Meddicorp Sales

The Forecast Force

Edward De Avila, Sai Surya Varun Bade, Sophia Ding, Lauren Newton, Sheetal Padmanabhan, Nathan Sone

Department of Information Systems & Decision Sciences, California State University Fullerton

ISDS 562: Advanced Regression Analysis for Business Applications

Dr. Sinjini Mitra

December 17, 2024

Table of Contents

Executive Summary

Meddicorp would like to know if the company experiences changes in sales that are significantly related to employee bonuses. This report follows our team's steps to investigate and analyze the overall significance of the relationship between Bonus and Sales while also evaluating the role that the regions in which these bonuses are awarded play along with their potential importance in the overall effectiveness of the final model.

For data exploration, we explored our data to uncover preliminary trends and correlations. We found the correlations between the predictor variables and sales were adequate and may indicate a relationship.

For our regression model(s) section, we evaluated a model that only included advertising and bonuses and a model that also included the region in which the bonus was awarded. While the primary model did indicate a significance between the variables, the second model deepened this relationship to show us that the region in which the bonus is assigned may have a statistically significant impact on sales rather than merely seeing if a bonus was given. We conducted nested F tests to decide if including nonsignificant variables was substantial to the model. Additionally, we explored the impact that transformations would have on the performance of our model.

Finally, in our model evaluation section, we ensured that our model validated the four linear model assumptions. We also evaluated the model to determine if outliers could impact model performance.

Introduction: background and problem description

Meddicorp supplies hospitals and medical facilities with medical supplies in the South, West, and Midwest regions of the United States. The corporation has developed a new program of awarding bonuses and would like to understand the overall effectiveness of the program and its impact on the performance of its salesforce. This program rewards salespeople with bonuses based on their performance and is overseen by the regional managers. As a group, we will be helping Meddicorp determine if there is a significant relationship between bonuses paid and sales. Additionally, we will be considering the regions to observe if and how bonuses offered in different regions contribute to changes in sales.

Questions of Interest and Variable Description

The Meddicorp case study evaluates the effectiveness of a new bonus program for sales personnel. The primary questions of interest are:

- Is there a relationship between bonuses paid in 2003 and sales performance?
- How does advertising spending affect sales?
- Are there regional differences in sales performance?

To address these questions, the study utilizes the following variables:

Variable	Description	Unit
Y (Sales)	Territory sales for 2003	Thousands ($)
X_1 (Advertising)	Amount spent on advertising per territory	Hundreds ($)
X_2 (Bonus)	Total bonus payments per territory	Hundreds ($)
Region	Sales Region coded as: 1 = South, 2 = West, 3 = Midwest	Categorical

Two indicator variables, D1 and D2, represent the qualitative region variable, with South as the reference category.

Data Exploration

Descriptive Statistics

First, let's look at the summary statistics for our quantitative variables:

Statistic	Sales (Y) (In Thousands of Dollars)	Advertising (X_1) (In Hundreds of Dollars)	Bonus (X_2) (In Hundreds of Dollars)
Sample Size	**25**	**25**	**25**
Mean	**1269.02**	**514.54**	**276.34**
Median	**1202.75**	**500.56**	**272.55**
Standard Deviation	**228.12**	**74.09**	**28.50**
Minimum	**893**	**374.27**	**230.98**
Maximum	**1634.75**	**637.60**	**332.64**

Distribution of Sales

The histogram in *Appendix Figure 1* shows the distribution of sales across territories. The histogram of sales distribution reveals several key characteristics:

- The distribution appears to have two distinct peaks - one around 1000-1200 thousand dollars and another at approximately 1600 thousand dollars in sales. This bimodal pattern could suggest two distinct groups of sales territories with different performance levels, possibly reflecting regional differences or varying effectiveness of sales strategies.
- There is a notable spread in the data suggesting considerable variation in sales performance across territories.

Relationship between Sales and Advertising and Sales and Bonus

From the plot for Sales vs. Advertising in *Appendix Figure 3*, we can say there is a strong positive linear correlation with tight clustering around the trend line, showing advertising expenditure consistently predicts sales performance, whereas for the Sales vs Bonus plot in *Appendix Figure 4*, we can see a weak positive correlation with significant scatter around the trend line, indicating bonuses are less reliable predictors of sales performance.

Correlation Analysis

	Sales	**Advertising**	**Bonus**
Sales	1	0.900	0.568
Advertising	0.900	1	0.419
Bonus	0.568	0.419	1

- The correlation between Sales and Advertising is 0.900, indicating a robust positive relationship where increased advertising spending is highly associated with higher sales.
- The correlation between Sales and Bonuses is 0.568, suggesting a moderate positive relationship with sales, though notably weaker than advertising's influence.
- The correlation between Advertising and Bonuses is 0.419, which indicates a moderate correlation that is not strong enough to raise significant multicollinearity concerns.

