Judo and Data Warehouses

Database Answers

Gemma Gibbons of the UK throws Audrey Tcheumeo of France on her way to a Silver Medal

Barry Williams

1. Management Summary

1.1 A trip to the Olympics

In this Paper we use a trip to the Olympics to discuss an approach to the implementation of a Reference Data Architecture and the design of a Data Warehouse.

A **Canonical Data Model** (CDM) is central to this and we discuss the benefits of using Design Patterns based on a CDM.

During the trip, I used the Ticket I had bought online, bought lunch and watched the Judo competition.

After I returned home I found myself thinking that the trip would provide a good opportunity to develop an interesting and 'User-Friendly' Tutorial on Data Warehouses.

The design of the Data Models reflects the scope and the fact that the overall aim is to provide data for Business Intelligence.

We also try to keep in mind that a well-designed Data Model should be good to look at and it should be possible to tell a story based on the Model.

1.2 The Approach

The Approach is to follow these Steps :-

Step 1 – Identify the **Events** involved

Step 2 – Define a **Design Pattern** based on the Event-driven **Canonical Data Model** from this page on our Database Answers Web Site :

- http://www.databaseanswers.org/data_models/canonical_data_model/index.htm

Step 3 - Define a **Message** Format for the data in each Event

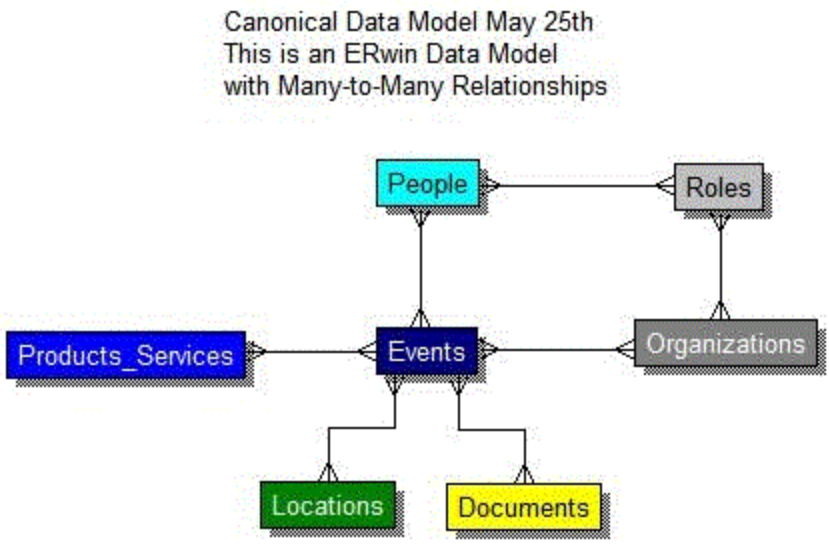
Step 4 - Design a 3rd Normal Form **Data Warehouse** (DWH) and update it for each Event.

Step 5 – Define the format for **loading data** into the DWH for each Message

2. Canonical Data Model

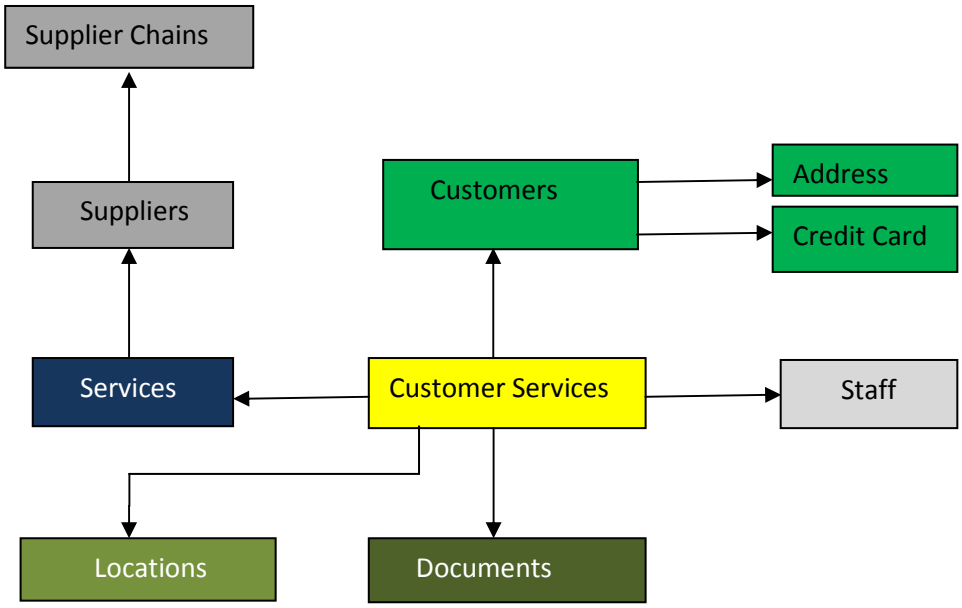
Step 1 – Use the Event-driven Canonical Data Model from this page in our Web Site :

