



PUBLIC HOUSING INSPECTION ANALYSIS

- Shivani Vellanki

Introduction:

Public housing in the United States plays a crucial role in providing affordable housing options for low-income individuals and families. The administration of public housing programs is decentralized, with local or regional Public Housing Agencies (PHAs) responsible for managing and maintaining housing developments within their respective jurisdictions. These PHAs operate under the oversight of the Federal Department of Housing & Urban Development (HUD). To ensure the quality and safety of public housing units, PHAs conduct regular inspections of these developments. During these inspections, PHAs assess various factors such as the physical condition of the properties, observance to safety standards, and overall habitability. Each inspection results in an inspection score, which is expressed as a ratio or percentage, reflecting the condition of the housing development. Additionally, there is a cost associated with performing these inspections, which is borne by the U.S. taxpayers. This data is important to HUD and senior management, as it helps in monitoring the effectiveness of public housing programs, allocating resources efficiently, and ensuring that taxpayer dollars are used judiciously to maintain safe and decent housing for those in need. In this assignment, we will develop a dimensional model for the inspection data and conduct key analyses to support informed decision-making.

1. Answer the questions below.

- How many facts are there in this dataset?
- Which facts do you identify?
- For the facts that you identify, what type of facts are they?

There are **2 facts** in this dataset, `COST_OF_INSPECTION_IN_DOLLARS` and `INSPECTION_SCORE`. However, `INSPECTION_ID` and the `PUBLIC_HOUSING_AGENCY_ID` are typically not considered facts or measures themselves; rather, they are identifiers or keys used to link data in the fact and dimension tables.

`COST_OF_INSPECTION_IN_DOLLARS` – Additive Fact
`INSPECTION_SCORE` – Non-Additive Fact

2. Answer the questions below.

- How many dimensions are there in this dataset?
- Which dimensions do you identify?

Dimension Table 1 (PUBLIC_HOUSING_AGENCY):

PUBLIC_HOUSING_AGENCY_NAME_ID (Primary Key)
PUBLIC_HOUSING_AGENCY_NAME

Dimension Table 2 (INSPECTION_DETAILS):

INSPECTION_ID (Primary Key)
INSPECTED_DEVELOPMENT_NAME
INSPECTED_DEVELOPMENT_ADDRESS
INSPECTED_DEVELOPMENT_CITY
INSPECTED_DEVELOPMENT_STATE
INSPECTION_DATE

3. Answer the question below.

Senior management is interested in viewing the facts identified above, at both the inspection level, as well as a periodic summary of inspection costs for each month. Based on this context, if you were to store these data in a set of fact tables, which type (or types) of fact tables would you use and why?

For periodic summary of inspection costs for each month,

Transaction Fact Table (Inspection-Level): This type of fact table is suitable for storing detailed, granular data at the inspection level. I would use it to store facts like "COST_OF_INSPECTION_IN_DOLLARS" and "INSPECTION_SCORE" for each individual inspection. Each row in this fact table represents a single inspection event with associated metrics. This allows you to analyze and report on specific inspections.

Snapshot Fact Table (Monthly Summary): This type of fact table is suitable for storing aggregated or summarized data over specific time periods, such as months. In this case, I would create a snapshot fact table to store the monthly summary of inspection costs. This table would include aggregated measures like total inspection cost for each month. Each row in this table represents a month, and the associated metrics provide a summary for that month.

4. Answer the question below.

Senior Management is also concerned with changes in the names and addresses of the public housing agency names since they tend to get merged with other agencies on a frequent basis.

Based on this context, how would handle this slowly changing dimension? Select from types 0,1,2, or 3 from the Kimball reading. Justify your answer.

In the context of changes in the names and addresses of public housing agency names, where these changes occur frequently due to mergers, we can handle this as a Type 2 Slowly Changing Dimension (SCD) according to the Kimball methodology. Type 2 Slowly Changing Dimension (SCD)/ Historical Dimension: This type of dimension handling allows us to maintain historical changes over time. In this case, when public housing agencies change names or addresses due to mergers, we want to keep a historical record of these changes.

