

# Ejercicio 1: Consumidor / Productor básico

---

## Productor

---

Realizar un programa en Python / Scala que simule eventos de sondas. El programa debe producir

un string en formato CSV en un topic de kafka cada cierto tiempo, por ejemplo 1 segundo

Ejemplo en formato CSV:

```
17,40.48184,-3.364497
20,40.48184,-3.364497
26,23.1234,-8.364497
```

## Consumidor.

---

Realizar un consumidor que, a partir de una llamada a un servicio web, consulte los elementos del un topic indicado a partir del path de la llamada

Ejemplo:

- Llamada `curl http://localhost:8080/this-is-a-topic`
- Topic a consultar: `this-is-a-topic`
- Resultado de la petición:

```
[
  {"temperatura":20,"location":{"latitud": 40.48184,"longitud": -3.364497}},
  {"temperatura":17,"location":{"latitud": 40.48184,"longitud": -3.364497}},
  {"temperatura":26,"location":{"latitud": 23.1234,"longitud": -8.364497}}
]
```

Nota: Hay un ejemplo de un servidor web en `ejercicios-py/ejercicios/finales/http_server.py`

## Ejercicio 2: KSQL

Dado los siguientes topics de kafka que reciben mensajes con el siguiente formato:

users:

```
{ "id": "1234", "userName": "aUser1", "country": "es" }
{ "id": "4312", "userName": "aUser2", "country": "en" }
{ "id": "1233", "userName": "aUser3", "country": "jp" }
{ "id": "1", "userName": "aUser4", "country": "es" }
{ "id": "3412", "userName": "aUser5", "country": "es" }
{ "id": "223", "userName": "aUser5", "country": "co" }
```

messages:

```
{ "id": 1504451510296, "userId": "1234", "text": "esto es un mensaje" }
{ "id": 1494273403520, "userId": "1233", "text": "hola mundo" }
{ "id": 1494273403520, "userId": "1233", "text": "hello world" }
{ "id": 1505275563904, "userId": "223", "text": "Mensaje del usuario 223" }
```

Buscamos enriquecer el contenido de `messages` con el `userName` y el `country`

Preguntas:

- 1) En qué tipo de estructura deberíamos almacenar usuarios y mensajes (KTable o KStream)
- 2) ¿Cuál es la query para crear las estructuras?
- 3) ¿Cuál sería la consulta para obtener en un topic los mensajes enriquecidos?

## Ejercicio 3: Spark Structured Streaming

Dado un topic de kafka con el siguiente tipo de entradas

```
1
234
1234
12
1
234
```

123  
412  
341  
234  
123  
41  
234  
123  
41  
234  
1234  
123  
41  
234  
1341  
34  
1234  
123  
4

Se debe hacer un programa que calcule la media de los números que hay en el topic.

## Proyecto

---

Exponer en un documento un escenario de aplicación en donde exista procesamiento streaming y batch.

Explicar varios casos de uso implicados la aplicación y justificar por qué hay aplicar o no streaming en ellos.

Se debe entregar un documento con los siguientes apartados

- 1) Exposición de la aplicación. Debe incluir una explicación general.
- 2) Detalle de las distintas piezas que forman parte del caso de uso.
- 3) Detalle de varios casos de uso de la aplicación, justificando por qué se aplica o no streaming para cada caso de uso.