

Refrigerator Sales Analysis

Yuetian Sun; Chen Jing; Keer Jiang; Ge Jin

10/8/2017

Introduction

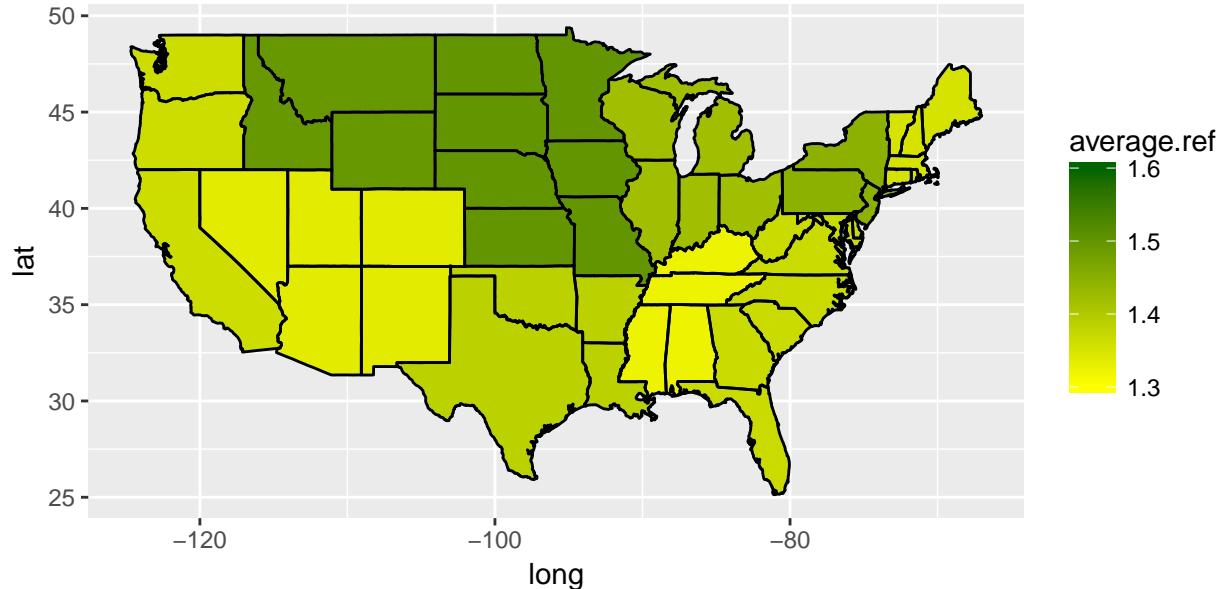
This report presents a competitive analysis of the US refrigerator market based on the 2015 EIA data. We want to explore how to improve the sale of refrigerators by using a more intuitional way with visualising, transforming and modeling the data. After searching some marketing information, we decide to devide the problem into five part. The first part is analyzing how region characteristics affect the refrigerator selling. More specificly we focus on how region, census division, metropolitan or micropolitan statistical area, urban type and climate zones affect the number, age, size and through-the-door ice of refrigerators. And we utilize Anova test to see if these factors does affect refrigerator selling. The second part is about how different household income affect the number, size, age and whether has through-the-door ice of refrigerators. In the third, we try to find how these factors affect whether people should purchase energy star qualified refrigerator. And in the fourth part, we focus on how the factors affect purchasing recycling refrigerators. The fifth part is to build statistics model by using Logistics regression to predict number, size, and age of refrigerator. We use AIC to do variable selection and use ROC curve and AUC method to evaluate the model. After data visualization, hypotheses testing, model selection and model evaluation, we can come to the conclusion of the place where we should develop our market in and what kinds of the refrigerator we should product most. We make specific analysis on selling different kinds of refrigerator to different people. Most importantly, we give specific suggestion on how to improve the sale of refrigerator.

1 How region characteristics affect the refrigerator selling

1.1 How region characteristics affect the number of refrigerator

1.1.1 How the census affects the number of refrigerator

Average of Refrigerator Based on U.S. Division



Total of Refrigerator Based on U.S. Division

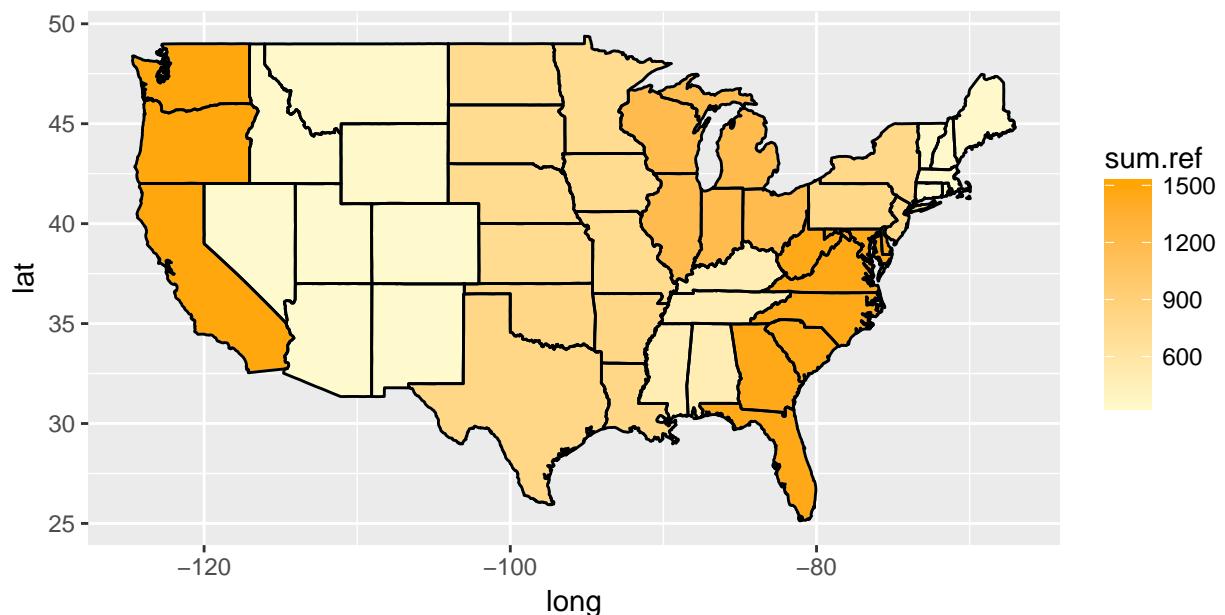


Table 1: Anova test of region

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
recs\$DIVISION	1	2.482549	2.482549	5.450808	0.0195939

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	5684	2588.754877	0.455446	NA	NA

As we can see from the plot above, there is big difference among the total number of refrigerators people bought from different census divisions. The Pacific and South Atlantic parts seem to have more refrigerators than the others. While the New England, Mountain North and Mountain South seem to have fewer. We all know that the South Atlantic and Pacific are near the ocean and the south part has higher temperature than the Northeast part. Thus, we can make an assumption that the number of refrigerator has correlation between census division regional characteristics like the weather. Another assumption is that it is simply because this survey included lots of people from South. So we use Anova to test the difference among the average value in different census division. As we can see from the table above, the result shows that there is significant difference among the mean. Thus, we can draw the conclusion that census division parts do have significant effect on buying refrigerator. People from mountain north and west north central would like to buy more refrigerator than the others.

1.1.2 How Metropolitan or Micropolitan affects the number of refrigerator

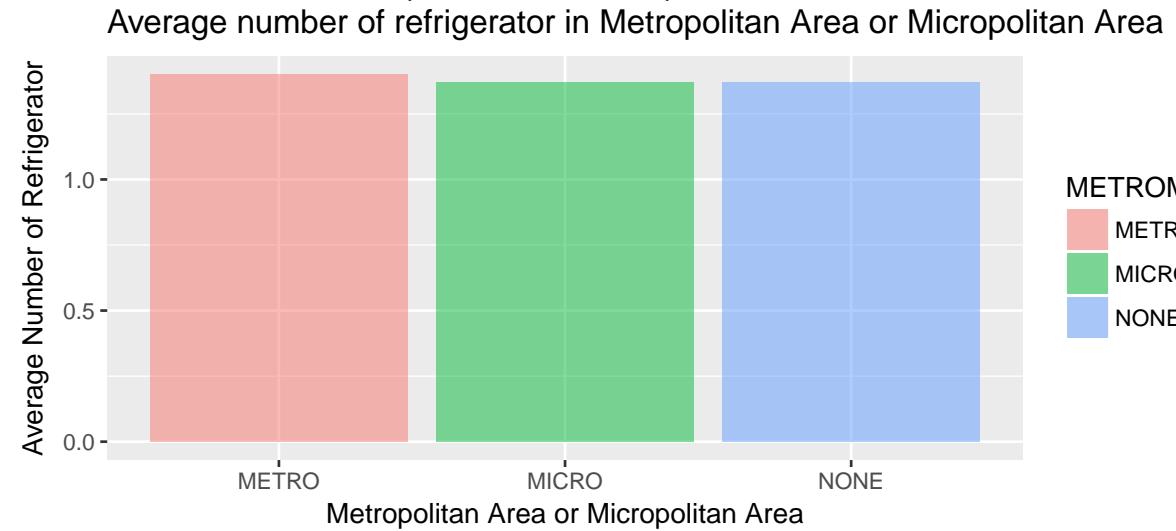
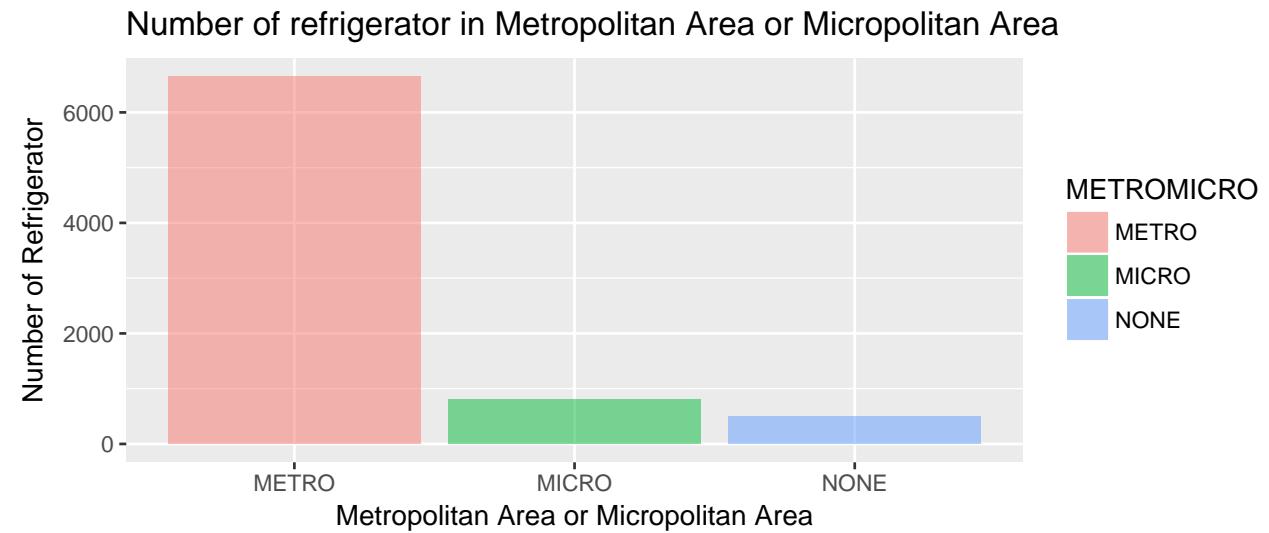


