Lab 2: Marketing & Fast Food Sales

Samuel Ha, Briana Hart, Nadia Tantsyura

Contents

1 Introduction

1.1 Motivation

The fast food market is very competitive, and chains and restaurants must continually release new menu items and market these to customers in order to stay relevant. Company X is one such fast food restaurant chain. This analysis is on Company X's latest new menu item and the promotion of that new menu item. Company X is unsure which of three marketing campaigns they should use to promote this new menu item. To understand which of the three promotions has the greatest effect on sales, the company ran A/B tests using three different marketing campaigns: Promotion 1, Promotion 2, and Promotion 3. Company X introduced the new item at different locations in randomly selected markets, with a different and randomly selected promotion used at each location. The weekly sales of the new item were recorded over the first four weeks of sales.

This analysis addresses the effect of each of the different promotions on sales of the new item, with the goal of determining which promotion leads to the greatest increase in sales. Company X's marketing team will use this information to devise a rollout strategy to optimize sales of the new item.

1.2 Research Question

Provided the goals of Company X and the data that was collected, our research question can be stated as the following:

Which promotion (1, 2, or 3) is the most effective for marketing the new menu item as measured by the total monthly sales?

2 Data & Methodology

2.1 About the data

For this research, we will be using the publicly available dataset from the IBM Watson Analytics community (https://www.kaggle.com/datasets/chebotinaa/fast-food-marketing-campaign-ab-test). There are a total of seven variables in the original dataset:

2.1.1 Outcome Variable

- SalesInThousands: This variable indicates sales amounts of the new item for a specific location, promotion, and week. Since we wanted to use cross-sectional data, this variable is not used as presented in the original dataset. Instead, we collapse the 4 observations per store (1 per week for a month) in the dataset into 1 observation and sum the sales into a total monthly value. The transformed monthly_sales variable serves as our outcome variable of interest. It should also be noted that the collapse of the 4 weekly observations per store into 1 month-long observation makes the unit of analysis stores without any time dependency.

2.1.2 Explanatory Variable

- Promotion: This is the primary causal variable and the main topic of our research. The possible values are 1, 2, and 3. For the analysis, we created dummy variables of promotion_1, promotion_2, and promotion_3 which contained a 1 or a 0 depending on the promotion implemented for each store.

2.1.3 Other Variables

- MarketID: MarketID uniquely identifies the store location's market. Based on the EDA, this variable appeared to explain some of the variation of the variation we observe in the distribution of sales. The metadata does not provide any additional information on this variable.
- MarketSize: This variable indicates the size of the market area by sales. The possible values are small, medium, and large. We are making the assumption that this variable is set prior to running this campaign, and it is based on the store's typical sales and not on sales of the new item. For the analysis, we created dummy variables of MarketSize_sm, MarketSize_md, and MarketSize_lg which contained a 1 or a 0 depending on the size of the market for each store.
- LocationID: This variable is the unique identifier for store location and thus each unit of analysis in the dataset.
- AgeOfStore: This variable presents the age of store in years. The values range from 1 year to 28 years. We consider this variable when evaluating the effect of the promotions on sales.
- Week: This variable indicates which of the four weeks of promotions the observation was collected. The possible values are 1, 2, 3, and 4, because the campaign was run over a span of 4 weeks. In our analysis, we combined the sales into the total sales over the month and drop this variable.

It is worth noting that we do not have any information on what fast food restaurant this is, what the new food item is, the area where these stores are located, or any other background information like what the promotions entailed or promotion delivery method. Due to the lack of information, we were forced to make assumptions based on our knowledge of other fast food chains. This is addressed later when we address the I.I.D. assumption.

2.2 Research Design

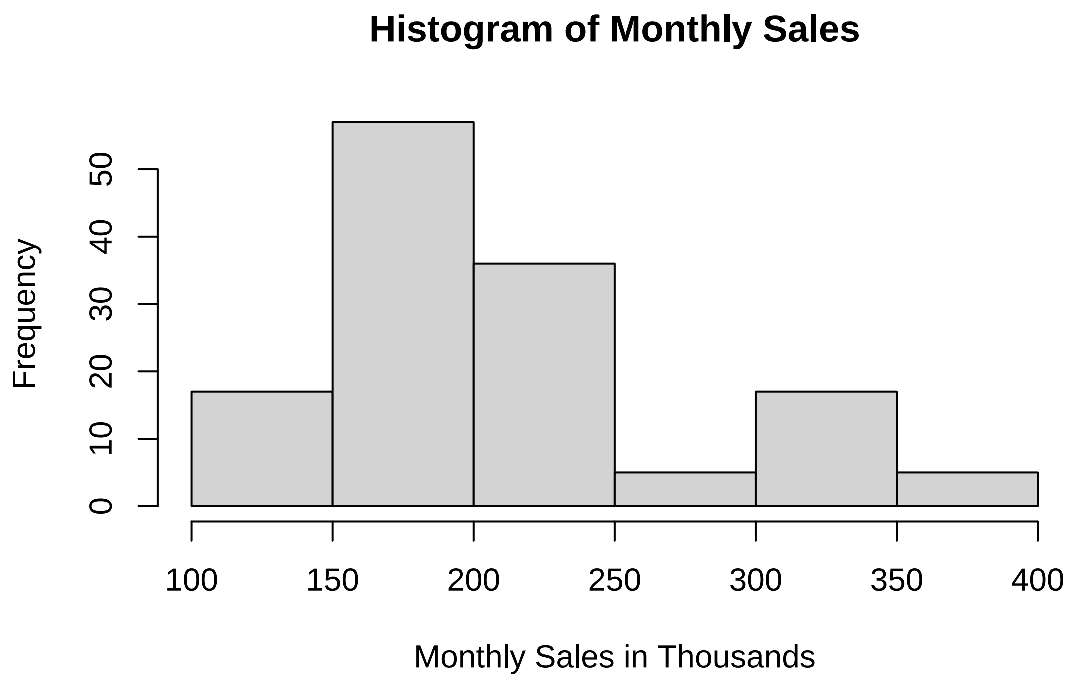
In this study, we aim to address the effect of three different promotions on the sales of a new item, in order to answer the research question: Which promotion (1, 2, or 3) is the most effective for marketing the new menu item as measured by the total monthly sales? In the study, promotion is the main variable of interest. In order to control for variables that may also affect monthly sales of the new item, we will include several additional explanatory variables and interaction terms. We will provide the additional contextual information in additional models. Given that we do not have any information on the promotions besides their ID number, we will not make any hypotheses on which promotion will have the greatest effect on sales.

