

Python - Analiza danych z modulem PANDAS

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

LAB - S07-L002 - metody append i concat

1. Zaimportuj moduł pandas i numpy nadaj im standardowe aliasy
2. Do zmiennej **mrbean** zaimportuj zawartość pliku **mrbean_facebook_statuses_join.csv**. Pobierz tylko kolumny **'status_message_mrbean', 'status_type', 'num_likes'**. To dane dotyczące statusów Jasia Fasoli (może znasz z TV ;))
3. Do zmiennej **nasa** zaimportuj zawartość pliku **nasa_facebook_statuses_join.csv**. Pobierz tylko kolumny **'status_message_nasa', 'status_type', 'num_likes'**. To dane dotyczące statusów NASA
4. Wyświetl nagłówki obu zmiennych
5. Do zmiennej **fb** zapisz wynik komendy **append**, która połączy dane z **mrbean** i z **nasa** (łączyć dane w tym i następnych poleceniach korzystaj z parametru **sort=True**)
6. Wyświetl początkowe i ostatnie wiersze **fb**. Zwróć uwagę na kolumny **status_message_mrbean** i **status_message_nasa**
7. Wyświetl wiersze **fb**, które mają indeks **0**. (Powinieneś zobaczyć 2 wiersze)
8. Zmień polecenie z kroku 5, tak aby obiekt **fb** miał wygenerowany nowy indeks
9. Wyświetl wiersze **fb**, które mają indeks **0**. Teraz powinien się znaleźć tylko jeden wiersz o indeksie 0
10. Zmień nazwę kolumny **status_message_mrbean** w data frame **mrbean** na **status_message**. Podobnie w data frame **nasa** zmień nazwę kolumny **status_message_nasa** na **status_message**.
11. Wykonaj ponownie połączenie z kroku 8. Potem wyświetl zmienną **fb**. Zauważ, że istnieje już tylko jedna kolumna z **status_message**.
12. Zmień poprzednie polecenie tak, aby korzystało z polecenia **concat** zamiast **append**

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]: mrbean = pd.read_csv('mrbean_facebook_statuses_join.csv',
                             usecols = ['status_message_mrbean', 'status_type', 'num_likes'])
nasa = pd.read_csv('nasa_facebook_statuses_join.csv',
                   usecols = ['status_message_nasa', 'status_type', 'num_likes'])
```

```
In [3]: mrbean.head()
```

Out[3]:

	status_message_mrbean	status_type	num_likes
0	It's time for Mr Bean and Teddy to get ready f...	video	16079
1	NaN	photo	114008
2	NaN	photo	218579
3	Mr Bean is ready to do some shopping but bewar...	video	20338
4	NaN	photo	215591

```
In [4]: nasa.head()
```

```
Out[4]:
```

		status_message_nasa	status_type	num_likes
0	This morning, the #RockOn mission with student...		photo	5142
1	Pluto's largest moon, Charon, is home to an un...		photo	19074
2	Our Restore-L mission is an endeavor to launch...		photo	7551
3	A new dark vortex was discovered in the atmo...		photo	24359
4		NaN	video	2731

```
In [5]: fb = mrbean.append(nasa, sort=True)
```

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

```
Out[6]:
```

	num_likes	status_message_mrbean	status_message_nasa	status_type
0	16079	It's time for Mr Bean and Teddy to get ready f...	NaN	video
1	114008		NaN	photo
2	218579		NaN	photo
3	20338	Mr Bean is ready to do some shopping but bewar...	NaN	video
4	215591		NaN	photo

```
In [7]: fb.tail()
```

```
Out[7]:
```

	num_likes	status_message_mrbean	status_message_nasa	status_type
758	39355	NaN	You've never seen Pluto like this before! New ...	photo
759	11835	NaN	Our cooperative field study in East Asia advan...	photo
760	15567	NaN	After spending a #YearInSpace, astronaut Scott...	photo
761	41584	NaN	Gravitational waves, predicted by Einstein and...	photo
762	12793	NaN	Have you ever used a GPS to help you get somew...	photo

```
In [8]: fb.loc[0]
```

```
Out[8]:
```

	num_likes	status_message_mrbean	status_message_nasa	status_type
0	16079	It's time for Mr Bean and Teddy to get ready f...	NaN	video
0	5142	NaN	This morning, the #RockOn mission with student...	photo

```
In [9]: fb = mrbean.append(nasa, sort=True, ignore_index=True)
```

```
In [10]: fb.loc[0]
```

```
Out[10]:
```

	num_likes	status_message_mrbean	status_message_nasa	status_type
0	16079	It's time for Mr Bean and Teddy to get ready f...	NaN	video

Name: 0, dtype: object

```
In [11]: mrbean.rename({"status_message_mrbean" : "status_message"}, axis=1, inplace=True)
```

```
In [12]: nasa.rename({"status_message_nasa" : "status_message"}, axis=1, inplace=True)
```

```
In [13]: fb = mrbean.append(nasa, sort=True, ignore_index=True)
```

```
In [14]: fb.head()
```

Out[14]:

	num_likes	status_message	status_type
0	16079	It's time for Mr Bean and Teddy to get ready f...	video
1	114008	NaN	photo
2	218579	NaN	photo
3	20338	Mr Bean is ready to do some shopping but bewar...	video
4	215591	NaN	photo

```
In [15]: fb = pd.concat([mrbean, nasa])
```

```
In [16]: fb.head()
```

Out[16]:

	status_message	status_type	num_likes
0	It's time for Mr Bean and Teddy to get ready f...	video	16079
1	NaN	photo	114008
2	NaN	photo	218579
3	Mr Bean is ready to do some shopping but bewar...	video	20338
4	NaN	photo	215591

```
In [ ]:
```