

## A Discussion Paper on Guidelines for Getting Started with Teradata

1. Management Summary .....	1
2. Getting Started.....	2
3. What Have We Learned? .....	11
4. The Author .....	11
Appendix A. Data Model Templates .....	12

### 1. Management Summary

There are three Phases in the lifecycle of a typical Teradata Data Warehouse :-

1. Getting Started
2. Assessment of an operational Data Warehouse
3. Business As Usual (BAU) – changes and enhancements

This Paper discusses Phase 1 – Getting Started – and has the objective of promoting discussion of Best Practice.

The later two Phases are described in a separate Paper which is available on demand from Database Answers at [barryw@databaseanswers.org](mailto:barryw@databaseanswers.org).

We use Transportation as the Application to discuss our Approach but it applies equally well to any other area for which a Teradata Industry Data Model is available.

#### Phase 1 – Getting Started

This incorporates a Blueprint and will incorporate any appropriate Teradata Industry Data Model.

It will then involve design of a Conceptual High-level Model and Subject Areas Models.

This Chapter provides material that is suitable for presentations, preparation of a business case, clarification of the requirements or to establish the scope of a Data Warehouse.

Presentations can be prepared for specific Industry Areas using this material, supplemented with 'Kick-Start' Data Model Templates in Appendix A for Banking, Communications, Retail, Transport and Web Visit Statistics.

## 2. Getting Started

## 2.1 What is this ?

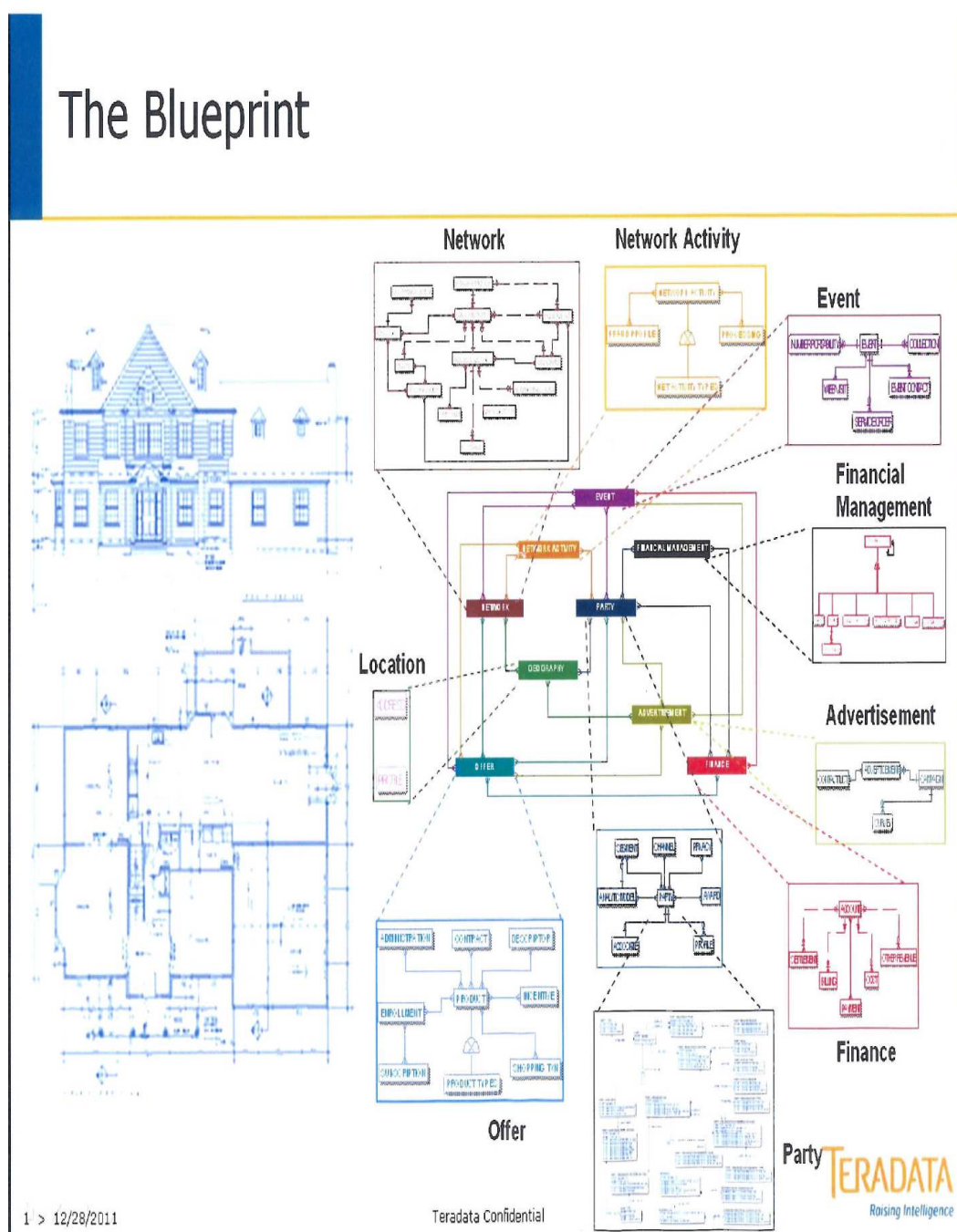
This Chapter describes how to get started with a Teradata Project.

## 2.2 Why is it important ?

It is important because it helps us to understand the overall process and concepts of a Data Warehouse.

## 2.3 A Blueprint

This is a Teradata Blueprint for an Architecture where rooms become Subject Areas in a Data Model.



## 2.4 A Transport Enterprise Data Model

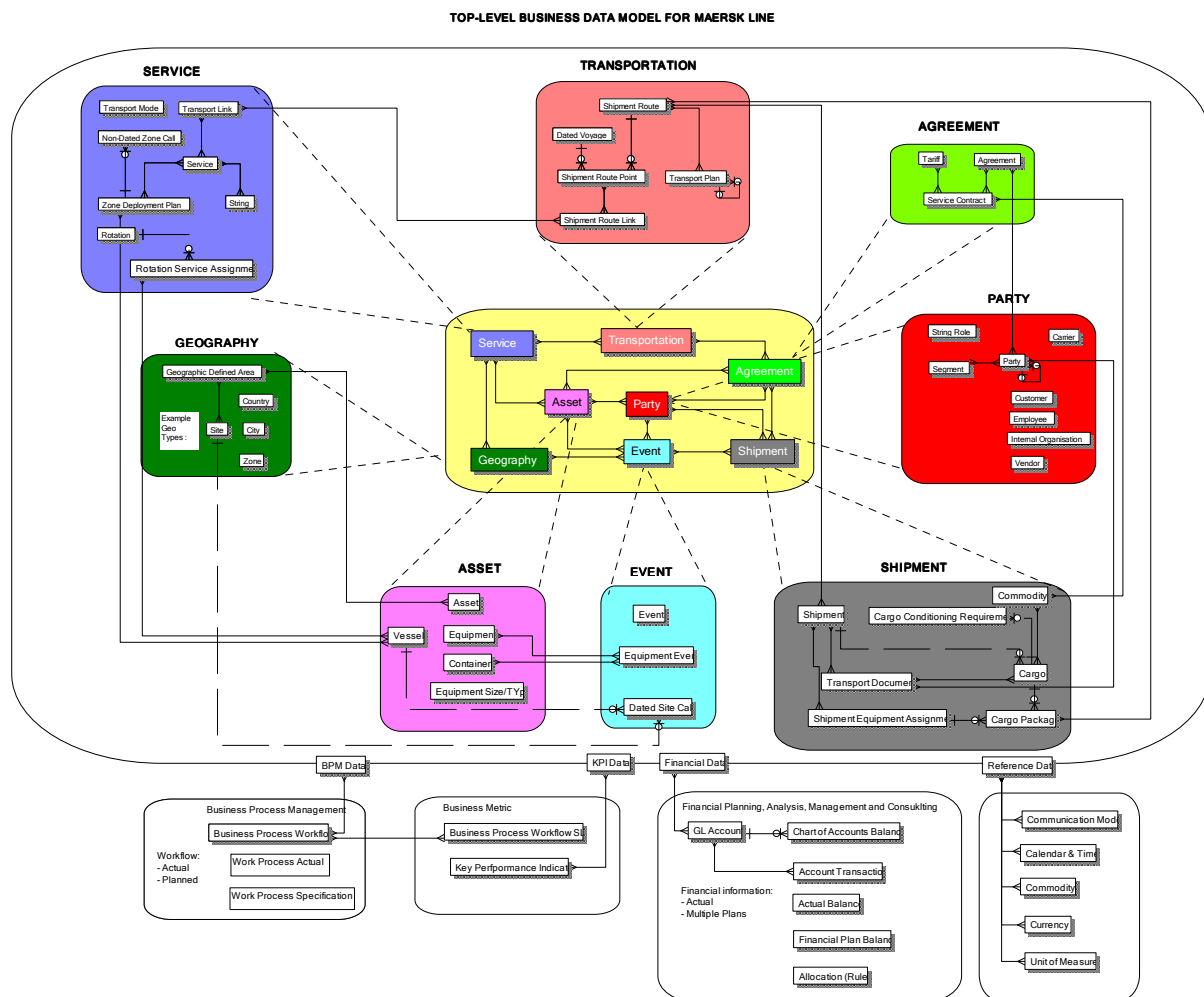
This shows an extract of the Teradata Transport and Logistics Data Model.

