# Smart Home Automation Database System - Course Work Documentation

## Table of Contents

1. [OLTP Database Context](#oltp-database-context)
2. [OLAP Database Context](#olap-database-context)
3. [Overall Description of Schemas](#overall-description-of-schemas)
4. [Implementation Instructions](#implementation-instructions)
5. [Power BI Report Description](#power-bi-report-description)

## 1. OLTP Database Context

### What We Store in OLTP Database

The OLTP (Online Transaction Processing) database stores operational data for a **Smart Home Automation System**. This system manages:

#### Core Business Entities:

* **Users**: Homeowners who use the smart home system
* **Rooms**: Physical spaces within homes that contain devices
* **Device Types**: Categories of smart devices (lighting, security, climate, etc.)
* **Devices**: Individual smart home devices (lights, sensors, thermostats, etc.)
* **Scenes**: Automation scenarios that control multiple devices
* **Events**: System activities and user interactions

#### Business Processes Supported:

1. **User Management**: Registration, authentication, profile management
2. **Device Management**: Installation, configuration, monitoring of smart devices
3. **Room Organization**: Logical grouping of devices by physical location
4. **Automation Control**: Creation and execution of smart home scenes
5. **Activity Tracking**: Logging of all system events and device status changes
6. **Status Monitoring**: Real-time tracking of device operational states

#### Data Characteristics:

* **High Transaction Volume**: Frequent device status updates and user interactions
* **Real-time Processing**: Immediate response to user commands and sensor triggers
* **Data Integrity**: Critical for home security and safety systems
* **Normalized Structure**: 3NF to minimize redundancy and ensure consistency

## 2. OLAP Database Context

### Analytical Questions We Want to Answer

The OLAP (Online Analytical Processing) database is designed to answer strategic business questions for smart home system optimization and business intelligence:

#### Energy Management Analytics:

* Which device types consume the most energy across different seasons?
* How does energy consumption vary by room type and user behavior?
* What are the trends in energy efficiency over time?
* Which manufacturers produce the most energy-efficient devices?

#### User Behavior Analytics:

* How do different user segments interact with their smart home systems?
* What are the peak usage hours for different device categories?
* Which automation scenes are most popular among users?
* How does user engagement vary by geographic location?

#### System Performance Analytics:

* What is the success rate of scene executions throughout the day?
* Which devices have the highest error rates and need maintenance?
* How does system response time vary during peak usage periods?
* What is the average lifespan of different device types?

#### Business Intelligence:

* Which device manufacturers have the largest market share?
* What room configurations generate the highest user satisfaction?
* How effective are different automation strategies?
* What are the seasonal patterns in smart home usage?

#### Operational Analytics:

* Which rooms have the highest device density and activity?
* How do device failure rates compare across manufacturers?
* What are the most common device status patterns?
* How does user activity correlate with external factors (time, weather)?

## 3. Overall Description of Schemas

### 3.1 OLTP Schema (smart\_home database)

#### Architecture: 3rd Normal Form (3NF)

The OLTP schema follows 3NF principles to eliminate redundancy and ensure data integrity.

#### Tables Description:

**Users Table** - **Purpose**: Store user account information - **Primary Key**: UserID - **Attributes**: Name, Email, Password - **Relationships**: One-to-many with Rooms and Scenes

**Rooms Table** - **Purpose**: Physical spaces within homes - **Primary Key**: RoomID - **Foreign Key**: UserID → Users - **Attributes**: RoomName - **Relationships**: One-to-many with Devices

**DeviceTypes Table** - **Purpose**: Categories of smart devices - **Primary Key**: DeviceTypeID - **Attributes**: TypeName, Description - **Relationships**: One-to-many with Devices

**Devices Table** - **Purpose**: Individual smart home devices - **Primary Key**: DeviceID - **Foreign Keys**: RoomID → Rooms, DeviceTypeID → DeviceTypes - **Attributes**: DeviceName, Manufacturer - **Relationships**: One-to-many with DeviceStatus, many-to-many with Scenes

**DeviceStatus Table** - **Purpose**: Historical log of device status changes - **Primary Key**: StatusID (auto-generated) - **Foreign Key**: DeviceID → Devices - **Attributes**: StatusTimestamp, StatusValue - **Relationships**: Many-to-one with Devices

