



Engineering Core

Business Analytics & Data Model

EN6001

Assignment-1

Sheikh Muhammed Tadeeb (AU19B1014)

❖ Problem Statement:

Students will interview the business owner to understand their business (photographs are mandatory). They will gather the information and will create a data model on the following criteria:

1. How does the business operate?
2. What are the problems and challenges faced by the business?
3. What is the expansion plan or wish list items?
4. What type of reports will help them in working on above two points? What would be the report formats?

❖ Building Data Model:

Step1) Data Extraction:

Q1) How does the Business Operate?

Ans: I interviewed the business owner and the workers to gather the data. The Drive link of interview is below:

Link: [Interview recording link](#)

Business Name: *Royal Ice-creams*

Business Type: *Sole Proprietorship (Edible items)*



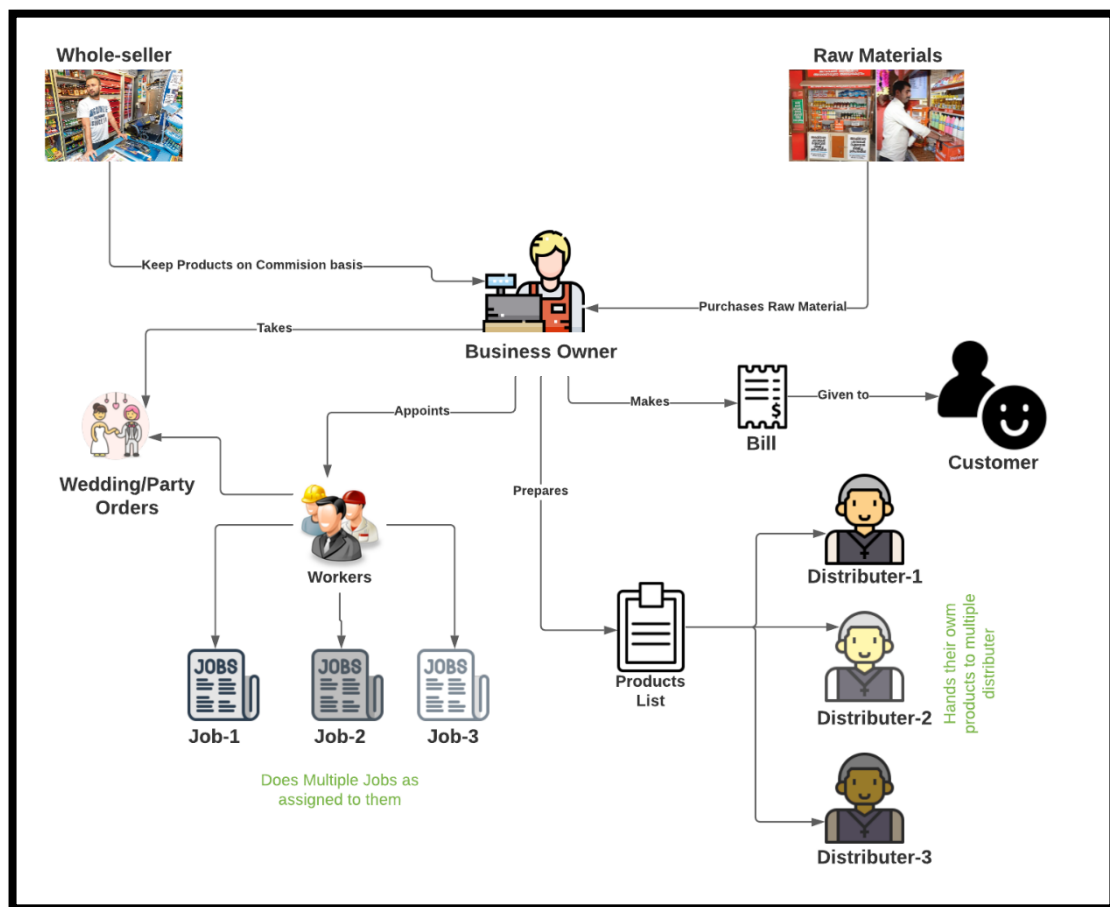


Fig1.0 Business Work Flow

➤ **Explanation of Fig 1.0)**

1. The Business owner produces many products like ice-cream, cold-drinks etc. at their own shop for which they purchase the raw materials and then process them.
2. The Business owner also keeps many other edible items like chocolates, coca-colas, food color etc. on commission basis.
3. The Business takes wholesale orders too for parties and weddings.
4. There are 3 workers in the shop each having a specific role i.e., waiter, ice-cream maker, Falooda maker.
5. The Business owner also supplies their own products to different small distributors in the city.

6. The bills, date of order is generated only for the big orders and are maintained in a register.

Q2) What are the problems and challenges faced by the business?

Ans: The problems and challenges faced by the business are:

1. The Shop is on a small lane.
2. The recently opened wine shop has affected the business as families avoid coming.
3. Covid is one of the major challenge due to which the ice-cream business suffered a lot as people avoided cold stuff and public gathers/parties.
4. Few trusted members who can take care of the shop.
5. Dependent on one ice-cream worker who process the raw materials without having any alternate choice.
6. Can't connect to food suppliers like Zomato or Swiggy because people want fresh beverages especially faluda, lassi and ice-cream.
7. The police station is just behind the shop which makes it difficult to open the shop late nights.

Q3) What is the expansion plan or wish list item?

Ans:

1. After Interacting with the owner, it was very clear that as of now he as no expansion plan due to lack of any trusted helping hand but he does wish to keep extra products like Nitrogen ice-cream and waffles.
