Iowa Liquor Sales

Group 3: Yuanyuan Liu, Erdong Xia, Shikuan Gao, Zhuohui Jiang

Executive Summary

We collect data about the liquor sales in Iowa, and our purpose is to build a model to help three stores in three major cities in Iowa maximize the profits. Firstly, we select the useful variables in our dataset and process raw data into two parts: cost-retail price information and bottle sold information. For cost-retail price information, we group them by category name and calculate the average bottle volume, average bottle cost and average bottle retail. For bottle sold information, we do the same thing as we did for cost- retail price information and get the sum of sale of every category. And then, we build our model by inputs (bottles sold, average bottle volume, average bottle cost, warehouse capacity and average bottle retail), and constraints(demand, capacity and budget). Finally, we get the result that the objective net income is $60166.45.

Introduction

State of problem

Every retail store pursues the goal of maximizing the profits, so they need to consider how many inventories left, how much profit that they can make by every unit of product, how much warehouse capacity that they have to store their products and the demand of consumers. It's very difficult for them to arrange their replenishment strategy. They don't know how to maximize net income without model and tend to replenish products by experience.

Objective

We choose three stores which located in three major cities in Iowa, and our objective is to build a model to optimize the net income of these three stores.

Hypothesis

With demand constraint and warehouse capacity constraint under consideration, net income goes up as total cost increases until it reaches to a maximum value and no longer increase. Firstly,

Data

Database Introduction

Data source of our project is from data.iowa.gov. The raw data contains more than two million sale records of liquor in Iowa state over past two years. Also, each record has 24 variables: invoice/item number, date, store number, store name, address, city, zip code, store location, county number, county, category (number form), category name, vendor number, vendor name, item number, item description, pack (it shows how many bottles in one pack), bottle volume (ml), state bottle cost, state bottle retail, bottle sold, sale (dollars), volume sold (liters) and volume sold (gallons).

Data Processing

Variable selection

As it is obviously that not all the variables are helpful for our modeling, we select useful variables and break the raw data into two parts: cost-retail price information and bottle sold information.

Cost-retail price contains five variables: date, category name, bottle volume (ml), state bottle cost and state bottle retail.

Bottle sold contains four variables: date, city, category name and bottle sold.

Deleting incomplete data of year 2017

During the data processing, we found that though there was no null value of missing value of variables which we selected for our model, the sale records of year 2017 only had 10 months' records which was lack of data of November and December. Considering the effect of season, we could not easily predict the sale number of last two months by the previous 10 months. However, data of year 2016 did not have such problems. Thus, we deleted all the data with year 2017 in variable date.

Classification of cost-retail information

After deleting incomplete data of year 2017, for cost-retail information, we grouped them by category name. In the meantime of grouping, average bottle volume (ml), average state bottle cost and average state bottle retail of each category name were also calculated. Thus, after

grouping, we had a table with four variables without date which was discarded in this step because it was not useful any more for our model. However, grouping only combined the exactly same category name into one row which means different category names but were considered as same kind of liquor by experience were remain separated after grouping. Then next, we classified category names by 9 categories: scotch and whiskey, tequila, rum, brandy, gin, vodka, cocktail, liqueur and spirit, and average bottle volume (ml), average state bottle cost and average state bottle retail of each category were also calculated. The cost-retail price information after classification are shown as follow.

	Average bottle volume (ml)	Average bottle cost	Average bottle retail
Scotch and Whiskey	885.49	$18.52	$27.80
Tequila	755.29	$13.65	$20.49
Rum	928.94	$9.12	$13.70
Brandy	673.41	$8.14	$12.21
Gin	910.72	$7.67	$11.52
Vodka	903.99	$9.18	$13.79
Cocktail	1,546.22	$7.28	$10.97
Liqueur	812.85	$8.25	$12.40
Spirit	744.12	$9.79	$14.70

Classification of bottle sold information

After deleting incomplete data of year 2017, we also grouped bottle sold information by the combination of city and category name. While grouping, we calculated the sum of bottle sold of each category name. The same situation as cost-retail price information, under the same city value, the category name was classified further into the same 9 categories as well. Also, in process of classification, sum of bottle sold of each category was calculated. Thus, after classification, we got a table with three variables: city, category, bottle sold (date was also discarded for the same reason as cost-retail information). Then, we calculated the total bottle sold for each city, and found the highest three sale records as well as the cities they referred to.

