotheminimumnumberof  
non-null  
valuesrequiredforanon-NAsumorproduct.  
Series.prod()  
hasbeenupdatedtobehavethesameas  
Series.sum()  
,returning  
1  
instead.  
In[5]:  
pd  
.  
Series([])  
.  
prod()  
Out[5]:  
1.0  
In[6]:  
pd  
.  
Series([np  
.  
nan])  
.  
prod()  
\\\\\\\\\\\\Out[6]:1.0  
In[7]:  
pd  
.  
Series([])  
.  
prod(min\_count  
=  
1  
)  
\\\\\\\\\\\\\\\\\\\\\\\\Out[7]:nan  
Thesechangesaffect  
DataFrame.sum()  
and  
DataFrame.prod()  
aswell.Finally,afewlessobviousplacesin  
pandasareaffectedbythischange.  
1.2.1.2  
GroupingbyaCategorical  
Groupingbya  
Categorical  
andsummingnowreturns  
0  
insteadof  
NaN  
forcategorieswithnoobservations.The  
productnowreturns  
1  
insteadof  
NaN  
.  
pandas0.21.x  
In[8]:  
grouper  
=  
pd  
.  
Categorical([  
˜  
a  
˜  
,  
˜  
a  
˜  
],categories  
=  
[  
˜  
a  
˜  
,  
˜  
b  
˜  
])  
In[9]:  
pd  
.  
Series([  
1  
,  
2  
])  
.  
groupby(grouper)  
.  
sum()  
Out[9]:  
a3.0  
bNaN  
dtype:float64  
pandas0.22  
In[8]:  
grouper  
=  
pd  
.  
Categorical([  
˜  
a  
˜  
,  
˜  
a  
˜  
],categories  
=  
[  
˜  
a  
˜  
,  
˜  
b  
˜  
])  
In[9]:  
pd  
.  
Series([  
1  
,  
2  
])  
.  
groupby(grouper)  
.  
sum()  
Out[9]:  
a3  
b0  
dtype:int64  
Torestorethe0.21behaviorofreturning  
NaN  
forunobservedgroups,use  
min\_count>=1  
.  
In[10]:  
pd  
.  
Series([  
1  
,  
2  
])  
.  
groupby(grouper)  
.  
sum(min\_count  
=  
1  
)  
Out[10]:  
a3.0  
bNaN  
dtype:float64  
1.2.1.3  
Resample  
Thesumandproductofall-  
NA  
binshaschangedfrom  
NaN  
to  
0  
forsumand  
1  
forproduct.  
pandas0.21.x  
44  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
In[11]:  
s  
=  
pd  
.  
Series([  
1  
,  
1  
,np  
.  
nan,np  
.  
nan],  
...:  
index  
=  
pd  
.  
date\_range(  
˜  
2017  
˜  
,periods  
=  
4  
))  
...:  
s  
Out[11]:  
2017-01-011.0  
2017-01-021.0  
2017-01-03NaN  
2017-01-04NaN  
Freq:D,dtype:float64  
In[12]:  
s  
.  
resample(  
˜  
2d  
˜  
)  
.  
sum()  
Out[12]:  
2017-01-012.0  
2017-01-03NaN  
Freq:2D,dtype:float64  
pandas0.22.0  
In[11]:  
s  
=  
pd  
.  
Series([  
1  
,  
1  
,np  
.  
nan,np  
.  
nan],  
....:  
index  
=  
pd  
.  
date\_range(  
˜  
2017  
˜  
,periods  
=  
4  
))  
....:  
In[12]:  
s  
.  
resample(  
˜  
2d  
˜  
)  
.  
sum()  
Out[12]:  
2017-01-012.0  
2017-01-030.0  
dtype:float64  
Torestorethe0.21behaviorofreturning  
NaN  
,use  
min\_count>=1  
.  
