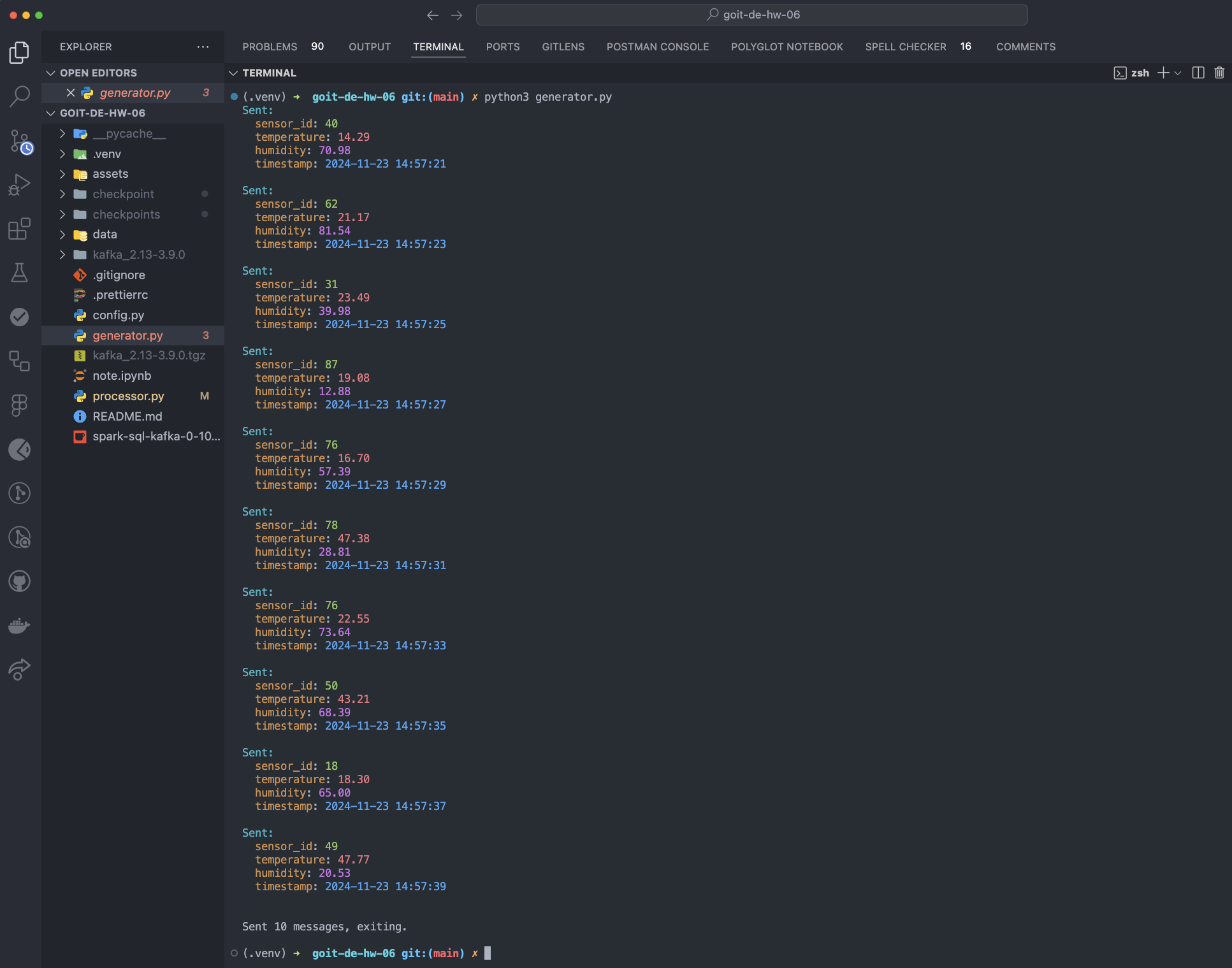
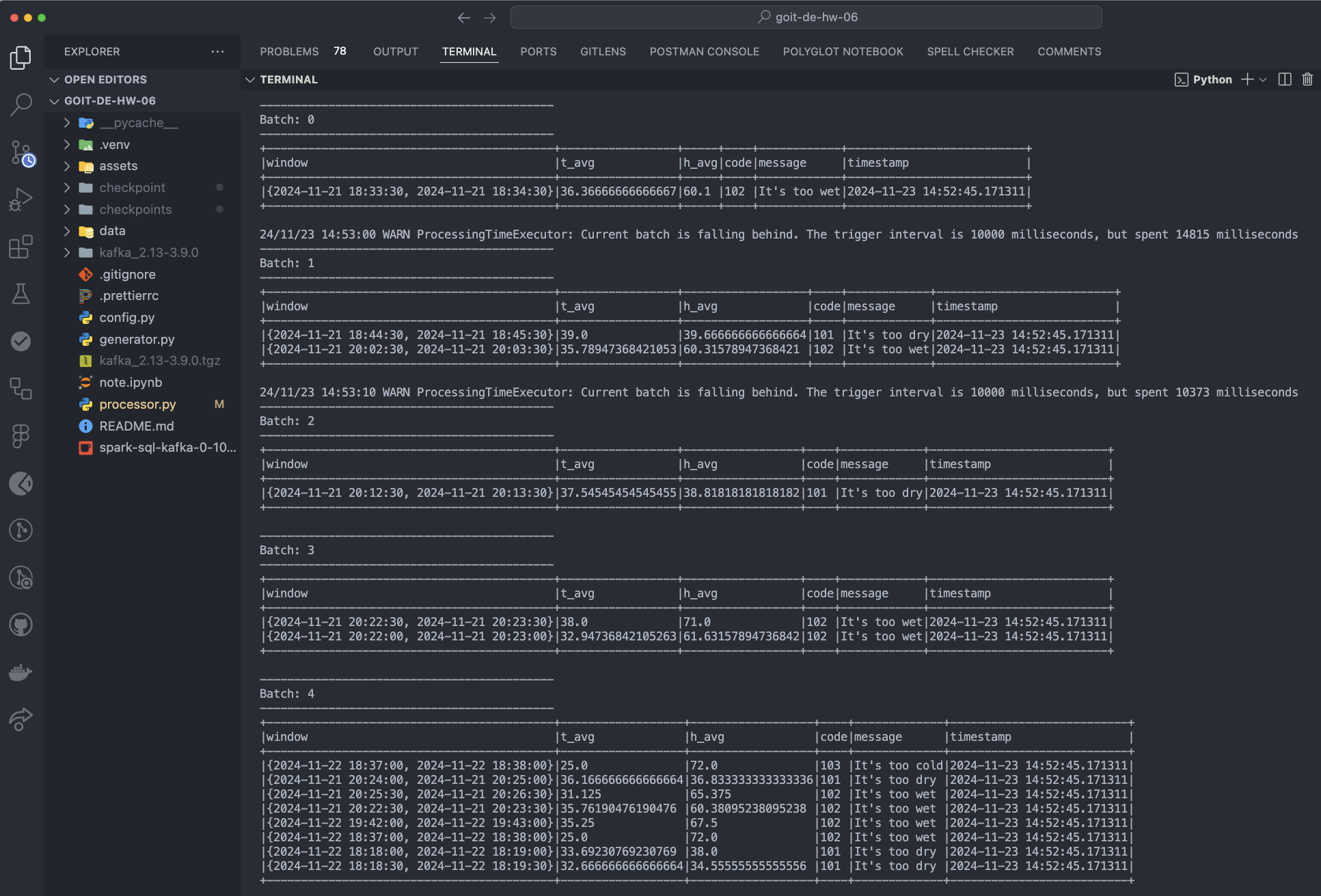
**Домашнє завдання до теми «Spark streaming»  
  