In the end, we grouped the raw data of these three cities by combination of city and store name, then counted the number of specific row for each city and the number of stores of each city was obtained.

The bottle sold information after classification were shown as follow.

	2017 three Major City bottles sold			
Categories	**Cedar Rapids**	**Davenport**	**Des Moines**	**City Total**
Scotch and Whiskey	131556	105203	264118	500877
Tequila	19209	22184	67179	108572
Rum	52787	36795	89324	178906
Brandy	35539	39202	67952	142693
Gin	27733	26851	62991	117575
Vodka	151968	125152	305824	582944
Cocktail	10722	7324	8451	26497
Liqueur	58132	48666	144068	250866
Spirit	9050	3282	3637	15969
Total				1924899
Number of Stores	182	102	187	

Modeling

Identity problem

In the modeling section. There are five subsections which are identity problem, inputs, variables, constraints and objective value. This section includes all the details of setting up a model and all the modeling techniques.

In this problem, our objective is to maximize the net income of the 2018 Iowa liquor store procurement plan. The procurement plan makes each category of liquor must be an integer. In another word, optimization models in which some or all of the variables are constrained to be integer are known as integer programming models.

Inputs

From the dataset we used in Iowa government official website, there are three parts we will be used as inputs in our model. There are three major city bottles sold, liquors information and warehouse capacity information.

First, three major city bottles sold in 2017 is the dataset including nine categories of liquor we classified in the data classification stage (scotch and whiskey, tequila, rum, brandy, gin, vodka, cocktail, liqueur and spirit), three highest sales record cities we chose (Cedar Rapids, Davenport and Des Moines) and the number of stores in each city.

	2017 three Major City bottles sold			
Categories	**Cedar Rapids**	**Davenport**	**Des Moines**	**City Total**
Scotch and Whiskey	131556	105203	264118	500877
Tequila	19209	22184	67179	108572
Rum	52787	36795	89324	178906
Brandy	35539	39202	67952	142693
Gin	27733	26851	62991	117575
Vodka	151968	125152	305824	582944
Cocktail	10722	7324	8451	26497
Liqueur	58132	48666	144068	250866
Spirit	9050	3282	3637	15969
Total				1924899
Number of Stores	182	102	187	

Second, from the origin dataset from Iowa government official website, we filtered and calculated three valuable cost-retail information, there are average bottle volume, average bottle cost and average bottle retail price. They will be use to calculate the total capacity used, total cost and total revenue.

	Average bottle volume (ml)	Average bottle cost	Average bottle retail
Scotch and Whiskey	885.49	$18.52	$27.80
Tequila	755.29	$13.65	$20.49
Rum	928.94	$9.12	$13.70
Brandy	673.41	$8.14	$12.21
Gin	910.72	$7.67	$11.52
Vodka	903.99	$9.18	$13.79
Cocktail	1,546.22	$7.28	$10.97
Liqueur	812.85	$8.25	$12.40
Spirit	744.12	$9.79	$14.70

Third and last, from the origin dataset from Iowa government official website, we take the warehouse capacity information for each city. The warehouse capacity is recorded in milliliters. This input is valuable in setting capacity constraint in the constraints subsection.

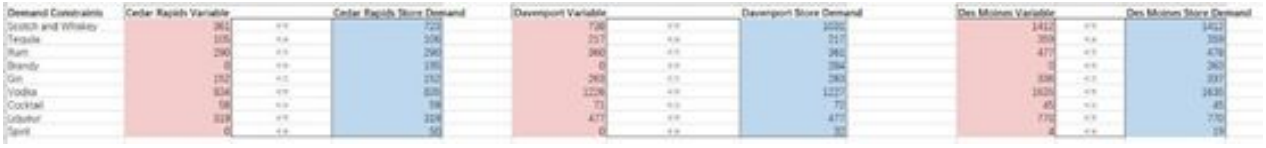
City	Warehouse Capacity (ml)
Cedar Rapids	2200000
Davenport	3000000
Des Moines	4500000

Variables

In this problem, variables are each store of three cities procurement plan in 2018. Because each store has nice categories of liquor to purchase, so the total variables are three cities times nice categories equal to twenty-seven variables.

