**DATA WAREHOUSE CONCEPTS AND DESIGN**

Contents

[1. Case Study: 3](#_Toc514666549)

[2. Defining the Business Process by E-R Model 3](#_Toc514666550)

[3. Data Warehouse Bus Matrix 4](#_Toc514666551)

[4. Assumptions 4](#_Toc514666552)

[5. Data Marts 4](#_Toc514666553)

[Business Process 1: Store Inventory 4](#_Toc514666554)

[Business Process 2: Vehicle Rental 5](#_Toc514666555)

[Business Process 3: Advertising Cost 6](#_Toc514666556)

[Business Process 4: Item Sales 7](#_Toc514666557)

[Business Process 5: VIP cards 8](#_Toc514666558)

[6. Fact Granularity 8](#_Toc514666559)

[7. Different Dimensions and attribute Hierarchy 8](#_Toc514666560)

[8. Other Design Features 8](#_Toc514666561)

[9. References 11](#_Toc514666563)

# Case Study:

CarHireOz is a large group of rental vehicle company over 1000 stores across all cities in Australia and lends different vehicles like vans, mini bus, cars etc. There are two options for customers to book the vehicle either through online or from rental stores. The store also sells posters, VIP cards, drinks, posters, maps and candies.

Revenue is generated by sales of side- items, rental fees and by any overdue fees. Every month the headquarter sends list with availability information, available vehicles from different suppliers, the categories of vehicles, the rating.

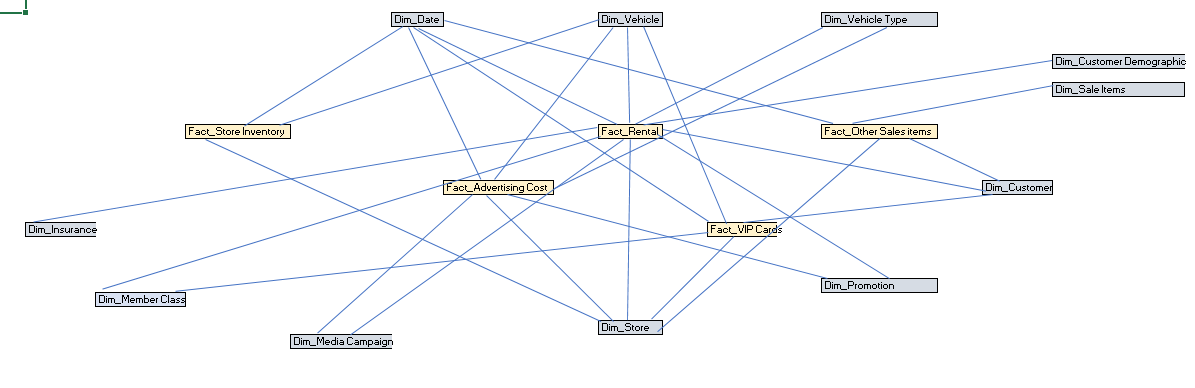
The vehicle company has different methods of payment. HQ analyses the usage of cards for all customers by stores. Each store has database to capture rental and sales figures.

Management wants detail analysis of performance for the company and decided to build a data warehouse to assist the business analysis and decision making for new car purchase.

Analysis for rental and sales were made to see historical data. They need to build daily/ monthly/quarterly/yearly list of top 10 vehicle category per store, per class of customers.

Revenues from rentals, VIP cards, other sale items need to be analysed. Management was looking for to find if the promotion were effective.

