Exploration 3B: Gender Discrimination Analysis with Pivot Tables

To help understand how pivot tables work and can help you analyze data to explore a problem context, we will consider a small private company called EnPact that produces environmental impact statements. (Basically, when a company wants to build in an area or manufacture a product, impact statements help predict the expected impact of this work on the local ecology.) Recently, the company has been sued by a group of female employees on the grounds that males have an unfair advantage in the salary process. By exploring this problem using pivot tables, you will learn a fundamental truth about data mining: the deeper you explore, the more you are forced to reconsider each and every piece of evidence you have.

The company salary data (and employee profiles) are in the file C03 EnPact.xls. Open this file. Using simple pivot tables (see the How To Guide for this section), answer the following questions.

1. How many male employees are there? How many female employees? What percentage of the employees is male? Female?
2. What is the average male salary? What is the average female salary?
3. Based on your answers to #1 and #2, write a sentence or two discussing the company’s lawsuit.
4. Cross-section the data by both Gender and Education Level. Look at the average salaries of the employees and discuss the company’s lawsuit.
5. Cross-section the data by both Gender and Job Level. What does the lawsuit look like now?
6. For a final look at just how complex this issue is, cross-section on three variables simultaneously. Set the pivot table up with Gender as the row variable, Education Level as the column variable, Job Level as the "page variable" (at the top of the pivot table) and average salary as the data.
7. Using the three-variable pivot table, pull down the "Job Level" menu and look at each job level separately in the pivot table. Are there any particular job levels where the male and female salaries, after accounting for education, are roughly the same? Are there any where the salaries are quite different?
8. Select one of the job levels that shows a large difference in salaries by gender. Go back to the original data. Can you account for these differences by looking at the numerical variables (Years of experience and Years Prior)?