Forecasting

Regression Methods
Linear Regression

Linear regression relates demand *(dependent variable)* to an *independent variable*

$$y = a + bx$$

$$a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

where:

$$\bar{x} = \frac{\sum x}{n} = \text{mean of the x data}$$

$$\bar{y} = \frac{\sum y}{n} = \text{mean of the y data}$$

Regression Methods
Simple Linear Regression Example

State University Athletic Department can you predict attendance based on the number of wins the team has experienced?

Wins	Attendance
4	36,300
6	40,100
6	41,200
8	53,000
6	44,000
7	45,600
5	39,000
7	47,500

x (wins)	y (attendance, 1,000s)	xy	x^2
4	36.3	145.2	16
6	40.1	240.6	36
6	41.2	247.2	36
8	53.0	424.0	64
6	44.0	264.0	36
7	45.6	319.2	49
5	39.0	195.0	25
7	47.5	332.5	49
49	346.7	2,167.7	311

Regression Methods
Linear Regression Example

$$\bar{x} = \frac{49}{8} = 6.125$$

$$\bar{y} = \frac{346.9}{8} = 43.34$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{(2{,}167.70 - (8)(6.125)(43.34)}{(311) - (8)(6.125)^2} = 4.06$$

$$a = \bar{y} - b\bar{x} = 43.34 - (.406)(6.125) = 18.46$$

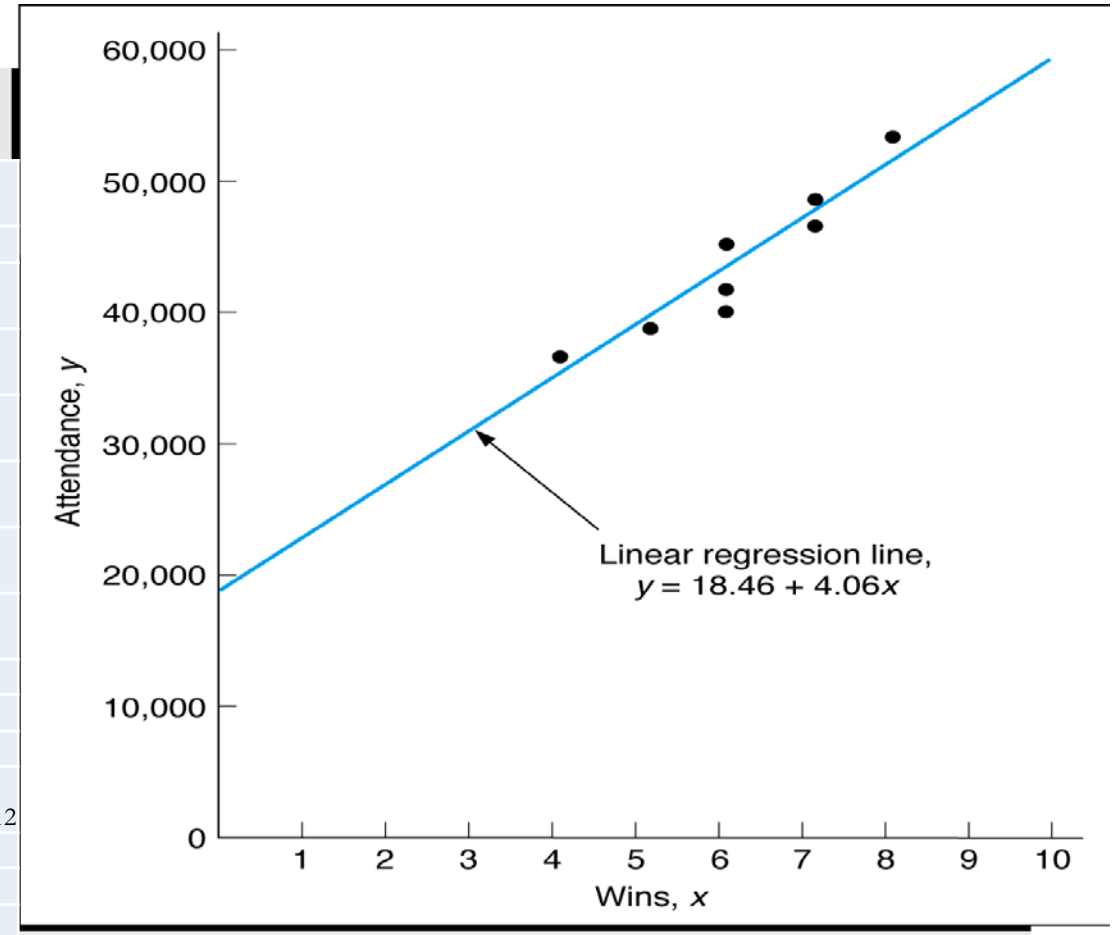
Therefore, $y = 18.46 + 4.06x$

Attendance forecast for x = 7 wins is
$y = 18.46 + 4.06(7) = 46.88$ or 46,880

Regression Methods
Linear Regression Example

Linear Regression Line

SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.9478			
R Square	0.898324			
Adjusted R Square	0.881378			
Standard Error	1.839311			
Observations	8			
ANOVA				
	df	SS	MS	F
Regression	1	179.3403592	179.3404	53.0112
Residual	6	20.2983908	3.383065	
Total	7	199.63875		
	Coefficients	Standard Error	t Stat	P-value
Intercept	18.46437	3.477568937	5.309562	0.001813
Wins	4.06092	0.557751243	7.280879	0.000342