# Defining the Business Process by E-R Model



Business Process1: Store Inventory

Business Process2: Advertising Cost

Business Process3: Rental

Business Process4: Item Sales

Business Process 5: VIP cards

# Data Warehouse Bus Matrix

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Business Process/Dimensions | Date | Vehicle | Vehicle Category | Customer | Customer Demographic | Store | Product Items | Promotion | Member Class | Payment Method | VIP Card | Media Campaign |
| Store Inventory | × |  | × |  |  | × |  |  |  |  |  |  |
| Vehicle Rental | × | × | × | × | × | × |  | × | × | × |  |  |
| Other Sales items | × |  |  | × |  | × | × |  |  |  |  |  |
| Advertising Cost | × | × | × | × |  | × |  | × |  |  |  | × |
| VIP Cards | × |  | × | × | × | × |  |  |  |  | × |  |

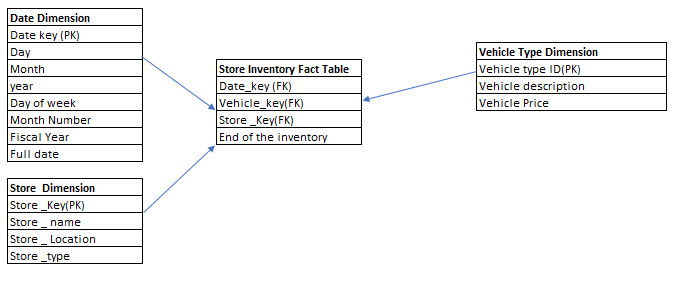
# Assumptions

Following assumptions where made and Data warehouse was built accordingly

1. I have assumed product to be other sale items.
2. I have considered drivers and customers as same entity.
3. I’m assuming all customers are local (who are living in Australia) as it was mentioned in case study that stores are in Australia. The data warehouse would be different if separate countries where involved.
4. I also considered the cost/revenue here is in Australian dollars.

# Data Marts

## Business Process 1: Store Inventory



*Please refer Tab 3 in excel.*

Store inventory is a **periodic snapshot**.

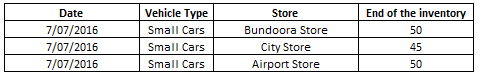
**Granularity**: Daily to check stock of vehicles in store.

**Measures:** End of inventory

**Dimensions**: Date dimension, Vehicle Type Dimension, Store Dimension

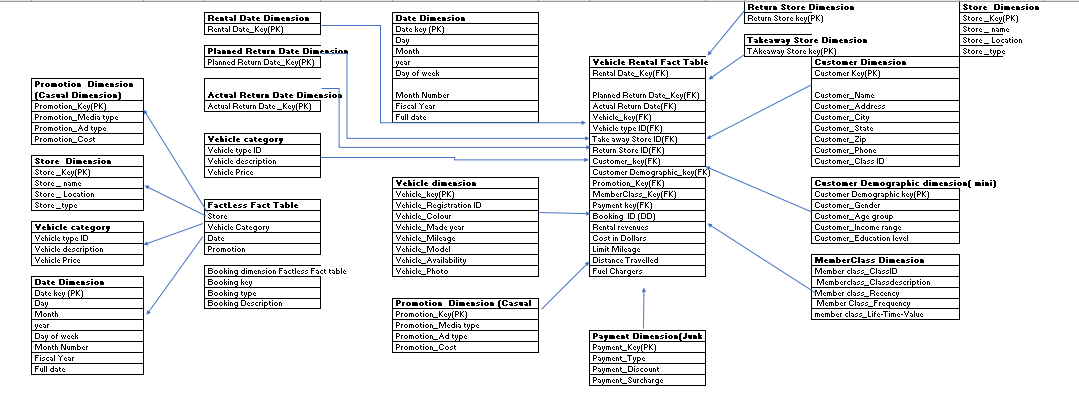
We have considered date, store and vehicle type dimensions to keep a track on the vehicle stock available in the store.

**Rows for fact table**:



This is helpful to find the information about daily inventory levels and also average inventory levels. The measure considered here is end of inventory and remain non- additive and grain is period not the individual transaction. For example: If there is no activity taking place during specific period a row is typically inserted in the fact table containing Null for each fact.

