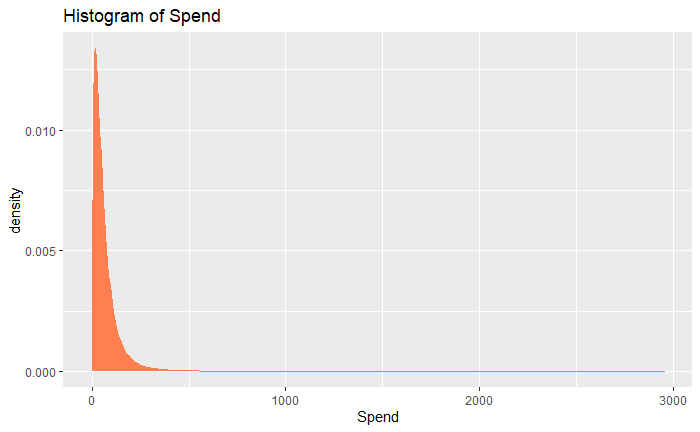
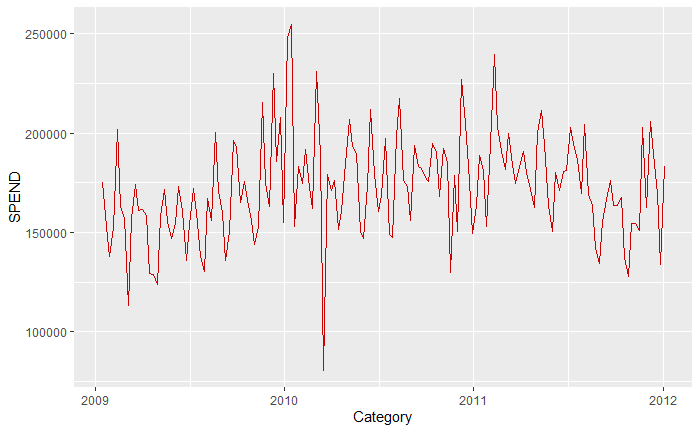
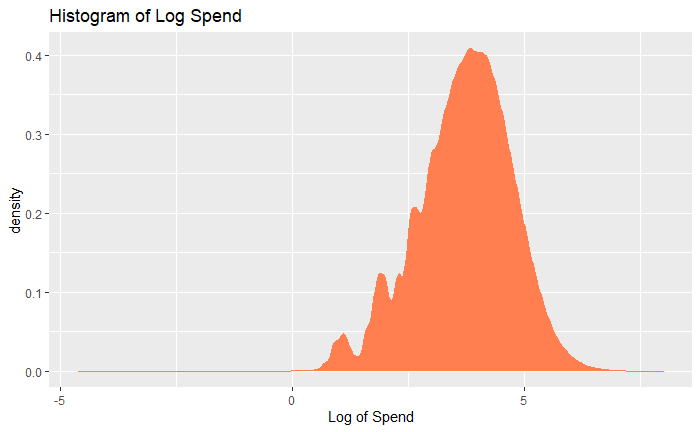
Assignment 8- Retail Chain

Submitted by: Hammad Muniem

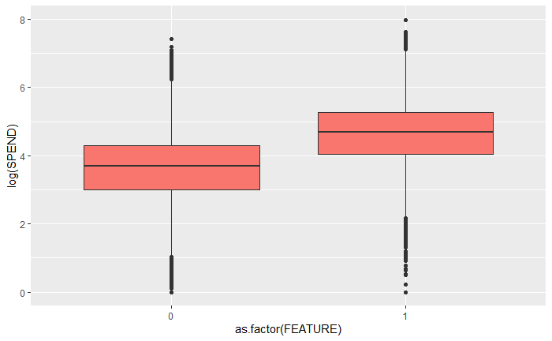
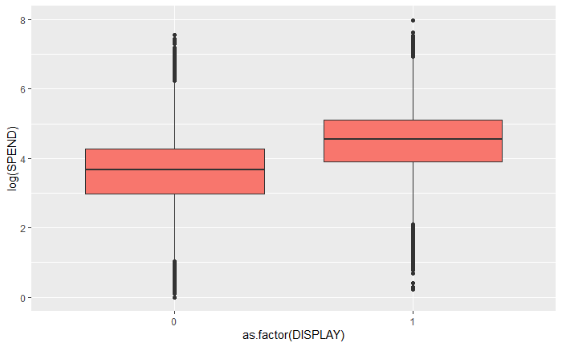
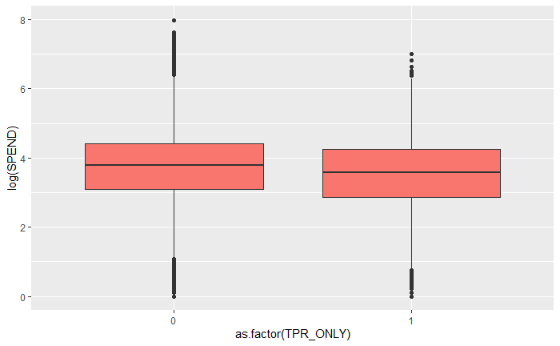
**Objective:** The objective of this exercise is to check the effectiveness of different marketing promotions and guaging the factors that effect sales, units and HHS.

