Student ID : S3857825
Name : Karthik Kolume

DATABASE CONCEPTS - ASSESSMENT 4- PROJECT

Section 1. Overview.

Scenario: The Retail company required a new database system to be built for the Sales team of the company. They required a database system for activities such as storing Sales information, monitoring sales records, and conducting a sales management process effectively by not having to face any anomalies with respect to the sold items that are spread across many branches. A simple database system where the data is accurate with no compromise in data integrity. The system which could be used to access the data easily and at the same time is secured. The relational Database model was built to satisfy all requirements of the company.

Final Database: Relational Database model was built with 6 entities namely Customers, Branch, Location, Time, Item, and Sales Information. The database is built considering the dependencies and relations between every entity present in the database. With this relational database model, Sales executives can now find the desired information using the software tool SQLite. Complicated information can be fetched from the database by passing on the SQL queries.

- **Customers**: Entity consists of all the basic details of the customer such as Customer ID, First Name, Last Name, Phone number, and e-mail address of the customer who bought items from the company. Every Sale belongs to a Customer and every customer is uniquely identified by Customer ID.
- **Branch**: The Company has its branches spread across the country. This entity consists of the details of every branch of the company. It consists of details such as Branch ID, Branch Name, and Owner of the branch. Every sale belongs to a particular Branch and Branches are uniquely defined by Branch ID.
- **Location**: The location or the address of the customer is being recorded in the Location entity. This entity consists of the complete address of the customer to where the item is delivered to the customer. This entity consists of Location ID, Location Name, state, and Pincode of the location of the customer. Every Customer belongs to a particular Location and Branches are uniquely defined by Location ID.
- **Time**: The time of delivery is recorded in this entity. This entity consists of all the details of the delivery date such as Time ID, Day, Month, Quarter, and Year of the delivery. Every sale belongs to a particular Time. Time is uniquely defined by Time ID.
- **Item**: The details of the product/Item delivered to the customer are recorded in this entity. It consists of the Item key, Item name, Brand, Sold by, and Item group to which it belongs to. Every sale belongs to a particular Item and Items are uniquely defined by the Item Key.
- **Sales Information**: This is the most prominent entity in this database which references most of the other tables. It consists of all the Sales information of the company such as quantity sold and the amount sold to the customers. Sales Information/details are uniquely defined by Sale Order ID.

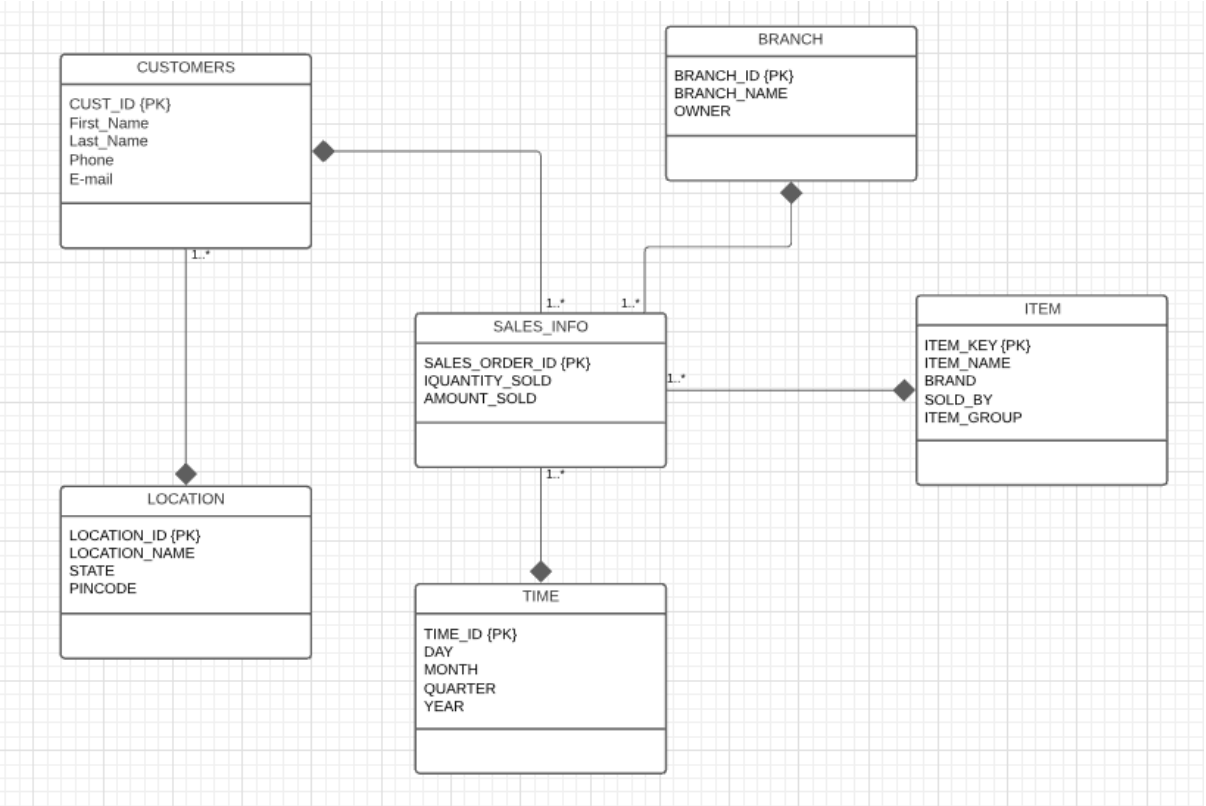
Design Improvements: Initially the database was designed by just considering the entity names and had created just the basic structure of each entity. From week 4 to week 13,