Опис виконання завдання:**

1. Генерація та відправка даних сенсорів у Kafka (generator.py):
   * За допомогою модуля confluent\_kafka створено Kafka Producer для відправки згенерованих даних про сенсори в Kafka.
   * Дані включають: sensor\_id, температуру, вологість і час вимірювання.
   * Генерація даних відбувається в циклі, де кожного разу генерується новий набір значень для кожного сенсора.
   * Виведення в консоль відбувається з використанням бібліотеки colorama, яка забарвлює текст для кращої візуалізації.



1. Обробка даних за допомогою PySpark (processor.py):
   * Використано Apache Spark для читання та обробки потоку даних з Kafka, з підключенням через конфігурацію Kafka, що включає SASL механізм для безпечної передачі.
   * Дані з потоку декодуються з JSON і обробляються для отримання середніх значень температури та вологості в певних вікнах часу.
   * Потім дані об’єднуються з умовами для алертів, що задані в окремому CSV файлі (alerts\_conditions.csv), і фільтруються за заданими порогами для температури та вологості.
   * Результати виводяться в консоль у вигляді форматованого JSON для кожного вікна часу, що дозволяє відслідковувати порогові значення для алертів.
   * Дані, що пройшли фільтрацію, готуються для відправки в Kafka.
2. Перевірка коректності роботи системи:
   * Створено кілька записів у Kafka, які демонструють правильність роботи генерації та обробки даних.
   * Виведення результатів фільтрації здійснюється через консоль, де видно, чи відповідають середні значення температури та вологості заданим умовам.

****