Table 2: Anova test of Metropolitan or Micropolitan area

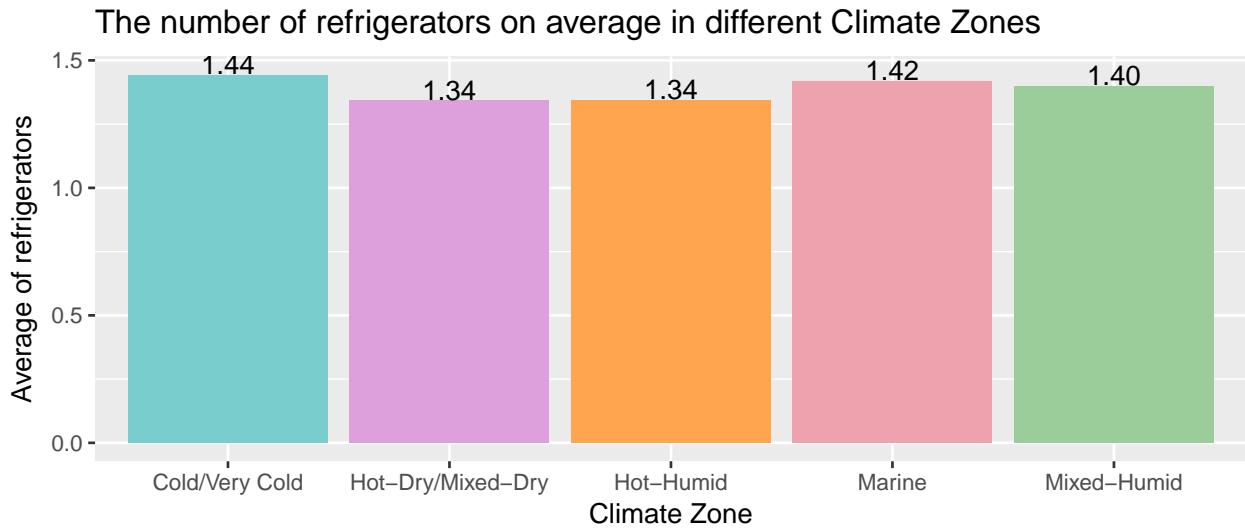
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
recs\$METROMICRO	2	0.7354019	0.3677009	0.8066562	0.4463992
Residuals	5683	2590.5020234	0.4558335	NA	NA

As we can see from the graph, the Metropolitan area has much more refrigerator than the others. We think it may because there are lots of people living in the Metropolitan area or they tend to have more income. To make it more precisely, we use anova to test the whether there are significant difference among the average number of refrigerator people bought in these area. As we can see from the table above, the p-value is much larger than the significant value. Thus we believe there is actually no difference in the number of buying refrigerator per person no matter the customer coming from big city. So we can develop our business in Metropolitan or the city which is booming and will have larger population in the future.

1.1.3 How do Climate Zones affect the number of refrigerators

Table 3: Total Number of Refrigerators in Different Climate Zones

CLIMATE_REGION_PUB	NUMFRIG
Cold/Very Cold	2811
Hot-Dry/Mixed-Dry	1006
Hot-Humid	1393
Marine	601
Mixed-Humid	2052

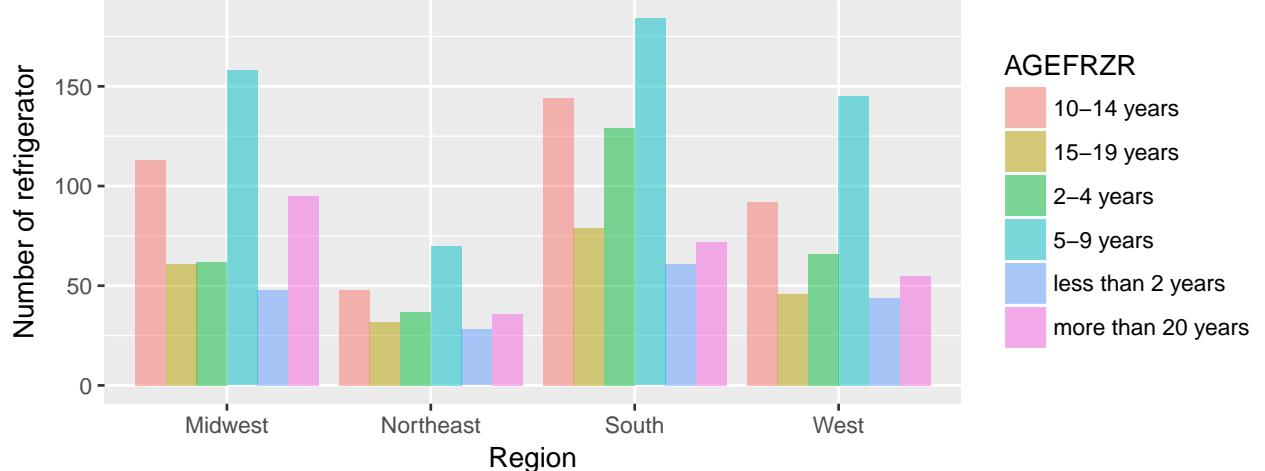


Note that the total number of refrigerators used in cold/very cold areas is the largest whereas the total number in Hot-Dry/Mixed-Dry areas is the least. And from the aspect of average number, the result is similar. So we can conclude that there is more demand for refrigerators in the cold/very cold, and mixed-humid areas.

1.2 How region characteristics affect the age of refrigerator

1.2.1 How region affects the age of most-used refrigerator

Number of most-used refrigerator in different age and region



Relevant frequency of most-used refrigerator in different age and region

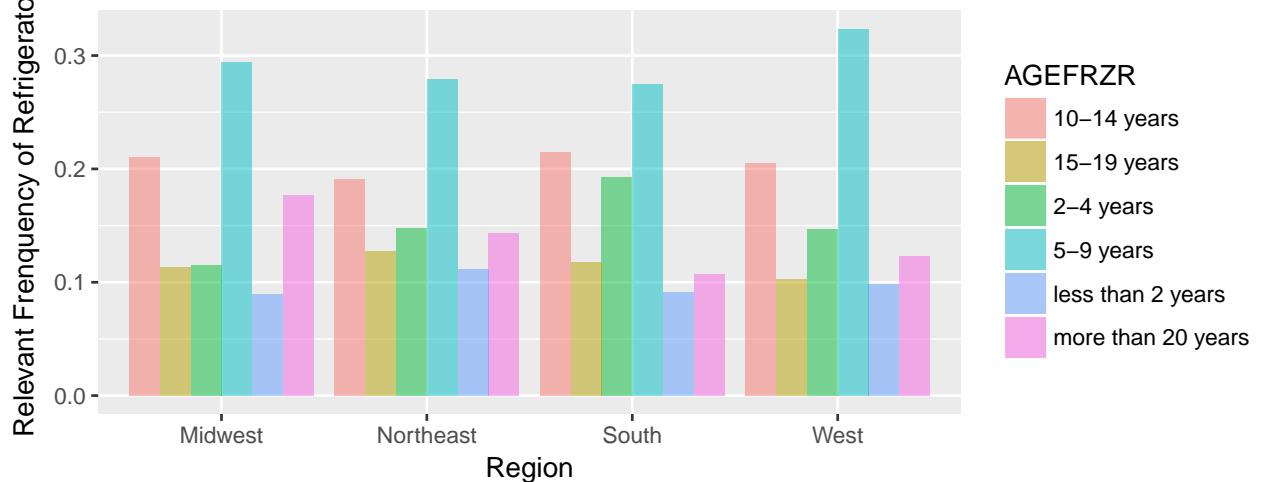


Table 4: Old refrigerators' relevant frequency in different regions

Relevant Frequency	
Northeast	0.2709163
Midwest	0.2905028
South	0.2257100
West	0.2254464

We find in the internet that the average usage year of refrigerator is about 15 years. So we define the refrigerator used more than 15 years is a old refrigerator meaning that it will be replace soon in the future. As we can see from the plot, the midwest region has more percentage of old refrigerator from the other regions. Combining the first region analysis result, we can conclude that we develop the market more in Midwest area. People living there not only tend to buy more refrigerator but also they will replace the lots of old one in the future.

1.2.2 How census division affects the age of most-used refrigerator

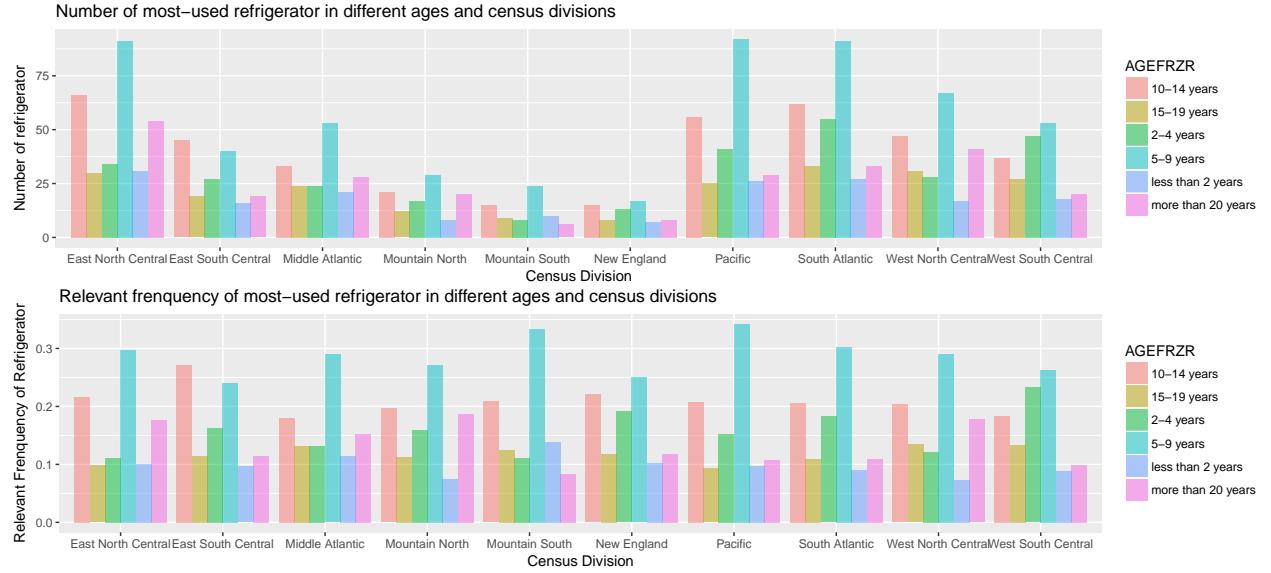


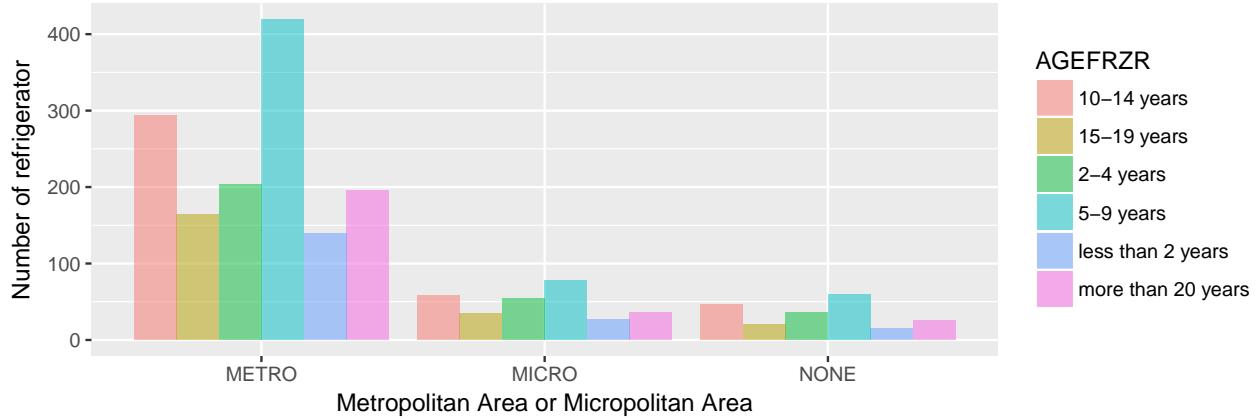
Table 5: Old refrigerators' relevant frequency in different census divisions

