

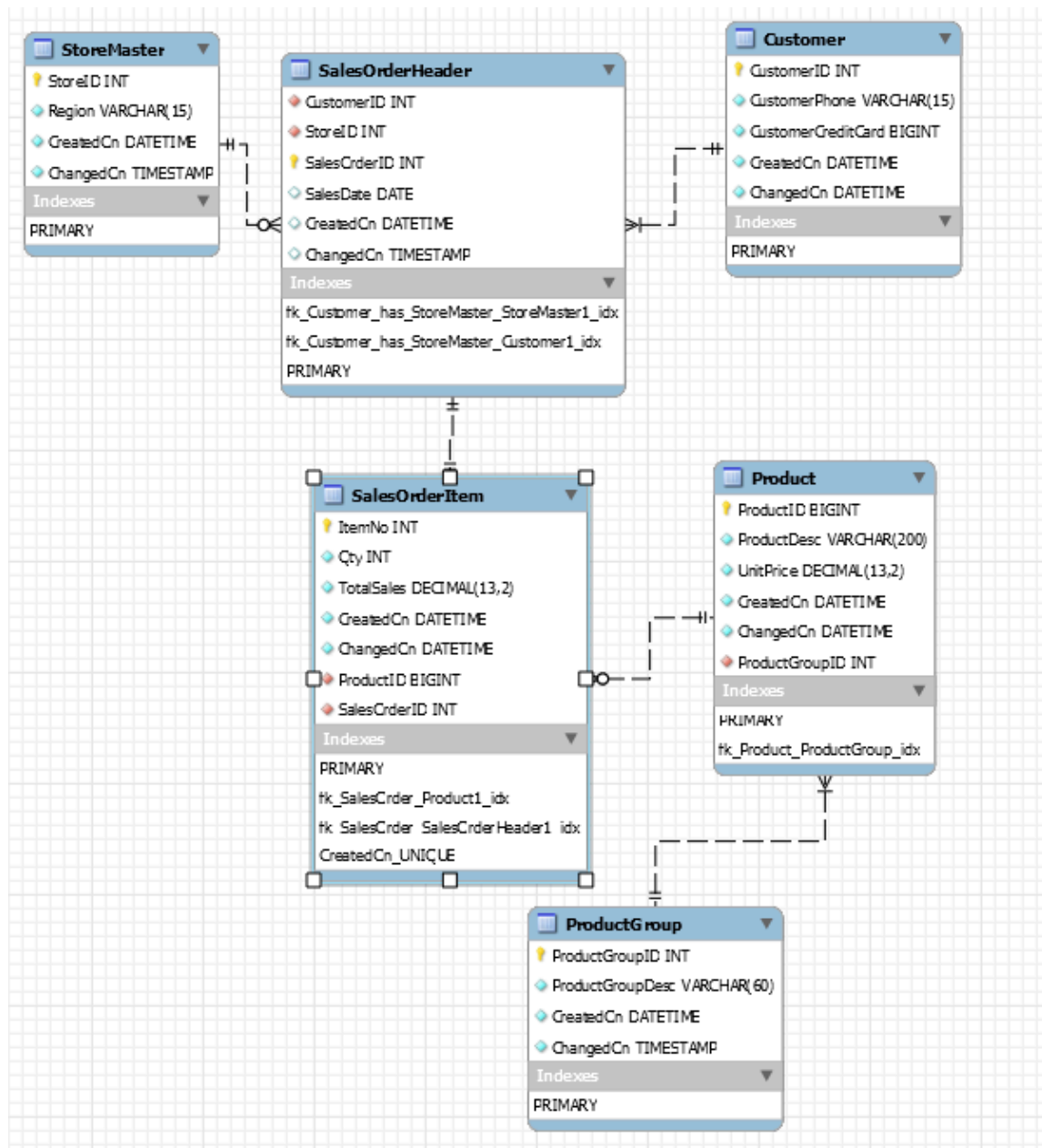
Database Design of a Convenience Store

Anber Bhuiyan

Professor Mitra

Database Management

1. EER Model



Changes from design phase:

- Fixed SalesOrderItem and added correct fields (i.e., Qty, TotalSales)
- Adjusted FKs and PKs for SalesOrderHeader and SalesOrderItem