**Scenes Table** - **Purpose**: Automation scenarios - **Primary Key**: SceneID - **Foreign Key**: UserID → Users - **Attributes**: SceneName, Description - **Relationships**: Many-to-many with Devices, one-to-many with Events

**SceneDevices Table** - **Purpose**: Bridge table for Scenes and Devices many-to-many relationship - **Composite Primary Key**: (SceneID, DeviceID) - **Foreign Keys**: SceneID → Scenes, DeviceID → Devices - **Attributes**: DesiredStatus

**Events Table** - **Purpose**: System activity log - **Primary Key**: EventID (auto-generated) - **Foreign Keys**: UserID → Users, DeviceID → Devices, SceneID → Scenes - **Attributes**: EventTimestamp, EventDescription

#### Key Constraints:

* **Primary Keys**: Ensure unique identification of records
* **Foreign Keys**: Maintain referential integrity
* **NOT NULL**: Critical fields cannot be empty
* **UNIQUE**: Email addresses must be unique per user

### 3.2 OLAP Schema (smart\_home\_olap database)

#### Architecture: Snowflake Schema

The OLAP schema implements a snowflake design with normalized dimensions and multiple fact tables.

#### Dimension Tables:

**DIM\_DATE** - **Purpose**: Time dimension for date-based analysis - **Attributes**: Year, Quarter, Month, Week, Day, IsWeekend - **Granularity**: Daily level

**DIM\_TIME** - **Purpose**: Time dimension for hour-based analysis - **Attributes**: Hour, Minute, TimeOfDay, IsBusinessHour - **Granularity**: Hourly level

**DIM\_USER (SCD Type 2)** - **Purpose**: User dimension with historical tracking - **SCD Attributes**: ValidFrom, ValidTo, IsCurrent - **Business Attributes**: Name, Email, UserType, City, Country - **Change Tracking**: Maintains history of user changes

**DIM\_MANUFACTURER (SCD Type 2)** - **Purpose**: Device manufacturer dimension with history - **SCD Attributes**: ValidFrom, ValidTo, IsCurrent - **Business Attributes**: ManufacturerName, Country, Founded\_Year - **Change Tracking**: Tracks manufacturer information changes

**DIM\_DEVICE\_TYPE** - **Purpose**: Device type classification - **Attributes**: TypeName, Category, Description, EnergyEfficiencyRating

**DIM\_ROOM** - **Purpose**: Room characteristics - **Attributes**: RoomName, RoomType, Area\_SqM, Floor, HasWindows

**DIM\_DEVICE** - **Purpose**: Device characteristics - **Attributes**: DeviceName, Model, InstallationDate, WarrantyExpiry, Status

**DIM\_SCENE** - **Purpose**: Scene characteristics - **Attributes**: SceneName, Category, Description, IsActive

#### Bridge Table:

**BRIDGE\_SCENE\_DEVICE** - **Purpose**: Handles many-to-many relationship between scenes and devices - **Attributes**: SceneKey, DeviceKey, DeviceGroupKey, DesiredStatus, Priority - **Weighting**: Priority field for device importance in scenes

#### Fact Tables:

**FACT\_DEVICE\_USAGE** - **Purpose**: Device usage metrics and energy consumption - **Grain**: Daily device usage per user per room - **Measures**: TotalActivations, TotalUsageMinutes, EnergyConsumption\_kWh, AvgResponseTime\_ms, ErrorCount, UsageScore - **Dimensions**: Date, Device, User, Room

**FACT\_SCENE\_EXECUTION** - **Purpose**: Scene execution performance metrics - **Grain**: Hourly scene execution per user - **Measures**: ExecutionCount, SuccessfulExecutions, FailedExecutions, AvgExecutionTime\_ms, DevicesAffected, SuccessRate - **Dimensions**: Date, Time, Scene, User

**FACT\_MONTHLY\_SUMMARY** - **Purpose**: Aggregated monthly metrics per user - **Grain**: Monthly summary per user - **Measures**: TotalDeviceActivations, TotalEnergyConsumption\_kWh, TotalSceneExecutions, AvgDevicesPerScene - **Dimensions**: Year, Month, User, MostUsedDevice, MostUsedScene

#### Indexes for Performance:

* **Fact Tables**: Indexed on all dimension keys and date fields
* **Dimension Tables**: Indexed on natural keys and SCD fields
* **Bridge Table**: Indexed on both scene and device keys

