CSE5DWD – Data Warehouse Concepts and Design Assignment, Semester 1 2020

Dimensional Modelling Business Cases

Submitted by: Bhargavi Singaravelu (20113574)

Your task is to design a data warehouse for Heathcote Winery Group using multidimensional Modelling. Your design needs to encompass the following steps:

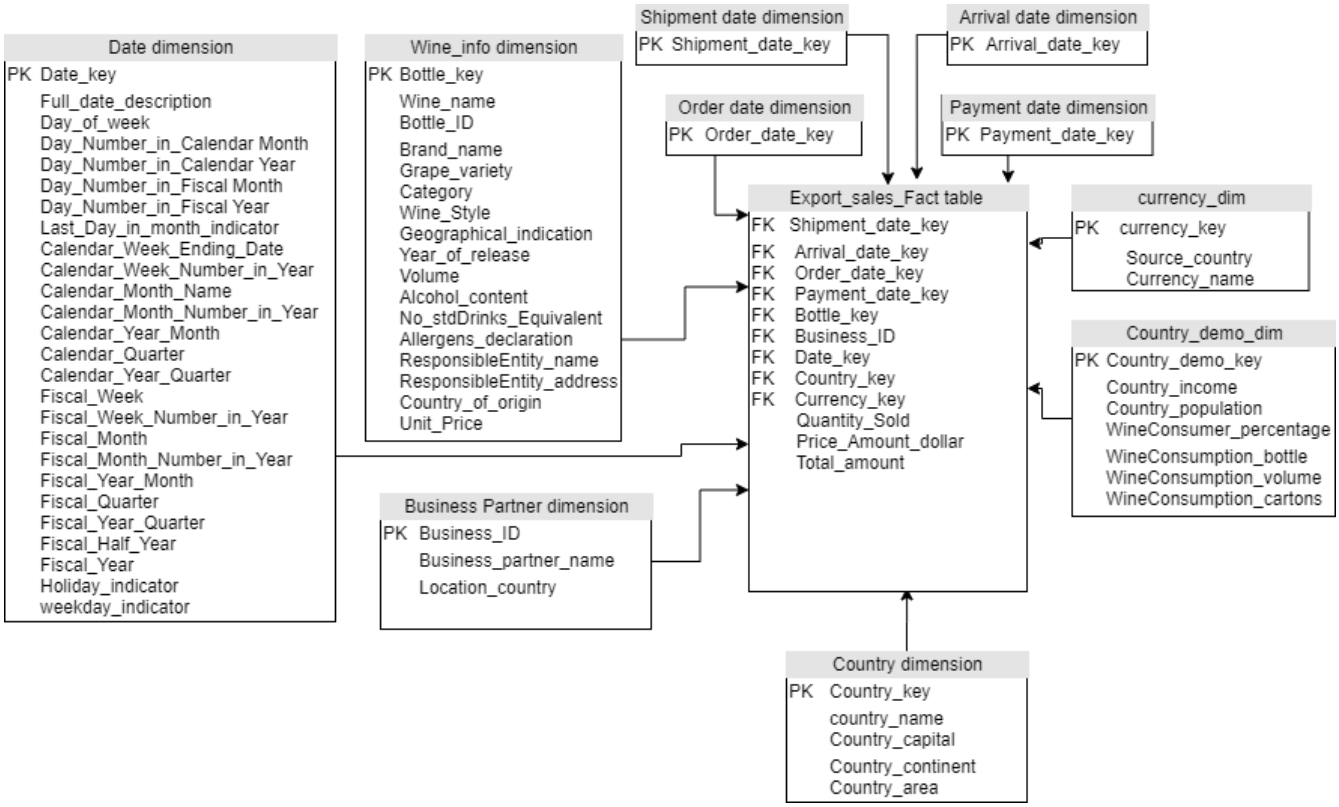
(i) First construct a Data Warehouse Bus Matrix to identify the company's business processes and any likely Data Marts. (10 marks)

	Business partners	Wine bottle	Country	Date	Year	Customer	Local owners	Restaurants	Promotion/ Discounts
Export sale Management	✓	✓	✓	✓	✓	✓			
Local sale management		✓		✓	✓	✓	✓		✓
Daily tours		✓		✓	✓	✓		✓	✓

