MSA 8110 Final Project

Yan Xu

Sam's Club Sales Assessment

March 10, 2018

Executive Summary

Per request from Sam's Club president, a detailed sales assessment has been performed. This assessment analyzed both performance of Sam's Club stores and membership purchasing behaviors from three tables (Member Index, Store Visits, Store Information) in Sam's Club database. Data integrity has been thoroughly assessed before the actual analysis to ensure the data quality. Results will be presented in table format; any related SQL statements will be included in the report appendix.

Part 1. Data Integrity Assessment

Entity Integrity: based on the entity integrity assessment results, there are no entity integrity issues for Member Index, Store Visits, and Store Information tables.

Referential Integrity: based on the referential integrity assessment results, no issue was found between Store Visits and Store Information tables. However, the assessment for referential integrity between Store Visits and Member Index indicated, there is an issue between those two tables. Not all the Membership nbr in Store Visits (FK table) matches an existing value in the Member Index table (Primary Table).

Fix the referential integrity issue: To avoid any confusion with the actual Null values in the Store Visits table, a dummy value was created in the primary table – Member Index via SQL statement in Appendix A. Then, the FK in the Store Visits was matched to the dummy value via SQL statement in Appendix A.

Part 2. Sam's Club Store Performance Assessment

Overall, the total sales amount is $79,797,751.11 during the month of January 2015 based on the assumption that Total Visit Amt is the actual amount every store received from customer included any applicable tax due to limited information provided. All associated SQL statements are included in Appendix B.

Table 1 is a summary of the top 5 performing stores, their corresponding region and total store sales during January 2015. Store #18 has the best performing of all the stores, the total store sales is $6,549,649.42. The store belongs to region # 23. Noted that region # 23 has two top performing stores in the same period.

Table 1. Top 5 Performing Stores Sales Summary

Store Number	Region Number	Total Store Sales
18	23	$ 6,549,649.42
28	6	$ 5,784,964.62
19	22	$ 5,605,419.32
15	23	$ 5,368,442.95
27	14	$ 5,302,792.14

Table 2 is a summary of the bottom 5 performing stores, their corresponding region and total store sales during January 2015. Store # 3 is the at bottom of the sales performance of all the stores, the total store sales is $2,843,630.47. It is roughly half of the best performing store's number.

Table 2. Bottom 5 Performing Store Sales Summary

Store Number	Region Number	Total Store Sales
3	44	$ 2,843,630.47
25	2	$ 3,075,330.28
4	50	$ 3,272,714.44
23	38	$ 3,376,051.51
17	38	$ 3,792,796.00

Table 3 is a summary of the regional sales. As region#23 has two of the top performing stores, it is also ranked #1 among all the regions. Some regions, such as #38 ranked high in the regional sales because of the # of stores in the region. But at the same time, this region has two of the worst performing stores.

Table 3. Regional Sales Summary

Region Number	Total Sales By Region
23	$ 15,913,900.80
32	$ 8,721,518.46
38	$ 7,168,847.51
6	$ 5,784,964.62
22	$ 5,605,419.32
14	$ 5,302,792.14
24	$ 5,265,693.39
49	$ 4,824,079.39
33	$ 4,248,969.39
5	$ 3,923,840.38
19	$ 3,846,050.52
50	$ 3,272,714.44
2	$ 3,075,330.28
44	$ 2,843,630.47

Table 4 summarized the sales on weekdays. Saturday, Sunday, and Friday are the top 3 revenue generating days as people usually do their shopping on weekend. The high number from Monday possibly indicated a national holiday during the month that also attracted many shoppers. As table showed, the total revenue from Saturday is double the amount from Tuesday.

Table 4. Weekday Sales Summary

Day of the week	Total Weekday Sales
Saturday	$ 17,706,294.32
Sunday	$ 12,634,820.27
Friday	$ 11,784,074.24
Monday	$ 11,518,853.13
Thursday	$ 9,262,431.10
Wednesday	$ 8,478,509.44
Tuesday	$ 8,412,768.61

Table 5 summarized the sales by membership type. Without further information provided, we can conclude that customer with V, W and A membership type are the top 3 group of people generating revenues. Those 3 types of members generated 88% of the store revenue.

Table 5. Sales Summary by Membership Type

Membership Type	Total Sales By Membership
V	$ 29,990,444.56
W	$ 22,936,850.65
A	$ 17,458,238.42
X	$ 9,071,849.81
E	$ 222,714.53
D	$ 56,599.15
Y	$ 14,781.15
G	$ 13,486.64
1	$ 11,344.31
3	$ 10,517.25
Z	$ 8,971.51
H	$ 1,953.13

Also, additional sales analysis based on customer payment type was summarized in Table 6. The two most common payment methods customers used are Check and Cash. Top 3 payment methods accounted for 88% of the transaction amount in the store.

