

Python - Analiza danych z modulem PANDAS

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

LAB - S05-L008- Multiindex Metadata

1. Zaimportuj moduł pandas i numpy, nadaj im standardowe aliasy. Do zmiennej **air** wczytaj zawartość pliku **Air_Traffic_Passenger_Statistics.csv**. Pobierz wszystkie kolumny. Wyświetl nagłówek obiektu data frame.
2. Wyświetl informacje o automatycznie utworzonym indeksie dla obiektu **air**
3. Zmień indeks na **"Operating Airline", "Activity Period"**. Wyświetl 10 pierwszych wierszy
4. Wyświetl:
 - informacje o indeksie
 - informacje o poziomie zerowym indeksu
 - informacje o poziomie pierwszym indeksu
 - informacje o nazwach segmentów indeksu
5. Zresetuj indeks. Wyświetl nagłówek
6. Zmień indeks na **"Activity Period", "Operating Airline"**. Wyświetl 10 pierwszych wierszy
7. Wyświetl:
 - informacje o indeksie
 - informacje o poziomie zerowym indeksu
 - informacje o poziomie pierwszym indeksu
 - informacje o nazwach segmentów indeksu

Dane pochodzą ze strony <https://data.sfgov.org/Transportation/Air-Traffic-Passenger-Statistics/rkru-6vcg>
<https://data.sfgov.org/Transportation/Air-Traffic-Passenger-Statistics/rkru-6vcg>

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
air = pd.read_csv("./Air_Traffic_Passenger_Statistics.csv")
air.head()
```

Out[1]:

	Activity Period	Operating Airline	Operating Airline IATA Code	Published Airline	Published Airline IATA Code	GEO Summary	GEO Region	Activity Type Code	Price Category Code	Terminal
0	200507	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Deplaned	Low Fare	Terminal 1
1	200507	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Enplaned	Low Fare	Terminal 1
2	200507	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Thru / Transit	Low Fare	Terminal 1
3	200507	Air Canada	AC	Air Canada	AC	International	Canada	Deplaned	Other	Terminal 1
4	200507	Air Canada	AC	Air Canada	AC	International	Canada	Enplaned	Other	Terminal 1

```
In [2]: air.index
```

```
Out[2]: RangeIndex(start=0, stop=18398, step=1)
```

```
In [3]: air.set_index(["Operating Airline", "Activity Period"], inplace=True)
air.head(10)
```

```
Out[3]:
```

		Operating Airline IATA Code	Published Airline	Published Airline IATA Code	GEO Summary	GEO Region	Activity Type Code	Price Category Code	Terminal
Operating Airline	Activity Period								
ATA Airlines	200507	TZ	ATA Airlines	TZ	Domestic	US	Deplaned	Low Fare	Terminal 1
	200507	TZ	ATA Airlines	TZ	Domestic	US	Enplaned	Low Fare	Terminal 1
	200507	TZ	ATA Airlines	TZ	Domestic	US	Thru / Transit	Low Fare	Terminal 1
Air Canada	200507	AC	Air Canada	AC	International	Canada	Deplaned	Other	Terminal 1
	200507	AC	Air Canada	AC	International	Canada	Enplaned	Other	Terminal 1
Air China	200507	CA	Air China	CA	International	Asia	Deplaned	Other	International
	200507	CA	Air China	CA	International	Asia	Enplaned	Other	International
Air France	200507	AF	Air France	AF	International	Europe	Deplaned	Other	International
	200507	AF	Air France	AF	International	Europe	Enplaned	Other	International
Air New Zealand	200507	NZ	Air New Zealand	NZ	International	Australia / Oceania	Deplaned	Other	International

```
In [4]: air.index
```

```
Out[4]: MultiIndex(levels=[['ABC Aerolineas S.A. de C.V. dba Interjet', 'ATA Airlines',
'Aer Lingus', 'Aeromexico', 'Air Berlin', 'Air Canada', 'Air Canada Jazz', 'Air
China', 'Air France', 'Air India Limited', 'Air New Zealand', 'Air Pacific Limit
ed dba Fiji Airways', 'AirTran Airways', 'Alaska Airlines', 'All Nippon Airways'
, 'Allegiant Air', 'American Airlines', 'American Eagle Airlines', 'Ameriflight'
, 'Asiana Airlines', 'Atlantic Southeast Airlines', 'Atlas Air, Inc', 'BelAir Ai
rlines', 'Boeing Company', 'British Airways', 'COPA Airlines, Inc.', 'Cathay Pac
ific', 'China Airlines', 'China Eastern', 'China Southern', 'Compass Airlines',
'Delta Air Lines', 'EVA Airways', 'Emirates ', 'Etihad Airways', 'Evergreen Inte
rnational Airlines', 'ExpressJet Airlines', 'Finnair', 'Frontier Airlines', 'Haw
aiian Airlines', 'Hong Kong Airlines Limited', 'Horizon Air ', 'Icelandair', 'In
dependence Air', 'Japan Airlines', 'Jazz Aviation', 'Jet Airways', 'JetBlue Airw
ays', 'KLM Royal Dutch Airlines', 'Korean Air Lines', 'LAN Peru', 'Lufthansa Ger
man Airlines', 'Mesa Airlines', 'Mesaba Airlines', 'Mexicana Airlines', 'Miami A
ir International', 'Midwest Airlines', 'Northwest Airlines', 'Pacific Aviation',
'Philippine Airlines', 'Qantas Airways', 'Republic Airlines', 'SAS Airlines', 'S
ervisair', 'Singapore Airlines', 'SkyWest Airlines', 'Southwest Airlines', 'Spir
it Airlines', 'Sun Country Airlines', 'Swiss International', 'Swissport USA', 'T
ACA', 'Thomas Cook Airlines', 'Trego Dugan Aviation', 'Turkish Airlines', 'US Ai
rways', 'United Airlines', 'United Airlines', 'US Airways', 'Virgin America', 'V

```