- http://www.databaseanswers.org/data_models/canonical_data_model/index.htm



3. Design Pattern

From the Canonical Data Model shown above, we can derive this Design Pattern :-



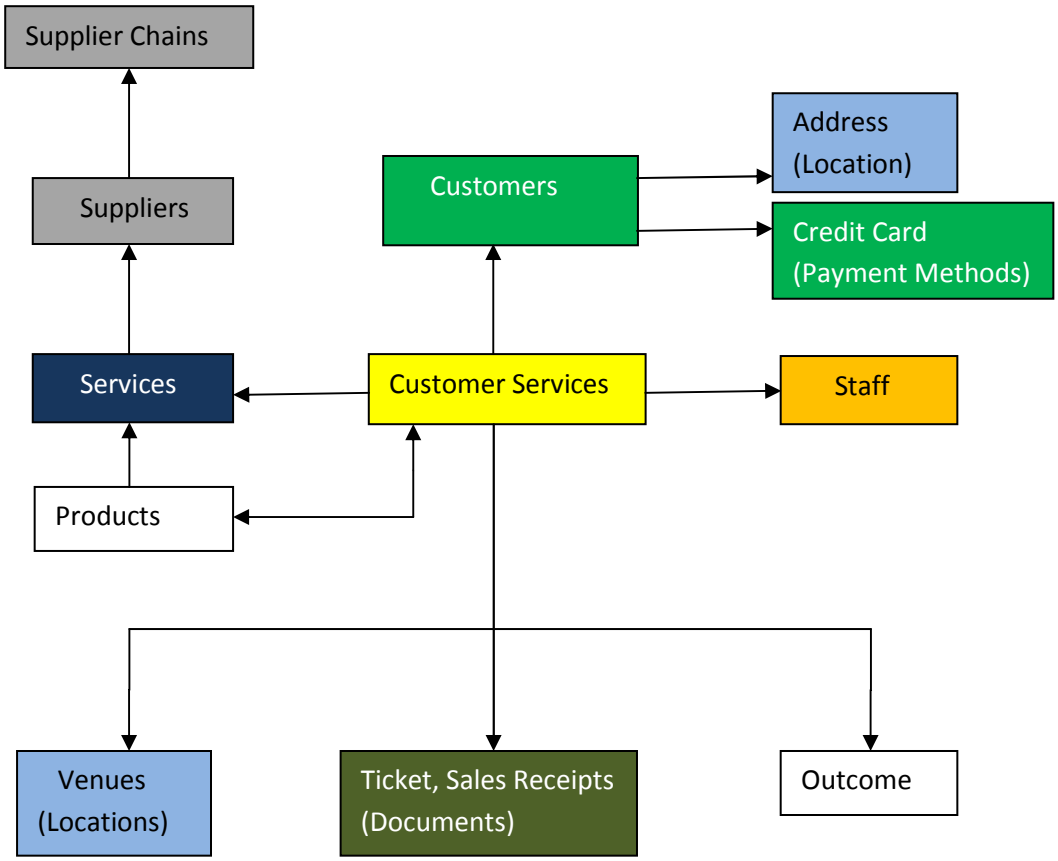
4. Generic Message Format

This shows the fields in the Generic Message Format :-

EVENT	DATE	LOCATION	PRICE	DETAILS

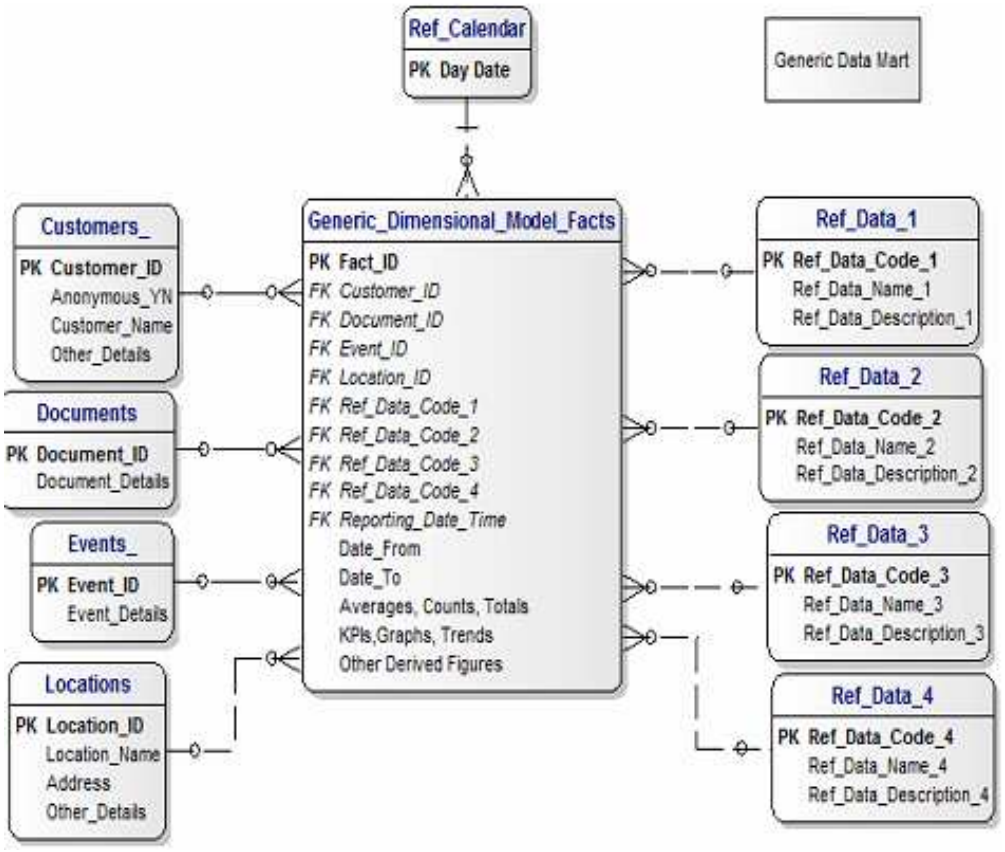
5. 3NF Data Warehouse

The design of this Data Warehouse is derived from the Design Patterns for the three Events that we discuss.
Every Event has an Outcome but it is not usually important and is taken for granted.
We show it here because it is very important for Judo competitions.
We show a many-to-many relationship between Products and Customer Services to provide for the situation where one Customer Services Order contains many Products.
In my case, this applied when I went to the Restaurant and ordered a main course, yoghurt and wine.



