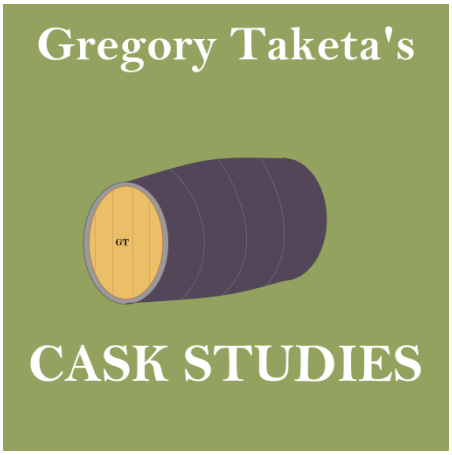
ANALYZE RESPONSIBLY

Welcome to Cask Studies, where you can properly age your skills without getting old. Even sour grapes can become fine wines here.



December 1, 2014

Weight Loss (Regression – Cumulative, Lag Variables)

A fictional case study by Gregory Taketa. For non-data Managers, this case illustrates that your analysts will not be able to solve anything as complicated as weight loss using typical linear regression methods. Regression practitioners can use a MS Excel Data Set to approach this problem by transforming the variables with respect to growth and time.

I'm sure you or somebody you know has an interest in a healthier diet or lifestyle (I've lost 25 lbs. in 19 months by the time of this writing).

A friend of mine has a fantastic web site for those of you who are interested in low-carb recipes:

findingtheweigh.com

This site has many visitors, and you gotta visit at least once just for the clever name!

Weight Lifting

As you know, weight lifting is very helpful for weight loss because an increase in muscle mass, all else equal, yields higher fat burn. This burn happens even when you are not actively exercising, so in the diet world, this is comparable to passive income!

Gregory Taketa is the Data Decanter, serving refined, well-breathed data analysis while keeping out the sediments.

E-mail: greg@gregariousconsulting.com
See more of The Data Decanter at: http://gregariousconsulting.com

The Case: Estimating the Impact of Weight-Lifting

Click Here to Download Data (MS Excel 2010+)

Our subject weighs 200 lbs. at the start of a diet. At the end of each week, he marks his weight and whether he did a diligent amount of weight-lifting (according to his health advisor) that week. As typical of any binary (dummy) variable, "1" indicates "yes," and a "0" indicates "no."

Although many other variables, including calorie intake versus metabolic requirements, influence body weight, we will assume for simplicity that these variables have remained fixed.

Key Case Facts

- For the first 8 weeks, our subject has lifted weights but notices no weight loss.
- After the 8th week, he gives up on lifting weights (he perceived no weight loss earlier).
- He exhibits some weight loss from weeks 9-16. This is because his gain in muscle mass takes time, about 8 weeks in this case.* He does not know about this.

**Disclaimer: I am not a fitness expert of any sort, and this case is simplified for the purpose of quick practice. I do not have factual evidence and do not suggest that 8 weeks is required to develop muscle mass.*

Did the weight-lifting help our subject lose weight? How could you demonstrate it? What was the rate of weight loss for the given effort?

Cask Questions:

1) Review the raw data for 30 weeks in Tab Q1. We know that our subject has accurate weight readings at the end of each week, and he accurately recorded the 8 weeks in which he diligently lifted weights.
    a. What is problematic about the current data set? If you are unsure, run a regression analysis to find out. Looking back at the raw data (especially the output), what do you think the pattern should be?
    b. How can we use the current input variable to create a better input variable which accounts for the gradual, cumulative weight loss? Do not worry about the delayed muscle mass growth for now.
2) Run a new regression using the improved input variable from 1b. Is the coefficient to your liking? The sign? The magnitude? Which of these two could be improved quickly?
3) We now invoke our knowledge that the muscle mass grows 8 weeks after the weight-lifting (e.g. week 1's effect materializes in week 9, and week 8's effect materializes in week 16). How can we change the input variable to account for this delay? Did your estimate improve?

Eis Igian!