ANOVA Project

Jack Hiegel, Riley Olson, and Long Tran

Introduction

When it comes to the food industry, there are a number of factors that determine the amount of sales that can be reached for particular types of food. Obviously, these can relate to the food being sold itself: serving sizes, ingredients, condiments, and so on and so forth. However, there are also a number of other, non-food related factors that also play a role in determining sales figures. These include location of the food being sold, population demographics in the particular area, and cultural background. With all of these in mind, we have decided to take a look at some of these factors and their potential interaction using a dataset from a marketing blitz for a new type of pizza located in the textbook.

The data, which can be found on page 160 of the textbook, deals with a fast food supply company trying to determine how a new pizza variety will sell and how certain factors affect pizza sales. These factors are threefold: crust size, toppings, and region. The levels for crust are thick (1) and thin (2). The levels for toppings are triple cheese (1), double cheese with beef (2), and cheese with beef and mushrooms (3). For the factor region, there are nine levels in the original data. For convenience's sake, we have consolidated the nine levels into three new levels. These levels are East (1), which includes data for NE, Mid-Atlantic, and SE, Midwest (2), which includes Upper Midwest, Midwest, and Lower Midwest, and West (3), which includes NW, Mid-Pacific, and SW.

Assumptions

Prior to conducting analytical tests on our data, it would be best to first verify our assumptions. These are the assumptions of independence, Normality, and equal variance. We shall look at independence first. This is the assumption that the observations in the data were not tied to one another in a way that would affect their results. Looking at the manner in which the data was collected, there does not seem to be any connections of this sort between the samples, so we can conclude that independence is met.

Moving on, we shall now test Normality. To test this, we will take a look at a Normal Probability Plot of the residuals, which can be seen as **Graph 1** in the appendix at the end of this report. Upon looking at the graph, we can see that the slope is not exactly linear. The p-value associated with the graph is also extremely small, smaller than any reasonable significance level. Therefore, we cannot assume that the Normality assumption is met. While we will continue our analysis, we must remember this and not wholeheartedly believe our results.

The final assumption is the equal variances assumption. In order to test this, we use Levene's Test (shown in **Table 1**), which gives us a p-value of 1. This is much greater than any reasonable significance level, so we can assume that this assumption is met.

Model Statement

With that out of the way, we shall now create a model statement for our data. The statement is shown below along with definitions of key terms:

$y_{jkl} = \mu + \alpha_j + \beta_k + \gamma_l + \alpha\beta_{jk} + \alpha\gamma_{jl} + \beta\gamma_{kl} + \alpha\beta\gamma_{jkl} + \varepsilon_{jkl}$

$j = 1, 2 \quad k = 1, 2, 3 \quad l = 1, 2, 3 \quad \varepsilon_{jkl} \sim$ independent $N(0, \sigma^2)$

$\sum\alpha_j = 0, \sum\beta_k = 0, \sum\gamma_l = 0, \sum\sum\alpha\beta_{jk} = 0, \sum\sum\alpha\gamma_{jl} = 0, \sum\sum\beta\gamma_{kl} = 0, \sum\sum\sum\alpha\beta\gamma_{jkl} = 0$

μ: The grand mean for pizza sales

α_j: The effect of level j for the factor Crust

β_k: The effect of level k for the factor Toppings γ_l: The effect of level l for the factor Region

Tests for Interaction

Now, it is time to test for three-way interaction. We will use a hypothesis test for this procedure. For all tests, we shall use a significance level of 5%. Our null and alternative hypotheses are written below:

Ho: $\alpha\beta\gamma_{jkl} = 0$ for all j, k, l (j = 1, 2, k = 1, 2, 3, l = 1, 2, 3)
Ha: $\alpha\beta\gamma_{jkl}$ does not equal 0 for some j, k, l (j = 1, 2, k = 1, 2, 3, l = 1, 2, 3)

The ANOVA table is located in the appendix as **Table 2**. Looking at the last p-value, we can see that it is equal to 0.898, which is greater than our alpha of 0.05. Because of this, we fail to reject the null hypothesis and lack statistically significant evidence of three-way interaction. Therefore, we must move on to testing two-way interaction. This involves testing interaction between Crust and Toppings, Crust and Region, and Toppings and Region. Our null and alternative hypotheses are shown below:

