

Харківський національний університет імені В. Н. Каразіна
Факультет комп'ютерних наук
Кафедра штучного інтелекту та програмного забезпечення

ЗВІТ
З ЛАБОРАТОРНОЇ РОБОТИ №3

дисципліна: «Крос-платформне програмування»

Виконав: студент групи КС21
Барбінов Віктор Андрійович

Перевірів: доцент кафедри ШІтаПЗ
Споров Олександр Євгенович

Харків
2024

Завдання №0

Із сайту з відкритими даними (<https://catalog.data.gov/dataset/popular-baby-names>) було отримано свіжий (від 3 березня, 2023), великий за розміром датасет в XML форматі з інформацією про популярні імена дітей у місті Нью-Йорк. Цей датасет складений за офіційною інформацією із служби реєстрації актів цивільного стану міста Нью-Йорка. Архів з цим датасетом має назву `Popular_Baby_Names_NY.zip` та розміщений в лекційному Гугл-класі в розділі Методичні вказівки з виконання лабораторних робіт. Кожен запис цього датасету представляє інформацію про дитину: вказано дату народження, гендер, етнічну приналежність мами, власне ім'я дитини, кількість (count) дітей з цим іменем та рейтинг (rating) імені у відповідній групі. Потрібно провести попередній аналіз цих даних та вибрати з них лише потрібну для подальшої роботи інформацію. Виконати наступні завдання:

- Написати програму для виведення на екран частини XML документу за допомогою SAX парсеру без валідації для вивчення його структури та вмісту; програмно отримати перелік всіх тегів, імена яких присутні в документі.
- За невеликим характерним фрагментом скласти xsd схему документу, створити валідатор та перевірити, чи правильно було зрозуміло структуру документу.
- Написати програмне рішення, що за допомогою SAX парсеру без валідації отримає назви всіх національних груп, що представлені в документі.
- Написати додаток, що з всього XML документу вибирає задану кількість найбільш популярних імен в заданій етнічній групі із зберіганням інформації про: ім'я, гендер, кількість імен та рейтинг імен, а також створює відповідні Java об'єкти для зберігання цієї інформації та сортує інформацію по збільшенню номеру в рейтингу. Зберегти вибрану та відсортовану інформацію до нового XML файлу за допомогою DOM парсеру.
- Прочитати цей новий документ за допомогою DOM парсеру та вивести інформацію, що в ньому зберігається, на екран.

Результати виконання завдання №1:

```
"C:\Program Files\Java\jdk-19\bin\java.exe" "-javaagent:C:\Program Files\JetBrains\Intel
Ter: response
Ter: row
Ter: row
Атрибути:
- _id: row-v4f5~xz3v~vr86
- _uuid: 00000000-0000-0000-E0ED-52E8592E0A7A
- _position: 0
- _address: https://data.cityofnewyork.us/resource/\_25th-nujf/row-v4f5~xz3v~vr86
Ter: brth_yr
Ter: gndr
Ter: ethcty
Ter: nm
Ter: cnt
Ter: rnk
Ter: row
Атрибути:
- _id: row-gdep~mr7x~dj3u
- _uuid: 00000000-0000-0000-68AD-7B741D1DF31B
- _position: 0
- _address: https://data.cityofnewyork.us/resource/\_25th-nujf/row-gdep~mr7x~dj3u
Ter: brth_yr
Ter: gndr
Ter: ethcty
Ter: nm
Ter: cnt
Ter: rnk
Ter: row
Атрибути:
```

Рисунок 1 – Результат виконання програми

```
Ter: brth_yr
Ter: gndr
Ter: ethcty
Ter: nm
Ter: cnt
Ter: rnk
Ter: row
Атрибути:
- _id: row-udhd_mkva_t8sv
- _uuid: 00000000-0000-0000-6871-0A34F7C891CC
- _position: 0
- _address: https://data.cityofnewyork.us/resource/\_25th-nujf/row-udhd\_mkva\_t8sv
Ter: brth_yr
Ter: gndr
Ter: ethcty
Ter: nm
Ter: cnt
Ter: rnk
Ter: row
Атрибути:
- _id: row-p6z5-9gbz~59v6
- _uuid: 00000000-0000-0000-2F97-2765BEFE798C
- _position: 0
- _address: https://data.cityofnewyork.us/resource/\_25th-nujf/row-p6z5-9gbz~59v6
Досягнуто ліміту елементів
Process finished with exit code 0
```

