

REGRESSION MODELS

TEAM	W	ERA	AVG
New York Yankees	103	4.26	0.283
Los Angeles Angels	97	4.45	0.285
Boston Red Sox	95	4.35	0.270
Minnesota Twins	87	4.50	0.274
Texas Rangers	87	4.38	0.260
Detroit Tigers	86	4.29	0.260
Seattle Mariners	85	3.87	0.258
Tampa Bay Rays	84	4.33	0.263
Chicago White Sox	79	4.14	0.258
Toronto Blue Jays	75	4.47	0.266
Oakland Athletics	75	4.26	0.262
Cleveland Indians	65	5.06	0.264
Kansas City Royals	65	4.83	0.259
Baltimore Orioles	64	5.15	0.268

MONTH	DJIA	STOCK 1	STOCK 2
1	11,168	48.5	32.4
2	11,150	48.2	31.7
3	11,186	44.5	31.9
4	11,381	44.7	36.6
5	11,679	49.3	36.7
6	12,081	49.3	38.7
7	12,222	46.1	39.5
8	12,463	46.2	41.2
9	12,622	47.7	43.3
10	12,269	48.3	39.4
11	12,354	47.0	40.1
12	13,063	47.9	42.1
13	13,326	47.8	45.2

- (c) Which of the two models is better for predicting the number of victories?
- (d) Develop a multiple regression model that includes both ERA and batting average. How does this compare to the previous models?

**Q:** 33

The closing stock price for each of two stocks was recorded over a 12-month period. The closing price for the Dow Jones Industrial Average (DJIA) was also recorded over this same time period. These values are shown in the following table:

- (a) Develop a regression model to predict the price of stock 1 based on the Dow Jones Industrial Average.
- (b) Develop a regression model to predict the price of stock 2 based on the Dow Jones Industrial Average.
- (c) Which of the two stocks is most highly correlated to the Dow Jones Industrial Average over this time period?

**Case Study**

**North-South Airline**

In January 2008, Northern Airlines merged with Southeast Airlines to create the fourth largest U.S. carrier. The new North-South Airline inherited both an aging fleet of Boeing 727-300 aircraft and Stephen Ruth. Stephen was a tough former Secretary of the Navy who stepped in as new president and chairman of the board.

Stephen's first concern in creating a financially solid company was maintenance costs. It was commonly surmised in the airline industry that maintenance costs rise with the age of the aircraft. He quickly noticed that historically there had been a significant difference in the reported B727-300 maintenance costs (from ATA Form 41s) both in the airframe and engine areas between Northern Airlines and Southeast Airlines, with Southeast having the newer fleet.

On February 12, 2008, Peg Jones, vice president for operations and maintenance, was called into Stephen's office and asked to study the issue. Specifically, Stephen wanted to know whether the average fleet age was correlated to direct airframe maintenance costs, and whether there was a relationship between average fleet age and direct engine maintenance costs.

Peg was to report back by February 26 with the answer, along with quantitative and graphical descriptions of the relationship.

Peg's first step was to have her staff construct the average age of Northern and Southeast B727-300 fleets, by quarter, since the introduction of that aircraft to service by each airline in late 1993 and early 1994. The average age of each fleet was calculated by first multiplying the total number of calendar days each aircraft had been in service at the pertinent point in time by the average daily utilization of the respective fleet to total fleet hours flown. The total fleet hours flown was then divided by the number of aircraft in service at that time, giving the age of the "average" aircraft in the fleet.

The average utilization was found by taking the actual total fleet hours flown on September 30, 2007, from Northern and Southeast data, and dividing by the total days in service for all aircraft at that time. The average utilization for Southeast was 8.3 hours per day, and the average utilization for Northern was 8.7 hours per day. Because the available cost data were calculated for each yearly period ending at the end of the first quarter, average fleet age was calculated at the same points in time.

The fleet data are shown in the following table. Airframe cost data and engine cost data are both shown paired with fleet average age in that table.

### Discussion Question

1. Prepare Peg Jones's response to Stephen Ruth.

**Note:** Dates and names of airlines and individuals have been changed in this case to maintain confidentiality. The data and issues described here are real.

#### North-South Airline Data for Boeing 727-300 Jets

NORTHERN AIRLINE DATA				SOUTHEAST AIRLINE DATA		
YEAR	AIRFRAME COST PER AIRCRAFT(\$)	ENGINE COST PER AIRCRAFT(\$)	AVERAGE AGE (HOURS)	AIRFRAME COST PER AIRCRAFT(\$)	ENGINE COST PER AIRCRAFT(\$)	AVERAGE AGE (HOURS)
2001	51.80	43.49	6,512	13.29	18.86	5,107
2002	54.92	38.58	8,404	25.15	31.55	8,145
2003	69.70	51.48	11,077	32.18	40.43	7,360
2004	68.90	58.72	11,717	31.78	22.10	5,773
2005	63.72	45.47	13,275	25.34	19.69	7,150
2006	84.73	50.26	15,215	32.78	32.58	9,364
2007	78.74	79.60	18,390	35.56	38.07	8,259

### Bibliography

Berenson, Mark L., David M. Levine, and Timothy C. Krehbiel. *Basic Business Statistics: Concepts and Applications*, 11th ed. Upper Saddle River, NJ: Prentice Hall, 2009.

Black, Ken. *Business Statistics: For Contemporary Decision Making*, 6th ed. John Wiley & Sons, Inc., 2010.

Draper, Norman R., and Harry Smith. *Applied Regression Analysis*, 3rd ed. New York: John Wiley & Sons, Inc., 1998.

Kutner, Michael, John Neter, Chris J. Nachtsheim, and William Wasserman. *Applied Linear Regression Models*, 4th ed., Boston; New York: McGraw-Hill/Irwin, 2004.

Mendenhall, William, and Terry L. Sincich. *A Second Course in Statistics: Regression Analysis*, 6th ed., Upper Saddle River, NJ: Prentice Hall, 2004.

### Appendix 1 Formulas for Regression Calculations

When performing regression calculations by hand, there are other formulas that can make the task easier and are mathematically equivalent to the ones presented in the chapter. These, however, make it more difficult to see the logic behind the formulas and to understand what the results actually mean.

When using these formulas, it helps to set up a table with the columns shown in Table 7, which has the Triple A Construction Company data that was used earlier in the chapter. The sample size ( $n$ ) is 6. The totals for all columns are shown, and the averages for  $X$  and  $Y$  are calculated. Once this is done, we can use the following formulas for computations in a simple linear regression model (one independent variable). The simple linear regression equation is again given as

$$\hat{Y} = b_0 + b_1X$$

Slope of regression equation:

$$b_1 = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$b_1 = \frac{180.5 - 6(4)(7)}{106 - 6(4^2)} = 1.25$$