Data MAPPING DOCUMENT

Data Warehouse – Yearly\_fact

Date: 20/11/2015

Author: Declan Barnes

**TABLE OF CONTENTS**

1. DOCUMENT HISTORY 3

1.1. Document Location 3

1.2. Revision History 3

1.3. Approvals 3

1.4. Distribution 3

2. Introduction 4

2.1. Purpose of this document 4

2.2. Glossary 4

2.3. Impact & Dependency Analysis (if changing a Table) 4

3. *Yearly\_fact* MappING 5

3.1. Extraction Criteria 5

3.2. Primary Index 5

3.3. Partitioning 5

3.4. Data Retention 5

3.5. Statistics gathering 5

3.6. Data Mapping 6

3.6.1. Column Name 6

# DOCUMENT HISTORY

## Approvals

This document requires the following approvals.

|  |  |  |
| --- | --- | --- |
| **Name** | **Title** | **Date of Issue** |
| Eamon Nolan | Lecturer | 23/11/2015 |

## Distribution

This document has been distributed to

| **Name** | **Title** | **Date of Issue** |
| --- | --- | --- |
| Declan Barnes | Project Manager | 20/11/2015 |
| Chris Doran | Data Administrator | 20/11/2015 |
| Sean McDermott | Data Architect | 20/11/2015 |
| Yousef O’Conner | Web Designer | 20/11/2015 |
| Daniel Gorman | Business Analyst | 20/11/2015 |
| John McNamara | Data Analyst | 20/11/2015 |

# Introduction

## Purpose of this document

*Enter a description of what mappings are being documented.. It should be possible from reading this section to understand in English where the data comes from for this table and the purpose of storing the data in the data warehouse.*

*For the Yearly\_ fact table the mappings are the following:*

*The Yearly\_fact table is storing the information from the Order\_dim and Payment\_Dim, the Yearly\_fact table will consist of the Yearly: Cust\_ID, Pay\_ID, Prod\_ID and Order\_ID etc. It will also be storing information from the CustomerProduct\_Dim Table which will also consist of the Cust\_ID and Product\_ID along with the Product Type, price, size of the product i.e. prefShoes\_size attributes etc. Whereas the Yearly\_fact table will consist of the following attributes: YearDate\_from, YearDate\_to, MostProductSoldByCountry, BusiestSoppingPeriod and several other Attributes. All this in regards to the Year\_fact will help us in our queries at a later stage and more.*

## Glossary

|  |  |
| --- | --- |
| **Abbrev.** | **Meaning** |
| \_DIM | ‘Dimension’ |
| Cust\_ID | The customer’s id |
| Prod\_ID | The product’s id |
| Order\_ID | The order’s id |
| Pay\_ID | The payment’s id |
|  |  |

# *TABLE Name* MappING

## Extraction Criteria

*Enter any relevant information regarding the data extraction:-*

* *Specify if there should be an initial load of data, and what the constraints are on the initial load.*
* *Specify the extraction criteria for the delta load and how often the delta load is to run.*

## Data Retention

*Define how long the data should be retained for (usually in months). This applies to fact tables or to dimension tables that contain history records, i.e. any table the will keep growing.*

\*\*\*\*\*\*\*\*\*\*\*Eamon said this does not apply to our project\*\*\*\*\*\*\*\*\*\*\*\*

## Data Mapping

*Notes:*

* *Ensure table names and column names adhere to the standards*
* *Ensure that columns which exist already in the warehouse are given the same name and datatype*
* *The Description column must be meaningful – we are not maintaining a separate data dictionary. Identify the list values for ID fields where a finite list exists.*
* *Include column src\_file where the source is a file source. Include src\_table where there are multiple tables that could be the source.*
* *Include loaded\_date.*

\*\*\*\*TABLE NAMES MUST NOT BE GREATER THAN 20 CHARACTERS IN LENTGH\*\*\*\*\*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Destination Column** | **Source Table(s)** | **Source column(s)** | **Data**  **Type** | **Transformation** | **Description** | **Default value** |
| Payment\_Dim | Yearly\_fact | Pay\_ID | Int(11) | N/A | The payment id of a payment that has been made | *-1* |
| Order\_Dim | Yearly\_fact | Order\_ID | Int(11) | N/A | The id of an order made. | -1 |
| CustomerProduct\_Dim | Yearly\_fact | Cust\_ID | int(11) | N/A | The id of Smart-Shops customers. | -1 |
| CustomerProduct\_Dim | Yearly\_fact | Prod\_ID | Int(11) | N/A | The id of each our products | -1 |

*Detail any additional rules or conditions for columns:*

*Pay\_ID = Primary Key, Order\_ID = Primary Key, Cust\_ID = Primary Key, Prod\_ID = Primary Key*