It acted as a starting-point for the creation of a version of the Model for Maersk, which was the Teradata Customer.

The yellow area in this diagram shows an EDM with separate Subject Areas :-

- Agreement, Asset, Event, Geography, Party, Service, Shipment and Transportation.

These are shown in larger sizes around the yellow rectangle.



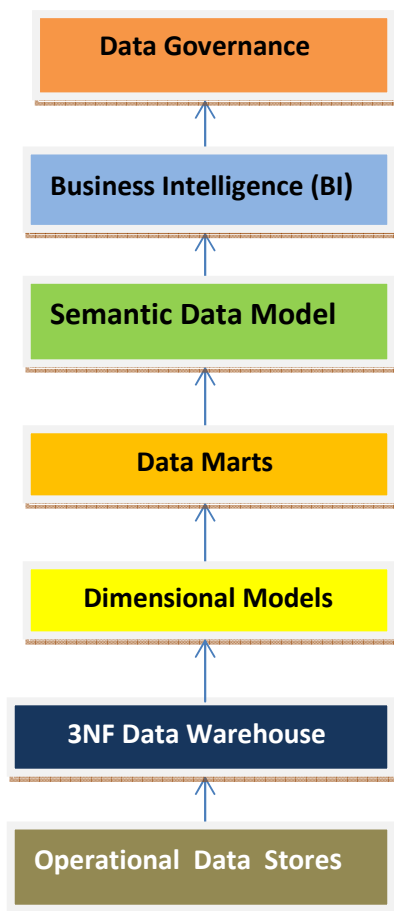
## 2.5 A Data Architecture and BIReady

This shows six Layers in a Data Architecture for Data Warehouses.

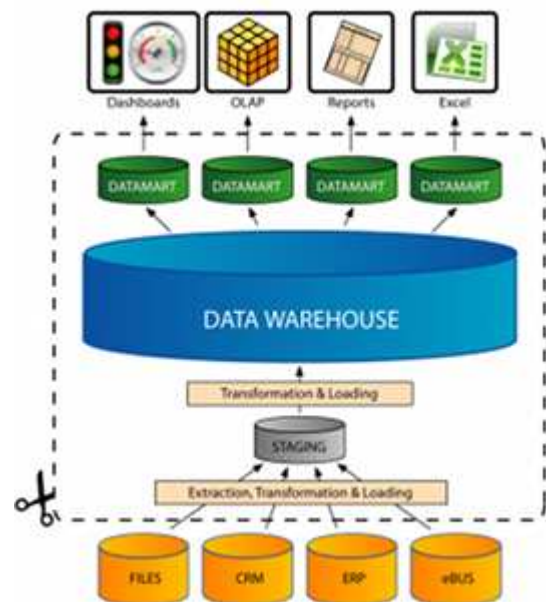
On the left-hand side we show the layers in a Teradata Data Architecture.

On the right-hand side, we use a product called BIReady that provides the functionality within the dotted lines which corresponds to the Data Integration and Data Marts Stages in the Road Map.

In other words, [BIReady](#) automates the implementation of ETL and ELT functions, the Data Warehouse and the Data Marts.



BIReady facilities are shown in the dotted square



## **2.6 Stages in loading Data into a Warehouse**

There is a logical sequence to the Stages of loading data into a Warehouse.

The lowest levels of data provide a foundation. The lowest level is Reference Data, such as Currencies and Languages.

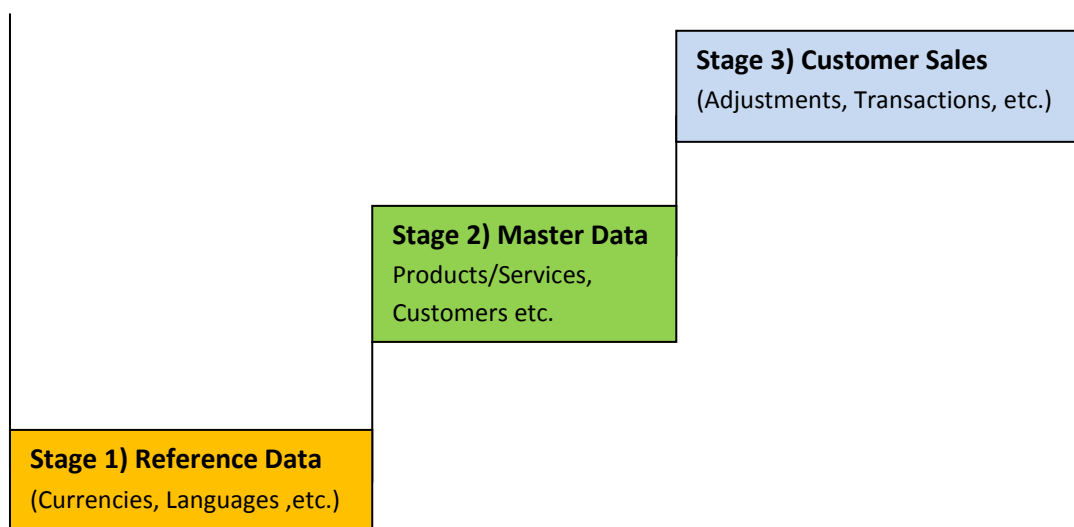
Then, after these have been identified, sources and loaded we can move to the Master Data.

This includes Products, Services, Customers and so on.

Then we come to the more dynamic data such as Sales, Transactions and so on.

Best Practice states that we should set up load procedures that can be run so that data at each Stage can be deleted and replaced repetitively.

This shows three Stages of data in a typical Data Warehouse.



