

# *Database Answers*

## **Tutorial in Generic Data Modelling by Example**



**Barry Williams**

1. Management Summary .....	3
Step 1 : Agree the User Scenario .....	4
Step 2 : Define a Generic Customers Model.....	5
Step 3 : Define a Generic Customer Services Models .....	12
Step 4 : Define a Generic Payments Models .....	15
Step 5 : Define a Generic Reservations Model .....	19
Step 6 : Define a Generic Travel Model.....	21
Step 7 : Review progress.....	24
Step 8 : Adopt the Canonical Data Model.....	25
Step 9 : Define Top-Level Model with Customers .....	26
Step 10 : Define Top-Level Model with Parties .....	27
Step 11 : Review the List of Generic Data Models.....	27
Step 12 : Review the List of Subject Area Data Models .....	27
Step 13 : Design the Data Warehouse .....	28
Step 14 : Design the Data Marts .....	29
Step 15 : Draw our Conclusions.....	33

## 1. Management Summary

### 1.1 Our Objective

Our objective is to explore the use of Subject Area and Generic Data Models to build a Top-Level Model with associated Data Warehouse and Data Mart.

### 1.2 Our Approach

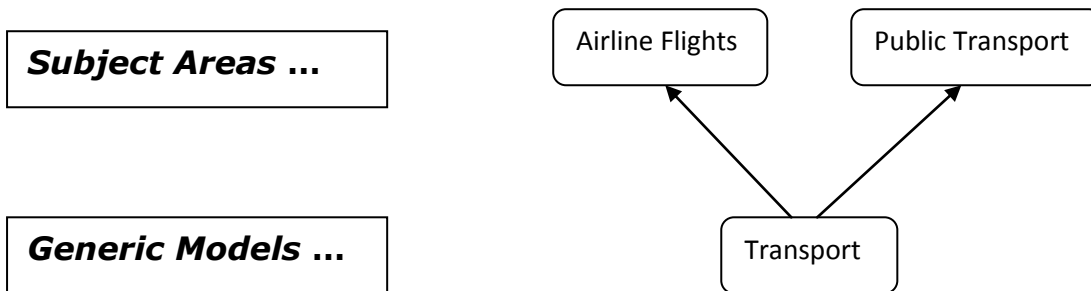
Our Approach is to use our Generic Models as Building Blocks to construct more complex Subject Area Models and then assemble the Subject Area Models into a Top-Level Model.

These Top-Level Models can then be expanded into Enterprise Data Models.

This Approach provides an incremental method of developing an Enterprise Data Model. with close user involvement along the way.

This has many benefits compared to the alternative method.

This shows Travel as an example of how our Approach works in practice :-



If you are interested in the details, you can check out these two Models :-

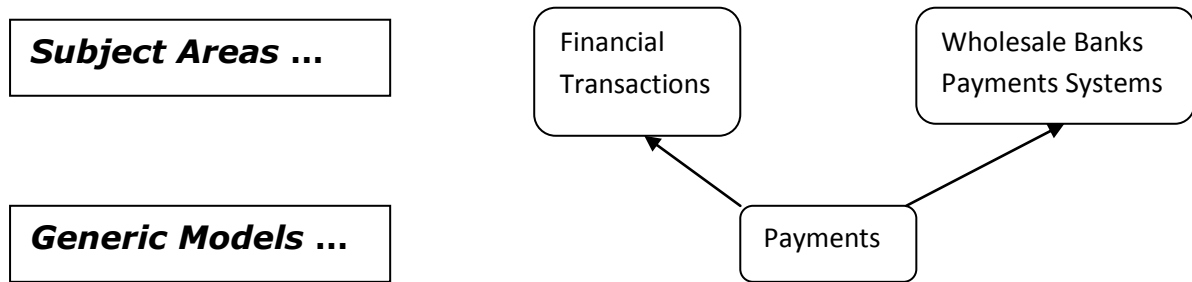
Airline Flights

- o [http://www.databaseanswers.org/data\\_models/airline\\_operations/index.htm](http://www.databaseanswers.org/data_models/airline_operations/index.htm)

Trains and Boats and Planes -

- o [http://www.databaseanswers.org/data\\_models/trains\\_and\\_boats\\_and\\_planes/index.htm](http://www.databaseanswers.org/data_models/trains_and_boats_and_planes/index.htm)

This shows Payments as another example of our Approach :-



If you are interested in the details, check out these Models :-

- Payments
  - [http://www.databaseanswers.org/data\\_models/payments/index.htm](http://www.databaseanswers.org/data_models/payments/index.htm)
- Financial Transactions
  - [http://www.databaseanswers.org/data\\_models/financial\\_transactions/index.htm](http://www.databaseanswers.org/data_models/financial_transactions/index.htm)
- Wholesale Banks Payments System :-
  - [http://www.databaseanswers.org/data\\_models/wholesale\\_banks\\_payment\\_systems/index.htm](http://www.databaseanswers.org/data_models/wholesale_banks_payment_systems/index.htm)

## Step 1 : Agree the User Scenario

In order to achieve our objective, we adopt a simple Scenario which is :-

“A Customer makes a Purchase, takes a Train journey, makes a Reservation for a Flight with an Airline and finally takes a Flight”.

We aim to identify the Generic and Subject Area Models that will help us to design the final Top-Level Data Model to match our User Scenario.

Finally, we will design a Data Warehouse and Data Mart.

We start by reviewing our Generic Data Models.

We define a Generic Data Model as a generic Solution to a commonly recurring requirement.

Wikipedia has an entry for Generic Data Models :-

- [http://en.wikipedia.org/wiki/Generic\\_data\\_model](http://en.wikipedia.org/wiki/Generic_data_model)

Our current list of 27 Generic Models appears on this page :-

- [http://www.databaseanswers.org/data\\_models/design\\_patterns\\_and\\_generic\\_models.htm](http://www.databaseanswers.org/data_models/design_patterns_and_generic_models.htm)

## Step 2 : Define a Generic Customers Model

### 2.1 : Setting the Scene

This cartoon shows Customers in Starbucks.

We can identify the 'Things of Interest' as being :-

- Customers
- A Starbucks Product in the form of a cup of coffee
- A member of staff
- Payment

Therefore our starting-point is to identify a Customer-related Data Model that includes these four 'Things of Interest'.



## 2.2 : Review available Customer Models

This Model is one of my early ones and dates from 2001.

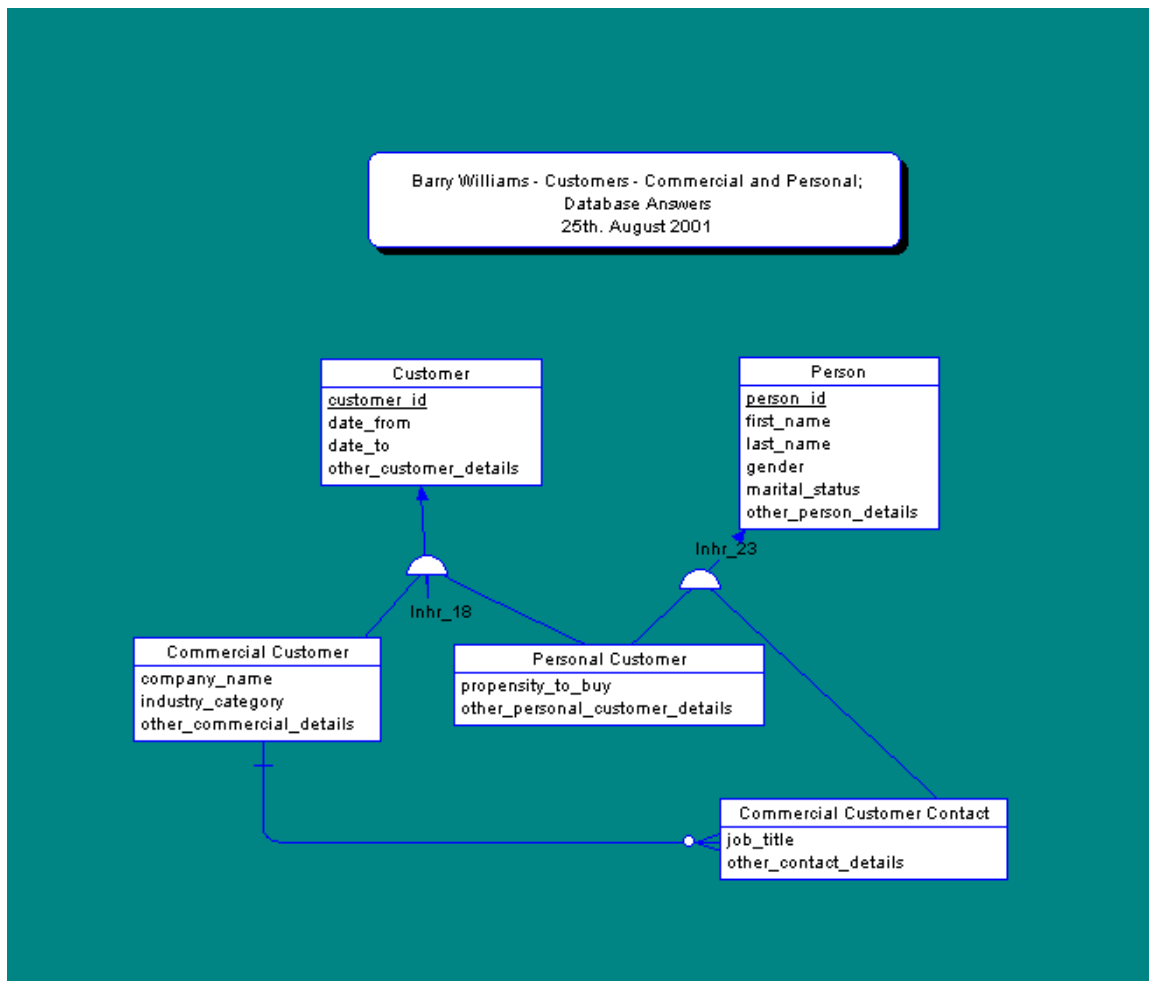
It is on this page :-

- [http://www.databaseanswers.org/data\\_models/customers\\_commercial\\_and\\_personal/index.htm](http://www.databaseanswers.org/data_models/customers_commercial_and_personal/index.htm)

It reflects my thinking and the style of Data Models I was using at that time.

It shows how Inheritance helps us to model both Commercial and Personal Customers. However it doesn't show any details of other things that we are interested in, like Payment Methods.

So we look for another Model which is a more suitable starting-point for a Customer who is going to make a Purchase.



### 2.3 : Customers and Addresses

This Model is a later one and dates from 2006.

It shows my evolution to a different Modelling Tool (Dezign) and style.

It is on this page :-

- [http://www.databaseanswers.org/data\\_models/customers\\_and\\_addresses/index.htm](http://www.databaseanswers.org/data_models/customers_and_addresses/index.htm)

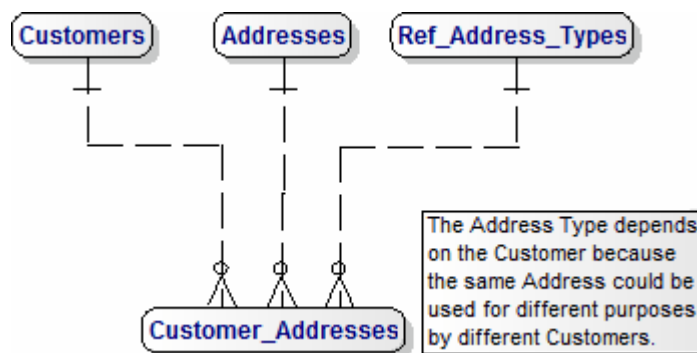
It shows a general solution for Customers and Addresses.

We will have to add more Customer-related items to the Customer Entity but it is fine for a starting-point for Address details.

It is useful to track a history of Address changes and also to be able to store multiple Addresses for the same Customer, such as Home, Work, Delivery and so on.

However it doesn't show any details of other things that we are interested in, like Purchases and Payment Methods.

So we look for another Model which is a more suitable starting-point for a Customer who is going to make a Purchase.



## 2.4 : Customers and Purchases

In this Step, a Customer makes a Purchase  
This Customers and Orders Model also dates from 2010.

It includes (Customer-related) Mailshots and (Product-related) Promotions.

