# Smart Home Automation Database System - Technical Documentation

## Table of Contents

1. [OLTP Database Context](#oltp-database-context)
2. [OLAP Database Context](#olap-database-context)
3. [Database Schemas and Relationships](#database-schemas-and-relationships)
4. [Execution Instructions](#execution-instructions)
5. [Power BI Report Analysis](#power-bi-report-analysis)

## OLTP Database Context

### Purpose

Operational database for real-time transaction processing of smart home e-commerce platform.

### What We Store

* **Customer Data**: User accounts, addresses, contact information
* **Product Catalog**: Smart home devices, specifications, pricing, inventory
* **Order Management**: Customer orders, order items, payment processing
* **Shopping Cart**: Active cart sessions
* **Order Status Tracking**: Complete order lifecycle history
* **Payment Processing**: Transaction records, payment methods

### Key Tables (8 tables)

| Table | Records | Primary Purpose |
| --- | --- | --- |
| users | 40+ | Customer accounts across US states |
| products | 45+ | Smart home device catalog |
| orders | 70+ | Customer orders with full lifecycle |
| order\_items | 100+ | Order line items and quantities |
| payments | 70+ | Financial transaction processing |
| shopping\_cart | Active | Real-time cart management |
| categories | 40+ | Product categorization hierarchy |
| order\_status\_history | 200+ | Order status change tracking |

## OLAP Database Context

### Purpose

Analytical data warehouse for business intelligence and strategic decision making.

### Analytical Questions We Answer

* **Sales Performance**: Which products generate the most revenue?
* **Customer Segmentation**: Who are our most valuable customers?
* **Geographic Analysis**: Which regions have highest sales potential?
* **Seasonal Trends**: How do sales patterns change over time?
* **Inventory Optimization**: What are optimal stock levels for each product?
* **Manufacturer Performance**: Which suppliers provide best ROI?
* **Customer Lifetime Value**: What is the long-term value of customer segments?

### Key Components (10 tables)

| Component | Tables | Purpose |
| --- | --- | --- |
| **Fact Tables** | fact\_sales, fact\_inventory | Sales metrics and inventory tracking |
| **Dimensions** | dim\_customer, dim\_product, dim\_date, dim\_time, dim\_location, dim\_category, dim\_manufacturer | Analysis dimensions |
| **SCD Type 2** | dim\_customer, dim\_manufacturer | Historical change tracking |
| **Bridge Table** | bridge\_product\_category | Many-to-many relationships |

## Database Schemas and Relationships

### OLTP Schema (3NF Normalized)

-- Core Entity Relationships  
users (1) ──→ (M) orders  
orders (1) ──→ (M) order\_items  
products (1) ──→ (M) order\_items  
categories (1) ──→ (M) products  
orders (1) ──→ (1) payments  
orders (1) ──→ (M) order\_status\_history  
users (1) ──→ (M) shopping\_cart

### OLAP Schema (Snowflake)

-- Fact-Dimension Relationships  
fact\_sales ──→ dim\_customer (customer\_key)  
fact\_sales ──→ dim\_product (product\_key)  
fact\_sales ──→ dim\_date (date\_key)  
fact\_sales ──→ dim\_time (time\_key)  
fact\_sales ──→ dim\_location (location\_key)  
fact\_sales ──→ dim\_manufacturer (manufacturer\_key)  
  
-- Bridge Relationships  
dim\_product ←──→ bridge\_product\_category ←──→ dim\_category

### Key Constraints

* **Primary Keys**: All tables have surrogate keys
* **Foreign Keys**: Enforced referential integrity
* **Check Constraints**: Price >= 0, Quantity > 0
* **Unique Constraints**: Email addresses, product model numbers
* **Not Null**: Critical business fields

## Execution Instructions

### 1. System Setup

# Start Docker containers  
docker-compose up -d  
  
# Verify containers are running  
docker ps

### 2. OLTP Database Setup

# Create OLTP schema  
psql -h localhost -p 5434 -U postgres -d smart\_home\_shop\_oltp -f sql/oltp/01\_create\_tables.sql  
  