1. Improved the structure of each entity and included a new entity called Customers. Addition of new entity Customers was done to get an understanding of items being sold to which customers and their corresponding details.
2. Updated Customer location by including another entity called Location to set up a proper channel of the items being delivered to the customer's location.
3. Designed an ER diagram using Unified modeling language. Using this model, we were able to achieve the exact relationship between entities.
4. Relational database schema with primary key and foreign key annotations was derived from the Entity relations in UML design. Verified that there is no redundancy present in any relations of the design. All the relations are in 3NF or BCNF.
5. From the derived schema the complete Sales database was built using SQL DDL.
6. Later the Data was manually imported to the database from flat files(.csv files).

Student ID : S3857825
Name : Karthik Kolume

Section 2. ER diagram and relational database schema. (SQL DDL is in the Appendix.)

1. ER diagram



2. The relational database schema with primary key and foreign key annotations.

TIME(TIME_ID, DAY, MONTH, QUARTER, YEAR)

ITEM(ITEM_KEY, ITEM_NAME, BRAND, SOLD_BY, ITEM_GROUP)

BRANCH(BRANCH_ID, BRANCH_NAME, OWNER)

CUSTOMERS(CUST_ID, First_name, Last_Name, Phone, E-mail, LOCATION_ID*)

LOCATION(LOCATION_ID, LOCATION_NAME, STATE, PINCODE)

SALES_INFO(SALES_ORDER_ID, QUANTITY_SOLD, AMOUNT_SOLD, CUST_ID*,BRANCH_ID*, ITEM_KEY*, TIME_ID*)

Note that every Relation is already in Normalized form with no redundancies. All the relations are in 3NF/BCNF.

Student ID : S3857825
Name : Karthik Kolume

Section 3. Data analysis tasks and SQL queries.

ANALYSIS TASKS	SQL QUERIES
1. Find the location details of all the customers who belong to Canberra. Output the Customer details.	SELECT CUST_ID, First_Name, Last_Name, Phone, Email FROM CUSTOMERS NATURAL JOIN LOCATION WHERE LOCATION_NAME like 'Canberra';
2. Find the Sales information of the list of the items delivered to customers. Output the Item and Customer details.	SELECT SALES_ORDER_ID, ITEM_KEY, ITEM_NAME, TIME_ID, CUST_ID, First_Name, Last_Name FROM SalesInfoTable NATURAL JOIN CUSTOMERS NATURAL JOIN ITEM;
3. Find and output the Time and Item details delivered to customer named Tyson Ashley.	SELECT First_Name, Last_Name, ITEM_KEY, ITEM_NAME, BRAND, DAY, MONTH, YEAR FROM CUSTOMERS NATURAL JOIN SalesInfoTable NATURAL JOIN ITEM NATURAL JOIN TIME WHERE CUSTOMERS.First_Name LIKE 'Tyson' AND CUSTOMERS.Last_Name LIKE 'Ashley';
4. Find the list of the Customers who bought the maximum Quantity of items from the company. Output the customer and Item details.	SELECT Cust_ID, First_Name, Last_Name, ITEM_KEY, ITEM_NAME, BRAND, Quantity_Sold FROM CUSTOMERS NATURAL JOIN SalesInfoTable NATURAL JOIN ITEM WHERE QUANTITY_SOLD IN (SELECT max(Quantity_sold) FROM SalesInfoTable);
5. Find and output the number of items delivered to location Melbourne in the year 2020.	SELECT count(ITEM_KEY) AS Number_of_ITEMS_Delivered FROM LOCATION NATURAL JOIN CUSTOMERS NATURAL JOIN SalesInfoTable

