

## 1. 2 Facts

Columns that represent quantitative, measurable values:

**COST\_OF\_INSPECTION\_IN\_DOLLARS:** The dollar amount spent on the inspection (a measure of cost).

**INSPECTION\_SCORE:** The score awarded for the inspection (a measure of quality/performance).

### Types of Facts:

#### **COST\_OF\_INSPECTION\_IN\_DOLLARS:**

- **Type:** Additive Fact.
- **Reason:** This metric can be summed across multiple dimensions, such as time, geography, or public housing agencies. For instance, you can calculate the total cost of inspections for all agencies in a given year or state.

#### **INSPECTION\_SCORE:**

- **Type:** Semi-Additive Fact.
- **Reason:** This metric cannot be meaningfully summed across all dimensions but can be averaged or used in specific contexts, such as calculating the average score for a housing agency or within a state over time.

## 2. 7 Dimensions:

Columns that provide descriptive information to contextualize the facts:

1. **INSPECTION\_ID:** A unique identifier for the event (used for identifying rows but not a measure).
2. **PUBLIC\_HOUSING\_AGENCY\_NAME:** The name of the housing agency (descriptive of the organization involved).
3. **INSPECTED\_DEVELOPMENT\_NAME:** The name of the inspected development (provides context for the location).
4. **INSPECTED\_DEVELOPMENT\_ADDRESS:** The address of the development (specific geographic location).
5. **INSPECTED\_DEVELOPMENT\_CITY:** The city where the inspection occurred (higher-level geographic context).
6. **INSPECTED\_DEVELOPMENT\_STATE:** The state where the inspection occurred (broader geographic context).

7. **INSPECTION\_DATE**: The date of the inspection.

### 3. Type of fact tables

Given the provided context and the requirements for storing and analyzing the inspection data, I would design two types of fact tables to cater to the needs of Senior Management:

#### **1. Inspection-Level Fact Table:**

- **Purpose:** This table would store granular, inspection-level data, capturing the detailed information for each inspection event.
- **Key Facts:**
  - COST\_OF\_INSPECTION\_IN\_DOLLARS (Additive Fact)
  - INSPECTION\_SCORE (Semi-Additive Fact)
- **Dimensions:**
  - DATE\_DIMENSION (to store the inspection date)
  - PHA\_DIMENSION (to store details about the Public Housing Agency)
  - DEVELOPMENT\_DIMENSION (to store development name, address, city, and state)
- **Why Needed:**
  - This table allows detailed reporting and analysis at the individual inspection level.
  - Senior Management can drill down into inspection details for any development, agency, or specific date.

#### **2. Periodic Summary Fact Table:**

- **Purpose:** This table would store monthly aggregated data to support periodic summary reporting.
- **Key Facts:**
  - TOTAL\_COST\_OF\_INSPECTIONS (Additive Fact): The total cost of inspections for a given month.
  - AVERAGE\_INSPECTION\_SCORE (Derived Semi-Additive Fact): The average inspection score for inspections performed during the month.
- **Dimensions:**
  - DATE\_DIMENSION (aggregated at the month level to allow summarization by year, quarter, and month).
  - PHA\_DIMENSION (to aggregate by Public Housing Agency if needed).
  - GEOGRAPHIC\_DIMENSION (optional for aggregating costs by region, state, or city).
- **Why Needed:**
  - Senior Management is explicitly interested in viewing periodic summaries of inspection costs by month.
  - Summarized tables improve performance for high-level reports and dashboards.

#### 4. Handle the slowly changing dimension

To handle the changes in public housing agency (PHA) names and addresses, Type 2 Slowly Changing Dimension (SCD) is the most appropriate choice. Here's why:

##### **Characteristics of the Problem:**

- Frequent Mergers: PHAs are often merged or reorganized, leading to changes in their names and possibly addresses.
- Need to Track History: Senior Management likely needs to analyze both historical data (based on how the agency was structured at the time of the inspection) and current data (based on the latest agency structure).
- Historical Integrity: Retaining historical context is essential to ensure the accuracy of reports over time.

##### **Why Type 2 is the Best Option:**

- Preserves History: Type 2 creates a new record for every change in the PHA's name or address. This allows the system to retain historical information about how the agency was structured during past inspections.
- Supports Time-Based Analysis: With Type 2, we can include a start date, end date, and possibly an "is current" flag to distinguish between historical and active records. This supports both historical and current-state reporting.
- Handles Mergers and Splits Effectively: If two agencies merge, new records can be created to reflect the updated agency structure while retaining details about the original entities.