Національний технічний університет України

«Київський політехнічний інститут імені Ігоря Сікорського»

Кафедра автоматизації проектування

 енергетичних процесів і систем

**Звіт**

**з циклу лабораторних робіт**

**з дисципліни**

**„Обробка потокової інформації”**

Лабораторна  робота № 5

Використання алгоритму Флажоле-Мартена

Виконала: **Юрченко Богдана**

Студентка групи:   **ТВ-01 мп**

Перевірив викладач: **Іван Матичин**

Київ - 2021

Завдання:

1. Згенерувати нескінчений потік чисел, що можуть повторюватися.

2. Реалізувати алгоритм Флажоле-Мартена для оцінки кількості різних елементів

потоку.

3. Порівняти результат, отриманий на попередньому кроці, з реальною кількістю

різних елементів потоку.

В рамках даної роботи було зроблено :

змінено код консьюмера, розробленого в рамках лабораторної роботи 2.

Кроки виконання :

1. Були запущені контейнери із zookeeper i kafka за допомогою докеру та докер-компоуз (аналогічно до першої роботи).

Виконання команд для створення топіку users\_auth.

kafka-topics.sh --create --zookeeper zookeeper:2181 --replication-factor 1 --partitions 1 --topic users\_auth

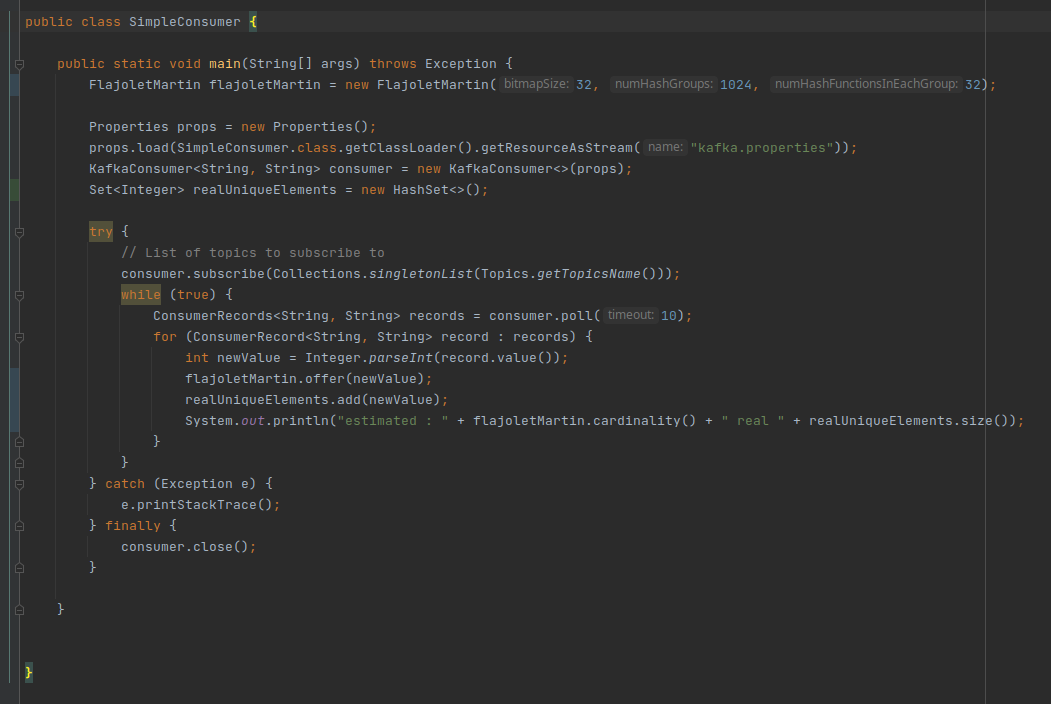
Код створення топіку із числами, що можуть повторюватися :



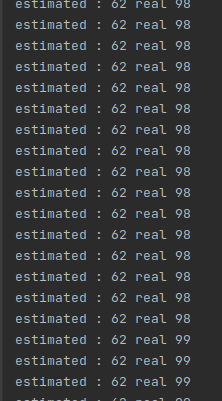
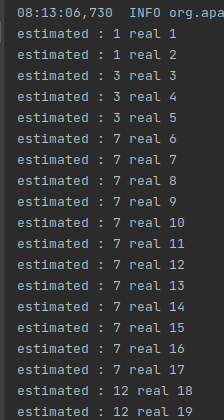
Перероблений продюсер, що відсилає в топік із числами випадкові числа від 0 до 100 (які будуть повторюватися із часом) :



Код консьюмера, що одначасно використовує алгоритм Флажоле-Мартена (уже готову реалізацію) та простий метод використання хеш-сетів (для визначення реальної кількості елементів).



Результат роботи консьюмера



Як видно зі скріншотів, на початку точність доволі таки висока. Із ростом реальної кількості елементів, точність зменшуєтьс.

Після виконання роботи - зупинити контейнери