# Load operational data  
psql -h localhost -p 5434 -U postgres -d smart\_home\_shop\_oltp -f sql/oltp/02\_load\_data.sql

### 3. OLAP Database Setup

# Create OLAP schema  
psql -h localhost -p 5433 -U postgres -d smart\_home\_shop\_olap -f sql/olap/01\_create\_tables.sql  
  
# Execute ETL process  
psql -h localhost -p 5433 -U postgres -d smart\_home\_shop\_olap -f sql/olap/02\_etl\_process.sql

### 4. Dataset Loading Process

The system loads data from CSV files located in /data/ directory: - 01\_users.csv → users table - 02\_categories.csv → categories table  
- 03\_products.csv → products table - 04\_orders.csv → orders table - 05\_order\_items.csv → order\_items table

### 5. ETL Process Details

The ETL pipeline (sql/olap/02\_etl\_process.sql) performs: - **Extract**: Data from OLTP using postgres\_fdw - **Transform**: Create dimensions with SCD Type 2 for historical tracking - **Load**: Populate fact tables with business metrics - **Validate**: Data quality checks and integrity constraints

### 6. Analytics Execution

# Run OLTP operational queries  
psql -h localhost -p 5434 -U postgres -d smart\_home\_shop\_oltp -f sql/03\_oltp\_queries.sql  
  
# Run OLAP analytical queries  
psql -h localhost -p 5433 -U postgres -d smart\_home\_shop\_olap -f sql/04\_olap\_queries.sql

## Power BI Report Analysis

### File Location

documentation/Smart Home Automation Online-Shop.pbix (436KB)

### Key Visualizations and Their Purpose

#### 1. Revenue Trend Analysis

* **What it shows**: Monthly sales performance over time
* **Business value**: Identifies growth patterns and seasonal trends
* **Data source**: fact\_sales.total\_sales\_amount by dim\_date.month\_name

#### 2. Geographic Sales Distribution

* **What it shows**: Sales performance by US regions and states
* **Business value**: Identifies high-potential markets for expansion
* **Data source**: fact\_sales aggregated by dim\_location.state

#### 3. Top Products Ranking

* **What it shows**: Best-selling products by revenue and quantity
* **Business value**: Guides inventory planning and marketing focus
* **Data source**: fact\_sales.total\_sales\_amount grouped by dim\_product.product\_name

#### 4. Customer Segmentation Analysis

* **What it shows**: Customer distribution by purchase behavior
* **Business value**: Enables targeted marketing and customer retention
* **Data source**: dim\_customer with calculated fields for segment classification

#### 5. Category Performance Matrix

* **What it shows**: Revenue and profit margins by product category
* **Business value**: Identifies most profitable product lines
* **Data source**: fact\_sales joined with dim\_category through bridge table

#### 6. Inventory Health Dashboard

* **What it shows**: Stock levels, reorder points, and turnover rates
* **Business value**: Prevents stockouts and reduces carrying costs
* **Data source**: fact\_inventory with calculated KPIs

### Key Performance Indicators (KPIs)

* **Total Revenue**: $50,000+ from smart home sales
* **Active Customers**: 40+ across major US regions
* **Average Order Value**: $600-800 per transaction
* **Product Portfolio**: 45+ devices from 20+ manufacturers
* **Inventory Turnover**: Calculated per product category
* **Customer Acquisition Cost**: Derived from sales and marketing data

## System Architecture Summary

### Technology Stack

* **Database**: PostgreSQL 15+
* **Containerization**: Docker & Docker Compose
* **ETL**: SQL with postgres\_fdw extension
* **Business Intelligence**: Microsoft Power BI
* **Automation**: PowerShell and Bash scripts

### Data Flow

CSV Data → OLTP Database → ETL Process → OLAP Database → Power BI Dashboard

### Performance Optimizations

* Indexed columns for frequent queries
* Partitioned fact tables by date
* Materialized views for complex aggregations
* SCD Type 2 for historical analysis without data loss