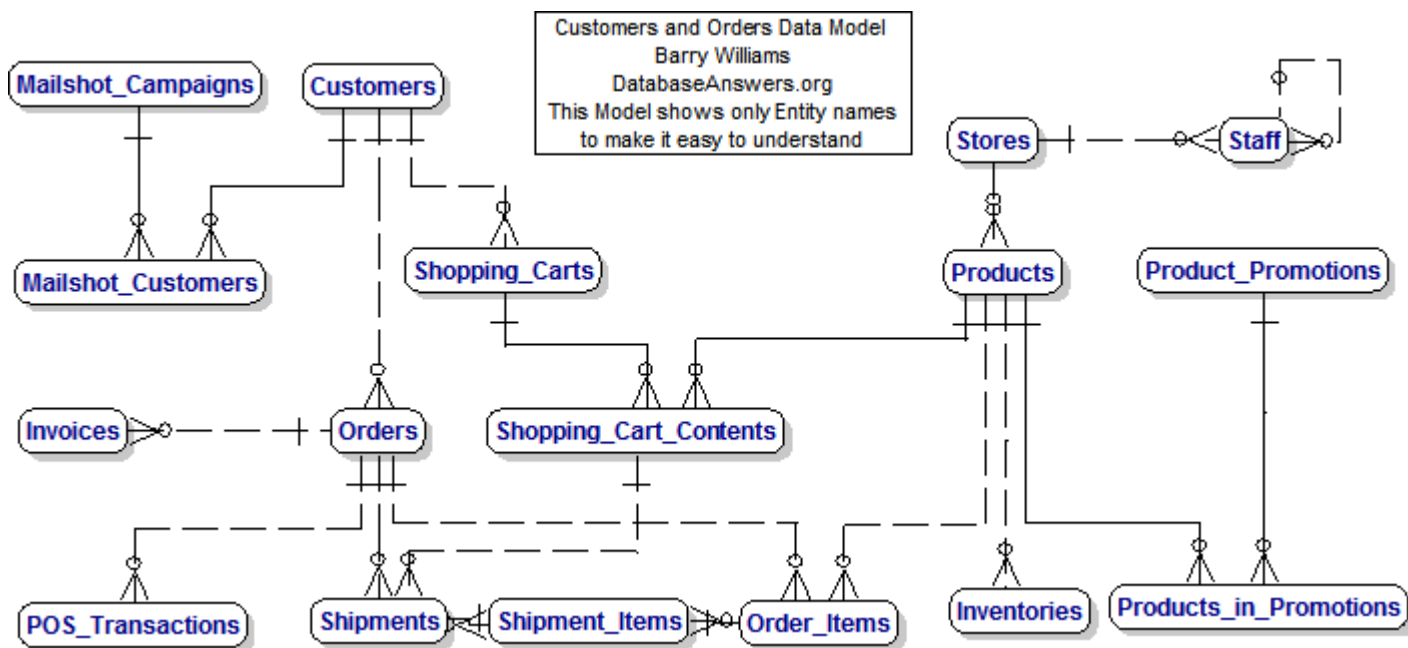
It is comprehensive so we decide to include it.

It is on this page :-

- [http://www.databaseanswers.org/data\\_models/retail\\_customers/customers\\_and\\_orders\\_model.htm](http://www.databaseanswers.org/data_models/retail_customers/customers_and_orders_model.htm)

However, it does not include Payments which we consider to be very important.

So we decide to look for a more suitable Data Model.





### 2.5 : A Customer makes a Purchase and a Payment

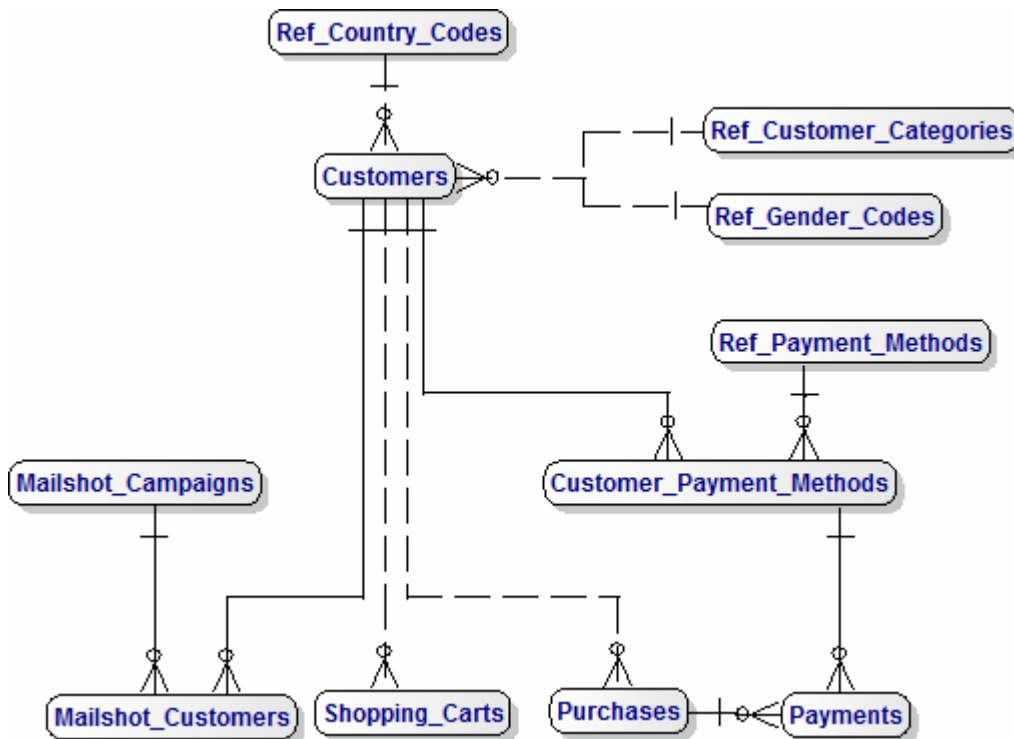
Then we find this Model which includes Payments in a way that is economical and natural.

We can see that it has added Payment Methods Reference Data, in other words, our standard Payment Methods, such as Credit and Debit Cards and Cash. The Customer Payment Methods Entity shows us that we can record specific Payments Methods for specific Customers. Then the relationship between the Customer Payment Methods Entity and Payments allows us to record which of the specific Payment Methods the Customer has chosen to use for this particular Purchase.

This Data Model is on this page :-

- [http://www.databaseanswers.org/data\\_models/retail\\_customers/customers\\_purchases\\_and\\_payments\\_model.htm](http://www.databaseanswers.org/data_models/retail_customers/customers_purchases_and_payments_model.htm)

We decide to include this Model because it includes Customers, Purchases and Payment Methods.

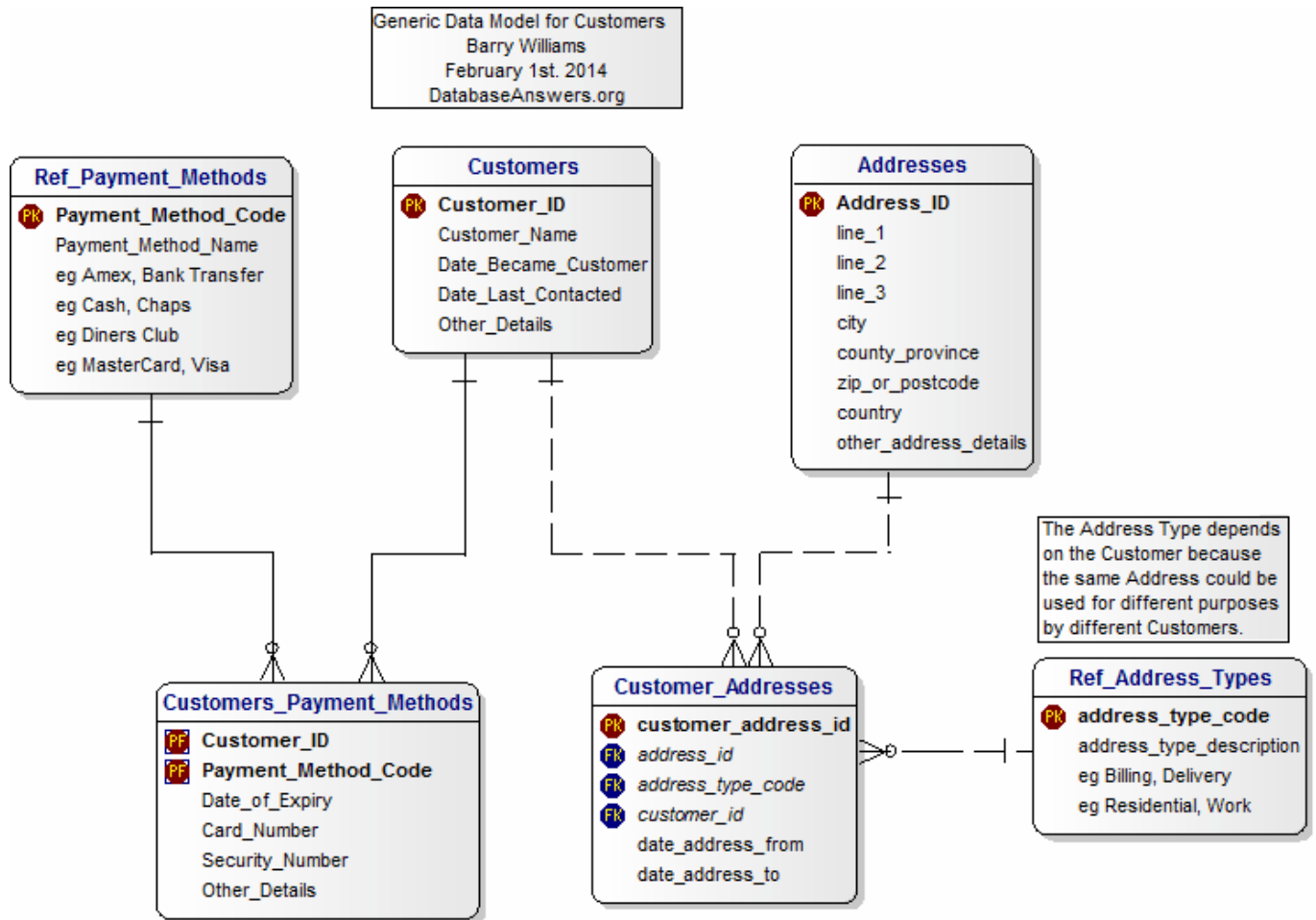


## 2.6 A Generic Customer Model

For our Generic Customer Model, we decide that we want something basic, that include Addresses and Payment Methods. We cannot find anything suitable, so we create one by combining two other Data Models.

We have a starting-point on this page :-

- [http://www.databaseanswers.org/data\\_models/customers\\_generic\\_model/index.htm](http://www.databaseanswers.org/data_models/customers_generic_model/index.htm)



### 2.7 Subject Area Model for Parties

We have almost thirty different Customer-related Models on our Web Site and when we think about the most suitable Subject Area for Customers we decide that the best one is Parties, Roles and Customers.

This is because it is the most general.

The Data Model is on this page :-

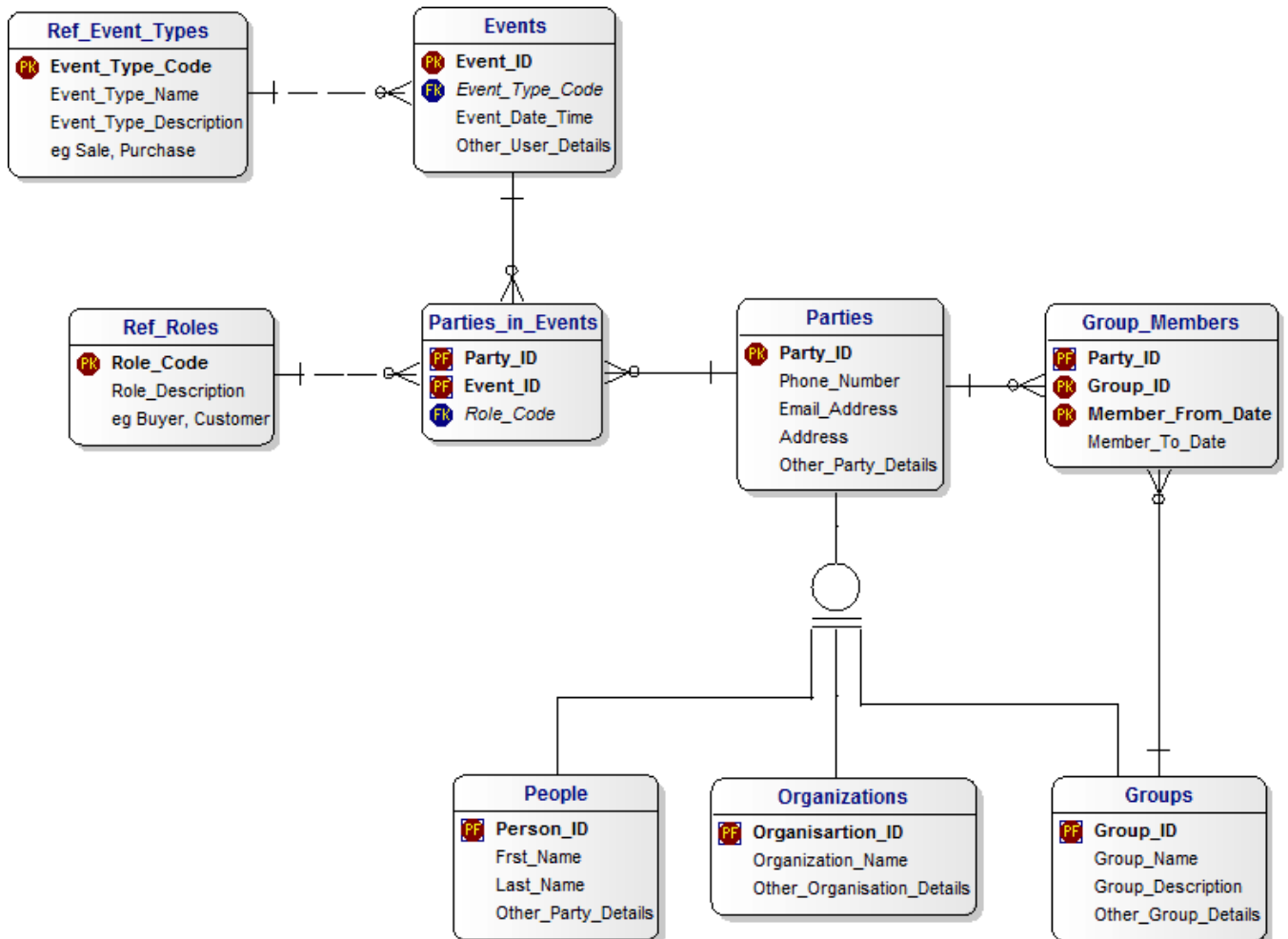
- [http://www.databaseanswers.org/data\\_models/parties\\_roles\\_and\\_customers/index.htm](http://www.databaseanswers.org/data_models/parties_roles_and_customers/index.htm)

In English, we would say that a Party is a Person, Group or Organisation.