2. Structure of Created Tables

desc group9_8eleven.customer

Field	Type	Null	Key	Default	Extra
Customer_ID	int	NO	PRI	NULL	auto_increment
CustomerPhone	varchar(15)	YES		NULL	
CustomerCreditCard	bigint	YES		NULL	
CreateDateTime	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
CreateTimeStamp	timestamp	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

desc group9_8eleven.product

Field	Type	Null	Key	Default	Extra
ProductID	bigint	NO	PRI	NULL	
ProductGrpID	int	NO	MUL	NULL	
ProductDesc	varchar(200)	YES		NULL	
UnitPrice	decimal(13,2)	YES		NULL	
CreatedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
ChangedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

desc group9_8eleven.productgroup

Field	Type	Null	Key	Default	Extra
ProductGrpID	int	NO	PRI	NULL	
ProductGrpDesc	varchar(60)	YES		NULL	
CreatedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
ChangedOn	timestamp	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

desc group9_8eleven.salesorderheader

Field	Type	Null	Key	Default	Extra
SalesOrderID	int	NO	PRI	NULL	auto_increment
Store_ID	int	NO	MUL	NULL	
SalesDate	date	YES		NULL	
Customer_ID	int	NO	MUL	NULL	
CreatedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
ChangedOn	timestamp	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

desc group9_8eleven.salesorderitem

ItemNo	int	NO	PRI	NULL	auto_increment
SalesOrderID	int	NO	MUL	NULL	
ProductID	bigint	NO	MUL	NULL	
Qty	int	YES		NULL	
TotalSale	decimal(13,2)	YES		NULL	
CreatedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
ChangedOn	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

desc group9_8eleven.storemaster

Field	Type	Null	Key	Default	Extra
Store_ID	int	NO	PRI	NULL	auto_increment
Region	varchar(15)	YES		NULL	
CreateDateTime	datetime	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED
CreateTimeStamp	timestamp	YES		CURRENT_TIMESTAMP	DEFAULT_GENERATED

3. Data from Tables

select * from group9_8eleven.customer

Customer_ID	CustomerPhone	CustomerCreditCard	CreateDateTime	CreateTimeStamp
1	(782) 355-5110	7302252569655990	2021-03-26 16:33:47	2021-03-26 16:33:47
2	(645) 230-6144	4647729499520270	2021-03-26 16:33:47	2021-03-26 16:33:47
3	(421) 031-2190	5007183667576290	2021-03-26 16:33:47	2021-03-26 16:33:47
4	(897) 326-6659	8258642766322370	2021-03-26 16:33:47	2021-03-26 16:33:47
5	(982) 923-7957	7785060118653330	2021-03-26 16:33:48	2021-03-26 16:33:48
6	(744) 946-3546	5975391689895320	2021-03-26 16:33:48	2021-03-26 16:33:48
7	(759) 720-1517	5141461049958770	2021-03-26 16:33:48	2021-03-26 16:33:48
8	(717) 661-2914	5044912916148140	2021-03-26 16:33:48	2021-03-26 16:33:48
9	(932) 811-2427	8344624340547840	2021-03-26 16:33:48	2021-03-26 16:33:48
10	(623) 254-2887	4071351627010740	2021-03-26 16:33:48	2021-03-26 16:33:48
11	(789) 140-5603	5191800939258900	2021-03-26 16:33:48	2021-03-26 16:33:48
12	(681) 474-3773	4844490406845920	2021-03-26 16:33:48	2021-03-26 16:33:48
13	(858) 112-4333	3995806099862940	2021-03-26 16:33:48	2021-03-26 16:33:48

