

To automate the synchronization of data for every single table in Hue with more complexity and comprehensiveness, you can create a Hue workflow that includes multiple actions, each action corresponding to a specific table. Each action will execute a Hue SQL query to export data from a table to a CSV file. Additionally, you can schedule the workflow to run periodically using Oozie coordinators for more automation.

Here's a detailed and comprehensive step-by-step guide:

### Step 1: Define SQL Queries for Each Table

Create SQL queries for each table that export data to CSV files. For example:

#### Patient Information Table

```
-- Export Patient Information to CSV
INSERT OVERWRITE DIRECTORY '/user/hue/data_sync/patient_info'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
SELECT * FROM PatientInformation;
```

#### Doctor Information Table

```
-- Export Doctor Information to CSV
INSERT OVERWRITE DIRECTORY '/user/hue/data_sync/doctor_info'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
LINES TERMINATED BY '\n'
SELECT * FROM DoctorInformation;
```

Repeat this for each table in your dataset.

### Step 2: Create a Hue Workflow

1. In Hue, navigate to the “Workflows” section.
2. Click on “Create Workflow” to start a new workflow.
3. Give your workflow a meaningful name, e.g., “Data Synchronization.”
4. Within the workflow, add individual actions for each table synchronization:
  - Action 1: Synchronize Patient Information Table
    - Type: Hive
    - Query: Insert the SQL query for exporting Patient Information.
    - Output Directory: Specify the HDFS path, e.g., `/user/hue/data_sync/patient_info`.
  - Action 2: Synchronize Doctor Information Table

- Type: Hive
- Query: Insert the SQL query for exporting Doctor Information.
- Output Directory: Specify the HDFS path, e.g., `/user/hue/data_sync/doctor_info`.
- Repeat this process for each table you want to synchronize.

### **Step 3: Configure Workflow Actions**

1. Configure the execution order of actions within the workflow, ensuring that they run sequentially.
2. Add error handling and notification actions to handle any issues that may occur during data synchronization.

### **Step 4: Schedule the Workflow with Oozie Coordinator**

1. To automate the workflow’s execution at regular intervals, set up an Oozie coordinator job. This coordinator job will trigger the workflow periodically.
2. Configure the coordinator job with a suitable frequency, such as daily or hourly, and specify the start and end times.
3. Ensure that the coordinator job triggers the “Data Synchronization” workflow.

### **Step 5: Save and Deploy**

1. Save the workflow and deploy it to your Hue environment.
2. The coordinator job will automatically trigger the “Data Synchronization” workflow based on your specified schedule, and each action within the workflow will export data from the corresponding table to CSV files.

This approach provides a more detailed, complex, and comprehensive solution for automating data synchronization for every single table in Hue. It includes error handling, scheduling, and organization of actions within the workflow to ensure data integrity and automation efficiency. Please adapt the SQL queries and paths according to your specific dataset and requirements.