Table 6. Sales Summary by Payment Method

Tender Type	Total Sales By Tender Type
1	$ 37,828,429.08
0	$ 24,544,361.17
3	$ 8,178,787.34
6	$ 4,566,349.39
5	$ 2,371,402.59
9	$ 1,510,160.18
8	$ 433,305.69
4	$ 364,779.69
7	$ 175.98

Part 3. Membership Purchasing Behavior Assessment

Membership purchasing behavior was another important aspect of performance assessment. An overall review of the customer purchasing will be analyzed. Followed with some more detailed look. All related SQL statement will be presented in Appendix C.

As summarized in Table 7, customer with membership #577535669 has the highest average spending and number of items per transaction. However, this customer does not belong to the Top 3 type of membership per Table 5 indicated. Both customer with member type X and W are among the top customers in average spending and items bought.

Table 7. Top 10 Average Spending Customer

Membership Number	Member Type	Average Spending	Average Number of items
577535669	E	$ 96,631.80	3240
382713918	X	$ 35,828.98	1200
390037978	X	$ 23,869.02	2184
574059584	W	$ 23,115.00	2640
474880590	X	$ 17,342.20	880
687440025	X	$ 12,449.85	605
630742214	X	$ 10,774.58	514
23300953	W	$ 8,610.29	1480
764110817	W	$ 7,478.69	946
15332364	X	$ 5,450.60	199

Like Table 4, Saturday and Sunday are still the top 2 days even in total number of customer visits per Table 8 showed. Monday attracted more customers, but with a slightly lower revenue than Friday possibly due to holiday On Sale price. Saturday again attracted most of the customer visits during a week, double the number on Wednesday.

Table 8. Customer Visit by Weekday

Day of the week	Total Visit
Saturday	217515
Sunday	159564
Monday	150659
Friday	144136
Thursday	117567
Tuesday	110580
Wednesday	107940

As Saturday is the most popular day for shopping, the results in Table 9 showed that customer likes to shop during the lunch hours. 12:00 pm – 14:00 pm are traditional lunch time for most people. Customers are less likely to shop during later night or early morning due to the factor that most people will be resting during that period.

Table 9. Visit by Hour of the Day

Hour of the day	Total Visit
13	117201
14	115165
12	113077
15	108966
16	101742
11	99686
17	86879
10	72278
18	69515
19	59188
9	26231
20	26024
8	8158
7	3003
6	455
5	168
21	165
4	56
2	3
	1

From the analysis of elite status data, it seems that customers with Elite Status 0 are the most active group people. They contributed 98% of the visits among all the elite status members.

Table 10.Visit by Elite Status

Elite Status Code	Visit by Elite status
0	990027
1	1327
2	10028
3	2520
4	4059

Top 3 customers who are most active in both sales and visits are summarized in Table 11 with some typical characteristics. Even though they don't share the exact characteristics, but they have something in common. They all belong to either Member Type V or W which is consistent with the results showed in Table 5 that Type V and W are the top 2 types in sales. Their Elite Status Code is 0 which is consistent with what Table 10 showed.

Table 11.Top 3 Most Active Customer in Sales and Visits

MEMBERSHIP NUMBER	ZIP CODE	MEMBER TYPE	ELITE STAT CODE
709993795	74910	V	0
377467502	56203	V	0
748269885	12887	W	0

Part 4. Conclusion and Recommendation

Overall store performance varies across the regions. Store with best performance can generate double the revenue than the bottom performers. Weekends are the peak revenue generating days which ties to the factor that people usually spend more time on shopping during weekend. Customers with member ship type V, W and A generated most of the sales. When customer shop, they tender to use Cash or Check as the main payment methods. Another interesting finding is customer likes to shop during lunch hours.

For bottom performing stores, more promotions or new advertising campaign might help to improve the store performance. Put more emphasis on the weekend Sales events to further increase the revenues. Promote Member Type V or W to more customers.

Appendix A

Data Integrity Assessment SQL statement

```sql
-- Assess the entity integrity for MEMBER INDEX table
select *
from member_index
where membership_nbr is null;

select membership_nbr, count(*)
from member_index
group by membership_nbr
having count(*) > 1;


-- Assess the entity integrity for STORE INFORMATION table
select *
from STORE_INFORMATION
where store_Nbr is null;

select store_nbr, count(*)
from store_information
group by store_nbr
having count(*) > 1;

-- Assess the entity integrity for STORE VISITS table
select *
from store_visits
where visit_Nbr is null;

select visit_Nbr, count(*)
from store_visits
group by visit_Nbr
having count(*) > 1;


-- Assess the referential integrity between Store Visits and Member Index
select membership_nbr
from store_visits
where membership_nbr not in (select membership_nbr from member_index);


-- Assess the referential integrity between Store Visits and Store Information
select store_nbr
from store_visits
where store_nbr not in (select store_nbr from store_information);

-- Fix the referential integrity issues between Store Visits and Member Index
-- Create a Dummy Variable in the primary table
insert into MEMBER_INDEX values(9999, 'unkno', 'unknow', 'A','A', 0, '2001-10-13
00:00:00.000', '2001-10-13 00:00:00.000', 0, 0, 0, 'unkw', 0, '2001-10-13 00:00:00.000')

-- Updated the unmatched FK values to dummy values.
update STORE_VISITS
set Membership_Nbr = 9999
where Membership_Nbr not in (select Membership_Nbr from MEMBER_INDEX);
```