6. Generic Data Mart

This shows our starting-point

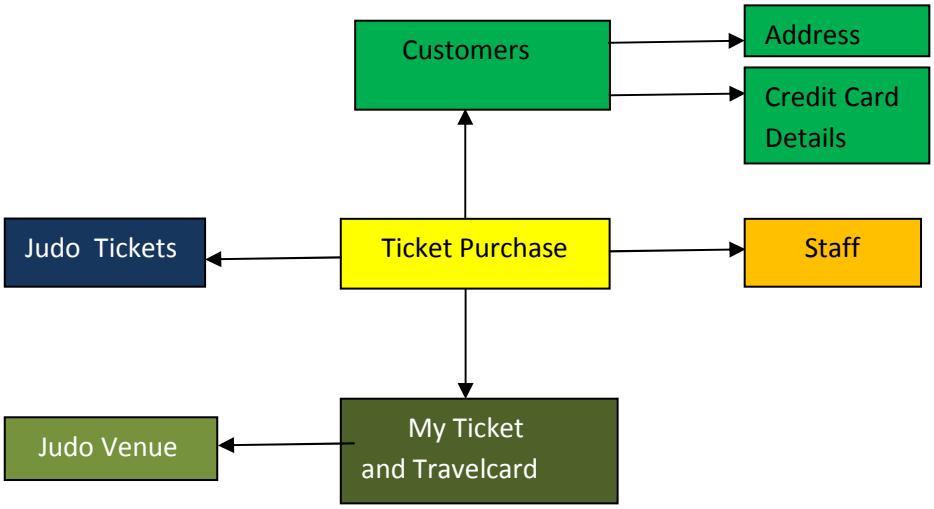


7. Event 1 – Buy Ticket for Judo Competition

This shows how we handle the first Event.

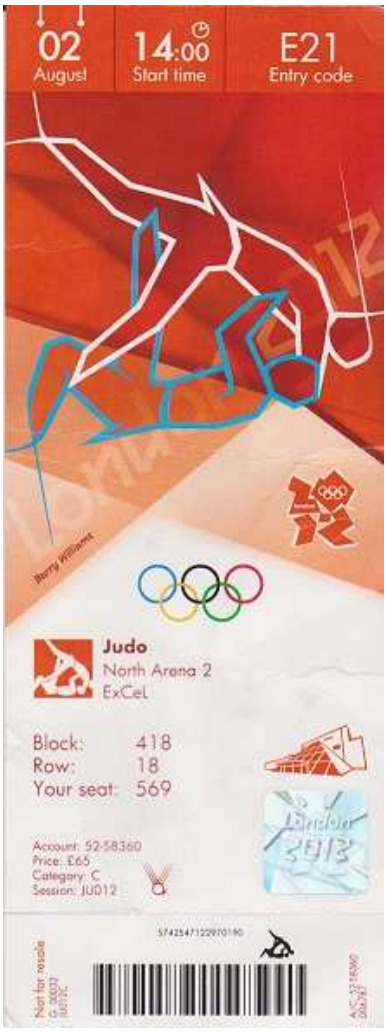
7.1 The Design Pattern

This shows how the Design Pattern applies to this Event.



7.2 Message Format

This shows the data items on the Ticket :-



This shows the fields in the Generic Message :-

EVENT	DATE	LOCATION	PRICE	DETAILS

This shows the fields in the Message for this Event :-

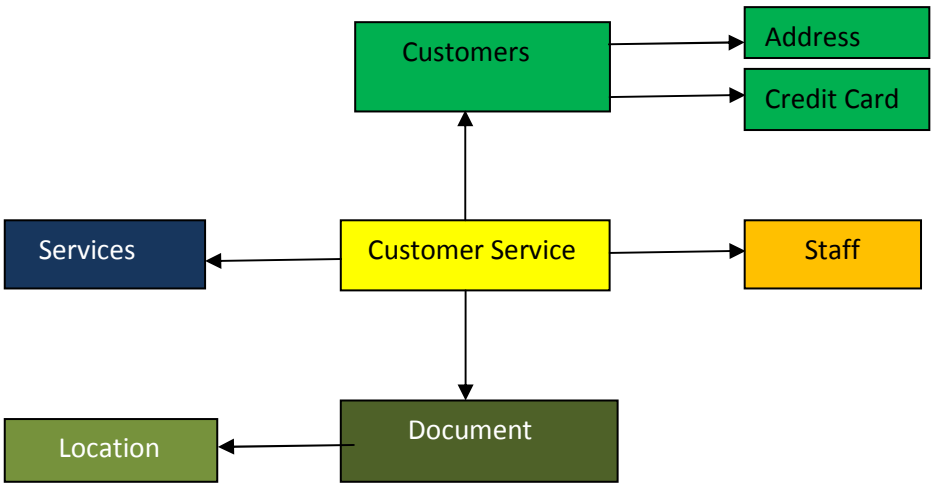
EVENT	DATE	LOCATION	PRICE	DETAILS
Purchase Ticket	Date of Purchase	Venue,Block, Row,Seat Nr	Seat Price	Date, Time of Competition

7.3 Data Warehouse

The benefit of adopting a Third-Normal Form ERD is that it enforces a 'Single View of the Truth'.

