

Convert Data Structures to Pandas Dataframes

Fan Wang

2023-02-16

Contents

1	Convert Data Structures to Panda Tables	1
1.1	Convert between a Nested Dictionary and a Pandas Dataframe	1

1 Convert Data Structures to Panda Tables

Go to the [RMD](#), [PDF](#), or [HTML](#) version of this file. Go back to [Python Code Examples](#) Repository ([bookdown site](#)) or the [pyfan](#) Package ([API](#)).

```
import pprint
import pandas as pd
```

1.1 Convert between a Nested Dictionary and a Pandas Dataframe

We have a doubly nested dictionary. The top layer has integer keys, and the values are dictionaries. The second layer has string keys, and the values are numeric.

First, we construct the nested-dictionary that we are interested in converting.

```
# The nested dictionary
dc_nested = {
  11: {
    'wkr': 1,
    'occ': 2,
    'wge': 1.2
  },
  202: {
    'wkr': 2,
    'occ': 2,
    'wge': None
  }
}
```

Second, we convert the nested dictionary, so that the number of observations is the number of key/values in the top nest layer, and the number of variables is the number of keys in the second layer of dictionary along with the key from the top layer. We will use [pandas.DataFrame.from_dict](#) to accomplish this.

```
st_key_var_name = "key_node"
# 1. convert to dataframe
df_from_nested = pd.DataFrame.from_dict(dc_nested, orient='index')
# 2. keys from top nest as variable and rename as key_node
df_from_nested = df_from_nested.reset_index()
df_from_nested.rename(columns={'index':st_key_var_name}, inplace=True)
```

```
# Print  
print(df_from_nested)
```

```
##      key_node  wkr  occ  wge  
## 0          11    1    2  1.2  
## 1          202    2    2  NaN
```

Third, now we convert the pandas dataframe we just created back to a nested dictionary. We will use the “index” option for the orient parameter for the `to_dict` function. We need to first convert the `key_node` variable created above to the index.

```
# 1. convert column to index  
df_from_nested = df_from_nested.set_index(st_key_var_name)  
# 2. Convert to dictionary  
dc_from_df = df_from_nested.to_dict(orient="index")  
# print  
pprint.pprint(dc_from_df)
```

```
## {11: {'occ': 2, 'wge': 1.2, 'wkr': 1}, 202: {'occ': 2, 'wge': nan, 'wkr': 2}}
```