A Party can then play different Roles in different Events.

For example, when we take a ride in a taxi, the taxi-driver is a Supplier, but when he (or she) buys petrol (or gas if you are in the States) then the taxi-driver is a Customer.

This situation is shown in this Model :-



### Step 3 : Define a Generic Customer Services Models

This is a very powerful Generic Model because it applies to wide range of situations.

It usually involves a Customer being provided with a Service by a Supplier.

This may involve a Payment which we discuss in the next Chapter,

#### 3.1 : Setting the Scene

We can identify the 'Things of Interest' for a Payment as being :-

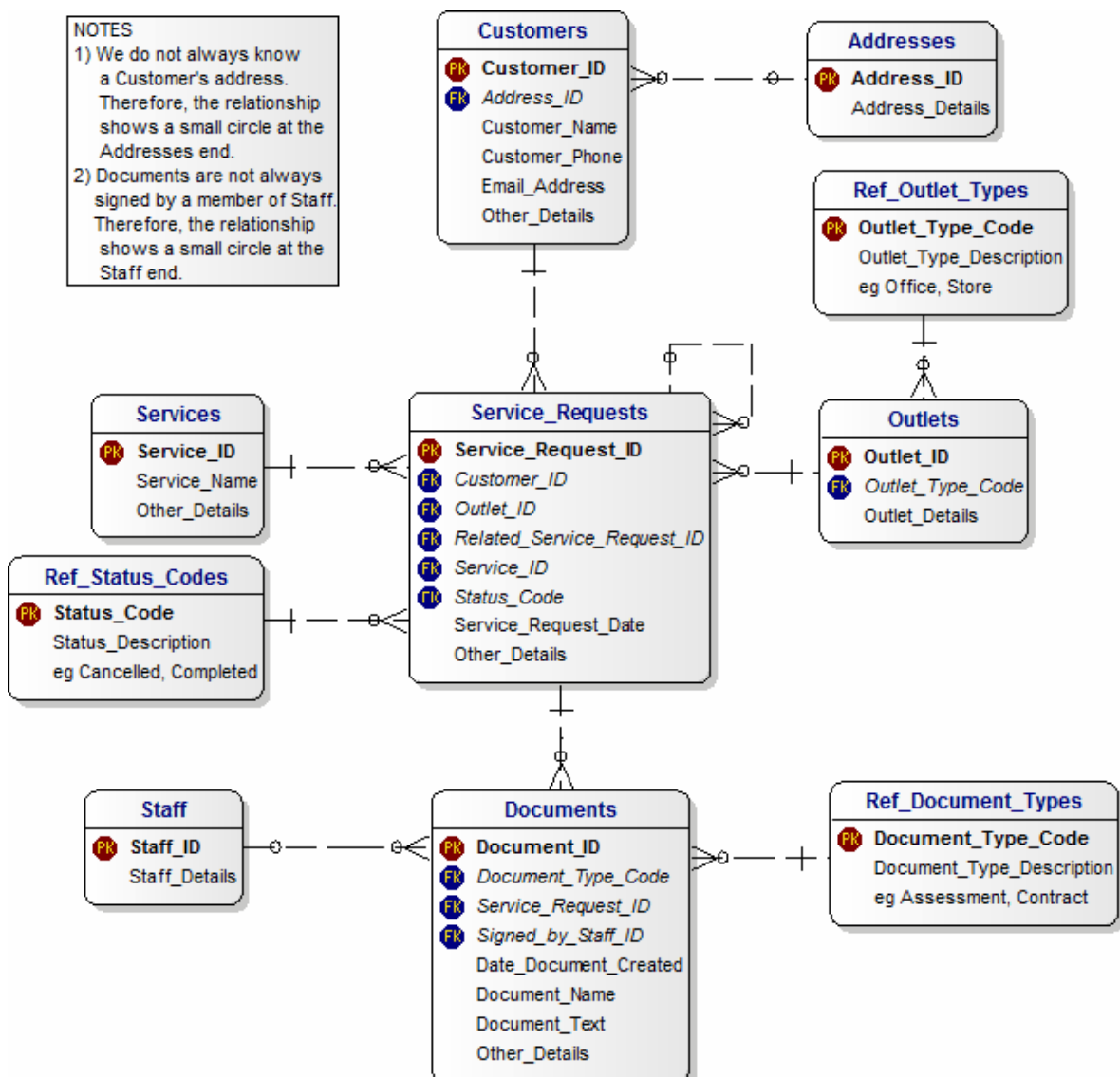
- The Person or organisation making the Payment
- The Person or organisation receiving the Payment
- The Payment Method – such as a Credit or Debit Card

#### 3.2 The Generic Model

Our Generic Customer Services Model on this page :-

- [http://www.databaseanswers.org/data\\_models/customers\\_and\\_services\\_generic/index.htm](http://www.databaseanswers.org/data_models/customers_and_services_generic/index.htm)

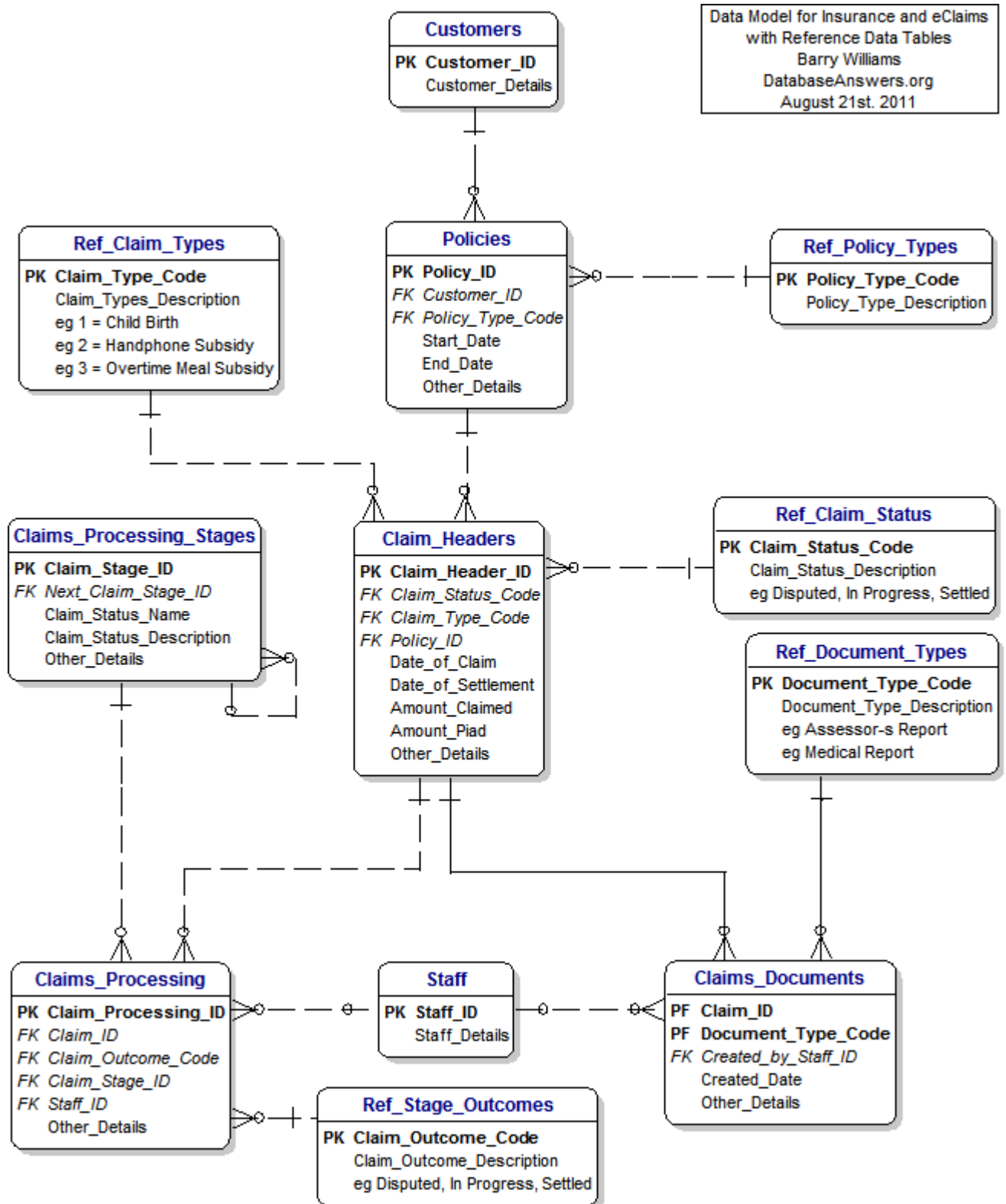
The Model looks like this :-



### 3.3 Subject Area Model for Insurance

This builds on the Generic Customer-Services Model :-

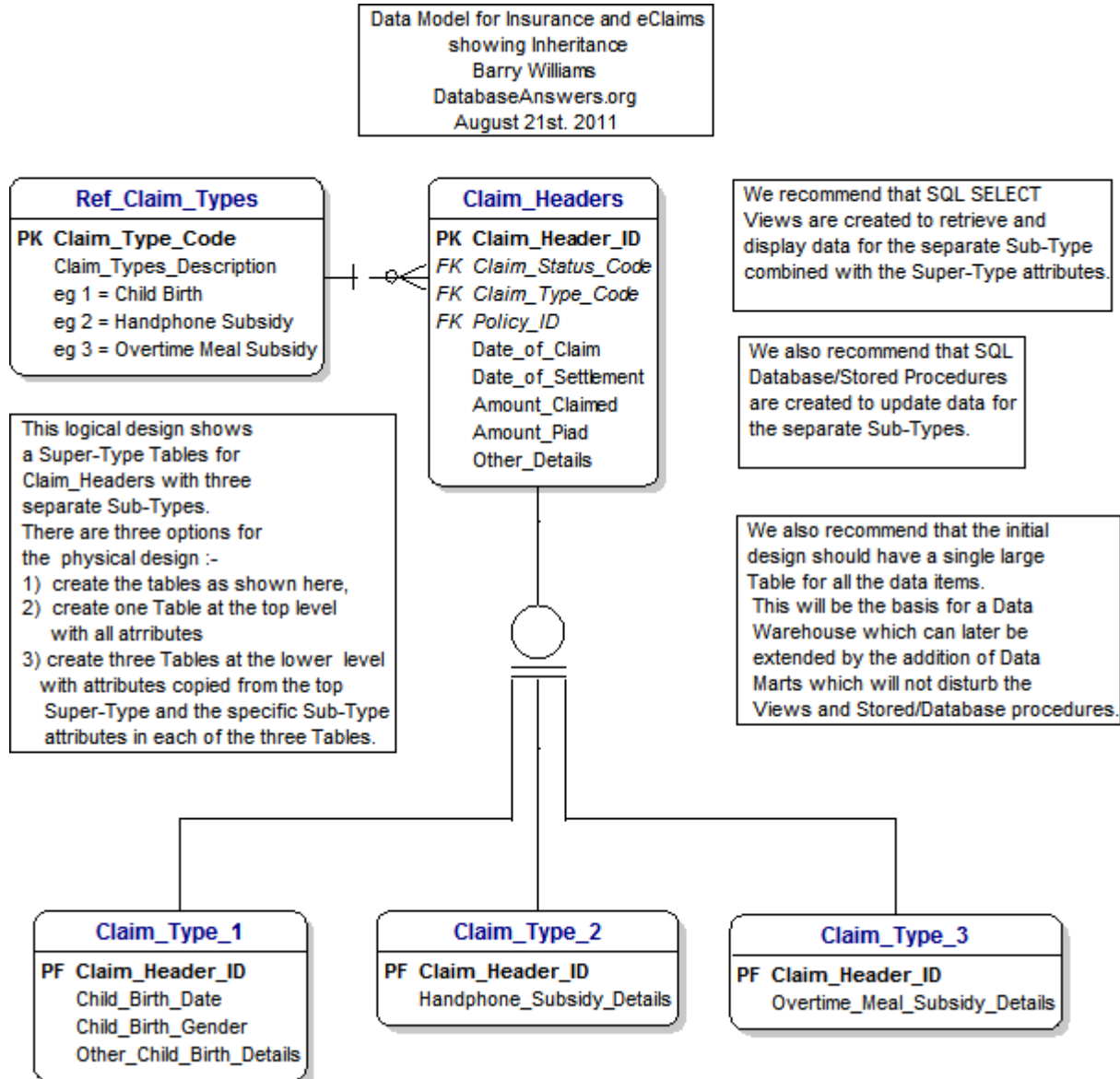
- [http://www.databaseanswers.org/data\\_models/insurance\\_and\\_eclaims/index.htm](http://www.databaseanswers.org/data_models/insurance_and_eclaims/index.htm)



### 3.4 Claims showing Inheritance for Claim Types

This is a variation on the versions above which emphasises the possibility of modelling Claims using Inheritance.

It is shown on the same page as the Model above.