Ho: $\alpha\beta_{jk} = 0$ for all j, k (j = 1, 2, k = 1, 2, 3)
Ha: $\alpha\beta_{jk}$ does not equal 0 for some j, k (j = 1, 2, k = 1, 2, 3)

Ho: $\alpha\gamma_{jl} = 0$ for all j, l (j = 1, 2, l = 1, 2, 3)
Ha: $\alpha\gamma_{jl}$ does not equal 0 for some j, l (j = 1, 2, l = 1, 2, 3)

Ho: $\beta\gamma_{kl} = 0$ for all k, l (k = 1, 2, 3, l = 1, 2, 3)
Ha: $\beta\gamma_{kl}$ does not equal 0 for some k, l (k = 1, 2, 3, l = 1, 2, 3)

Looking back at our ANOVA table, we see that the p-values associated with these tests are 0.193, 0.742, and 0.916, all of which are larger than 0.05. Therefore, we fail to reject the null in all cases and lack statistically significant evidence of two-way interaction between any of the variables.

Interaction Plots

To verify our conclusions from the tests, we have decided to create interaction plots for all of the two-way interactions. These can be seen in **Graph 2**. Looking at the six graphs displayed, we can see that virtually all of them display roughly parallel lines. This is a sign of a lack of interaction, confirming what we have discovered in our tests.

Main Effects Tests

Having failed to find evidence of interaction, we must test to see if the effects of each singular variable upon our response are significant. We will run tests for each variable (Crust, Toppings, and Region). Our hypotheses for these three tests can be seen below:

Ho: $\alpha_j = 0$ for all j (j = 1, 2)
Ha: α_j does not equal 0 for some j (j = 1, 2)

Ho: $\beta_k = 0$ for all k (k = 1, 2, 3)
Ha: β_k does not equal 0 for some k (k = 1, 2, 3)

Ho: $\gamma_l = 0$ for all l (l = 1, 2, 3)
Ha: γ_l does not equal 0 for some l (l = 1, 2, 3)

Looking back at our ANOVA table, we see that the p-values associated with these tests are equal to 0.000, 0.004, and 0.002, which are all less than 0.05. Because of this, we reject the null hypotheses and have statistically significant evidence that all three variables have a significant effect on the Amount of Sales.

Final Model Statement

Having determined the significant variables, we can now create our final model statement, which is shown below:

$y_{jkl} = \mu + \alpha_j + \beta_k + \gamma_l + \varepsilon_{jkl}$
$j = 1, 2 \quad k = 1, 2, 3 \quad l = 1, 2, 3 \quad \varepsilon_{jkl} \sim$ independent $N(0, \sigma^2)$
$\sum\alpha_j = 0, \sum\beta_k = 0, \sum\gamma_l = 0$

Multiple Comparisons

Now that we have determined that each of our variables has a significant effect on the Amount of Sales, we will move on to seeing how the levels of each factor affect the response. To do this, we have used Tukey's HSD test. Through Minitab, we attained the table displayed in **Table 3**. By using the grouping information, we are able to create line graphs for each factor. These are displayed below.

Crust:

$\bar{y}_1$ $\bar{y}_2$ (Both means are statistically different)

Toppings:

$\bar{y}_3$ $\underline{\bar{y}_2 \quad \bar{y}_1}$

Region:

$\bar{y}_3$ $\underline{\bar{y}_1 \quad \bar{y}_2}$

Upon looking at the information displayed for Crust, we can see that the mean amounts of sales for the two levels are significantly different, and that the one for thick crust is larger than the one for thin crust. Looking at Toppings, we observe that the mean for cheese with beef and mushrooms is larger than and significantly different from the other two means. Finally, when looking at Region, we see that the mean amount of sales for the West is also larger than and significantly different from the other two means associated with this factor.

