

Apache Kafka'da Veri İşleme Uygulaması

Apache Kafka'da Veri İşleme Uygulaması

Adım 1: Kafka üzerinde 2 parçaya sahip ve replikasyon faktörü 1 olan **atscale** adında bir topic oluşturunuz.

Adım 2: Kafka üzerindeki mevcut topicleri listeleyiniz.

Adım 3: **atscale** topic özelliklerini ekrana yazdırınız.

Adım 4: data-generator kullanarak **Churn_Modelling.csv** veri setini Kafka **churn** topic'e aşağıdaki koşulları sağlayarak produce/consume ediniz.

- Kafkaya gönderdiğiniz mesajın anahtarı **CustomerId** olsun.
- Üç farklı consumer'a sahip **churn_group** adındaki bir consumer group altında mesajları consume ediniz.
- Her bir consumer için farklı bir terminal kullanınız.
- Consumer olarak **kafka-console-consumer.sh** kullanınız.
- Tüm veri setinin bitmesini beklemeyiniz. Yaklaşık 500 satırı gözlemleyiniz.

Adım5: atscale ve **churn** topic'lerini siliniz.

Apache Kafka'da Veri İşleme Uygulaması

Adım 6: Python Kafka client kullanarak aşağıdaki işleri yapınız:

- Producer tarafında Türkiye'nin **coğrafi bölgelerinin adlarını** her birinin başında belirleyeceğiniz rakamları **key** olarak kullanınız. Örneğin 1 Marmara, 2 Ege şeklinde.
- Consumer tarafında ekranda beklenen çıktı (**key, value, partition, timestamp**) aşağıdaki gibi olmalıdır.

Key: 1, Value: Marmara, Partition: 0, TS: 1613224639352

Key: 4, Value: İç Anadolu, Partition: 1, TS: 1613224654849

Key: 3, Value: Akdeniz, Partition: 2, TS: 1613224661486

Key: 2, Value: Ege, Partition: 2, TS: 1613224667044

Not: Data generator indirmek için aşağıdaki komutu kullanabilirsiniz.

git clone https://github.com/erkansirin78/data-generator.git

miuul

miuul.com