Since the data comes from a true experiment, we will estimate one-equation structural models in order to infer a causal effect of the promotions. Given that there are only 137 stores in the dataset, we will use a Classic Linear Model to assess the effectiveness of the promotions. We will have to assess the CLM assumptions in order to justify the use of this model. This is addressed later in the section: CLM Assumptions.

When evaluating the data, we will combine the weekly sales across the four weeks during which the promotions were implemented into one row representing a month-long sales. The benefit of summing the weekly sales data is that it removes any time dependency and clusters by stores. This simplification is justified because the goal of the promotions was to increase sales overall, so by looking at the summation we see the total sales throughout the month. An additional benefit to this simplification is that it allows us to manipulate the dataset so there is only one row per store.

It should also be noted that all of the stores in the dataset participated in one of the three promotions, so we do not have a control group in this research design. Therefore, the only conclusions we can draw from this research is how the promotions impacted sales in comparison to each other. We cannot draw any findings about whether the effect of any of the promotions on monthly sales is different from the effect of no promotions on monthly sales.

The monthly sales (in thousands) have a bimodal distribution with a skew to the right, so it is not normally distributed. However, since we are using the Classic Linear Model a normal distribution is not required of the outcome or predictor variables–just of the errors. Nevertheless, the bimodal distribution is an indication that there may be other variables at work that are creating the two modes.



3 Modeling

In order to model the causal effect of the three different promotions of the new item on monthly sales of the new item, we plan to use monthly sales of the new item as the outcome variable and the promotion variable as the primary explanatory variable. We will consider the use of the age of the store, market within which the store is located, and size of the market area by sales as control variables in several model specifications.

The promotion variable contains an indication of which of the three promotions a store participated in. We transform each of the three categories into dummy variables and exclude one of the dummy variables to prevent perfect collinearity with the understanding that the excluded dummy variable will serve as the reference point to understand the coefficients of the dummy variables. Since there are three promotions, we are interested in the following comparisons of the promotions on monthly sales:

1. Promotion 1 vs Promotion 2
2. Promotion 1 vs Promotion 3
3. Promotion 2 vs Promotion 3

We will need to have at least two regression models to make all three of the above comparisons, since excluding only one of the promotions from a regression would only allow us to make two comparisons. For example, excluding Promotion 3 allows us to compare the estimated effect on monthly sales of Promotion 1 vs Promotion 3 and Promotion 2 vs Promotion 3 complete with t-tests to note whether the difference between the promotions was statistically significant, but not Promotion 1 vs Promotion 2. In order to make the Promotion 1 vs Promotion 2 comparison, we would also have to run a regression that excludes either

Promotion 1 or Promotion 2. As such, we will show two base models that make the above three comparisons, and choose one of the base models to serve as the restricted model onto which we add covariates for subsequent models.

In order to test whether heterogeneous effects on monthly sales exist between the promotion and covariates such as the age of the store and market size, we will include models that have interaction terms of promotion and age of store and promotion and market size.

3.1 First Model (Base)

In the base models, we include only the key variables for which we are trying to derive a model without any covariates. The outcome variable of the base model is monthly sales, and the primary explanatory variable is the promotion variable. As stated in the discussion above comparing the effect of three different promotions on monthly sales, we will need to have at least two base models in order to make the three possible combinations of comparisons between the three promotions.

The first base model will include the dummy variables for Promotion 1 and Promotion 2, leaving Promotion 3 as the reference point for the coefficients of Promotion 1 and 2. The first base model is specified in the following way and is run with cluster-robust standard errors:

$$\text{Model 1: MonthlySales} = \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion2}$$

3.2 Second Model (Base)

In the second base model, we again only include the key variables of the model without any covariates, but we include the dummy variables for Promotion 1 and Promotion 3, leaving Promotion 2 as the reference point for the coefficients of Promotion 1 and 3. This second model is assessed as follows and is run with cluster-robust standard errors:

$$\text{Model 2: MonthlySales} = \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3}$$

3.3 Third Model

For the third model, we settle on using Promotion 1 and 2 as the dummy variables for promotion because the second base model shows that the effect of Promotion 2 on monthly sales is significantly lower than that of Promotion 1 and 3, but we do not know whether Promotion 1 and 3 are different from each other. Excluding Promotion 3 allows us to make the comparison between Promotion 1 and 3 and understand whether they are significantly different from each other and retains a comparison against Promotion 2.