	NATURAL JOIN ITEM NATURAL JOIN TIME WHERE location_name LIKE 'Melbourne' AND Year = 2020;
6. Find and output the Branch details which has made least amount in the year 2018	SELECT year, Branch_ID, min(Total) AS Least_Amount FROM (SELECT year, BRANCH_ID, sum(AMOUNT_SOLD) AS Total FROM salesInfoTable NATURAL JOIN TIME GROUP BY year) NATURAL JOIN Time WHERE year = 2018;
7. Find and arrange the Items in decreasing order of Quantity sold.	SELECT DISTINCT ITEM_KEY, ITEM_NAME, BRAND, Quantity FROM ITEM NATURAL JOIN (SELECT sum(QUANTITY_SOLD) AS Quantity, ITEM_KEY FROM SalesInfoTable GROUP BY ITEM_key) GROUP BY ITEM_KEY ORDER BY Quantity DESC;
8. Find and output the item details which are not bought by Shivakumar Kholi using Nested method.	SELECT item_key, ITEM_NAME, BRAND FROM ITEM WHERE ITEM_KEY NOT IN (SELECT ITEM_KEY FROM SalesInfoTable WHERE Sales_Order_ID IN (SELECT Sales_Order_ID FROM SalesInfoTable WHERE CUST_ID IN (SELECT CUST_ID FROM CUSTOMERS WHERE First_Name LIKE 'Shivakumar' AND Last_Name LIKE 'Kholi')));
9. Find and output all the branches where the items have atleast 1 customer who belong to state Victoria.	SELECT * FROM BRANCH WHERE BRANCH_ID IN (SELECT DISTINCT Branch_ID FROM SalesInfoTable WHERE CUST_ID IN (SELECT CUST_ID FROM CUSTOMERS NATURAL JOIN LOCATION NATURAL JOIN

	SalesInfoTable WHERE STATE LIKE 'Victoria' GROUP BY CUST_ID HAVING count(*) >= 1));
10. Find the customer from the state victoria who has made the maximum Amount of sale for the branch name Carlton in the year 2020. Output the details Customer name,Item details in the ascending order.	SELECT CUST_ID, ITEM_KEY FROM LOCATION NATURAL JOIN CUSTOMERS NATURAL JOIN SalesInfoTable NATURAL JOIN TIME NATURAL JOIN BRANCH WHERE STATE LIKE 'Victoria' AND Year = 2020 AND BRANCH_NAME LIKE 'Carlton' AND AMOUNT_SOLD IN (SELECT max(AMOUNT_SOLD) FROM SalesInfoTable);

Section 4. Discussions.

Limitations of the Current Database and the Advancement possibility for Future work.

- The Designed Database considers the recorded data of Sales management in a company and not the operation/functionality of ordering, selling, and Delivering Items to Items to the Customers which can be used in Real-time. The Database can be improved by adding the operations such as Selling and Ordering. For example, Weak links have to be added to ensure that customer is required for a sale to happen.

- The Database does not consider the financial aspect of a Sale such as Mode of payment. It can be improved by introducing a new Class or Entity called Payment and include the payment methods.

- Inventory from where the Items are picked for Delivery to Customers. This can be done by having another entity called Inventory and this inventory must be linked to Items.

- Other advancements like Automatic update of Data, Security of the Data by giving Authorization to desired Sales executives.

- User interface can be changed by creating a new webpage using HTML & CSS and establish a connection with the SQL database. This helps the User to analyse the data effectively by clicking on the desired options on the interface.

Student ID : S3857825
Name : Karthik Kolume

Appendix. DDL and sample data

1. DDL

```
CREATE TABLE TIME(
TIME_ID INTEGER,
DAY VARCHAR(3),
MONTH VARCHAR(3),
QUARTER INTEGER,
YEAR INTEGER,
PRIMARY KEY (TIME_ID));

CREATE TABLE ITEM(
ITEM_KEY VARCHAR(10),
ITEM_NAME VARCHAR(40),
BRAND VARCHAR(20),
SOLD_BY VARCHAR(30),
ITEM_GROUP CHAR,
PRIMARY KEY (ITEM_KEY));

CREATE TABLE BRANCH(
BRANCH_ID INTEGER,
BRANCH_NAME VARCHAR(40),
OWNER VARCHAR(30),
PRIMARY KEY (BRANCH_ID));

CREATE TABLE CUSTOMERS(
CUST_ID INTEGER,
First_Name VARCHAR(15),
Last_Name VARCHAR(15),
Phone INTEGER,
Email VARCHAR(30),
LOCATION_ID INTEGER,
PRIMARY KEY(CUST_ID),
FOREIGN KEY(LOCATION_ID) REFERENCES LOCATION);

CREATE TABLE LOCATION(
LOCATION_ID INTEGER,
LOCATION_NAME VARCHAR(40),
STATE VARCHAR(30),
PINCODE INTEGER,
PRIMARY KEY (LOCATION_ID));

CREATE TABLE SalesInfoTable(
SALES_ORDER_ID INTEGER,
ITEM_KEY VARCHAR(10),
TIME_ID INTEGER,
BRANCH_ID INTEGER,
CUST_ID INTEGER,
QUANTITY_SOLD INTEGER,
AMOUNT_SOLD INTEGER,
PRIMARY KEY (SALES_ORDER_ID),
FOREIGN KEY(CUST_ID) REFERENCES CUSTOMERS,
FOREIGN KEY(TIME_ID) REFERENCES TIME,
FOREIGN KEY(ITEM_KEY) REFERENCES ITEM,
FOREIGN KEY(BRANCH_ID) REFERENCES BRANCH);
```