Constraints

In this problem, there are three major constraints. We set the constraints base on the reality situation in Iowa liquor business. First is the demand constraint. The numbers of demand in each city are based on the 2017 bottles sold input. We use each category divided by the number of stores, in this way we get each city demand for each category of liquor. Furthermore, we set the variables to be smaller or equal than demands because in this problem we wish to reduce leftover as low as possible to cut off the storage fees.



Second is capacity constraint. In the reality situation, the store volume cannot excess the warehouse capacity. For the store volume, it can be calculated as the sum product of each city variables and average bottle volume.

Cedar Rapids Store Volume		Warehouse Capacity (ml)
1909699.78	<=	2200000

Davenport Store Volume		Warehouse Capacity (ml)
2997134.02	<=	3000000

Des Moines Store Volume		Warehouse Capacity (ml)
4447047.23	<=	4500000

Third is budget constraint. The total cost for procurement plan cannot excess the budget plan. For the total cost, it can be calculated as the sum product of total variables and average bottle cost.

Total Cost		Budget
$119,998.49	<=	$120,000

Objective Value

The objective value is the maximum net income of the 2018 Iowa liquor procurement plan. The simplify objective function is net income equal to total revenue minus total cost. For the total

revenue, it's the same formula as total cost, it can be calculated as the sum product of total variables and average bottle retail price.

Total Cost
$119,998.49

Total Revenue
$180,164.93

Objective Value
Maximum Net Income
$60,166.45

Results

Solver results - current restocking plan

Solver results provided a restocking plan of how much bottles should be restocked for each category in each store.

Demand Constraints	Cedar Rapids Variable	Davenport Variable	Des Moines Variable
Scotch and Whiskey	361	738	1412
Tequila	105	217	359
Rum	290	360	477
Brandy	0	0	0
Gin	152	263	336
Vodka	834	1226	1635
Cocktail	58	71	45
Liqueur	319	477	770
Spirit	0	0	4

Total volume used for each warehouse of the store is shown below.

Cedar Rapids Store Volume
1909699.78

Davenport Store Volume
2997134.02

Des Moines Store Volume
4447047.23

At last, the most importantly, the total cost is $119,998.49, and maximum net income is $60,166.45.

Budget Constraint
Total Cost
$119,998.49
Total Revenue
$180,164.93
Objective Value
Maximum Net Income
$60,166.45

Solver Table 1 - Sensitivity of net income to warehouse capacity

The purpose of creating a Solver Table for warehouse capacity is to see the potential growth of net income if we expand the warehouse.

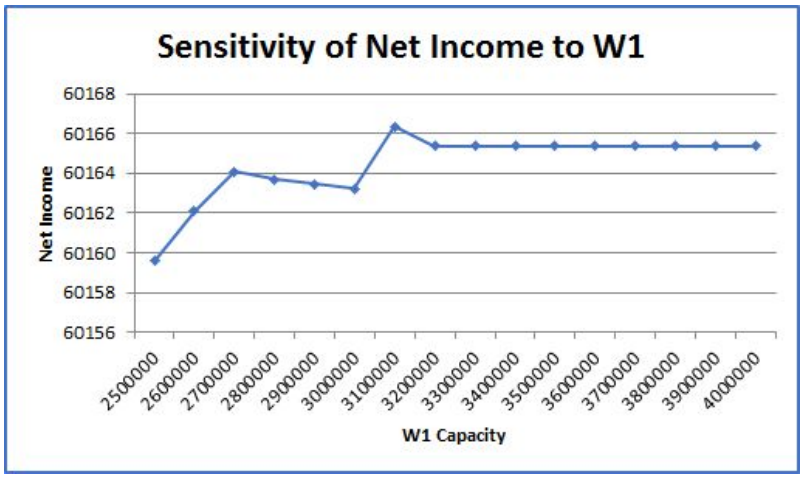
Since Davenport and Des Moines met their warehouse capacity constraints, but Cedar Rapids didn't meet its constraint from Solver results, only Davenport and Des Moines will be considered to expand.