In addition to the Promotion 1 and 2 dummy variables, we add additional covariates related to the characteristics of stores which we think may affect monthly sales:

- Age of store: The age of the store is given in number of years it has been in operation. We note that the ages of the stores in the dataset have moderate positive skew, but while we could log-transform the variable to make it more normal, we do not because the interpretation of the coefficient for log of age is less intuitive than the interpretation of the coefficient of age untransformed as the predicted effect of one additional year on monthly sales. (eda visual)
- Marketing ID: There are 10 markets for the 137 stores which are only identified by their number. The model includes MarketID as a factor and drops one to prevent perfect colinearity.

The third model is as follows and is estimated with cluster-robust standard errors:

$$\begin{aligned}\text{Model 3: MonthlySales} &= \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3 } + \beta_3 \text{ AgeOfStore} \\ &+\beta_4 \text{ Market2 } + \beta_5 \text{ Market3 } + \beta_6 \text{ Market4 } + \beta_7 \text{ Market5 } + \beta_8 \text{ Market6} \\ &+\beta_9 \text{ Market7 } + \beta_{10} \text{ Market8 } + \beta_{11} \text{ Market9 } + \beta_{12} \text{ Market10}\end{aligned}$$

3.4 Third Model (Robust)

The third model (robust) is the same as the third model described above, but instead of using cluster-robust standard errors, we show the same regression with robust standard errors. We do this to demonstrate the lack of change in findings when using cluster-robust standard errors compared with using robust standard errors. In order to test how steady our findings are to different choices in standard errors, we actually ran all the models specifications with clustered standard errors and with rare exception, the significance levels were unchanged. However, for the sake of space, we only show one model with robust standard errors.

The third model with robust standard errors is shown below:

$$\begin{aligned}\text{Model 3 (Robust): MonthlySales} &= \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3 } + \beta_3 \text{ AgeOfStore} \\ &+\beta_4 \text{ Market2 } + \beta_5 \text{ Market3 } + \beta_6 \text{ Market4 } + \beta_7 \text{ Market5 } + \beta_8 \text{ Market6} \\ &+\beta_9 \text{ Market7 } + \beta_{10} \text{ Market8 } + \beta_{11} \text{ Market9 } + \beta_{12} \text{ Market10}\end{aligned}$$

3.5 Fourth Model

The fourth model contains all of the explanatory variables included in Model 3 and we interact the promotions with the age of the store which allows the effect of the different promotions to vary differently with regard to age. That is, interacting the terms allows each promotion to have a different slope with respect to age.

As such, the fourth model is specified in the following way and is estimated with cluster-robust standard errors:

$$\begin{aligned}\text{Model 4: MonthlySales} &= \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3 } + \beta_3 \text{ AgeOfStore} \\ &+\beta_4 \text{ Market2 } + \beta_5 \text{ Market3 } + \beta_6 \text{ Market4 } + \beta_7 \text{ Market5 } + \beta_8 \text{ Market6} \\ &+\beta_9 \text{ Market7 } + \beta_{10} \text{ Market8 } + \beta_{11} \text{ Market9 } + \beta_{12} \text{ Market10} \\ &+ \text{ Promotion1 } * \text{ AgeOfStore } + \text{ Promotion2 } * \text{ AgeOfStore}\end{aligned}$$

3.6 Fifth Model

The fifth model includes all of the explanatory variables included in Model 3, but also includes interactions between the promotion variable and the market size variable.

- Market size: The market size of the randomly selected market is categorized as either small, medium, or large. Since it is a categorical variable, we convert it into three dummy variables and exclude the small market size variable in our regression model.

Our sixth model is assessed as follows with cluster-robust standard errors:

$$\begin{gathered}\text{Model 5: MonthlySales} = \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3 } + \beta_3 \text{ AgeOfStore} \\ +\beta_4 \text{ Market2 } + \beta_5 \text{ Market3 } + \beta_6 \text{ Market4 } + \beta_7 \text{ Market5 } + \beta_8 \text{ Market6} \\ +\beta_9 \text{ Market7 } + \beta_{10} \text{ Market8 } + \beta_{11} \text{ Market9 } + \beta_{12} \text{ Market10} \\ + \text{ Promotion1 } * \text{ MarketSizeMd } + \text{ Promotion1 } * \text{ MarketSizeLg} \\ + \text{ Promotion2 } * \text{ MarketSizeMd } + \text{ Promotion2 } * \text{ MarketSizeLg}\end{gathered}$$

3.7 Cluster-Robust standard errors

Cluster-robust standard errors and robust standard errors are used to address non-independence between fast food stores. From the market ID and the market size variables, we know that fast food stores were organized by market and were considered to all be part of a market of a certain size, but we do not know whether there was geographic clustering or whether the stores were all managed by a regional branch which would make them more similar to one another. Given this lack of clarity on the nature of the clustering, we plan to run our models with cluster-robust standard errors as well as robust standard errors. By default cluster-robust standard errors will be shown unless results with robust standard errors are substantially different.

4 Results

Both the first and second base models indicate that the effect of Promotion 2 on monthly sales of the new item is significantly less than that of Promotion 1 and Promotion 3. As shown in the first model, Promotion 2 is predicted to result in $32,140 less in monthly sales of the new item than Promotion 3. The second model also shows that Promotion 2 is predicted to result in $32,140 less in monthly sales than Promotion 3, but in addition, the second model estimates that stores with Promotion 2 will have $43,078 less in sales than Promotion 1.