Regression model(s) – Interpretations and Summary of Findings

<u>Fitted Regression Equation & Interpretation</u>

The fitted regression equation for the relationship between Meddicorp's sales, advertising expenditures, and bonuses is:

$$\widehat{\boldsymbol{Sales}} = -516.444+2.473Adv+1.856Bonus$$

<u>Interpretation of Coefficients</u>

I. Intercept (-516.444):
The intercept represents the estimated sales when advertising expenditures (*Adv*) and bonuses (*Bonus*) are $0. In this case, the model predicts sales would be approximately –516.44 thousand dollars. However, this negative value is likely not practical or meaningful in a real-world context, as Meddicorp would not operate without spending on advertising or bonuses. It primarily serves as a mathematical artifact of the regression model.

II. Slope Coefficient for *Adv* (2.473):
For every additional $100 (1 unit) spent on advertising, sales are expected to increase by approximately $2,473, assuming the amount spent on bonuses remains constant. This positive relationship indicates that increasing the advertising budget is associated with higher sales.

III. Slope Coefficient for *Bonus* (1.856):
For every additional $100 (1 unit) paid in bonuses, sales are expected to increase by approximately $1,856, assuming advertising expenditures remain constant. This suggests that bonuses positively correlate with sales, though their impact is slightly less pronounced than advertising expenditures.

In summary, both advertising and bonuses are positively associated with Meddicorp's sales, highlighting the importance of these factors in driving revenue. Advertising has a slightly more substantial effect than bonuses, indicating that investing in promotional activities may yield a greater return in sales.

The regression analysis output from SPSS and scatterplots with the regression line have been created to visualize the relationship between the predictors (advertising and bonuses) and sales. See *Appendix Figure 2* for the regression analysis output from SPSS, *Appendix Figure 3* for the scatterplot of sales against advertising, and *Appendix Figure 4* for the scatterplot of sales against bonuses.

<u>Model Fit and Hypothesis Testing</u>

<u>R Squared</u>

The R-squared value for our model is 0.855, indicating that approximately 85.5% of the variability in Meddicorp's sales can be explained by the independent variables, advertising spending (*Adv*) and bonuses (*Bonus*). This suggests a strong relationship between the predictors and the dependent variable, implying that the model captures a significant portion of the factors influencing sales.

<u>Standard Error of Estimate</u>

The standard error of the estimate is 90.75, meaning that, on average, the predicted sales values differ from the actual sales values by approximately $90,748.51. While this is a relatively large margin, it reflects the inherent variability in sales data. It highlights the need to further assess the predictor variables' significance through hypothesis testing.

<u>Hypothesis Testing</u>

<u>Global Significance Test (F-test)</u>

To test the overall significance of the regression model:

- **<u>Null Hypothesis (H0)</u>:** $b_1 = b_2 = 0$
  Our null hypothesis states that neither advertising nor bonuses significantly impact sales.
- **<u>Alternative Hypothesis (HA)</u>:** *At least one* $b_i \neq 0$
  Our alternative hypothesis states that at least one predictor significantly impacts sales.

Because our associated p-value is equal to less than 0.001, this indicates that the overall regression model is highly significant. This provides strong evidence to reject the null hypothesis and conclude that at least one of the predictors significantly contributes to explaining sales.

<u>Individual Predictor Significance (T-Tests)</u>

The t-statistic for advertising (*Adv*) is 8.98, with a p-value of < 0.001. This confirms that advertising spending has a significant positive effect on sales. For every additional $100 spent on advertising, sales are expected to increase significantly, underscoring its strong relationship with the dependent variable.

The t-statistic for bonuses (*Bonus*) is 2.59, with a p-value of 0.017. This indicates that bonuses also have a positive and statistically significant effect on sales. For every additional $100 bonus paid, sales are expected to increase, though the effect size is less pronounced than that of advertising, as reflected by the t-statistic.

Overall, both predictor variables are statistically significant, with advertising having a more substantial relative effect on sales than bonuses. These findings reinforce the utility of the regression model in explaining sales variability and highlight the importance of Meddicorp's spending decisions in both areas.

<u>Updated Estimated Regression Equation with Regional Dummy Variables</u>

The new estimated regression equation, incorporating regional dummy variables *D1* (West) and *D2* (Midwest), is as follows:

$$\widehat{\boldsymbol{Sales}}: \mathbf{177.21 + 1.37Adv + 0.98Bonus + 257.89D1 + 48.15D2}$$

This model accounts for differences in sales across the three regions, with the South region as the reference category.

<u>Region-Specific Regression Equations</u>

The inclusion of dummy variables allows us to generate separate regression equations for each region:

<u>South (Reference Category):</u>

For the South, both *D1* and *D2* equal 0:

$$\widehat{\boldsymbol{Sales}}: \mathbf{177.21} + \mathbf{1.37}\boldsymbol{Adv} + \mathbf{0.98}\boldsymbol{Bonus}$$

<u>West:</u>

For the West, *D1 = 1* and *D2 = 0:*

$$\widehat{\boldsymbol{Sales}}: \mathbf{177.21} + \mathbf{1.37}\boldsymbol{Adv} + \mathbf{0.98}\boldsymbol{Bonus} + \mathbf{257.89}$$

<u>Midwest:</u>

For the Midwest, *D1 = 0* and *D2 = 1:*

$$\widehat{\boldsymbol{Sales}}: \mathbf{177.21} + \mathbf{1.37}\boldsymbol{Adv} + \mathbf{0.98}\boldsymbol{Bonus} + \mathbf{48.15}$$

These equations demonstrate how sales predictions vary based on region, with distinct adjustments for the West and Midwest compared to the South.