Main Effects Plots

In order to confirm our findings from the multiple comparisons tests, we have created main effects plots. The plot for Crust can be seen in **Graph 3**, the one for Toppings can be seen in **Graph 4**, and the one for Region can be seen in **Graph 5**. These plots seem to verify what we have already discovered. In terms of crust size, thick sells better than thin. When it comes to toppings, cheese with beef and mushrooms sells the best. For region, the pizzas sell better in the West than in the other areas.

Summary & Conclusion

Before conducting any actual analysis, we checked all three of our assumptions. We found that while independence and equal variance were met, Normality was not, meaning that we should not view our results as absolutely certain. At the beginning of our analysis, we found out that there is no three-way interaction between the 3 factors (Crust, Toppings, Region). As a result, we proceeded to test 3 separate two-way interactions between the factors. However, none of the interaction tests provide any statistical evidence of interaction between any of the factors. This leads us to the main effect tests to see whether each factor is significant or not. We found that each of the factors has significant effect on the amount of sales. Next, we used the Tukey's HSD Test to determine how the levels of each factor affect the response. For the Crust factor, thick crust sells better than thin crust, reflected by the significant difference in the mean amount of sales between the two. In term of Toppings, cheese with beef and mushrooms sells better than the other two, also reflected by the significant difference between its mean and those of the other two. Lastly, the mean amount of Sales for the West region is significantly different than that of the East and Midwest regions as well.