Appendix B

Sam's Club Store Performance Assessment

```sql
-- 1. Overall assessment of store sales
-- a. Provide a concise summary of total sales.
select sum(Total_Visit_Amt) as 'Total Store Sales'
from store_visits;



--b. Provide a concise summary of total sales by each store.
-- i. Top 5 performing stores
select top 5 store_visits.Store_Nbr, region_nbr, sum(Total_Visit_Amt) as  'Total Store
Sales'
from store_visits, store_information
where store_visits.store_nbr = store_information.store_nbr
group by store_visits.Store_Nbr, region_nbr
order by sum(Total_Visit_Amt) desc;



-- ii. Bottom 5 performing stores
select top 5 store_visits.Store_Nbr, region_nbr, sum(Total_Visit_Amt) as  'Total Store
Sales'
from store_visits, store_information
where store_visits.store_nbr = store_information.store_nbr
group by store_visits.Store_Nbr, region_nbr
order by sum(Total_Visit_Amt);



-- iii. By Region.
select region_nbr,sum(Total_Visit_Amt) as 'Total Sales By Region'
from store_visits v, store_information i
where v.store_nbr = i.store_nbr
group by region_nbr
order by sum(Total_Visit_Amt) desc;

-- c. Provide a concise summary of total sales breakdowns:
-- i. By the day of the week.
select datename(weekday, Transaction_Date) as 'Day of the week', sum(Total_Visit_Amt) as
'Total Weekday Sales'
from store_visits
group by datename(weekday,Transaction_Date)
order by sum(Total_Visit_Amt) desc;

-- ii. By membership types.
select member_type, sum(Total_Visit_Amt) as 'Total Sales By Membership'
from store_visits,member_index
where store_visits.membership_nbr = member_index.membership_nbr
group by member_type
order by sum(Total_Visit_Amt) desc;

-- iii. By Tender types.
select tender_type, sum(Total_Visit_Amt) as 'Total Sales By Tender Type'
from store_visits
group by tender_type
order by sum(Total_Visit_amt) desc;
```

Appendix C

Membership Purchasing Behavior Assessment

```sql
-- 2. Assessment of member buying behavior
-- a. Provide a concise summary of the typical purchase patterns (amount and number of
-- items) by our members per visit.
-- Top 10 members who spend the most on average
select top 10 store_visits.membership_nbr, member_type,  avg(Total_Visit_Amt) as 'Average
Spending', avg(Tot_Scan_Cnt) as 'Average Number of items'
from store_visits, member_index
where store_visits.membership_nbr = member_index.membership_nbr and
store_visits.membership_nbr != 9999
group by store_visits.membership_nbr, member_type
order by avg(Total_Visit_Amt) desc;

-- b. Provide a concise summary of member visits breakdowns:
-- i. By the day of the week.
select  datename(weekday, Transaction_Date) as 'Day of the week', count(visit_nbr) as
'Total Visit'
from store_visits
group by datename(weekday, Transaction_date)
order by count(visit_nbr) desc;

--ii. By hours during a day.
select convert(varchar, transaction_time) as 'Time_in_String', visit_nbr
into stringtime
from store_visits

select substring(Time_in_String, -5, len(Time_in_String)) as 'Hour of the day',
count(visit_nbr) as 'Total Visit'
from stringtime
group by substring(Time_in_String,-5, len(Time_in_String))
order by count(visit_nbr) desc;

--iii.  By Elite Stat Code.
select elite_stat_code, count(visit_nbr) as 'Visit by Elite status'
from store_visits v, member_index m
where v.membership_nbr = m.membership_nbr
group by elite_stat_code;

-- c. Identify the characteristics of the most active (in sales and in frequent visits)
-- members.
select top 3 membership_nbr,count(visit_nbr) as 'Total Number of Visit',
sum(Total_Visit_amt) as 'Total Spending'
into mostactivecustomer
from store_visits
where membership_nbr != 9999
group by membership_nbr
order by count(visit_nbr) desc;

select m.*
from mostactivecustomer a, member_index m
where m.membership_nbr = a.membership_nbr
```