Cedar Rapids Store Volume		Warehouse Capacity (ml)
1909699.78	<=	2200000

Davenport Store Volume		Warehouse Capacity (ml)
2997134.02	<=	3000000

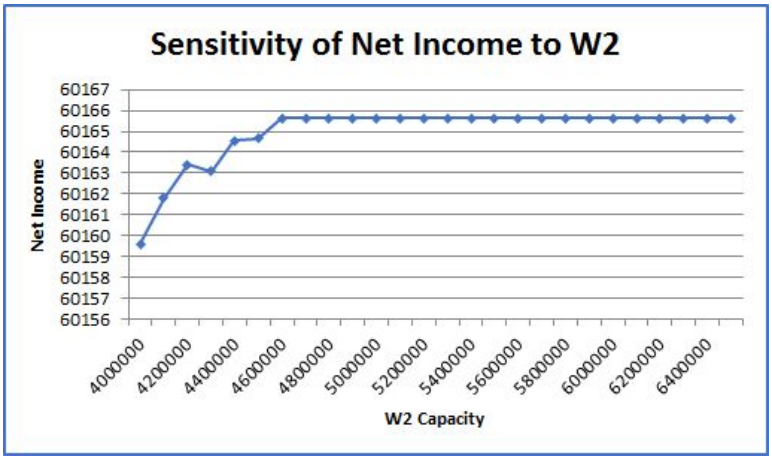
Des Moines Store Volume		Warehouse Capacity (ml)
4447047.23	<=	4500000

The figure below is showing that how net income changes if Davenport warehouse capacity changes and the volume of Des Moines stays at 4,000,000 mL. The maximum net income is $60,166.36.



Davenport

The figure below is showing that how net income changes if Des Moines warehouse capacity changes and the volume of Davenport stays at 2,500,000 mL. The maximum net income is $60,165.65.



Des Moines

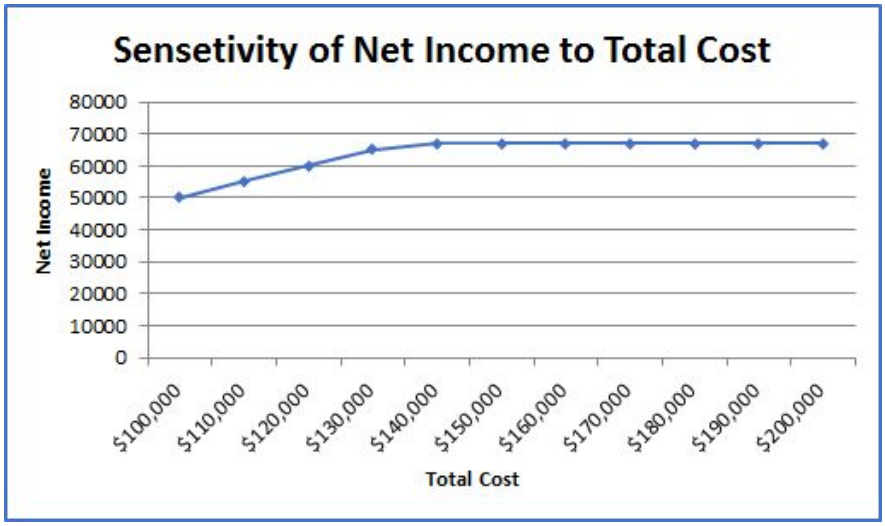
The two figures above only present the information in the red triangle of the table below, their maximum net income is shown in yellow cells. By looking at the table, the maximum income take place when both warehouse capacity change, which is 2,700,000 mL for Davenport, 4,500,000 mL for Des Moines. In this condition, the maximum net income is $60,166.87 (shown in green cell).

Warehouse Capacity	4000000	4100000	4200000	4300000	4400000	4500000	4600000	4700000	4800000
2500000	$60,159.60	$60,161.79	$60,163.42	$60,163.07	$60,164.55	$60,164.67	$60,165.65	$60,165.65	$60,165.65
2600000	$60,162.07	$60,162.30	$60,165.72	$60,166.41	$60,163.22	$60,165.25	$60,165.25	$60,165.25	$60,165.25
2700000	$60,164.10	$60,165.48	$60,164.53	$60,164.59	$60,165.15	$60,166.87	$60,166.87	$60,166.87	$60,166.87
2800000	$60,163.69	$60,162.61	$60,164.05	$60,165.15	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45
2900000	$60,163.44	$60,163.22	$60,165.15	$60,165.70	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45
3000000	$60,163.22	$60,165.15	$60,165.38	$60,165.70	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45
3100000	$60,166.36	$60,165.38	$60,165.38	$60,165.70	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45
3200000	$60,165.38	$60,165.38	$60,165.38	$60,165.70	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45
3300000	$60,165.38	$60,165.38	$60,165.38	$60,165.70	$60,166.44	$60,166.45	$60,166.45	$60,166.45	$60,166.45

