

Python - Analiza danych z modulem PANDAS

www.udemy.com (<http://www.udemy.com>) (R)

LAB - S07-L008 - merge wielu obiektów data frame

1. Zaimportuj moduł pandas i numpy nadaj im standardowe aliasy
2. Do zmiennej **suppliers** zaimportuj zawartość pliku **suppliers.csv**. Pobierz tylko kolumny **SupplierID**, **CompanyName**, **City**
3. Do zmiennej **products** zaimportuj zawartość pliku **products.csv**. Pobierz tylko kolumny **ProductID**, **ProductName**, **SupplierID**, **CategoryID**
4. Do zmiennej **categories** zaimportuj zawartość pliku **categories1.csv**. Pobierz tylko kolumny **CategoryID** i **CategoryName**. Kolumna **CategoryID** ma być indeksem.
5. Połącz **products** i **categories** korzystając z **CategoryID**. Wynik zapisz w nowej zmiennej **prod_cat**
6. Połącz **prod_cat** z **suppliers** korzystając z **SupplierID**. Wynik zapisz w nowej zmiennej **prod_cat_sup**
7. Wykonaj złączenia opisane w krokach (5) i (6), ale bez stosowania pomocniczych zmiennych

Rozwiązania:

Poniżej znajdują się propozycje rozwiązań zadań. Prawdopodobnie istnieje wiele dobrych rozwiązań, dlatego jeżeli rozwiązujesz zadania samodzielnie, to najprawdopodobniej zrobisz to inaczej, może nawet lepiej :) Możesz pochwalić się swoimi rozwiązaniami w sekcji Q&A

```
In [1]: import pandas as pd
import numpy as np
```

```
In [2]: suppliers = pd.read_csv("./northwind-mongo-master/suppliers.csv",
                                usecols=["SupplierID", "CompanyName", "City"])
```

```
In [3]: products = pd.read_csv("./northwind-mongo-master/products.csv",
                                usecols=['ProductID', 'ProductName', 'SupplierID', "CategoryID"])
```

```
In [4]: categories = pd.read_csv("./northwind-mongo-master/categories1.csv",
                                usecols=['CategoryID', 'CategoryName'],
                                index_col="CategoryID")
```

```
In [5]: prod_cat = products.merge(categories, left_on="CategoryID", right_index=True)
```

```
In [6]: prod_cat.head()
```

Out[6]:

	ProductID	ProductName	SupplierID	CategoryID	CategoryName
0	1	Chai	1	1	Beverages
1	2	Chang	1	1	Beverages
23	24	Guaraná Fantástica	10	1	Beverages
33	34	Sasquatch Ale	16	1	Beverages
34	35	Steeleye Stout	16	1	Beverages

```
In [7]: prod_cat_sup = prod_cat.merge(suppliers, left_on="SupplierID", right_on="SupplierID")
```

```
In [8]: prod_cat_sup.head()
```

Out[8]:

	ProductID	ProductName	SupplierID	CategoryID	CategoryName	CompanyName	City
0	1	Chai	1	1	Beverages	Exotic Liquids	London
1	2	Chang	1	1	Beverages	Exotic Liquids	London
2	3	Aniseed Syrup	1	2	Condiments	Exotic Liquids	London
3	24	Guaraná Fantástica	10	1	Beverages	Refrescos Americanas LTDA	Sao Paulo
4	34	Sasquatch Ale	16	1	Beverages	Bigfoot Breweries	Bend

```
In [9]: products.merge(categories, left_on="CategoryID", right_index=True  
).merge(suppliers, left_on="SupplierID", right_on="SupplierID").head()
```

Out[9]:

	ProductID	ProductName	SupplierID	CategoryID	CategoryName	CompanyName	City
0	1	Chai	1	1	Beverages	Exotic Liquids	London
1	2	Chang	1	1	Beverages	Exotic Liquids	London
2	3	Aniseed Syrup	1	2	Condiments	Exotic Liquids	London
3	24	Guaraná Fantástica	10	1	Beverages	Refrescos Americanas LTDA	Sao Paulo
4	34	Sasquatch Ale	16	1	Beverages	Bigfoot Breweries	Bend

```
In [ ]:
```