(ii) Identify the grain of each star/snowflake schema to design the dimensional model for business processes you have identified, ensuring your dimensions are conformed, primary and foreign keys are clearly labelled, and that your attributes are named using verbose textual descriptions. You also need to provide three rows of the fact table your designed to explain the meaning of the rows. (40 marks)

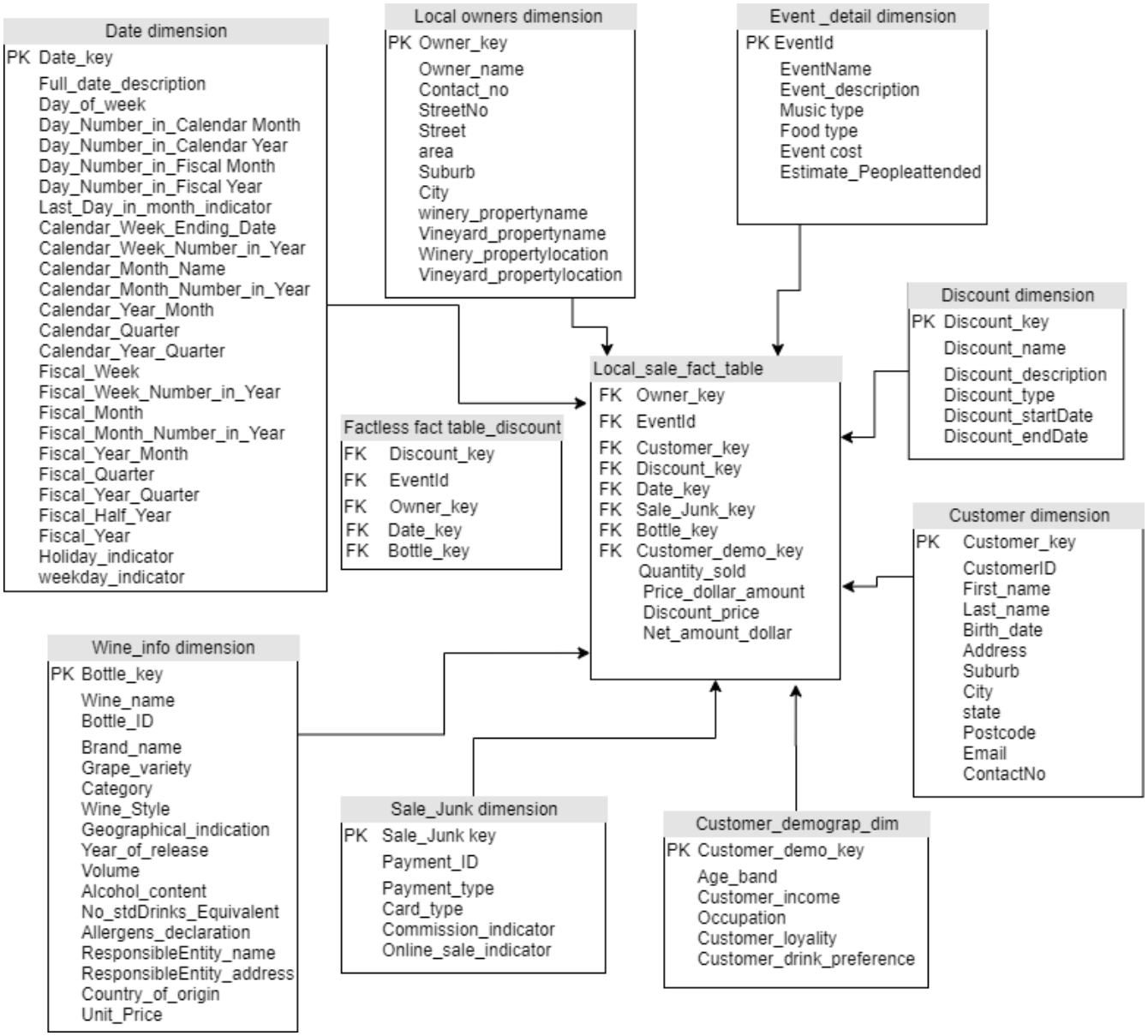
1. Export_sale_fact_table:

