

IE6700 19106 Data Management for Analytics SEC 04 Fall 2022  
Project Group 28

(Siddhartha Setty, Shrunali Salian)

## Use Case: Milestone #3

- List each query, describe their analytic purpose
  - Include actual well-formatted query and limited output(10rows)
  - In a recent news it was said that there's swing flu in Colorado hence all meat products especially chicken have been contaminated, hence the supermarket has decided to recall all the chicken products sold to customers that came from Warehouse in Colorado. Retrieve the list of all the customers the supermarket should contact.

```
5
6 • select user_ID, user_name, user_PhoneNo, user_email
7   from DMA_Project_1.Registered_User2
8 ⊕ where user_ID in(
9           select user_ID from DMA_Project_1.Warehouse_Order as WO inner join DMA_Project_1.Final_Order as FO
10          where WO.finorder_ID = FO.finorder_ID and
11            WO.WH_ID in (select WH_ID from DMA_Project_1.Warehouse where WH_Name like "%colorado%") and
12            FO.finorder_ID in ( select finorder_ID from DMA_Project_1.Order_Product as OP inner join DMA_Project_1.Product as P
13            where OP.prod_ID = P.prod_ID and P.prod_name like "%chicken%"));
14
```

- Retrieve average age of employee based on their job type

```
57 • select emp_jobtype, avg(age) from
58     (SELECT *, DATE_FORMAT(FROM_DAYS(DATEDIFF(NOW(), emp_DoB)), '%Y') + 0 AS age
59      FROM DMA_Project_1.Employee e1) as t1
60  group by emp_Jobtype;
61
62
63
```

100% 22:60 |

Result Grid Filter Rows: Search Export:  

emp_Jobtype	avg(age)
Warehouse Worker	36.76923076923077
Manager	47.63636363636363
Delivery Associate	37.3
Retail Store Worker	37.375

Result Grid Form Editor

- Registered user Samantha received her order however, there was one item missing in the order. She informed the customer care about the same. Retrieve the contact details of the warehouse the customer care should contact in order to check about Samantha's order.

```

20 •   select * from DMA_Project_1.Warehouse
21   ○ where WH_ID in(
22     ○   select WH_ID
23       from DMA_Project_1.Delivery_Associate
24     where finorder_ID in (
25       ○   select finorder_ID
26         from DMA_Project_1.Final_Order
27       where user_ID in (
28         ○   select user_ID
29           from DMA_Project_1.Registered_User2
30         where user_name like '%samantha%')));
31

```

Result Grid

WH_ID	WH_Name	WH_Phone	WH_address
33-7488679	Rio Grande	461-656-7003	Nevada

Form

- Number of people who are choosing for 'pick up' or 'delivery' depending on the number of items they purchased.

```

66 •   SELECT
67     finorder_numofitems,
68     COUNT(IF(finorder_type = 'Standard', 1, NULL)) Delivery,
69     COUNT(IF(finorder_type = 'PickUp', 1, NULL)) Pickup,
70     COUNT(IF(finorder_type = 'Standard', 1, NULL))/COUNT(IF(finorder_type = 'PickUp', 1, NULL)) as ratio
71   FROM
72     DMA_Project_1.Final_Order
73   group by finorder_numofitems
74   order by finorder_numofitems;
75

```

Result Grid

finorder_numofitems	Delivery	Pickup	ratio
1	5	3	1.6667
2	4	7	0.5714
3	4	7	0.5714
4	2	5	0.4000
5	3	17	0.1765
6	4	2	2.0000
7	6	8	0.7500
8	5	7	0.7143
9	6	4	1.5000
10	5	6	0.8333

Form Editor

Field Types

- The management wants to look at the inventory to evaluate the value of each product and product category to allocate proper inventory space. They ask us to calculate the worth of the inventory for each product as well as each product category.

```

36
37 • select
38     prod_type as "Product Type",
39     prod_name as "Product Name",
40     prod_price*prod_quantity as "Inventory Value"
41     from DMA_Project_1.Product
42     order by prod_price*prod_quantity desc;
43
44 • select
100% 40:42

Result Grid  Filter Rows: Search Export: Result Grid
Result 7  Read Only

```

Product Type	Product Name	Inventory Value
Womens Fashion	Lilac Pants	9000000
Womens Fashion	Neon Pants	9000000
Womens Fashion	Yellow Pants	2700000
Mens Fashion	Yellow Shirt	2700000
Womens Fashion	Pink Dress	2700000
Kids	Hamleys toy	1200000
Womens Fashion	Yellow Pants	900000
Kids	Story Book	809000
Kids	Story Book	809000
Kids	Boys shirt	328000
Kids	Barbie Toy	320000
Mens Fashion	Blue shoe	306000
Womens Fashion	Blue Dress	270000

```

40     prod_price*prod_quantity as "Inventory Value"
41     from DMA_Project_1.Product
42     order by prod_price*prod_quantity desc;
43
44 • select
45     prod_type as "Product Category",
46     sum(prod_quantity) as "Total Quantity",
47     sum(prod_quantity) * sum(prod_price) as "Total Value"
48     from DMA_Project_1.Product
49     group by prod_type
50     order by sum(prod_quantity) * sum(prod_price) desc;
51
52
53
100% 52:50

Result Grid  Filter Rows: Search Export: Result Grid
Result 7  Read Only

