The `TelcoTowerDataV14.py` script connects to a Vertica database to retrieve tower location data, simulating real-time data related to telecom tower locations. This data can then be used for various analyses, such as tracking and monitoring tower activity within specific geographic bounds. Here's a combined README for both the `TelcoCDRDataKafkaV11.py` and `TelcoTowerDataV14.py` files.

README for Telecom Data Simulation Programs

Overview

This repository contains two Python scripts that simulate telecom data streams: `TelcoCDRDataKafkaV11.py` and `TelcoTowerDataV14.py`. These scripts are designed to generate realistic telecom data for testing and analytics purposes, publishing it to Kafka topics for real-time processing. Together, these programs simulate Call Detail Records (CDRs) and telecom tower locations within specified geographic boundaries, providing a rich data stream for telecom analytics or network monitoring applications.

Scripts

1. `TelcoCDRDataKafkaV11.py`

Description: This script generates and publishes synthetic Call Detail Records (CDRs) to a Kafka topic. Each record represents a telecom interaction (call, SMS, or data session) with metadata such as start time, end time, duration, carrier, type, and status.

- **Data Generation**: Uses the Faker library to create realistic CDRs, assigning values to fields like Call ID, timestamps, duration, carrier, and interaction type.
- **Data Publishing**: Publishes the generated records to a Kafka topic, structured according to an Avro schema for consistency.
- **Usage**:
  1. Ensure Kafka is running and accessible on the specified IP and port.
  2. Run the script with `python TelcoCDRDataKafkaV11.py`, and it will continuously generate CDR data and send it to Kafka.

Example Data:

- `CallID`: Unique identifier for each call or data session.
- `StartTime` / `EndTime`: Timestamps for session start and end.
- `Duration`: Duration in seconds.
- `Type`: Type of interaction (Voice, SMS, Data).

- `Carrier`: Name of the telecom carrier.
- `Status`: Interaction status (Answered, Missed, Dropped).

2. `TelcoTowerDataV14.py`

Description: This script connects to a Vertica database to retrieve and simulate telecom tower location data within the geographic bounds of Massachusetts. It retrieves tower information from the database and publishes it to a Kafka topic, simulating real-time tower activity for telecom network monitoring.

- **Database Connection**: Connects to a Vertica database to retrieve tower location data based on predefined geographic boundaries.
- **Geographic Filtering**: Restricts locations to Massachusetts, simulating data from towers within specific latitude and longitude ranges.
- **Data Publishing**: Publishes tower data, including coordinates and carrier information, to a Kafka topic for real-time tracking.
- **Usage**:
  1. Ensure Vertica is accessible with the correct connection details.
  2. Run the script with `python TelcoTowerDataV14.py`, and it will fetch and publish tower data within the Massachusetts bounds.

Example Data:

- `TowerID`: Unique identifier for each telecom tower.
- `Latitude` / `Longitude`: Coordinates of the tower.
- `Location`: Description of the tower's geographic location.
- `Carrier`: Telecom carrier operating the tower.

Prerequisites

- **Kafka**: Ensure Kafka and Schema Registry are accessible for data publishing.
- **Vertica**: Ensure the Vertica database is accessible for `TelcoTowerDataV14.py`.
- **Python Packages**: Install required packages:

```
pip install faker confluent-kafka avro-python3 vertica-python requests
```

Notes

- Modify `CARRIERS`, `STATUSES`, or `TYPES` constants in `TelcoCDRDataKafkaV11.py` to adjust the data for different telecom scenarios.
- Update database connection details as needed to connect to your environment.