2. Secondly, he has to write down all the data in the register manually so he wants a desktop application where he can easily store all the data in a single platform.

Q4) What type of reports will help them in working on above two points? What would be the report formats?

Ans: The Shopkeeper was not fixed with any one type of report and he was flexible enough to ask me the suggestions on the report type and format. Being a data analyst student, I suggested and confirmed with him that a dashboard containing different graphs based on various transactional details and KPI's could really help the owner in doing the analysis.

Step2) Data Cleaning:

Well, I Interviewed the business owner & workers and after interviewing them I got the data in very lame format and therefore I converted it into functional requirements that could be easily understood by a technical person or an Analyst.

➤ *Functional Requirements:*

- 1) The Business consist of many workers with different roles and tasks.
- 2) Workers are identified by labour_id. Other details like name, D.O.B, address, shift, designation, salary and roles assigned must be recorded.
- 3) The Business owners could take up any role in which he is specialized or has experience.
- 4) Some products are sold on commission basis that should have information about supplier-Id, supplier name, phone number, address.
- 5) While other major beverage products are manufactured in the shop itself. These must have information about different ingredients name and their associated cost price and total manufacturing cost.
- 6) Every product has a name, category, unique id, flavors, discount, cost price, selling price, profit, quantity sold or purchased, purchase date or manufacturing date and sold date or expiry date.

- 7) Each table has a unique table number associated with it and no's of chairs are different for different table slots.
- 8) The billing information must have bill id, bill amount, date and time associated with it.
- 9) The products manufactured in the shop are also supplied to various other retailers or distributors so distributor id, discount%, distributor name, phone nos, address must be recorded.

Q5) Students will identify any type of transaction or event and create at least 5 analysis questions for the same.

Ans:

- Transaction or Event Identification:

S.nos	Transaction Type	Event
1.	Purchase of products (<i>cost price associated</i>).	Party/Program orders
2.	Selling of products (<i>selling price associated</i>).	Machine maintenance
3.	Distribution to retailers (<i>sales & profit associated</i>).	New Machine purchase
4.	Salary distribution to workers (<i>Net profit associated</i>).	
5.	Credit amount on Business (<i>Net profit associated</i>).	
6.	Credit amount on Distributer (<i>Net profit associated</i>).	

- Analysis Questions:

Q1. Profit wise top 5 products over last 3 years?

