

Pandas 模块速查表

数据输入 / 输出	
read_pickle()	读取 pickle 文件
read_table()	将带分隔符的常规文件读入 DataFrame 对象
read_csv()	将 CSV(逗号分隔) 文件读入 DataFrame 对象
read_fwf()	将固定宽度的格式化行表读入 DataFrame 对象
read_clipboard()	从剪贴板读取文本并传递到 read_table
read_excel()	将一个 Excel 表读入 DataFrame
ExcelFile.parse()	将 Excel 表读入 DataFrame
read_json()	将 JSON 字符串转换为 Pandas 对象
read_html()	将 HTML 表读入 DataFrame 对象
read_hdf()	读取 HDF5 文件
HDFStore.put()	将对象存储在 HDFStore 中
HDFStore.append()	附加到文件中的表。节点必须是已经存在的表
HDFStore.get()	检索存储在文件中的 Pandas 对象
HDFStore.select()	检索存储在文件中的 Pandas 对象，并根据位置进行选择
read_sql_table()	将 SQL 数据库中的表读入 DataFrame 对象
read_sql_query()	将 SQL 查询读入 DataFrame 对象
read_sql()	将 SQL 查询或数据库中的表读入 DataFrame 对象
read_gbq()	从谷歌 BigQuery 表（Google 推出的一项 Web 服务）加载数据
to_gbq()	向谷歌 BigQuery 表写入 DataFrame 对象
read_stata()	将 Stata（统计学软件）文件读入 DataFrame 对象
StataReader.data()	从 Stata 文件中读取观察结果，并将其转换为数据流
StataReader.data_label()	返回 Stata 文件的数据标签
StataReader.value_labels()	返回一个关联每个变量名的字典
StataReader.variable_labels()	以字典形式返回变量标签
StataWriter.write_file()	写入 Stata 文件
处理日期时间	
to_datetime()	将数据转换为日期时间格式
to_timedelta()	计算两个日期数据之间的时间差
date_range()	生成指定频率的日期时间索引，默认是 day（日历）
bdate_range	生成一个固定频率的日期时间索引
period_range()	根据指定频率创建日期时间范围

数据操作	
melt()	将数据从宽表转换为长表
pivot()	根据 DataFrame 数据中的 3 列生成透视表
pivot_table()	数据透视表
crosstab()	交叉表是用于统计分组频率的特殊透视表
cut()	数据面元化（即将数据按照一定的区间进行分割）
qcut()	把一组数字按大小区间进行分割
merge()	按列名相同的列合并 DataFrame 数据
concat()	根据不同方式合并 DataFrame 数据
get_dummies()	实现 one-hot 编码（例如，性别男、女转换为 0、1）
factorize()	当有多个变量出现时，将输入值编码为枚举类型或分类变量
缺失数据	
isnull()	检测缺失数据（数据是否为空）
notnull()	检测缺失数据（数据是否不为空）
移动窗口功能	
rolling_count()	在提供的窗口内对非 NaN 观测值进行移动计数
rolling_sum()	移动窗口数据的和
rolling_mean()	移动窗口数据的均值
rolling_median()	移动窗口数据的中位数
rolling_var()	移动窗口数据的方差
rolling_std()	移动窗口数据的标准差
rolling_min()	移动窗口数据的最小值
rolling_max()	移动窗口数据的最大值
rolling_corr()	移动窗口数据的相关系数
rolling_corr_pairwise()	配对数据的相关系数
rolling_cov()	移动窗口数据的协方差
rolling_skew()	移动窗口数据的偏度
rolling_kurt()	移动窗口数据的峰度
rolling_apply()	对移动窗口数据应用数组函数
rolling_quantile()	移动窗口数据的分位数
rolling_window()	移动窗口