In[13]:  
s  
.  
resample(  
˜  
2d  
˜  
)  
.  
sum(min\_count  
=  
1  
)  
Out[13]:  
2017-01-012.0  
2017-01-03NaN  
dtype:float64  
Inparticular,upsamplingandtakingthesumorproductisaffected,asupsamplingintroducesmissingvaluesevenif  
theoriginalserieswasentirelyvalid.  
pandas0.21.x  
In[14]:  
idx  
=  
pd  
.  
DatetimeIndex([  
˜  
2017-01-01  
˜  
,  
˜  
2017-01-02  
˜  
])  
In[15]:  
pd  
.  
Series([  
1  
,  
2  
],index  
=  
idx)  
.  
resample(  
˜  
12H  
˜  
)  
.  
sum()  
Out[15]:  
2017-01-0100:00:001.0  
2017-01-0112:00:00NaN  
2017-01-0200:00:002.0  
Freq:12H,dtype:float64  
pandas0.22.0  
In[14]:  
idx  
=  
pd  
.  
DatetimeIndex([  
˜  
2017-01-01  
˜  
,  
˜  
2017-01-02  
˜  
])  
In[15]:  
pd  
.  
Series([  
1  
,  
2  
],index  
=  
idx)  
.  
resample(  
"  
12H  
"  
)  
.  
sum()  
Out[15]:  
2017-01-0100:00:001  
(continuesonnextpage)  
1.2.v0.22.0(December29,2017)45

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
(continuedfrompreviouspage)  
2017-01-0112:00:000  
2017-01-0200:00:002  
Freq:12H,dtype:int64  
Onceagain,the  
min\_count  
keywordisavailabletorestorethe0.21behavior.  
In[16]:  
pd  
.  
Series([  
1  
,  
2  
],index  
=  
idx)  
.  
resample(  
"  
12H  
"  
)  
.  
sum(min\_count  
=  
1  
)  
Out[16]:  
2017-01-0100:00:001.0  
2017-01-0112:00:00NaN  
2017-01-0200:00:002.0  
Freq:12H,dtype:float64  
1.2.1.4  
RollingandExpanding  
Rollingandexpandingalreadyhavea  
min\_periods  
keywordthatbehavessimilarto  
min\_count  
.Theonlycase  
thatchangesiswhendoingarollingorexpandingsumwith  
min\_periods=0  
.Previouslythisreturned  
NaN  
,when  
fewerthan  
min\_periods  
non-  
NA  
valueswereinthewindow.Nowitreturns  
0  
.  
pandas0.21.1  
In[17]:  
s  
=  
pd  
.  
Series([np  
.  
nan,np  
.  
nan])  
In[18]:  
s  
.  
rolling(  
2  
,min\_periods  
=  
0  
)  
.  
sum()  
Out[18]:  
0NaN  
1NaN  
dtype:float64  
pandas0.22.0  
In[17]:  
s  
=  
pd  
.  
Series([np  
.  
nan,np  
.  
nan])  
In[18]:  
s  
.  
rolling(  
2  
,min\_periods  
=  
0  
)  
.  
sum()  
Out[18]:  
00.0  
10.0  
dtype:float64  
Thedefaultbehaviorof  
min\_periods=None  
,implyingthat  
min\_periods  
equalsthewindowsize,isunchanged.  
1.2.2  
Compatibility  
Ifyoumaintainalibrarythatshouldworkacrosspandasversions,itmaybeeasiesttoexcludepandas0.21fromyour  
requirements.Otherwise,allyour  
sum()  
callswouldneedtocheckifthe  
Series  
isemptybeforesumming.  
Withsetuptools,inyour  
setup.py  
use:  
install\_requires  
=  
[  
˜  
pandas!=0.21.  
\*  
˜  
,  
.  
.  
.  
]  
Withconda,use  
requirements:  
run:  
-pandas!=0.21.0,!=0.21.1  
46  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
Notethattheinconsistencyinthereturnvalueforall-  
NA  
seriesisstillthereforpandas0.20.3andearlier.Avoiding  
pandas0.21willonlyhelpwiththeemptycase.  
1.3  
v0.21.1(December12,2017)  
Thisisaminorbreleaseinthe0.21.xseriesandincludessomesmallregressiones,bugesandperformance  
improvements.Werecommendthatallusersupgradetothisversion.  
