

Project Brief: Data Streaming with Kafka

Background: Telecommunications Mobile Money Data Engineering with Kafka

In this project, you will work with telecommunications mobile money data to build a Kafka data engineering solution. You will be provided with a dummy json file containing sample data that you will use to test your solution.

The project aims to build a Kafka pipeline that can receive real-time data from telecommunications mobile money transactions and process it for analysis. The pipeline should be designed to handle high volumes of data and ensure that the data is processed efficiently.

To complete this project, you will need to follow these steps:

1. **Set up a Kafka cluster:** You must set up a Kafka cluster that can handle high volumes of data. You can use either a cloud-based or on-premises Kafka cluster.
2. **Develop a Kafka producer:** You must develop a Kafka producer that can ingest data from telecommunications mobile money transactions and send it to the Kafka cluster. The producer should be designed to handle high volumes of data and ensure that the data is sent to the Kafka cluster efficiently.
3. **Develop a Kafka consumer:** You must develop a Kafka consumer to receive data from the Kafka cluster and process it for analysis. The consumer should be designed to handle high volumes of data and ensure that the data is processed efficiently.
4. **Process the data:** Once you have set up the Kafka pipeline, you must process the data for analysis. This may involve cleaning and aggregating the data, performing calculations, and creating visualizations.
5. **Test the solution:** You must test your solution using the provided dummy json file. The file contains sample data that you can use to ensure that your Kafka pipeline is working correctly.

Here's the dummy JSON file that represents our mobile money data.

```
{
  "transaction_id": "12345",
  "sender_phone_number": "256777123456",
  "receiver_phone_number": "256772987654",
  "transaction_amount": 100000,
  "transaction_time": "2023-04-19 12:00:00"
}
```