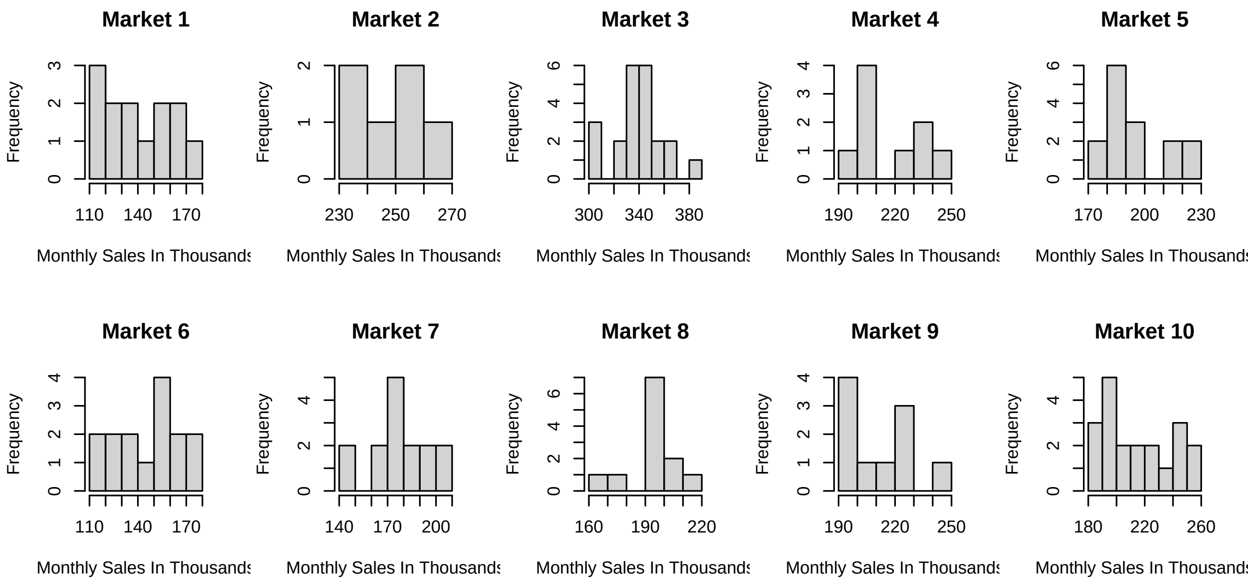
Given the two base models, Promotion 1 and 3 then are the leading contenders in promotions that may drive sales of the new item. However, in a direct comparison of Promotion 1 and Promotion 3 in the first model, a statistically significant difference in their effects on monthly sales is not observed.

The third model includes control variables about the store that may impact monthly sales: age of stores and dummy variables of market ID. With the inclusion of these control variables, we find the model estimates an effect size of Promotion 1 ($19,723) that is significantly greater than that of Promotion 3. In this third model, Promotion 3 continues to have a larger effect size than that of Promotion 2. Promotion 2 is estimated to be associated with $19,143 less in monthly sales than Promotion 3. Age of store is not shown to be a significant predictor of monthly sales, but belonging to a given market does. Each of the dummy variables for market ID included in the model is estimated to be a significant predictor of monthly sales as shown in the table above. It should also be noted that the inclusion of the market ID and age results in a large increase in the Adjusted R^2 value, from 0.067 in the base models to 0.974, which means that the model explains 97.4 percent of the total variance of monthly sales. This finding is consistent with our data exploration of monthly sales.

Previously, we found that monthly sales had a bimodal distribution which suggested that other variables may be generating the two modes. In visualizing monthly sales by market ID, we found that while all markets had approximately normally distributed sales, their means varied widely, as shown below. In particular, Market 3, which contains 22 of the 137 stores in the dataset, had high monthly sales. Therefore, both the significance and increase in adjusted R^2 after the inclusion of market ID as a control variable is consistent with our visualizations.