Relevant Frequency	
New England	0.2352941
Middle Atlantic	0.2841530
East North Central	0.2745098
West North Central	0.3116883
South Atlantic	0.2192691
East South Central	0.2289157
West South Central	0.2326733
Mountain North	0.2990654
Mountain South	0.2083333
Pacific	0.2007435

As we can see from the plot and the percentage of old refrigerator in the different census division, the West North Central has more old refrigerator than the other parts. Also, Mountain north part, Middle Atlantics and East North Central get large percentage of old refrigerator. Thus, we can sell more product in these area. This result makes sense because the West North Central and East North Central is just subdivision of the Middle part.

1.2.3 How Metropolitan or Micropolitan affects the age of most-used refrigerator

Number of different ages of most-used refrigerator in Metropolitan types



Relevant frequency of different ages most-used refrigerator in Metropolitan types

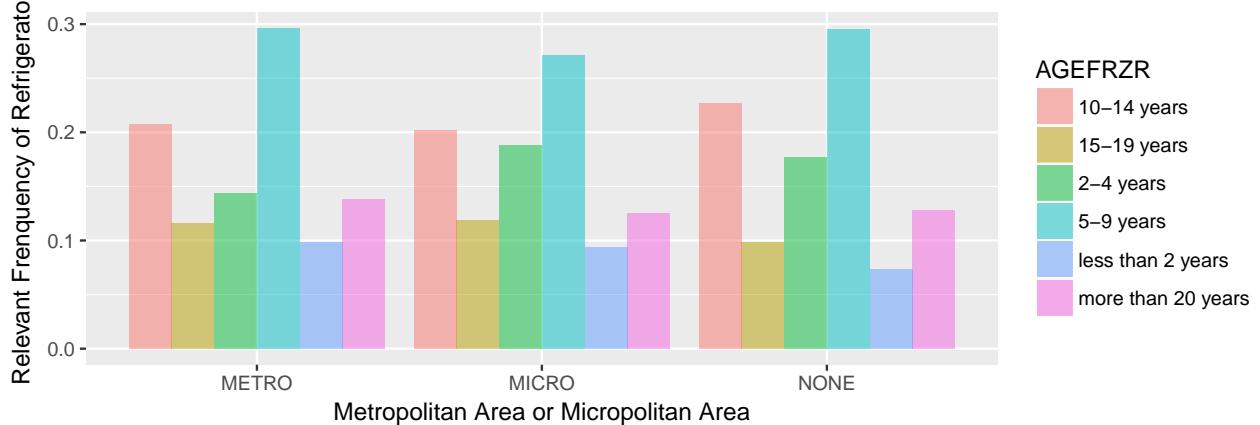


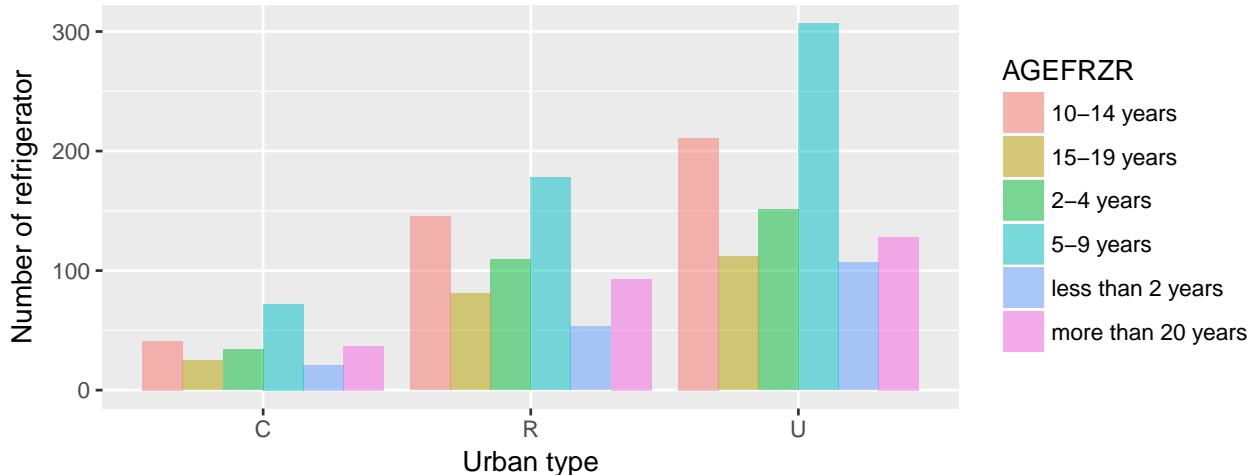
Table 6: Old refrigerators' relevant frequency in Metropolitan or Micropolitan Area

Relevant Frequency	
Metropolitan	0.2544170
Micropolitan	0.2439024
None	0.2266010

From the above plots and results, we find that the percentage of old refrigerator in Metropolitan Area and Micropolitan Area are nearly equal but more than the “None” part. Combining the conclusion that Metropolitan or Micropolitan Area doesn’t have significant effect on the average number of refrigerator people buying, we believe the larger population the more old refrigerator in Metropolitan and Micropolitan Area. As a result of these, we can produce more new refrigerator in city where the population is large because lots of people will buy new refrigerator and many old refrigerators need to be replaced.

1.2.4 How urban type affects the age of most-used refrigerator

Number of different age and urban type of most-used refrigerator



Relevant frequency of most-used refrigerator in different urban type

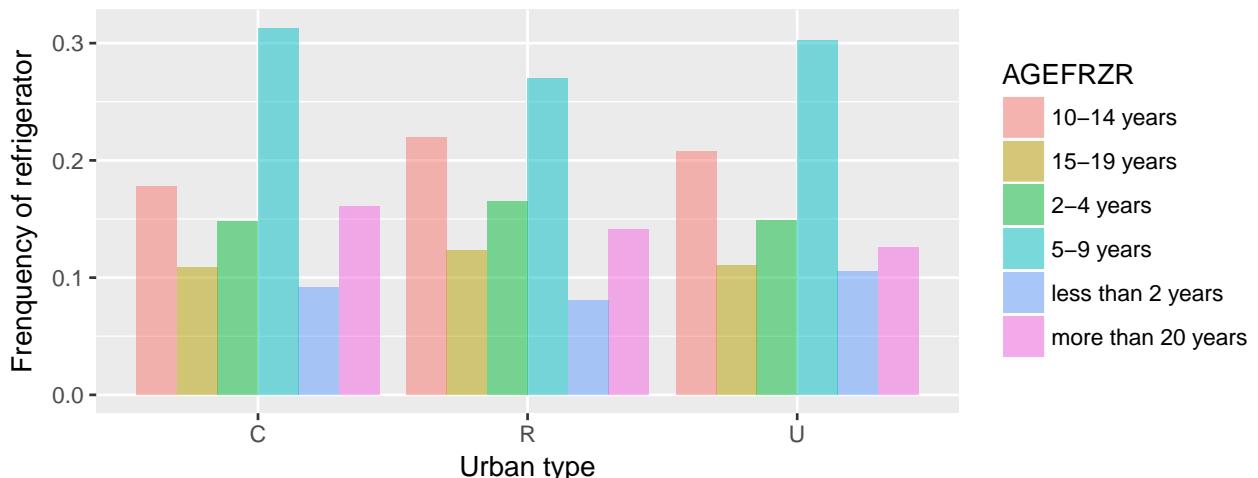


Table 7: Old refrigerators' relevant frequency in different urban types

	Relevant Frequency
Urban Area	0.2362205
Rural Area	0.2640364
Urban Cluster	0.2695652

As we can see from the above graph, we find that the both rural area and urban cluster has higher rate of old refrigerator. Combining the result with metropolitain types, we can conclude that the we should develop our business in the metropolitain city's rural and urban cluster area. In other words, we should focus on the city with very large poplution and inside the city we put more product in urban cluster and rural area.