```

Product Categ...	Total Quantit...	Total Value
Womens Fashion	12900	216720000
Kids	37180	57257200
Deli(seafood)	28700	10619000
Mens Fashion	5400	9558000
Grocery	24400	4806800
Deli(meat)	7100	781000
Electronics	88	160184.63999999998

- Retrieve the average amount spent by a customer at the supermarket.

```

79 • Select
80   Ⓛ case
81     when year(user_DoB) between 1945 and 1964 then "Baby Boom Generation"
82     when year(user_DoB) between 1965 and 1980 then "Gen X"
83     when year(user_DoB) between 1981 and 1996 then "Millennial"
84     when year(user_DoB) between 1997 and 2010 then "Gen Z"
85     else "Silent Generation"
86
87   end as user_DoB , Avg(F0.finorder_totalprice) as Average_Money_Spent_Generation_wise
88   From DMA_Project_1.Registered_User2 as RU inner join DMA_Project_1.Final_Order_Data as F0
89   on F0.user_ID = RU.user_ID
90   Group by
91   Ⓛ case
92     when year(user_DoB) between 1945 and 1964 then "Baby Boom Generation"
93     when year(user_DoB) between 1965 and 1980 then "Gen X"
94     when year(user_DoB) between 1981 and 1996 then "Millennial"
95     when year(user_DoB) between 1997 and 2010 then "Gen Z"
96     else "Silent Generation"
97   end
98 ;

```

Result Grid

user_DoB	Average_Money_Spent_Generation_wise
Gen X	333.8181818181818
Gen Z	323.2857142857143
Millennial	304

- Retrieve the list of products that will expire in the next 7 days

```

103 • select * from DMA_Project_1.Product where prod_expiry_date < date_add(now(), interval 7 day);
104

```

Result Grid

prod_ID	prod_type	prod_name	prod_price	prod_Expiry_date	prod_quant...
193054321	Deli(seafood)	Salmon	20	2022-11-27	3000
523054321	Grocery	Yogurt	5	2022-11-27	1000
623054321	Grocery	Watermelon	5	2022-11-24	1000
666666321	Grocery	Spinach	5	2022-11-23	1000
812304321	Deli(meat)	Chicken Breasts	10	2022-11-25	1000
913334321	Grocery	Spinach	5	2022-11-21	1000
917054006	Grocery	Spinach	15	2022-11-21	100

- Retrieve names of customers and their credit card information encrypted with only the last 4 digits un-encrypted as a matter of privacy.

```

100
101 Ⓛ Select ru.user_name, concat(repeat('*',length(rup.card_no)-4),
102   Substring(rup.card_no,-4)) as CreditCardNumber
103   From DMA_Project_1.Registered_User2 as ru
104   Inner join DMA_Project_1.Registered_User_Payment as rup using (user_ID);
105
106
107
108
109
110
111
112
113
114
115
116

```

Result Grid

user_name	CreditCardNumb...
Ynes Polland	*****3065
Abran Chatenet	*****3890
Reggi Sandever	*****3870
Markus Ambresin	*****3001
Hugo Gerard	*****3002
Ammamaria Bagot	*****3003
Cleon Rawcliffe	*****3005

\*\*\*\*\*THE END\*\*\*\*\*

# IE6700 19106 Data Management for Analytics SEC 04 Fall 2022

## Project Group 28

(Siddhartha Setty, Shrunali Salian)

### Milestone #1

*(Reference Data and Transactional Data included)*

#### Milestone Report #1:

- Real world business problem definition
- Must include 7-9 business entities and their relationships
- A conceptual data model using EER
- Must include both reference and transactional data