Highlightsinclude:  
Ł  
Temporarilyrestorematplotlibdatetimeplottingfunctionality.Thisshouldresolveissuesforuserswhoimplic-  
itlyreliedonpandastoplotdatetimeswithmatplotlib.See  
here  
.  
Ł  
ImprovementstotheParquetIOfunctionsintroducedin0.21.0.See  
here  
.  
What'snewinv0.21.1  
Ł  
RestoreMatplotlibdatetimeConverterRegistration  
Ł  
Newfeatures  
Œ  
ImprovementstotheParquetIOfunctionality  
Œ  
OtherEnhancements  
Ł  
Deprecations  
Ł  
PerformanceImprovements  
Ł  
BugFixes  
Œ  
Conversion  
Œ  
Indexing  
Œ  
I/O  
Œ  
Plotting  
Œ  
Groupby/Resample/Rolling  
Œ  
Reshaping  
Œ  
Numeric  
Œ  
Categorical  
Œ  
String  
1.3.1  
RestoreMatplotlibdatetimeConverterRegistration  
Pandasimplementssomematplotlibconvertersfornicelyformattingtheaxislabelsonplotswith  
datetime  
or  
Period  
values.Priortopandas0.21.0,thesewereimplicitlyregisteredwithmatplotlib,asasideeffectof  
import  
pandas  
.  
Inpandas0.21.0,werequireduserstoexplicitlyregistertheconverter.Thiscausedproblemsforsomeuserswho  
reliedonthoseconvertersbeingpresentforregular  
matplotlib.pyplot  
plottingmethods,sowe'retemporarily  
revertingthatchange;pandas0.21.1againregisterstheconvertersonimport,justlikebefore0.21.0.  
1.3.v0.21.1(December12,2017)47

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
We'veaddedanewoptiontocontroltheconverters:  
pd.options.plotting.matplotlib.  
register\_converters  
.Bydefault,theyareregistered.Togglingthisto  
False  
removespandas'formattersand  
restoreanyconvertersweoverwrotewhenregisteringthem(  
GH18301  
).  
We'reworkingwiththematplotlibdeveloperstomakethiseasier.We'retryingtobalanceuserconvenience(auto-  
maticallyregisteringtheconverters)withimportperformanceandbestpractices(importingpandasshouldn'thavethe  
sideeffectofoverwritinganycustomconvertersyou'vealreadyset).Inthefuturewehopetohavemostofthedate-  
timeformattingfunctionalityinmatplotlib,withjustthepandas-sconvertersinpandas.We'llthengracefully  
deprecatetheautomaticregistrationofconvertersinfavorofusersexplicitlyregisteringthemwhentheywantthem.  
1.3.2  
Newfeatures  
1.3.2.1  
ImprovementstotheParquetIOfunctionality  
Ł  
DataFrame.to\_parquet()  
willnowwritenon-defaultindexeswhentheunderlyingenginesupportsit.  
Theindexeswillbepreservedwhenreadingbackinwith  
read\_parquet()  
(  
GH18581  
).  
Ł  
read\_parquet()  
nowallowstospecifythecolumnstoreadfromaparquet(  
GH18154  
)  
Ł  
read\_parquet()  
nowallowstospecifykwargswhicharepassedtotherespectiveengine(  
GH18216  
)  
1.3.2.2  
OtherEnhancements  
Ł  
Timestamp.timestamp()  
isnowavailableinPython2.7.(  
GH17329  
)  
Ł  
Grouper  
and  
TimeGrouper  
nowhaveafriendlyreproutput(  
GH18203  
).  
1.3.3  
Deprecations  
Ł  
pandas.tseries.register  
hasbeenrenamedto  
pandas.plotting.  
register\_matplotlib\_converters()  
(  
GH18301  
)  
1.3.4  
PerformanceImprovements  
Ł  
Improvedperformanceofplottinglargeseries/dataframes(  
GH18236  
).  