1.2.5 How do Climate Zones affect the age of refrigerators

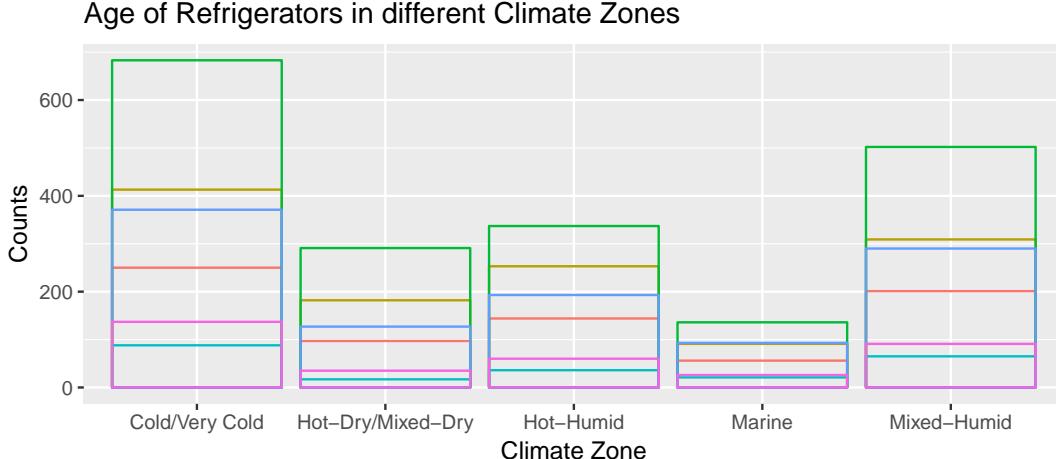


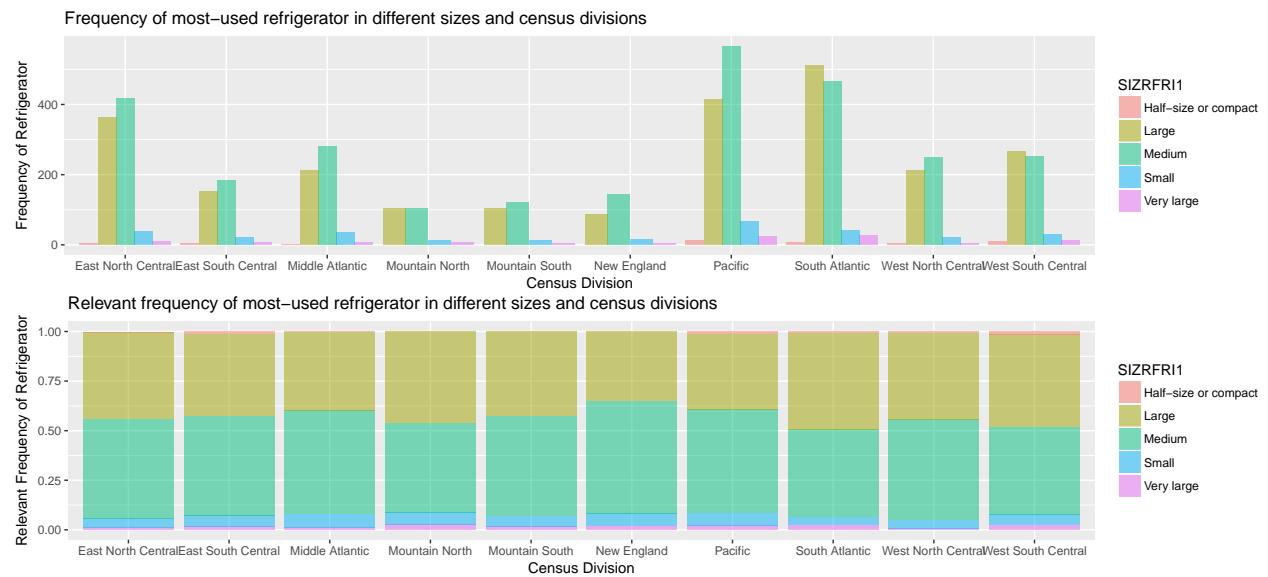
Table 8: The Proportion of Age of Refrigerators in different Climate Zone

	< 2 years	2~4 years	5~9 years	10~14 years	15~19 years	> 20 years
Cold/Very Cold	0.1287333	0.2126674	0.3516993	0.0453141	0.1910402	0.0705458
Hot-Dry/Mixed-Dry	0.1295060	0.2429907	0.3885180	0.0226969	0.1695594	0.0467290
Hot-Humid	0.1407625	0.2473118	0.3294233	0.0351906	0.1886608	0.0586510
Marine	0.1323877	0.2151300	0.3215130	0.0496454	0.2198582	0.0614657
Mixed-Humid	0.1378601	0.2119342	0.3443073	0.0445816	0.1989026	0.0624143

Similarly, the climate zone also does not have very significant effect on the age of refrigerators. But there is an obvious tendency that most people in different climate zone all prefer to change refrigerators in 5 ~ 9 years. We can see that the upgrade of refrigerator products is not very frequent. Hence, developing new product and accelerating the update of product are important means for refrigerator enterprise to continuous operation and development.

1.3 How region characteristics affect the size of refrigerator

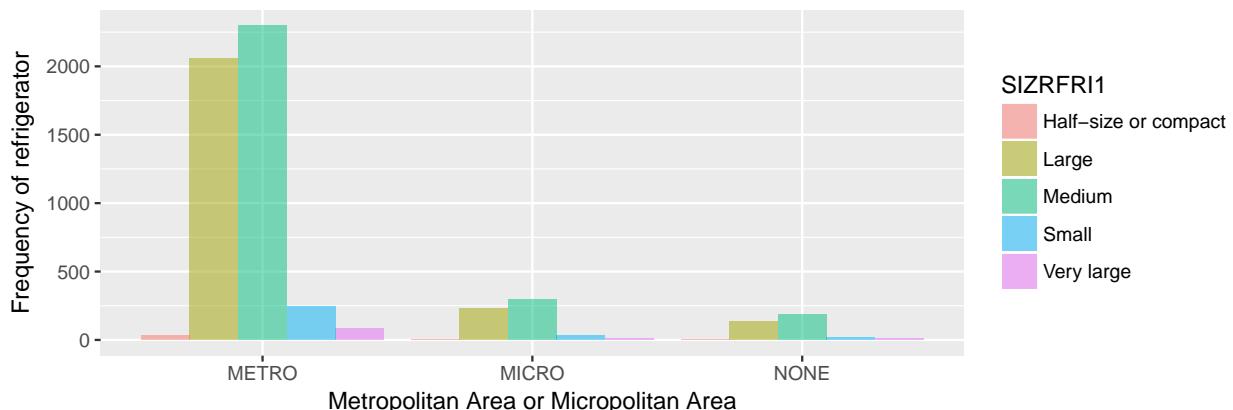
1.3.1 How census division affects the size of most-used refrigerator



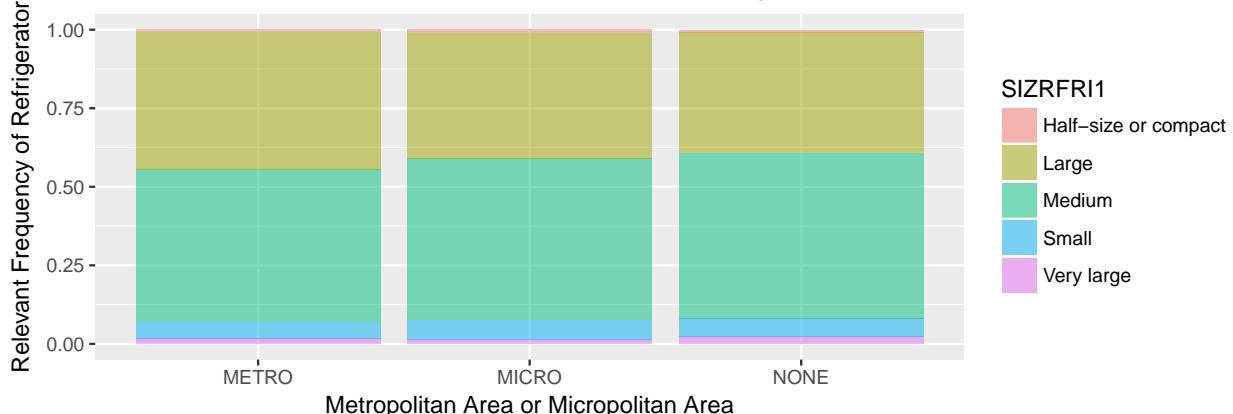
As we can see from the plot above, the medium and large size refrigerators are very popular in all the census division. Thus, we can produce much more medium and large refrigerator than the other products. And all the regions except the South has more medium refrigerator than the large one. Maybe it is because of the hot weather in South part of US, people need more space in the refrigerator. Thus, we can produce a little bit more large refrigerators in South.

1.3.2 How Metropolitan or Micropolitan division affects the size of most-used refrigerator

Relevant frequency of different sizes most–used refrigerator in Metropolitan or Micropolitan A



Relevant frequency of different sizes most–used refrigerator in Metropolitan or Micropolitan A



As we can see from the frequency and relevant plots, it seems no matter what kind of region division, all types of areas tend to have much more medium and large refrigerator than the others. It seems that Medium and large refrigerators are more commonly used and cost efficient. Thus, we can make conclusion that we should produce more large and medium size refrigerator than the other sizes in all areas.

1.3.3 How do Climate Zones affect the size of refrigerators

Different Size of Refrigerators in different Climate Zones

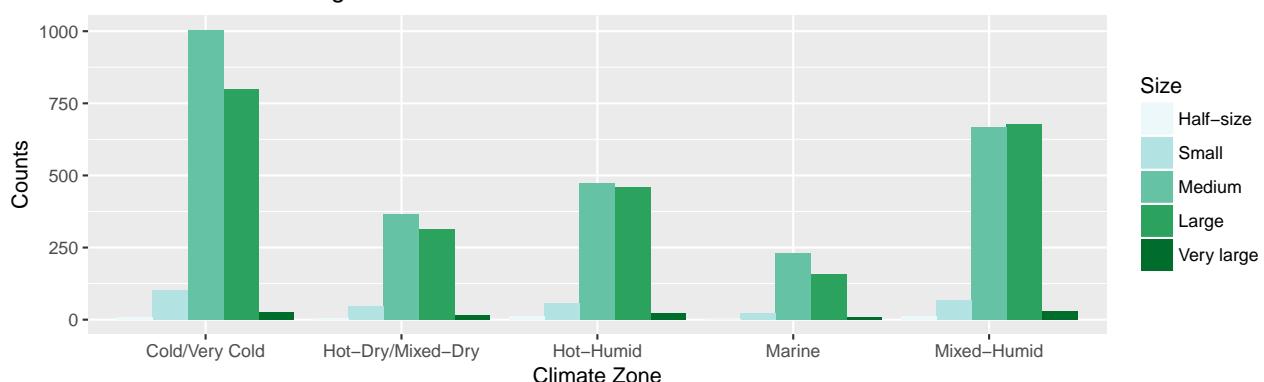


Table 9: The Proportion of Different Size of Refrigerators in different Climate Zone

	Half-size or compact	Small	Medium	Large	Very large
Cold/Very Cold	0.0051493	0.0525232	0.5175077	0.4119464	0.0128733
Hot-Dry/Mixed-Dry	0.0080107	0.0627503	0.4886515	0.4192256	0.0213618
Hot-Humid	0.0127077	0.0547410	0.4623656	0.4486804	0.0215054
Marine	0.0070922	0.0543735	0.5437352	0.3735225	0.0212766
Mixed-Humid	0.0075446	0.0473251	0.4588477	0.4650206	0.0212620

The climate zone does not have very significant effect on the size of refrigerators. However, it is clearly that every climate zone prefers medium and large refrigerators. So this result identifies medium and large refrigerators to be one of the major trends that will have a positive impact on the growth of the refrigerator market in the coming years.