select * from group9_8eleven.product

ProductID	ProductGrpID	ProductDesc	UnitPrice	CreatedOn	ChangedOn
794567898569	7000710	TYLENOL CHILD COLD DAY SUS 4 OZ	5.07	2021-04-12 18:55:22	2021-04-12 18:55:22
794567898570	29002910	GUD VANILLA SHAMPOO	4.64	2021-04-12 18:55:23	2021-04-12 18:55:23
794567898571	58005850	XMAS SISTER & BROTHER-IN-LAW	5.07	2021-04-12 18:55:25	2021-04-12 18:55:25
794567898572	91009180	HEALWELL AFO NIGHT WRAP LARGE / XLARGE	3.80	2021-04-12 18:55:26	2021-04-12 18:55:26
794567898573	91009180	HEALWELL AFO NIGHT WRAP LARGE / XLARGE	7.60	2021-04-12 18:55:30	2021-04-12 18:55:30
794567898574	58005850	BUNNY G/BAGS	5.49	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898575	58005850	BUNNY G/BAGS	4.64	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898576	58005850	BUNNY G/BAGS	8.44	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898577	58005850	CHRISTIAN GREETING CARDS	2.11	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898580	58005850	ALBUM KITCHEN CLASSIC REC	6.75	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898581	58005850	CHRISTMAS GIFT WRAP	3.80	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898582	15001520	ANKLE WRAP FLA 3XL BLK	8.02	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898583	15001520	ANKLE WRAP FLA 3XL BLK	1.27	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898584	7000710	ROBITUSSIN NGHT C&C&FL SYR 4 OZ	4.64	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898585	7000710	ROBITUSSIN NGHT C&C&FL SYR 4 OZ	6.75	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898586	7000710	ROBITUSSIN NGHT C&C&FL SYR 4 OZ	8.44	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898587	58005850	CARD CHRSTMS DAUGHTER	1.69	2021-04-12 18:57:39	2021-04-12 18:57:39
794567898588	58005850	CARD CHRSTMS DAUGHTER	2.53	2021-04-12 18:57:39	2021-04-12 18:57:39

select * from group9_8eleven.productgroup

ProductGrpID	ProductGrpDesc	CreatedOn	ChangedOn
1000110	PAIN RELIEF	2021-03-30 16:11:15	2021-03-30 16:11:15
1000120	EXTERNAL PAIN RELIEF	2021-03-30 16:11:15	2021-03-30 16:11:15
1000130	SLEEP AIDS, STIMULANTS & MOTION SICKNESS	2021-03-30 16:11:15	2021-03-30 16:11:15
1000140	CHILDREN'S PAIN & FEVER RELIEF	2021-03-30 16:11:15	2021-03-30 16:11:15
3000310	STOMACH REMEDIES	2021-03-30 16:11:15	2021-03-30 16:11:15
3000320	ACID CONTROLLERS	2021-03-30 16:11:15	2021-03-30 16:11:15
3000330	LAXATIVES	2021-03-30 16:11:16	2021-03-30 16:11:16
3000340	ANTI-DIARRHEALS	2021-03-30 16:11:16	2021-03-30 16:11:16
3000360	ANTI-GAS	2021-03-30 16:11:16	2021-03-30 16:11:16
3000380	CHILDREN'S DIGESTIVE HEALTH	2021-03-30 16:11:16	2021-03-30 16:11:16
5000520	BABY HEALTH, BEAUTY & WELLNESS	2021-03-30 16:11:16	2021-03-30 16:11:16
5000540	BABY FEEDING ACCESSORIES	2021-03-30 16:11:16	2021-03-30 16:11:16
7000710	COUGH, COLD, FLU & SINUS	2021-03-30 16:11:16	2021-03-30 16:11:16
7000720	COUGH DROPS, SORE THROAT RELIEF	2021-03-30 16:11:16	2021-03-30 16:11:16
7000730	ALLERGY	2021-03-30 16:11:16	2021-03-30 16:11:16
7000740	NASAL	2021-03-30 16:11:16	2021-03-30 16:11:16

select * from group9_8eleven.salesorderheader

SalesOrderID	Store_ID	SalesDate	Customer_ID	CreatedOn	ChangedOn
2857	978171	2016-03-15	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2858	978171	2016-03-15	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2859	978171	2016-03-15	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2860	978171	2016-03-15	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2861	978171	2016-03-15	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2862	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2863	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2864	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2865	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2866	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2867	978076	2018-06-01	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2868	978105	2020-02-16	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2869	978105	2020-02-16	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2870	978105	2020-02-16	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2871	978105	2020-02-16	308	2021-04-12 18:47:19	2021-04-12 18:47:19
2872	978220	2016-03-06	448	2021-04-12 18:47:19	2021-04-12 18:47:19
2873	978220	2016-03-06	448	2021-04-12 18:47:19	2021-04-12 18:47:19
2874	978220	2016-03-06	448	2021-04-12 18:47:19	2021-04-12 18:47:19