Solver Table 2 - Sensitivity of net income to total cost

In order to see how the change of total cost influente maximum net income, we created Solver Table 2. According to the table and graph shown below, when total cost reaches to $140,000, the maximum net income $66,931.76.

Total Cost	Net Income
$100,000	$50,151.70
$110,000	$55,159.07
$120,000	$60,166.45
$130,000	$65,171.59
$140,000	$66,931.76
$150,000	$66,931.76
$160,000	$66,931.76
$170,000	$66,931.76
$180,000	$66,931.76
$190,000	$66,931.76
$200,000	$66,931.76



Conclusion

The current restocking plan is to spend $119,998.49 on purchasing alcohol, and the maximum net income is $60,166.45.

Since the net income has more potential to grow when purchasing more alcohol than expanding warehouse. If more budget were spent, purchasing more alcohol to fulfill the demand would be the preferable way to increase net income.

Appendix

SAS code for bottle sold information part of the project.

After import the bottle sold information which was already grouped by category name as city.

```
proc sql;
create table city1 as
select *
from city
where city is not null and Category_Name is not null
order by city;
run;


%macro cat_city(cat, name1, name2);
proc sql;
create table &cat as
select *
from city1
where Category_Name like &name1 or Category_Name like &name2
```

```
order by city;
run;

proc sql;
update &cat
set Category_Name = "&cat";
run;
%mend cat_city;

%cat_city(spirit, '%Spirit%', '%SPIRIT%');
%cat_city(rum, '%RUM%', '%Rum%');
%cat_city (tequila, '%TEQUILA', '%Tequila%');
%cat_city(brandy, '%BRANDIES%', '%Brandies%');
%cat_city(Vodka, '%VODKA%', '%Vodka%');
%cat_city(Gin, '%GIN%', '%Gin%');
%cat_city(Cocktail, '%COCKTAIL%', '%Cocktail%');
%cat_city(schnapps, '%SCHNAPPS%', '%Schnapps%');
%cat_city(Liqueur, '%LIQUEUR%', '%Liqueur%');
%cat_city(whiski, '%WHISK%', '%Whisk%');

%macro diss(cat, cat1);
proc sql;
create table &cat1 as
select distinct city, Category_Name, sum(Bottles_Sold) as Bottles_Sold
from &cat
```

```
group by city;
run;
%mend diss;

%diss(spirit, spirit1);
%diss(rum, rum1);
%diss(tequila, tequila1);
%diss(brandy, brandy1);
%diss(Vodka, Vodka1);
%diss(Gin, Gin1);
%diss(Cocktail, Cocktail1);
%diss(schnapps, schnapps1);
%diss(Liqueur, Liqueur1);
%diss(whiski, whiski1);

proc sql;
create table liq_sch as
select *
from liqueur1;
run;

proc sql;
insert into liq_sch
select *
from schnapps1;
```

```
run;

proc sql;
update liq_sch
set Category_Name = 'liqueur';
run;

%diss(liq_sch, liq_sch1);

proc sql;
create table countcity as
select distinct city
from city1;
run;

proc sql;
create table cc1 as
select *
from spirit1;
run;

%macro addcat(cat);
proc sql;
insert into cc1
select *
```

```
from &cat;
run;
%mend addcat;


%addcat(rum1);
%addcat(tequila1);
%addcat(brandy1);
%addcat(Vodka1);
%addcat(Gin1);
%addcat(Cocktail1);
%addcat(whiski1);
%addcat(liq_sch1);


proc sql;
create table cc2 as
select *
from cc1
order by city;
run;
```

Table cc2 was the result of bottle sold information,

References

Aleksey Bilogur. 2017. “Iowa liquor sales,” Website:
https://www.kaggle.com/residentmario/iowa-liquor-sales/contributors
State of lowa. 2017. Website:
https://data.iowa.gov/