Pandas 模块速查表——Series 对象

属性和底层数据

index()	行 / 列标签
values()	返回序列数组
dtype()	返回序列的数据类型

数据转换

astype()	对序列数据类型强制转换
copy()	复制序列数据
isnull()	以布尔值返回空值序列
notnull()	以布尔值返回非空值序列

字符串处理

str.cat()	用给定的分隔符连接字符串数组
str.center()	居中，用额外的空格填充左右两边
str.contains()	检查给定的模式是否包含在数组的每个字符串中
str.count()	计算每个字符出现的次数
str.decode()	使用指定的编码将字符串解码为 unicode 编码格式
str.encode()	使用指定的编码将字符串编码为其他编码格式
str.endswith()/str.startswith()	返回布尔值，是否以指定的子字符串结尾 / 开头
str.extract()	使用传递的正则表达式在每个字符串中查找组
str.findall()	查找所有出现的模式或正则表达式
str.get()	从数组中每个元素的列表、元组或字符串中提取元素
str.join()	拼接字符串，类似于字符串函数 String 中的 join()
str.len()	计算数组中每个字符串的长度
str.lower()/str.upper()	将数组中的字符串转换为小写 / 大写字母
str.lstrip()/str.rstrip()	去除字符串左边 / 右边的空格
str.match()	使用传递的正则表达式在每个字符串中查找组
str.pad()	带空格的填充字符串
str.repeat()	按指定的次数复制数组中的每个字符串
str.replace()	替换字符串
str.slice()	按下标截取字符串
str.slice_replace()	按下标替换字符串
str.split()	分割字符串
str.strip()	删除数组中每个字符串中的空格（包括换行符）
str.title()	将字符串转换为带标题的版本
str.get_dummies()	拆分 Series 中以 “ ” 分隔的字符串，然后返回一个 DataFrame 对象

Pandas 模块速查表——DataFrame 对象

构造函数	
DataFrame()	创建表格数据结构，带有标记的轴（行和列）

属性和底层数据	
index	行标签
columns	列标签
as_matrix()	将 DataFrame 转换为数字数组矩阵形式
dtypes	返回 DataFrame 的数据类型
ftypes	返回每一列的数据类型
get_dtype_counts()	返回数据类型的数量
get_ftype_counts()	返回每一列的数据类型的数量
values	返回数组
ndim	返回 DataFrame 的维度
shape	返回 DataFrame 的形状
size	返回 DataFrame 元素的个数

类型转换	
astype()	对 DataFrame 数据类型强制转换
copy()	深度复制数据
isnull()	以布尔值返回空值数据
notnull()	以布尔值返回非空值数据

数据选择 / 迭代 / 筛选	
head()/tail()	返回前 n 行数据 / 返回最后 n 行数据
at 属性	以行名和列名获取单个数据
iat 属性	以行索引获取单个数据
loc 属性	以行名和列名获取多个数据
iloc 属性	以行列索引获取多个数据
insert()	在指定位置插入一列数据
iteritems()	返回列名和序列的迭代器
iterrows()	返回索引和序列的迭代器
itertuples()	以元组的形式迭代 DataFrame 行，并带有索引值
isin()	返回符合条件的 DataFrame。例如：A 列包含 8，df[df['A'].isin([8])]
filter()	通过指定条件筛选数据
where()	通过指定条件筛选数据

索引设置	
idxmax()	最大值的索引
idxmin()	最小值的索引
lookup()	基于标签的“花式索引”功能的 DataFrame
reindex()	重新设置索引
reset_index()	常用于数据清洗后，对数据重新设置连续行索引
set_index()	设置索引

GroupBy 分组与函数应用	
apply()	应用函数
applymap()	将函数应用于要操作的 DataFrame
groupby()	数据分组

数据计算与描述性统计	
abs()	绝对值
corr()	相关系数
corrwith()	行或列两两相关
count()	计算非空值或空值数量
cov()	协方差
cummax()	累积最大值
cummin()	累积最小值
cumprod()	累乘，主要用于观察数据的变化趋势
cumsum()	累加
describe()	描述性统计
diff()	数据移位（一阶差分操作）
eval()	在调用 DataFrame 的上下文中对表达式求值
kurt()	偏度
mad()	平均绝对偏差
max()/min()	最大值 / 最小值
mean()	平均值
median()	中位数
mode()	众数
pct_change()	百分比
prod()	指定轴的值的乘积
quantile()	分位数
rank()	数据排名
skew()	峰度
sum()	求和
std()	标准差
var()	方差