Model Comparison and Analysis

R-squared and Standard Error

The updated model has an R-squared of 0.947, an improvement over the previous model's R-squared of 0.855. This means the updated model explains 94.7% of the variance in sales compared to 85.5% in the original model.

The standard error of estimate has decreased from 90.75 to 57.63, indicating that the updated model's predictions are closer to the actual sales values, thus providing a better fit.

F-Statistic

The new F-statistic is 89.03 with a p-value of < 0.001, confirming that the overall model is statistically significant and represents an improvement over the previous model.

Individual Predictor Significance

The updated regression analysis shows that advertising maintains strong statistical significance (t=5.22, p<0.001), while bonuses show marginal significance (t=2.03, p=0.056). Regional analysis reveals the West has significantly higher sales compared to the South (t=5.33, p<0.001), while the Midwest shows no significant difference (t=1.47, p=0.158). The model demonstrates improved predictive capability with enhanced R-squared values and reduced standard error, with advertising emerging as the strongest sales predictor.

Nested Model F-Test for Variables D2 and BONUS

To determine whether the variables D2 (Midwest) and BONUS are helpful in the model in addition to the other variables, we perform a nested model F-test. This test compares the full model (including all variables) to a reduced model that excludes D2 and BONUS. The SPSS output for the reduced model is shown in *Appendix Figure 16*.

Hypotheses

- **Null Hypothesis (H0):** *b2 = b4 = 0*
  Our null hypothesis states that the coefficients for D2 and BONUS are equal to zero, indicating that these variables do not contribute to the model.
- **Alternative Hypothesis (HA):** *at least one of b2 or b4 is not equal to zero*
  Our alternative hypothesis states that at least one of these variables significantly contributes to the model.

F-Statistic Calculation

The F-statistic is calculated using the following formula:

$$\mathrm{F} = \frac{\frac{(RSS_R - RSS_C)}{(k-r)}}{\frac{RSS_C}{(n-k-1)}}$$

RSS represents the residual sum of squares, and df represents the degree of freedom.
From the SPSS output, the residual sum of squares for the whole model is 91,352.80, and for the reduced model is 66,413.84. Substituting these values into the formula:

$$F = \frac{\frac{91352.80 - 66413.84}{(4-2)}}{\frac{66413.84}{(25 - 4 - 1)}} \qquad F = \frac{\frac{24938.96}{2}}{\frac{66413.84}{20}} \qquad F = \frac{12469.48}{3320.692} = 3.755$$

The critical F-value for this test is 3.49, calculated using a significance level of 5%, the 95th percentile of the F-distribution with *k – r = 2* numerator degrees of freedom and *n – k –1 = 20* denominator degrees of freedom.

Since our calculated F-statistic 3.755 is greater than the critical F-value of 3.49, we reject the null hypothesis. This suggests that at least one of the variables, D2 and BONUS, significantly impacts the model and cannot be excluded. In conclusion, the inclusion of D2 and BONUS substantially improves the model's ability to predict Meddicorp's sales. Therefore, we retain these variables in the final model as they provide meaningful contributions to the prediction of sales.

Additional Transformations

Seeing that there was a lack of significance for the variables bonus and D2, we decided to conduct some transformations on the variables Adv and Bonus to see if they would yield better significance and predictive performance.

Starting with exponential transformations, conducting a nested F-test on a model that had the original variable (Adv, Bonus, D1, and D2) while adding AdvSquared and Bonus Squared yielded an F-statistic of 1.238 which was not more significant than the critical value of 3.55. Additionally, the nested F-statistic of the original variables with Bonus Squared added once again yielded an F-statistic, 2.298, which was not more significant than the critical value, 4.38.

We tried additional logarithmic and inverse transformations following the same pattern of trying on adv and bonus, then just on bonus. Each of these models yielded similar results to the untransformed ones. The inverse transformation of bonus yielded the highest Adjusted R square of .938 along with the lowest standard error of 56.85; however, we decided not to proceed with this model, considering Meddicorp wants to see the impact that bonuses play on sales and adding a transformation that yields such a minute increase in performance is not worth the increased complexities in explanation this would enact.

Final Model Selection and Justification

After conducting a nested model F-test to evaluate the significance of the variables D2 (Midwest) and Bonus, we found that both terms significantly contribute to the model's predictive power. The null hypothesis was rejected, indicating that the inclusion of D2 and Bonus improves the model's ability to explain variation in sales.