1.3.5  
BugFixes  
1.3.5.1  
Conversion  
Ł  
Bugin  
TimedeltaIndex  
subtractioncouldincorrectlyovwwhen  
NaT  
ispresent(  
GH17791  
)  
Ł  
Bugin  
DatetimeIndex  
subtractingdatetimelikefromDatetimeIndexcouldfailtoovw(  
GH18020  
)  
Ł  
Bugin  
IntervalIndex.copy()  
whencopyingand  
IntervalIndex  
withnon-default  
closed  
(  
GH18339  
)  
Ł  
Bugin  
DataFrame.to\_dict()  
wherecolumnsofdatetimethataretz-awarewerenotconvertedtorequired  
arrayswhenusedwith  
orient=˜records˜  
,raising  
TypeError  
(  
GH18372  
)  
Ł  
Bugin  
DateTimeIndex  
and  
date\_range()  
wheremismatchingtz-aware  
start  
and  
end  
timezones  
wouldnotraiseanerrif  
end.tzinfo  
isNone(  
GH18431  
)  
Ł  
Bugin  
Series.fillna()  
whichraisedwhenpassedalongintegeronPython2(  
GH18159  
).  
48  
Chapter1.What'sNew

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.3.5.2  
Indexing  
Ł  
Buginabooleancomparisonofa  
datetime.datetime  
anda  
datetime64[ns]  
dtypeSeries(  
GH17965  
)  
Ł  
Bugwherea  
MultiIndex  
withmorethanamillionrecordswasnotraising  
AttributeError  
whentrying  
toaccessamissingattribute(  
GH18165  
)  
Ł  
Bugin  
IntervalIndex  
constructorwhenalistofintervalsispassedwithnon-default  
closed  
(  
GH18334  
)  
Ł  
Bugin  
Index.putmask  
whenaninvalidmaskpassed(  
GH18368  
)  
Ł  
Buginmaskedassignmentofa  
timedelta64[ns]  
dtype  
Series  
,incorrectlycoercedto(  
GH18493  
)  
1.3.5.3  
I/O  
Ł  
Buginclass:  
~pandas.io.stata.StataReader  
notconvertingdate/timecolumnswithdisplayformattingaddressed  
(  
GH17990  
).Previouslycolumnswithdisplayformattingwerenormallyleftasordinalnumbersandnotcon-  
vertedtodatetimeobjects.  
Ł  
Bugin  
read\_csv()  
whenreadingacompressedUTF-16encoded(  
GH18071  
)  
Ł  
Bugin  
read\_csv()  
forhandlingnullvaluesinindexcolumnswhenspecifying  
na\_filter=False  
(  
GH5239  
)  
Ł  
Bugin  
read\_csv()  
whenreadingnumericcategorywithhighcardinality(  
GH18186  
)  
Ł  
Bugin  
DataFrame.to\_csv()  
whenthetablehad  
MultiIndex  
columns,andalistofstringswaspassed  
infor  
header  
(  
GH5539  
)  
Ł  
Buginparsingintegerdatetime-likecolumnswithformatin  
read\_sql  
(  
GH17855  
).  
Ł  
Bugin  
DataFrame.to\_msgpack()  
whenserializingdataofthe  
numpy.bool\_  
datatype(  
GH18390  
)  
Ł  
Bugin  
read\_json()  
notdecodingwhenreadinglinedelimintedJSONfromS3(  
GH17200  
)  
Ł  
Bugin  
pandas.io.json.json\_normalize()  
toavoidof  
meta  
(  
GH18610  
)  
Ł  
Bugin  
to\_latex()  
whererepeatedmulti-indexvalueswerenotprintedeventhoughahigherlevelindex  
differedfromthepreviousrow(  
GH14484  
)  
Ł  
BugwhenreadingNaN-onlycategoricalcolumnsin  
HDFStore  
(  
GH18413  
)  
Ł  
Bugin  
DataFrame.to\_latex()  
with  
longtable=True  
wherealatexmulticolumnalwaysspannedover  
threecolumns(  
GH17959  
)  
1.3.5.4  
Plotting  
Ł  
Bugin  
DataFrame.plot()  
and  
Series.plot()  
with  
DatetimeIndex  
whereageneratedby  
themisnotpickleableinPython3(  
GH18439  
)  
1.3.5.5  
Groupby/Resample/Rolling  
Ł  
Bugin  
DataFrame.resample(...).apply(...)  
