Jan Popko

Python Advanced

Business Trends Academy (BTA) GmbH



Nestorstraße 3 D-10709 Berlin

> Tel.: +49 (0) 30 894 087 57 Fax: +49 (0) 30 895 429 94

Geschattsturrer: Gabriele Feischmann-Hahn Maxi-Marlen Fleischmann Hauptsilz des Url drenhmens: Nestorstraße 36, D-10709 Berlin HRB 115251 B / Amtsgericht Berlin Charlottenburg

Vektorisierte String Operationen:

Mit Hilfe der str funktionen kann man sehr einfach Texte in pandas Series Objekten bearbeiten:

```
s = pd.Series(['A', 'B', 'C'])
s = s.str.lower()

print(s)

0    a
1    b
2    c
dtype: object
Jan Popko
```

Funktionen auf DataFrames:

```
df.apply(func, axis= 0)
```

- axis = 0: func wird Zeilenweise angewandt
- axis = 1: func wird Spaltenweise angewandt

```
df.aggregate([func1, func2], axis = 0)
```

- führt eine beliebige Anzahl an Funktionen über eine Bestimmte Achse aus Auch möglich nur auf bestimmte Spalten mit Dictionary:

```
df.agg({'column1' : [func1, func2], 'column2' : [func3, func4]})
```

Jan Popko

Zeilen und Spalten können mehr als nur einen Index haben.

MultiIndex Konstruktoren

```
i = pd.MultiIndex.from_arrays([['a', 'a', 'b', 'b'], [1, 2, 1, 2]])
i = pd.MultiIndex.from_tuples([('a', 1), ('a', 2), ('b', 1), ('b', 2)])
i = pd.MultiIndex.from_product([['a', 'b'], [1, 2]])
i = pd.MultiIndex(levels=[['a', 'b'], [1, 2]], codes=[[0, 0, 1, 1], [0, 1, 0, 1]])
```

Erstellen des DataFrames:

```
data = np.arange(16).reshape(4,4)
 Als Zeilenindex:
                  pd.DataFrame(data, index=i)
 Als Spaltenindex: pd.DataFrame(data, columns=i)
Jan Popko
```

Methoden für MultiIndexing

reindex(index) – erstellt eine Kopie des DataFrames mit neuem Index

stack() - füge ein Spaltenindexlevel zum Index hinzu (default = -1)

unstack() - entferne eine Ebene aus einem MultiIndex und füge sie als Spalte an (default = -1)

index.names – man kann den Indices Namen geben

reset index() - entfernt MultiIndexing (oder bestimmte Indices) und stellt Default her

set_index() - setzt den Index eines DataFrames und nutzt dafür vorhandene Spalten

df.index.get_level_values(0) – gibt die Werte für den Zeilenindex im angegebenen Level an df.columns.get_level_values(0) - gibt die Werte für den Spaltenindex im angegebenen Level an

Jan Popko

bto a