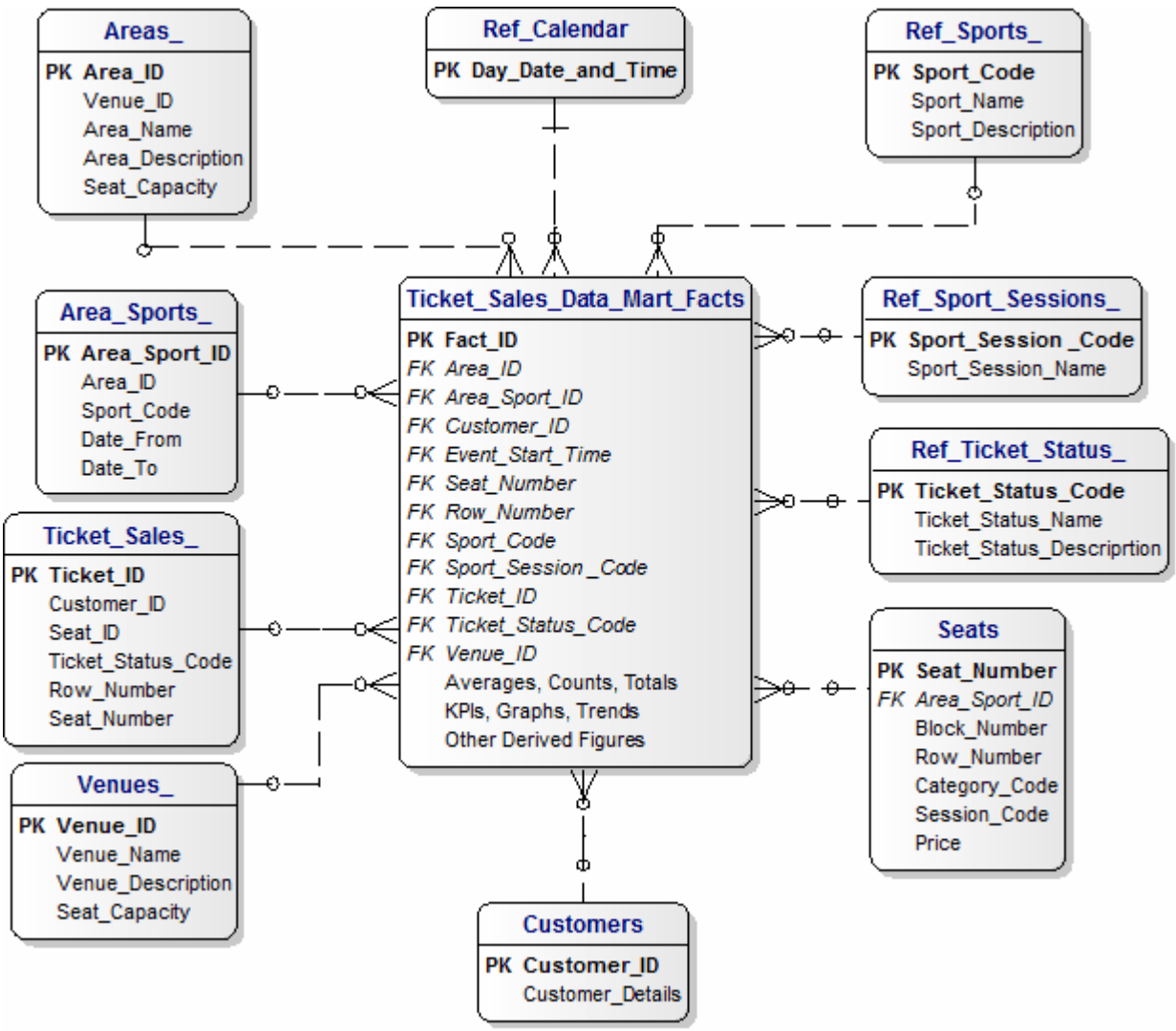
If we adopt a Dimensional Model it is not so easy to achieve this.

Thiis shows the design of the Data Warehouse (DWH) after the first Event of Purchasing a Ticket



7.4 Data Mart

This shows the Data Mart for Ticket Sales.

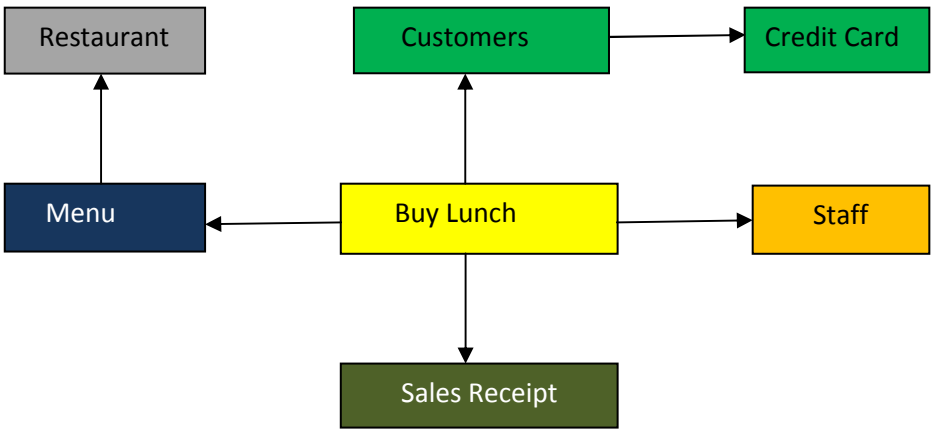


8. Event 2 – Get Lunch

This shows how we handle the Second Event.

8.1 The Design Pattern

This shows how the Design Pattern applies to this Event.



8.2 Message Format

This shows the fields in the Generic Message :-

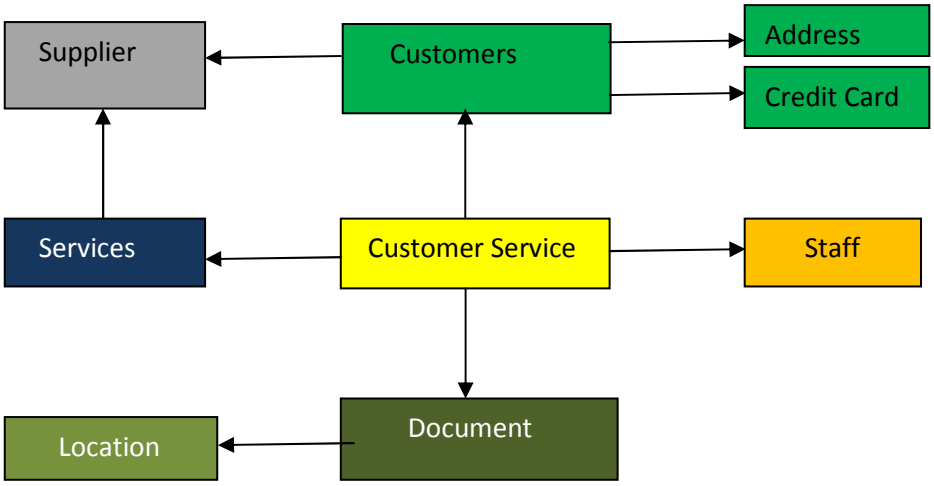
EVENT	DATE	LOCATION	PRICE	DETAILS

This shows the fields in the Message for this Event :-

EVENT	DATE	LOCATION	PRICE	DETAILS
Buy Lunch	Date & Time of Lunch	Restaurant	Total Price	Chicken and Mushroom Pie, Wine

