

Python - Analiza danych z modulem PANDAS

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

LAB - S03-L011-LAB Metody where i query

1. Zaimportuj moduł pandas oraz numpy i nadaj im standardowe aliasy. Do zmiennej **animals** wczytaj zawartość pliku **sleep_time.csv**. Pobierz wszystkie kolumny, a kolumnę ID wybierz jako indeks. Podczas wczytywania zamień wszystkie wartości **NaN** na 0. Wyświetl nagłówek zaimportowanego pliku.
2. W kolumnie **vore** znajduje się informacja o tym czy zwierzę jest roślinożerne (herbi), mięsożerne (carni) i wszystkożerne (omni). Korzystając z metody **where** wybierz wiersze opisujące zwierzęta roślinożerne. Wynik zapisz w zmiennej **herbi**
3. Ile wierszy jest w zmiennej **herbi**?
4. Ile wierszy jest w zmiennej **herbi** po usunięciu wartości **NaN**?
5. Utwórz zmienną **carni** i zapisz w niej wynik filtrowania obiektu **animals** tak, że pobrane zostaną tylko wiersze zwierząt mięsożernych. Wyświetl informację o ilości wierszy w zmiennej **carni** oraz o ilości wierszy po usunięciu wartości **NaN**.
6. Utwórz zmienną **omni** i zapisz w niej wynik filtrowania obiektu **animals** tak, że pobrane zostaną tylko wiersze zwierząt wszystkożernych. Wyświetl informację o ilości wierszy w zmiennej **omni** oraz o ilości wierszy po usunięciu wartości **NaN**.
7. Wyświetl średni czas snu (kolumna **sleep_total**) dla zwierząt roślino-, mieso- i wszystkożernych. Czy można próbować postawić hipotezę, że rodzaj pokarmu wpływa na czas snu?
8. Korzystając z polecenia **query** wybierz te wiersze, które opisują zwierzęta śpiące więcej niż 14 godzin na dobę (kolumna **sleep_total**)
9. Do warunku z poprzedniego zadania dodaj: rodzaj zwierzęcia (kolumna **vore**) ma być roślinożerny (wartość **herbi**)
10. Do warunku z poprzedniego zadania dodaj: waga ciała (kolumna **bodywt**) ma być większa niż 1

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
animals = pd.read_csv("sleep_time.csv", index_col="ID").fillna(0)
animals.head()
```

Out[1]:

	name	genus	vore	order	conservation	sleep_total	sleep_rem	sleep_cycle	awake	brainw
ID										
1	Cheetah	Acinonyx	carni	Carnivora	lc	12.1	0.0	0.000000	11.9	0.0000
2	Owl monkey	Aotus	omni	Primates	0	17.0	1.8	0.000000	7.0	0.0155
3	Mountain beaver	Aplodontia	herbi	Rodentia	nt	14.4	2.4	0.000000	9.6	0.0000
4	Greater short-tailed shrew	Blarina	omni	Soricomorpha	lc	14.9	2.3	0.133333	9.1	0.0002
5	Cow	Bos	herbi	Artiodactyla	domesticated	4.0	0.7	0.666667	20.0	0.4230

```
In [2]: herbi = animals.where( animals['vore'] == 'herbi')
```

```
In [3]: len(herbi)
```

```
Out[3]: 83
```

```
In [4]: len(herbi.dropna())
```

```
Out[4]: 32
```

```
In [5]: carni = animals.where( animals['vore'] == 'carni')
print(len(carni))
print(len(carni.dropna()))
```

```
83
```

```
19
```

```
In [6]: omni = animals.where( animals['vore'] == 'omni')
print(len(omni))
print(len(omni.dropna()))
```

```
83
```

```
20
```

```
In [7]: print('carni mean sleep time:', carni['sleep_total'].mean())
print('herbi mean sleep time:', herbi['sleep_total'].mean())
print('omni mean sleep time:', omni['sleep_total'].mean())
```