## Business Process 2: Vehicle Rental



*Please refer Tab 4 in Excel*

Vehicle rental is an **accumulative snapshot**

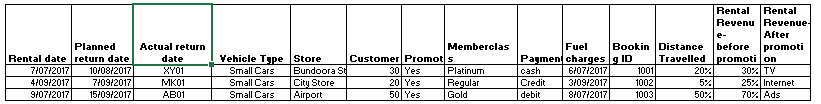
**Granularity**: Each vehicle per rental

**Measures**: Rental revenue, Cost in dollars, Limit mileage, Distance Travelled, Fuel Chargers

**Dimension**: Rental date dimension, Planned return date dimension, Actual return date dimension, Date dimension, Vehicle category dimension, Vehicle dimension, Store dimension, Customer dimension, Customer demographic dimension, Promotion dimension, Payment dimension, Member class dimension, Fact less fact table

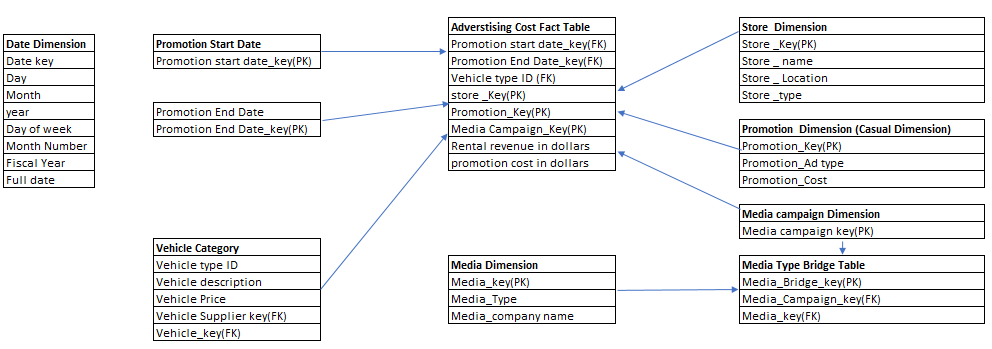
The Date dimension which act as a role-playing dimension have different dates and views for three roles. The store dimension had views for two roles. The promotion dimension describes promotion condition in which the product was sold, this type of dimension is called casual dimension where the related factors thought to cause a change. There are few types of vehicles which are not under promotion in CarHireOz company, thus to avoid NULL in the fact table a row is included in the promotion dimension to identify “no promotion in effect”. There are two fact-less fact tables consist of store dimension, Vehicle category dimension, date dimension and booking dimension fact-less fact table. These fact-less fact table helpful to analyse that there are vehicles in store which didn’t go for rental, these helps to record complete information and it does not contain any measures. The Booking ID is considered as Degenerated dimension which does not have any corresponding dimension table. It plays an integral role in fact table. Payment dimension acts a junk dimension and consists of several low cardinality flags and attributes in single dimension table. Customer Demographic acts as mini dimension which gives complete information about the customers. Thus, this model forms a fact constellation which helps to measure analytical process with multiple fact tables.

**Rows of Fact table**



*Please refer Tab 4 in excel*

## Business Process 3: Advertising Cost



*Please refer Tab 5 in Excel*

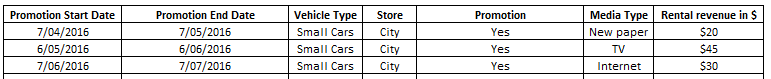
Advertising cost **Transactional Snapshot**

**Granularity:** Per Vehicle per promotion revenue analysis

**Measures:** Rental revenue in dollars, promotion cost in dollars

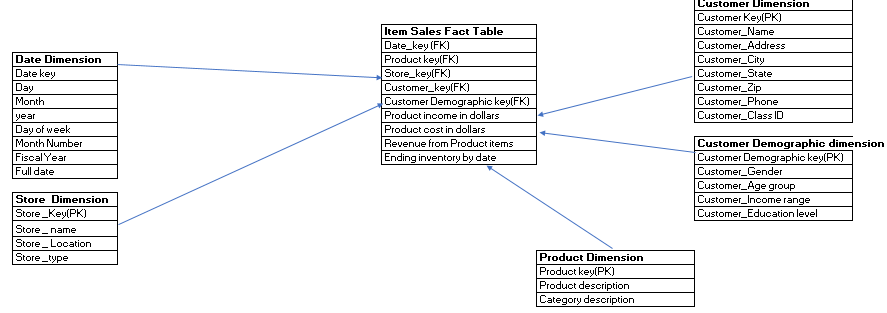
**Dimensions:** Promotion start date, Promotion end date, Store dimension, Promotion Dimension, Media campaign dimension, Media type bridge table, Vehicle category

Advertising cost here is considered as transactional as we are considering per promotion per transaction. Media bridge dimension is used where media campaign acts as a multi valued dimension and attached to fact table through a group dimension key with help of bridge table with one row of each media in a group. There are two types of dates considered as promotion start date and promotion end date.