Therefore, the full regression model, which includes Advertising (*Adv*), Bonuses (*Bonus*), and the regional dummy variables *D1 (West)* and *D2 (Midwest)*, is retained for further analysis. The final regression model can be expressed as follows:

$$\widehat{Sales} = 177.21 + 1.37Adv + 0.98Bonus + 257.89D1 + 48.15D2$$

This model incorporates all the significant variables, with ***Adv*** and ***Bonus*** representing the influence of advertising and employee bonuses on sales, while ***D1*** and ***D2*** account for the regional variations in sales performance. The whole model is now considered the best fit for the data and will be used for further analysis and predictions.

Model Evaluation – Assumptions, Outliers, High Leverage Values, and Multicollinearity

Model Assumptions

To evaluate the assumptions of the final regression model, we conducted several checks to ensure that the key assumptions were met:

I. **Zero Means Assumption:**
The zero means assumption suggests that the residuals should be centered around zero. To evaluate this, we plotted the residuals against the predicted values. From the residuals scatterplot (*Appendix Figure 7*), we observe that the residuals are symmetrically distributed above and below the zero line. There is no noticeable trend, and the residuals are evenly spread, which indicates that the zero mean assumption is reasonably satisfied.

II. **Constant Variance Assumption (Homoscedasticity):**
The assumption of constant variance, or homoscedasticity, requires that the residuals maintain a constant spread across all levels of predicted values. Plotting the studentized residuals (*Appendix Figure 8*) shows they are evenly scattered around zero without a noticeable cone or funnel shape. This visual inspection suggests that the constant variance assumption is reasonably met for the model.

III. **Normality Assumption:**
To test for the normality of the residuals, we first performed a Shapiro-West test (*Appendix Figure 9*), which returned a statistic of 0.894. Since the p-value is greater than 0.05, we fail to reject the null hypothesis, suggesting that the residuals are normally distributed. Further supporting this, the histogram (*Appendix Figure 10*) shows a bell-shaped normal distribution, and the Q-Q plot (*Appendix Figure 11*) demonstrates that the residuals closely follow the diagonal line, indicating normality.

IV. **Independence Assumption:**
The independence assumption requires that the residuals are not autocorrelated. We checked for autocorrelation by examining the Durbin-Watson statistic, which is 1.959. Since the value is close to 2, it suggests that there is no significant autocorrelation in the residuals; thus, the independence assumption holds. The SPSS model summary of the Durbin-Watson statistic is found in *Appendix Figure 12*.

Model Diagnostics – Outliers, High Leverage Values, and Multicollinearity

Outliers

To check for outliers, we examined the standardized residuals. Standardized residuals with an absolute value greater than 2 or 3 are typically considered potential outliers. Our analysis of the standardized residuals shows no records exceeding this threshold. The computed standardized residual can be found in *Appendix Figure 13*, labeled as ZRE_1.

High Leverage Values

Leverage measures the influence of individual data points on the regression model. We calculated the leverage values and used the criterion. $\frac{2(p+1)}{n} = 0.40$, where p is the number of predictors and n is the sample size. Since no points have leverage values greater than 0.40, our model has no high leverage values. The computed leverage values are presented in *Appendix Figure 14*, labeled as LEV_1.

Multicollinearity

To assess multicollinearity, we examined the Variance Inflation Factor (VIF) values for the predictor variables. A VIF greater than 5 or 10 would indicate a potential multicollinearity issue. In our case, none of the VIFs exceeded these thresholds, suggesting that multicollinearity is not a concern in this model. The SPSS output for the VIFs can be found in *Appendix Figure 15*.

Conclusions

Our findings from this model indicate that while bonus expenditure is only partially the backbone of regional success, it is an essential factor to consider, as its omission causes our model to perform less efficiently. Meddicorp may also want to consider the effects of their advertising campaigns, as we found that Advertising contributes almost double that of the Bonus variable. We also discovered that this effect is more pronounced upon introducing regional insight, specifically the West.

While analyzing this data, some challenges included whether to factor in interaction variables to clean up our regression model assumptions possibly. This did not need to be performed as our initial models had already passed. This investigation might also inspire Meddicorp to reconduct a sample of more specific, categorical factors the next time.

Meddicorp may want to consider restructuring its advertising and bonus efforts by region, as we have discovered the Western region contributes the most to its sales. Additional investigators may also be suggested on external factors within these regions that Meddicorp does not directly control. This may include competitor activity, customer feedback, or perceived value. Finally, investigating the internal effects of Meddicorp's bonus program may prove a worthwhile next step. This can include questions about internal efficiency and employee morale based on the bonus program. A tailored, regional strategy may improve efficiency around the company and drive ongoing sales.