### 2.7 Business Intelligence (Top-down)

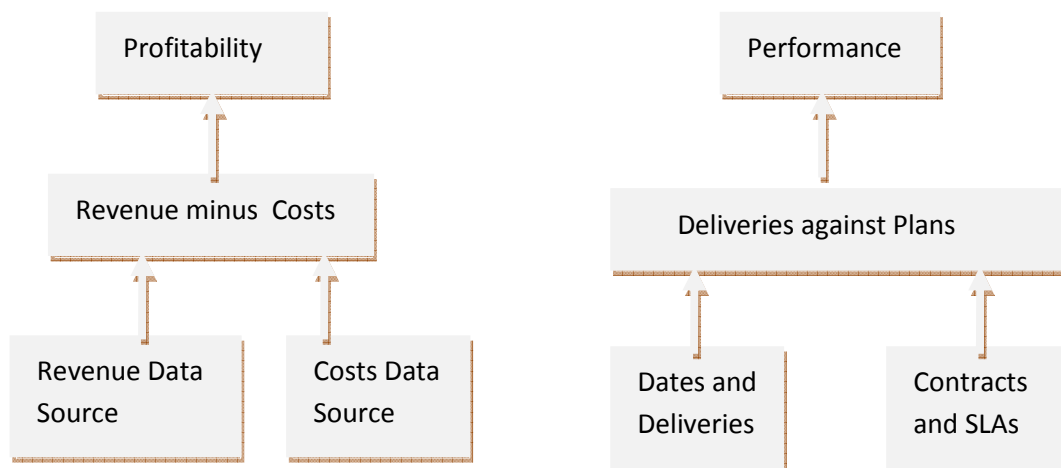
Key Performance Indicators ('KPIs') provide an excellent starting-point for Business Intelligence.

Basic KPIs can be defined for Profitability and Operational Performance.

These can be enhanced in collaboration with the business users to establish commitment.

Profitability can initially be defined as a basic difference between Profits and Revenue.

Performance can be measurement of Deliveries being 'On-Time' against plans agreed in Service Level Agreements (SLAs).



The initial Data Architecture will include this basic structure of KPI and Source Data.

This will provide a simple but flexible framework which can be enhanced in a controlled manner.

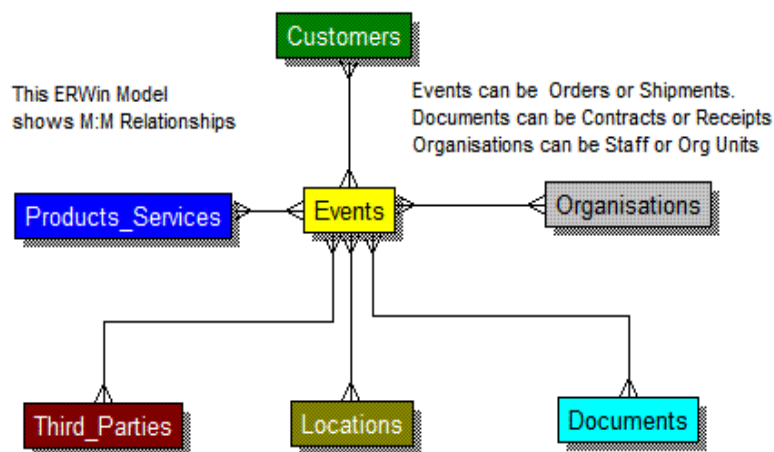
We will use this as a starting-point for discussions to establish the requirements with the business users.

### 2.8 Build the Data Warehouse Incrementally

A Common Data Model can be used to design a Data Warehouse in a series of incremental Steps that provide a validation with each Step.

#### 2.8.1 Canonical Data Model

This shows a Common Data Model that provides an excellent Starting Point for building an Enterprise Data Warehouse.



#### 2.8.2 Event Message

We can define Events with corresponding Messages.

For each Message, we can add to the design of our Data Warehouse.

The strength of this approach is that it validates the design of the Data Warehouse.

This Common Data Model defines a Message for each Event.

### 2.8.3 Example Event Message (1)

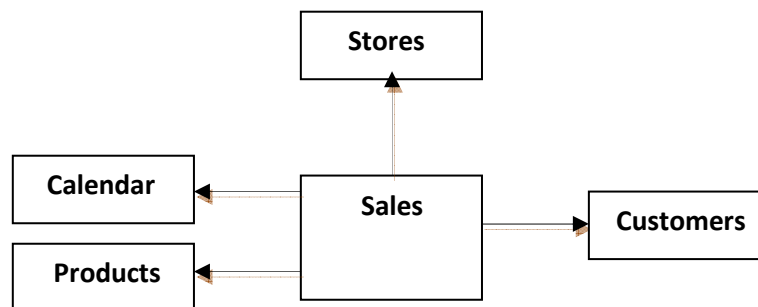
For example, a Customer purchases one Product at a Store :-

Date	Store ID	Customer Name	Product	Quantity	Amount
------	----------	---------------	---------	----------	--------

The Business Rules would say :-

1. Each Sale is made to one Customer
2. Each Sale can involve just one Product
3. Each Sale is made at a specific Store
4. Each Sale is made at a specific valid Date and Time which is defined in a Calendar

Therefore we can define this fragment of a normalised Data Model :-





## A Discussion Paper on Guidelines for Getting Started with Teradata

### 2.8.4 Event Message (2)

In this Event, the Customer buys two Products. Therefore our initial Data Model must be enhanced to have a separate Table for 'Products Sold'.

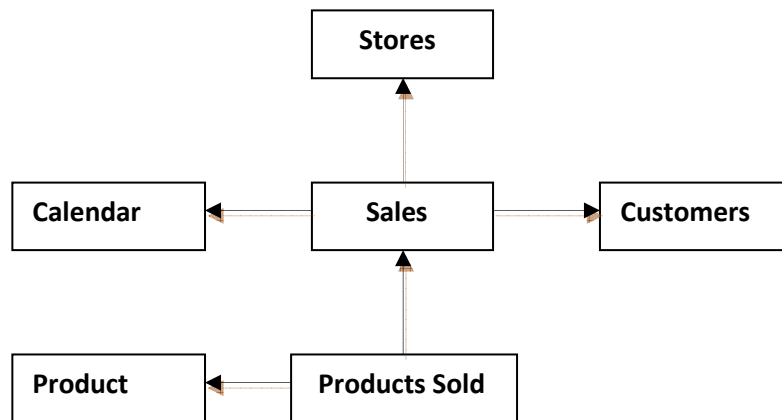
When we continue to include more Events, we enhance our Model until it is complete.

Date	Store ID	Customer Name	Product 1	Quantity 1	Amount 1	Product 2	Quantity 2	Amount 2	Total Amount
------	----------	---------------	-----------	------------	----------	-----------	------------	----------	--------------

This Event changes one of the Business Rules :-

1. Each Sale can involve multiple Products

Therefore we add a separate entity called 'Products Sold' to our Data Model :-

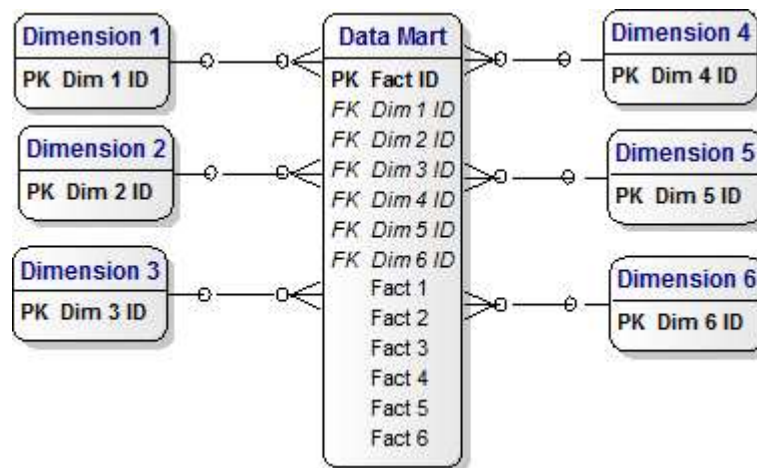


## 2.9 Generic Dimensional Model

The existence of Dimensions and Facts makes it possible to define a Generic Design for a range of Dimensional Models.

This example shows six Dimensions and six Facts.

Each of these Dimensions and Facts can be translated from the generic names to specifics, such as Customer, Product, Location, Time Period and so on.

