

Data:

Imię i nazwisko:

Zadanie 1.

Dokonując przekształceń ciągów znaków w każdej ze zmiennych zadeklarowanych na *Listing 1*. pozostaw tekst "Jana III Sobieskiego". Wielkość liter ma znaczenie. Do manipulacji użyj wycięć zakresów (*slice*), podmian (*replace*), zmian wielkości liter (*lower*, *upper*, *title*).

Listing 1.

```
a = 'UL. JANA III SOBIESKIEGO 1/2'
b = 'ulica Jana III Sobieskiego 1 apt 2'
c = 'os. Jana III Sobieskiego'
d = 'plac Jana III SobieSKiego 1/2'
e = 'aleja JANA III Sobieskiego'
f = 'alei Jana III Sobieskiego 1/2'
g = 'ul. jana III SOBIESKIEGO 1 m. 2'
h = 'os. Jana Iii Sobieskiego 1 apt 2'
```

Zadanie 2.

Stwórz pustą listę *features* oraz *species*. Iteruj po *DATA* z *Listing 2*. i dla każdego wiersza za pomocą *slice* wydrebnij pomiary oraz nazwę gatunku. Pomiary dodaj do *features* a nazwę gatunku do *species*. Na końcu *features* powinno być listą tupli, a *species* listą ciągów znaków (*str*).

Listing 2.

```
DATA = [
    (5.8, 2.7, 5.1, 1.9, 'virginica'),
    (5.1, 3.5, 1.4, 0.2, 'setosa'),
    (5.7, 2.8, 4.1, 1.3, 'versicolor'),
    (6.3, 2.9, 5.6, 1.8, 'virginica'),
    (6.4, 3.2, 4.5, 1.5, 'versicolor'),
    (4.7, 3.2, 1.3, 0.2, 'setosa'),
    (7.0, 3.2, 4.7, 1.4, 'versicolor'),
    (7.6, 3.0, 6.6, 2.1, 'virginica'),
    (4.9, 3.0, 1.4, 0.2, 'setosa'),
    (4.6, 3.1, 1.5, 0.2, 'setosa'),
]
```

Zadanie 3.

Na podstawie *DATA* z *Listing 3*. wyświetl na ekranie nazwy gatunków zaczynające się na "v".

Listing 3.

```
DATA = [
    ('Sepal length', 'Sepal width', 'Petal length', 'Petal width', 'Species'),
    (5.8, 2.7, 5.1, 1.9, {'species': 'virginica'}),
    (5.1, 3.5, 1.4, 0.2, {'species': 'setosa'}),
    (5.7, 2.8, 4.1, 1.3, {'species': 'versicolor'}),
    (6.3, 2.9, 5.6, 1.8, {'species': 'virginica'}),
    (6.4, 3.2, 4.5, 1.5, {'species': 'versicolor'}),
    (4.7, 3.2, 1.3, 0.2, {'species': 'setosa'}),
    (7.0, 3.2, 4.7, 1.4, {'species': 'versicolor'}),
    (7.6, 3.0, 6.6, 2.1, {'species': 'virginica'}),
    (4.6, 3.1, 1.5, 0.2, {'species': 'setosa'}),
]
```

Zadanie 4.

Iterując po *DATA* z *Listing 4*. wygeneruj zbiór unikalnych kluczy słowników. Wynik zapisz do pliku w formacie *JSON*.

Listing 4.

```
DATA = [
    {'Sepal length': 5.1, 'Sepal width': 3.5, 'Species': 'setosa'},
    {'Petal length': 4.1, 'Petal width': 1.3, 'Species': 'versicolor'},
    {'Sepal length': 6.3, 'Petal width': 1.8, 'Species': 'virginica'},
    {'Petal length': 1.4, 'Petal width': 0.2, 'Species': 'setosa'},
    {'Sepal width': 2.8, 'Petal length': 4.1, 'Species': 'versicolor'},
    {'Sepal width': 2.9, 'Petal width': 1.8, 'Species': 'virginica'},
]
```

Data:

Imię i nazwisko:

Zadanie 5A.

Stwórz klasę *Iris* z polami: *sepal_length*, *sepal_width*, *petal_length*, *petal_width*. Stwórz klasy *Virginica*, *Versicolor*, *Setosa*, które będą dziedziczyły po *Iris*. Klasy mają mieć pole *species* ustawionym odpowiednio do nazwy klasy.

Zadanie 5B.

Napisz metody *total()* oraz *average()*, które będą wyliczały odpowiednio sumę lub średnią z wszystkich pomiarów *sepal_length*, *sepal_width*, *petal_length*, *petal_width* dla danego obiektu.

Zadanie 5C.

Iterując po *DATA* z *Listing 5*. twórz obiekty klasy odpowiedniej dla nazwy gatunku (ostatni rekord każdej z krotek). Obiekt inicjalizuj danymi z pomiarów. Na ekranie wyświetlaj nazwę gatunku oraz sumę i średnią z pomiarów.

Listing 5.

```
DATA = [  
    ('Sepal length', 'Sepal width', 'Petal length', 'Petal width', 'Species'),  
    (5.8, 2.7, 5.1, 1.9, 'virginica'),  
    (5.1, 3.5, 1.4, 0.2, 'setosa'),  
    (5.7, 2.8, 4.1, 1.3, 'versicolor'),  
    (6.3, 2.9, 5.6, 1.8, 'virginica'),  
    (6.4, 3.2, 4.5, 1.5, 'versicolor'),  
    (4.7, 3.2, 1.3, 0.2, 'setosa'),  
    (7.0, 3.2, 4.7, 1.4, 'versicolor'),  
    (7.6, 3.0, 6.6, 2.1, 'virginica'),  
    (4.9, 3.0, 1.4, 0.2, 'setosa'),  
    (4.6, 3.1, 1.5, 0.2, 'setosa'),  
]
```

Zadanie 6A.

Wykorzystując *numpy* dla ziarna losowości równego 0 wygeneruj macierz o nazwie *A*, 50x50 losowych liczb całkowitych z zakresu od 0 do 1024 włącznie. Stwórz macierz *B*, która będzie zawierała liczby z macierzy *A* będące potęgami dwójki. Pozostaw tylko i wyłącznie unikalne wartości. Uporządkuj macierz *B* w kolejności malejącej (od największej do najmniejszej).

Zadanie 6B.

Wczytaj macierz *A* za pomocą *pandas* i:

- A) podmień wartości większe lub równe 10 na *NaN*,
- B) usuń wiersz, jeżeli wszystkie wartości będą zawierały *NaN* (rezultat: *DataFrame* 17x50),
- C) Wypełnij wszystkie *NaN* w wynikowym *DataFrame* zerami,
- D) Wylicz podstawowe statystyki opisowe dla danych.

Zadanie 7.

Za pomocą *beautifulsoup4* ze strony <https://github.com/AstroMatt/book-python/blob/master/numerical-analysis/data/iris-dirty.csv> pobierz zbiór Irysów. Parsując kod *HTML* oczyść dane tak, aby nadawały się do obróbki w *pandas*. Następnie skasuj pierwszy wiersz, nazwij kolumny: *Sepal length*, *Sepal width*, *Petal length*, *Petal width*, *Species* i wylicz statystyki opisowe dla *DataFrame* powstałego po odfiltrowaniu wyników posiadających *Sepal length* mniejsze od 6 cm.