Appendix

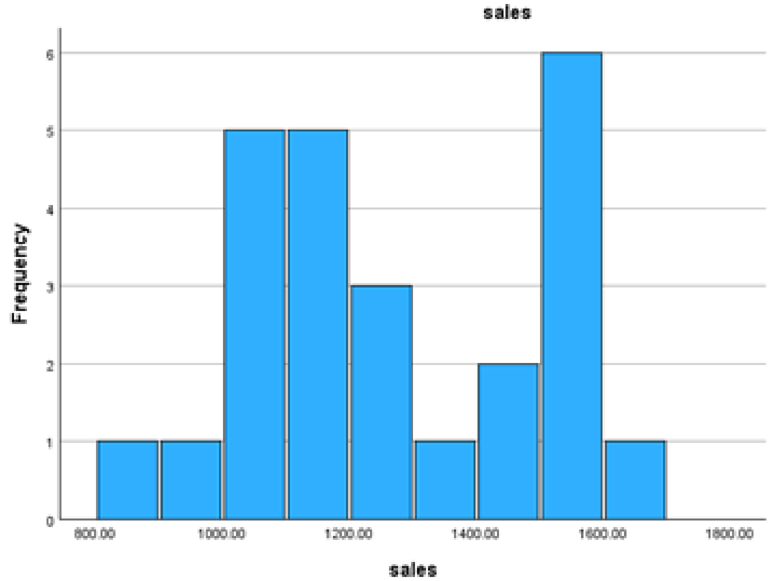

Figure 1

Histogram for Sales

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.925[a]	.855	.842	90.74851

a. Predictors: (Constant), bonus, adv

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1067797.321	2	533898.660	64.831	<.001[b]
	Residual	181176.419	22	8235.292		
	Total	1248973.740	24			

a. Dependent Variable: sales

b. Predictors: (Constant), bonus, adv

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-516.444	189.876		-2.720	.013
	adv	2.473	.275	.803	8.983	<.001
	bonus	1.856	.716	.232	2.593	.017

a. Dependent Variable: sales

Figure 2

SPSS Regression Analysis Output for Advertising & Bonuses

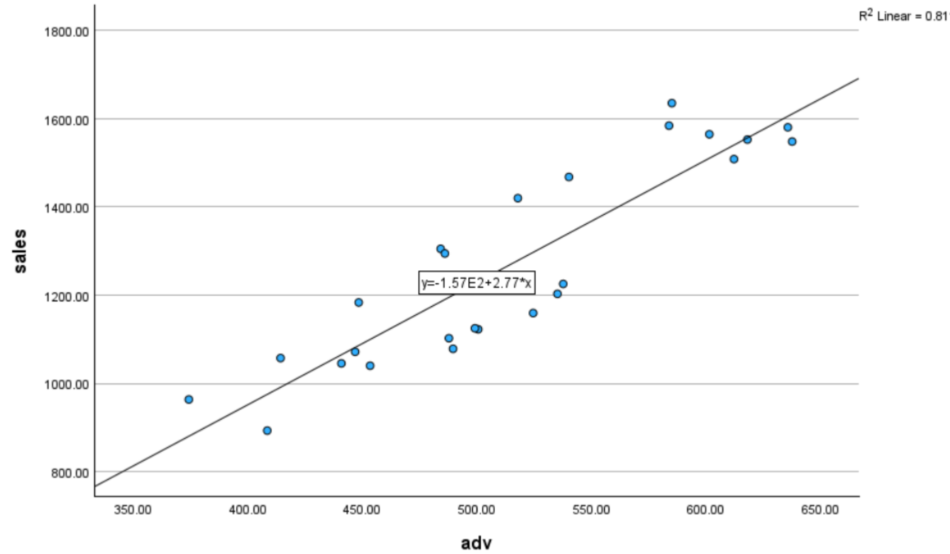

Figure 3

Scatterplot of Sales (Y) versus Advertising (X)