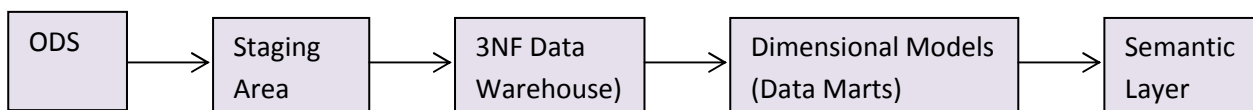


## 2.10 Simple Data Architecture

A Data Architecture should be composed of separate Components with separate and clearly-defined functionality.

The 3NF Data Warehouse implements data integration and provides a 'Single View of the Truth' (SVOT).

The objective is to show the major Building Blocks in a simple clean Architecture.



A more detailed version is shown in the Data Warehouse Bus Architecture in Chapter 3.

## 2.11 Sample Kick-Start Data Warehouses

Some sample Data Warehouses are included in Appendix A for a starting-point in discussions.

### **3. What Have We Learned?**

We have learned some very important rules that will help us to get started in an organized manner and put in place a sound foundation that we can build on in a safe and stable manner.

### **4. The Author**

Barry Williams is the founder and principal consultant with Database Answers.



His company has been providing advice and assistance to blue-chip clients for over 20 years.

His particular interest is in advancing the role of data models as a way of improving communication between the business user community and data management professionals.

As part of this role he publishes best practice on his Database Answers Web site at

<http://www.databaseanswers.org/>

## Appendix A. Data Model Templates

### A.1 What is this ?

These Templates provide introductory Data Models for some specific Applications.

### A.2 Why is it important ?

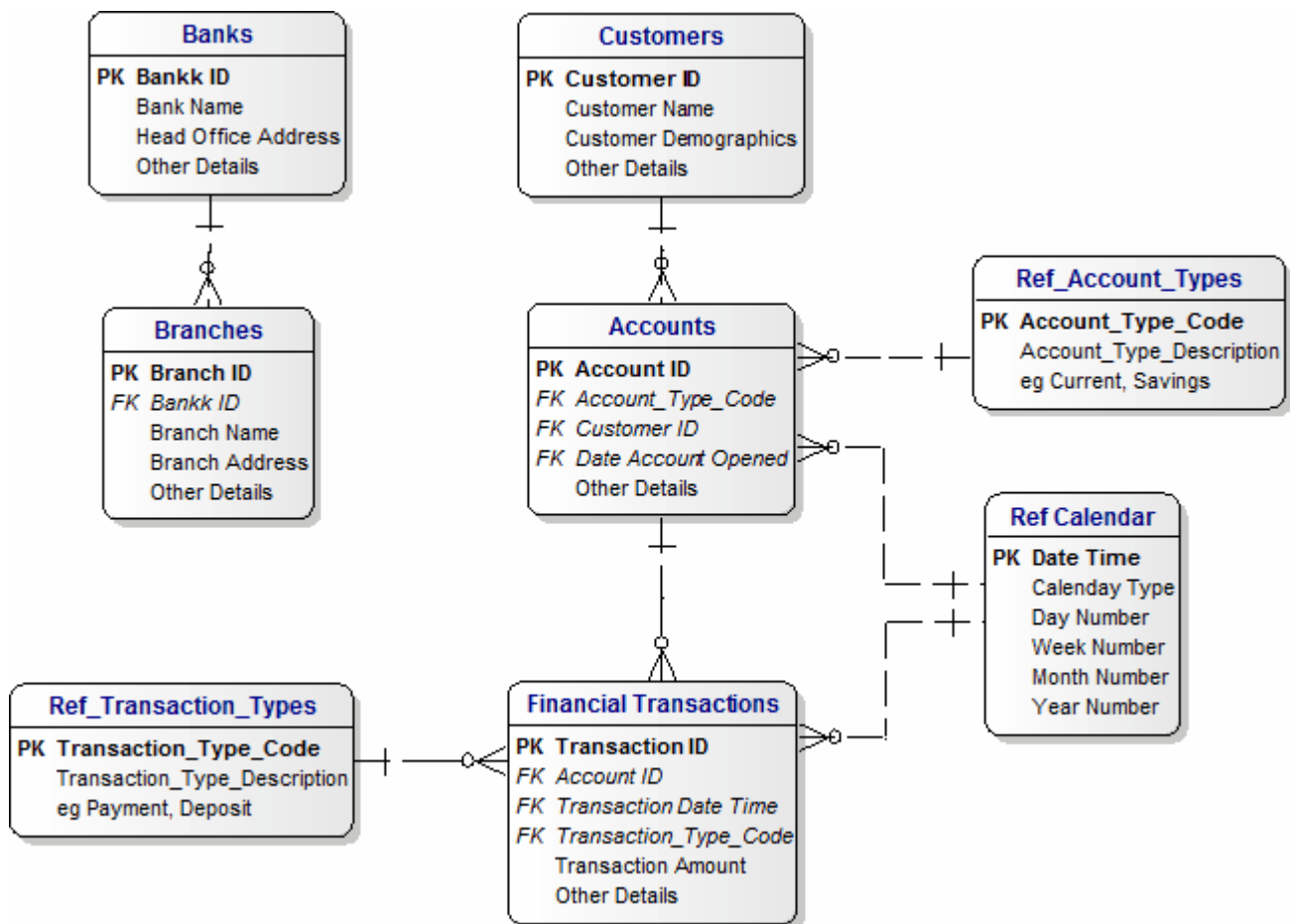
These Templates provide 'Kick-Start' samples to get up and running quickly.

They also provide samples that can be used to assess existing Data Models for 'fit for purpose'.

### A.3 Banking

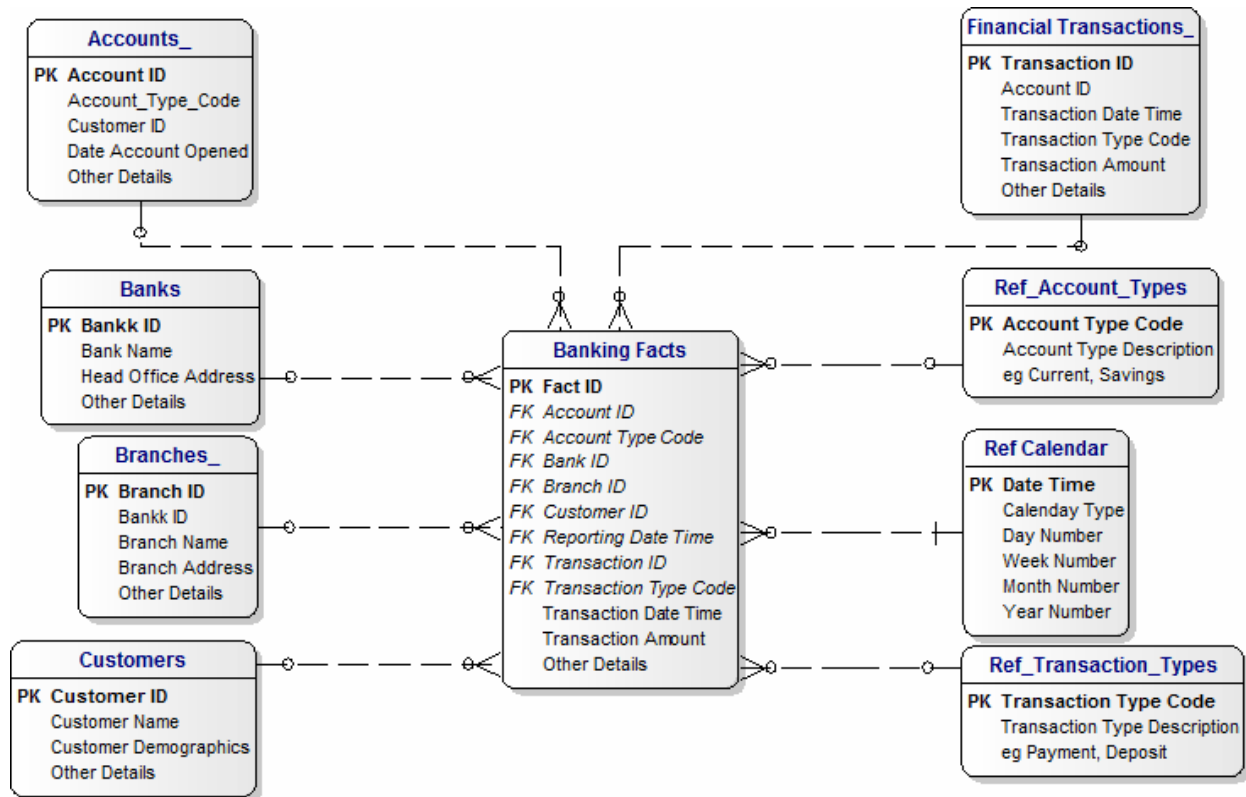
#### A.3.1 Data Warehouse

This shows theTemplate EDW for the analysis of retail Banking.