1.4 How do Climate Zones affect whether has through-the-door ice of refrigerators

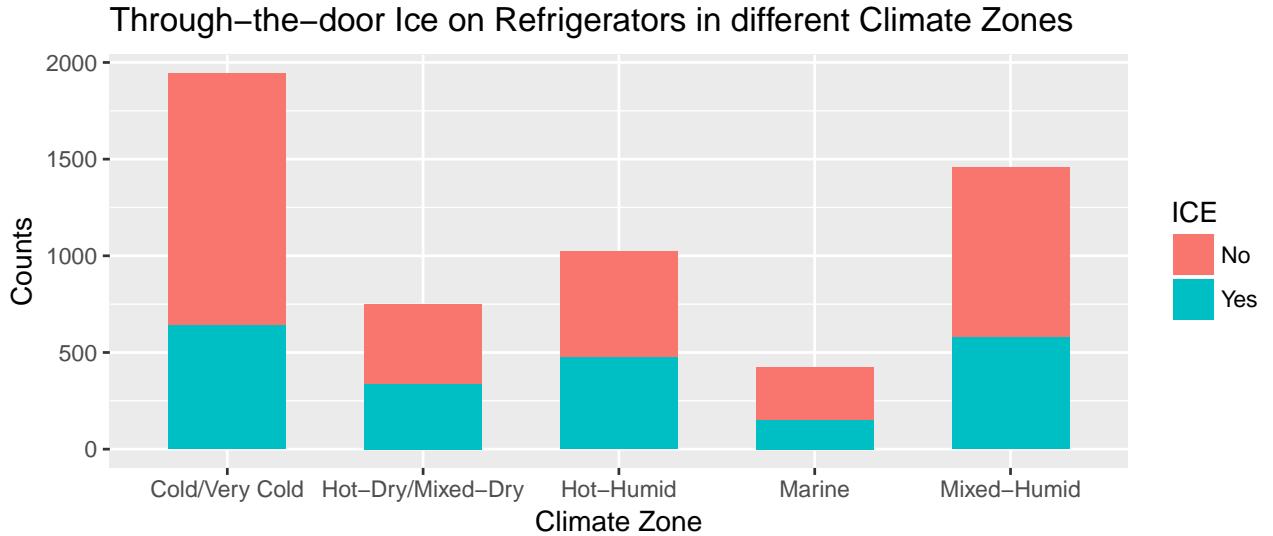


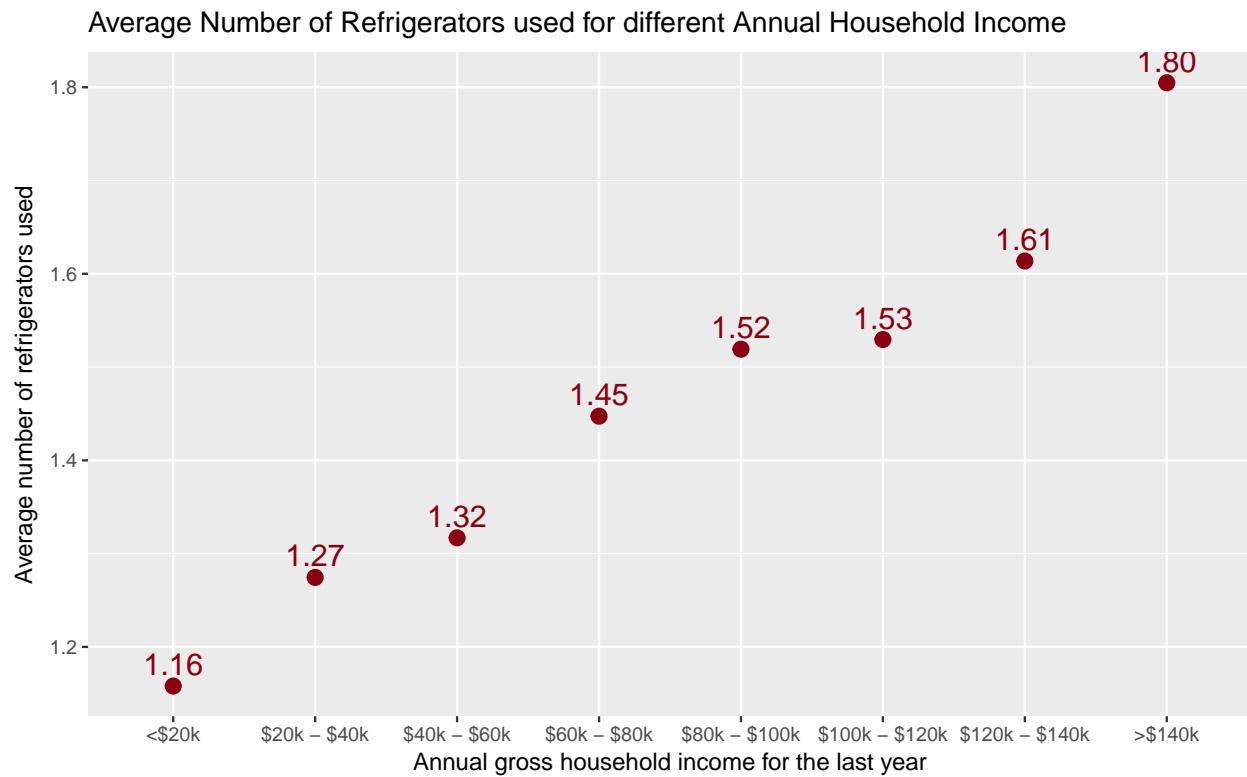
Table 10: The Proportion of having Through-the-door Ice on Refrigerators in different Climate Zone

	No	Yes
Cold/Very Cold	0.6688980	0.3311020
Hot-Dry/Mixed-Dry	0.5500668	0.4499332
Hot-Humid	0.5356794	0.4643206
Marine	0.6453901	0.3546099
Mixed-Humid	0.6008230	0.3991770

From the figure and table above, the result is reasonable that clearly people in hot-dry/mixed-dry, and hot-humid areas tend to this feature on refrigerators. With through-the-door ice on refrigerators, people will have more space for frozen foods. The rest three climate zones are relatively less in percentage.

2 How household income affect the refrigerator selling

2.1 How does different household income affect the number of refrigerators



From the figure above, we can see that as income increases, the average number of refrigerators used increases from 1.16 to 1.80. This indicates that households with higher-income are more likely to use refrigerators, and they may prefer quickly renewed products with such characteristics as advanced technology, complicated function.

2.2 How does different household income affect the size of refrigerators

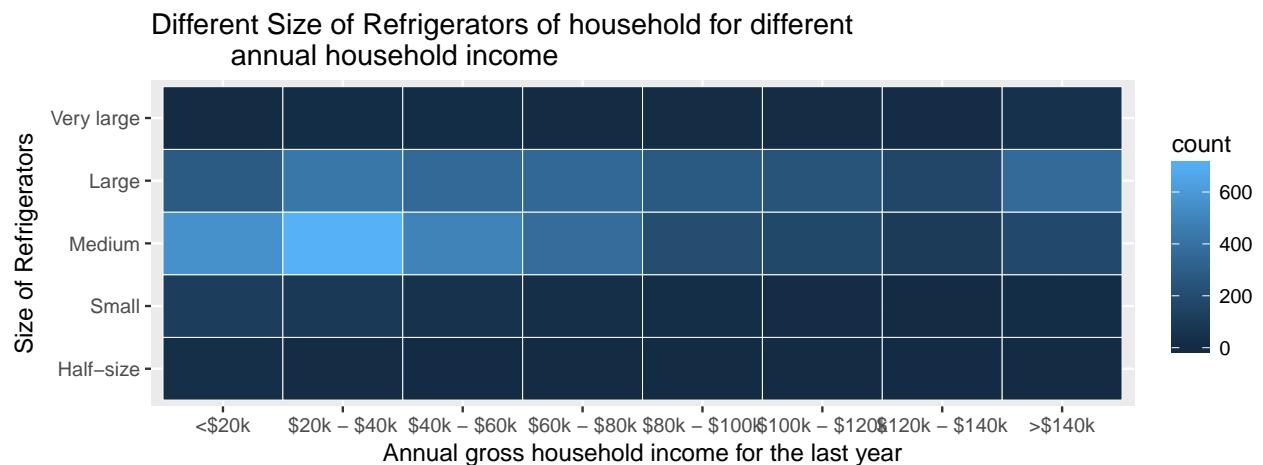


Table 11: The Proportion of Different Size of Refrigerators for each income level

	Half-size or compact	Small	Medium	Large	Very large
<\$20k	0.0247678	0.1093911	0.5758514	0.2837977	0.0061920
\$20k - \$40k	0.0087510	0.0692124	0.5719968	0.3389021	0.0111376
\$40k - \$60k	0.0033520	0.0446927	0.5441341	0.3944134	0.0134078
\$60k - \$80k	0.0000000	0.0330251	0.4980185	0.4610304	0.0079260
\$80k - \$100k	0.0020243	0.0323887	0.4008097	0.5425101	0.0222672
\$100k - \$120k	0.0000000	0.0212766	0.3924350	0.5626478	0.0236407
\$120k - \$140k	0.0108696	0.0144928	0.3586957	0.5869565	0.0289855
>\$140k	0.0034247	0.0205479	0.3030822	0.6113014	0.0616438

The effect of household income on refrigerators isn't so obvious as effects on number. But like climate zone, there is a clear phenomenon that people all prefer to use medium and large refrigerators regardless of income. What we need to notice is that the percentage of medium refrigerators declines to 30.3% from 57.5% as income increases, which is the opposite of the percentage of large refrigerators. The percentage of large refrigerators increases from 28.3% to 61.1% as income increases. Moreover, the percentage of very large refrigerators also increases with higher-income.

2.3 How does different household income affect whether has through-the-door ice of refrigerators

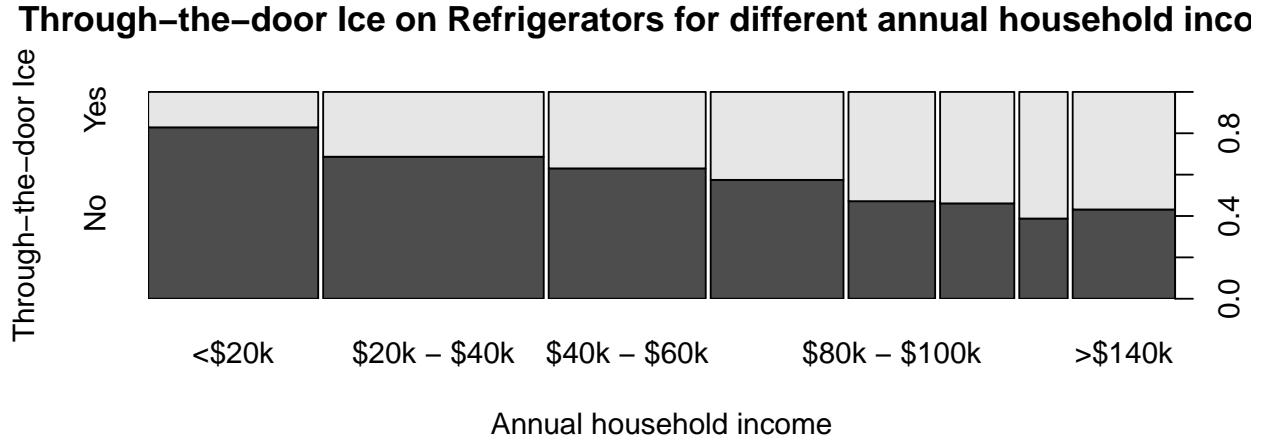


Table 12: The Proportion of having Through-the-door Ice on Refrigerators for each income level

	No	Yes
<\$20k	0.8286894	0.1713106
\$20k - \$40k	0.6865553	0.3134447
\$40k - \$60k	0.6301676	0.3698324
\$80k - \$100k	0.4716599	0.5283401
\$100k - \$120k	0.4609929	0.5390071
\$120k - \$140k	0.3876812	0.6123188
>\$140k	0.4315068	0.5684932

The result is clearly that the percentage of having through-the-door ice on refrigerators increases as household income increases. With the high speed of the development of science and technology, traditional design method can no longer satisfy the quickly responding requirement of the refrigerator market. High-earning households require products with diversification and high-tech. So companies need to focus on technological innovation to serve higher-income groups, also they could add some discount on the products to those consumers with less income.