Regression Methods
Correlation

- *Correlation is a measure of the strength of the relationship* between independent and dependent variables

Formula:

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[\, n\Sigma x^2 - (\Sigma x)^2 \,][\, n\Sigma y^2 - (\Sigma y)^2 \,]}}$$

- Value lies *between +1 and -1*
- Value of *zero indicates little or no relationship* between variables
- Values near 1.00 and -1.00 indicate strong linear relationship

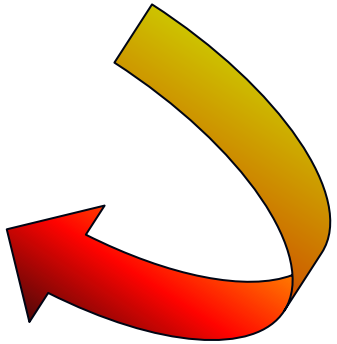
Regression Methods Correlation

Value for State University example:

$$r = \frac{(8)(2{,}167.7) - (49)(346.7)}{\sqrt{[(8)(311) - (49^2)][(8)(15{,}224.7) - (346.7)^2]}} = 0.948$$

SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.9478			
R Square	0.898324			
Adjusted R Square	0.881378			
Standard Error	1.839311			
Observations	8			
ANOVA				
	df	SS	MS	F
Regression	1	179.3403592	179.3404	53.0112
Residual	6	20.2983908	3.383065	
Total	7	199.63875		
	Coefficients	Standard Error	t Stat	P-value
Intercept	18.46437	3.477568937	5.309562	0.001813
Wins	4.06092	0.557751243	7.280879	0.000342

Or, you could simply read your regression output from Excel



Regression Methods
Coefficient of Determination

- The *coefficient of determination* is the *percentage of the variation in the dependent variable* that results from the independent variable

- Computed by *squaring the correlation coefficient, r*

For State University:

r = 0.948 => r^2 = 0.899

SUMMARY OUTPUT				
Regression Statistics				
Multiple R	0.9478			
R Square	0.898324			
Adjusted R Square	0.881378			
Standard Error	1.839311			
Observations	8			
ANOVA				
	df	SS	MS	F
Regression	1	179.3403592	179.3404	53.0112
Residual	6	20.2983908	3.383065	
Total	7	199.63875		
	Coefficients	Standard Error	t Stat	P-value
Intercept	18.46437	3.477568937	5.309562	0.001813
Wins	4.06092	0.557751243	7.280879	0.000342

- This value indicates that 89.9% of the amount of variation in attendance can be attributed to the number of wins by the team, with the remaining 10.1% due to other, unexplained, factors

Regression Analysis with QM for Windows

Forecasting Results

State University Athletic Department Summary

Measure	Value
Error Measures	
Bias (Mean Error)	.001
MAD (Mean Absolute Deviation)	1,412.644
MSE (Mean Squared Error)	2,537,298.0
Standard Error (denom=n-2-0=6)	1,839.311
MAPE (Mean Absolute Percent Error)	.033
Regression line	
Dpndnt var, Y = 18,464.37	
+ 4,060.919 * X1	
Statistics	
Correlation coefficient	.9478
Coefficient of determination (r^2)	.8983

Multiple Regression with Excel

Multiple regression relates demand to *two or more independent* variables

General form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

where: β_0 = the intercept

$\beta_1 \ldots \beta_k$ = parameters representing contributions of the independent variables

$x_1 \ldots x_k$ = independent variables

Multiple Regression with Excel

State University example:

Given the # of wins and the amount spent on promotion by the football department (which could have gone to purchase lab equipment instead <sigh>), can you predict attendance?

Wins	Promotion ($)	Attendance
4	29,500	36,300
6	55,700	40,100
6	71,300	41,200
8	87,000	53,000
6	75,000	44,000
7	72,000	45.600
5	55,300	39,000
7	81,600	47,500

Multiple Regression with Excel

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.949183
R Square	0.900949
Adjusted R Square	0.861329
Standard Error	1.988687
Observations	8

ANOVA

	df	SS	MS	F	Significance F
Regression	2	179.8644	89.93218	22.73956	0.003088
Residual	5	19.77439	3.954877		
Total	7	199.6388			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	19.09442	4.139282	4.61298	0.005772	8.454061	29.73479	8.454061	29.73479
Promotion (K)	0.03689	0.101346	0.363999	0.730744	-0.22363	0.297409	-0.22363	0.297409
Wins	3.560996	1.499981	2.374027	0.063634	-0.29483	7.416821	-0.29483	7.416821

Problem Time!
Building Products Store

For the data on the following slide:

- Develop a linear regression model - forecast amount of lumber needed, given some knowledge of permits issued
- Determine the strength of the linear relationship using correlation
- Determine a forecast for lumber given 10 building permits in the next quarter

Example Problem Solution
Building Products Store

Quarter	Building Permits, x	Lumber Sales (1,000s of bd ft), y
1	8	12.6
2	12	16.3
3	7	9.3
4	9	11.5
5	15	18.1
6	6	7.6
7	5	6.2
8	8	14.2
9	10	15.0
10	12	17.8