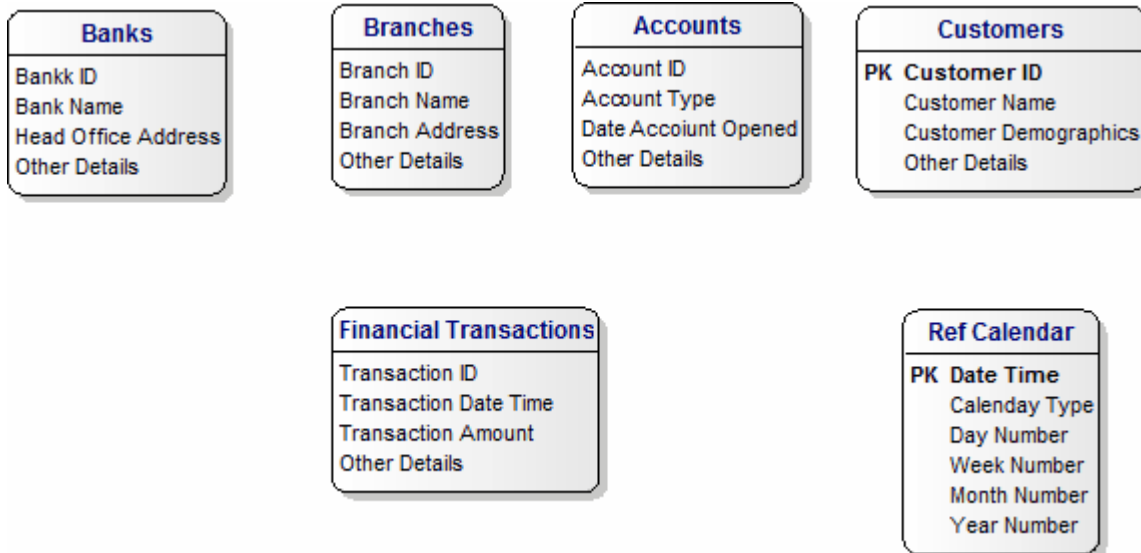
### A.3.2 Dimensional Model

This shows the Template Dimensional Model for retail Banking.



### A.3.3 Operational Data Stores (ODS)

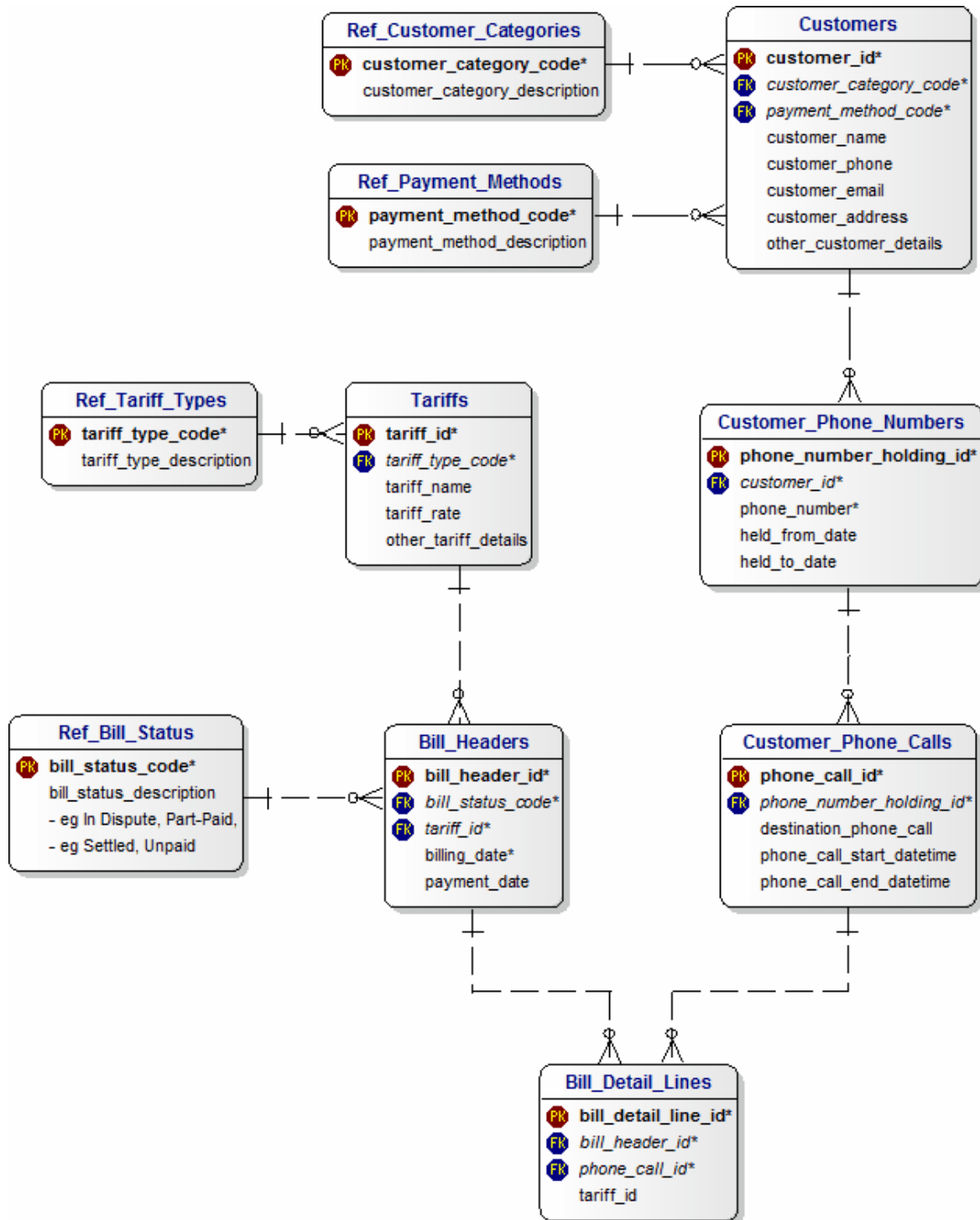
This shows the ODS for the analysis of retail Banking.