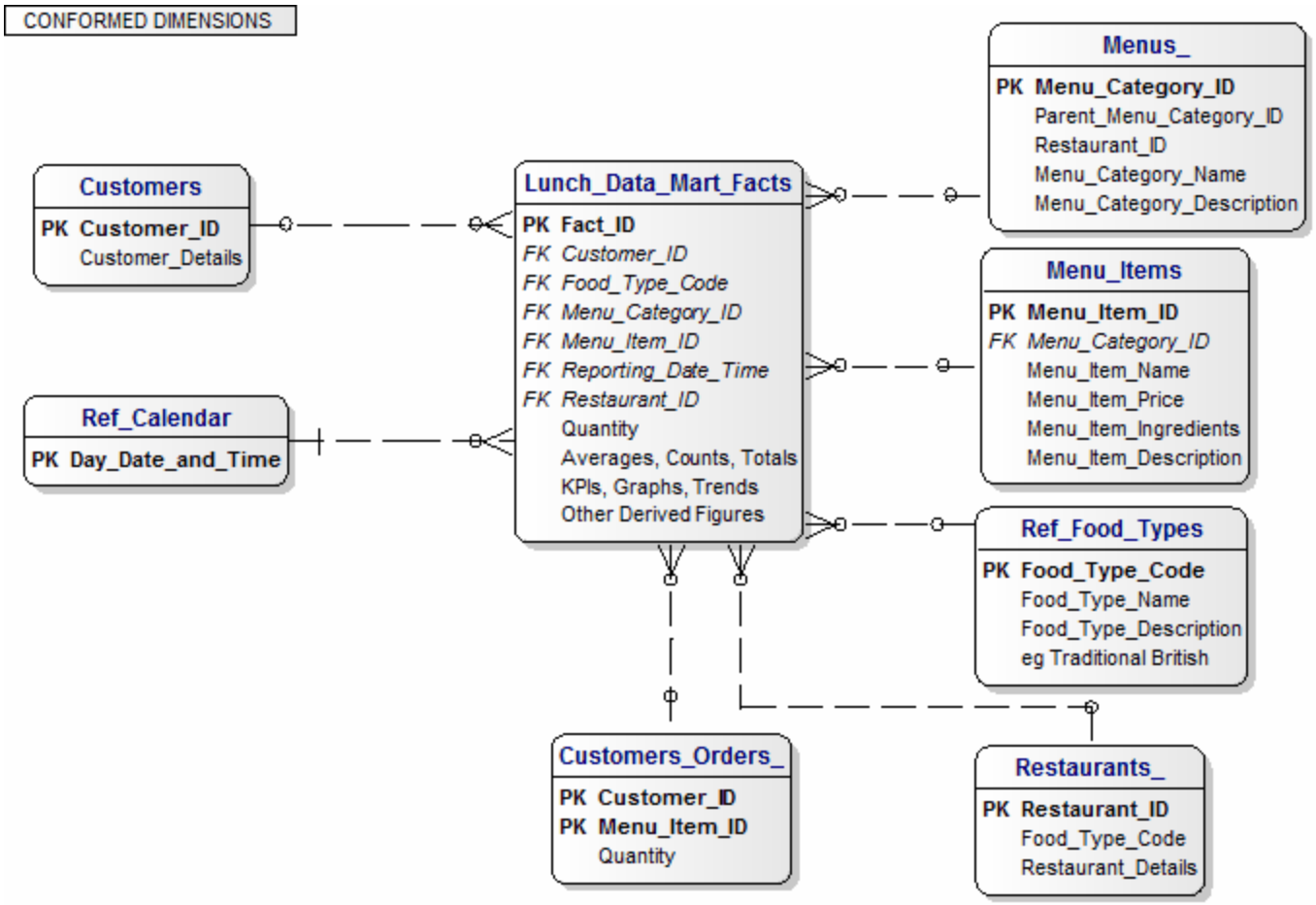
8.3 Data Warehouse

This shows the design of the Data Warehouse (DWH) after the second Event of Buying Lunch.



8.4 Data Mart

This shows the Data Mart for Restaurant data.



9. Event 3 – Watch the Judo Competition

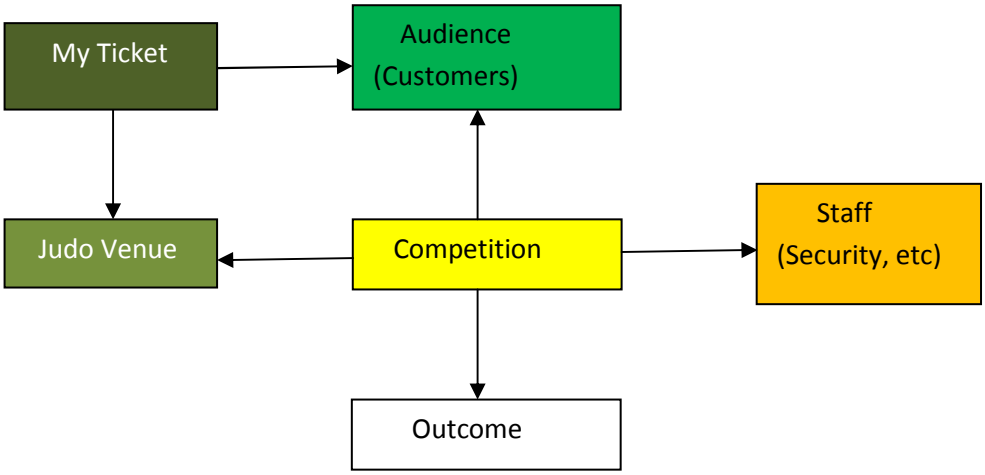
This shows how we handle the Third Event.

