

Python - Analiza danych z modulem PANDAS

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

LAB - S03-L003 - Pobieranie danych z Data Frame

1. Zaimportuj moduł pandas i nadaj mu standardowy alias. Do zmiennej **fuel** wczytaj zawartość pliku **fuel.csv**. Podczas wczytywania skorzystaj z dodatkowego argumentu **low_memory=False**, pobierz tylko następujące kolumny: 'Vehicle ID','Year','Make','Model','Class','Fuel Type','Combined MPG (FT1)'. Pobierając dane oznacz 'Vehicle ID' jako indeks. Wyświetl nagłówki tak utworzonego Data Frame
2. Wyświetl nagłówek serii zawierającej tylko dane z kolumny **Make**
3. Korzystając z funkcji wyświetlającej dla serii danych ilościowe informacje o wartościach z kolumny wyświetl 5 najczęściej ocenianych w pliku fuel producentów samochodów (kolumna **Make**)
4. Wyświetl informacje o samochodzie znajdującym się pod indeksem Vehicle ID **1873**
5. Wyświetl informacje o liczbie mil jakie można przejechać na galonie paliwa (kolumna **Combined MPG (FT1)**) dla samochodu pod indeksem Vehicle ID **1873**
6. Sprawdź ile najwięcej mil da się przejechać samochodem na jednym galonie paliwa (innymi słowy poszukaj maksymalnej wartości w kolumnie **Combined MPG (FT1)**)
7. Znajdź indeks opisujący samochód z wynikiem znalezionym w poprzednim punkcie...
8. ...a następnie posiadając ten indeks, wyświetl wszystkie dostępne w zmiennej **fuel** informacje o tym samochodzie.
9. Połącz polecenia, które zbudowałeś w 2 poprzednich punktach, tak, aby potrzebna informacja była uzyskiwana w pojedynczej instrukcji.
10. Utwórz nowy obiekt data frame nazwany **shortFuel**, który posiada tylko 2 kolumny skopiowane z obiektu **fuel**: **Make** i **Model**.
11. Wyświetl nagłówek zmiennej **shortFuel**

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
fuel = pd.read_csv("fuel.csv", usecols=['Vehicle ID', 'Year', 'Make',
                                         'Model', 'Class', 'Fuel Type',
                                         'Combined MPG (FT1)'],
                index_col = 'Vehicle ID')
fuel.head()
```

```
Out[1]:
```

	Year	Make	Model	Class	Fuel Type	Combined MPG (FT1)
Vehicle ID						
26587	1984	Alfa Romeo	GT V6 2.5	Minicompact Cars	Regular	20.0
27705	1984	Alfa Romeo	GT V6 2.5	Minicompact Cars	Regular	20.0
26561	1984	Alfa Romeo	Spider Veloce 2000	Two Seaters	Regular	21.0
27681	1984	Alfa Romeo	Spider Veloce 2000	Two Seaters	Regular	21.0
27550	1984	AM General	DJ Po Vehicle 2WD	Special Purpose Vehicle 2WD	Regular	17.0

```
In [2]: fuel['Make'].head()
```

```
Out[2]:
```

```

Vehicle ID
26587    Alfa Romeo
27705    Alfa Romeo
26561    Alfa Romeo

```

```
In [3]: fuel['Make'].value_counts().head()
```

```

Out[3]: Chevrolet    3389
        Ford         2721
        Dodge        2361
        GMC          2174
        Toyota       1599
        Name: Make, dtype: int64

```

```
In [4]: fuel.loc[1873]
```

```

Out[4]: Year                1986
        Make                Mitsubishi
        Model              Starion
        Class      Subcompact Cars
        Fuel Type      Premium
        Combined MPG (FT1)      19
        Name: 1873, dtype: object

```

```
In [5]: fuel.loc[1873, 'Combined MPG (FT1)']
```

```
Out[5]: 19.0
```

```
In [6]: fuel["Combined MPG (FT1)"].max()
```

```
Out[6]: 112.0
```

```
In [7]: fuel["Combined MPG (FT1)"].idxmax()
```

```
Out[7]: 31673
```

```
In [8]: fuel.loc[31673]
```

```

Out[8]: Year                2012
        Make                Mitsubishi
        Model              i-MiEV
        Class      Subcompact Cars
        Fuel Type      Electricity
        Combined MPG (FT1)      112
        Name: 31673, dtype: object

```

```
In [9]: fuel.loc[fuel["Combined MPG (FT1)"].idxmax()]
```

```

Out[9]: Year                2012
        Make                Mitsubishi
        Model              i-MiEV
        Class      Subcompact Cars
        Fuel Type      Electricity
        Combined MPG (FT1)      112
        Name: 31673, dtype: object

```

```
In [10]: fuelShort = fuel[['Make', 'Model']]
```

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

```

Out[11]:
           Make      Model
Vehicle ID
26587  Alfa Romeo  GT V6 2.5
27705  Alfa Romeo  GT V6 2.5
26561  Alfa Romeo  Spider Veloce 2000

```

Vehicle ID	Make	Model
------------	------	-------