

Python - Analiza danych z modulem PANDAS

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

LAB - S02-L014 - Reindex i intersection

1. Zaimportuj moduł pandas i nadaj mu standardowy alias.
2. Utwórz serię **countries** importując z pliku **countries.csv** kolumny **Symbol** i **Name**. Kolumna **Symbol** powinna stać się indeksem serii. Skorzystaj z polecenia:

```
countries = pd.read_csv("countries.csv", usecols=['Symbol','Name'], index_col='Symbol', squeeze=True)
```

3. Usuń niekompletne dane korzystając z polecenia:

```
countries.dropna(inplace=True)
```

4. Wyświetl 20 pierwszych wierszy tej serii.
5. Utwórz listę **toFind** zawierającą symbole krajów: BB, AA, BS (kraju AA nie ma w wykazie państw, bo to ... Atalantyda...
6. Korzystając z metody loc i zmiennej wyświetl wynik poszukiwania krajów o symbolach z listy **toFind**. (Oprócz wyniku powinno zostać wyświetlone ostrzeżenie)
7. Wyszukaj krajów **toFind** korzystając z metody reindex.
8. Wyświetl listę będącą częścią wspólną indeksu serii **countries** oraz listy **toFind**.
9. Korzystając z metody loc oraz listy wygenerowanej w poprzednim punkcie wyświetl kraje o symbolach z listy **toFind**

Rozwiązania:

Poniżej znajdują się propozycje rozwiązań zadań. Prawdopodobnie istnieje wiele dobrych rozwiązań, dlatego jeżeli rozwiążesz 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
```

```
In [2]: countries = pd.read_csv("countries.csv", usecols=['Symbol', 'Name'], index_col='Symbol',
```

```
In [3]: countries.dropna(inplace=True)
```

```
In [4]: countries.head(20)
```

```
Out[4]: Symbol
AF      Afghanistan
AL      Albania
DZ      Algeria
AD      Andorra
AO      Angola
AI      Anguilla
AQ      Antarctica
AG      Antigua and Barbuda
SA      Saudi Arabia
AR      Argentina
AM      Armenia
AW      Aruba
AU      Australia
AT      Austria
AZ      Azerbaijan
BS      Bahamas
BH      Bahrain
BD      Bangladesh
BB      Barbados
BE      Belgium
Name: Name, dtype: object
```

```
In [5]: toFind = ['BB', 'AA', 'BS']
```

```
In [6]: countries.loc[toFind]
```

```
C:\Users\rafal\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: FutureWarning:
Passing list-likes to .loc or [] with any missing label will raise
KeyError in the future, you can use .reindex() as an alternative.
```

See the documentation here:

<https://pandas.pydata.org/pandas-docs/stable/indexing.html#deprecate-loc-reindex-listlike> (<https://pandas.pydata.org/pandas-docs/stable/indexing.html#deprecate-loc-reindex-listlike>)

"""Entry point for launching an IPython kernel.

```
Out[6]: Symbol
BB      Barbados
AA      NaN
BS      Bahamas
Name: Name, dtype: object
```

```
In [7]: countries.reindex(toFind)
```

```
Out[7]: Symbol
BB      Barbados
AA      NaN
BS      Bahamas
Name: Name, dtype: object
```

```
In [8]: countries.index.intersection(toFind)
```

```
Out[8]: Index(['BS', 'BB'], dtype='object')
```

```
In [9]: countries.loc[countries.index.intersection(toFind)]
```

```
Out[9]: BS      Bahamas  
        BB      Barbados  
        Name: Name, dtype: object
```