9.1 The Design Pattern

This shows how the Design Pattern applies to this Event.

In this case, the Event is the Competition between two Judo experts and the Outcome is very important

It is quite common for Event to have an Outcome, but so far, it has not been important enough to justify appearing at the top level.



9.2 Message Format

This shows the fields in the Generic Message :-

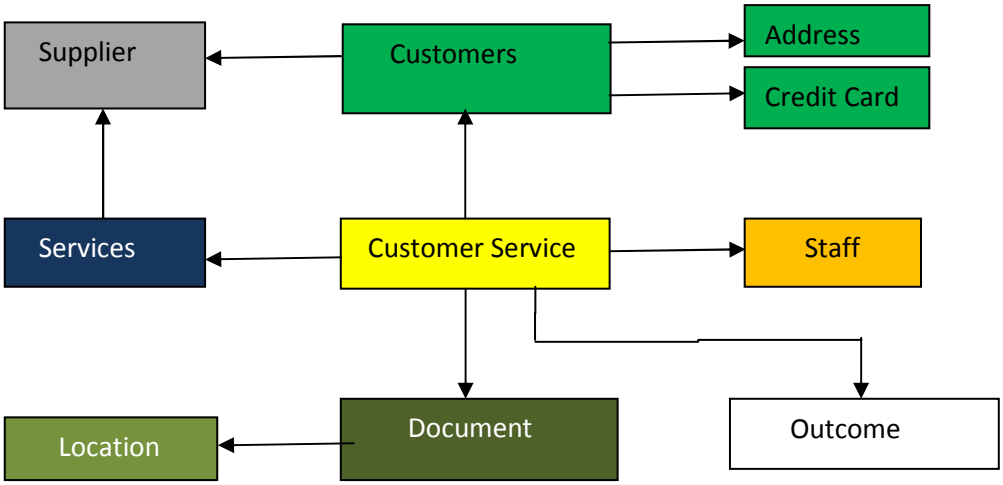
EVENT	DATE	LOCATION	PRICE	DETAILS

This shows the fields in the Message for this Event :-

EVENT	DATE	LOCATION	PRICE	DETAILS
Watch the Judo	Event Date	Judo Venue	Ticket Price	Outcome / Result

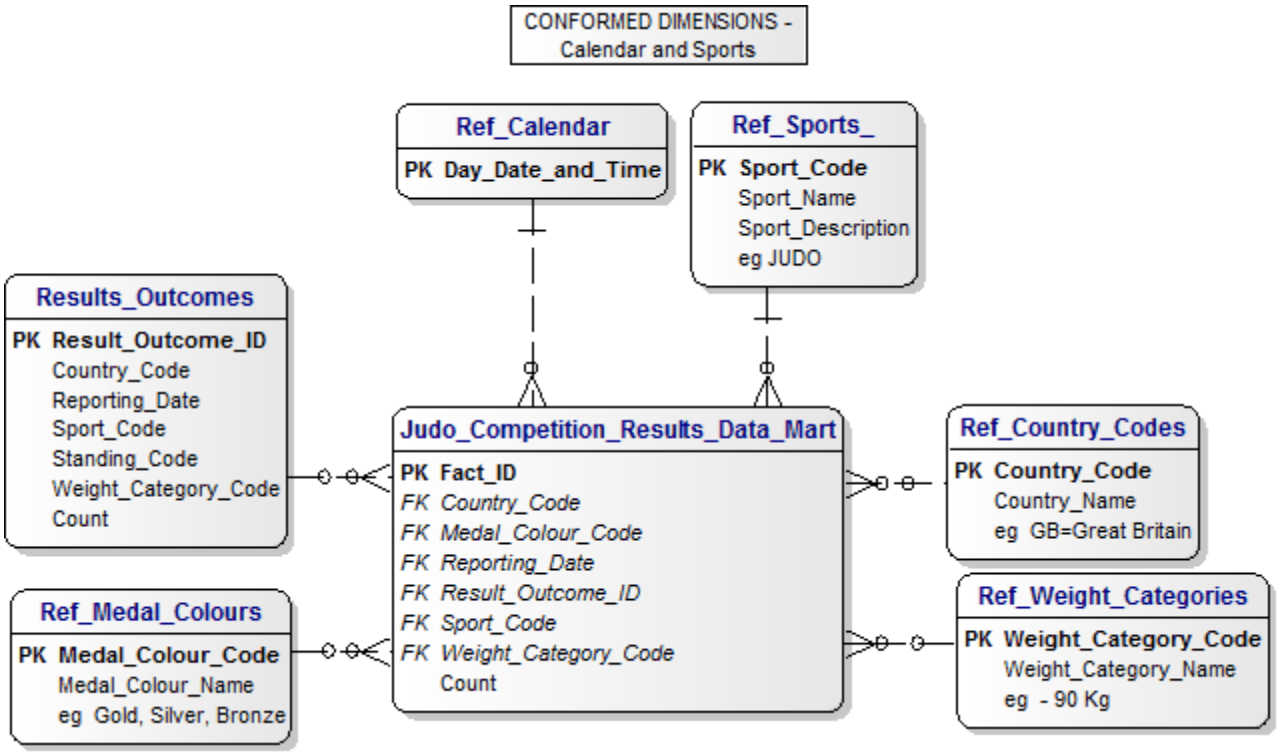
9.3 Data Warehouse

This shows the design of the Data Warehouse (DWH) after the third Event of watching the Judo Competition.



