

Zaawansowane programowanie w Języku Python

Praca samodzielna, Projekt 1

Temat: Przetwarzanie i prezentacja danych, Analiza pliku z danymi NOP udostępnianymi przez serwis RP.

Na [podstronach serwisu www.gov.pl](http://podstronach.serwisu.www.gov.pl) można znaleźć raport dotyczący aktualnego stanu „niepożądanych odczynów poszczepiennych” (NOP) dotyczących szczepień na COVID-19 w Polsce. Spis jest udostępniony w pliku docx a jego zawartość jest kiepsko usystematyzowana. W związku z tym na potrzeby projektu przygotowałem plik nopy160221.json (dostępny na moodle), który jest przerobioną i poprawioną wersją pliku docx. Poprawiłem nazwy województw, usunąłem niepotrzebne znaki nowych linii i poprawiłem wiele innych drobiazgów.

Plik proponuję wczytać w taki sposób:

```
import json
with open('160221.json', 'r') as f:
    data = json.load(f)
```

Zadania:

1 (2pkt) Wczytaj dane z pliku nopy160221.json do Twojego programu.

Celem jest przechowywanie wszystkich NOPów na jednej liście w programie. Ma to być lista Twoich własnych obiektów.

Proszę utworzyć klasę, która będzie reprezentowała „NOP”. Wymagane pola:

- data
- województwo
- powiat
- płeć
- objawy

Uwaga! Dla uproszczenia uznajemy następującą mapę objawów:

```
"Zaczerwienienie i bolesność" => 'zaczerwienienie i krótkotrwała bolesność'
"Gorączka" => 'temp', 'temperatura', 'gorączka'
"Drgawki" => 'drgawki'
"Wymioty" => 'wymioty'
"Omdlenie" => 'omdlenie, utrata przytomności'
```

Dodatkowe informacje:

Dobrym pomysłem może być utworzenie klasy reprezentującej objaw i trzymanie w klasie NOP listy objawów. Ale to jest tylko propozycja, nie jest to obowiązkowe.

W klasie można trzymać dowolne inne informacje, z zaznaczeniem, że obowiązkowe są tylko te powyżej.

2 (1 pkt) Proszę uzyskać następujące informacje:

- Ile łącznie NOPów zostało wczytanych.
- Ile z nich dotyczy kobiet.
- Ile z nich dotyczy mężczyzn.

Wartości liczbowe należy wprowadzić do pliku odpowiedzi.txt.

3 (1 pkt) Proszę wyznaczyć liczbę osób, która w województwie pomorskim miała gorączkę.

Odpowiedź proszę podać do pliku odpowiedzi.txt

4 (1,5 pkt) Proszę przygotować zestawienie przedstawiające ile NOPów pojawiło się w każdym z szesnastu województw.

dolnośląskie: 37

kujawsko-pomorskie: 45

lubelskie: 21

itd.

Swoje zestawienie proszę przekleić do pliku zestawienie.txt

5 (1,5 pkt) Statystyka w województwie:

Dla województwa wskazywanego przez Twój numer indeksu (na końcu dokumentu znajduje się instrukcja jak je wyznaczyć) przygotuj statystykę ilościową dla objawów wypisanych w zadaniu pierwszym. Statystyka ma być posortowana od najczęściej występującego objawu. Statystykę proszę zapisać do pliku <województwo>.txt.

6 (1 pkt) Wykres kołowy:

Proszę przygotować wykres kołowy, na którym będzie widać:

- Procent mężczyzn, którzy doświadczyli NOP.
- Procent kobiet, które doświadczyły NOP.

Proszę dobrać inne kolory dla kobiet i mężczyzn dla odróżnienia. Proszę ustawić tytuł wykresu. Na wykresie mają być widoczne wartości procentowe.

7 (1 pkt) Wykres słupkowy:

Proszę przygotować wykres słupkowy, na którym ilościowo zostaną przedstawione wszystkie objawy wypisane w zadaniu 1 ("Zaczerwienienie i bolesność", "Gorączka", "Drgawki", "Wymioty", "Omdlenie"). Tak więc 5 słupków na osi X, a na osi Y ilość. Proszę ustawić tytuł wykresu, opisać osie, zadbać o estetykę wykresu.

8 (1 pkt) Wykres liniowy z dwiema seriami:

Proszę przygotować wykres liniowy, na którym będzie przedstawiona zmiana ilości NOPów w czasie. Na osi X daty (dni), na osi Y ilość NOPów danego dnia. Wykres proszę przygotować z dwiema seriami – osobną dla mężczyzn, osobną dla kobiet. Serie mają być oznaczone osobnymi kolorami. Wymagana jest legenda. Proszę ustawić tytuł wykresu, opisać osie, zadbać o estetykę wykresu.

Wykresy proszę zapisać za pomocą metody `savefig('wykres_zadanie_X.png')`

Jak oddać zadanie:

- Zadanie przysyłamy na moodle do odpowiedniej sekcji (Projekt 1)
- Należy przysłać ARCHIWUM **ZIP**. Nazwa pliku to <twój_numer_indeksu>.zip
- W archiwum mają znaleźć się:
 - Katalog „KOD”, a w nim:
 - Pliki .py, które napisaliśmy
 - Plik requirements.txt jeśli korzystaliśmy z zewnętrznych modułów
 - W katalogu **nie mogą znaleźć się**: wirtualne środowiska, pliki pyc, katalogi .idea itp.
 - Plik odpowiedzi.txt (zadanie 2 i 3)
 - Plik zestawienie.txt (zadanie 4)
 - Plik <województwo>.txt (np. dolnośląskie.txt) (zadanie 5)
 - Pliki wykres_zadanie_6.png, wykres_zadanie_7.png, wykres_zadanie_8.png

- Proszę przysłać przynajmniej jeden screen na którym widać, że kod z katalogu KOD się uruchamia.

Przysłanie plagiatu -10 pkt.

Oddajemy swój kod. Będzie wyrywkowo sprawdzane.

Przysłanie pliku odpowiedzi.txt, zestawienie.txt albo wykres*, który nie ma przełożenia na kod to oszustwo.

Detale implementacyjne:

- Do rysowania wykresów proszę użyć modułu matplotlib.
- Pliki z odpowiedziami (.txt) nie muszą być generowane przez kod programu. Mogą być przeklejone z outputu ręcznie.

Wyznaczenie swojego województwa:

Weź dwie ostatnie cyfry swojego numeru indeksu. Powstałą liczbę podziel modulo 16. Tak uzyskana reszta z dzielenia wskaże województwo:

dolnośląskie
kujawsko-pomorskie
lubelskie
lubuskie
mazowieckie
małopolskie
opolskie
podkarpackie
podlaskie
pomorskie
warmińsko-mazurskie
wielkopolskie
zachodniopomorskie
łódzkie
śląskie
świętokrzyskie