### Step 4 : Define a Generic Payments Models

The usual pattern of a Pattern is that it involves Payment from one Party to another Party using one of an appropriate number of available Payment Methods.

For Retail Payments, this will include Credit or Debit Cards, or of course, Cash.

#### 4.1 : Setting the Scene

We can identify the 'Things of Interest' for a Payment as being :-

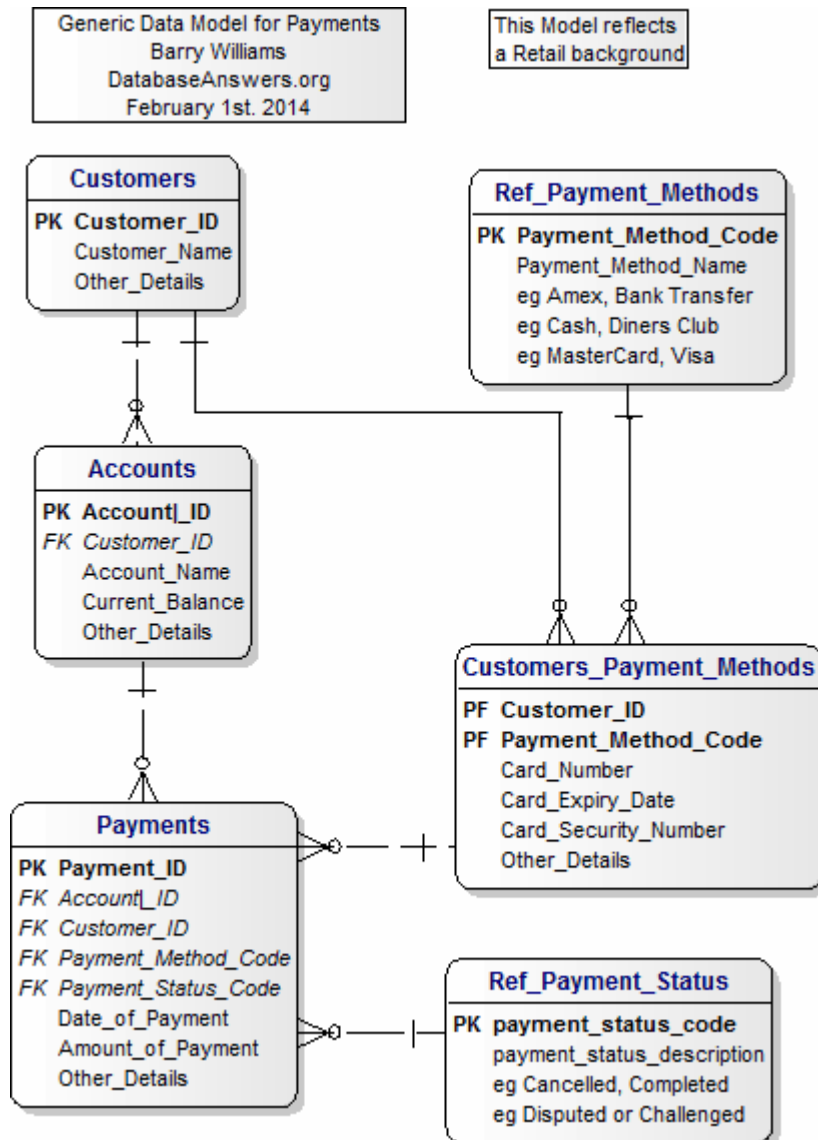
- The Person or organisation making the Payment
- The Person or organisation receiving the Payment
- The Payment Method – such as a Credit or Debit Card

#### 4.2 The Generic Model

Our Generic Payments Model on this page :-

- [http://www.databaseanswers.org/data\\_models/payments\\_generic\\_model/index.htm](http://www.databaseanswers.org/data_models/payments_generic_model/index.htm)

The Model reflects our orientation to Retail Customers and looks like this :-



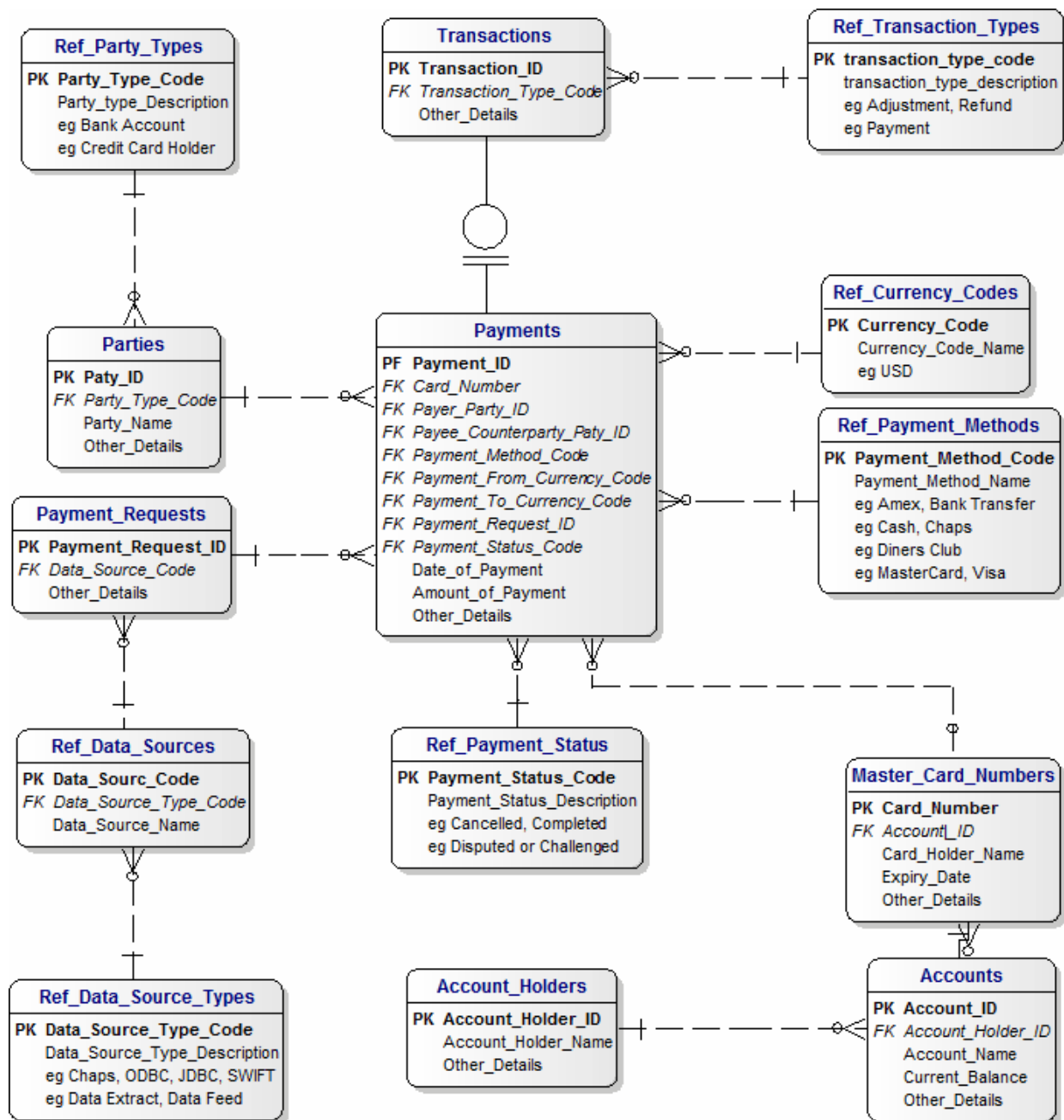
You can see that it shows a Customer who makes Payments using a Credit or Debit Card which is associated with a Bank Account.

It is a simple Model but it provides a good starting-point for the design of Subject Area Payments Models.

### 4.3 Subject Area Model for Payments

This builds on the Generic Model and is appropriate to a wide range of situations. It provides an overview of the Entity-Relationship Diagram (ERD) on this page :-

- [http://www.databasanswers.org/data\\_models/payments/index.htm](http://www.databasanswers.org/data_models/payments/index.htm)

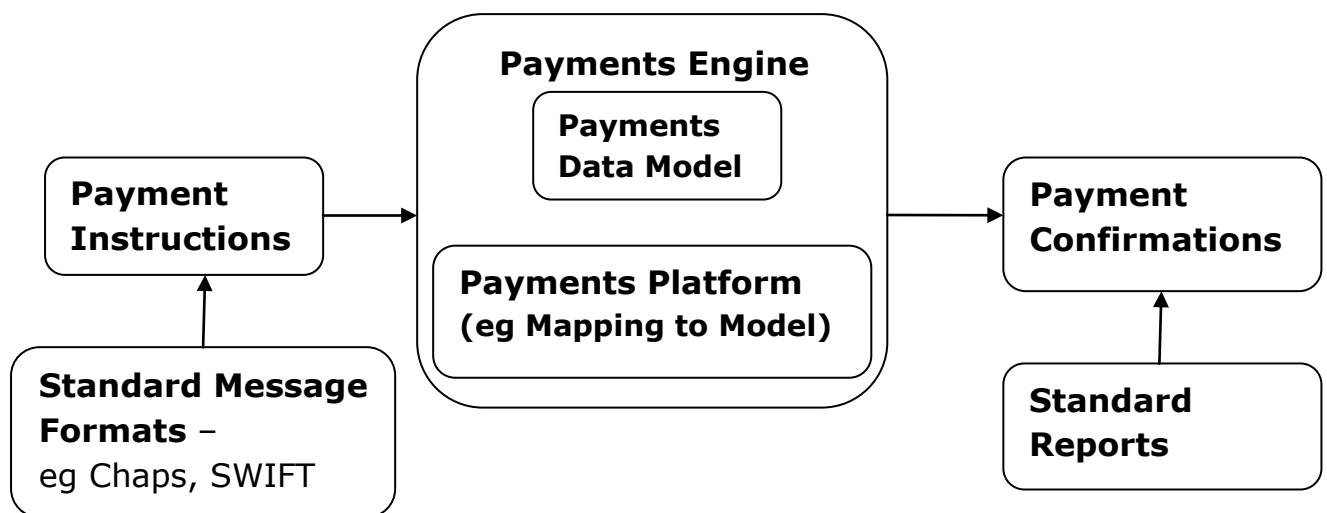




#### 4.4 Data Architecture

This Data Architecture shows that Payment Instructions are processed by a Payments Engine, which includes the Payments Platform and Payments Data Model. The Platform can process Messages in standard Formats and load the data in accordance with the design of the Data Model.

After Payments are processed, Payment Conformations are generated and standard Reports can be generated.



#### 4.5 Wholesale Banks Payments Subject Area Model

Wholesale Banking is defined by Wikipedia on this page :-

- [http://en.wikipedia.org/wiki/Wholesale\\_banking](http://en.wikipedia.org/wiki/Wholesale_banking)

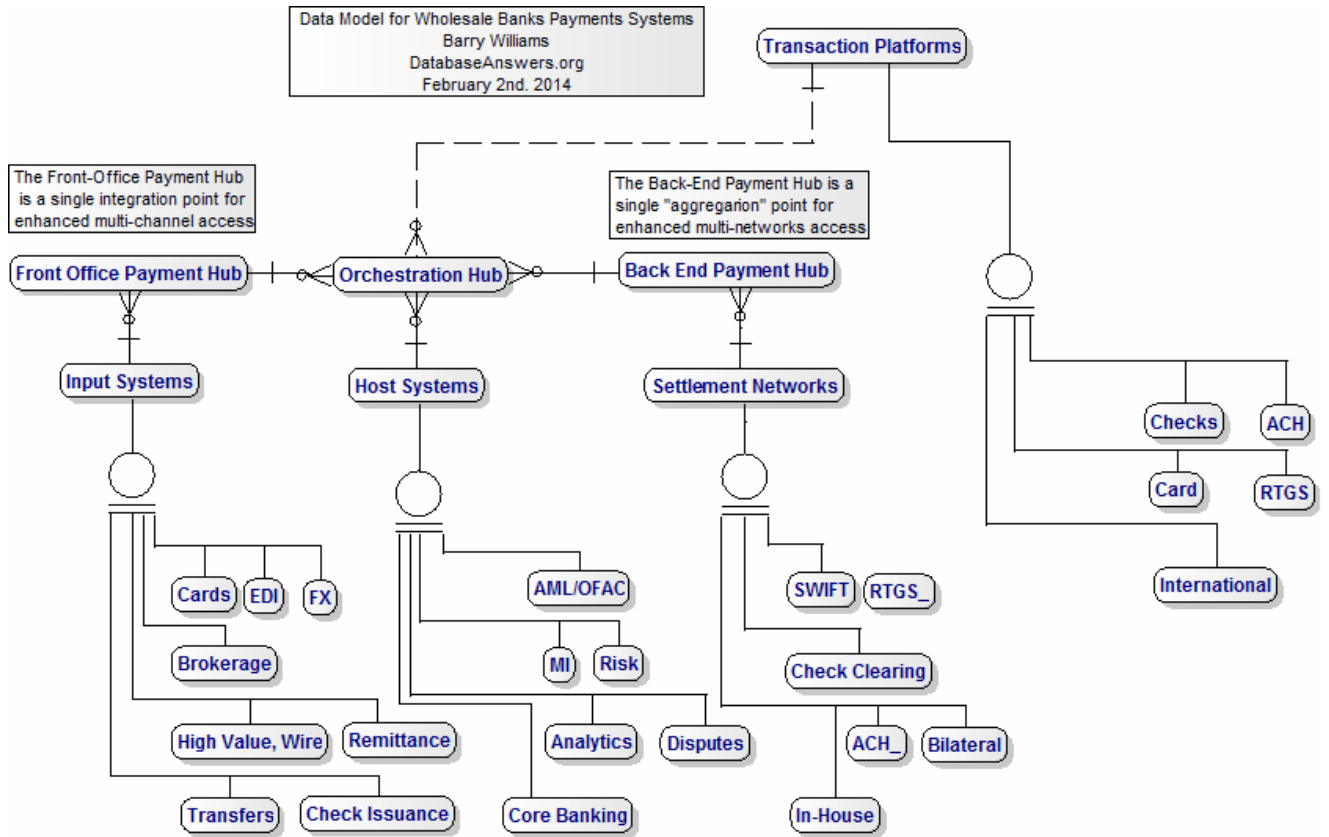
It says :-

“**Wholesale banking** is the provision of services by banks to organisations such as [Mortgage Brokers](#), large corporate clients, mid-sized companies, [real estate developers](#) and [investors](#), international trade finance businesses, [institutional customers](#) (such as [pension funds](#) and government entities/agencies), and services offered to other banks or other financial institutions.

Modern wholesale banks are engaged in: finance wholesaling, underwriting, market making, consultancy, mergers and acquisitions, fund management”.

We have a Payments Model for Wholesale Banks on this page :-

- [http://www.databaseanswers.org/data\\_models/wholesale\\_banks\\_payment\\_systems/index.htm](http://www.databaseanswers.org/data_models/wholesale_banks_payment_systems/index.htm)

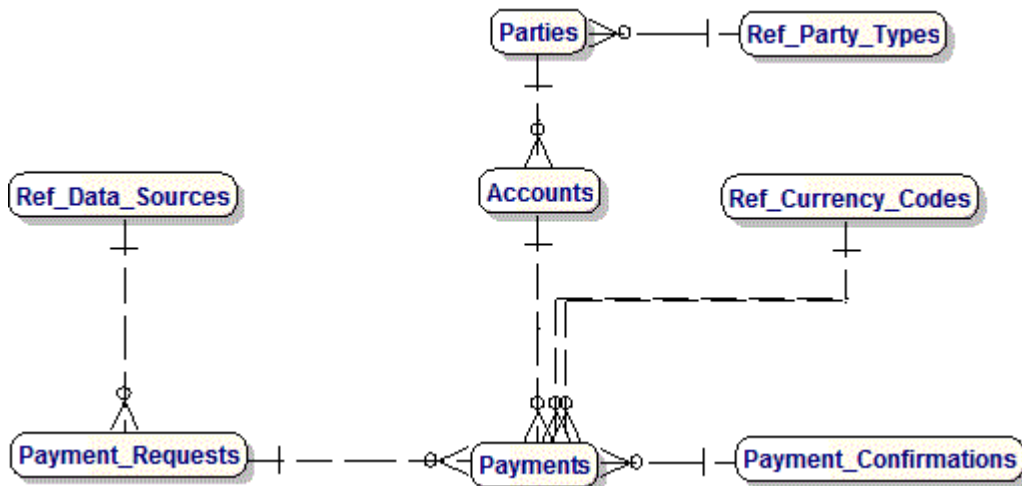


### 4.6 Wholesale Banks Payments Canonical Model

This Canonical Model is on this page :-

- [http://www.databaseanswers.org/data\\_models/wholesale\\_banks\\_payment\\_systems/canonical\\_data\\_model.htm](http://www.databaseanswers.org/data_models/wholesale_banks_payment_systems/canonical_data_model.htm)

There are two relationships between the Currency Codes and the Payments Entities. This is because a Payment might involve two Currencies.



## Step 5 : Define a Generic Reservations Model

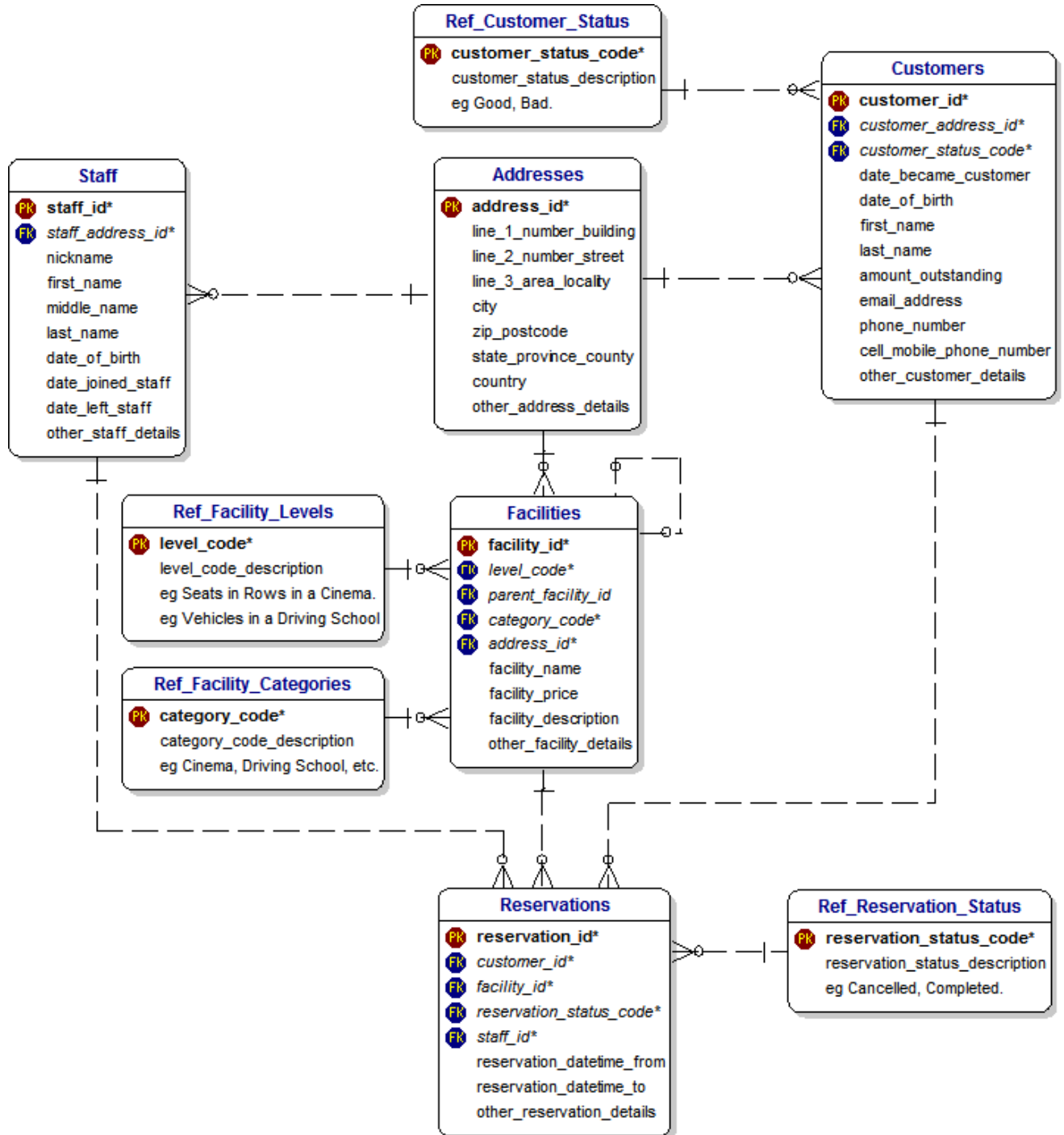
### 5.1 The Generic Reservations Model

In this Step of the User Scenario, our Customer makes a Reservation for a Flight

Our Generic Model is on this page :-

- [http://www.databaseanswers.org/data\\_models/generic\\_reservations/index.htm](http://www.databaseanswers.org/data_models/generic_reservations/index.htm)

This Model provides for Reservations for a Facility at a specific Date and Time.



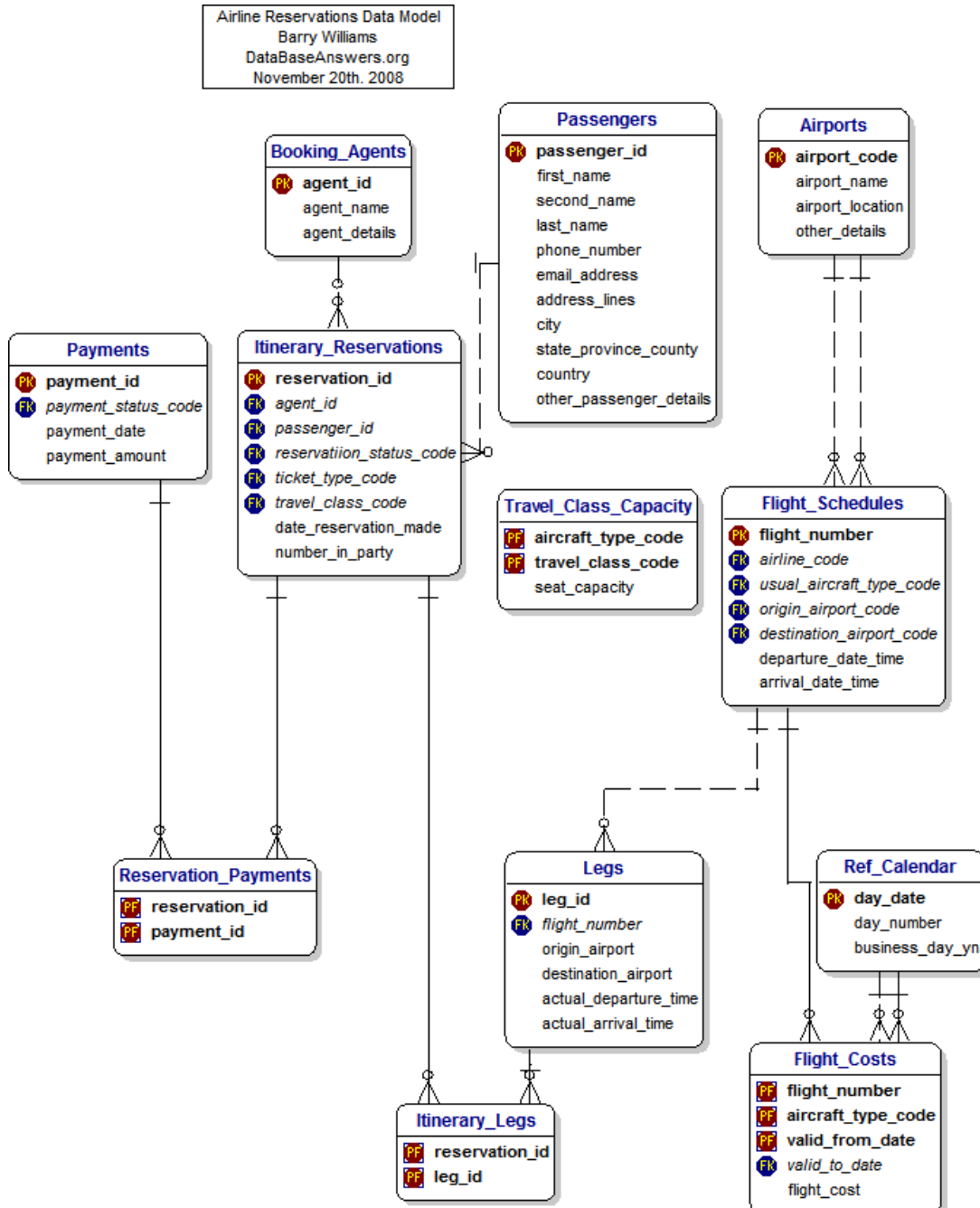
## 5.2 Subject Area Airline Reservations Model

This Airline Reservations Model dates from 2008 and is shown on this page :-

- [http://www.databasanswers.org/data\\_models/airline\\_reservations/index.htm](http://www.databasanswers.org/data_models/airline_reservations/index.htm)

We have added some things that are common with Air Travel which are not common with Generic Reservations described above.

These include Bookings Agents, Legs in an Itinerary, Costs and Payments.



## Step 6 : Define a Generic Travel Model

### 6.1 Public Transport Model

This Model dates from 2009, and is included in this page :-

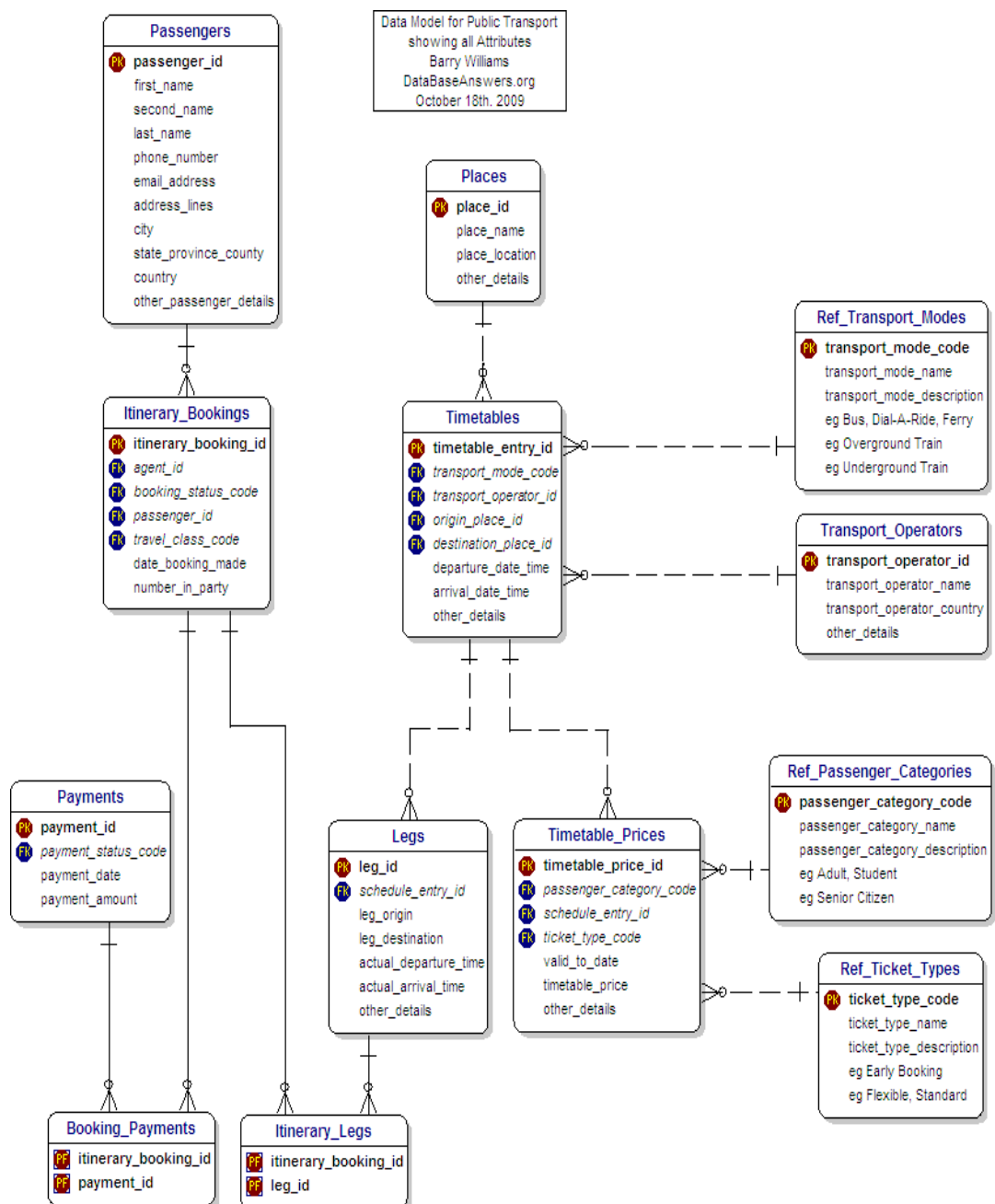
- [http://www.databasanswers.org/data\\_models/public\\_transport/index.htm](http://www.databasanswers.org/data_models/public_transport/index.htm)

This Model allows us to define an Itinerary consisting of multiple journeys by Train, Bus and any other type of Public Transport.

However, we like this because it means that this economical Model is more powerful than we need in a way for our current requirement that will certainly be useful in a Subject Area Model.

So we decide to include this Model.

It includes Payments but not with the detail in our previous Customers, Purchases and Payments Model.



## 6.2 The Customer travels by Train

This shows a Passenger getting on the Heathrow Express Train that travels between Heathrow Airport and London :-



The 'Things of Interest' for us are :-

- Passenger (or Customer) with Baggage
- Train (a Service)

She probably has already purchased a ticket, but she could possibly do that on the train.

She has a Bag, but will not have a reservation for a specific Seat.

The Queen of England used to have a private train, but for the rest of us, Train travel is always Public Transport.

Therefore, we can meet this requirement using the Generic Model of Public Transport that we discussed above.

## 6.3 Subject Area Model for Air Transport

In this Scenario, the Customer travels by plane.

This Model is shown on this page :-

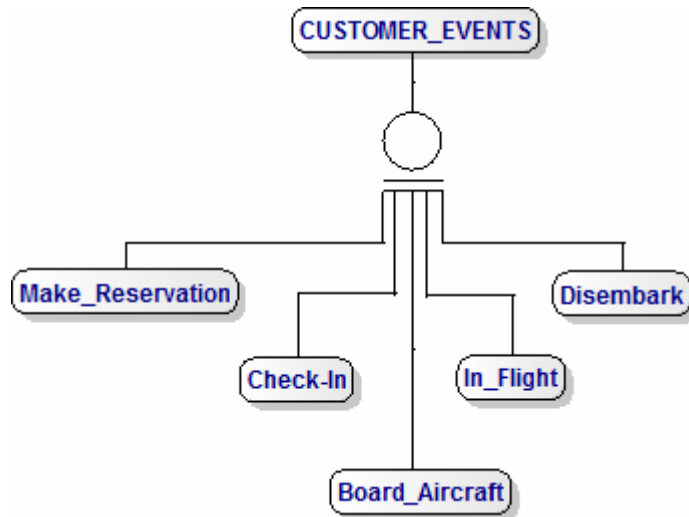
- [http://www.databaseanswers.org/data\\_models/airline\\_operations/index.htm](http://www.databaseanswers.org/data_models/airline_operations/index.htm)

It shows the Events that occur when a Passenger takes a Flight :-

1. Make a Reservation – we have discussed this earlier.
2. Check-In for the Flight, including Baggage.
3. Board the Aircraft
4. Travel in Flight
5. Disembark on arrival at the destination.

This Model is very small, economical and powerful.

It does not show all the details of the data involved but these can easily be derived and we discuss them when we come to look at the design of the Data Mart.



#### 6.4 Subject Area Model for Trains and Boats and Planes

At this point, we think about integrating the Public Transport Model with the Airline Operations Model.

For example, we consider replacing the two Models by the one **for Trains and Boats and Planes**

This Model dates from 2013 and is on this page :-

- [http://www.databaseanswers.org/data\\_models/trains\\_and\\_boats\\_and\\_planes/index.htm](http://www.databaseanswers.org/data_models/trains_and_boats_and_planes/index.htm)

When we look at this Model, we realise that we can use a general word 'Journeys' for Train Journey and a Flight with an Airline.

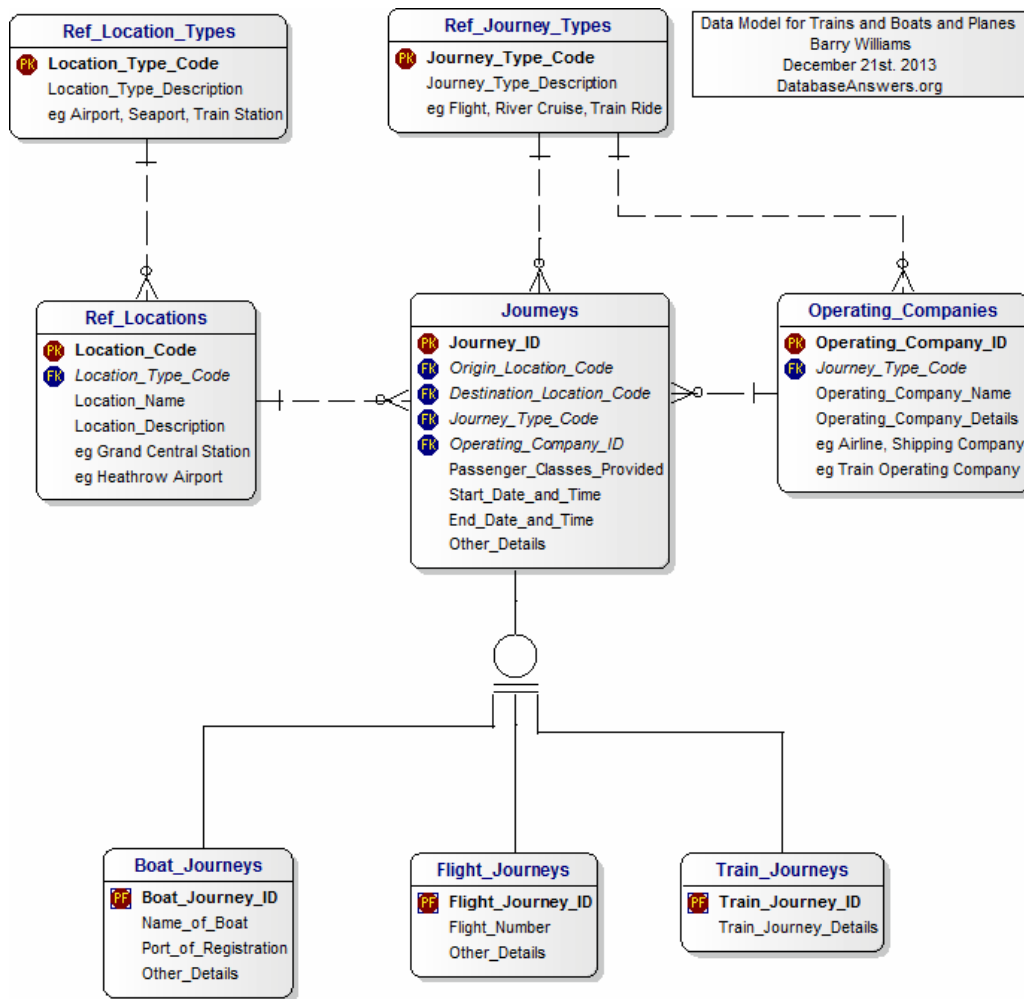
For a Flight, the Airline is an Operating Company and the Locations are Airports.

We use Inheritance to define a Journeys 'Super-Type' and Boat, Flight and Train Journeys as 'Sub-Types'.

Using this approach we put the common data for all Journeys into the Journeys Entity and the separate data that applies to specific types of Journeys into the appropriate Entities that we have called Boat, Flight and Train Journeys.

For example, all Journeys have an Origin and a Destination.

However we realise that we have not included Reservations so this Model is useful but limited..



However, we come to the conclusion that making a Flight Reservation and travelling by train do not have much in common.

What we have in this 'Trains and Boats and Planes' Model is simply the fact that all Journeys have something in common. Therefore, we decide to keep the Public Transport Model for the Train journey and compare the Airline Reservations Model with the "Generic Reservations" Model.

### Step 7 : Review progress

At this point, we review our progress and realise that we have a complex situation involving many Data Models that we want to integrate into one Model.

We have Customers, Passengers, Products, Purchases, Payments, Airline Flight Reservations and Travel by Train and Plane.

However, I realised that travelling by Train and making a Reservation for a Flight do not have much in common except that they are both related to Transport.



Therefore the natural solution involves two separate Data Models.

Our overall objective in this document is to talk through the process of design an integrated Data Model that has as small number of Entities as possible for the User Scenario that we defined at the beginning.

We also want to review and validate our candidates for Generic Data Models and create new ones as appropriate.

For example, we have reviewed our Generic Reservations Model and decided it needs work to incorporate Flight Reservations because it started out as a simple Model to book Theatre seats.

### Step 8 : Adopt the Canonical Data Model

At this point we decide to use our Canonical Data Model because its Event-Oriented Approach is very economical and very powerful.

It is on this page :-

- [http://www.databaseanswers.org/data\\_models/canonical\\_data\\_models/index.htm](http://www.databaseanswers.org/data_models/canonical_data_models/index.htm)

The type of Documents include Sales Receipts for our Purchases and Train Tickets. Organisations include Airlines, Train Companies

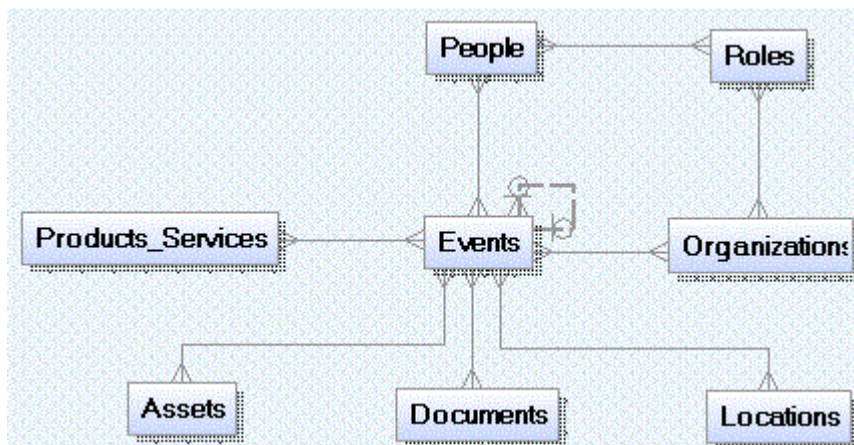
Products include our retail Purchases and Services include Train Journeys and Airline Flights.

The People are, of course, individuals and the Roles they can play include Customers, Staff of Organisations, such as Shops, Airlines or Train Companies.

The Types of Events can include making a Purchase, buying a Train Ticket, and making a Flight Reservation and so on.

From this brief description, we can see that the Event approach provides us with a Data Model that is both powerful and economical.

In other words, with a small number of Entities we can cover a wide range of situations.



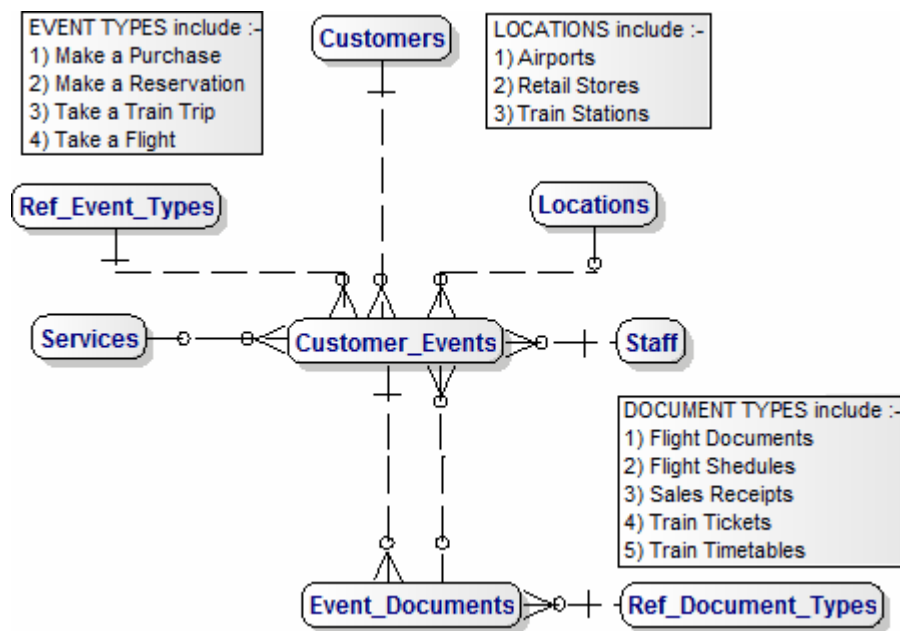
### Step 9 : Define Top-Level Model with Customers

Below, we show the Top-Level Data Model with Customers.  
 In the next Model, we replace Customers with Parties.

Logically, these are equivalent but the Parties approach is more general and favoured by professional Data Modellers.

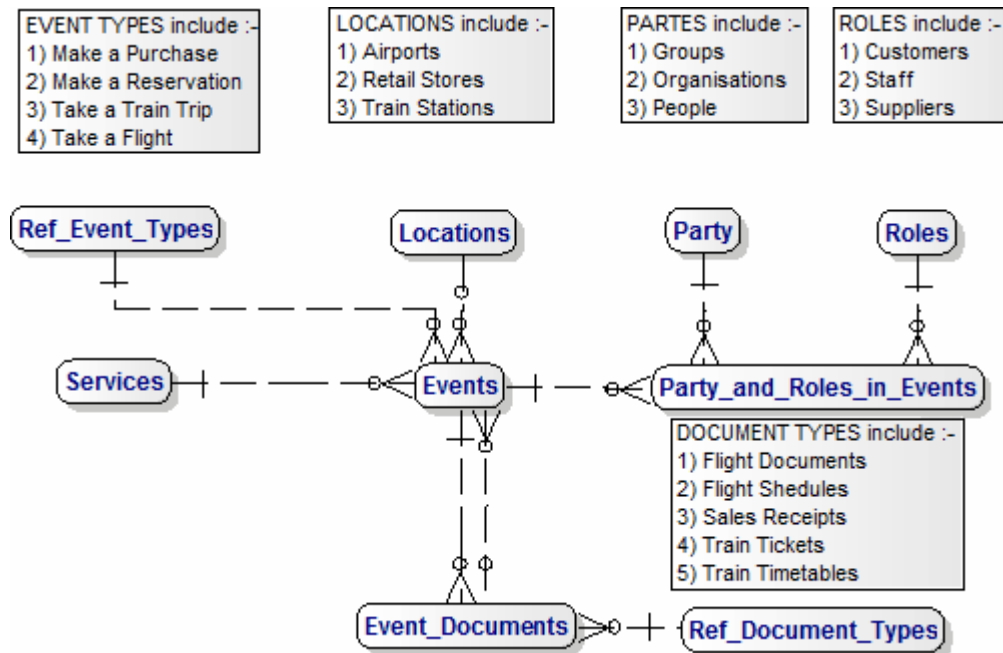
You can choose the one which is more appropriate to your situation.  
 The Parties approach helps to avoid discrepancies in analysis and BI where the same organisation might be counted twice.

When we use Events, the details of the individual Events are concealed.



## Step 10 : Define Top-Level Model with Parties

In this Model, we show Parties instead of Customers as an Entity.



## Step 11 : Review the List of Generic Data Models

We have seen the benefits of Generic Data Models

At this final stage, we have identified the following Subject Areas for Generic Models :-

- Parties, Roles and Customers
  - [http://www.databaseanswers.org/data\\_models/parties\\_roles\\_and\\_customers/index.htm](http://www.databaseanswers.org/data_models/parties_roles_and_customers/index.htm)

## Step 12 : Review the List of Subject Area Data Models

We have seen the benefits of Generic Data Models

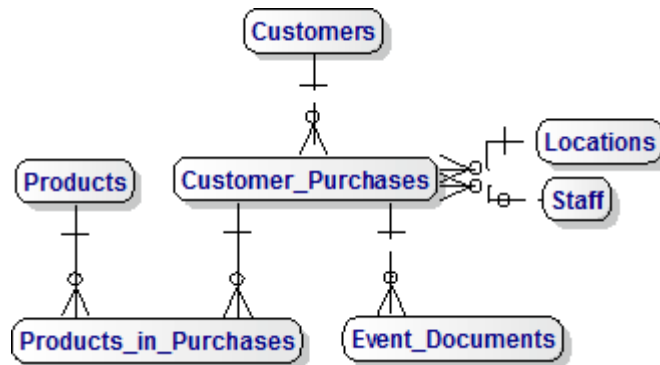
At this final stage, we have identified the following Subject Area Models :-

- Customers, Purchases and Payment
  - [http://www.databaseanswers.org/data\\_models/retail\\_customers/customers\\_purchases\\_and\\_payments\\_model.htm](http://www.databaseanswers.org/data_models/retail_customers/customers_purchases_and_payments_model.htm)
- Events (Canonical Data Model)
  - [http://www.databaseanswers.org/data\\_models/canonical\\_data\\_models/index.htm](http://www.databaseanswers.org/data_models/canonical_data_models/index.htm)
- Public Transport
  - [http://www.databaseanswers.org/data\\_models/public\\_transport/index.htm](http://www.databaseanswers.org/data_models/public_transport/index.htm)
- Reservations (Generic) -
  - [http://www.databaseanswers.org/data\\_models/generic\\_reservations/index.htm](http://www.databaseanswers.org/data_models/generic_reservations/index.htm)

## Step 13 : Design the Data Warehouse

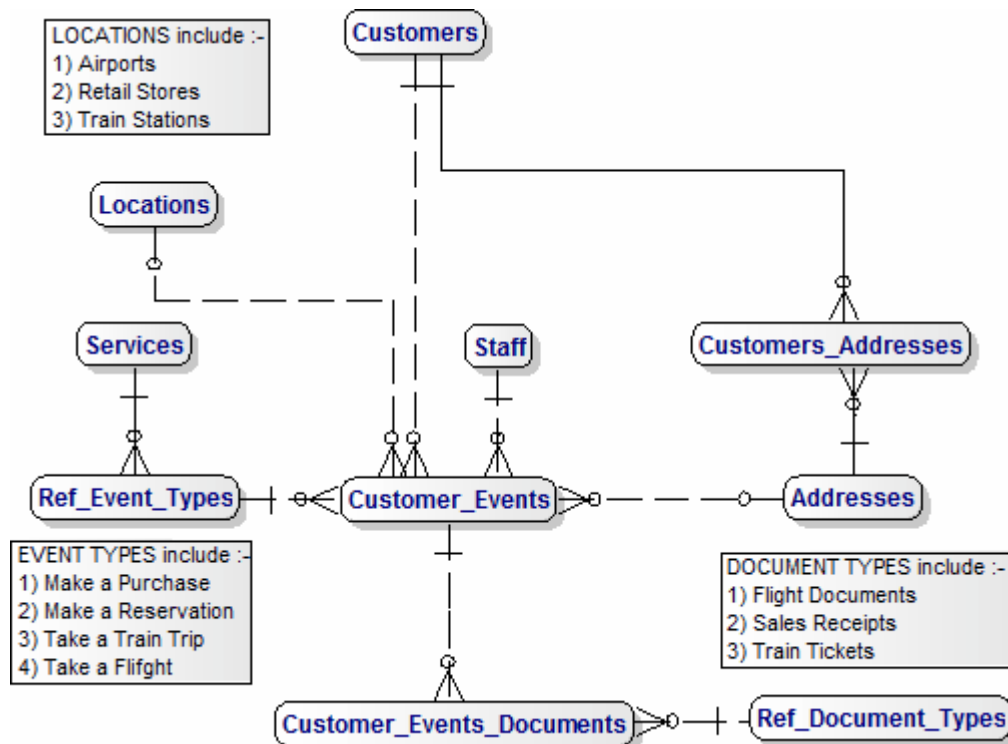
### 13.1 A Generic Data Warehouse

Our Data Warehouse is a Generic Third-Normal Form design.



### 13.2 Our Specific Data Warehouse

Our Data Warehouse has a Third-Normal Form design and is the same as our final Top-Level Model with Customers.

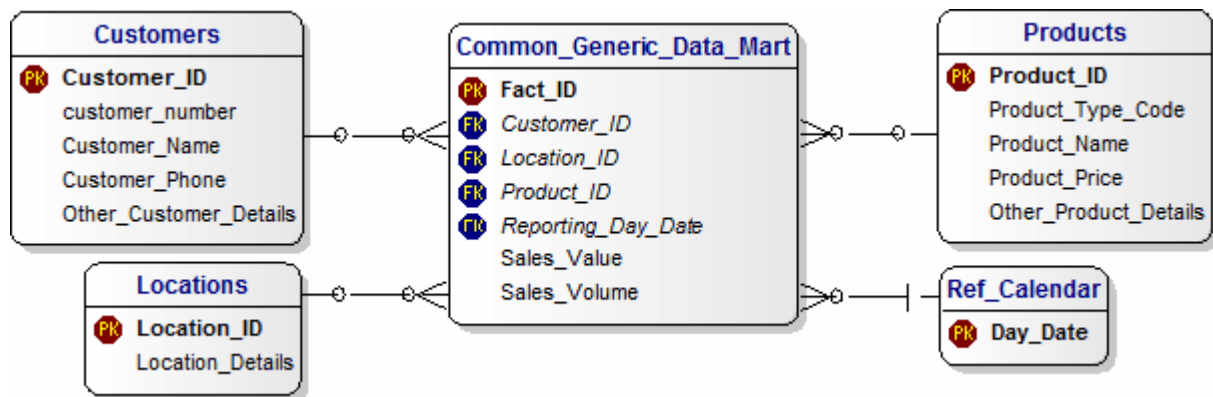


## Step 14 : Design the Data Marts

### 14.1 A Generic Data Mart

This simple design is one of the most commonly quoted Generic Data Mart. It has only four Dimensions :-

- Customers
- Locations
- Products
- Time-Periods



### 14.2 Translating Generic Terms

We can use this common design to create a range of Generic Reports and then translate the Common Terms into specific ones for our area of interest.

In this case we are interested in Travel and our translation would look like this :-

GENERIC TERM	TRAVEL-SPECIFIC TERM
Customers	Passengers
Locations	Airports and Train Stations
Products	Flights and Train Journeys
Time Periods	Time Periods

### 14.3 Specific Data Marts

#### 14.3.1 Requirements for our Specific Data Marts

When we look at the full list of possible data items that could be Dimensions in our Data Mart we come up with the following list :-

- Airlines
- Customers
- Documents
- Events
- Journeys
- Locations
- Mailshots
- Payment Methods
- Products
- Purchases
- Services
- Staff

- Stores
- Suppliers
- Time-Periods
- Travel Classes

Next we consider the fact that our Data Mart is going to be used to provide data for Business Intelligence, Performance Reports and a range of Analytics. Then we look at the list of possible data items to decide if we are they are going to be included in any analysis that we think is a realistic requirement.

We decide to leave out the following items:-

- Customers – because we are not interested in specific individuals
- Documents – because they are always a by-product of something like a Ticket for a Train Journey that we are going to include anyway.
- Staff – because we are not interested in that level of analysis

Because we have areas of data which are not related we decide to design two Data Marts which share Conformed Dimensions.

We define a simple example here with just a few Dimensions to explain the basics.

### 14.3.2 Phase 1 - Two Data Marts

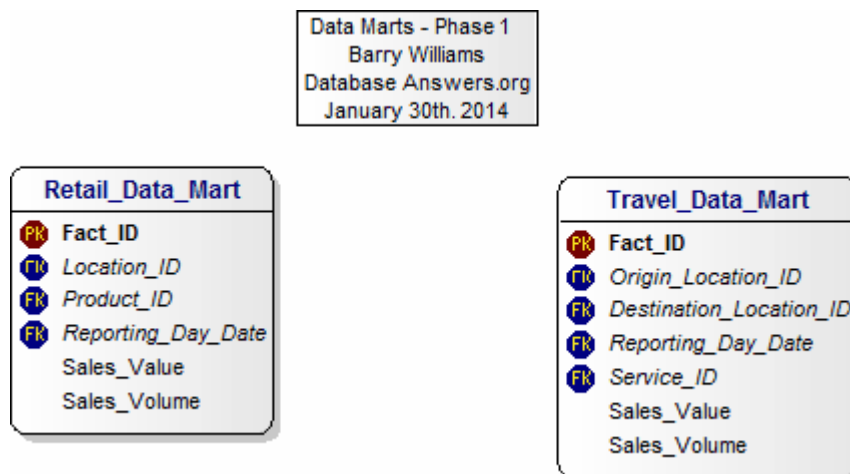
We start with two specific Data Marts – one for Retail Sales and the other for Travel.