Рисунок 2 – Результат виконання програми

Результати виконання завдання №2:

```
Checking document Popular_Baby_Names_NY.xml validation in accordance with Popular_Baby_Names_NY.xsd
The document is confirmed normally.
```

```
Process finished with exit code 0
```

Рисунок 3 – Результат виконання програми

Результати виконання завдання №3:

```
-----
Ethnic groups in alphabetical order:
ASIAN AND PACI
ASIAN AND PACIFIC ISLANDER
BLACK NON HISP
BLACK NON HISPANIC
HISPANIC
WHITE NON HISP
WHITE NON HISPANIC
-----
```

```
Ethnic groups found in the dataset and their counts:
WHITE NON HISPANIC: 14019
BLACK NON HISP: 1740
WHITE NON HISP: 3441
HISPANIC: 14530
BLACK NON HISPANIC: 7213
ASIAN AND PACI: 1767
ASIAN AND PACIFIC ISLANDER: 6799
-----
```

```
Total entries: 49509
-----
```

```
Process finished with exit code 0
```

Рисунок 4 – Результат виконання програми

Результати виконання завдання №4:

```
All 500 names were found
Babie{gender='MALE', name='JAYLEN', nums=50, rating=100}
Babie{gender='MALE', name='DEAN', nums=50, rating=100}
Babie{gender='MALE', name='JAYREN', nums=50, rating=100}
Babie{gender='MALE', name='JONAS', nums=50, rating=100}
Babie{gender='MALE', name='KAIDEN', nums=50, rating=100}
Babie{gender='MALE', name='ROMAN', nums=50, rating=100}
Babie{gender='MALE', name='ELVIN', nums=50, rating=100}
Babie{gender='MALE', name='ADAN', nums=55, rating=99}
Babie{gender='MALE', name='MATTEO', nums=55, rating=99}
Babie{gender='MALE', name='SKYLER', nums=55, rating=99}
Babie{gender='MALE', name='IZAIAH', nums=55, rating=99}
Babie{gender='MALE', name='Marc', nums=10, rating=99}
Babie{gender='MALE', name='Francis', nums=10, rating=99}
Babie{gender='MALE', name='Ronaldo', nums=10, rating=99}
Babie{gender='MALE', name='CAMILO', nums=60, rating=98}
Babie{gender='MALE', name='EDDY', nums=60, rating=98}
Babie{gender='MALE', name='SAUL', nums=60, rating=98}
```

Рисунок 5 – Результат виконання програми

```
Babie{gender='MALE', name='Dylan', nums=4204, rating=2}
Babie{gender='MALE', name='Jacob', nums=3740, rating=2}
Babie{gender='FEMALE', name='Sophia', nums=3349, rating=2}
Babie{gender='MALE', name='Noah', nums=3341, rating=2}
Babie{gender='MALE', name='JACOB', nums=2980, rating=2}
Babie{gender='FEMALE', name='Emma', nums=2788, rating=2}
Babie{gender='MALE', name='JUSTIN', nums=2625, rating=2}
Babie{gender='FEMALE', name='MIA', nums=2180, rating=2}
Babie{gender='FEMALE', name='SOPHIA', nums=2059, rating=2}
Babie{gender='MALE', name='Liam', nums=4300, rating=1}
Babie{gender='FEMALE', name='Isabella', nums=3972, rating=1}
Babie{gender='MALE', name='Jayden', nums=3868, rating=1}
Babie{gender='MALE', name='JAYDEN', nums=3160, rating=1}
Babie{gender='FEMALE', name='Mia', nums=3084, rating=1}
Babie{gender='FEMALE', name='ISABELLA', nums=2963, rating=1}

Process finished with exit code 0
```

Рисунок 6 – Результат виконання програми

Результати виконання завдання №5:

Стать : MALE
Ім'я : Liam
Кількість : 4300
Рейтинг : 1

Стать : FEMALE
Ім'я : Isabella
Кількість : 3972
Рейтинг : 1

Стать : MALE
Ім'я : Jayden
Кількість : 3868
Рейтинг : 1

Стать : MALE
Ім'я : JAYDEN
Кількість : 3160
Рейтинг : 1

Стать : FEMALE
Ім'я : Mia
Кількість : 3084
Рейтинг : 1

Стать : FEMALE
Ім'я : ISABELLA
Кількість : 2963
Рейтинг : 1

Рисунок 7 – Зчитування документа за допомогою DOM парсеру