*Please refer Tab 5 Excel*

## Business Process 4: Item Sales



*Please refer Tab 6 excel*

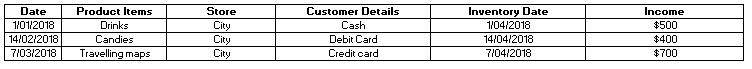
Item Sales **Transactional Snapshot**

**Granularity:** One item per sale

**Measures**: Product income in dollars, Product cost in dollars, revenue from product item, Ending inventory by date.

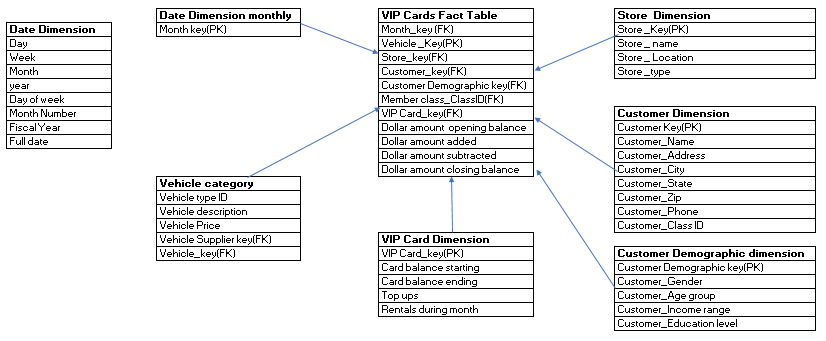
**Dimensions**: Date dimension, Store dimension, Customer dimension, Customer demographic dimension, Product dimension

Products are the other items like candies, drinks, posters, travelling maps etc. Here, this dimension is considered as Slow changing dimension as it changes slowly unpredictably. This helps to record actual SKU sold. With help of this we can determine what products were sold. On the other hand, inventory generate snapshot table.



*Please refer Tab 6 in excel*

## Business Process 5: VIP cards



*Please refer Tab 7 in excel*

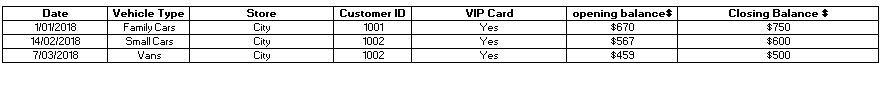
VIP cards **Transactional Snapshot**

**Granularity**: Per month transaction per VIP card per customer

**Measures**: Dollar amount opening balance, Dollar amount closing balance, Dollar amount added and Dollar amount subtracted

**Dimensions:** Monthly date dimension, Vehicle category, VIP card dimension, Store dimension, Customer dimension, Customer demographic dimension.

VIP cards are given to customers who wants to get benefited from discounts, different surcharges and special top ups over the card. Here customer demographic is demographic dimension which acts as a mini dimension. This fact table helps to understand that in particular month how many VIP card holders have transactions.



*Please refer Tab 7 in excel*

# Fact Granularity

*Please refer Excel Tab 8*

1. Different Dimensions and attribute Hierarchy

*Please refer Excel Tab 9*

1. Other Design Features

*Please refer Excel Tab 10*

# References

Dr Jinli Cao. (2018). Data Warehouse concepts and Design. Melbourne. Retrieved from https://lms.latrobe.edu.au/course/view.php?id=54465

Ralph Kimball, M. R. (2nd Edition). *The data Warehouse Tool Kit.* Wiley Computer Publishing.