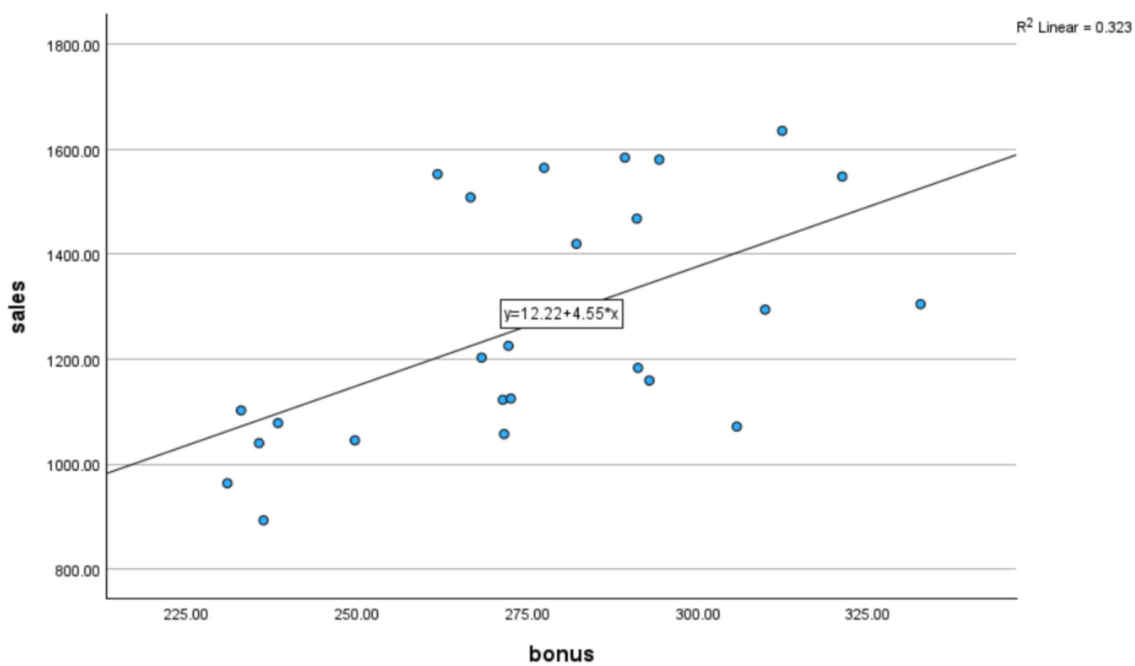

Figure 4

Scatterplot of Sales (Y) versus Bonuses (X)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.973[a]	.947	.936	57.62545

a. Predictors: (Constant), D2, bonus, adv, D1

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1182559.896	4	295639.974	89.030	<.001[b]
	Residual	66413.844	20	3320.692		
	Total	1248973.740	24			

a. Dependent Variable: sales

b. Predictors: (Constant), D2, bonus, adv, D1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	177.207	170.116		1.042	.310	-177.648	532.063
	adv	1.368	.262	.444	5.216	<.001	.821	1.915
	bonus	.975	.481	.122	2.028	.056	-.028	1.978
	D1	257.892	48.413	.565	5.327	<.001	156.904	358.879
	D2	48.146	32.801	.100	1.468	.158	-20.277	116.568

a. Dependent Variable: sales

Figure 5

SPSS Regression Analysis Output for Advertising, Bonuses, and Dummy Regional Variables (D1 & D2)

Statistic	Previous Model	New Model
R squared	0.855	0.947
Standard Error of Estimate	90.74851	57.62545
F-statistic	64.831	89.030
p-value (F-statistic)	<0.001	<0.001
t-statistic – Adv	8.983	5.216
p-value – Adv	<0.001	<0.001
t-statistic – Bonus	2.593	2.028
p-value – Bonus	0.017	0.056
t-statistic – D1	NA	5.327
p-value – D1	NA	<0.001
t-statistic – D2	NA	1.468
p-value – D2	NA	0.158

Figure 6

Table Describing Statistics from the Previous Model (Advertising & Bonuses) and the New Model (Advertising, Bonuses, Dummy Regional Variables)

