Exploration 9A: Production Line Data

The WheelRight company manufactures parts for automobiles. The factory manager wants a better understanding of overhead costs at her factory. She knows that the total overhead costs include labor costs, electricity, materials, repairs, and various other quantities, but she wants to understand how the total overhead costs are related to the way in which the assembly line is used. For the past 36 months, she has tracked the overhead costs along with two quantities that she suspects are relevant (see data file “C09 Production.xls”):

* MachHrs is the number of hours the assembly machines ran during the month
* ProdRuns is the number of different production runs during the month

MachHrs directly measures the amount of work being done. However, each time a new part needs to be manufactured, the machines must be re-configured for production. This starts a new production run, but it takes time to reset the machine and get the materials prepared.

Your task is to assist the manager in understanding how each of these variables affects the overhead costs in her factory.

1. First, formulate and estimate two simple regression models to predict overhead, once as a function of MachHrs and once as a function of ProdRuns. Which model is better?
2. Would you expect that the combination of both variables will do a better job predicting overhead? Why or why not? How much better would you estimate the multiple regression model to be?
3. Formulate and estimate a multiple regression model using the given data. Interpret each of the estimated regression coefficients. Be sure to include the units of each coefficient.
4. Compute and interpret the standard error of estimate  and the coefficient of determination . Examine the diagnostic graphs “Fitted vs. Actual” and “Residuals vs. Fitted”. What do these tell you about the multiple regression model?
5. Explain how having this information could help the manager in the future.