Shipment_date_key	Arrival_date_key	Order_date_key	Payment_date_key	Bottle_key	Business_ID	Date_key	Country_demo_key	Country_key	Currency_key	Quantity_Sold	Price_Amount_dollar	Total_amount
127	133	120	134	274519	21	138	198670	127	54	2000	100000	100000
128	134	121	135	927451	22	138	944565	89	67	2000	60000	60000
129	135	122	136	745192	23	138	176908	32	90	5000	150000	150000



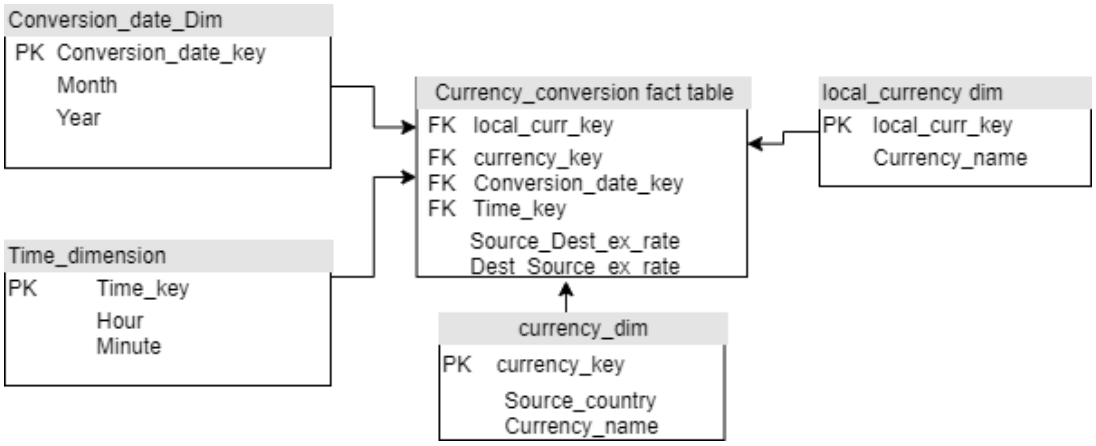
2. Local sales Fact Table :

Owner_key	EventId	Customer_key	Discount_key	Date_key	Sale_Junk_key	Customer_demo_key	Price_dollar_amount	Discount_price	Net_dollar_amount
98	8097	23456	31	138	22	234561	800	40	760
21	9087	43214	24	138	21	432143	1780	70	1710
51	7895	56432	32	138	12	564322	2100	130	1970



