

Study of History Management of Data: Slowly Changing Dimension to Independent Data Mart

Arvind Singh, Surya Prakash Pandey

Department of Computer Science, Awadhesh Pratap Singh University Rewa, Madhya Pradesh, India

ABSTRACT

Data History plays vital role on management of historical data on data mart such as real time updating and recording difficulties especially when many operations are performing. The difficulties of data management include timely update and strong storage system. In this paper, we describe few methods of managing data history in independent data marts as well as databases. And also illustrate the concept of slowly changing dimension that this how it can be utilized in the independent data mart of institutions to update and maintain campaign records of customers.

There are many kinds of dealing with the history of the data. In this article we will present some examples, advantages and disadvantages of each of the method and use of possible prospective.

Keywords : Independent Data Mart, Managing Data History, Databases

Article Info

Volume 8, Issue 5

Page Number : 575-578

Publication Issue

September-October-2021

Article History

Accepted : 01 Oct 2021

Published : 10 Oct 2021

I. INTRODUCTION

In current time, nearly everyone use data in where they are doing some work from their home to institutions. We can define data it as a set or collection of values and variables which may be belonging to a set of items. It is very often presented in a grid form (rows and columns), data tree (parent and child organization) or may be in a graphical structure such as tables models with graphical representation [3]. The data we keep does not have to be in a text type form. We can put it as a number and even an image. The data store in our databases or in independent data marts, its quality gives us a huge advantage for the data visualization and their history management. We need to remember in our mind that the dimension data is not a stable entity and it may be change time to

time so, to manage these types of data history Slowly Changing Dimensions was Discovered.

1. SLOWLY CHANGING DIMENSIONS

First time Slowly Changing Dimensions (SCD) was discovered by Ralph Kimball, who was supposed as a one of the architects of data warehousing. Their methodology and methods became a standard. Slowly changing dimensions (SCD) are a set (collection) of methods for managing data history in dimension tables [1].

The Financial institutions like bank generate large amount of data daily due multiple transactions and operations. These data are analyzed by the institutions to provide valuable customer insights such as the customer behavior, needs and desires.

Institutions also use such insights for strategic campaign management.

Several campaigns management may result in customer resentment when customers are overly contacted by simultaneous marketing campaigns.

There are two categories of campaigns in the institution like bank – sales-based and non-sales based. The sales based campaigns covers selected group of customers which may leads to market and sell products and services.

Non sales based campaigns provide customers with notification and fulfillment service.

1.1. Data loading process to data mart

Data loading may be change time to time this should take it into account while developing our system

1. **Source data**– Data source should be delivered in a proper manner.
2. **ETL (Extract Transform Load)** – ETL follows a mechanism to operate with data incoming from source to the target system.
3. **Structure Database**– This based on the method of data history management and the structure of database has to be conditioned.

Below figure shows that the how source data goes directly to the data mart for storage purposes.

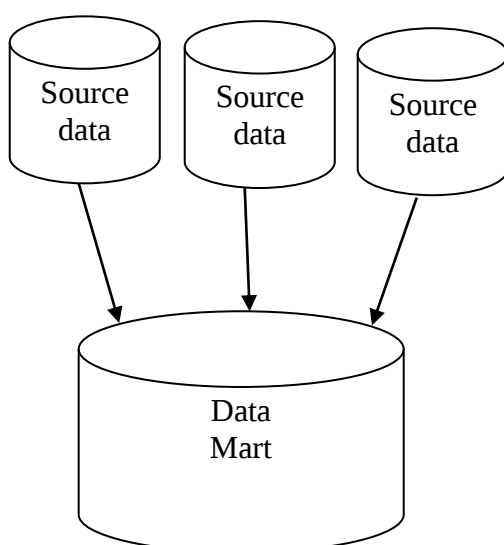


Figure 2.1 Independent Data Mart

Independent data marts are independently (alone) repositories that are controlled by a single institution and unit. Metadata (data about data) for the data mart may independent from metadata captured for data marts by other units.

Source data for the independent data mart is extract from multiple operational units as well as systems that may be external to the business unit for which the data mart constructed [4].

To better explanation of the way of data history management we need to study more on the main methods.

1.2. Types of Slowly Changing Dimensions (SCD)

The Slowly Changing Dimensions (SCD) types are to go through all of them with more examples. We have three (3) main types of Slowly Changing Dimensions: Type 1, Type 2, and Type 3

The most commonly practiced SCD types are 1, 2 and 3. Below we present the more details of the different types of SCD [2].

Type 1 - Type one SCD overwrites old records with new records as shown in tables given below. This is easily can be to maintain.

Before Change 2019

C_ID	C-Name	Contact	Date
1001	Jai Singh Gaharwar	7024070467	04-04-2019

After Change 2020

C_ID	C-Name	Contact	Date
1001	Jai Singh Gaharwar	8989869983	04-05-2020

The change in Contact and Date is a result of an error or may be change of structure. Even if sometimes does not meet according to business requirements. The main advantage is for this type of method of development is the simplicity of database structure as

well as ETL system. And main disadvantage of this method is the fact that we can miss the data history.

Type 2 – Type two SCD updates the record by inserting new records with new ID or Serial numbers while keeps the historical data which is already stored before. Using this method we can maintain lot of historical data.

Before change

C_ID	C-Name	Contact	Date	flag
1001	Jai Singh Gaharwar	7024070467	04-04-2019	1

After change

C_ID	C-Name	Contact	Date	flag
1001	Jai Singh Gaharwar	7024070467	04-04-2019	0
1002	Jai Singh Gaharwar	7024070467	24-06-2020	1

Type 3 – Type three SCD updates the record by creating new dimensions to the table structure. It preserves limited history - only the previous record could be preserve. In other words using this method we can add new dimension column to keep the historical value of the item.

Before change

C_ID	C-Name	City	Date
1001	Rajesh	Rewa	04-04-2019

After Change

C_ID	C-Name	City	Effective Date	End Date
1001	Rajesh	Rewa	04-04-2019	04-04-2020

II. IMPROVEMENT

The SCD Type two is that the data is regularly, inserting on a one database may cause the database to expand fast overtime. My suggestion is that diverting historical data onto another Data Archive Table. Flag management is one of the best options to manage historical data. In Type 3 if you add new column dimensions it may causing fast growing of records.



Figure 3.1. Improved Data Management

III. CONCLUSION

Study of Historical data and independent data mart we have studied tree types of managing historical data in Dimensional tables. Each of them has their advantages and disadvantages and it can be used in totally different institution requirements. It is easy to implement, maintain and develop. The most popular method from those three is SCD Type two where flag is implemented on example table, which gives us a history of a Dimension data and helps us to build reports not only on current but also on the historical data.

Management of multiple institution campaigns and their records of a large amount of customer base is a difficult work for the any institution.

IV. REFERENCES

- [1]. Altaf, Md, Ralph Kimball (2008, September 22). Slowly Changing Dimensions, Part 2. Kimball Group, January 20, 2013
- [2]. Margy Ross (2008, September 29). Practical Steps for Designing a Dimensional Model. Kimball Group. January 20, 2013

- [3]. Shauna Trainor (2013, February 13). Market In Modesty To Prevent Resentment. Covenant Group. February 13,2013
- [4]. Inmon, W.H., Building the Data Warehouse, New York: Wiley Computer Publishing, 1996, Second Edition

Cite this article as :

Arvind Singh, Surya Prakash Pandey , "Study of History Management of Data: Slowly Changing Dimension to Independent Data Mart", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 8 Issue 5, pp. 575-578, September-October 2021.

Journal URL : <https://ijsrst.com/IJSRST218698>