**Exploratory Data Analysis:**

**SPEND**



We can see that there is some seasonality in the spend data so that must be incorporated. The distribution suggests that modeling log(spend) would yield better results since the distribution is relatively normal.



We can also see that the promotions seem to work in increasing spend except for TPR only but that needs more investigation.

**The same trend that is being observed in Spend is also observed in HHS and Units with relevant visualizations. They are not being posted here to reduce clutter.**

**The hypothesized effects are listed down for predicting all 3 (Spend, units, HHS):**

|  |  |  |
| --- | --- | --- |
| Predictor Variable | Expected sign of effect | Rationale |
| Feature | + | Since it was featured in the store circular, then we can guess that there might be some people walking in the store to purchase that specific item. So overall spend, units and HHS should be higher if a product is featured. |
| Display | + | Since it was featured in the store display, it would easily enter the consideration set of customers. So there are chances that overall spend, units and HHS should be higher if a product is featured. |
| TPR only | + | Since it was not advertised, Spend, units and HHS would not be higher than that of feature and display, but it should still be higher than not having the price reduction. |
| Year | + | As year progresses, chances are that the 3 variables would increase because of economic growth of the store and of the customers as well. |
| Month | Both | We expect some months like the holiday months to have higher spend, units and HHS because people would be shopping for that specific holiday. |
| Week of month | Both | We expect the 2nd week and 1st week of the month to have higher sales because of payment periods by companies would affect purchase power of the customer. |
| Segment | Both | We expect upscale stores to have higher spend, HHS and units because they have the perception of better quality and have higher prices. Moreover, people with more purchase power would be going there so their spending would probably be more. |
| Category | Both | We expect cold cereal category to have higher spend because compared to the other categories it is a breakfast item which makes it a bit more irreplaceable. |
| Discount | + | We expect discounted products to sell more because of customers perceiving more value while purchasing. This would have more effect on units and HHS but less on spend since spend is a function of price. |
| Price | - | We expect higher prices to reduce the number of units purchased because the same item might be available for cheaper at another store and since these are all snacking products, they might have competition in the snacking industry, so the customer would purchase something else. This would lower HHS and Spend as well. |
| State | Both | States might differ based on snacking habits. |
| Store name  (random ef) | Both | There would still be unexplained variation between the stores due to store location, service etc. that we do not have data for. |
| Product (random ef) | Both | There would still be unexplained variation between the products due to quality, taste etc. that we do not have data for. |

All Item level variables have not been included because the client wants an upper level analysis without any information on Items.  
Store variables such as Outlet\_size, Age of store and Outlet\_Year have not been included because the client wants store level information i.e. we have to include Outlet\_ID which includes size and age of the outlet.

Among the models posted, model 3 was the best according to anova tests and AIC.

**Type of outlet:**The type of outlet with the highest sales keeping everything else constant is Supermarket Type 3. Supermarket Type 3 will get 290% more sales than grocery store, 54% more sales than Supermarket Type 2 and 44% more sales than Supermarket Type 1 keeping all other factors constant.

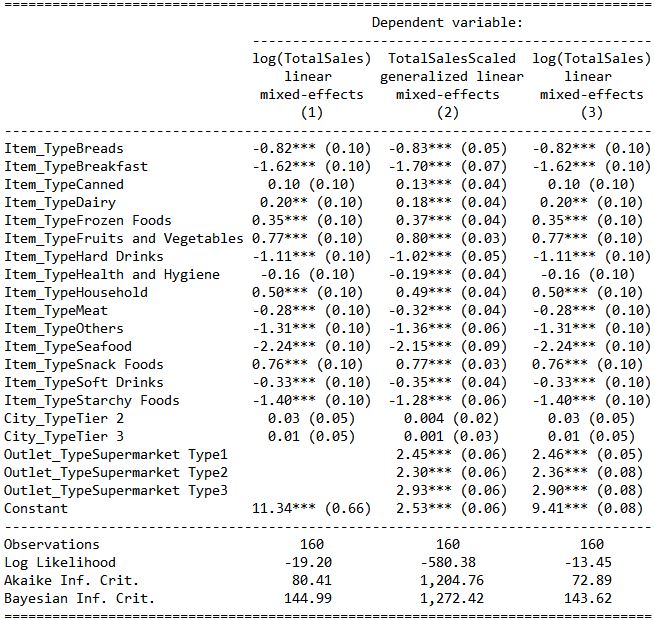
**City Tier:**City Tier 2 had the highest sales keeping all other factors constant. City Tier 2 had 3% higher sales than City Tier 1 and 2% higher sales than City Tier 3 keeping all other factors constant.

**Highest and lowest Performing Outlets:**At the store level, the economic significance is really low meaning that most of the variation has been accounted for by the other variables. Therefore, the variation caused by the lower level outlet itself is really low. However, I have still written down the best performing and least performing stores.  
The best performing outlets keeping all other factors constant were: (in order of best performing)

1. Outlet 35
2. Outlet 10
3. Outlet 49

The least performing outlets keeping all other factors constant were: (in order of least performing)

1. Outlet 13
2. Outlet 19
3. Outlet 17



**Recommendations:**

It is recommended that the entrepreneur invest in Supermarket Type 3 in city that is in City Tier 2. A bonus bit of advice would be stock up more on Snack foods and Fruits and Vegetables since they have the most sales as opposed to seafood which is expected to have the lowest sales.