select * from group9_8eleven.storemaster

Store_ID	Region	CreateDateTime	CreateTimeStamp
978003	North	2021-03-30 16:19:41	2021-03-30 16:19:41
978004	North	2021-03-30 16:19:41	2021-03-30 16:19:41
978015	North	2021-03-30 16:19:41	2021-03-30 16:19:41
978020	North	2021-03-30 16:19:41	2021-03-30 16:19:41
978022	North	2021-03-30 16:19:41	2021-03-30 16:19:41
978023	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978031	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978034	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978037	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978039	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978043	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978044	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978047	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978048	North	2021-03-30 16:19:40	2021-03-30 16:19:40
978049	North	2021-03-30 16:19:39	2021-03-30 16:19:39
978052	North	2021-03-30 16:19:39	2021-03-30 16:19:39

select * from group9_8eleven.salesorderitem

ItemNo	SalesOrderID	ProductID	Qty	TotalSale	CreatedOn	ChangedOn
1	308	794567902501	12	30.39	2021-04-13 16:12:36	2021-04-13 16:12:36
2	308	794567903669	40	270.16	2021-04-13 16:12:36	2021-04-13 16:12:36
3	308	794567902233	6	22.79	2021-04-13 16:12:36	2021-04-13 16:12:36
4	308	794567899164	17	64.58	2021-04-13 16:12:36	2021-04-13 16:12:36
5	308	794567907614	22	167.16	2021-04-13 16:12:36	2021-04-13 16:12:36
6	308	794567902788	41	103.84	2021-04-13 16:12:36	2021-04-13 16:12:36
7	308	794567905296	11	74.29	2021-04-13 16:12:36	2021-04-13 16:12:36
8	308	794567898768	30	240.61	2021-04-13 16:12:36	2021-04-13 16:12:36
9	308	794567900796	19	104.26	2021-04-13 16:12:36	2021-04-13 16:12:36
10	308	794567907921	26	43.90	2021-04-13 16:12:36	2021-04-13 16:12:36
11	308	794567901481	6	22.79	2021-04-13 16:12:36	2021-04-13 16:12:36
12	308	794567906432	40	287.04	2021-04-13 16:12:36	2021-04-13 16:12:36
13	308	794567899789	41	121.15	2021-04-13 16:12:36	2021-04-13 16:12:36
14	308	794567907403	17	21.53	2021-04-13 16:12:36	2021-04-13 16:12:36
15	308	794567902437	44	204.31	2021-04-13 16:12:36	2021-04-13 16:12:36
16	448	794567908236	28	212.75	2021-04-13 16:12:36	2021-04-13 16:12:36
17	448	794567901799	17	107.64	2021-04-13 16:12:36	2021-04-13 16:12:36
18	448	794567899300	37	203.04	2021-04-13 16:12:36	2021-04-13 16:12:36
19	448	794567899601	38	208.53	2021-04-13 16:12:36	2021-04-13 16:12:36
20	448	794567899878	23	165.05	2021-04-13 16:12:36	2021-04-13 16:12:36
21	448	794567908356	31	196.29	2021-04-13 16:12:36	2021-04-13 16:12:36
22	448	794567900055	25	110.00	2021-04-13 16:12:36	2021-04-13 16:12:36

4. Count Entries

select count(Customer_ID) as NumberofCustomers from group9_8eleven.customer

NumberofCustomers
497

select count(ProductID) as NumberofProducts from group9_8eleven.product

NumberofProducts
8883

select count(ProductGrpID) as NumberofProductGroups from group9_8eleven.productgroup

NumberofProductGroups
117

select count(SalesOrderID) as NumberofSales from group9_8eleven.salesorderheader

NumberofSales
5993

select count(Store_ID) as NumberofStores from group9_8eleven.storemaster

NumberofStores
72

select count(ItemNo) as NumberofSalesOrderItems from group9_8eleven.SalesOrderItem

NumberofSalesOrderItems
5993

5. Data Analytics

- a. Could you recommend specific products that stores in your region should always have in stock? If so, what are the top three in descending priority?