2.4 How does different household income affect the age of refrigerators

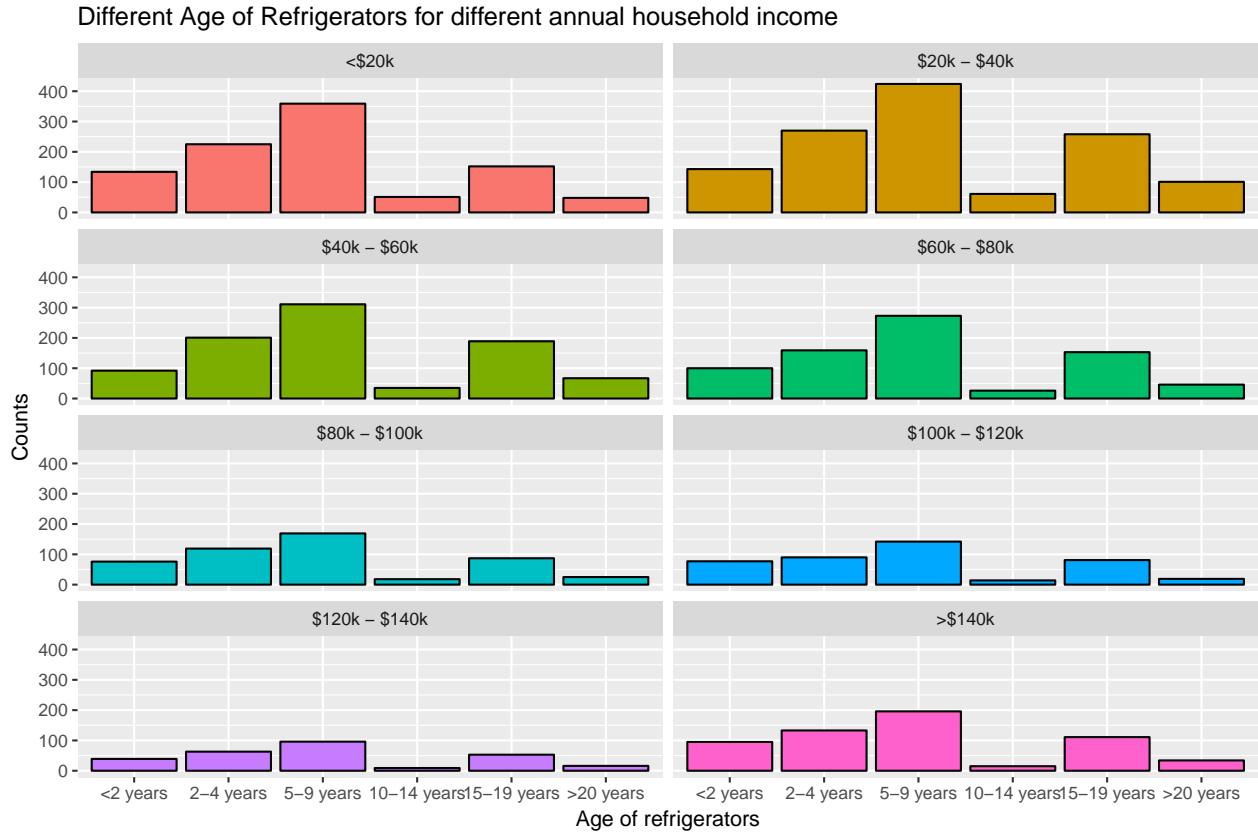


Table 13: The Proportion of different age of Refrigerators for each income level

	< 2 years	2~4 years	5~9 years	10~14 years	15~19 years	> 20 years
<\$20k	0.1382869	0.2321981	0.3704850	0.0526316	0.1568627	0.0495356
\$20k - \$40k	0.1137629	0.2147971	0.3373111	0.0485282	0.2052506	0.0803500
\$40k - \$60k	0.1027933	0.2245810	0.3474860	0.0391061	0.2111732	0.0748603
\$60k - \$80k	0.1321004	0.2100396	0.3606341	0.0343461	0.2021136	0.0607662
\$80k - \$100k	0.1538462	0.2408907	0.3421053	0.0364372	0.1761134	0.0506073
\$100k - \$120k	0.1820331	0.2127660	0.3356974	0.0330969	0.1914894	0.0449173
\$120k - \$140k	0.1413043	0.2282609	0.3478261	0.0326087	0.1920290	0.0579710
>\$140k	0.1626712	0.2277397	0.3356164	0.0256849	0.1900685	0.0582192

From the figure above, we can see a similar pattern in different household income, which is, most people prefer using a refrigerator for 5~9 years regardless of income. The percent of age (5~9 years) in different income groups maintain around 35%, and The percent of age (2~4 years) is about 22%. Hence, more than

half of households is willing to change their refrigerators in the range of 2 to 9 years.

3 How factors affect purchasing energy star qualified refrigerator

3.1 Expolore the relation between whether people would purchase energy star qualified refrigerator and where they live.

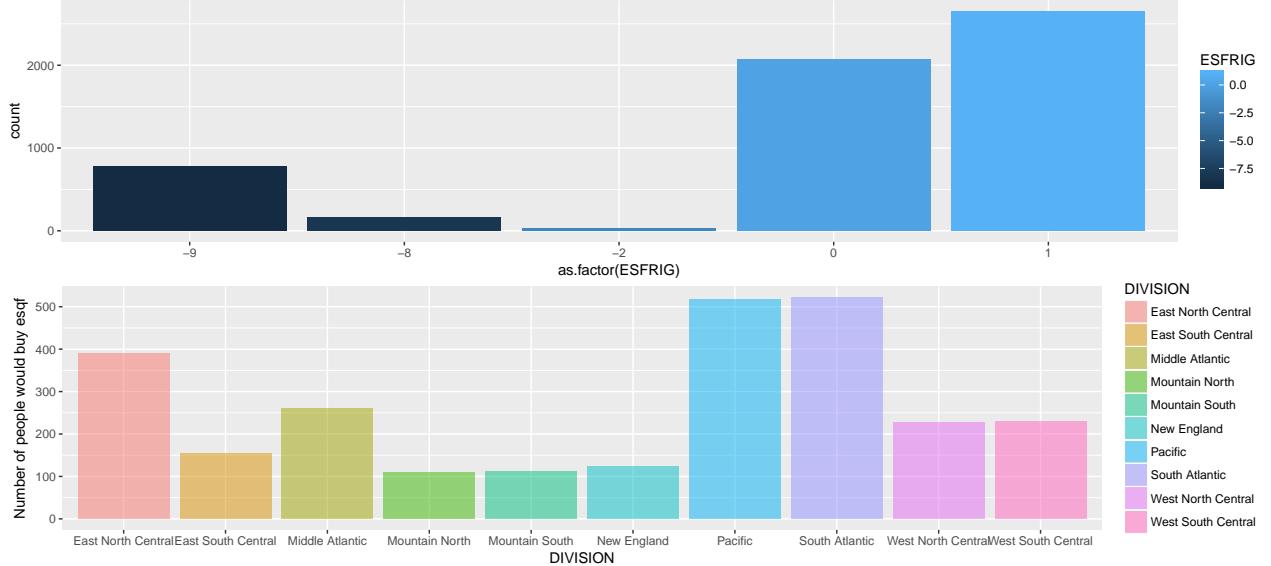
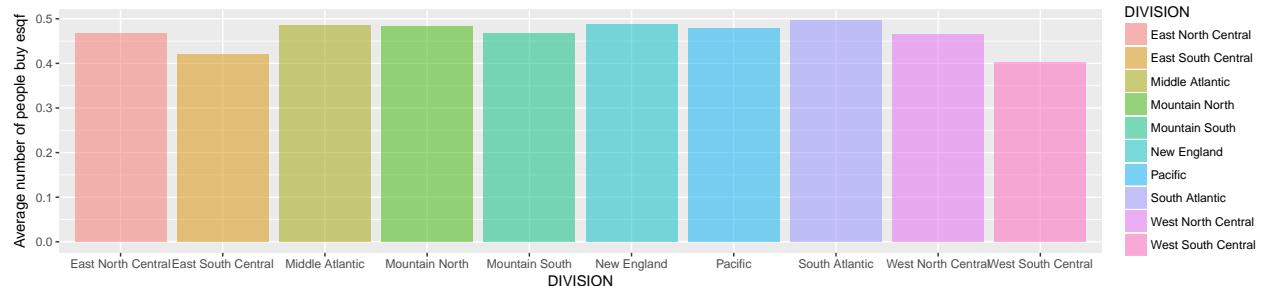


Table 14: anova test of DIVISION and ESFRIG

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DIVISION	1	118.0734	118.07344	9.567619	0.00199
Residuals	5684	70145.9202	12.34094	NA	NA



Firstly, we plot a histogram to summarise the responses of whether they would use energy star refrigerator. The histogram indicates that the answer of lots of people are yes, so it is meaningful to explore more details of it. It is reasonable that region is a factor on the selling of energy star refrigerator. First, we take a look at the number of people would buy energy star refrigerator in different divisons. From the plot above we can see that South Atlantic and Pacific have the greatest number of people would like to purchase energy star qualified refrigerator while the Mountain South, Mountain North and New England have the smallest. Base on the common sense that the temperature in South Atalntic and Pacific are great than other area, so they prefer purchasing the energy star qualified refrigerator to save more money. Then we can conclude that the marketing people should focus more on South Atlantic and Pacific than other area. Secondly, we plot the average number of people would buy energy star refrigerator in different division. And the result

indicates that there are no significant difference between each division. To test our conclusion, we use anova to test whether there is a significant relation between the number of people would like to buy energy star qualified refrigerator and the area they live. The p-value is 0.003 here so we can conclude that division does effect the number of energy star refrigerator.

3.2 The relation between whether people would purchase energy star qualified refrigerator and their income.

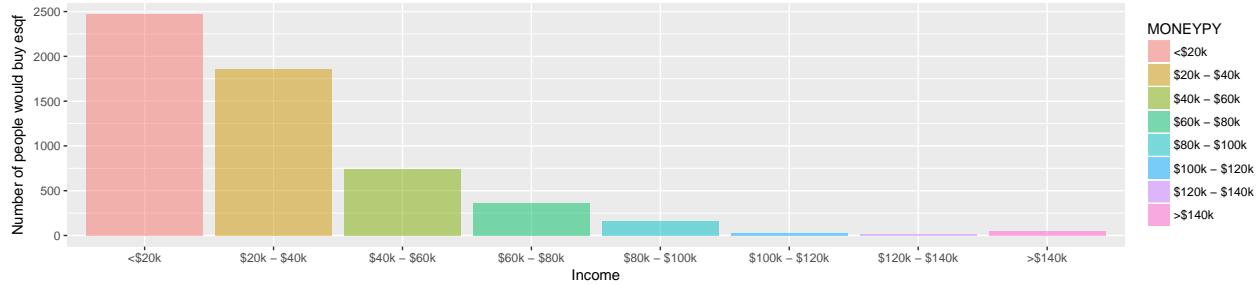
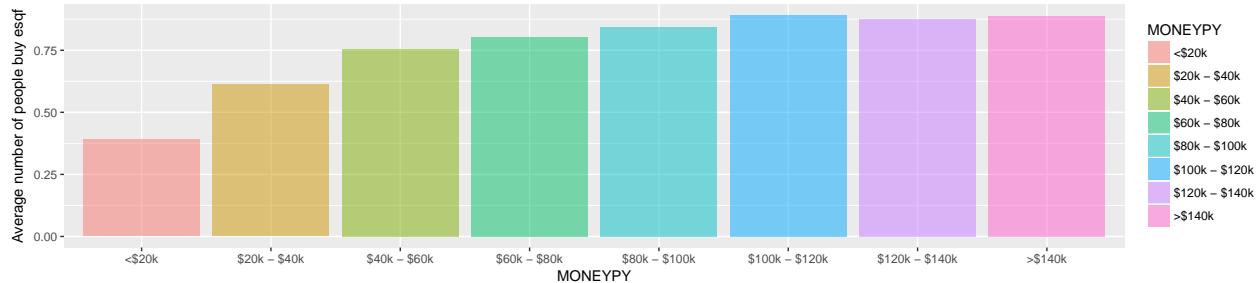