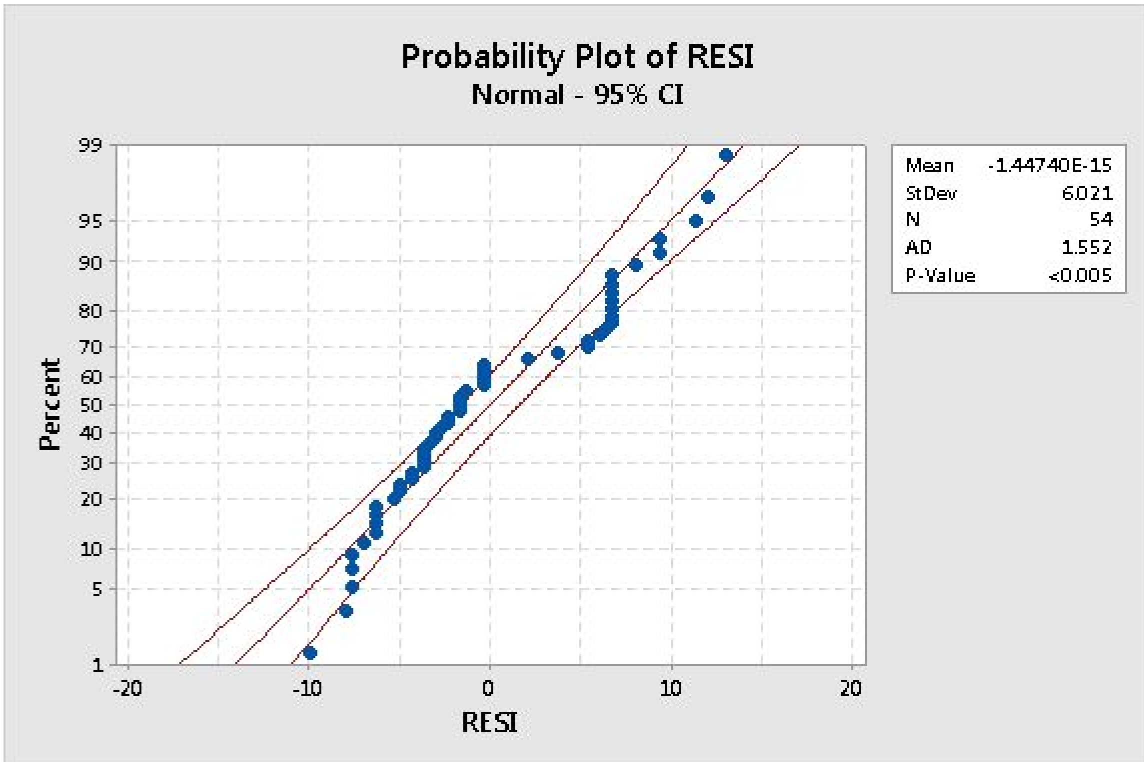
Recommendations

While certain questions can be answered based on the analysis of the given data, we probably should not generalize the results of this study to the general population. Data collected from 6 convenience stores in each region is a very small sample to begin with. Even though the assumption of "all the stores are as nearly as alike as can be expected" exists, there are many other factors that affect the amount of sales in a region. The stores can have the same community size, but factors like age or cultural background (race, religion, type of community) certainly affect the data. For the sake of future studies with the same objectives, some designs should be changed for a better outcome. For example, taking data from many stores from each region to represent different settings will make the data more representative of the given region as a whole.

Credits

Our group consists of three individuals. We worked together on finding the dataset and the proposal. For the final report, Riley worked on the introductory paragraph, explaining the study itself and how it was conducted. Jack conducted the ANOVA studies and made the graphs on Minitab as well as the Analysis explanation and the Appendix. Long worked on the summary and the recommendations part of the report.

Appendix



Graph 1

Tests

Method	Test Statistic	P-Value
Multiple comparisons	—	0.985
Levene	0.19	1.000

Table 1

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Crust	1	3297.85	3297.85	61.79	0.000
Toppings	2	676.93	338.46	6.34	0.004
Region	2	782.37	391.19	7.33	0.002
Crust*Toppings	2	183.59	91.80	1.72	0.193
Crust*Region	2	32.15	16.07	0.30	0.742
Toppings*Region	4	50.52	12.63	0.24	0.916
Crust*Toppings*Region	4	56.74	14.19	0.27	0.898
Error	36	1921.33	53.37		
Total	53	7001.48			

Table 2

Graph 2

Comparisons for Amount of Sales ($)

Tukey Pairwise Comparisons: Crust

Grouping Information Using the Tukey Method and 95% Confidence

Crust	N	Mean	Grouping	
1	27	71.2963	A	
2	27	55.6667		B

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Crust Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
2 - 1	-15.63	1.99	(-19.66, -11.60)	-7.86	0.000

Individual confidence level = 95.00%

Tukey Pairwise Comparisons: Toppings

Grouping Information Using the Tukey Method and 95% Confidence

Toppings	N	Mean	Grouping	
3	18	68.3889	A	
2	18	61.8889		B
1	18	60.1667		B

Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

Difference of Toppings Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
2 - 1	1.72	2.44	(-4.24, 7.68)	0.71	0.761
3 - 1	8.22	2.44	(2.26, 14.18)	3.38	0.005
3 - 2	6.50	2.44	(0.54, 12.46)	2.67	0.030

