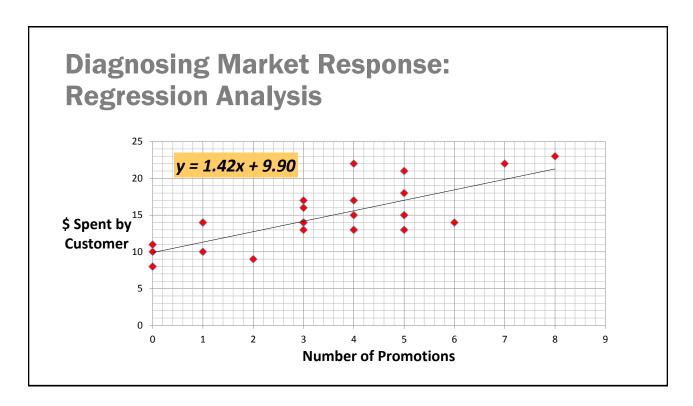
Marketing Analytics Module 4 Slides These materials are for your personal use while participating in this course. Please do not share or distribute them. Regression Basics

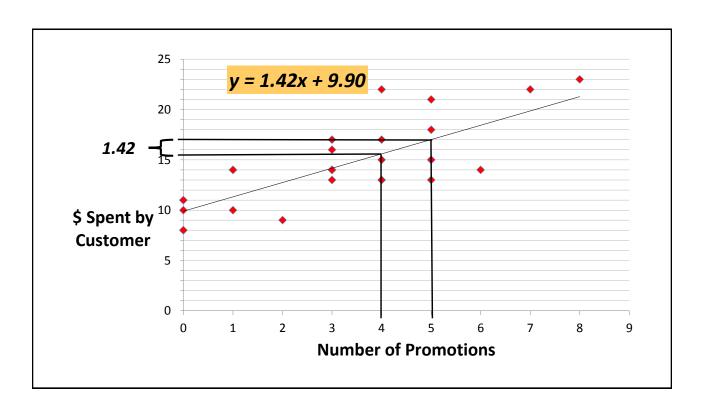
Introduction

- Regression is an important part of an analytics tool kit that allows us to understand how two variables are related.
- In this module we will
 - Discuss how to interpret regression outputs
 - Explore confounding effects and the biases introduced by missing variables
 - Distinguish between economic and statistical significance
- At the end of this module, you will be able to make inferences about customer behavior from regressions and connect them to business decisions.

These materials are for your personal use while participating in this course.

Please do not share or distribute them.





Developed by Raj Venkatesan for the University of Virginia's Darden School of Business Coursera Course: Marketing Analytics

These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Diagnosing Market Response: Regression Analysis

Regression Statistics				
Multiple R	0.775			
R-Squared	0.601			
Adjusted R-Squared	0.586			
Standard Error	2.566			
Observations	29			

ANOVA					
	df	SS	MS	F	Sig F
Regression	1	267.28	267.28	40.60	0.00
Residual	27	177.75	6.58		
Total	28	445.03			

	Coefficients	Standard Error		P-value
Intercept	9.90	0.85	11.60	0.00
Number of Promotions	1.42	0.22	6.37	0.00

