Multiple Regression Replaces the Need for a Tailor

Allison Davidson

Muhlenberg College, 2400 Chew St, Allentown PA 18104

Abstract

One of the challenges with online clothing shopping is the inability to try on the clothing. CottonBrew aims to use the extraordinary power of statistics to provide the luxurious service of a tailor-made suit through a few simple clicks of a mouse in the comfort of your home. After the author was approached by the CEO of CottonBrew seeking advice on enhancing his model to predict suit sizes for his online custom-suit company, a class project emerged. Ideally, CottonBrew would like to have a small handful of known measurements for the customer to input. From those measurements the site could extrapolate the additional measurements needed to create a suit made to fit the customer of interest. Students were trained and calibrated how to take 13 different measurements using a mannequin provided by CottonBrew; creating potential models for a customized suit.

Key words: Multiple Regression, Class Project, Education, e-commerce

1. Introduction

In Fall of 2015, I completed a class project with my multivariable modeling statistics class through collaboration with George Li, the CEO of CottonBrew; an online menswear clothing company. The course uses the textbook Stat2: Building Models for a World of Data, beginning in Unit A, linear regression. The class project I discuss was intended to use the first unit of the course to allow students to use what they've been learning in the classroom to answer a real problem.

I was first approached by George while I was developing the course over the summer of 2015. George was looking for assistance on creating a model for predicting clothing sizes and had found online materials I had developed while I was a graduate student. After discussing the scope of the request with George, we decided to implement the request as a class project. We collaborated on what the project would entail, got IRB approval for data collection, and implemented the project at the beginning of the semester.

1.1 Overview of CottonBrew

CottonBrew is an online mensware clothing company, specializing in customized design, fabric, and fit of professional attire. The goal of CottonBrew is to create clothing that feels custom-made for each customer.

> *"Our mission, and my life's passion, is to help all people find their personal fit and experience the lost art of tailoring so they too can look and feel great, effortlessly."*
> *George Li, CEO of CottonBrew.*

1.2 Outlining the Question

As an online company, ease of use and speed are essential. Ideally, the amount of time a customer would need to spend on the site before completing their transaction would be minimized. In addition, their experience should be easy and enjoyable. As such, the information the customer provides CottonBrew should be minimal and accurate in order to create a customized suit the customer will be happy with. Two variables that meet these criteria are height and weight, as people tend to know these measurements and they tend to be relatively accurate. Additionally, the customer is asked for their fit preference; classic, slim, or extra slim. To develop a suit that would best fit the customer, 13 different measurements need to be predicted, measurements typically taken by a tailor (e.g. neck circumference, shoulder width, sleeve length, etc.). The goal of the class project was therefore to use height and weight, as well as additional predictor variables that are both widely known and accurate, in order to develop different statistical models to predict the 13 different measurements used to make the customized suit.

2. Beginning the Project

The first step of the project was to decide what predictor variables would be needed to answer the question of interest, and prepare to obtain the necessary data. Students were required to brainstorm potential predictor variables (in addition to height and weight), and discuss their ideas for determining how to predict measurements to make a customized suit.

After each group of students had a solidified plan, all students were trained to obtain good and accurate data. George frequently uses the phrase "Garbage in – Garbage out", meaning that if the measurements taken are not accurate (garbage in), any suit created based on these measurements will not meet the standards of CottonBrew (garbage out).

To train students to take good measurements, George provided professional videos online which the students watched to learn how each measurement is taken. George mailed a mannequin from Los Angeles to Allentown PA, sporting handsome CottonBrew clothing. He provided predetermined measurements he had taken on the mannequin to serve as an "answer key" for effective calibration. The students practiced taking measurements on the mannequin and calibrated their accuracy using to George's measurements. Once students calibrated their measurements, they were ready to collect data.

The anticipated output for the project included 13 different multiple regression models, one for each of the measurements to be predicted, with explanatory variables including height, weight, other predictor variables, linear combinations of variables or interaction terms.

There were 3 different groups, and as a result 3 different multiple regression approaches. One group created 13 models as anticipated; one group created 13 models but used the intended response variables as predictor variables (i.e. in one model they are intending to estimate the shirt length, and in another model they use shirt length to predict shoulder width). The last group decided that chest circumference was the one response variable that would best give an overall suit size, and developed one model, with height and weight as predictor variables, as well as any measurements that were intended to be response variables.

3. Sample Models

The following represents one of the 13 models developed by Group 1 using the intended predictor variables as potential options for explanatory variables. This model has one of the highest adjusted R^2 values of the models they developed.

$$Thigh = 309 + 2.7(Wt) - 6.9(Ht) - 1.4(BMI) - 0.05(Wt*BMI)$$
$$Adj.\ R^2 = 0.7361$$

The following represents one of the 13 models developed by Group 2. Of note is that chest circumference (Chest), hip circumference (Hips), sleeve length (Sleeve), and waist circumference (Waist), were all intended to be response variables, but in this model were used as predictor variables. As the model was not as intended, additional metrics are not included.

$$Neck = 12.6 + 0.2(Chest) - 0.2(Hips) - 0.1(Sleeve) + 0.2(Waist)$$

The following represents the model developed by Group 3 estimating chest circumference, which would then be used to identify the overall suit size. Of note is that the three predictor variables, shoulder width (Shoulder), stomach circumference (Stomach), and pant length (Pant), are all intended to be response variables. As the model was also not as intended, additional metrics are not included.

$$Chest = 0.007 + 0.4(Shoulder) + 0.6(Stomach) + 0.3(Pant)$$

4. Conclusion and Future Ideas

At the conclusion of the project students held a Skype conference with George. During the call, George explained the goals and history of CottonBrew, and the students discussed their models. The students also provided ideas regarding where to obtain additional data to be used in future models.

Both the class project and CottonBrew ended the collaboration successfully! CottonBrew currently has 98% satisfaction, and the project, I believe, was also a success. If I were to do the project again, I would give the students more time to collect data, and perhaps additional incentives or accountability in collecting data. Additionally, I would implement more communication throughout the project between myself and the students, students and other groups, as well as between students and George. I believe if students had additional contact with George early on, they would have a deeper investment in the project. They would know who they were working for and have a clearer idea of what was being requested. Additionally, I think having classroom discussions would have aided in some of the models that were not quite accurate. The ideas may have been a bit more creative had they been given the opportunity to give and receive more feedback from other groups.

Acknowledgements

Intermediate Statistics class for all of their efforts on the project. Furthermore, I'd like to acknowledge Muhlenberg College and the supportive teaching initiatives, particularly the *Faculty Center for Teaching*; providing resources to create the course.

References

Li, George. Your private tailor. Retrieved September 28, 2016, from https://www.cottonbrew.com/