5. Address the scenario below.

Finally, Senior Management is interested in a subset of this data, for only those PHAs that saw an *increase* in the \$\$ cost of performing an inspection in their jurisdiction. Since none of them are SQL programmers, they've asked your help in performing this analysis by providing a file as your final deliverable with the following columns:

[Output.csv](#)

| PUBLIC_HOUSING_AGENCY_NAME | MR_INSPECTION_DATE | SECOND_MR_INSPECTION_DATE | MR_INSPECTION_COST | SECOND_MR_INSPECTION_COST | CHANGE_IN_C |
|-----------------------------------|--------------------|---------------------------|--------------------|---------------------------|-------------|
| Akron Metropolitan Housing Autho | 2014-10-09 | 2014-10-08 | 25593 | 15626 | 9967 |
| Alachua County | 2015-01-22 | 2014-05-01 | 37345 | 17019 | 20326 |
| Alaska Housing Finance Corporati | 2014-11-14 | 2014-11-13 | 26342 | 21366 | 4976 |
| Albany Housing Authority | 2015-01-12 | 2015-01-09 | 31115 | 30247 | 868 |
| Alexander County Housing Authori | 2014-11-18 | 2014-04-24 | 31272 | 18855 | 12417 |
| Alexandria Redevelopment & Housi | 2014-05-09 | 2014-04-18 | 29123 | 14767 | 14356 |
| ALLEGHENY COUNTY HOUSING AUTHORI | 2015-02-02 | 2015-02-02 | 37108 | 36454 | 654 |
| Allentown Housing Authority | 2014-11-17 | 2014-11-14 | 34040 | 18989 | 15051 |
| ALTOONA HOUSING AUTHORITY | 2014-11-24 | 2014-09-15 | 25750 | 24813 | 937 |
| ANNISTON HA | 2014-12-30 | 2014-08-21 | 31506 | 10785 | 20721 |
| Area Housing Commission | 2013-06-25 | 2013-06-24 | 28713 | 19114 | 9599 |
| Asbury Park Housing Authority | 2014-06-03 | 2014-05-21 | 35723 | 14987 | 20736 |
| Ashland Housing Authority | 2014-04-29 | 2014-04-29 | 29106 | 17510 | 11596 |
| ASHTABULA METROPOLITAN HOUSING A | 2014-06-03 | 2014-04-24 | 37948 | 13920 | 24028 |
| Athens Metropolitan Housing Auth | 2014-05-22 | 2014-05-21 | 21816 | 10996 | 10820 |
| Aurora Housing Authority | 2015-02-02 | 2014-06-24 | 14683 | 12831 | 1852 |
| Aurora Housing Authority of the C | 2014-07-03 | 2013-06-11 | 14908 | 14570 | 338 |
| Austin Housing Authority | 2014-06-30 | 2014-06-26 | 36672 | 25920 | 10752 |
| Barre Housing Authority | 2014-06-18 | 2014-06-16 | 19254 | 16757 | 2497 |
| Batavia Housing Authority | 2015-01-28 | 2014-12-30 | 26365 | 14576 | 11789 |
| Battle Creek Housing Commission | 2015-01-29 | 2015-01-27 | 34258 | 15344 | 18914 |
| Bay City Housing Commission | 2014-11-12 | 2014-01-28 | 35900 | 16470 | 19430 |
| Bayonne Housing Authority | 2014-09-12 | 2014-09-11 | 25407 | 16280 | 10127 |
| Belmont Metropolitan Housing Aut | 2013-07-10 | 2013-07-09 | 35736 | 26915 | 8821 |
| Beloit Housing Authority | 2014-04-30 | 2013-05-14 | 35276 | 14461 | 20815 |

References:

Dimensional Modeling Techniques. Kimball Group. (n.d.). <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/>

Quintas, R. (2015, November 30). *The Data Warehouse Toolkit, 3rd Edition*. Academia.edu. https://www.academia.edu/19222417/The_Data_Warehouse_Toolkit_3rd_Edition