Table 1: Relationship between Monthly Sales and Promotion

	Dependent variable:				
	monthly_sales				
	(1)	(2)	(3)	(4)	(5)
promotion_1	10.938	43.078^{***}	19.723^{***}	26.651^{***}	22.105^{***}
	(6.672)	(5.414)	(2.223)	(3.311)	(3.779)
promotion_2	-32.140^{***}		-19.143^{***}	-10.406^{*}	-9.044^{**}
	(9.286)		(1.856)	(4.614)	(3.121)
promotion_3		32.140^{***}			
		(9.286)			
AgeOfStore			0.051	0.662^{*}	0.071
			(0.159)	(0.325)	(0.154)
factor(MarketID)2			103.277^{***}	101.604^{***}	102.561^{***}
			(1.148)	(0.944)	(1.894)
factor(MarketID)3			198.700^{***}	199.014^{***}	198.614^{***}
			(0.351)	(0.563)	(2.040)
factor(MarketID)4			77.568^{***}	77.704^{***}	71.809^{***}
			(0.206)	(0.372)	(3.591)
factor(MarketID)5			62.667^{***}	62.874^{***}	62.115^{***}
			(0.772)	(0.775)	(0.767)
factor(MarketID)6			6.508^{***}	5.748^{***}	6.423^{***}
			(0.088)	(0.390)	(0.138)
factor(MarketID)7			37.592^{***}	37.553^{***}	37.541^{***}
			(0.390)	(0.356)	(0.563)
factor(MarketID)8			50.428^{***}	48.808^{***}	50.620^{***}
			(0.373)	(0.665)	(0.467)
factor(MarketID)9			69.518^{***}	68.523^{***}	69.581^{***}
			(0.457)	(0.441)	(0.611)
factor(MarketID)10			77.650^{***}	77.338^{***}	78.249^{***}
			(0.131)	(0.400)	(2.095)
promotion_1:AgeOfStore				-0.778^{**}	
				(0.295)	
promotion_2:AgeOfStore				-1.033^{*}	
				(0.419)	
MarketSize_md					
MarketSize_lg					
promotion_1:MarketSize_md					-3.881
					(4.800)
promotion_1:MarketSize_lg					0.189
					(4.852)
promotion_2:MarketSize_md					-9.129^{*}
					(4.075)
promotion_2:MarketSize_lg					-13.455^{***}
					(3.012)
Constant	221.458^{***}	189.318^{***}	139.760^{***}	134.585^{***}	139.799^{***}
	(26.975)	(21.065)	(1.832)	(3.383)	(2.317)
Observations	137	137	137	137	137
R^2	0.080	0.080	0.976	0.978	0.977
Adjusted R^2	0.067	0.067	0.974	0.976	0.974
Residual Std. Error	62.586 (df = 134)	62.586 (df = 134)	10.401 (df = 124)	10.034 (df = 122)	10.378 (df = 120)
Note:					*p<0.05; **p<0.01; ***p<0.001



Controlling for the effects of belonging to a market and the age of the store, the effect size of Promotion 1 compared to that of Promotion 3 increases from \$10,938 (in the first model) to \$19,723, and the effect size of Promotion 2 compared to that of Promotion 3 shrinks from -\$32,140 (in the first model) to -\$19,143.

We also run the third model with robust standard errors instead of cluster-robust standard errors. In the table below, the first column presents Model 3 with cluster-robust standard errors and the second column presents Model 3 with robust standard errors. The effect sizes remain the same, as would be expected for a recalculation with different standard errors, and while the standard errors change, the significance levels reached by the coefficient estimates ($p < 0.05$, $p < 0.01$, and $p < 0.001$) remain largely the same with only Market 6 changing from a significant predictor of monthly sales in the cluster-robust standard errors model to an insignificant predictor in the robust standard errors model.

The fourth model includes all the variables in the third model and includes interaction terms of promotion and the age of the store to allow heterogeneous effects of age on promotion. In this specification, the age of store variable and the interaction terms of promotion with age are statistically significant, suggesting that the slope (change of monthly sales/change in age) for the different promotions are different from one another. Both coefficients for the interaction term are negative suggesting stores with promotion 1 and 2 are estimated to make fewer sales of the new item the older in age they are. Including these interaction terms increases the effect size of Promotion 1 compared to Promotion 3 from \$19,723 (in the third model) to \$26,651 and the finding remains significant.

The fifth model has all the predictor variables of the third model and includes interaction terms of promotion and the size of the market. The interaction terms are not statistically significant for promotion 1, but the interaction terms between promotion 2 and market size are statistically significant. The former finding suggests that the effect of the promotion on monthly sales does not vary by market size in stores that had promotion 1. The effect size of Promotion 1 compared to Promotion 3 from \$19,723 (in the third model) increased to \$22,105.

Overall, we find that the greater effect size of Promotion 1 on monthly sales of the new item compared to other promotions is robust to various model specifications and the findings hold with both robust and clustered standard errors. In order to determine whether the inclusion of the interaction terms produces better predictions of monthly sales, we run two F-tests to compare Model 3 with Model 4 and Model 5. The null hypothesis of the F-test is that the restricted model (Model 3) does as well explaining the outcome variable as the unrestricted model (Models 4 and 5). We find that the null hypothesis is rejected for Model 4, but not for Model 5. The F-test provides evidence that Model 4 explaining the variance in monthly sales significantly better than Model 3. Given the findings of the F-test, we select Model 4 as our final model.

Table 2: Model 3 with Clustered Standard Errors and Robust Standard Errors

	Dependent variable:	
	monthly_sales	
	(1)	(2)
promotion_1	19.723***	19.723***
	(2.223)	(2.337)
promotion_2	−19.143***	−19.143***
	(1.856)	(2.571)
AgeOfStore	0.051	0.051
	(0.159)	(0.171)
factor(MarketID)2	103.277***	103.277***
	(1.148)	(5.143)
factor(MarketID)3	198.700***	198.700***
	(0.351)	(3.536)
factor(MarketID)4	77.568***	77.568***
	(0.206)	(3.871)
factor(MarketID)5	62.667***	62.667***
	(0.772)	(3.654)
factor(MarketID)6	6.508***	6.508
	(0.088)	(4.636)
factor(MarketID)7	37.592***	37.592***
	(0.390)	(3.752)
factor(MarketID)8	50.428***	50.428***
	(0.373)	(3.623)
factor(MarketID)9	69.518***	69.518***
	(0.457)	(5.472)
factor(MarketID)10	77.650***	77.650***
	(0.131)	(3.518)
Constant	139.760***	139.760***
	(1.832)	(3.233)
Observations	137	137
R^2	0.976	0.976
Adjusted R^2	0.974	0.974
Residual Std. Error (df = 124)	10.401	10.401
Note:	*p<0.05; **p<0.01; ***p<0.001	