Individual confidence level = 98.06%

Tukey Pairwise Comparisons: Region

Grouping Information Using the Tukey Method and 95% Confidence

Region	N	Mean	Grouping	
3	18	68.7778	A	
1	18	61.6667		B
2	18	60.0000		B

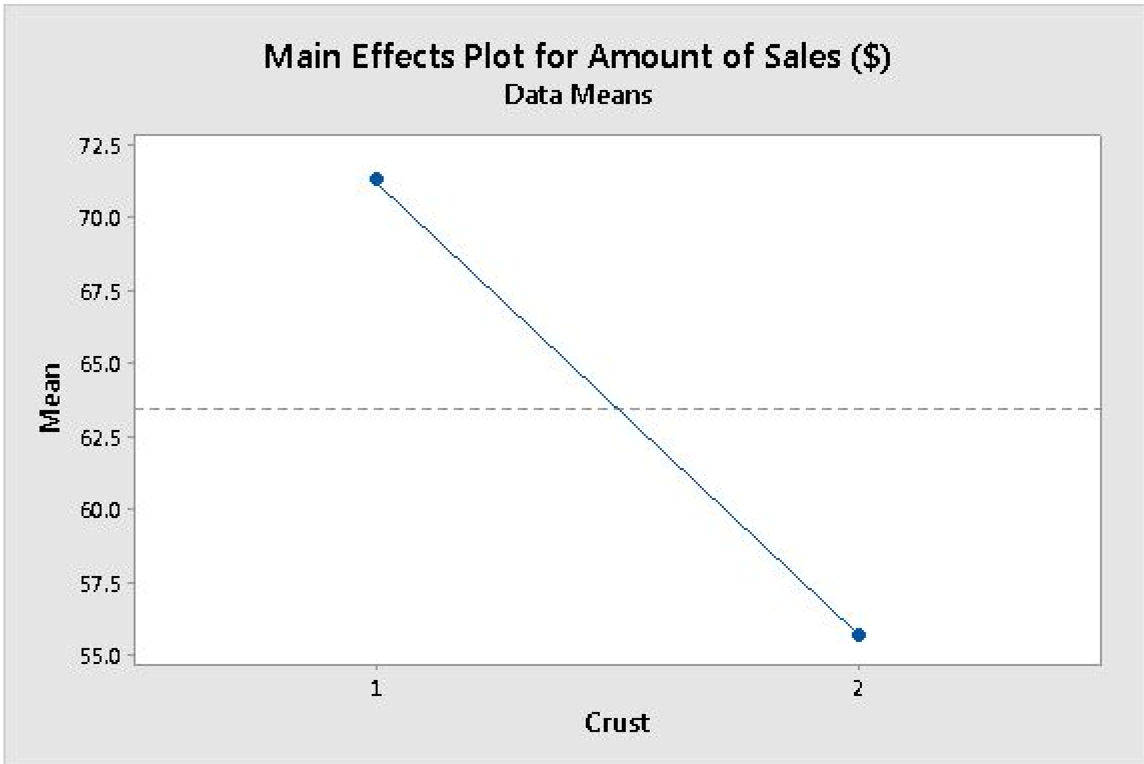
Means that do not share a letter are significantly different.

Tukey Simultaneous Tests for Differences of Means

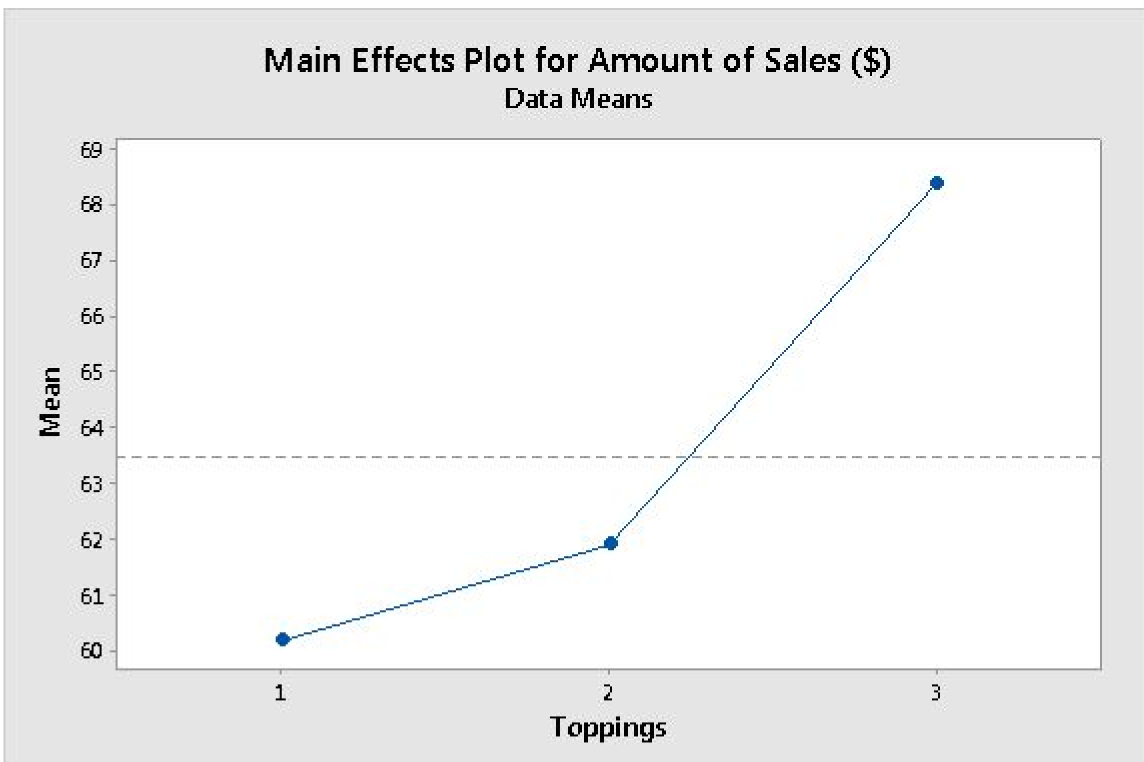
Difference of Region Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
2 - 1	-1.67	2.44	(-7.62, 4.29)	-0.68	0.774
3 - 1	7.11	2.44	(1.15, 13.07)	2.92	0.016
3 - 2	8.78	2.44	(2.82, 14.74)	3.60	0.003