## A.4 Communications

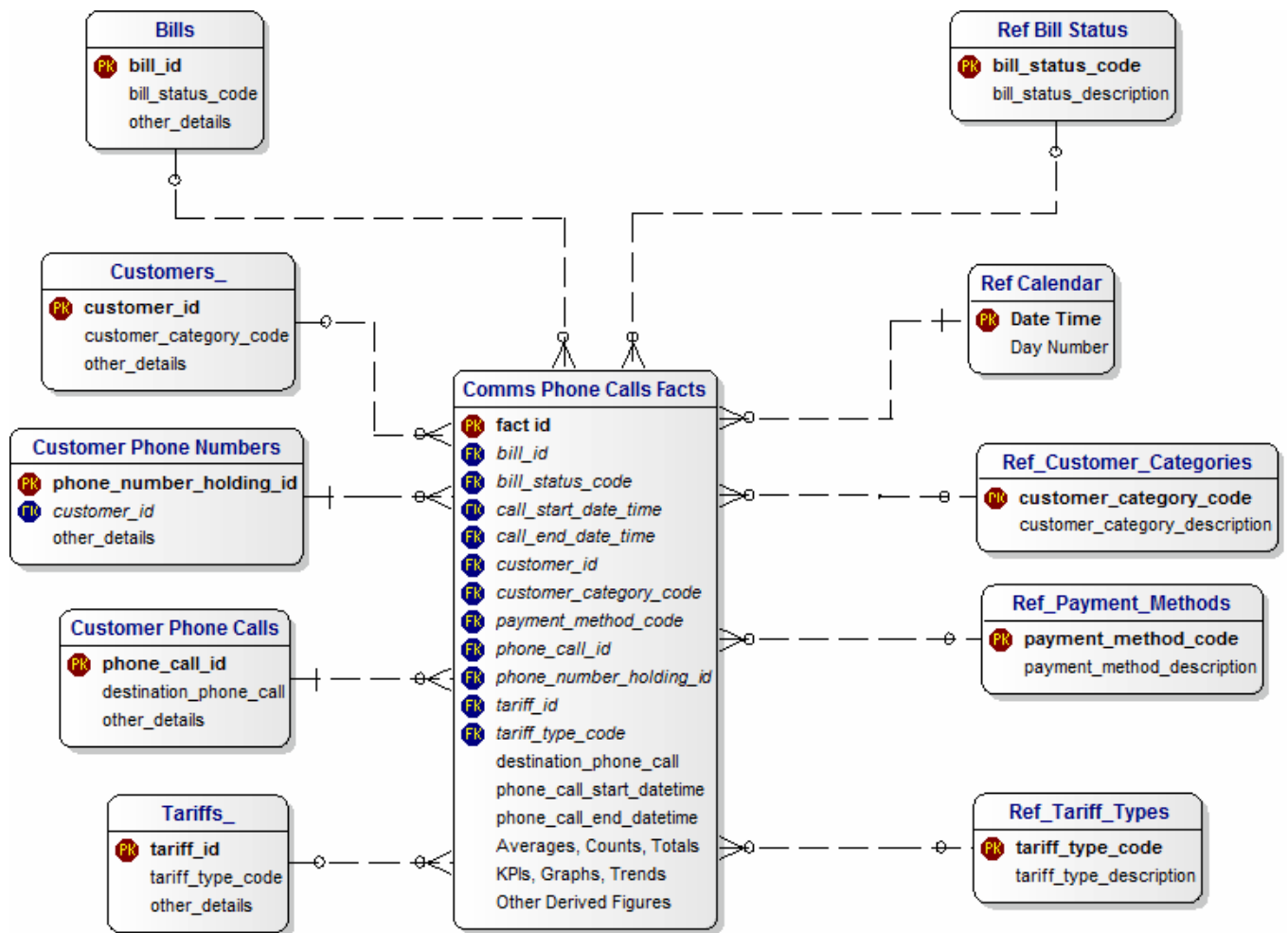
### A.4.1 Data Warehouse

This shows the Template EDW for Communications and, in particular, Telephone Billing Systems.



### A.4.2 Dimensional Model

This shows the Template Dimensional Model for Communications.





### A.4.3 Operational Data Stores (ODS)

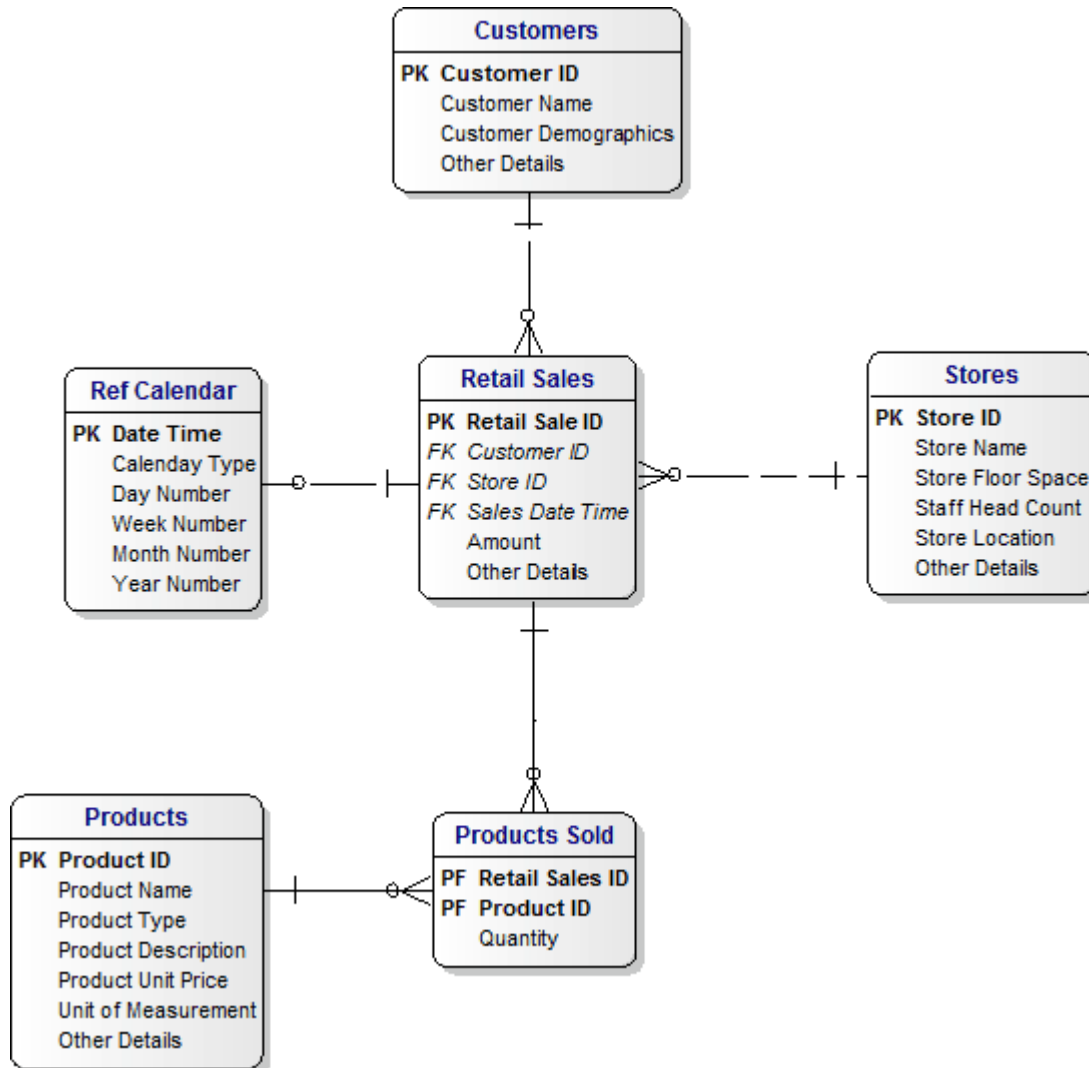
This shows the ODS for Communications.