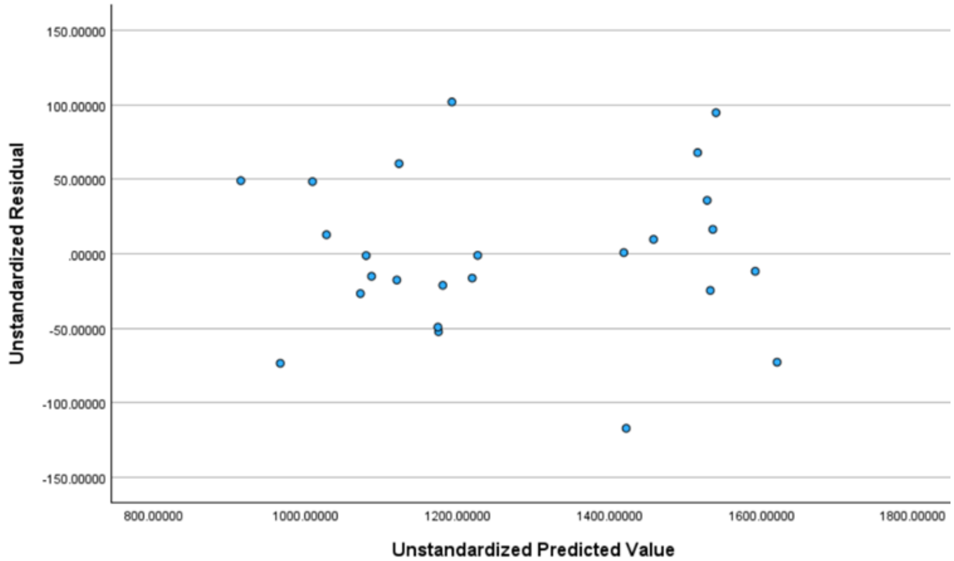

Figure 7

Residuals Scatterplot for the Zero Mean Assumption

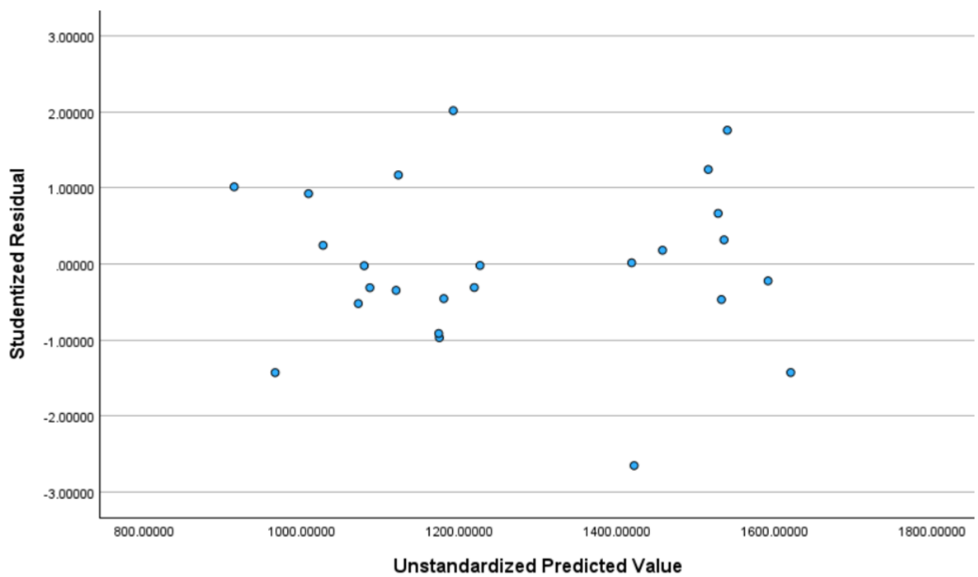

Figure 8

Studentized Residuals Scatterplot for the Constant Variance Assumption

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.107	25	.200*	.980	25	.894

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 9

SPSS Output of the Shapiro-Wilk Test

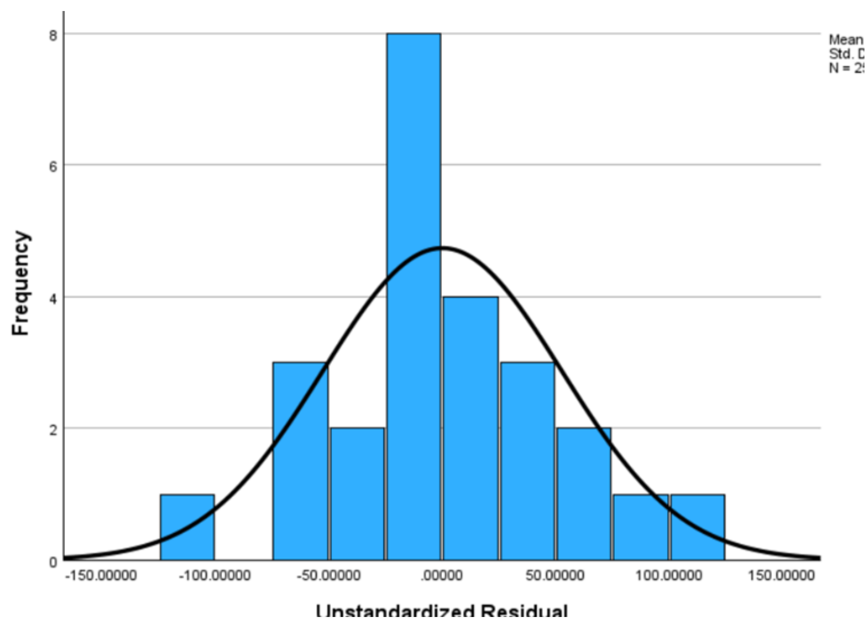

Figure 102

Histogram for the Normality Assumption

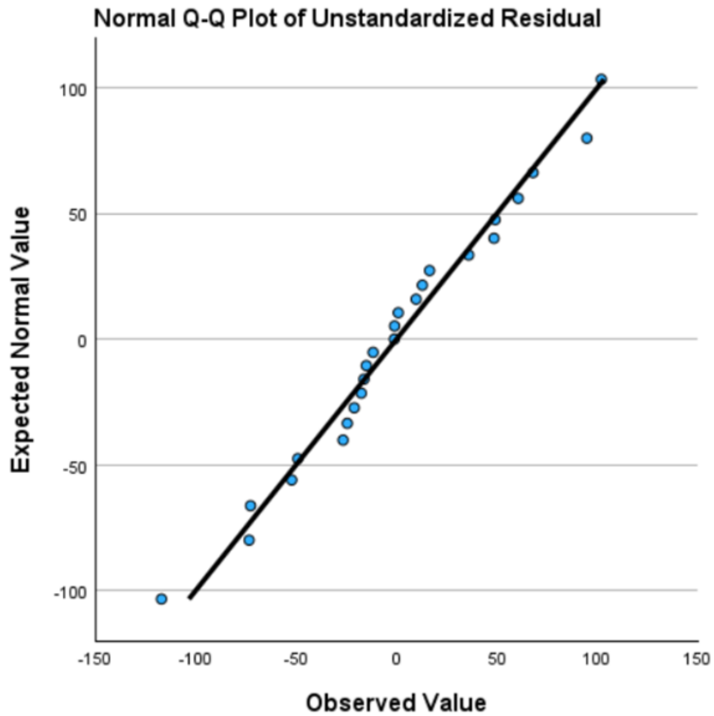

Figure 11

Q-Q Plot for the Normality Assumption

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.973[a]	.947	.936	57.62545	1.959