标签操作 / 数据删除操作	
rename()	修改行列名称
drop()	删除数据返回新的 DataFrame 对象
drop_duplicates()	删除重复行的 DataFrame 数据
pop()	返回删除的项目

缺失数据处理	
dropna()	缺失值删除
fillna()	缺失值填充
replace()	数据替换

重塑 / 排序 / 转换	
pivot()	重塑数据（数据透视表）
reorder_levels()	重新排列索引级别
sort()	数据排序
sort_index()	根据索引排序
sortlevel()	根据选择的轴和索引进行多级别排序
stack()	列索引转换成最内层的行索引
swaplevel()	更改索引的层级
T 属性	行列数据转换
unstack()	最内层的行索引转换成列索引

数据增加与合并	
align()	将轴上的两个对象与每个轴索引的指定连接方法连接
append()	增加数据
merge()	按列名相同的列合并 DataFrame 数据
concat()	根据不同方式合并 DataFrame 数据
transpose()	转置索引和列
update()	通过传入非空值修改数据

日期处理与时间序列	
asfreq()	将指定的日期数据转换为一定频率的数据
first()	基于日期偏移量来设置时间序列数据的初始时段的便捷方法
last()	基于日期偏移量来划分时间 series 数据的最终时段的便捷方法
resample()	数据重采样
shift()	用一个可选的时间频率按所需的周期数移动索引
to_period()	将时间戳转换为时期
to_timestamp()	将时期转换为时间戳
tz_convert()	将轴转换为目标时区
tz_localize()	时区定位

DataFrame 绘图	
boxplot()	绘制箱形图
hist()	绘制直方图
plot()	绘制折线图、柱形图、箱线图、密度图、饼图等
plot.area()	面积图
plot.bar()	垂直条形图
plot.barh()	水平条形图
plot.box()	箱形图
plot.density()	核密度
plot.hexbin()	热力图
plot.hist()	直方图
plot.kde()	核密度图
plot.line()	折线图
plot.pie()	饼形图
plot.scatter()	散点图

数据输入输出	
from_csv()	输出 csv 文件，使存储的数据不变
from_dict()	通过字典创建 DataFrame 数据
from_items()	将（键、值）对转换为数据格式。其中，键就是轴
from_records()	将结构化或记录数组转换为 DataFrame 数据
to_pickle()	生成 pickle 文件对数据进行永久储存
to_csv()	输出数据为 csv 文件
to_hdf()	输出数据为 hdf 文件（可以存储不同类型的图像和数码数据的文件格式）
to_sql()	将 DataFrame 数据写入 SQL 数据库
to_dict()	将 DataFrame 数据转换为字典
to_excel()	输出数据为 Excel 文件 s
to_json()	将 DataFrame 数据转换为 JSON 字符串
to_html()	将数据输出为 HTML 网页格式
to_stata()	将 DataFrame 数据输出为 Stata（统计学软件）文件
to_gbq()	向谷歌 BigQuery 表写入一个 DataFrame
to_records()	将数据格式转换为记录数组

