

Python - Analiza danych z modulem PANDAS

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

LAB - S10-L003 - excel

1. Zaimportuj moduł pandas i numpy nadaj im standardowe aliasy
2. Do zmiennej **mcdonalds** zaimportuj plik **mcdonalds.csv**, pobierz tylko kolumny **Category**, **Item**, **Calories** i **TotalFat**
3. Zadeklaruj zmienną **excelWriter**, która pozwoli na zapisywanie danych do skoroszytu **mcdonalds-excel.xls**
4. Zapisz zawartość zmiennej **mcdonalds** do **excelWriter**. Nowy arkusz ma się nazywać **Menu**. Indeks nie powinien być zapisywany
5. Zapisz **excelWriter** na dysku. (Dobry pomysł sprawdzić, jak wygląda teraz plik, ale po takim sprawdzeniu pamiętaj aby zamknąć Excela, bo inaczej dostęp do pliku może być zablokowany)
6. Do zmiennej **imported1** zaimportuj ze skoroszytu **mcdonalds-excel.xls** arkusz **Menu**.
7. Zaktualizuj w **imported1** zawartość kolumny **Calories** dzieląc ją przez 2.
8. Zapisz zawartość **imported1** w **excelWriter**, dodając arkusz **Menu Light**
9. Zaktualizuj w **imported1** zawartość kolumny **Calories** dzieląc ją ponownie przez 2.
10. Zapisz zawartość **imported1** w **excelWriter**, dodając arkusz **Menu Extra Light**
11. Zapisz **excelWriter** na dysku.
12. Do zmiennej **final** wczytaj zawartość wszystkich arkuszy ze skoroszytu **mcdonalds-excel.xls**
13. Sprawdź ile masz obiektów data frame w zmiennej **final**, wyświetlając jej długość
14. Wyświetl nagłówek zerowego, pierwszego i drugiego arkusza. Czy widzisz zmniejszającą się ilość kalorii w kolejnych arkuszach?

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]: mcdonalds = pd.read_csv("Mcdonalds.csv", usecols=['Category', 'Item', 'Calories', 'TotalFat'])
mcdonalds.head()
```

```
Out[2]:
```

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	300	13.0
1	Breakfast	Egg White Delight	250	8.0
2	Breakfast	Sausage McMuffin	370	23.0
3	Breakfast	Sausage McMuffin with Egg	450	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	400	23.0

```
In [3]: excelWriter = pd.ExcelWriter("mcdonalds-excel.xls")
```

```
In [4]: mcdonalds.to_excel(excelWriter, sheet_name="Menu", index=False)
```

```
In [5]: excelWriter.save()
```

```
In [6]: imported1 = pd.read_excel("mcdonalds-excel.xls", sheet_name="Menu")
```

```
In [7]: imported1["Calories"] = imported1["Calories"] / 2
```

```
In [8]: imported1.head()
```

Out[8]:

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	150.0	13.0
1	Breakfast	Egg White Delight	125.0	8.0
2	Breakfast	Sausage McMuffin	185.0	23.0
3	Breakfast	Sausage McMuffin with Egg	225.0	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	200.0	23.0

```
In [9]: imported1.to_excel(excelWriter, sheet_name="Menu Light")
```

```
In [10]: imported1["Calories"] = imported1["Calories"]/2
```

```
In [11]: imported1.head()
```

Out[11]:

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	75.0	13.0
1	Breakfast	Egg White Delight	62.5	8.0
2	Breakfast	Sausage McMuffin	92.5	23.0
3	Breakfast	Sausage McMuffin with Egg	112.5	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	100.0	23.0

```
In [12]: imported1.to_excel(excelWriter, sheet_name="Menu Extra Light")
```

```
In [13]: excelWriter.save()
```

```
In [14]: final = pd.read_excel("mcdonalds-excel.xls", sheet_name=[0,1,2])
```

```
In [15]: len(final)
```

Out[15]: 3

```
In [16]: final[0].head()
```

Out[16]:

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	300	13.0
1	Breakfast	Egg White Delight	250	8.0
2	Breakfast	Sausage McMuffin	370	23.0
3	Breakfast	Sausage McMuffin with Egg	450	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	400	23.0

```
In [17]: final[1].head()
```

```
Out[17]:
```

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	150.0	13.0
1	Breakfast	Egg White Delight	125.0	8.0
2	Breakfast	Sausage McMuffin	185.0	23.0
3	Breakfast	Sausage McMuffin with Egg	225.0	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	200.0	23.0

```
In [18]: final[2].head()
```

```
Out[18]:
```

	Category	Item	Calories	TotalFat
0	Breakfast	Egg McMuffin	75.0	13.0
1	Breakfast	Egg White Delight	62.5	8.0
2	Breakfast	Sausage McMuffin	92.5	23.0
3	Breakfast	Sausage McMuffin with Egg	112.5	28.0
4	Breakfast	Sausage McMuffin with Egg Whites	100.0	23.0

```
In [ ]:
```