## 4. Implementation Instructions

### 4.1 Prerequisites

* Docker installed on your system
* Docker Compose available
* Minimum 4GB RAM available
* Port 5433 and 8080 available

### 4.2 Initial Setup

#### Step 1: Clone and Navigate

git clone git@github.com:ArKoWol/smart\_home\_automation\_SQL.git  
cd smart\_home\_automation\_SQL

#### Step 2: Start the System

docker-compose up -d

#### Step 3: Wait for Initialization

* Wait approximately 30-60 seconds for database initialization
* Check logs: docker-compose logs

### 4.3 Database Access

#### PostgreSQL Direct Access:

* **Host**: localhost
* **Port**: 5433
* **OLTP Database**: smart\_home
* **OLAP Database**: smart\_home\_olap
* **Username**: admin
* **Password**: admin123

#### PgAdmin Web Interface:

* **URL**: http://localhost:8080
* **Email**: admin@example.com
* **Password**: admin123

### 4.4 Script Execution Order

#### OLTP Database Setup:

1. **Table Creation**: init/01\_create\_tables.sql (auto-executed)
2. **Data Loading**: init/02\_load\_data.sql (auto-executed)

#### OLAP Database Setup:

1. **Schema Creation**: olap/01\_create\_olap\_schema.sql
2. **Reference Data**: olap/02\_populate\_reference\_data.sql
3. **ETL Process**: olap/03\_etl\_process.sql
4. **Demo Data**: olap/05\_generate\_demo\_data.sql

#### Manual Execution Commands:

# Connect to OLTP database  
docker exec -it smart\_home\_db psql -U admin -d smart\_home  
  
# Connect to OLAP database (after creation)  
docker exec -it smart\_home\_db psql -U admin -d smart\_home\_olap  
  
# Run specific script  
docker exec -i smart\_home\_db psql -U admin -d smart\_home\_olap < olap/01\_create\_olap\_schema.sql

### 4.5 ETL Process Execution

#### Full ETL Setup:

-- 1. Create OLAP schema  
\i olap/01\_create\_olap\_schema.sql  
  
-- 2. Populate reference data  
\i olap/02\_populate\_reference\_data.sql  
  
-- 3. Run ETL process  
\i olap/03\_etl\_process.sql  
  
-- 4. Generate demo data  
\i olap/05\_generate\_demo\_data.sql

#### ETL Process Features:

* **Incremental Loading**: Only new/changed data is processed
* **Data Validation**: Quality checks before loading
* **Error Handling**: Graceful handling of data issues
* **SCD Type 2**: Automatic historical tracking
* **Rerunnable**: Can be executed multiple times safely

### 4.6 Analytical Queries Execution

#### Run Analysis Queries:

# Execute all analytical queries  
docker exec -i smart\_home\_db psql -U admin -d smart\_home < analytical\_queries.sql  
  
# Execute demonstration queries  
docker exec -i smart\_home\_db psql -U admin -d smart\_home\_olap < demonstration\_queries.sql

### 4.7 Troubleshooting

#### Common Issues:

1. **Port Conflicts**: Change ports in docker-compose.yml if needed
2. **Memory Issues**: Ensure sufficient RAM available
3. **Permission Issues**: Check Docker permissions

#### Reset Command:

# Complete reset  
docker-compose down -v  
docker-compose up -d

## 5. Power BI Report Description

### 5.1 Report Overview

The Power BI report (Course work.pbix) provides comprehensive business intelligence dashboards for smart home system analysis.

### 5.2 Data Connection

* **Source**: OLAP database (smart\_home\_olap)
* **Connection Type**: Direct Query to PostgreSQL
* **Refresh Schedule**: Can be configured for automatic updates

### 5.3 Data Transformations Applied

#### Data Type Optimizations:

* **Decimal to Integer**: Converted activation counts and device counts to whole numbers
* **Text Formatting**: Standardized manufacturer names to proper case
* **Date Formatting**: Unified date formats across all time dimensions

#### Filtering and Cleaning:

* **Current Records Only**: Filtered SCD Type 2 dimensions to show only current records
* **Active Devices**: Excluded decommissioned devices from analysis
* **Valid Date Ranges**: Filtered data to relevant business periods