Matplotlib 模块速查表——图表绘制专属函数

A

acorr()	绘制 x 的自相关图
angle_spectrum()	绘制角度谱

B

bar()	绘制柱形图
barh()	绘制水平条形图
boxplot()	绘制箱形图
broken_barh()	绘制画水平断条图

C

contour()	绘制等高线图（不填充）
contourf()	绘制等高线图（填充）
csd()	绘制交叉谱密度图

F

fill()	绘制填充多边形
---------------	---------

H

hexbin()	绘制二维六角形多色柱状图
hist()	绘制直方图

M

magnitude_spectrum()	绘制震级谱
-----------------------------	-------

P

pcolor()	绘制二维阵列的伪彩色图
pcolormesh	绘制四边形网格
phase_spectrum	绘制相位谱
pie()	绘制饼形图
plot()	绘制折线图
plot_date()	绘制时间序列图
polar()	绘制极坐标图
psd()	绘制功率谱密度图

S

scatter()	绘制散点图
spectrogram()	绘制声谱图
stackplot()	绘制堆叠面积图
step()	绘制阶梯图
subplot()/	绘制子图表
subplots()	绘制子图表

T

table()	绘制表格
----------------	------

V

violinplot()	绘制小提琴图
---------------------	--------

Matplotlib 模块速查表——图表设置函数

A

annotate()	注释
arrow()	在坐标轴上添加一个箭头
axex()	添加一个轴
axis()	获取或设置axis属性的方法

B

box()	打开/关闭“坐标轴”框
--------------	-------------

C

cla()	清除当前轴
clabel()	等高线图标签
clf()	清除当前画布
colorbar()	在绘图中添加一个颜色条

D

draw()	重新绘制当前图形
---------------	----------

E

errorbar()	绘制误差线
-------------------	-------

F

figtext()	向图中添加文本
figure()	创建新的画布

G

grid()	为图表设置网格线
---------------	----------

H

hlines()	绘制水平线（横线）
-----------------	-----------

I

imread()	读取图像
imsave()	保存图像
imshow()	显示图像

L

legend()	图例
-----------------	----

S

savefig()	保存当前画布（图表）
show()	显示所有画布（图表）
subplots_adjust()	图表与画布边缘间距
suptitle()	子图标题

T

text()	文本标签
title()	图表标题
twinx()	共享x轴
twiny()	共享y轴

V

vlines()	绘制垂直线（竖线）
-----------------	-----------

X

xcorr()	绘制相关性
xlabel()/ylabel()	x轴标题/y轴标题
xlim()/ylim()	x轴/y轴的坐标范围
xscale()/yscale()	x轴/y轴刻度
xticks()/yticks()	x轴/y轴刻度位置与标签