## A.5 Retail Sales

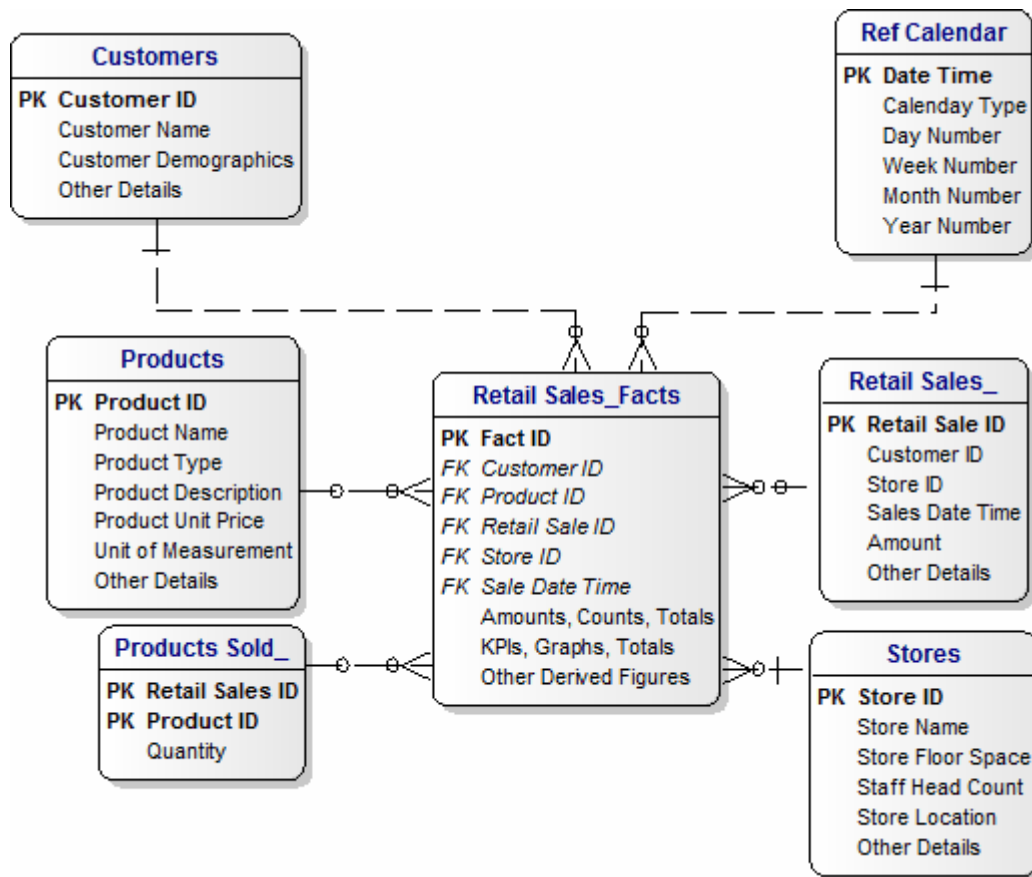
### A.5.1 Data Warehouse

This shows the EDW for the analysis of Retail Sales.



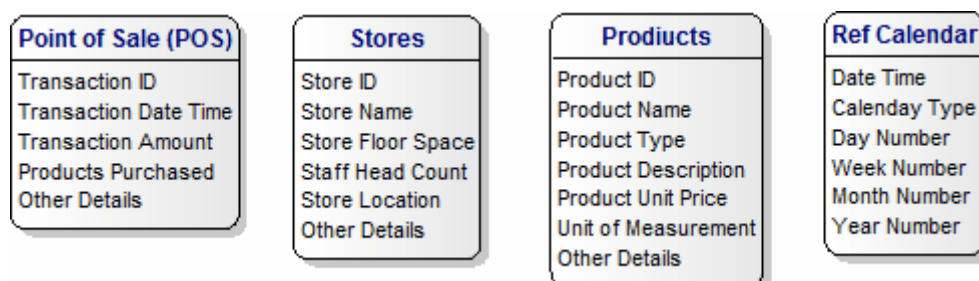
### A.5.2 Dimensional Model

This shows the Dimensional Model for the Retail Sales.



### A.5.3 Operational Data Stores (ODS)

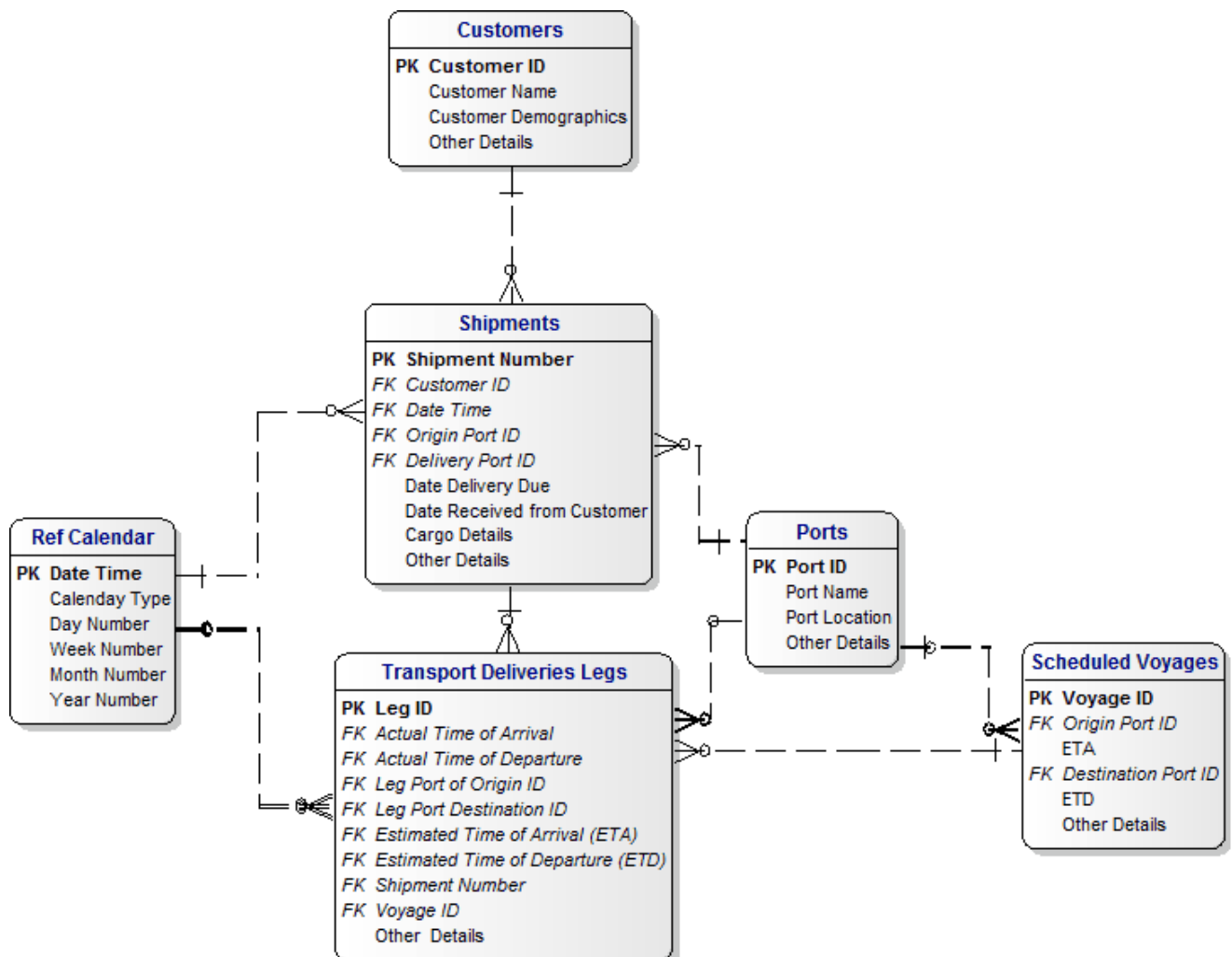
This diagram shows a sample of extract data for Retail Sales.



