

索引 使用 label 使用位置
~~.at~~ ~~.iat~~ ~~.loc~~ ~~.iloc~~ ~~.ix~~

No. _____

Date _____

DataFrame(data, index=xx, columns=xx)

Series(data, index=xx)

{	Series	1维	数据类型相同
	DataFrame	2维	大小可变, 每一列的数据类型可以不同
	Panel	3维	大小可变

属性:

dtypes, index, columns, values, T.

方法:

head(), tail()

describe()

sort_index(axis=0, ascending=True)

sort_values(by=xx, axis=0, ascending=True)

copy()

dropna(how='any') 丢弃所有缺失数据的行

缺失数据

fillna(value=5) 填充缺失数据

mean()

apply(f)

value_counts() Series专用?

hist()

Series专用? 绘制直方图



使用label索引 ^{index}1. `df.loc[dates[0]]` 单行数据

2. 多列数据

`df.loc[:, ['A', 'B']]`

3. 多行多列数据

`df.loc['20130102': '20130104', ['A', 'B']]`

4. 得到标量值:

`df.loc[dates[0], 'B']`☺ `df.at[dates[0], 'B']` 速度更快

使用位置索引

1. `df.iloc[3]` 单行数据

2. 多行多列数据

`df.iloc[0:2, 0:2]`

3. 索引不连续部分:

`df.iloc[[1, 2, 4], [0, 2]]`

4. 索引整行:

`df.iloc[0:2, :]`

5. 索引整列:

`df.iloc[:, 1:3]`

6. 标量值

`df.iloc[0, 0]`☺ `df.iat[0, 0]`

Date

布尔型索引:

1. `df[df.A > 0]`

2. `df[df > 0]`

3. `df2[df2['E'].isin(['two', 'four'])]`