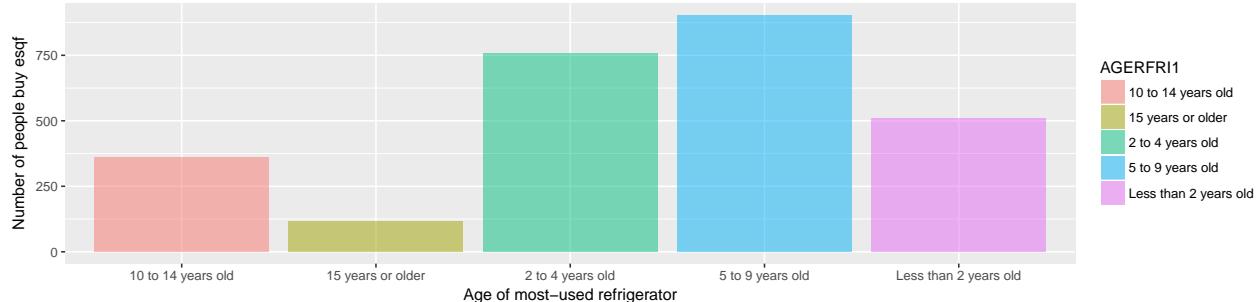
Table 15: anova test of MONEYPY and ESFRIG

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
MONEYPY	7	4056.946	579.56370	49.70411	0
Residuals	5678	66207.048	11.66028	NA	NA



From the plot above we can see that the household with average annual gross income that more than \$100,000 prefer the energy star qualified refrigerator than other people. The possible reason might be they can afford the energy star qualified refrigerator and want to save some money through the refrigerator. Therefore, one useful marketing suggestion is to advertise the energy star refrigerators to high-income people, and at the same time let the relative low income people know more about the energy star qualified refrigerator so they may buy it in the future. To test our conclusion, we use anova to test whether there is a significant relation between income and purchasing energy star qualified refrigerator. From the table above, the p-value indicates a significant relation between the two variables.

3.3 The relation between age of most-used refrigerator and whether people would buy energy star qualified refrigerator.



The plot above shows that the household have age of most-used refrigerator between 5 to 9 years old prefer change their current refrigerators to energy star qualified refrigerators. One reasonable reason is that they have already used their current refrigerators for a while and want to change a energy star qualified refrigerator to save more money in the future. Therefore, the marketing people should focus more on those people in energy star qualified refrigerator selling.

4 How factors affect purchasing recycling refrigerators

4.1 The relation between receiving recycling refrigerators and where they live

Sometimes some people receiving free recycling of old refrigerator instead of purchasing new refrigerator. We want to explore the relation between using old refrigerator and where they live.

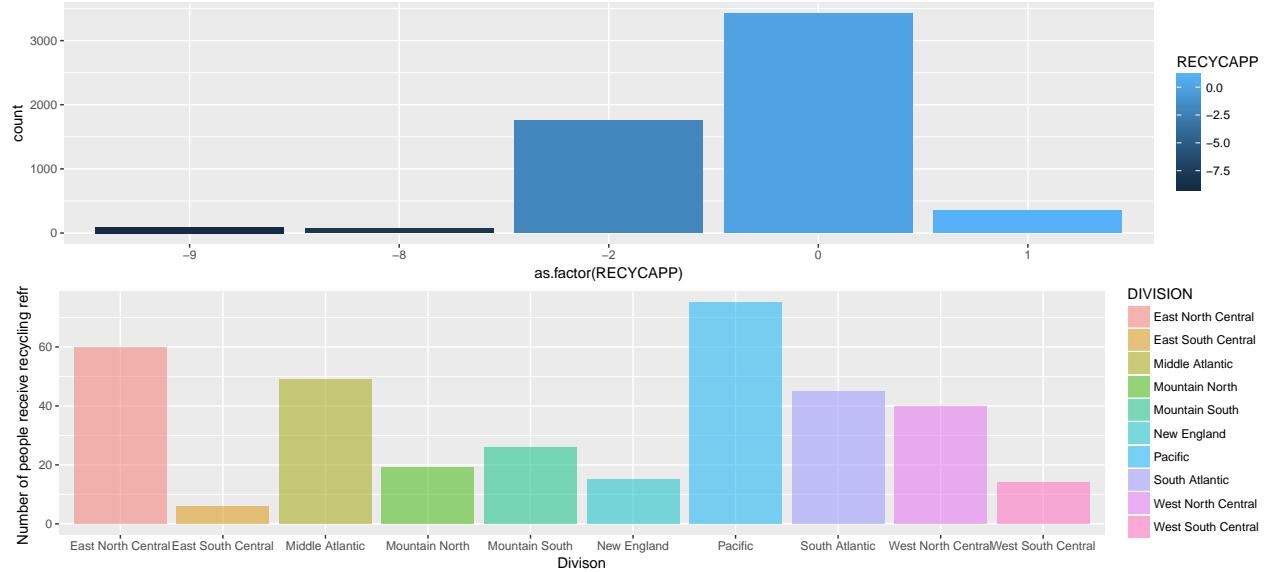
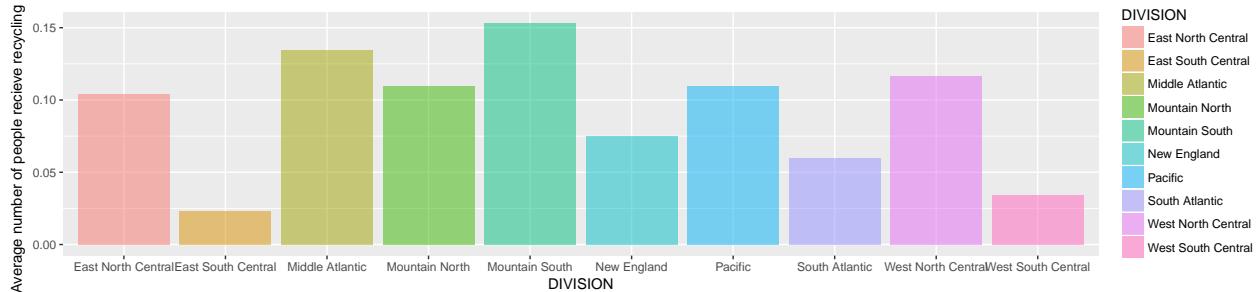


Table 16: anova test of RECYCAPP and DIVISION

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DIVISION	1	14.17879	14.178794	5.156433	0.0231977
Residuals	5684	15629.46120	2.749729	NA	NA



Firstly, let's get a first impression of roughly how many people would receive recycling refrigerator. According to histogram, some people receive old refrigerators so it is useful to explore more details. From the histogram above we see that people live in Pacific area prefer receiving old recycling refrigerator. A marketing suggestion is that a little discount to the refrigerator in Pacific so they may not receive old refrigerator but buy a new one instead. Using anova again to test our conclusion. The p-value here verifies that there is a significant relation between receiving old recycling refrigerator and where they live.

4.2 The relation between receiving recycling refrigerator and household income

Another factor that might be a reason of receiving recycling refrigerator is household income.

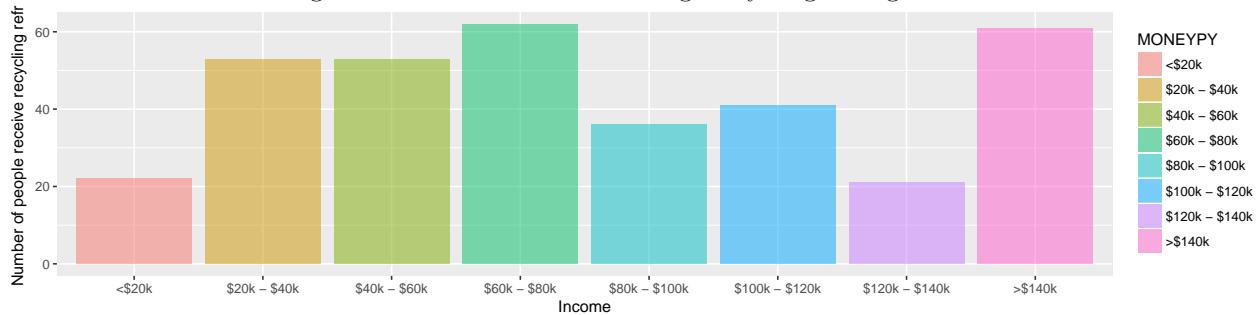
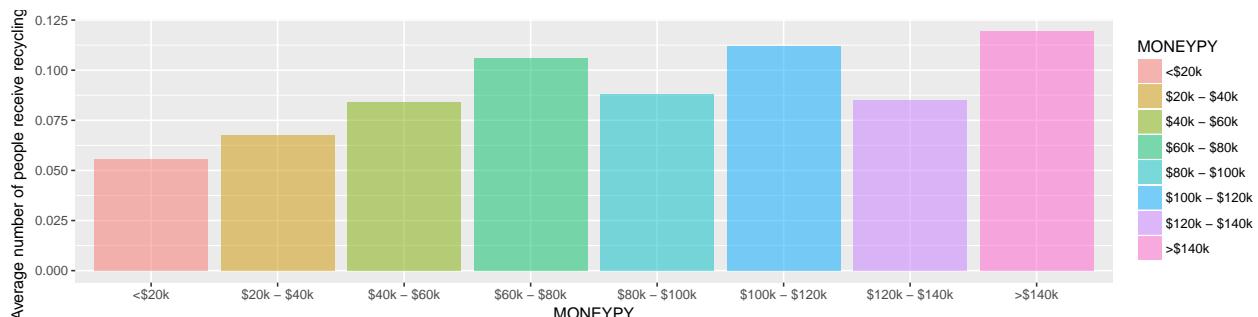


Table 17: anova test of MONEYPY and RECYCAPP

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
MONEYPY	7	784.1741	112.024874	42.8062	0
Residuals	5678	14859.4659	2.617025	NA	NA



According to the plots, households with income less than \$20,000 would not receive recycling refrigerator. One of the reasons might be they do not want to spend money on old refrigerators. The marketing suggestion is that encouraging high-income household to buy new refrigerator and give them discount on refrigerators. To test our conclusion, use anova again to test whether there is a significant relation between income and receiving old recycling refrigerator. The p-value implies there is a significant relation between income and

receiving recycling refrigerator.