图表颜色值速查表

颜色名称	颜色说明	颜色代码	RGB颜色值	实色效果
LightPink	浅粉色	#FFB6C1	255,182,193	
Pink	粉红	#FFC0CB	255,192,203	
Crimson	深红色	#DC143C	220,20,60	
LavenderBlush	淡紫色	#FFF0F5	255,240,245	
PaleVioletRed	浅紫红\苍紫罗蓝色	#DB7093	219,112,147	
HotPink	桃红色\艳粉色\亮粉	#FF69B4	255,105,180	
DeepPink	深粉色	#FF1493	255,20,147	
MediumVioletRed	适中的紫罗兰红色	#C71585	199,21,133	
Orchid	兰花 淡紫色	#DA70D6	218,112,214	
Thistle	蓟 苍紫色	#D8BFD8	216,191,216	
plum	李子 紫红色	#DDA0DD	221,160,221	
Violet	紫罗兰色	#EE82EE	238,130,238	
Magenta	洋红	#FF00FF	255,0,255	
DarkMagenta	深洋红色	#8B008B	139,0,139	
Purple	紫色	#800080	128,0,128	
MediumOrchid	中紫色	#BA55D3	186,85,211	
DarkVoilet	深紫罗兰色	#9400D3	148,0,211	
DarkOrchid	暗紫色	#9932CC	153,50,204	
Indigo	靛蓝色	#4B0082	75,0,130	
BlueViolet	蓝紫色	#8A2BE2	138,43,226	
MediumPurple	中紫色	#9370DB	147,112,219	
MediumSlateBlue	中石板蓝色	#7B68EE	123,104,238	
SlateBlue	石蓝色	#6A5ACD	106,90,205	
DarkSlateBlue	暗灰蓝色	#483D8B	72,61,139	
Lavender	薰衣草 淡紫色	#E6E6FA	230,230,250	
GhostWhite	幽灵白色	#F8F8FF	248,248,255	
Blue	纯蓝色	#0000FF	0,0,255	
MediumBlue	中蓝色	#0000CD	0,0,205	
MidnightBlue	午夜蓝色	#191970	25,25,112	
DarkBlue	深蓝色	#00008B	0,0,139	
Navy	海军 深蓝色 藏青色	#000080	0,0,128	
RoyalBlue	宝蓝色	#4169E1	65,105,225	
CornflowerBlue	矢菊花蓝 浅蓝色	#6495ED	100,149,237	
LightSteelBlue	淡钢蓝色	#B0C4DE	176,196,222	
LightSlateGray	浅蓝灰色	#778899	119,136,153	
SlateGray	石板灰 灰石色	#708090	112,128,144	
DoderBlue	道奇蓝色	#1E90FF	30,144,255	
AliceBlue	爱丽丝蓝色	#F0F8FF	240,248,255	
SteelBlue	钢蓝色 铁青色	#4682B4	70,130,180	
LightSkyBlue	浅天蓝色	#87CEFA	135,206,250	
SkyBlue	天蓝色	#87CEEB	135,206,235	
DeepSkyBlue	深天蓝色	#00BFFF	0,191,255	
LightBLue	浅蓝色	#ADD8E6	173,216,230	
PowDerBlue	火药青 粉蓝色	#B0E0E6	176,224,230	

颜色名称	颜色说明	颜色代码	RGB颜色值	实色效果
CadetBlue	军蓝色	#5F9EA0	95,158,160	
Azure	蔚蓝色	#00FFFF	240,255,255	
LightCyan	淡青色	#E1FFFF	225,255,255	
PaleTurquoise	苍白的宝石绿色	#AFEEEE	175,238,238	
Cyan	蓝绿色	#00FFFF	0,255,255	
DarkTurquoise	深宝石绿色	#00CED1	0,206,209	
DarkSlateGray	墨绿色	#2F4F4F	47,79,79	
DarkCyan	深青色	#008B8B	0,139,139	
Teal	水鸭色 青色 蓝绿色	#008080	0,128,128	
MediumTurquoise	中宝石绿色	#48D1CC	72,209,204	
LightSeaGreen	浅海蓝色 淡绿	#20B2AA	32,178,170	
Turquoise	绿松石 蓝绿色	#40E0D0	64,224,208	
MediumAquamarine	中碧绿色	#00FA9A	0,250,154	
MediumSpringGreen	中亮绿色 春天的绿色	#5FFFA	245,255,250	
MintCream	薄荷奶油色	#00FF7F	0,255,127	
SpringGreen	春天的绿色 春绿色	#3CB371	60,179,113	
SeaGreen	海藻绿 海绿色	#2E8B57	46,139,87	
Honeydew	蜜瓜 蜜色	#00FF0	240,255,0	
LightGreen	淡绿色	#90EE90	144,238,144	
PaleGreen	苍绿色	#98FB98	152,251,152	
DarkSeaGreen	深绿色 青绿色	#8FBC8F	143,188,143	
LimeGreen	橙绿色	#32CD32	50,205,50	
Lime	绿黄色	#00FF0	0,255,0	
ForestGreen	森林绿 葱绿色	#228B22	34,139,34	
Green	纯绿色	#008000	0,128,0	
DarkGreen	深绿色 暗绿色	#006400	0,100,0	
Chartreuse	淡黄绿色	#7FFF0	127,255,0	
LawnGreen	草绿色	#7CFC00	124,252,0	
GreenYellow	黄绿色	#ADFF2F	173,255,47	
OliveDrab	橄榄土褐色	#556B2F	85,107,47	
Beige	米色	#F5F5DC	245,245,220	
LightGoldenrodYellow	浅金黄色	#FAFAD2	250,250,210	
Ivory	象牙色 乳白色	#FFFFFF0	255,255,240	
LightYellow	浅黄色 鹅黄	#FFFFE0	255,255,224	
Yellow	纯黄色	#FFFF00	255,255,0	
Olive	橄榄色	#808000	128,128,0	
DarkKhaki	暗卡其色	#BDB76B	189,183,107	
LemonChiffon	柠檬纱色	#FFFACD	255,250,205	
PaleGodenrod	灰秋麒麟色	#EEE8AA	238,232,170	
Khaki	卡其色	#F0E68C	240,230,140	
Gold	金色	#FFD700	255,215,0	
CornisIk	玉米色	#FFF8DC	255,248,220	
GoldEnrod	秋麒麟色	#DAA520	218,165,32	
OldLace	浅米色	#FDF5E6	253,245,230	