	ProductID	count(SalesOrd... ▾	
▶	794567898660	4	
	794567899072	4	
	794567899716	4	
	794567899744	4	
	794567899973	4	
	794567898602	3	
	794567898624	3	
	794567898639	3	
	794567898657	3	
	794567898738	3	
	794567898768	3	
	794567898816	3	
	794567898872	3	

After running:

```
1 • select ProductID, count(SalesOrderID) from Group9_eleven.SalesOrderItem
2     group by ProductID;
3 • select ProductDesc from Group9_eleven.Product
4     where ProductID = 794567898660;
5 • select ProductDesc from Group9_eleven.Product
6     where ProductID = 794567899072;
7 • select ProductDesc from Group9_eleven.Product
8     where ProductID = 794567899716;
```

We recommend stores in the North region should always have:

3. Activa Stockings

2. Rezamid Acne Treatment Lotion 20oz

1. Bag Paper WHT 4.5"x2.25"x11" 3000CT G-S

These products had the most sales orders.

- b. Could you recommend specific products that stores in your region should consider not carrying? If so, what are the top three in descending priority that should be eliminated?

ProductID	count(SalesOrd...
794567898572	1
794567898573	1
794567898576	1
794567898581	1
794567898582	1
794567898588	1
794567898589	1
794567898590	1
794567898591	1
794567898596	1

After running:

```
1 • select ProductID, count(SalesOrderID) from Group9_8eleven.SalesOrderItem
2     group by ProductID;
3 • select ProductDesc from Group9_8eleven.Product
4     where ProductID = 794567898572;
5 • select ProductDesc from Group9_8eleven.Product
6     where ProductID = 794567898576;
7 • select ProductDesc from Group9_8eleven.Product
8     where ProductID = 794567898581;
```

We recommend stores in the North region should consider not carrying:

3. Healwell AFO Night Wrap Large/XLarge

2. Bunny G/Bags

1. Christmas Gift Wrap

These products had the least sales orders (only purchased in 1 sales order).

- c. Think of three more business critical decisions that you would recommend to the owners from your group. Write the MySQL query and explain your group's analytical reasoning.
- i. What items are most popular (qty wise) during Black Friday sales in 2019? We recommend 8/11 should run promotions and keep much supply for these items for Black Friday in order to meet demand.
1. LBL SLCN BRACELET NOT ASHAMED
 2. HLMK ANNIV CARD 4.49
 3. ULTRA FIT WRIST BRACE RIGHT

```

4  #Most popular during Black Friday week 2019
5  • SELECT ProductID, Qty
6      from SalesOrderItem as I inner join SalesOrderHeader as H
7      on H.SalesOrderID = I.SalesOrderID
8      where SalesDate >= '2019-11-23' and #Using Operators over Between gives you more control
9      '2019-11-30' <= SalesDate;
10 • select ProductDesc from Group9_8eleven.Product
11     where ProductID = 794567904287;
12 • select ProductDesc from Group9_8eleven.Product
13     where ProductID = 794567904421;
14 • select ProductDesc from Group9_8eleven.Product
15     where ProductID = 794567904885;

```

ProductID	Qty
794567904287	45
794567905387	45
794567904421	45
794567904885	45
794567904824	45
794567899986	45
794567902185	45
794567901448	45
794567908308	45
794567905108	45

- ii. What are the most popular Product Groups? We recommend 8/11 should focus on keeping items from these groups in stock. 8/11 may also consider expanding the selection in these groups.

1. ALL OTHER GREETING CARD LINES
2. BRACES
3. MEDICAL SUPPORT - HIGH COMPRESSION (OVER 8-15 MMHG)

```
7      #Most popular product groups
8      SELECT PG.ProductGrpID, count(I.SalesOrderID)
9          from SalesOrderItem as I inner join Product as P
10             on I.ProductID = P.ProductID
11             inner join ProductGroup as PG
12             on PG.ProductGrpID = P.ProductGrpID
13             group by PG.ProductGrpID;
14      select * from Group9_eleven.ProductGroup
15          where ProductGrpID = 58005850;
16      select * from Group9_eleven.ProductGroup
17          where ProductGrpID = 15001520;
18      select * from Group9_eleven.ProductGroup
19          where ProductGrpID = 16001620;
```

ProductGrpID	count(I.SalesOrd... ▾
58005850	3545
15001520	505
16001620	350
15001560	124
91009180	59
90009050	58
29002920	58
1000110	55
79007900	55
7000710	52
60006010	50

iii. Who were the most frequent customers in 2019? What were their phone numbers? 8/11 should consider sending promotional information to these customers or rewarding their frequent shopping.