```
carni mean sleep time: 10.378947368421054
```

```
herbi mean sleep time: 9.509375
```

```
omni mean sleep time: 10.924999999999999
```

```
In [8]: animals.query("sleep_total > 14")
```

```
Out[8]:
```

	name	genus	vore	order	conservation	sleep_total	sleep_rem	sleep_cycle	awak
ID									
2	Owl monkey	Aotus	omni	Primates	0	17.0	1.8	0.000000	7.
3	Mountain beaver	Aplodontia	herbi	Rodentia	nt	14.4	2.4	0.000000	9.
4	Greater short-tailed shrew	Blarina	omni	Soricomorpha	lc	14.9	2.3	0.133333	9.
6	Three-toed sloth	Bradypus	herbi	Pilosa	0	14.4	2.2	0.766667	9.
18	Long-nosed armadillo	Dasypus	carni	Cingulata	lc	17.4	3.1	0.383333	6.
20	North American Opossum	Didelphis	omni	Didelphimorphia	lc	18.0	4.9	0.333333	6.
22	Big brown bat	Eptesicus	insecti	Chiroptera	lc	19.7	3.9	0.116667	4.
27	Western american chipmunk	Eutamias	herbi	Rodentia	0	14.9	0.0	0.000000	9.
37	Thick-tailed opossum	Lutreolina	carni	Didelphimorphia	lc	19.4	6.6	0.000000	4.
39	Mongolian gerbil	Meriones	herbi	Rodentia	lc	14.2	1.9	0.000000	9.
40	Golden hamster	Mesocricetus	herbi	Rodentia	en	14.3	3.1	0.200000	9.
43	Little brown bat	Myotis	insecti	Chiroptera	0	19.9	2.0	0.200000	4.
44	Round-tailed muskrat	Neofiber	herbi	Rodentia	nt	14.6	0.0	0.000000	9.
47	Northern grasshopper mouse	Onychomys	carni	Rodentia	lc	14.5	0.0	0.000000	9.
51	Tiger	Panthera	carni	Carnivora	en	15.8	0.0	0.000000	8.
62	Giant armadillo	Priodontes	insecti	Cingulata	en	18.1	6.1	0.000000	5.
70	Arctic ground squirrel	Spermophilus	herbi	Rodentia	lc	16.6	0.0	0.000000	7.
72	Golden-mantled ground squirrel	Spermophilus	herbi	Rodentia	lc	15.9	3.0	0.000000	8.
76	Eastern american chipmunk	Tamias	herbi	Rodentia	0	15.8	0.0	0.000000	8.
78	Tenrec	Tenrec	omni	Afrosoricida	0	15.6	2.3	0.000000	8.

```
In [9]: animals.query("sleep_total > 14 and vore == 'herbi'")
```

```
Out[9]:
```

	name	genus	vore	order	conservation	sleep_total	sleep_rem	sleep_cycle	awake	brainwt
ID										
3	Mountain beaver	Aplodontia	herbi	Rodentia	nt	14.4	2.4	0.000000	9.6	0.0000
6	Three-toed sloth	Bradypus	herbi	Pilosa	0	14.4	2.2	0.766667	9.6	0.0000
27	Western american chipmunk	Eutamias	herbi	Rodentia	0	14.9	0.0	0.000000	9.1	0.0000
39	Mongolian gerbil	Meriones	herbi	Rodentia	lc	14.2	1.9	0.000000	9.8	0.0000
40	Golden hamster	Mesocricetus	herbi	Rodentia	en	14.3	3.1	0.200000	9.7	0.0010
44	Round-tailed muskrat	Neofiber	herbi	Rodentia	nt	14.6	0.0	0.000000	9.4	0.0000
70	Arctic ground squirrel	Spermophilus	herbi	Rodentia	lc	16.6	0.0	0.000000	7.4	0.0050
72	Golden-mantled ground squirrel	Spermophilus	herbi	Rodentia	lc	15.9	3.0	0.000000	8.1	0.0000
76	Eastern american chipmunk	Tamias	herbi	Rodentia	0	15.8	0.0	0.000000	8.2	0.0000

```
In [10]: animals.query("sleep_total > 14 and vore == 'herbi' and bodywt > 1")
```

```
Out[10]:
```

	name	genus	vore	order	conservation	sleep_total	sleep_rem	sleep_cycle	awake	brainwt	b
ID											
3	Mountain beaver	Aplodontia	herbi	Rodentia	nt	14.4	2.4	0.000000	9.6	0.0	
6	Three-toed sloth	Bradypus	herbi	Pilosa	0	14.4	2.2	0.766667	9.6	0.0	

```
In [ ]:
```