## A.6 Transport

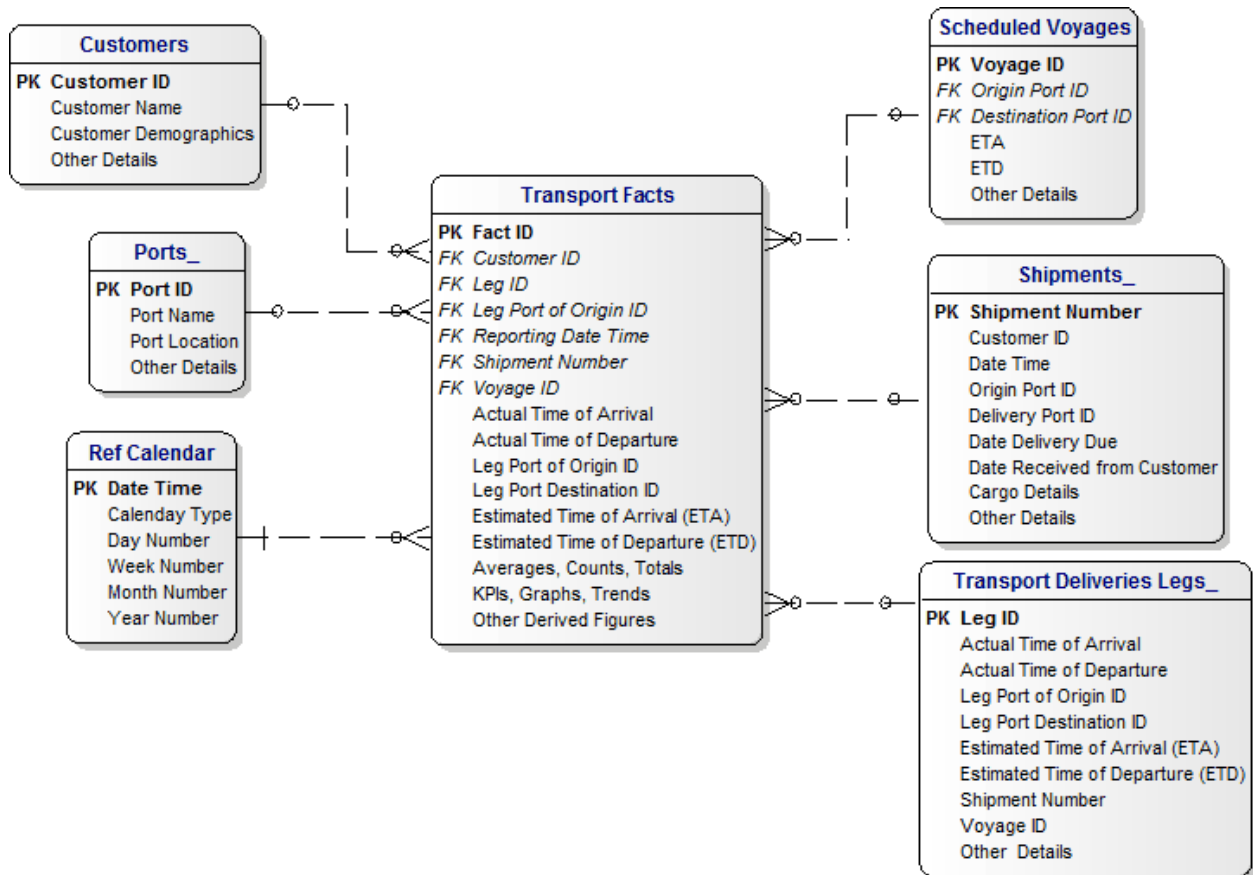
### A.6.1 Data Warehouse

This shows the EDW for the analysis of Transport.



### A.6.2 Dimensional Model

This shows the Dimensional Model for the Transport Subject Area.



### A.6.3 Operational Data Stores (ODS)

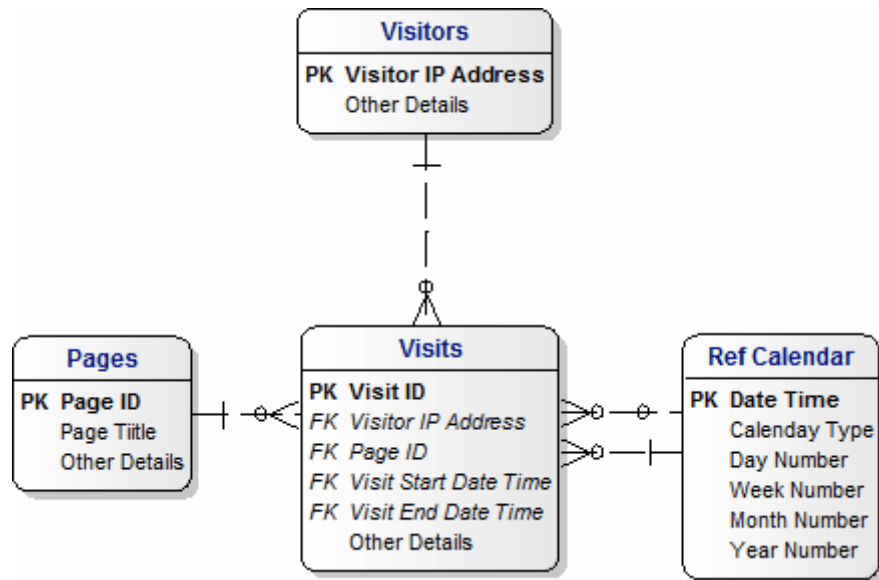
This diagram shows a sample of extract data for Transport.



## A.7 Web Visit Statistics

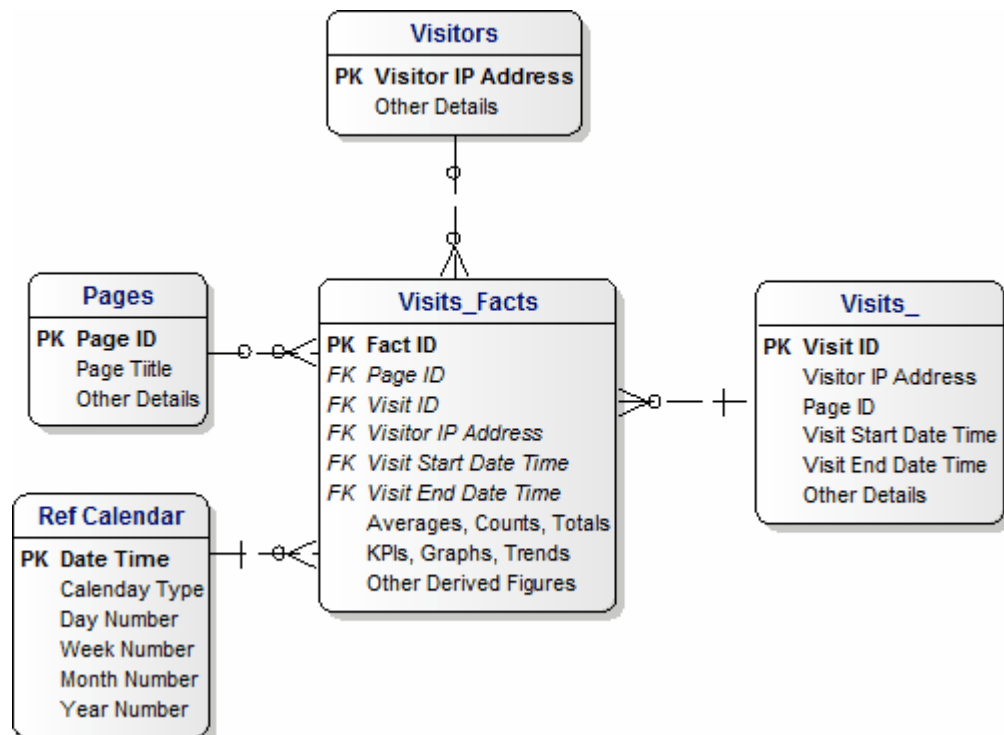
### A.7.1 Data Warehouse

This shows the EDW for the analysis of Transport.



### A.7.2 Dimensional Model

This diagram shows a Dimensional Model for Web Visit Stats.



### A.7.3 Operational Data Stores (ODS)

This diagram shows a sample of extract data for Web Visit Stats.

