

Case - Absenteeism

Hi, I'm Pam Poovey, the director of Human Resources for Ingels Sherman International Shipping. Recently I have been concerned about how often our workers have been taking days off for either personal or family health reasons (i.e. Family and Medical Leave). The company has a very generous leave policy, but I have noticed that many more people are using the leave policy than they used to. I suspect that at least some of the employees have been taking advantage of the company's policy but I would like to see if there is some sort of pattern in what kind of employee is more likely to take days off. Our department has recorded whether or not each employee took a leave day last quarter along with some basic information that we think may be related for each employee. We recorded their current salary, whether the employee is full time or part time, age, and whether or not the employee got a raise or promotion in the last five years.

Because this project is sensitive, we would like you to be an external consultant to help us analyze this data. Primarily, we would like to know if any of the factors we have access to can help us identify employees who are most likely to use a leave day. Secondly, we would like to know how accurately the data can be used to predict whether or not an employee is likely to take a leave day during a given three-month period. As you might imagine, we want to avoid talking to employees to further investigate this issue unless we are very confident that they are likely to use the leave policy in three months. For that reason, before we use the model you develop to start investigating possible abuse the policy, is there a way to tune the model so that at most 10% of the people who don't use the leave policy are incorrectly predicted by the model to use the leave policy? If the model can be tuned that way, how good will the model be at accurately identifying people who actually will use the leave policy in three months if we do that?