Tables Used

- transaction_types
- merchant_categories
- age_groups
- states
- banks
- device_types
- Network_types
- Transaction_statuses
- fact_upi_transactions

Relationships

- fact_upi_transactions.transaction_type_id → transaction_types.type_id
- fact_upi_transactions.merchant_category_id → merchant_categories.category_id
- fact_upi_transactions.sender_age_group_id and receiver_age_group_id → age_groups.age_group_id
- fact_upi_transactions.sender_state_id → states.state_id
- fact_upi_transactions.sender_bank_id and receiver_bank_id → banks.bank_id
- fact_upi_transactions.device_type_id → device_types.device_type_id
- fact_upi_transactions.network_type_id → network_types.network_type_id
- fact_upi_transactions.transaction_status_id → transaction_statuses.status_id

Key Calculations / Measures

- Total Transactions – COUNT(fact_upi_transactions[transaction_id])
- Total Amount – SUM(fact_upi_transactions[amount_inr])
- Success Transactions – CALCULATE([Total Transactions], transaction_status = "Success")
- Success Rate % – DIVIDE([Success Transactions], [Total Transactions])

- Average Transaction Value – DIVIDE([Total Amount], [Total Transactions])
- Failed Transactions – CALCULATE([Total Transactions], transaction_status = "Failed")
- Top Merchant Category – TOPN(1, merchant_categories, [Total Amount])
- Top Performing State – Top Performing State

Notes

The transaction_date_text field in the raw data was converted into a proper date format using the SQL function TO_DATE(transaction_date_text, 'DD-MM-YY') to ensure consistency in date-based analysis.

All dimension tables were created by extracting distinct values from the upi_raw table. This approach eliminates duplication, ensures data normalization, and creates a clean, consistent reference for analytical queries.

The database schema uses foreign key constraints to maintain referential integrity between the fact table and dimension tables. This ensures that every record in the fact table corresponds to valid entries in the related dimension tables.

The fact_upi_transactions table is the central component of the model, functioning as the main fact table within a star schema design. It consolidates all transaction-level information and connects to multiple dimension tables, making it optimized for reporting and analytical purposes.

Several derived fields, such as is_weekend, day_of_week, and hour_of_day, were added to support time-based analysis. These fields allow for insights based on daily, hourly, and weekend transaction patterns.

The fraud_flag column is a Boolean indicator that helps identify potentially fraudulent transactions, supporting fraud detection and monitoring efforts.

Overall, this data model is designed for scalability and flexibility, allowing easy expansion with additional dimensions or calculated measures as future analytical needs evolve.

Diagram