

**Data Warehouse Design:****Design Requirements:****1. The Processes and identity of Fact Tables****2. The Grain of each Fact Table**

The grain of fact tables are as follows:

1. Car Sales: A single sales record
2. Car Service: A single service record
3. Car Part: A single part listing

**3. The dimensions of each fact table**

The dimensions of each fact table are as follows:

1. Car Sales: 3 (**Customer\_Id, Car\_Id and Sale\_Date**)
2. Car Service: 7(**Booking\_Id, Part\_Id, Mechanic\_Id, Defect\_Id, Booking\_Date, Contact\_Id and Model\_Code**)
3. Car Part: 4 (**Part\_Id, Car\_Id, Supplier\_Id and Part\_Manufacturer\_Code**)

**4. The facts, including pre-calculated facts**

All required facts are included. However, some precalculated facts to fulfil business questions are as follows:

1. Car sales are ordered by day to find out the most popular cars.
2. Car sales are ordered by a count of each particular model to find customers buying a particular car.
3. Car sales are ordered by features to identify the most popular features.
4. Car service records are ordered by the manufacturer to find which manufacturer's cars are mostly serviced.
5. Car service records ordered by mechanic Id to find mechanics who serviced most cars.
6. Car part listings are ordered by count to find the most demanded part, grouping this by the model will give the most demanded part per model.

**5. The dimension attributes with complete descriptions and proper terminology**

All dimensions identified in part (3) and bolded in the diagram provide the basis for all drill-down operations on the database. Each dimension is a foreign key in the facts table and points to complete tables that provide further information. However, as star schema implementation is used so there are no further summarizing after-fact tables and hence aggregate tables identified in part (4) are required to solve some required business questions.

**6. How to track slowly changing dimensions**

Different types of changes may be required based on the following scenarios:

**Type 1:** This may be required for updating records of customers' details, and correcting any errors in entries. Mostly applicable on customer dimension table. This change may be required as per the needs of the transactional system. These changes will be usually minor and old values will not be retained by the system.

**Type 2:** These types of changes will be required when a car has a transfer of ownership, a manufacturer changes the configuration for a particular model or discontinues producing a particular model. Another scenario will be changing a supplier for a part. To apply such changes, additional columns will need to be added to the database. Here original information may still be retained with new rows being added to the database with new surrogate keys.

**Type 3:** These types of changes may be rare but not completely inexistent. These changes may be required in scenarios when we may need to compare performance across two competitors over a given period. This may be analyzing sales of two car models or comparing the longevity of parts from two different manufacturers. Applying these changes will be easier by constructing separate aggregate fact tables where the original data is not altered.

**7. The historical duration of the database**

The required business questions in this case do not require any complex historical analysis or future prediction. So, for given purpose data for the past 1-5 years would be enough to answer most of the given business questions.

**8. The urgency with which the data is extracted and loaded into the data warehouse.**

The business questions in this case do not require extensive daily based analysis. So, considering this the data can be transformed and loaded into the data warehouse from the transactional system once or twice every month. This will provide adequate system performance as the usual analysis will be done on monthly basis to assess performance in the past month, providing information for future decisions.