9.4 Data Mart

This shows the Data Mart for Judo Competition Results data.

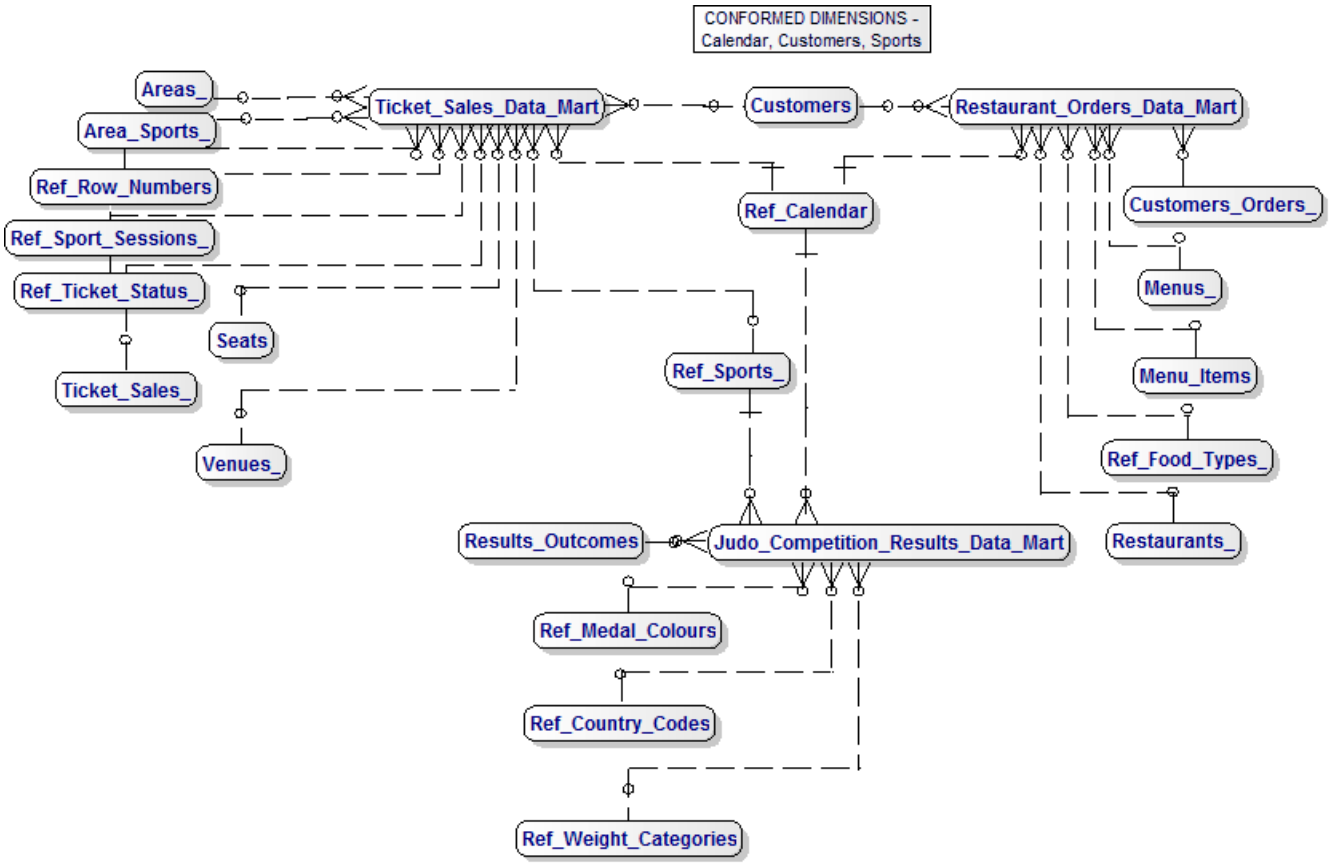


10. Combined Data Mart

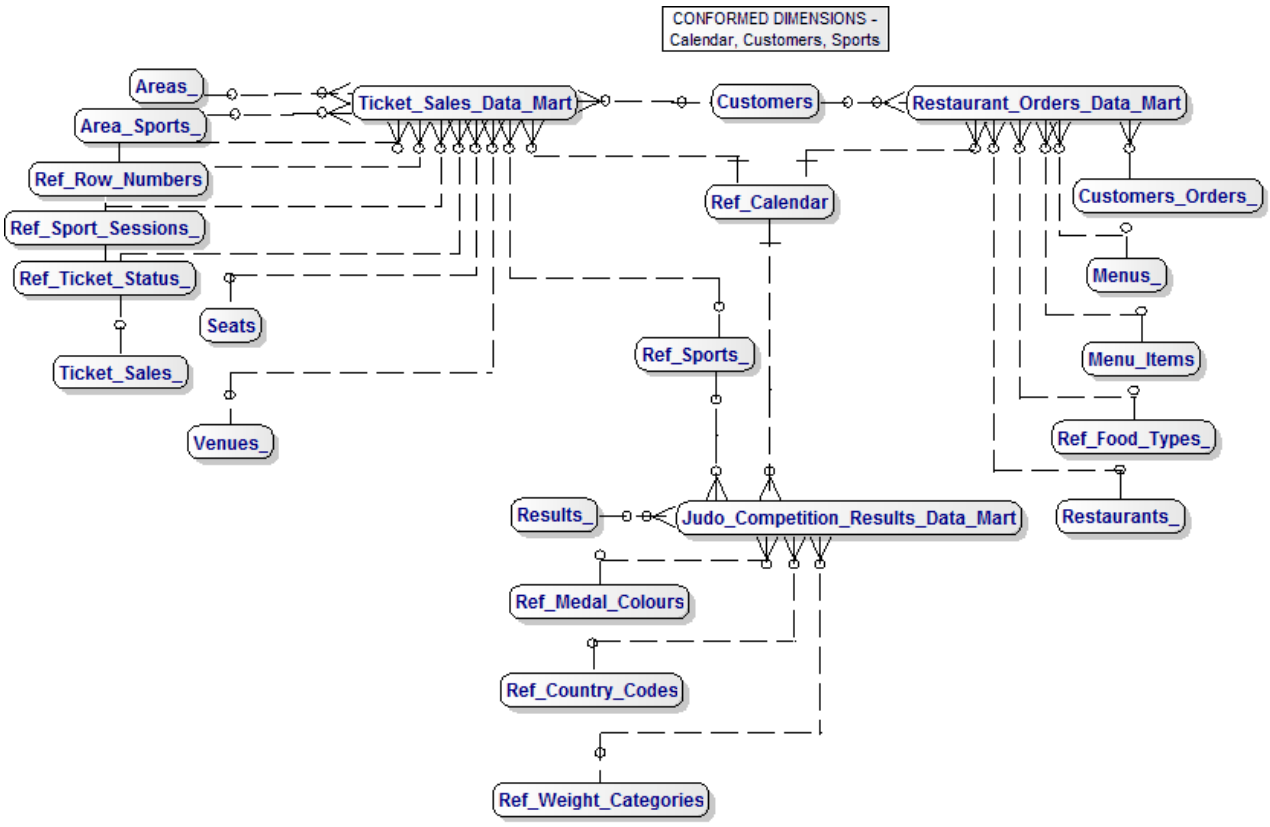
This shows the three Data Marts :-

- Judo Competition Results
- Restaurant Orders
- Ticket Sales

These share Conformed Dimensions of the Calendar, Customers and Sports.



Judo and Data Warehouses



11. Business Intelligence

11.1 A BI Layer

The reason for all the work that we have done to get to this point is, of course, to produce Business Intelligence ('BI').

Here is a simple example to show how this works in practice.

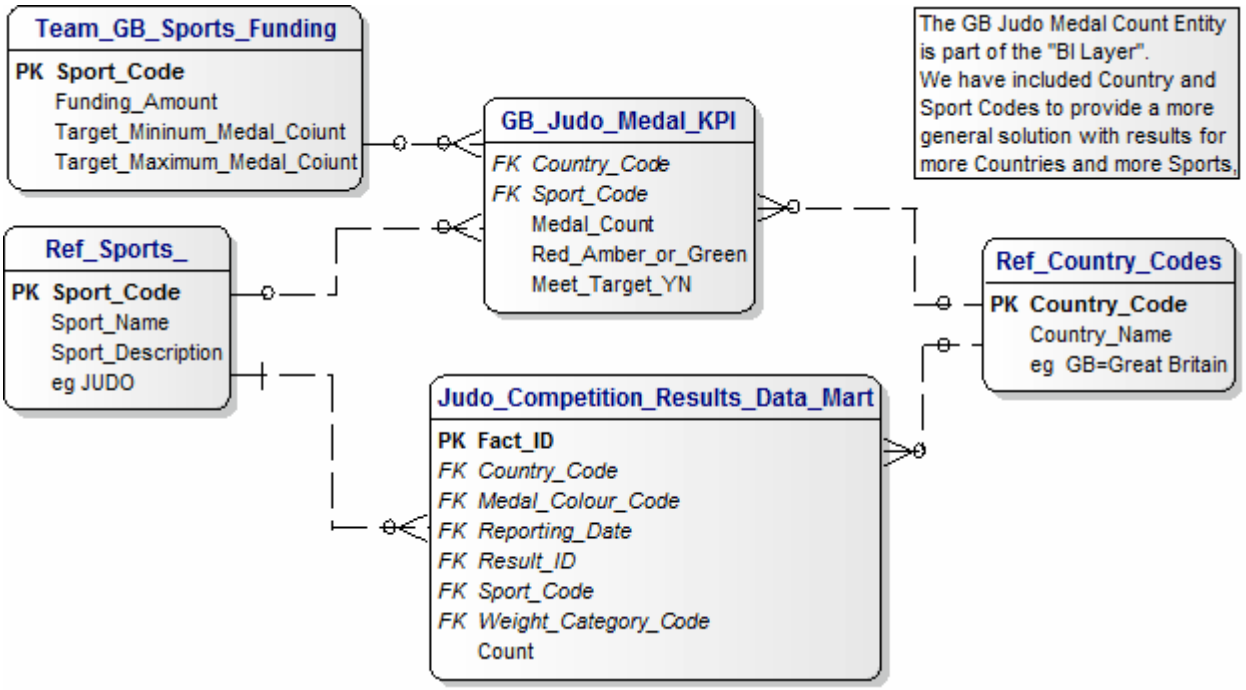
One of the key aspects of the Olympics was an analysis of the number of medals that each country won.

For Great Britain, funding was made available and targets were set for medals to be achieved by each sport.

For Judo, the funding was almost £8 million and the target was 1 or 2 medals.

In order to analyse this and produce the appropriate KPIs (Key Performance Indicators) we define a Bi Layer that takes data from the Judo Competition Results Data Mart.

This shows the Data Model for the Judo Medal Total BI Layer :-



11.2 BI Output

- This is a simple example of how this data could be displayed using a Green Traffic light :-

DISCIPLINE	BUDGET	TARGET	ACHIEVEMENT	TRAFFIC LIGHT (RED/AMBER/GREEN)
Archery				
Judo	£8 million	0-2 medals	2 medals	
Wrestling				

12. Conclusion

This Paper has presented a Method for designing a Data Warehouse following a Canonical Data Model and Messages.

We have validated the Method by designing a Data Warehouse for a Day at the Olympics.

I would be pleased to have your comments and you can email me at barryw@databaseanswers.org.

zzz