5 Statistics model for refrigerator data prediction and variable selection for significant factors

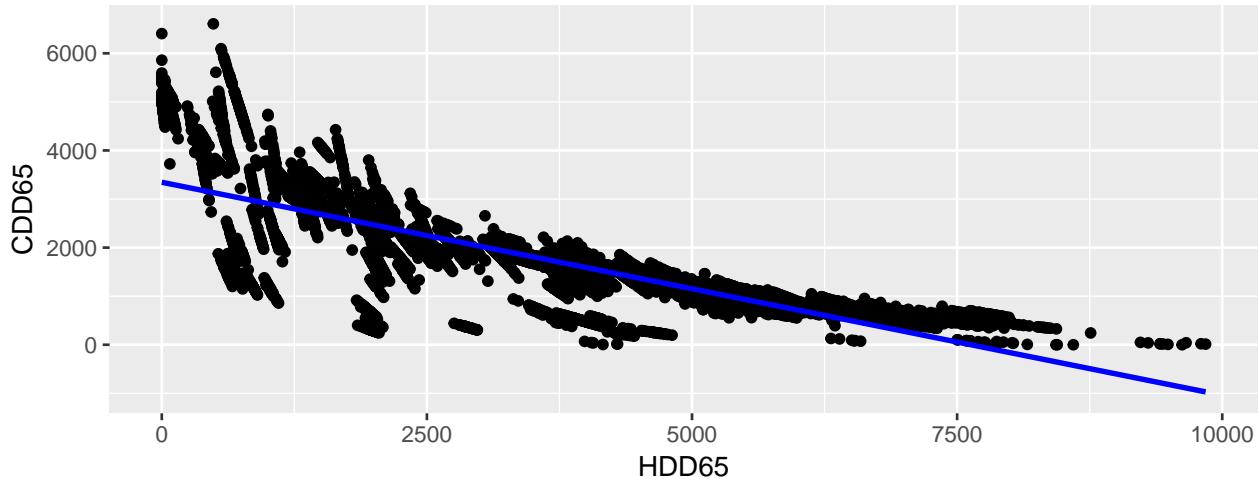
5.1 Create new variable

As we could observe from the number of refrigerator, over 90% of people have either 1 or 2 refrigerator. Because of such feature, new variable ‘numF’ is created in order to apply logistic regression. For people who have multiple refrigerator (more than 1 refrigerator), value ‘1’ will be assign for variable ‘numF’. For people who have 1 or less refrigerator, value ‘0’ will be assign for variable ‘numF’.

5.2 Degree day explantion

Degree day is a quantitative index demonstrated to reflect demand for energy to heat or cool houses and businesses. This index is derived from daily temperature observations at nearly 200 major weather stations in the contiguous United States. The “heating year” during which heating degree days are accumulated extends from July 1st to June 30th and the “cooling year” during which cooling degree data are accumulated extends from January 1st to December 31st. A mean daily temperature (average of the daily maximum and minimum temperatures) of 65°F is the base for both heating and cooling degree day computations. Heating degree days are summations of negative differences between the mean daily temperature and the 65°F base; cooling degree days are summations of positive differences from the same base. For example, cooling degree days for a station with daily mean temperatures during a seven-day period of 67,65,70,74,78,65 and 68, are 2,0,5,9,13,0, and 3, for a total for the week of 32 cooling degree days.

Relation between cooling and heating degree day



As the explantion of the degree day, there should be a correlation exists bewtween heating and cooling degree day. And from the result above, there is a strong negative correlation (-0.7885685) bewtween heating and cooling degree day. To avoid the collinearity, only one of the degreee day variable will be use in the following model.

5.3 Overall Statistic, Model Selection and Model Evaluation

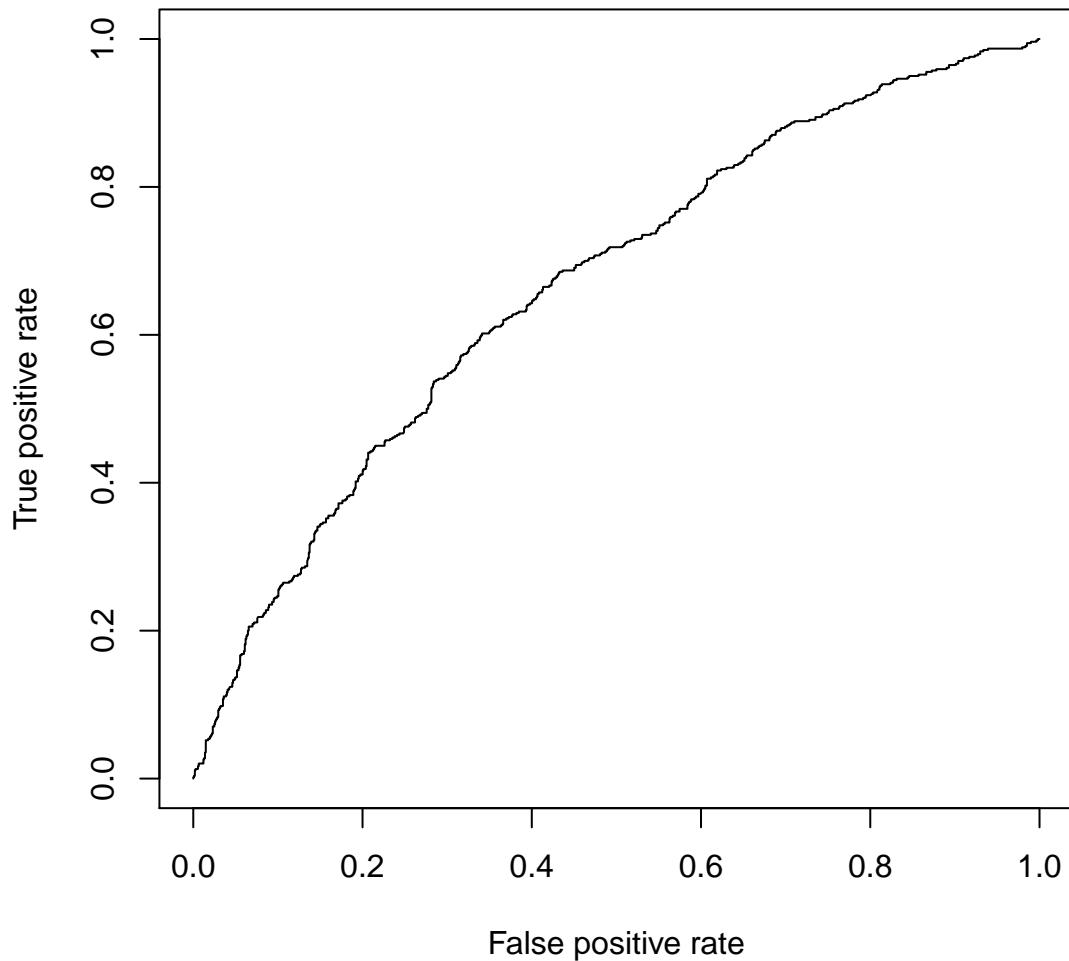
5.3.1 Number of refrigerator

```
selection<-step(numR)

## Start: AIC=4696.22
## numF ~ HDD65 + CDD65 + as.factor(CLIMATE_REGION_PUB) + as.factor(MONEYPY) +
##       as.factor(REGIONNC) + as.factor(DIVISION) + as.factor(METROMICRO)
##
##
## Step: AIC=4696.22
## numF ~ HDD65 + CDD65 + as.factor(CLIMATE_REGION_PUB) + as.factor(MONEYPY) +
##       as.factor(DIVISION) + as.factor(METROMICRO)
##
##                               Df Deviance   AIC
## - as.factor(METROMICRO)    2   4644.7 4692.7
## - CDD65                   1   4644.7 4694.7
## <none>                     4644.2 4696.2
## - as.factor(DIVISION)      9   4666.8 4700.8
## - HDD65                   1   4653.3 4703.3
## - as.factor(CLIMATE_REGION_PUB) 5   4664.8 4706.8
## - as.factor(MONEYPY)       7   4966.5 5004.5
##
## Step: AIC=4692.73
## numF ~ HDD65 + CDD65 + as.factor(CLIMATE_REGION_PUB) + as.factor(MONEYPY) +
##       as.factor(DIVISION)
##
##                               Df Deviance   AIC
## - CDD65                   1   4645.3 4691.3
## <none>                     4644.7 4692.7
## - as.factor(DIVISION)      9   4667.7 4697.7
## - HDD65                   1   4654.6 4700.6
## - as.factor(CLIMATE_REGION_PUB) 5   4665.1 4703.1
## - as.factor(MONEYPY)       7   4968.1 5002.1
##
## Step: AIC=4691.26
## numF ~ HDD65 + as.factor(CLIMATE_REGION_PUB) + as.factor(MONEYPY) +
##       as.factor(DIVISION)
##
##                               Df Deviance   AIC
## <none>                     4645.3 4691.3
## - as.factor(DIVISION)      9   4667.9 4695.9
## - HDD65                   1   4656.7 4700.7
## - as.factor(CLIMATE_REGION_PUB) 5   4665.4 4701.4
## - as.factor(MONEYPY)       7   4968.1 5000.1
```

In order to decide the significant variables that effect the number of refrigerator, our group did a binary logistic regression and variable selection based on the model AIC. Our result indicate that variable heating degree day (HDD65), Climate Zone (CLIMATE_REGION_PUB), Income (MONEYPY), Division (DIVISION) are significant factor for the number of refrigerator model.

ROC curve for logistics model



We also did a prediction by using ROC curve and AUC method, based on the result above, our final model has Prediction Accuracy 0.664.

5.3.2 Size of refrigerator

In order to decide the significant variables that effect the size of refrigerator, our group did a Ordinal Logistic Regression and variable selection based on the model AIC. Our result indicate that variable heating degree day (HDD65), Climate Zone (CLIMATE_REGION_PUB), Income (MONEYPPY), Division (DIVISION) are significant factor for the size of refrigerator model.

5.3.3 Age of refrigerator

In order to decide the significant variables that effect the age of refrigerator, our group did a Ordinal Logistic Regression and variable selection based on the model AIC. Our result indicate that variable heating degree day (HDD65), Climate Zone (CLIMATE_REGION_PUB), Income (MONEYPPY), Division (DIVISION) and Housing unit in Census Metropolitan or Micropolitan Statistical Area (METROMICRO), are significant factor for the age of refrigerator model.

Conclusion

In conclusion, marketing department should focus on city with large population. Although region characteristics do affect average number of refrigerator, the larger the population is, the more refrigerator people will buy and the more old refrigerator they will replace. Beside, people in metropolitain part tend to have higher income and households with higher income have more average of refrigerators. Additionally, all parts of people from US regardless of income are likely buy medium or large size of refrigerator. But the percentage of large refrigerators is increasing as income increases. Furthermore, households with higher income enjoy using refrigerators equipped with high-tech feature (through-the-door ice). In exploring purchasing energy star qualified refrigerator, we found out that people live in South Atlantic and Pacific prefer purchasing energy star qualified refrigerators. The marketing suggestion aim at this is to focus more on people live in South Atlantic and Pacific in energy star qualified refrigerators. Moreover, we believe the households with average annual gross more than \$100,000 prefer purchasing the energy star qualified refrigerator. Therefore, the suggestion is to advertise the energy star refrigerator to high-income people, and let the relative low-income people know more about this product so they may buy it in the future. Furthermore, households have the most-used refrigerator with age between 5 to 9 years old prefer changing their current refrigerator to an energy starred one. Therefore, the marketing people should focus more on those households. In analyzing the what may cause whether receiving recycling refrigerator, we find that people live in Pacific area prefer receiving recycling refrigerator. Therefore, a little discount can be applied in these area to encourage people buy a new one. Finally, households with income less than \$20000 would not receive recycling refrigerators. The reasonable suggestion is give them discount and thus encourage them to buy a new refrigerator.