The final model selected is specified in the following way:

$$
\begin{aligned}
\text{Model 4: MonthlySales} = \beta_0 + \beta_1 \text{ Promotion1 } + \beta_2 \text{ Promotion3 } + \beta_3 \text{ AgeOfStore} \\
+\beta_4 \text{ Market2 } + \beta_5 \text{ Market3 } + \beta_6 \text{ Market4 } + \beta_7 \text{ Market5 } + \beta_8 \text{ Market6} \\
+\beta_9 \text{ Market7 } + \beta_{10} \text{ Market8 } + \beta_{11} \text{ Market9 } + \beta_{12} \text{ Market10} \\
+ \text{ Promotion1 } * \text{ AgeOfStore } + \text{ Promotion2 } * \text{ AgeOfStore}
\end{aligned}
$$

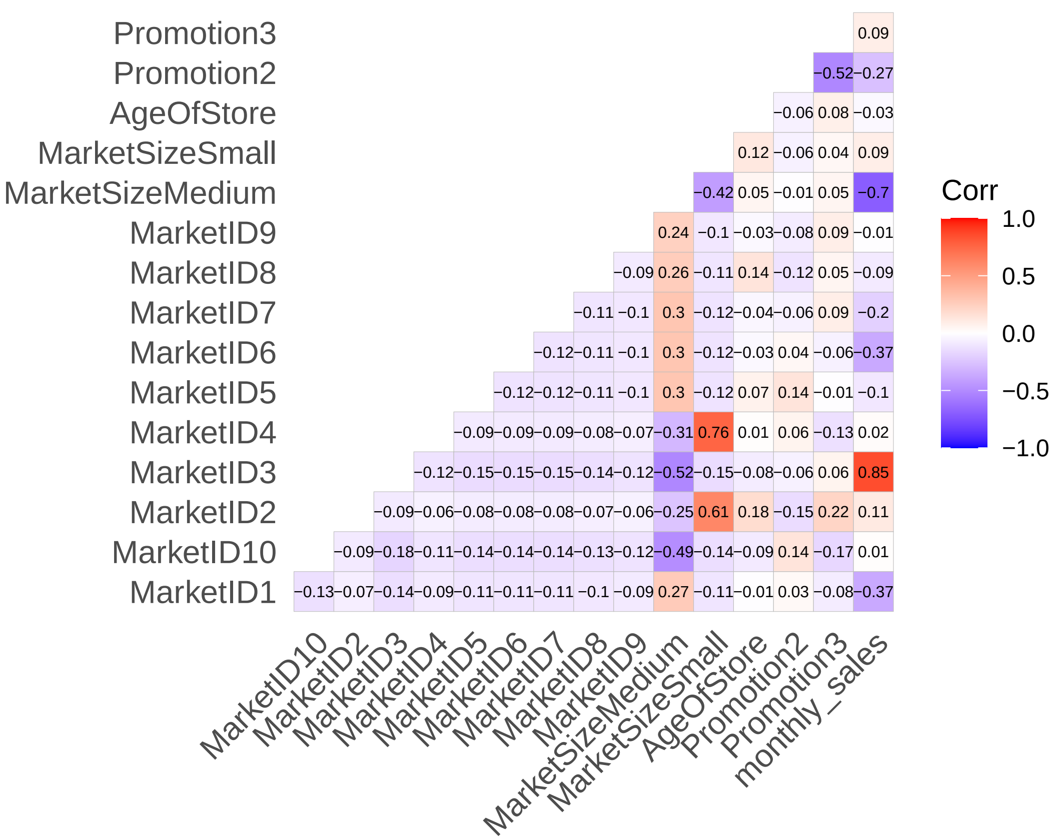
5 CLM Assumptions

5.1 IID Data:

In assessing the I.I.D. assumption, we look to see if the data is idendependent and identically distributed. We do not have background information on the selection process but the dataset is described as coming from an A/B test. Since A/B tests require randomized assignment of a variant, we think it is a reasonably safe assumption that stores in selected market were randomly selected to implement one of the three promotions. Therefore, we can say that this model meets the I.I.D. assumption.