颜色名称	颜色说明	颜色代码	RGB颜色值	实色效果
Wheat	小麦色	#F5DEB3	245,222,179	
Moccasin	鹿皮色	#FFE4B5	255,228,181	
Orange	橙色	#FFA500	255,165,0	
PapayaWhip	番木瓜色	#FFefd5	255,239,213	
BlanchedAlmond	白杏色	#FFEBCD	255,235,205	
NavajoWhite	纳瓦白	#FFDEAD	255,222,173	
AntiqueWhite	古董白色	#FAEBD7	250,235,215	
Tan	棕褐色 黝黑色	#D2B48C	210,180,140	
BrulyWood	结实的树	#DEB887	222,184,135	
Bisque	橘黄色	#FFE4C4	255,228,196	
DarkOrange	深橙色	#FF8C00	255,140,0	
Linen	亚麻色	#FAF0E6	250,240,230	
Peru	秘鲁	#CD853F	205,133,63	
PeachPuff	桃色	#FFDAB9	255,218,185	
SandyBrown	黄褐色 沙棕色	#FA4A60	244,164,96	
Chocolate	巧克力色	#D2691E	210,105,30	
SaddleBrown	马鞍棕色 重褐色	#8B4513	139,69,19	
SeaShell	海贝壳色	#FFF5EE	255,245,238	
Sienna	黄土赭色	#A0522D	160,82,45	
LightSalmon	浅肉色	#FFA07A	255,160,122	
Coral	珊瑚	#FF7F50	255,127,80	
OrangeRed	橙红色	#FF4500	255,69,0	
DarkSalmon	深肉色	#E9967A	233,150,122	
Tomato	番茄色	#FF6347	255,99,71	
MistyRose	浅玫瑰色	#FFE4E1	255,228,225	
Salmon	浅橙红色	#FA8072	250,128,114	
Snow	雪白色	#FFFAFA	255,250,250	
LightCoral	浅珊瑚色	#F08080	240,128,128	
RosyBrown	玫瑰棕色	#BC8F8F	188,143,143	
IndianRed	印度红色	#CD5C5C	205,92,92	
Red	纯红色	#FF0000	255,0,0	
Brown	棕色	#A52A2A	165,42,42	
FireBrick	砖红色	#B22222	178,34,34	
DarkRed	深红色	#8B0000	139,0,0	
Maroon	栗色	#800000	128,0,0	
White	纯白色	#FFFFFF	255,255,255	
WhiteSmoke	烟白色	#F5F5F5	245,245,245	
Gainsboro	亮灰色	#DCDCDC	220,220,220	
LightGray	浅灰色	#D3D3D3	211,211,211	
Silver	银灰色	#C0C0C0	192,192,192	
DarkGray	深灰色	#A9A9A9	169,169,169	
Gray	灰色	#808080	128,128,128	
DimGray	暗灰色	#696969	105,105,105	
Black	纯黑色	#000000	0,0,0	