```
In [5]: air.index.get_level_values(0)
```

```
Out[5]: Index(['ATA Airlines', 'ATA Airlines', 'ATA Airlines', 'Air Canada',
              'Air Canada', 'Air China', 'Air China', 'Air France', 'Air France',
              'Air New Zealand',
              ...
              'Virgin America', 'Virgin America', 'Virgin America', 'Virgin Atlantic',
              'Virgin Atlantic', 'Volaris Airlines', 'Volaris Airlines', 'WOW Air',
              'WOW Air', 'WOW Air'],
              dtype='object', name='Operating Airline', length=18398)
```

```
In [6]: air.index.get_level_values(1)
```

```
Out[6]: Int64Index([200507, 200507, 200507, 200507, 200507, 200507, 200507, 200507,
                    200507, 200507,
                    ...
                    201803, 201803, 201803, 201803, 201803, 201803, 201803, 201803,
                    201803, 201803],
                    dtype='int64', name='Activity Period', length=18398)
```

```
In [7]: air.index.names
```

```
Out[7]: FrozenList(['Operating Airline', 'Activity Period'])
```

```
In [8]: air.reset_index(inplace=True)
air.head()
```

```
Out[8]:
```

	Operating Airline	Activity Period	Operating Airline IATA Code	Published Airline	Published Airline IATA Code	GEO Summary	GEO Region	Activity Type Code	Price Category Code	Terminal	i
0	ATA Airlines	200507	TZ	ATA Airlines	TZ	Domestic	US	Deplaned	Low Fare	Terminal 1	
1	ATA Airlines	200507	TZ	ATA Airlines	TZ	Domestic	US	Enplaned	Low Fare	Terminal 1	
2	ATA Airlines	200507	TZ	ATA Airlines	TZ	Domestic	US	Thru / Transit	Low Fare	Terminal 1	
3	Air Canada	200507	AC	Air Canada	AC	International	Canada	Deplaned	Other	Terminal 1	
4	Air Canada	200507	AC	Air Canada	AC	International	Canada	Enplaned	Other	Terminal 1	

```
In [9]: air.set_index(["Activity Period", "Operating Airline"], inplace=True)
air.head(10)
```

Out[9]:

		Operating Airline IATA Code	Published Airline	Published Airline IATA Code	GEO Summary	GEO Region	Activity Type Code	Price Category Code	Terminal
Activity Period	Operating Airline								
200507	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Deplaned	Low Fare	Terminal 1
	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Enplaned	Low Fare	Terminal 1
	ATA Airlines	TZ	ATA Airlines	TZ	Domestic	US	Thru / Transit	Low Fare	Terminal 1
	Air Canada	AC	Air Canada	AC	International	Canada	Deplaned	Other	Terminal 1
	Air Canada	AC	Air Canada	AC	International	Canada	Enplaned	Other	Terminal 1
	Air China	CA	Air China	CA	International	Asia	Deplaned	Other	International
	Air China	CA	Air China	CA	International	Asia	Enplaned	Other	International
	Air France	AF	Air France	AF	International	Europe	Deplaned	Other	International
	Air France	AF	Air France	AF	International	Europe	Enplaned	Other	International
	Air New Zealand	NZ	Air New Zealand	NZ	International	Australia / Oceania	Deplaned	Other	International

```
In [10]: air.index.get_level_values(0)
```

```
Out[10]: Int64Index([200507, 200507, 200507, 200507, 200507, 200507, 200507, 200507,
                    200507, 200507,
                    ...,
                    201803, 201803, 201803, 201803, 201803, 201803, 201803, 201803, 201803,
                    201803, 201803],
                    dtype='int64', name='Activity Period', length=18398)
```

```
In [11]: air.index.get_level_values(1)
```

```
Out[11]: Index(['ATA Airlines', 'ATA Airlines', 'ATA Airlines', 'Air Canada',
                'Air Canada', 'Air China', 'Air China', 'Air France', 'Air France',
                'Air New Zealand',
                ...,
                'Virgin America', 'Virgin America', 'Virgin America', 'Virgin Atlantic',
                'Virgin Atlantic', 'Volaris Airlines', 'Volaris Airlines', 'WOW Air',
                'WOW Air', 'WOW Air'],
                dtype='object', name='Operating Airline', length=18398)
```

```
In [12]: air.index.names
```

```
Out[12]: FrozenList(['Activity Period', 'Operating Airline'])
```

In []: