Solutions

Problem 1

Dimensions for this problem: Franchise, Service, Customer, Merchandise, Calendar.

* Franchise: FranchId, FranchRegion, FranchPostalCode, FranchModelType.

FranchPostalCode (hierarchical Country code -> State Code -> District code -> City Code) (Source: ERD)

* Merchandise: MerchId, MerchName, MerchType, (Source: ERD)
* Service: ServId, ServCatName, (Source: ERD) EventypeCode, EventName. (Source: Spread Sheet)
* Customer: MembrId, MmbrName, MmbrZipCode, MmbrEmail, MemType, (Source: ERD) CustId, CorporateName (Source: Spread Sheet)
* Calendar: EventDate (Source: Spread Sheet), ServPurchDate, SaleDate, MmbrDate, (Source: ERD)

Problem 2

Measures for this problem, mainly come from merchandise sales and service purchases. Customer purchases of merchandise and services/ events are also considered.

* ServicePrice: ServCatPrice in ServiceCategory in ERD.
* MerchUnitPrice: MerchPrice in Merchandise in ERD.
* Quantity: Qty in Contains in ERD.
* EventAmount: From Special Spread Sheet.
* MemtypePrice: From MemberType from ERD.

Problem 3

For detailed grain, the following estimations are taken into considerations.

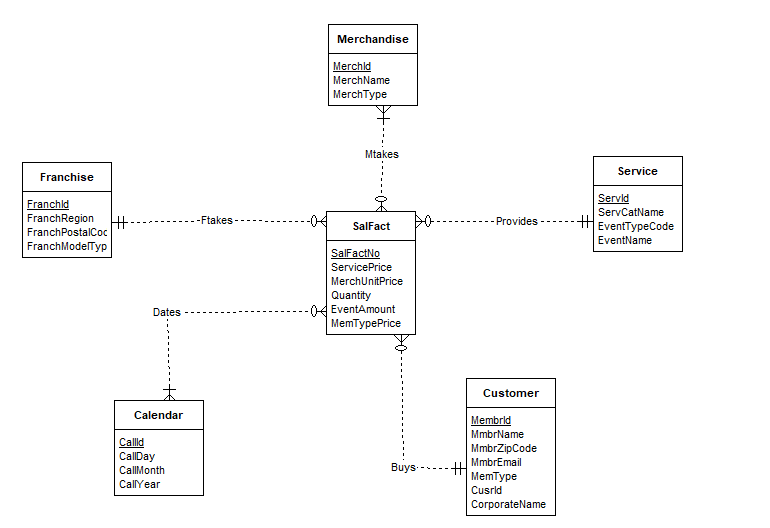
* 350 Franchises, 500 Merchandises, 320 (300 Spcl Spread sheets + 20 Service categories), Customers 50000(Members) + 150(Unique customers) = 50150. There are estimated values for measures.
* Merchandise sales are qunatities multiplied by unitprice

= 450000\*150000\*500

Services Sales = 100000

* Sparsity estimate:
  + 1- (fact table / product of dimensions)
  + 1- ( (33750000000000+100000)/(350\*500\*320\*50150\*365))
  + = 0.967075291.

Problem 4



The above star schema, provides a solution for the given datawarehouse. The Customer entity takes in values for the fact table SalFact through Buys. The merchandise sales are tracked with multiple values through Mtakes relationship. This total sale value is product of Quantity and MerchUnitPrice. Calender Entity tracks all the dates for Service Purchases, Merchandise Purchases and Customer Prices.

Problem 5

There are 2 summarizability problems in the mentioned schema. Both of them are Non-Strict Dimension Fact Relationship. It is present in Mtakes and Dates Relations. Both of them are M-N relations.

Solutions for both of them are, to replace both of them with 2 1-M relations. This implies that Mtakes will be replaced with 2 new relations Mtakes1 and Mtakes2. Similarly Dates is split into Pdate, Mdate. Pdate tracks sales for Services/ events. Mdate tracks sales of merchandise.

Problem 6

Sample data for Franchise

|  |  |  |  |
| --- | --- | --- | --- |
| FranchId | FranchRegion | FranchPostalCode | FranchModelType |
| F1 | Northwest | 98011 | Full |
| F2 | Mountain | 80111 | Medium |
| F3 | Central | 45236 | Limited |
| F4 | Mountain | 80112 | Limited |
| F6 | Valley | 14256 | Limited |
| F7 | Plains | 74125 | Full |
| F8 | North | 97568 | Medium |

Sample data for Merchandise

|  |  |  |
| --- | --- | --- |
| MerchId | MerchName | MerchType |
| MC1 | Wilson balls | Balls |
| MC2 | Wilson racket | Racket |
| MC3 | Adidas shoes | Shoes |
| MC4 | Racket stringing | Racket |
| MC5 | Nike Shoes | Shoes |
| MC6 | Wilson Caps | Caps |
| MC7 | Adidas Skipping Ropes | Skipping Ropes |

Sample data for Service

|  |  |  |  |
| --- | --- | --- | --- |
| ServId | ServCatName | EventTypeCode | EventName |
| SC1 | Ball machine | L-A | Adult Social |
| SC2 | Private lesson | L-B | Pioneer Social |
| SC3 | Adult class | L-C | Team Practise |
| SC4 | Child class | L-D | Child Social |
| SC5 | Grass cutter | L-E | Nutrition Care |
| SC7 | Serv coaching |  |  |
| SC9 | Stamina build-up |  |  |

Sample data for Calender

|  |  |  |  |
| --- | --- | --- | --- |
| CallId | CallDay | CallMonth | CallYear |
| C10000211 | 10 | 2 | 2013 |
| C10000212 | 11 | 2 | 2013 |
| C10000213 | 12 | 2 | 2013 |
| C10000214 | 13 | 2 | 2013 |
| C10000215 | 14 | 2 | 2013 |
| C10000216 | 15 | 2 | 2013 |
| C10000217 | 16 | 2 | 2013 |
| C10000218 | 17 | 2 | 2013 |
| C10000219 | 17 | 2 | 2013 |
| C10000220 | 18 | 2 | 2013 |
| C10000221 | 18 | 2 | 2013 |
| C10000222 | 19 | 2 | 2013 |
| C10000223 | 19 | 2 | 2013 |
| C10000224 | 20 | 2 | 2013 |
| C10000225 | 21 | 2 | 2013 |
| C10000226 | 21 | 2 | 2013 |
| C10000227 | 21 | 2 | 2013 |

Sample data for Customer

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| MembrId | MmbrName | MmbrZipCode | MmbrEmail | MemType | CusId | Corporate Name |
| 1111 | Joe | 98011 | joe@serv1.com | M1 | CC1 | First Data, Greenwood Village |
| 2222 | Mary | 80112 | mary@serv2.com | M2 | CC2 | DU Tennis, Denver |
| 3333 | Sue | 45327 | sue@serv3.com | M3 | CC3 | Creek High School, Greenwood  Village |
| 4444 | George | 45236 | george@serv4.com |  |  |  |
| 5555 | John | 98012 | [john@serv5.com](mailto:john@serv5.com) |  | CC4 | MAQ Software,  Hyderabad |
| 8888 | Nadal | 45698 | [nadal@serv2.com](mailto:nadal@serv2.com) | M5 |  |  |
| 9999 | Federer | 74125 | [fed@serv1.com](mailto:fed@serv1.com) |  | CC7 | Narayana Technologies,  Delhi |

Sample Data for SalFact

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SalFactNo | ServicePrice | MerchUnitPrice | Quantity | EventAmount | MemType Price |
| SF100001 | $15 | $3 | 2 | $1,000 | $1,000 |
| SF100002 | $75 | $200 | 1 | $500 | $800 |
| SF100003 | $150 | $100 | 1 | $200 | $300 |
| SF100005 | $125 | $40 | 1 |  |  |
| SF100006 | $65 | $175 | 1 | $1,750 | $1,800 |
| SF100007 | $15 | $45 | 4 | $450 | $1700 |
| SF100008 | $145 | $200 | 3 | $360 |  |