whenthereisacallablethatreturnsdifferentcolumns  
(  
GH15169  
)  
Ł  
Bugin  
DataFrame.resample(...)  
whenthereisatimechange(DST)andresamplingfrequecyis12hor  
higher(  
GH15549  
)  
Ł  
Bugin  
pd.DataFrameGroupBy.count()  
whencountingoveradatetimelikecolumn(  
GH13393  
)  
Ł  
Bugin  
rolling.var  
wherecalculationisinaccuratewithazero-valuedarray(  
GH18430  
)  
1.3.v0.21.1(December12,2017)49

pandas:powerfulPythondataanalysistoolkit,Release0.23.0  
1.3.5.6  
Reshaping  
Ł  
Errormessagein  
pd.merge\_asof()  
forkeydatatypemismatchnowincludesdatatypeofleftandrightkey  
(  
GH18068  
)  
Ł  
Bugin  
pd.concat  
whenemptyandnon-emptyDataFramesorSeriesareconcatenated(  
GH18178  
GH18187  
)  
Ł  
Bugin  
DataFrame.filter(...)  
when  
unicode  
ispassedasaconditioninPython2(  
GH13101  
)  
Ł  
BugwhenmergingemptyDataFrameswhen  
np.seterr(divide=˜raise˜)  
isset(  
GH17776  
)  
1.3.5.7  
Numeric  
Ł  
Bugin  
pd.Series.rolling.skew()  
and  
rolling.kurt()  
withallequalvalueshasissue  
(  
GH18044  
)  
1.3.5.8  
Categorical  
Ł  
Bugin  
DataFrame.astype()  
wherecastingto`category'onanempty  
DataFrame  
causesasegmentation  
fault(  
GH18004  
)  
Ł  
Errormessagesinthetestingmodulehavebeenimprovedwhenitemshavedifferent  
CategoricalDtype  
(  
GH18069  
)  
Ł  
CategoricalIndex  
cannowcorrectlytakea  
pd.api.types.CategoricalDtype  
asitsdtype  
(  
GH18116  
)  
Ł  
Bugin  
Categorical.unique()  
returningread-only  
codes  
arraywhenallcategorieswere  
NaN  
(  
GH18051  
)  
Ł  
Bugin  
DataFrame.groupby(axis=1)  
witha  
CategoricalIndex  
(  
GH18432  
)  
1.3.5.9  
String  
Ł  
Series.str.split()  
willnowpropagate  
NaN  
valuesacrossallexpandedcolumnsinsteadof  
None  
(  
GH18450  
)  
1.4  
v0.21.0(October27,2017)  
Thisisamajorreleasefrom0.20.3andincludesanumberofAPIchanges,deprecations,newfeatures,enhancements,  
andperformanceimprovementsalongwithalargenumberofbuges.Werecommendthatallusersupgradetothis  
version.  
Highlightsinclude:  
Ł  
Integrationwith  
ApacheParquet  
,includinganewtop-level  
read\_parquet()  
functionand  
DataFrame.  
to\_parquet()  
method,see  
here  
.  
Ł  
Newuser-facing  
pandas.api.types.CategoricalDtype  
forspecifyingcategoricalsindependentof  
thedata,see  
here  
.  
Ł  
Thebehaviorof  
sum  
and  
prod  
onall-NaNSeries/DataFramesisnowconsistentandnolongerdependson  
whether  
bottleneck  
isinstalled,and  
sum  
and  
prod  
onemptySeriesnowreturnNaNinsteadof0,see  
here  
.  
Ł  
Compatibilityesforpypy,see  
here  
.  
Ł  
Additionstothe  
drop  
,  
reindex  
and  
rename  
APItomakethemmoreconsistent,see  
here  
.  
50  
Chapter1.What'sNew