Student ID : S3857825
Name : Karthik Kolume

2. Sample Data

- **TIME**

	TIME_ID	DAY	MONTH	QUARTER	YEAR
1	120	FRI	SEP	2	2020
2	330	MON	FEB	1	2018
3	445	THU	MAY	2	2019
4	455	FRI	FEB	1	2021
5	545	MON	MAR	2	2021
6	730	TUE	AUG	1	2020
7	756	MON	FEB	1	2018
8	824	THU	MAY	2	2019
9	930	TUE	AUG	1	2020
10	1010	WED	JAN	1	2021

- **ITEM**

	ITEM_KEY	ITEM_NAME	BRAND	SOLD _BY	ITEM_ GROUP
1	LTUB	Laundry Tub	Tupperware	Karthik	A
2	STAND	Bicycle stand	Decathlon	Kruthika	A
3	GABBY	Football player statue	Star	Shivakumar	D
4	STANK	Septic tank	Fixit	Manjula	B
5	GNOME	Garden Gnome	Pepperfry	Tyson	C
6	LAPTOP	Apple Macbook	Apple	Chris	D
7	TVLED	LED Televison	Samsung	Jack	B
8	IBOX	Iron Box	Philips	Ashley	C
9	TBLGT	Tube light	Rexon	Downey	A
10	OCHAIR	Office Chair	Feather	Hoping	D

- **BRANCH**

	BRANCH_ID	BRANCH_NAME	OWNER
1	10061	Carlton	JACK
2	10062	Docklands	JACK
3	10063	Docklands	EMMA
4	10067	Carlton	JACK
5	10075	Docklands	EMMA
6	12371	Kensington	CHRIS
7	12372	Kensington	EMMA
8	13276	Docklands	JOHN
9	23759	Kensington	CHRIS
10	23760	Carlton	CHRIS

Student ID : S3857825
Name : Karthik Kolume

- **LOCATION**

	LOCATION_ID	LOCATION_NAME	STATE	PINCODE
1	1603	Canberra	Australian Capital Territory	2601
2	1621	Melbourne	Victoria	3005
3	1631	Melbourne	Victoria	3005
4	1722	Sydney	New South Wales	2016
5	1723	Sydney	New South Wales	2016
6	1850	Brisbane	Queensland	4012
7	3468	Darwin	Northern Territory	176310
8	3469	Brisbane	Queensland	4012
9	5876	Brisbane	Queensland	4012
10	5879	Darwin	Northern Territory	176310

- **CUSTOMERS**

	CUST_ID	First_Name	Last_Name	Phone	Email	LOCATION_ID
1	1001	Karthik	Kolume	2775696	ksk@gmail.com	1722
2	1002	Manjula	SK	3987689	manj@gmail.com	3468
3	1003	Satish	Shanbogh	2679056	sat@gmail.com	5876
4	1004	Tyson	Ashley	2568909	tyson.a@outlook.com	8710
5	1005	Kruthika	Reddy	2776464	kruthika_reddy@outlook.com	1631
6	1006	Adithya	Kumar	8765489	adi.kumar12@gmail.com	8710
7	1007	Avinash	Honey	2768953	avi_honey@gmail.com	5879
8	1008	Shivakumar	Kholi	2367098	shivu.k@outlook.com	3469
9	1009	Shanmu	Dhawan	7683653	shanmu_d@gmail.com	1603
10	1010	Nag	Sharma	4780943	nag.sharma321@outlook.com	1850

- **SalesInfoTable**

	SALES_ORDER_ID	ITEM_KEY	TIME_ID	BRANCH_ID	CUST_ID	QUANTITY_SOLD	AMOUNT_SOLD
1	1449	GABBY	1010	10063	1007	15	1776
2	1621	LTUB	330	10063	1008	171	4571
3	1777	STAND	120	10061	1009	624	12016
4	2653	GABBY	1010	10063	1003	15	1776
5	2991	STAND	120	10061	1006	624	12016
6	3452	GABBY	1010	10063	1006	15	1776
7	5431	STANK	445	12371	1004	75	2365
8	5750	STANK	445	12371	1007	75	2365
9	6297	LTUB	330	10063	1005	171	4571
10	7371	STANK	445	12371	1008	75	2365