### 5.4 Visual Components Description

#### 5.4.1 Title and Headers

* **Main Title**: “Smart Home Analytics Dashboard”
* **Section Headers**: Clear labeling for each analytical section
* **KPI Summary**: High-level metrics prominently displayed

#### 5.4.2 Slicers (Interactive Filters)

1. **Time Period Slicer**
   * **Purpose**: Filter data by date ranges (monthly, quarterly, yearly)
   * **Type**: Date range picker
   * **Impact**: Affects all visuals dynamically
2. **Device Category Slicer**
   * **Purpose**: Focus analysis on specific device types
   * **Options**: Lighting, Security, Climate, Entertainment, etc.
   * **Type**: Multi-select dropdown
3. **User Segment Slicer** (Bonus)
   * **Purpose**: Analyze different user groups
   * **Options**: Power User, Regular User, Casual User
   * **Type**: Button-style slicer

#### 5.4.3 Visual Components

1. **Energy Consumption Trend Line Chart**
   * **Purpose**: Shows energy consumption trends over time
   * **X-axis**: Month/Year
   * **Y-axis**: Total Energy Consumption (kWh)
   * **Series**: Broken down by device category
   * **Insights**: Identifies seasonal patterns and consumption trends
2. **Device Usage Heat Map**
   * **Purpose**: Visualizes device activity by time of day and day of week
   * **X-axis**: Hour of day (0-23)
   * **Y-axis**: Day of week
   * **Color Scale**: Intensity represents usage frequency
   * **Insights**: Shows peak usage patterns for optimization
3. **Manufacturer Market Share Pie Chart**
   * **Purpose**: Shows market share distribution among device manufacturers
   * **Values**: Count of devices by manufacturer
   * **Labels**: Manufacturer names with percentages
   * **Insights**: Identifies dominant vendors and market concentration
4. **Scene Success Rate Bar Chart**
   * **Purpose**: Compares automation scene performance
   * **X-axis**: Scene categories (Security, Comfort, Energy Saving, etc.)
   * **Y-axis**: Success rate percentage
   * **Colors**: Green (high success) to Red (low success)
   * **Insights**: Identifies which automation types work best
5. **Room Efficiency Scatter Plot**
   * **Purpose**: Analyzes room utilization efficiency
   * **X-axis**: Room area (square meters)
   * **Y-axis**: Energy consumption per square meter
   * **Bubble Size**: Number of devices in room
   * **Insights**: Identifies over/under-utilized spaces
6. **User Engagement Funnel Chart**
   * **Purpose**: Shows user activity levels
   * **Stages**: Registered Users → Active Users → Power Users
   * **Values**: Count of users at each engagement level
   * **Insights**: User adoption and engagement patterns

#### 5.4.4 KPI Cards

* **Total Active Users**: Current count of active system users
* **Average Energy Savings**: Percentage reduction vs. traditional homes
* **System Uptime**: Overall system reliability percentage
* **Total Devices Managed**: Count of all active devices in system

### 5.5 Business Value Delivered

#### For Homeowners:

* Energy consumption optimization insights
* Device performance monitoring
* Usage pattern analysis for cost savings

#### For Smart Home Companies:

* Market share analysis by manufacturer
* Product performance benchmarking
* User behavior insights for product development

#### For System Administrators:

* Performance monitoring dashboards
* Maintenance scheduling optimization
* Capacity planning insights

### 5.6 Report Features

#### Interactive Elements:

* **Cross-filtering**: Selecting data in one visual filters others
* **Drill-down**: Click through from yearly to monthly to daily views
* **Tooltips**: Hover for additional context and details

#### Export Capabilities:

* **PDF Export**: For sharing with stakeholders
* **Excel Export**: For further analysis
* **PowerPoint Export**: For presentations

## 6. Conclusion

This Smart Home Automation Database System provides a comprehensive solution for both operational management (OLTP) and business intelligence (OLAP). The implementation follows industry best practices with proper normalization, dimensional modeling, and performance optimization.

### Key Achievements:

* Fully normalized OLTP system with 8 tables
* Snowflake OLAP schema with 2 facts, SCD Type 2, and bridge table
* Comprehensive ETL processes with data quality controls
* Rich set of analytical queries for business insights
* Professional Power BI dashboard with multiple visualizations
* Complete documentation and deployment automation

The solution is production-ready and can be easily deployed, maintained, and extended for real-world smart home automation scenarios.