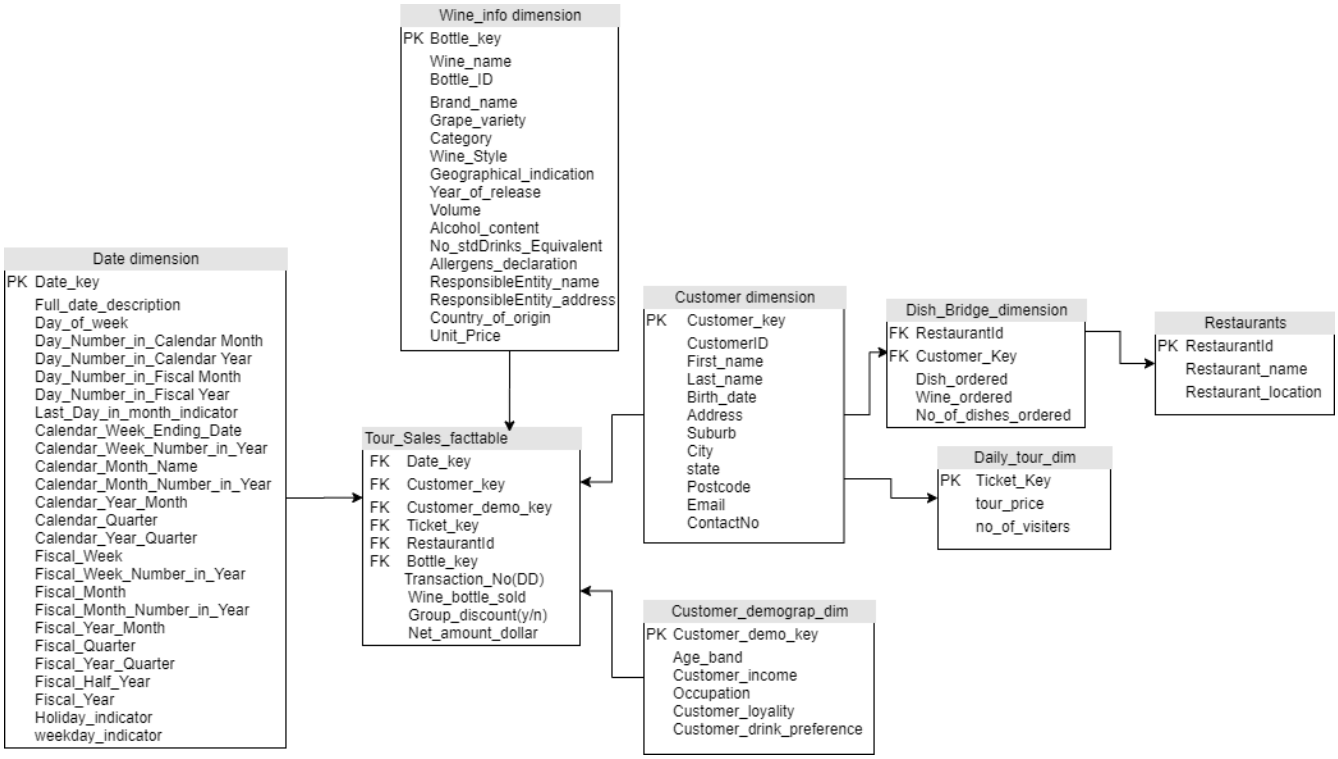
3. Currency Conversion fact table:

local_curr_key	currency_key	Conversion_date_key	Time_key	Dest_local_ex_rate
10	54	134	890	0.22
10	67	135	567	0.015
10	90	136	423	0.64



4. Tour sales fact table:

Date_key	Customer_key	Customer_demo_key	Ticket_key	RestaurantId	Bottle_key	Transaction_No(DD)	Wine_bottle_sold	Group_discount(y/n)	Net_amount_dollar
137	23456	234561	9876	3	274519	67890765	3	y	120
137	43214	432143	4566	9	927451	56432587	8	n	234
137	56432	564322	2134	21	745192	56754245	2	n	432



(iii) Create the following table with a row for each fact table in your design, indicating the granularity of each fact and a brief justification for choosing that granularity. (12 marks).

Fact Table name	Fact Granularity	Fact Table Type	Brief Justification
Export sale fact table	Quantity exported (wine)	Accumulating fact table	This type of fact table is used to show the activity of export process that has a well-defined beginning and end. An order moves through specific steps until it is fully processed.
Local sales Fact Table	Quantity sold(wine)	Transaction fact table	This fact table gives the details of the total sale made in the local area
Currency Conversion fact table	Destination source exchange rate	Periodic snapshot fact table	This fact table changes the values from time to time as the exchange rate of the currency changes very frequently
Tour sales fact table	Net amount of the visitor expenses per tour	Transaction Fact table	This fact table gives the details of the expenditure for each visitor to the vineyard which is given by their transaction number.

(iv) Create the following table with a row for each dimension table in your design, giving a brief justification for choosing that dimension, and indicating any attribute hierarchies that exist within the dimension. (12 marks).

Dimension Table name	Brief justification	Attribute Hierarchies
Date_dimension	Gives information about the period in which the day occurs, and whether is a holiday and whether is a weekday or weekend	-
Wine_info_dimension	Gives the bottle information such as the name, brand , type , volume, alcohol content etc.	-