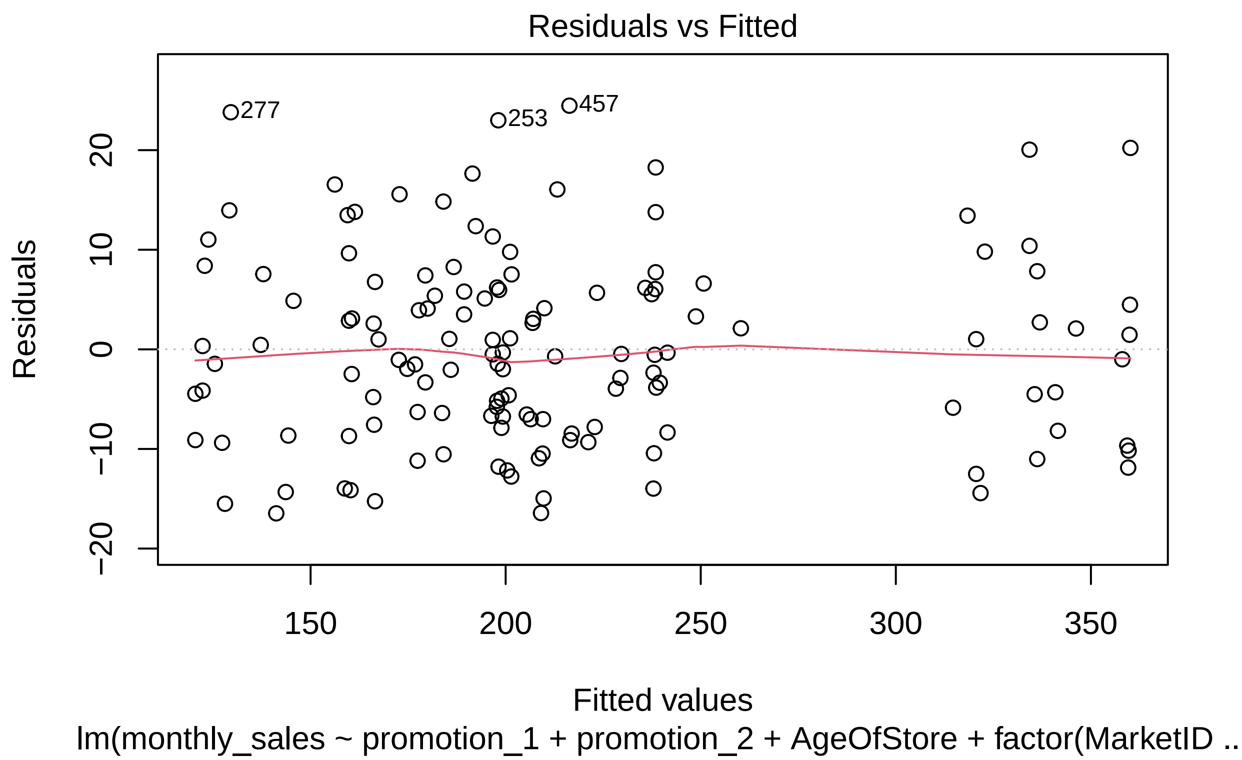
5.2 No Perfect Colinearity:

When assessing perfect collinearity, we are looking to see if any of the inputs can be written as a linear combination of another input. That is, we are looking to see if any of the inputs are fully encompassed by another input. To do so, we evaluate a correlation plot. In the plot below, we can see that none of the variables are a perfect linear combination of any others.



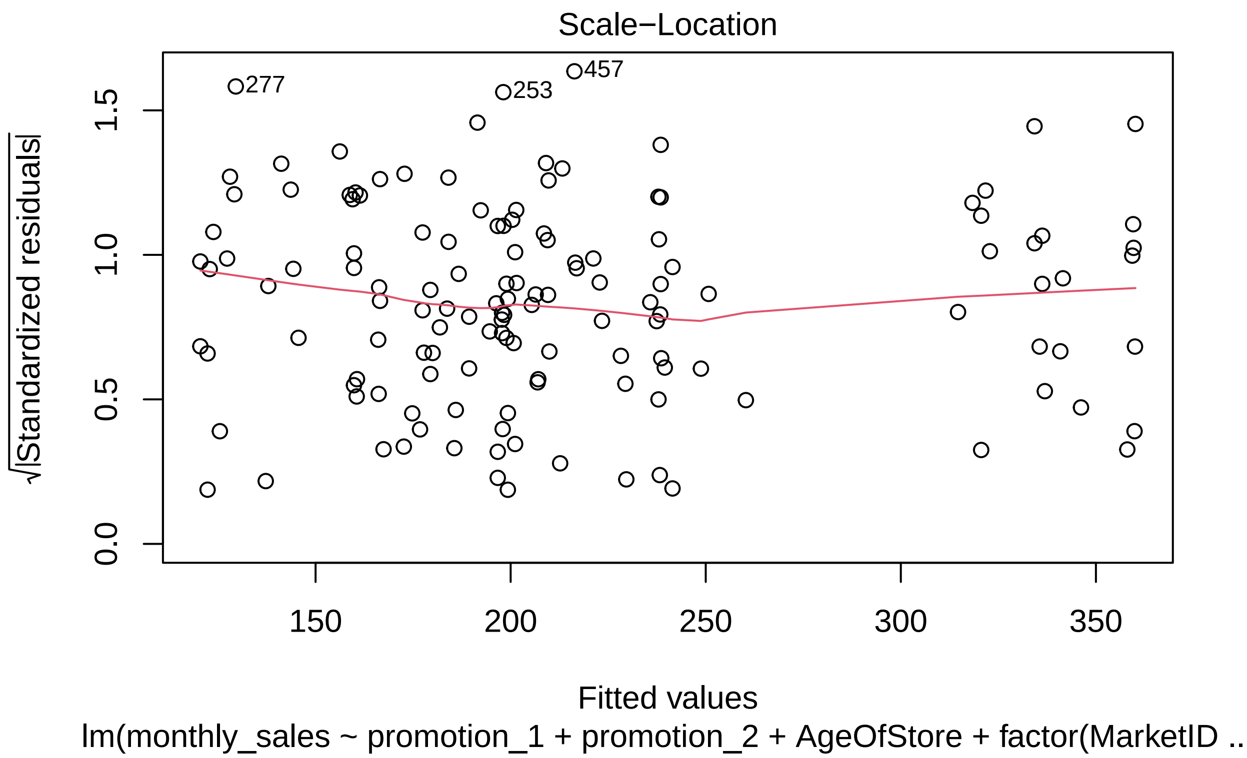
5.3 Linear Conditional Expectation:

When addressing the assumption of linear conditional expectation, we are looking to determine if the residuals are consistent around the BLP line. To do this, we evaluate the Residuals vs Fitted plot. Looking at the plot, we see an even distribution of the residuals - this can be seen by looking at the data points but also by evaluating the red line that runs through the middle of the plot. This indicates that expected error is essentially zero over the range of the input data because the red line is nearly on top of the dashed line at 0. Therefore, we can say that this model satisfies the assumption of a linear conditional expectation.



