Python - Analiza danych z modułem PANDAS

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

LAB - S04-L003-Dodawanie i usuwanie wierszy

1. Zaimportuj moduł pandas i numpy, nadaj im standardowe aliasy. Do zmiennej **professions** wczytaj zawartość pliku **Prestige.csv**. Wyświetl nagłówek obiektu data frame.
2. Zdefiniuj słownik w zmiennej **dict**, który ma klucze takie, jak nazwy kolumn w **professions**. Wpisz do wartości dla klucza **name** - **data scientist**, pozostałe wartości zdefiniuj jako NaN. Wyświetl obiekt w celu weryfikacji.
3. Dodaj do **professions** nowy wiersz zdefiniowany przez słownik **dict**. W celu weryfikacji wyświetl 5 ostatnich wartości z **professions**
4. Odszukaj w **professions** wiersza, w którym name to **taxi.drivers**. Zapisz go do zmiennej **taxi** i wyświetl.
5. Usuń z **professions** wiersz z indekstem 98 (to właśnie **taxi.drivers**). Wyświetl 5 ostatnich wartości z **professions**
6. Dodaj wiersz znajdujący się w **taxi** do **professions**, w taki sposób, aby wiersz na stałe znajdował sie w zmiennej **professions**. ***Wyświetl ostatnich 5 wierszy*** ***professions** w celu weryfikacji
7. Usuń z **professions** kolumnę **census**
8. Usuń z **professions** kolumnę **type**. Wykorzystaj do tego inną metodę niż wykorzystana w poprzednim zadaniu

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
professions = pd.read_csv("Prestige.csv")
professions.head(5)
```

Out[1]:

	name	education	income	women	prestige	census	type
0	gov.administrators	13.11	12351	11.16	68.8	1113	prof
1	general.managers	12.26	25879	4.02	69.1	1130	prof
2	accountants	12.77	9271	15.70	63.4	1171	prof
3	purchasing.officers	11.42	8865	9.11	56.8	1175	prof
4	chemists	14.62	8403	11.68	73.5	2111	prof

In [2]:
```
dict = { 'name'            :   'data scientist',
         'education'       :   np.NaN,
         'income'          :   np.NaN,
         'women'           :   np.NaN,
         'prestige'        :   np.NaN,
         'census'          :   np.NaN,
         'type'            :   np.NaN}
dict
```

Out[2]:
```
{'name': 'data scientist',
 'education': nan,
 'income': nan,
 'women': nan,
 'prestige': nan,
 'census': nan,
 'type': nan}
```

In [3]:
```
professions = professions.append(dict, ignore_index=True)
professions.tail()
```

Out[3]:

	name	education	income	women	prestige	census	type
98	taxi.drivers	7.93	4224.0	3.59	25.1	9173.0	bc
99	longshoremen	8.37	4753.0	0.00	26.1	9313.0	bc
100	typesetters	10.00	6462.0	13.58	42.2	9511.0	bc
101	bookbinders	8.55	3617.0	70.87	35.2	9517.0	bc
102	data scientist	NaN	NaN	NaN	NaN	NaN	NaN

In [4]:
```
taxi = professions[professions["name"]=='taxi.drivers']
taxi
```

Out[4]:

	name	education	income	women	prestige	census	type
98	taxi.drivers	7.93	4224.0	3.59	25.1	9173.0	bc

In [5]:
```
professions.drop(98, inplace=True)
professions.tail()
```

Out[5]:

	name	education	income	women	prestige	census	type
97	bus.drivers	7.58	5562.0	9.47	35.9	9171.0	bc
99	longshoremen	8.37	4753.0	0.00	26.1	9313.0	bc
100	typesetters	10.00	6462.0	13.58	42.2	9511.0	bc
101	bookbinders	8.55	3617.0	70.87	35.2	9517.0	bc
102	data scientist	NaN	NaN	NaN	NaN	NaN	NaN

In [6]:
```
professions = professions.append(taxi)
professions.tail()
```

Out[6]:

	name	education	income	women	prestige	census	type
99	longshoremen	8.37	4753.0	0.00	26.1	9313.0	bc
100	typesetters	10.00	6462.0	13.58	42.2	9511.0	bc
101	bookbinders	8.55	3617.0	70.87	35.2	9517.0	bc

	name	education	income	women	prestige	census	type
102	data scientist	NaN	NaN	NaN	NaN	NaN	NaN

In [7]:
```
del professions ["census"]
professions.tail()
```

Out[7]:

	name	education	income	women	prestige	type
99	longshoremen	8.37	4753.0	0.00	26.1	bc
100	typesetters	10.00	6462.0	13.58	42.2	bc
101	bookbinders	8.55	3617.0	70.87	35.2	bc
102	data scientist	NaN	NaN	NaN	NaN	NaN
98	taxi.drivers	7.93	4224.0	3.59	25.1	bc

In [8]:
```
professions.drop(columns='type', inplace=True)
professions.tail()
```

Out[8]:

	name	education	income	women	prestige
99	longshoremen	8.37	4753.0	0.00	26.1
100	typesetters	10.00	6462.0	13.58	42.2
101	bookbinders	8.55	3617.0	70.87	35.2
102	data scientist	NaN	NaN	NaN	NaN
98	taxi.drivers	7.93	4224.0	3.59	25.1

In [ ]:
```

```