




Tarea módulo data streaming

 Add comment

Objetivo

El objetivo de esta tarea es fijar los contenidos desarrollados durante el módulo de Data Streaming, incluyendo la aplicación de herramientas como Redpanda y KSQLDB para la solución de un problema.

Instrucciones

La tarea consiste en construir un servicio que consuma la API en tiempo real de Finnhub: <https://finnhub.io/docs/api/websocket-trades> y consuma actualizaciones para los siguientes símbolos:

- AAPL
- AMZN
- BINANCE:BTCUSDT

De modo a procesar las actualizaciones, deben seguirse los siguientes pasos:

1. Instalar un cluster Redpanda de manera local utilizando un archivo *docker-compose.yml*
2. Implementar un producer utilizando [kafka-python](#), de acuerdo a la [documentación](#) de Redpanda y similar al ejemplo desarrollado en clase, para suscribirse a los eventos de la API.
3. Instalar [KSQLDB](#) modificando el archivo *docker-compose.yml*, de acuerdo a la [documentación](#) de Redpanda.
4. Ejecutar ksqldb-cli, definir los *streams* y *tablas* necesarios para responder a las siguientes preguntas:
 1. ¿Cuál fue el promedio ponderado de precio de una unidad por cada uno de los símbolos procesados? (e.j. AAPL)
 2. ¿Cuántas transacciones se procesaron por símbolo?
 3. ¿Cuál fue el máximo precio registrado por símbolo?

4. ¿Cuál fue el mínimo precio registrado por símbolo?

Presentación

El proyecto debe desarrollarse en los grupos previamente formados. El entregable debe incluir:

- Enlace a repositorio público de Github incluyendo:
 - docker-compose.yml
 - código fuente del consumer
 - archivos .sql con las 4 consultas utilizadas para responder las preguntas de la sección anterior
- Una grabación de pantalla mostrando el sistema en funcionamiento y explicando brevemente los pasos que se siguieron para su implementación.

Fecha límite de entrega: 21-08-2023 23:59