Example: Simulated Shopper Card Data

Units purchased = a+b1*price paid + b2*feature + b3*display + error

Customer Price Paid Feature Display Units Purchased 1 1.50 0 0 3 1 2.56 1 1 1 1 2 2.41 1 0 3 3 2 2.37 0 1 1 1 1 1 2 2.23 0 1 1 1 1 1 2 2.265 0 0 0 0 0 0 0 2 2 2.06 1 0 2 2 2.06 1 0 2 2 2.206 1 0 2 2 2.212 1 1 2 2 3 2.31 0 1 1 1 3 3 1.69 1 1 3 3 1.82 0 0 2 2 3 1.54 0 1 3 3 1.54 0 1 3 3					•	
1 2.56 1 1 1 1 1 1.62 1 0 3 2 2.41 1 0 1 1 2 2.237 0 1 1 1 2 2.655 0 0 0 0 2 2.655 0 0 0 0 2 2.665 1 0 2 2 2.206 1 0 2 3 2.31 0 1 1 3 1.69 1 1 3 3 1.82 0 0 2 3 1.82 0 0 2 3 1.54 0 1 3 3 1.99 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1	Customer	Price Paid	Feature	Display	Units Purchased	
1 1.62 1 0 3 2 2.41 1 0 1 2 2.37 0 1 1 2 2.23 0 1 1 2 2.65 0 0 0 2 2.06 1 0 2 2 2.12 1 1 2 3 2.31 0 1 1 3 3 1.69 1 1 3 3 1.37 1 1 4 3 1.82 0 0 2 3 1.54 0 1 3 3 1.96 1 0 2 3 2.20 0 0 1 3 2.20 0 0 1 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1	1	1.50	0	0		3
2 2.41 1 0 1 1 2 2 2.37 0 1 1 1 1 2 2 2.23 0 0 1 1 1 1 2 2 2.65 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1	2.56	1	1		1
2 2.37 0 1 1 1 2 2 2.23 0 1 1 1 1 2 2 2.65 0 0 0 0 0 0 2 2 2.12 1 1 2 2 3 3 1.54 0 1 3 3 1.29 1 1 4 3 3 1.55 1 0 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 2 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 1 2 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 1 1 2 2 4 2.79 1 0 0 0 1 1 1 2 2 4 2.79 1 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	1.62	1	0		
2 2.23 0 1 1 1 2 2 2.65 0 0 0 0 0 0 2 2 2.06 1 0 1 1 2 3 3 2.31 0 1 1 1 3 3 3 1.69 1 1 1 3 3 1.54 0 1 3 3 1.96 1 0 2 2 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 2.20 0 0 1 3 3 3 2.20 0 0 1 3 3 3 2.20 0 0 1 3 3 3 2.20 0 0 1 3 3 3 2.20 0 0 1 3 3 3 2.20 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 0 1 3 3 3 2.20 0 0 0 0 1 3 3 3 2.20 0 0 0 0 1 3 3 3 2.20 0 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 3 3 2.20 0 0 0 1 3 2 4 2.79 1 0 0 0 0 4 2.79 1 0 0 0 0 4 2.15 0 0 0 0 1 1	2	2.41	1	0		1
2 2.65 0 0 0 0 2 2 2.06 1 0 2 2 2.12 1 1 1 2 2 3 2.31 0 1 1 3 3 1.69 1 1 1 3 3 1.82 0 0 2 2 3 1.54 0 1 3 3 1.29 1 1 4 4 3 1.96 1 0 2 3 2.20 0 0 1 3 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 1 2 4 2.15 0 0 0 1 1	2	2.37	0	1		
2 2.06 1 0 2 2 2.12 1 1 2 3 2.31 0 1 1 3 1.69 1 1 3 3 1.37 1 1 4 3 1.82 0 0 2 3 1.54 0 1 3 3 1.29 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.20 0 0 1 3 2.20 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 0 4 2.15 0 0 1						1
2 2.12 1 1 2 2 3 2.31 0 1 1 1 3 3 1.69 1 1 1 3 3 3 1.37 1 1 1 4 4 3 1.82 0 0 0 2 2 3 1.54 0 1 3 3 1.96 1 0 2 2 3 2.20 0 0 1 3 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 1 2 4 4 2.15 0 0 0 1 1	2	2.65	0	0		
3 2.31 0 1 3 1 3 3 1.69 1 1 1 3 3 3 1.69 1 1 1 3 3 3 1.82 0 0 0 2 2 3 3 1.54 0 1 1 3 3 3 1.29 1 1 1 4 4 3 1.96 1 0 2 3 2.20 0 0 0 1 3 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 2 4 2.79 1 0 0 0 4 2.15 0 0 0 1 1 4 2.15 0 0 0 1 1						2
3 1.37 1 1 4 3 1.82 0 0 0 2 3 1.54 0 1 3 3 1.29 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1			1	1		2
3 1.37 1 1 4 3 1.82 0 0 0 2 3 1.54 0 1 3 3 1.29 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1						1
3 1.82 0 0 2 3 1.54 0 1 3 3 1.29 1 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.79 1 0 0 0			1	1		3
3 1.54 0 1 3 3 1.29 1 1 4 3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1				_		4
3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1			-			2
3 1.96 1 0 2 3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1			0			3
3 2.20 0 0 1 3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1						4
3 1.55 1 0 3 3 2.01 0 1 2 4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0 1						2
4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0			-			1
4 2.07 0 1 2 4 2.79 1 0 0 4 2.15 0 0						3
4 2.79 1 0 0 4 2.15 0 0 1						
4 2.15 0 0 1						
4 2.15 0 0 1 4 2.50 1 0 1						
4 2.50 1 0 1						1
	4	2.50	1	0		1