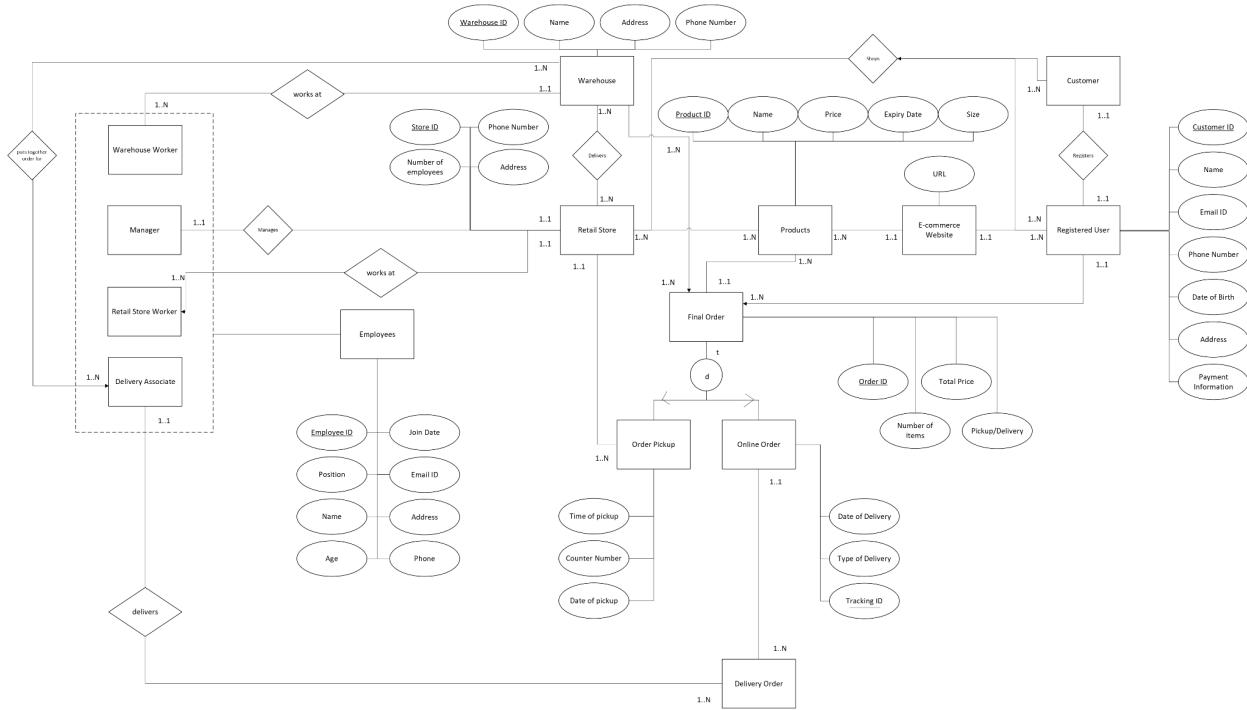
#### Problem Definition:

The business problem that is being implemented here is for a database system of a retail and e-commerce store. The database will include information about the Warehouse, Retail Store, the employees that are part of the entire operation. This database will also contain information about the e-commerce website, the products that are available at the retail and the e-commerce website, the final order and will contain information regarding the type of order the customer has, whether it is an online order or an in person pick up of their order. The aim of this project is to have a functional database that can maintain information regarding the products being offered by the Retail store and the e-commerce website, it can also keep track of the order, have information regarding the different employees involved in the operation and contain the information about the registered users who use the services of the company.

#### Entities:

1. Warehouse: This contains information regarding the details about the warehouse, its ID, name, address and phone number.
2. Retail store: This contains information regarding the number of employees working at the store, and the details of the store such as address, its name and number.
3. Products: This entity will have all the information related to the products that are being offered by the retail store and the e-commerce website, such as the product ID, the name, price, size, and expiry date of the products.
4. Registered User: This entity will hold all the information related to the customer who orders products online through the ecommerce website, such as their customer ID, name, age, email ID, phone number, Date of Birth, their address, and their Payment Information.
5. Employees: This is another major entity in this database, this will contain all the details regarding the employees who are working on the entire operation. These details include their employee ID, the position they work at, their date along with all their personal information.
6. Final Order: This entity contains all the information regarding a registered user's final order. This includes, order ID, total price, number of items and whether the customer wants to opt for pickup or delivery.

## Conceptual data model using EER diagram:



For a detailed view of the EER diagram please use the link below:

[https://drive.google.com/file/d/1atO1I3f73XtX--5F9soTDm3yPp0x\\_hF5/view?usp=sharing](https://drive.google.com/file/d/1atO1I3f73XtX--5F9soTDm3yPp0x_hF5/view?usp=sharing)

(Please note this EER diagram has been made using Microsoft Visio. Since Visio is not supporting arcs currently, we have indicated specialization using arrows for the entity – Final Order and its subclasses Order pickup and Online order)

### REFERENCE DATA & TRANSACTIONAL DATA

Reference data:

1. Order PickUp and Online Order both refer to Final Order for the finorder\_ID
2. Delivery Order refers to Online Order for the finorder\_ID
3. Order PickUp refers to Retail Store for the RS\_ID
4. Different types of employees i.e., warehouse worker, manager, retail store worker, delivery associate refer to employee table for emp\_ID
5. Manager refers to Retail Store for RS\_ID
6. Warehouse worker refers to Warehouse for WH\_ID
7. Final Order refers to Product table for prod\_ID

Transactional data:

1. Final Order table contains transactional data in terms of the total price for the particular order\_ID.
2. User credit card details are encrypted for privacy in the separate Registered\_User\_Payment table.
3. Retail Store contains the number of employees working in each store.
4. Product table contains the prod\_price for each product, hence we can calculate the average price of a product at the supermarket.
5. Employee table contains the emp\_DoB, thus Age of every employee can be found.
6. Warehouse\_Worker contains the WH\_WWaisles\_managed thus the total number of aisles managed by each employee working at that warehouse can be found.

*Thank You*