For each Mart, the Foreign Keys (marked with 'FK') are the Dimensions, and the Sales Value and Volumes are the Facts.

These are the basic Facts and we can expect to have other derived Facts, such as percentages, changes, Key Performance Indicators (KPIs) and so on.

For the Retail Data Mart, we have one Location, which is the shop or outlet where the Retail Purchase was made.

For the Travel Data Mart, we have two Locations, which are the Origin and Destination of the Travel.

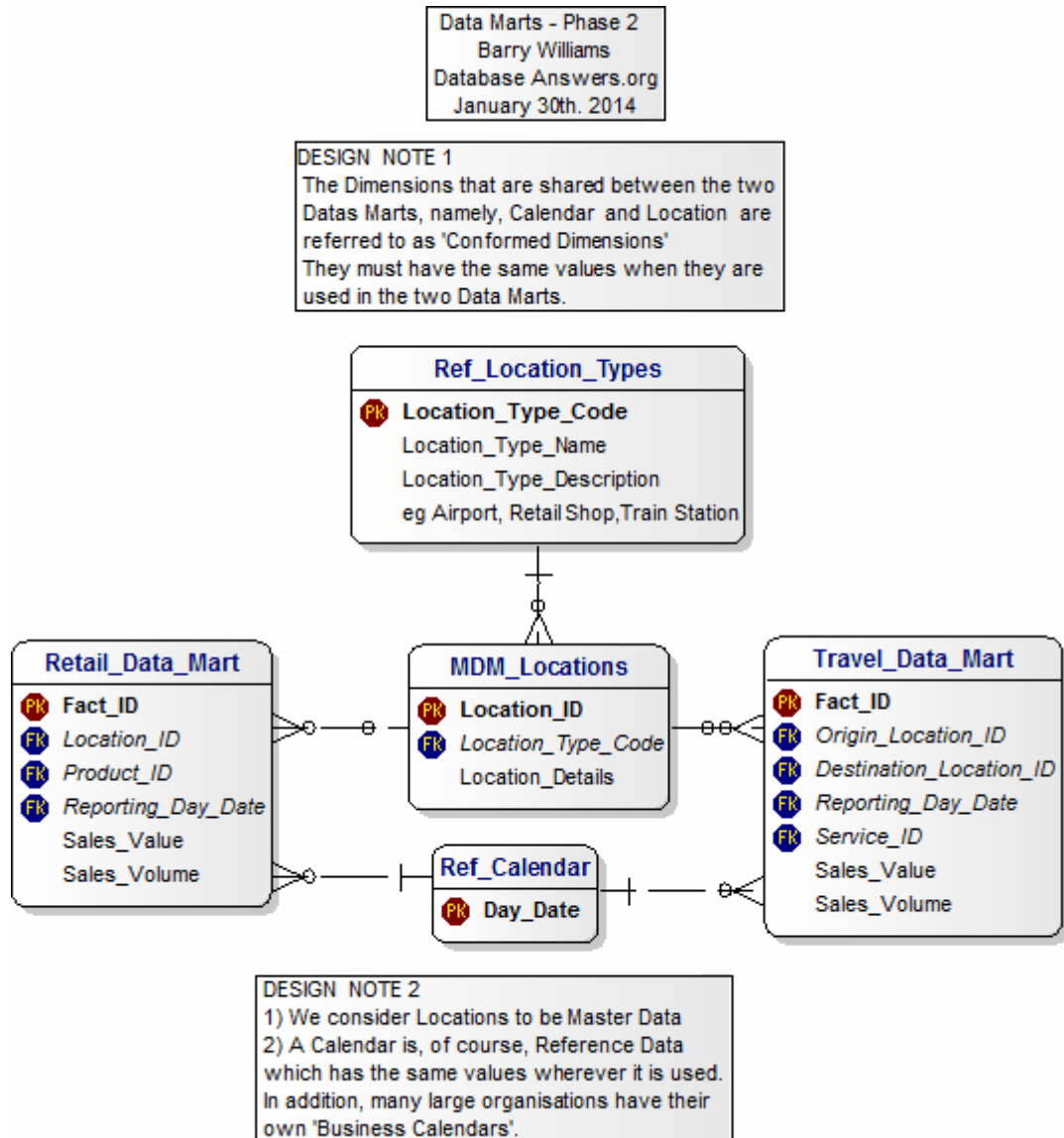


### 14.3.3 Phase 2 – Add Conformed Dimensions

The two Marts will share some Dimensions, in this Case, Location and Dates.

This means, of course, that the Marts will use the same values for Locations and Dates.

In this case, we refer to these as Conformed Dimensions



### 14.3.4 Phase 3 – Produce the complete Data Model

Finally, we add details of the Products and Services.

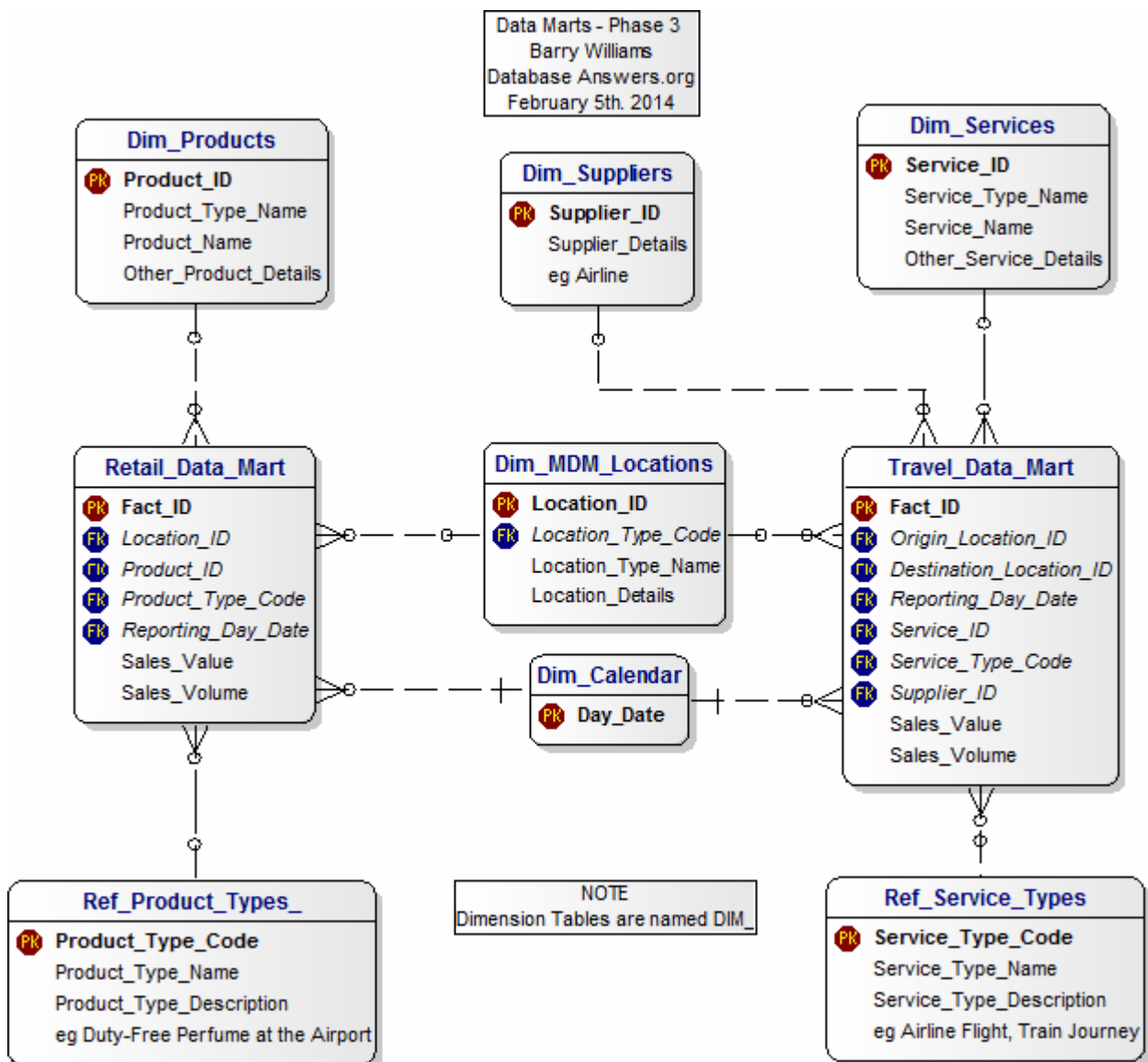
This Model is shown towards the bottom of this page :-

- [http://www.databaseanswers.org/data\\_models/bmews\\_design\\_pattern\\_1/index.htm](http://www.databaseanswers.org/data_models/bmews_design_pattern_1/index.htm)

Details of Products purchased are stored in the Retail Data Mart and details of Travels, which are Services, are stored in the Travel Data Mart.

Names of Dimension Entities all begin with 'Dim\_'.

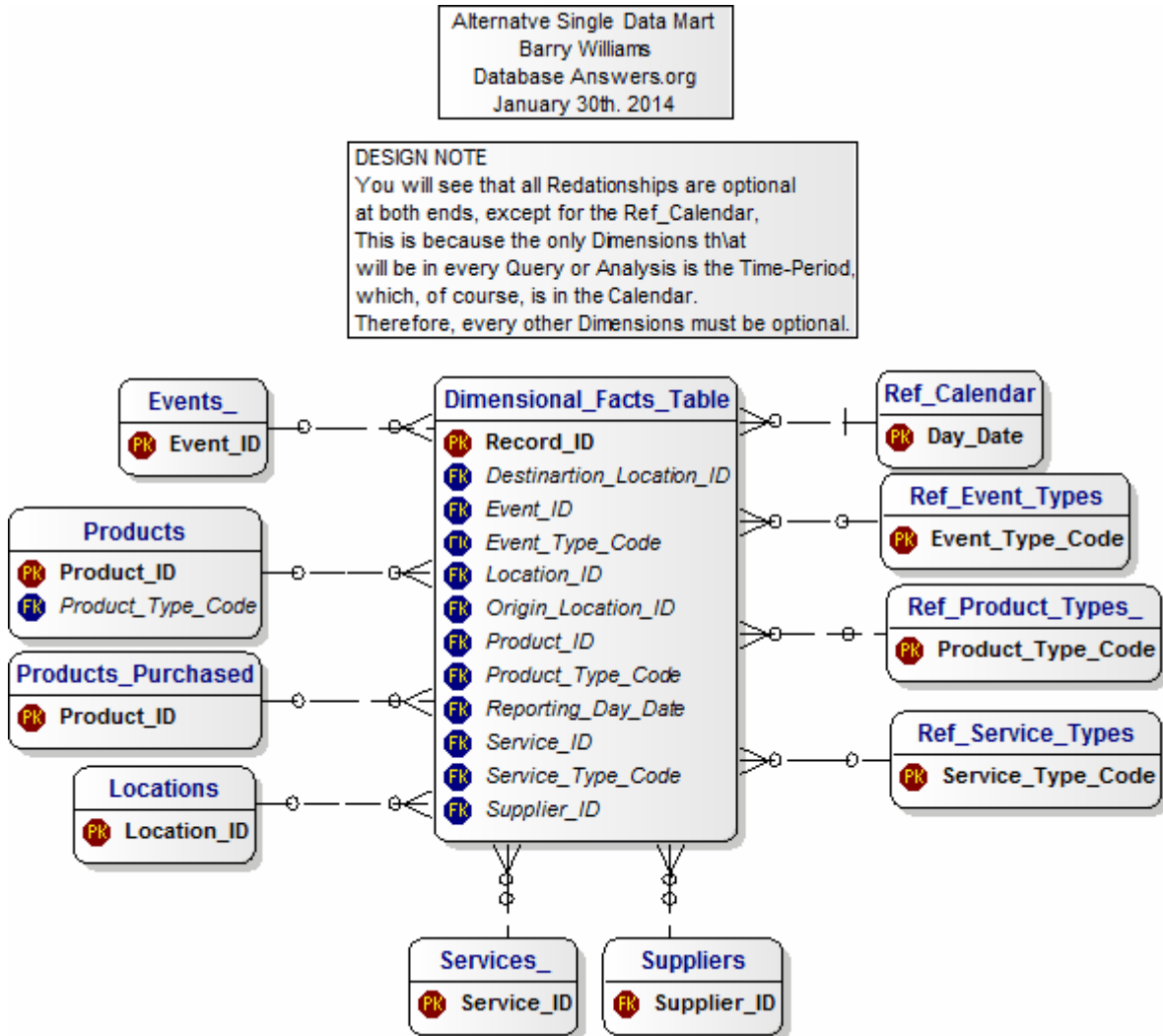
MDM in the name of the Locations means that Locations are part of Master Data Management, ('MDM').





### 14.3.5 Phase 4 – Produce an alternative Design

This is an alternative design with just one Data Mart with all Dimensions.



### Step 15 : Draw our Conclusions

When we review what we have done, we come to the following conclusion that it is feasible to design a Top-Level Data Model by following this Approach

- Establish a User Scenario to define the Requirements
- Use a library of Generic and Subject Area Models to build on previous work.
- Produce a draft of our Top-Level Model using our Library.
- Validate all the Models that we use and add more to our Library as appropriate.