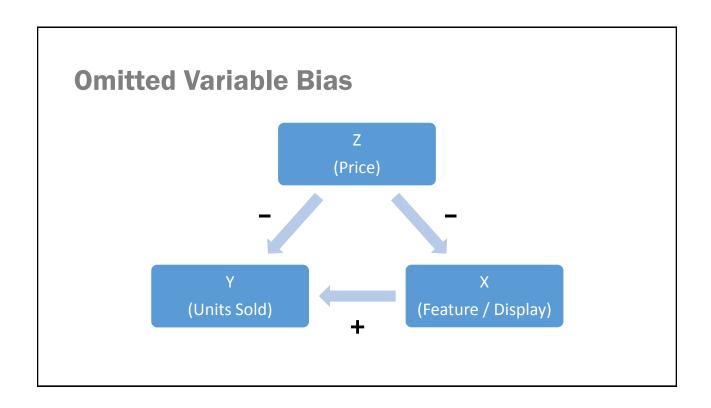
- Feature and Display:
 - 1 = Yes
 - 0 = No

These materials are for your personal use while participating in this course. Please do not share or distribute them.

Simulated Shopper Card Data: Regression Output

	True Model	Estimated Model
Intercept	6.28	1.34
Price	-2.31	_
Feature	0.38	0.822
Display	0.48	0.687
R-Squared	0.93	0.188

Why are the coefficients of feature and display different in the true and estimated models?

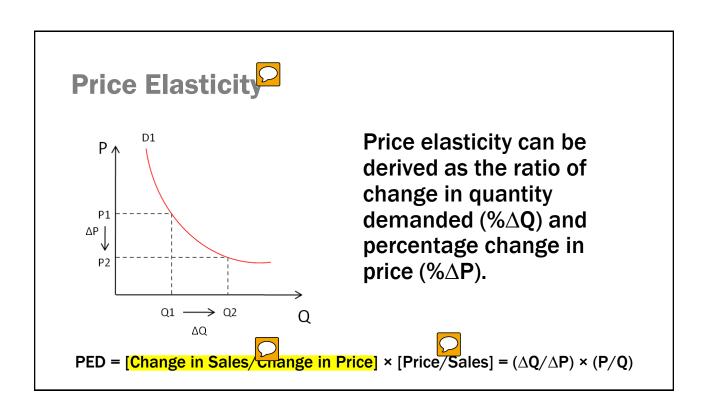


These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Simulated Shopper Card Data: Correlation Matrix

	Price	Feature	Display	Units Purchased
Price	1	-0.25	-0.24	-0.98
Feature		1	-0.09	0.45
Display			1	0.32
Units Purchased				1



These materials are for your personal use while participating in this course.

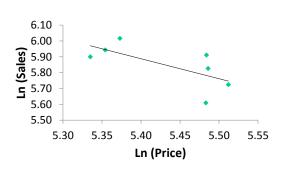
Please do not share or distribute them.

Belvedere Vodka

Year	Sales (units)	Ln(Sales)	Price (US dollars)	Ln(Price)	Advertising (US dollars)	Ln(Advertising)
2007	410	6.016	215.44	5.373	20486.1	9.93
2006	381	5.943	211.45	5.354	2923.5	7.98
2005	365	5.900	207.45	5.335	4826.3	8.48
2004	369	5.911	240.87	5.484	13726.6	9.53
2003	339	5.826	241.33	5.486	10330.2	9.24
2002	306	5.724	247.55	5.512	13473.6	9.51
2001	273	5.609	240.48	5.483	9264.6	9.13

Belvedere Price Elasticity

Regression Statistics				
Multiple R	0.67536			
R-Squared	0.45611			
Adjusted R				
Square	0.34733			
Observations	7			



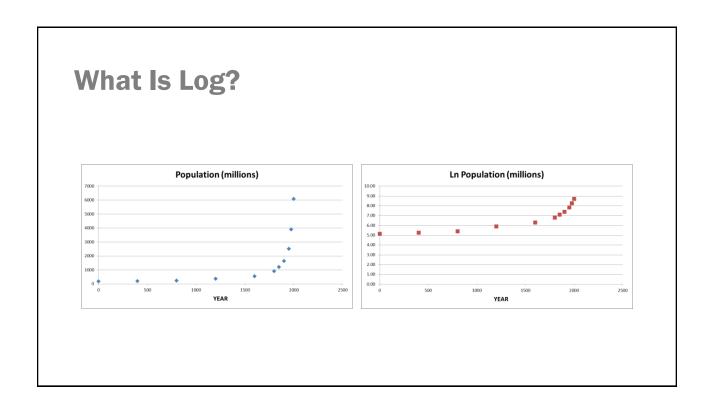
	Coefficients	Standard Error	t Stat	P-value
Intercept	12.686	3.340	3.798	0.013
Ln(Price)	-1.259	0.615	-2.048	0.096

These materials are for your personal use while participating in this course.