5.4 Homoskedastic Errors:

When looking for homoskedastic errors, we are looking if there is constant error variance across the range of the x's. To assess homoskedasticity, we can do a visual check and a Breusch-Pagan test. The visual check is done with the Scale-Location plot. In this plot, we see some minor clustering on the left. However, with such few data points it is hard to determine whether there is significant clustering. Turning to the Breusch-Pagan test, we get a p-value of 0.4615 Therefore, we fail to reject the null hypothesis and conclude that there is no evidence of heteroskedasticity. The model satisfies this assumption. However, it should be noted that the Breusch-Pagan test only shows that there is no evidence of heteroskedasticity, but that does not mean that there is no heteroskedasticity. As we run the promotions on more stores and gather more data points, we should keep an eye on this.




```
##
##  studentized Breusch-Pagan test
##
## data:  model_4
## BP = 13.842, df = 14, p-value = 0.4615
```

5.5 Normally Distributed Errors:

Assessing the assumption of normally distributed errors is done by determining if the residuals (errors) of the model are normally distributed. This is done with a Normal Q-Q plot. On the Q-Q plot, we see that the errors are not normally distributed at the tails, which indicates that the errors are not normally distributed. To test for formally distributed errors, we use the Shapiro-Wilk test which has as its null hypothesis that residuals are distributed normally, and we find that the null hypothesis is rejected. Therefore, the model does not satisfy this assumption of normally distributed errors.

The implication for our estimate of the effect of promotions not meeting all 5 classical linear model assumptions is that the estimate will be unbiased and according to the Guass-Markov theorem, will have the minimum variance out of all estimators which are unbiased and linear.

5.6 Omitted Variables

When addressing omitted variables, we are looking to see if there are any additional variables that could affect both promotion and monthly sales. However, because the data is gathered from an A/B testing process, promotions were assigned at random. As such, the assignment of promotion to a store is independent of all other variables that affect monthly sales. Since omitted variables are variables that affect both promotion and monthly sales, we assert that there are no omitted variables to consider.

6 Conclusion

6.1 Conclusion

An A/B test of three promotions of the new menu item was conducted to understand which of the three promotions ought to be deployed in order to most effectively increase sales of the new item. The stores included in the A/B test came from a range of markets, market sizes, and ages. We estimated one-equation structural models to understand the effect of various promotions while controlling for various covariates and find that Promotion 1's effect on monthly sales is robust to various model specifications and various ways to calculate standard errors. The model that we select as the best explanatory model for monthly sales (Model 4) of the new menu item has as its control variables age of the store, market ID, and interaction terms between the promotion and age of the store.

From the model building exercise, we find evidence that Promotion 1 was the most effective promotion as measured by monthly sales. From the selected Model 4, we find that Promotion 1 is estimated to increase sales by ~$26,651 over Promotion 3 and ~$37,057 over Promotion 2. Given that the model draws data from a true experiment, meets the assumptions of the Gauss-Markov theorem, and explains a high degree of the variance of monthly sales in the data, we recommend that the marketing team pursue the implementation of Promotion 1 more broadly to increase sales of the new menu item.

With regard to the limitations of our findings, because we do not have a control group, we can only show the comparison of sales from one promotion to another. We cannot say anything about what sales would have looked like should there be no promotion and the menu item was released. The marketing team should use Promotion 1 to market the new menu item. In addition, we assumed that the selection of stores was conducted randomly and as such the data is I.I.D., however the metadata is not clear on that point, only explaining that markets were randomly selected. If stores within the markets were purposively selected, then the first CLM assumption would not be met and our estimates cannot be considered to be unbiased, and we would not confidently recommend using Promotion 1. Finally, since we find that errors are not normally distributed, the fifth CLM assumption is not met and we cannot say that uncertainty estimates such as standard errors are unbiased. If the standard errors are sufficiently biased, then the effect size of Promotion 1 may not be significantly greater than Promotion 3 or even that of Promotion 2. With further study, we hope to address the limitations and provide further confidence to our marketing team.

6.2 Further Study

The top priority for further study is to rerun this A/B test but include a control group that does not receive any promotion. This would allow for a baseline of what sales of the new item are predicted and can inform how the promotions in general affect sales. To rerun the test, we would need to select a new subset of store locations, ideally in a different geographical location, so that the customer base has no prior knowledge of the item. Company X's marketing team would be able to use this information to more effectively build out a marketing strategy.

Additionally, in further studies it would be beneficial to have additional contextual information on the store, the new item, what the promotions entailed, promotion delivery method, geographical location, and target customer base. This additional information would allow us to build out a more robust and accurate model and suggestion for the marketing team.

Finally, the last suggestion for further study is to continue to monitor the promotions as the new item is rolled out to all stores. The A/B test was run on a small subset of stores (137) and the data science team should continually be analyzing the results of promotions to see if our predictions are accurate or whether the marketing team should change course.