Business_Partner dimension	The basic information of the organization's business partners	-
Country_dimension	The country details of where the wine is being exported	country_name-> Country_capital-> Country_area-> Country_continent
Country_demo_dimension	This is a demographic dimension which stores the attributes that tend to change often.	-
Currency_dim	Has the record of the multiple currencies of the business partners of different countries	-
Shipment_date_key	Stores the date of shipment of the product	-
Arrival_date_key	Stores the date of arrival of the shipment	-
Order_date_key	Stores the date of order of the product	-
Payment_date_key	Stores the date of payment of the product arrived.	-
Local_owners dimension	Stores the information about the local owner and the winery and vineyard owned .	StreetNo->Street->area->Suburb->city
Event_detail dimension	Stores the information on the special event days for the management reference	-
Discount Dimension	Has the list of discounts and promotions along with its description	-
Customer dimension	Stores the details of the customer who buy products online or through cellar door	Address->Suburb->City->State->Postcode
Customer_demograp_dim	Stores the frequently changing attributes of the customer dimension.It is a mini dimension for the customer dimension	-
Sale_Junk dimension	Indicates the payment method type and sales type(online or cellar door)	-

Dish_Bridge_dimension	It is a bridge dimension to hold the details of the food and wine ordered by the vistors of the daily tour in the local restaurants and eateries.	-
Daily_tour_dim	Holds the tour details	-
Restaurants	Stores the restaurant and eateries details.	

(v) Create the following table with a row for each design feature you have used, such as handling of possible null foreign keys, and the inclusion of any fact-less fact tables, degenerate dimensions, role playing dimensions, junk dimensions, outriggers, mini-dimensions, bridge dimens or any other design techniques discussed in the lectures. Provide a brief description of each design feature used (how and where it is used – not the theory behind the concept), any possible problems & solutions with a justification for their use.

Design Feature	Brief Description	Brief Justification
Role-playing dimensions	Shipment_date_key, Payment_date_key, Order_date_key, Arrival_date_key,Conversion_date_key	An order is taken on a specific date; the order is packed at the warehouse on one or more dates; the order is shipped to the customer on a date; the order has a delivery date; the order's payment received on a date; the payment involves multiple currencies and require conversion. All of these dates need to be recorded because they all play a role in the retail sale.
Factless Fact table	Factless_fact_table_Discount	The table has foreign keys of dimensions like discount, wine_info etc. which will have details of all the wine bottles regardless of whether they are sold or not and can be used to extract the wine bottles that were not sold during the special event promotion

Junk Dimension	Sale_Junk dimension has records of sale type and payment type	To classify the sale type and payment type we need to create a dimension for each type, which is a lot of dimension table in the schema. In order to eliminate small dimensions, we create a single “junk” dimension which cross joins all possible attributes into a single dimension which will be used in the fact table.
Mini dimension	Customer_demograp_dim record the frequently changing attributes	The age, income , preference of each customer tend to change often and these attributes cannot be stored in the customer dimensions as they are rapidly changing .
Bridge Dimension	Dish_Bridge_dimension to store number of dishes ordered by the customers	The number of dishes ordered by customers in the restaurant has no fixed limit and has there can be multiple values associated with the grain of the fact table. Hence, we use a bridge dimension to store the number of dishes ordered by each customer in different restaurants.

(vi) Identify which fields from your facts/dimensions are required to answer 7 business questions listed as the Bold words and marked with (*) above.

The management wants to analyse the data every quarter to see

• the trends of the market such as the average value of the exports to each country/area (*)

Dimensions : Country Dimension – to denote which country

Fact: Quantity_sold and Total_amount – gives the Value of the exports

• the arrived date between the ordered date and the payment date between the shipment date, how soon the payment can be received after the shipment? (*)

Role-playing dimensions:

Shipment_date_key, Payment_date_key, Order_date_key, Arrival_date_key

Managers want to analysis the event sales:

• Compare the total sales of the Shiraz wine style in last quarter of 2018 with the same period of quarter in 2019 (*)

Dimension: Date dimension - Calender_year_quarter

• What wines are not sold during the event discount? (*)

Factless Fact table : Factless_fact_table_Discount

Managers want to analyse visitors' habits in a day tour. They want to

• find out the most popular group of dishes are ordered by the visitors (*)

Bridge dimension: Dish_Bridge_dimension

• find the most popular wine glasses ordered by the visitor (*)

Measure - Wine_bottle_sold

Wine_info_dimension – Wine_name, Category

• Has the holiday period in Christmas month increased the tourists number comparing with the same period in last year? (*)

Date Dimension- Holiday_indicator