

Python - Analiza danych z modulem PANDAS - S03-L002

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

LAB - Metody i atrybuty 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)'. Wyświetl nagłówek tak utworzonego Data Frame
2. Wyświetl podstawowe informacje o obiekcie korzystając z metody **info()**
3. Wyświetl informacje o typach każdej z poszczególnych kolumn
4. Ile razy jest wykorzystywany każdy z typów? Odpowiadając na pytanie skorzystaj z dedykowanej do tego funkcji.
5. Kolumna 'Make' zawiera informacje o producentach samochodów. Wyświetl 10 producentów wraz z informacją o ilości samochodów tego producenta opisanych w obiekcie frame. Skorzystaj z dedykowanej do tego celu metody serii danych.
6. Wyświetl informacje o indeksie obiektu **fuel**
7. Wyświetl informacje o kolumnach obiektu **fuel**
8. Wyświetl informacje o wartościach obiektu **fuel**
9. Wyświetl informacje o "osiach" obiektu **fuel**
10. Wyświetl informacje o "kształcie" obiektu **fuel** - ile ma wierszy i ile kolumn
11. Porównaj co zwraca metoda **count** wywołana na rzecz **frame**, a...
12. co zwraca funkcja **len**

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
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.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 31684 entries, 26587 to 32106
Data columns (total 6 columns):
Year                31684 non-null int64
Make                31684 non-null object
Model              31683 non-null object
Class              31683 non-null object
Fuel Type          31683 non-null object
Combined MPG (FT1) 31683 non-null float64
dtypes: float64(1), int64(1), object(4)
memory usage: 1.7+ MB
```

```
In [3]: fuel.dtypes
```

```
Out[3]: Year                int64
Make                object
Model              object
Class              object
Fuel Type          object
Combined MPG (FT1)  float64
dtype: object
```

```
In [4]: fuel.get_dtype_counts()
```

```
Out[4]: float64      1
int64      1
object      4
dtype: int64
```

```
In [5]: fuel["Make"].value_counts().head(10)
```

```
Out[5]: Chevrolet      3389
Ford                  2721
Dodge                 2361
GMC                   2174
Toyota               1599
BMW                   1180
Nissan                1082
Mercedes-Benz         948
Mitsubishi            937
Pontiac               893
Name: Make, dtype: int64
```

```
In [6]: fuel.index
```

```
Out[6]: Int64Index([26587, 27705, 26561, 27681, 27550, 28426, 27549, 28425, 27593,
                    28455,
                    ...,
                    31673, 32062, 32059, 32061, 32058, 32060, 32063, 32065, 32064,
                    32106],
                    dtype='int64', name='Vehicle ID', length=31684)
```

```
In [7]: fuel.columns
```

```
Out[7]: Index(['Year', 'Make', 'Model', 'Class', 'Fuel Type', 'Combined MPG (FT1)'], dtype
              = 'object')
```

```
In [8]: fuel.values
```

```
Out[8]: array([[1984, 'Alfa Romeo', 'GT V6 2.5', 'Minicompact Cars', 'Regular',
                20.0],
               [1984, 'Alfa Romeo', 'GT V6 2.5', 'Minicompact Cars', 'Regular',
                20.0],
               [1984, 'Alfa Romeo', 'Spider Veloce 2000', 'Two Seaters',
                'Regular', 21.0],
               ...,
               [2012, 'Mitsubishi', 'Lancer Evolution', 'Compact Cars',
                'Premium', 19.0],
               [2012, 'Mitsubishi', 'Lancer Evolution', 'Compact Cars',
                'Premium', 19.0],
               [2012, 'Mitsubish', nan, nan, nan, nan]], dtype=object)
```

```
In [9]: fuel.axes
```

```
Out[9]: [Int64Index([26587, 27705, 26561, 27681, 27550, 28426, 27549, 28425, 27593,
                    28455,
                    ...
                    31673, 32062, 32059, 32061, 32058, 32060, 32063, 32065, 32064,
                    32106],
                    dtype='int64', name='Vehicle ID', length=31684),
         Index(['Year', 'Make', 'Model', 'Class', 'Fuel Type', 'Combined MPG (FT1)'], dtype='object')]
```

```
In [10]: fuel.shape
```

```
Out[10]: (31684, 6)
```

```
In [11]: fuel.count()
```

```
Out[11]: Year                31684
         Make                31684
         Model              31683
         Class              31683
         Fuel Type          31683
         Combined MPG (FT1) 31683
         dtype: int64
```

```
In [12]: len(fuel)
```

```
Out[12]: 31684
```