1. (798)014-7456
2. (603)574-0934
3. (692)803-3764

```
1  #Most frequent shopper during 2019
2  • SELECT H.Customer_ID, count(I.SalesOrderID)
3      from SalesOrderItem as I inner join SalesOrderHeader as H
4          on H.SalesOrderID = I.SalesOrderID
5          where SalesDate >= '2019-01-01' and #Using Operators over Between gives you more control
6              '2019-12-31' <= SalesDate
7          group by H.Customer_ID;
8  • select CustomerPhone from Group9_eleven.Customer
9      where Customer_ID = 92;
10 • select CustomerPhone from Group9_eleven.Customer
11     where Customer_ID = 227;
12 • select CustomerPhone from Group9_eleven.Customer
13     where Customer_ID = 308;
```

Customer_ID	count(I.SalesOrderl...
92	132
227	109
308	94
343	93
172	59

6. Project Summary

a) Timeline of your development process (actual key dates)

- Create and upload the first three tables (Customer, ProductGroup, StoreMaster) - March 26th
- Create the next three tables (SalesOrderHeader, Product, SalesOrderItem) - April 5th
- Load data for SalesOrderHeader, Product, SalesOrderItem - April 5th-10th
- Finishing data analytics and cleaning of Project Report- April 12th

b) Were there any indications that the Excel Data received from the Franchise was dirty?

Importing our data into our tables gave us indications that the Excel data was dirty. The data contained multiple duplicate entries in almost every data table. There were also certain entries that were inconsistent and would cause an error in load; for example, when loading ProductDesc, some entries had very large pieces of data, which is why varchar had to be adjusted to allow 200 characters.

c) How you went about loading just the sales data for your region.

The sequence of creating/loading tables was the first important step. The following is the sequence of creation/load for tables related to sales data:

1. StoreMaster
2. ProductGroup
3. SalesOrderHeader
4. Product
5. SalesOrderItem

Steps in loading sales data (after loading StoreMaster, ProductGroup, Product):

1. Customer table was exported (CustomerID is a FK)
2. StoreMaster table was exported (StoreID is a FK)
3. Sales data from Master data was put in sheet
4. Added column for CustomerID using Vlookup on sorted ascending Customer Phone Number, used Customer export
5. Added column for StoreID using Vlookup on sorted ascending, used StoreMaster export
6. SalesOrderHeader data cleaned to match table structure then imported - now SalesOrderHeader holds only North sales
7. Data (i.e, Qty, TotalSale) needed for SalesOrderItem from Master data was put in sheet
8. Added column for SalesOrderID using Vlookup on sorted ascending Sales Date, used SalesOrderHeader export
9. SalesOrderItem cleaned to match table structure then imported

d) The top two project takeaways for your team about an industrial strength database design process.

1. There is great importance in having a comprehensive EER model, normalizing the data, and cleansing the data. These processes can help prevent any update, deletion, and insertion anomalies, especially due to the high volume of data. Upgrading the quality of the raw data by cleansing it helps us understand data relationships and uncover potential quality issues to comply with business requirements. Overall, these processes minimize the duplication of data, decrease modification issues, and simplify queries.

2. Another important factor when running an industrial strength database design is proper scheduling and time management. Deadlines are important to minimize time spent and overall cost on a project, and sticking to timelines and establishing when key steps in the process should be done allows for the team to troubleshoot. Sometimes, the sheer amount of data can be a burden on systems as well as hardware. Giving a buffer period during the project to wait for data to load is important. A team cannot assume a smooth process and wait until the deadline approaches to initiate data loads. The firm needs to consider this to minimize unexpected costs and delays during the project.