Title: Using R in Data Analysis: Examples Based on the Economic and Finance of the Entertainment Industry

Faculty: Thomas M. Smith, Ph.D. Goizueta Business School; Sloan Kennedy, Vice President, HBO

Location: Atlanta Based Module

From the Walking Dead to Trouble With the Curve, Georgia (and Atlanta) are becoming the Hollywood West. Studios are responding to the incentives put in place to provide for 30% tax incentives for qualified spending on television, movies and other entertainment purposed. As a result, Georgia is spending additional resources to understand the economic impact of these activities. This mid-semester module will be an opportunity for Emory student to learn more about the economic and finance of the entertainment industry, examine industry data and try their hands at learning and utilizing R.

The Data and Exercises

Much like data in the sports industry, entertainment data is at our fingertips. As a result, we have the opportunity to build models to understand consumer behavior and to forecast consumer spending and activity. For this mid-semester module, students will download data from boxofficemojo.com to build models to gain a better insight into both studio decisions as well as consumer decision. During this module, students will examine the following problems:

1. Replication of Nate Silver’s estimate of the studio losses to SONY for not releasing The Interview.
2. Predicting the revenue from Neighbors and Neighbors 2 based on pre-production concept hypotheses;
3. Predicting the revenue for running season 8 of Game of Thrones as a duo or trilogy of movies rather than 6 television episodes it is slated for.

First, students rarely replicate the research and analysis of other industry experts. The replication of Nate Silver’s estimates will provide the students the opportunity to understand the fundamentals of the models. The students will then be asked to improve upon the model. For the Neighbors, I have retrieved from Sony the original Seth Rogan proposal with the set of comparables (movies upon which Rogan relies to ‘set the stage’ on the success of his movie). The students, knowing this information, will be asked to develop the model that the studio ‘likely’ used in its decision to produce the movie. Then, we can examine the underlying assumptions that the studio must have used in order to understand how assumptions are made and used in model building. Lastly, the students will be asked to build a model to predict how a trilogy of Game of Thrones movies will fare in the theaters relative to television broadcasts. This last section will be team-taught by Sloan Kennedy, Vice President of Development for Home Box Office.

Learning and Using R

R is an open-source program that can be used to analyze and examine data. Most of our full-time MBA students will be taught using excel and JMP. Both excel and JMP are terrific and offer some helpful model building techniques. But, the JMP license expires when our students leave Emory and excel has limited power with respect to Markov Chains, factor analysis, logit and probit models and other more
sophisticated model building. This mid-module will provide many of these students their first glimpse at R and it’s power as a tool in statistical analysis.

**Exposure to Industry Expertise in a Growing Industry**

Sloan Kennedy is a terrific resource for Emory. She has more than 15 years in the entertainment industry. She has spent a great deal of her career in international development and is working with distributing HBO’s cadre of original programing. Her perspective will provide our students will an understanding of a growing industry from a company that has been at the forefront of original and groundbreaking programming.

**Final Deliverable**

The students will be asked to identify a problem, find the data and build a model. An example of this problem is to identify what kinds of books make the best movies. Students will build a model and run the data in R and then present their finding to the class.

Using data provided by some of our Alumni in Los Angeles, we will run models examining several of the animated features produced by Sony Pictures. We will be testing audience response to elements in animated features such as female lead versus male lead, name recognition of lead characters (voiced by Seth Rogan or Adam Sandler) and run time.