Q2. Find out the Category wise maximum sales and average discount % for that category?

Q3. Give the Count of orders per product per distributor over last 5 years?

Q4. Top 3 products by flavour and type during summer season for last 3 years?

Q5. Week with least sales in year 2019 & 2020. (Considering COVID situation)?

Q6. Generate monthly labour report based on working hours for past 3 years?


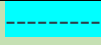
Q6) Identify the fact and dimension attributes based on the star schema concept?

Ans:

- Identification of Entities:

1. Worker	2. Supplier	3. Distributer	4. Bill
5. Product/Goods	6. Credit	7. Business	

- Identifying Attributes & primary key:

Primary Key	
Weak Key	

- 1) Worker (labour_id, Worker_name, D.O.B, address, shift, designation, salary).

Labour_id	Worker_name	DoB	Address	Shift	Designation	Salary
-----------	-------------	-----	---------	-------	-------------	--------

- 2) Supplier (Supplier_id, supplier name, phone_nos, address).

Supplier_id	Supplier_Fname	S_Mname	S_Lname	Phone_nos	S_Address
-------------	----------------	---------	---------	-----------	-----------

- 3) Distributer (Distributor_id, discount%, distributer_name, phone_nos, D_Address).

D_id	D_Fname	D_Mname	D_Lname	Discount%	Phone_nos	S_Address
------	---------	---------	---------	-----------	-----------	-----------

- 4) Bill (Bill_id, bill_amount, date, time).

Bill_id	Bill_amt	Date	Time
---------	----------	------	------

- 5) Product (**P_id**, P_name, category, flavors, discount, cost_price, selling_price, profit, qty_sold or qty_purchased, purchase_date or manufacturing_date and sold_date or expiry_date). *NOTE: We can do specialisation in products.*

P_id	P_name	P_category	Flavors	Discount	Cost_price	Selling_price	Profit	qty_purchased	qty_sold	purchase_date	sold_date
P_id	P_name	P_category	Flavors	Discount	Cost_price	Selling_price	Profit	qty_purchased	qty_sold	Manufacturing_date	expiry_date

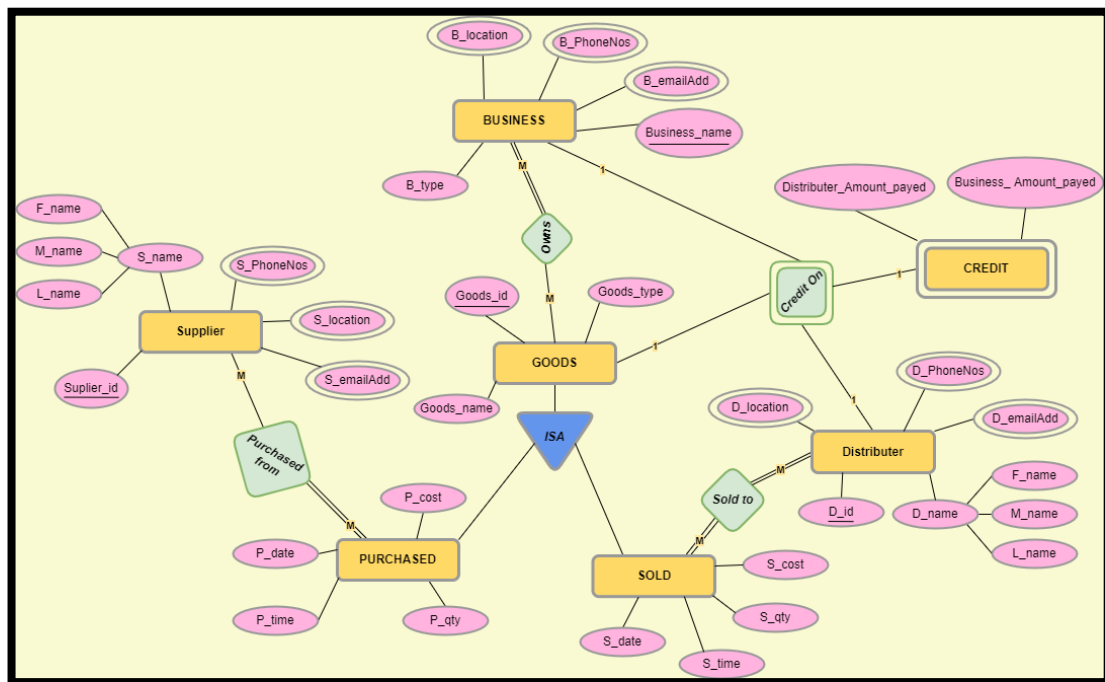
- 6) Credit (**Credit_on_business**, Credit_on_Supplier).