Individual confidence level = 98.06%

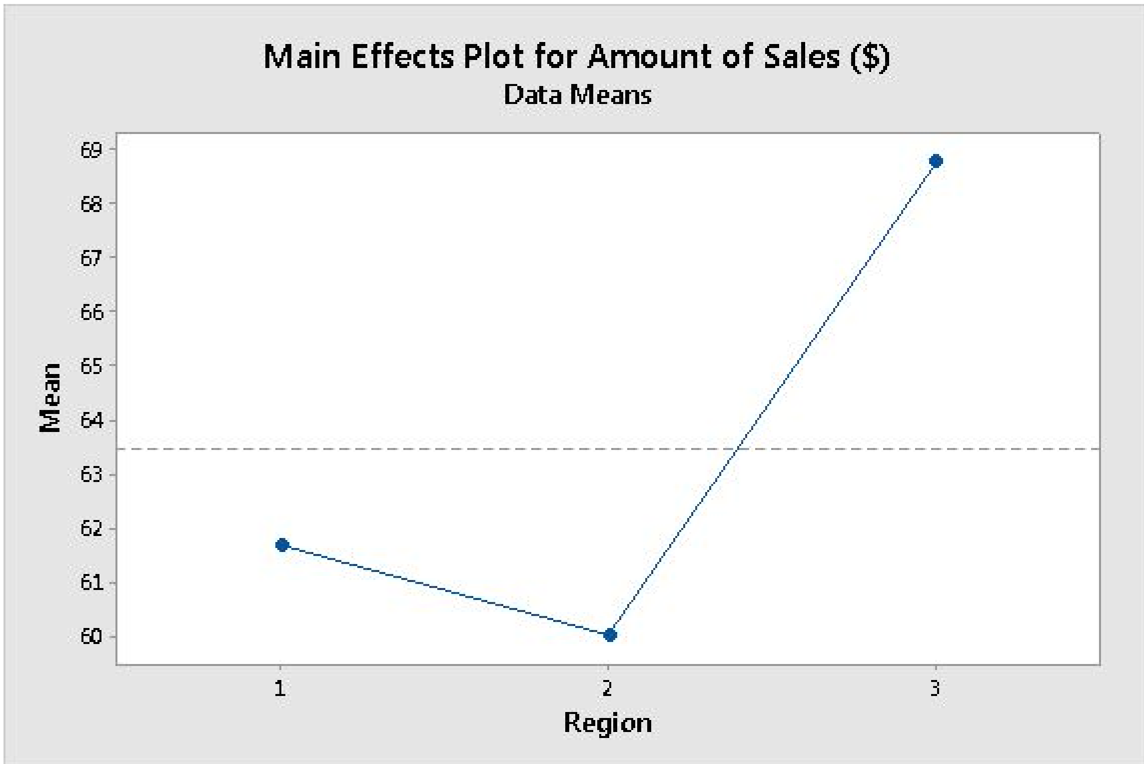
Table 3



Graph 3



Graph 4


Main Effects Plot for Amount of Sales ($)
Data Means
Mean
69
68
67
66
65
64
63
62
61
60
1
2
3
Region

Graph 5