a. Predictors: (Constant), D2, bonus, adv, D1

b. Dependent Variable: sales

Figure 12

SPSS Model Summary with Durbin-Watson Statistic for the Independence Assumption

ZRE_1
.85250
-1.27309
.84212
1.05134
.01539
-1.26051
-.20204
-.26058
-.01916
-.90508
-2.03008
.28565
.22336
-.46138
-.30373
-.01673
-.42505
.62349
1.64447
-.36516
-.28134
1.76926
.16935
1.17934
-.85236

Figure 13

Computed Standardized Residuals for Outliers

LEV_1
.25569
.16364
.13373
.15399
.15105
.17805
.12061
.25603
.17528
.08654
.37272
.15655
.14211
.16903
.18577
.12828
.12858
.08523
.08652
.31221
.12437
.19109
.09570
.06101
.08622

Figure 14

Computed Leverage Values for High Leverage Values

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	Upper Bound	Collinearity Statistics Tolerance	VIF
1	(Constant)	177.207	170.116		1.042	.310	-177.648	532.063		
	adv	1.368	.262	.444	5.216	<.001	.821	1.915	.367	2.728
	D1	257.892	48.413	.565	5.327	<.001	156.904	358.879	.236	4.235
	bonus	.975	.481	.122	2.028	.056	-.028	1.978	.737	1.357
	D2	48.146	32.801	.100	1.468	.158	-20.277	116.568	.567	1.763

a. Dependent Variable: sales

Figure 15

SPSS Output of VIFs for Multicollinearity

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.963[a]	.927	.920	64.43912

a. Predictors: (Constant), D1, adv

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1157620.943	2	578810.471	139.392	<.001[b]
	Residual	91352.797	22	4152.400		
	Total	1248973.740	24			

a. Dependent Variable: sales

b. Predictors: (Constant), D1, adv

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	369.032	128.191		2.879	.009
	adv	1.565	.271	.508	5.784	<.001
	D1	237.079	40.090	.520	5.914	<.001

a. Dependent Variable: sales

Figure 16

SPSS Output of Reduced Model: Adv and D1

Transformations: Nested F-Tests

Squared Avd and Squared bonus ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1182559.896	4	295639.974	89.030	.000[b]
	Residual	66413.844	20	3320.692		
	Total	1248973.740	24			
2	Regression	1190589.144	6	198431.524	61.177	.000[c]
	Residual	58384.596	18	3243.589		
	Total	1248973.740	24			

a. Dependent Variable: sales

b. Predictors: (Constant), bonus, D2, adv, D1

c. Predictors: (Constant), bonus, D2, adv, D1, adv2, bonus2

$$\text{F-Stat}=\frac{\frac{(RSS_R - RSS_C)}{(k-r)}}{\frac{RSS_C}{(n-k-1)}} = \frac{\frac{(66413.844 - 58384.596)}{(6-4)}}{\frac{58384.596}{(25-6-1)}} = 1.238$$

df1=n-r =6-4= 2

df2=n-k-1=25-6-1= 18

CV=3.55

Adv and Squared Bonus ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1182559.896	4	295639.974	89.030	.000[b]
	Residual	66413.844	20	3320.692		
	Total	1248973.740	24			
2	Regression	1189727.168	5	237945.434	76.308	.000[c]
	Residual	59246.572	19	3118.241		
	Total	1248973.740	24			

a. Dependent Variable: sales

b. Predictors: (Constant), bonus, D2, adv, D1

c. Predictors: (Constant), bonus, D2, adv, D1, bonus2

$$\text{F-Stat}=\frac{\frac{(RSS_R - RSS_C)}{(k-r)}}{\frac{RSS_C}{(n-k-1)}} = \frac{\frac{(66413.844 - 59246.57)}{(5-4)}}{\frac{59246.57}{(25-5-1)}} = 2.298$$

df1=n-r=5-4= 1

df2=n-k-1=25-5-1= 19

CV=4.38

Transformations: Model Summaries

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.925[a]	.855	.842	90.74851

a. Predictors: (Constant), bonus, adv

Full Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.973[a]	.947	.936	57.62545

a. Predictors: (Constant), D2, bonus, adv, D1

Loge Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.972[a]	.945	.934	58.42418

a. Predictors: (Constant), logBonus, D2, logAdv, D1

Inverse Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.970[a]	.942	.930	60.35682

a. Predictors: (Constant), invadv, D2, invbonus, D1

b. Dependent Variable: sales

Inverse Bonus Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.974[a]	.948	.938	56.85208

a. Predictors: (Constant), invbonus, D2, adv, D1

b. Dependent Variable: sales

loge Bonus Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.973[a]	.948	.937	57.19653

a. Predictors: (Constant), logBonus, D2, adv, D1

b. Dependent Variable: sales