Please do not share or distribute them.

What Is Log?

Year	Population (millions)	Ln(population)
1	170	5.14
400	190	5.25
800	220	5.39
1200	360	5.89
1600	545	6.30
1800	900	6.80
1850	1200	7.09
1900	1625	7.39
1950	2500	7.82
1975	3900	8.27
2000	6080	8.71

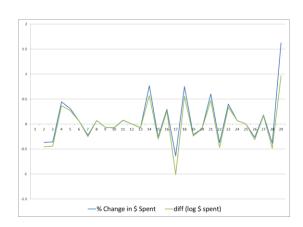


These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Log and Percentage Change

- First difference of natural LOG = percentage change:
 - Logging converts absolute differences into relative (i.e., percentage) differences.
 - The series DIFF(LOG(Y)) represents the percentage change in Y from period to period.



Elasticity – Log/Log Models

Dependent Variable : In (\$ Spent)

	Coefficients	Standard Error	t Stat	P-value
Intercept	2.24	0.07	32.06	0.00
log(num promo + 1)	0.32	0.05	6.44	0.00

0.317 = change in log (\$ spent) when log(num promo) increases by 1 unit



log(\$ spent) when log(num promo) is 0 = 2.236 (1)

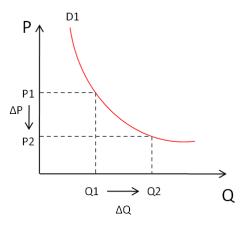
log(\$ spent) when log(num promo) is 1 = 2.553 (2)

(1) - (2)0.317

These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Price Elasticity



Price elasticity can be derived as the ratio of change in quantity demanded ($\%\Delta Q$) and percentage change in price ($\%\Delta P$).

PED = [Change in Sales/Change in Price] × [Price/Sales] = $(\Delta Q/\Delta P)$ × (P/Q)

Common Variables to Consider in Marketing Mix Models □

	Bias in	Bias in
Factor	Price Elasticity	Advertising Elasticity
Product quality	+	
Distribution 🔎	_	
Brand life cycle – early	+	
Absolute sales	+	
Time series	-	-
Include carryover 💭		+

These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Statistical and Economic Significance

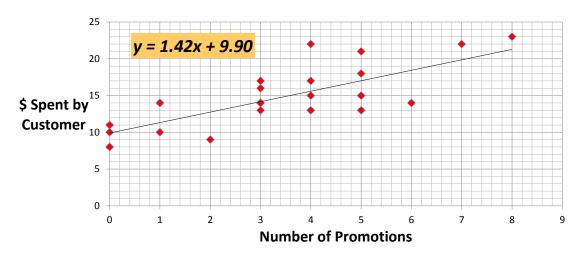
Statistical significance:

- Is the relationship observed in the sample likely to be observed in the population as well
- **■** Look for p-value < .10 for the coefficient of interest

Economic significance:

■ Does the benefit from a marketing intervention (i.e., the size of the coefficient) justify the expense?





These materials are for your personal use while participating in this course.

Please do not share or distribute them.

Diagnosing Market Response: Regression Analysis

Regression Statistics					
Multiple R	0.775				
R-Squared	0.601				
Adjusted R-Squared	0.586				
Standard Error	2.566				
Observations	29				

ANOVA					
	df	SS	MS	F	Sig F
Regression	1	267.28	267.28	40.60	0.00
Residual	27	177.75	6.58		
Total	28	445.03			

	Coefficients	Standard Error	t Stat	P-value
Intercept	9.90	0.85	11.60	0.00
Number of Promotions	1.42	0.22	6.37	0.00

Economic Significance

- A unit increase in number of promotions increases units purchased by 1.42
- Assume gross profit per unit is \$5
- Cost of promotion is \$0.50
- Profit = (units purchased * gross profit) (cost of promotion * number of promotions)
- Profit = (1.42 * 5 0.50 * 1) = (7.1 0.5) = 6.6

These materials are for your personal use while participating in this course. Please do not share or distribute them.

Conclusion

- Regressions are about what you include and also what you DON'T include in the model.
- Logarithm is a useful transformation for calculating elasticity from regression.
- Connecting regression to business decisions requires understanding economic significance.