Credit_on_business	Credit_on_Supplier
--------------------	--------------------

- 7) BUSINESS (**Business_name**, B_type, B_location, B_PhoneNos, B_emailAdd).

Business_name	B_type	B_location	B_PhoneNos	B_emailAdd
---------------	--------	------------	------------	------------

- Proposed ER Model:

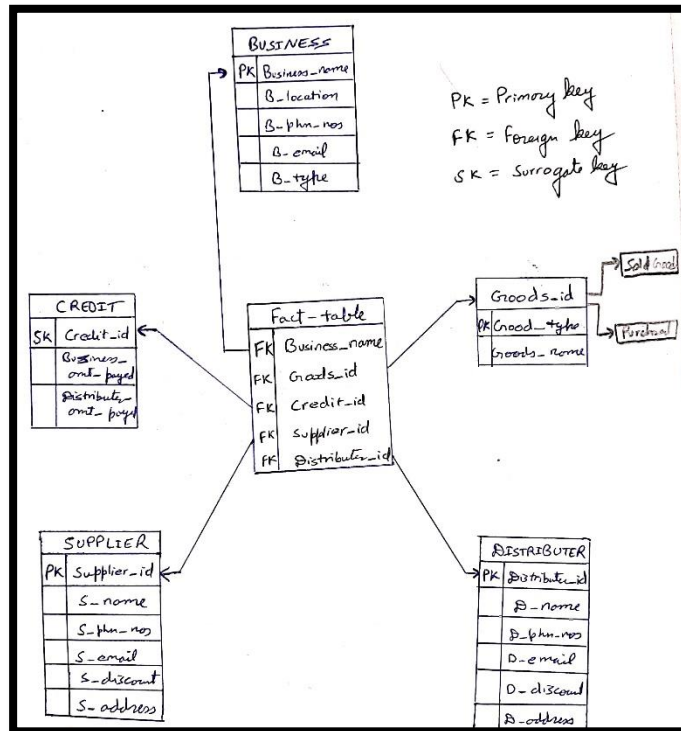


Step3) Evaluate and Deploy Data-Model:

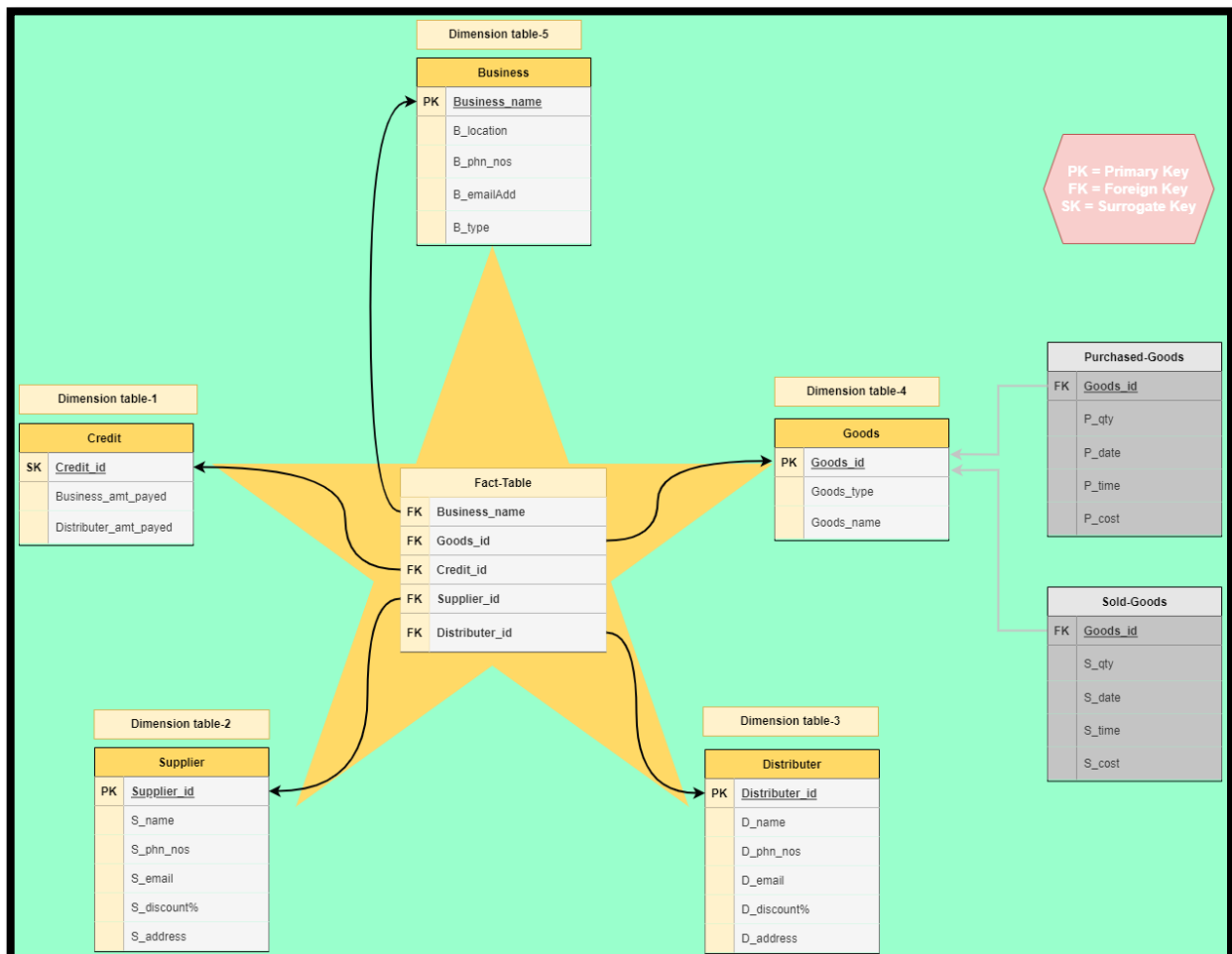
- Q7) Prepare the data model as a star schema?

Ans:

○ Hand-Drawn:



○ Digitally on Software:



❖ Conclusion:

The following are the 3 major inferences after doing the assignment:

1. A data model determines how data is exposed to the end user. Optimally creating and structuring database tables to answer business questions is the desired role of data modeling, setting the stage for the best data analysis possible by exposing the end user to the most relevant data they require.
2. The fundamental objective of data modeling is to only expose data that holds value for the end user. Clearly delineating what questions a table should answer is essential, and deciding on how different types of data will be modeled creates optimal conditions for data analysis.
3. Advances in cloud data warehousing have tied the capability to source both structured and unstructured data directly to an end user's ability to analyze and implement that data within minutes. A comprehensive, pragmatic data model further facilitates BI tools such as tableau helps turning raw data into business value.
4. Making assumptions is not a good process while creating data model. If user gives us the flexibility to decide then we must confirm our assumption with our user/customer.
5. I found that queries will be simpler in star schema as star-schema join-logic is generally simpler than the join logic required to retrieve data from a highly normalized transactional schema.
6. Query performance gains as star schemas can provide performance enhancements for read-only reporting applications when compared to highly normalized schemas.
7. Fast aggregations as the simpler queries against a star schema can result in improved performance for aggregation operations.

Business Analytics & Data Model

