Assignment 1

Bremians's articles talk mostly about the two-way of using the statistical modeling in order to reach their desired outcomes. In one section he talks about the data modeling culture where the culture starts with a lot of collected data that is already created into represented boxes to best make a goodness-to-fit-test. While the Algorithmic Modeling culture does not have the known data as it uses substitute variables in order to report the responses. In either case, the use of these depends on what the desired goals of the program are in order to best understand and utilize these models. He speaks mainly about his own experiences and how some of these techniques are greatly underutilized, and that there needs to be more reach in universities and world places alike in order to better understand statistical research. The problems of current data holding come from the issues of goodness-to-fit-test as models are very little power unless they are particularly the focus of said study. This draws more misleading conclusions that pass the goodness-to-fit-tests and only further limits the data models use as multivariate tools.

Shmueli's articles in similar in that it also talks about two types of statical modeling that are explanation and prediction and what uses these provide for analyzers. He also talks about how these premises have been misused and further misunderstood in how they work and what they are made for. He explains that Explanatory modeling is to be used exclusively testing causal theory and that there does not need an abundance of data for these. This is mad to explain what data they have and have a deep dive into eh theories that they have created around it. The predictive model on the other hand is considered to use data to make future predictions in order to understand where the program is going. This requires far more data in order to get as accurate data as possible and make it useful. He also laments the lack of validity that the predictive model has when compared to the explanatory model. As the explanatory model is used frequently for scientific reach to justify their theories but the predictive model is not seen in the same value.

Shmueli's article had far more examples in his understanding of how the statical modeling of explanation and prediction have been mistreated within the statical field than Bremians's. Bremians spoke of personal experiences and did have in-depth examples to support his argument and while I did not find anything amiss in his argument, I found Shmueli mot to be more compelling. Shmueli's multiple examples and clear explanations gave way to way to the average reader to understand their point of view. While both articles made good points that there are shortcomings within the data gathering field, I think that there should be more concern with Shmueli's arguments the lack of understanding seems a far more detrimental to the field. The lack of statical models is still a great issue however if the models can be understood and found to be befitting of the study that is a different argument from completely not utilizing the statical models.

Review the sample projects in ISLR and suggest data project examples using your own experience and data (e.g. cancer gene identification, spam classification, income etc.)

a. Present a proposed hypothesis
    I. Movie ticket sales, like saying what accounts to increase pricing

b. How data can be collected
    I. This data ca be collected by using the movie ticket online pricing and selling's, as all purchases are recorded it would be easy to find out the when, where, and how much.
c. What methods could be considered?
    I. It would be best to use a straightforward approach in order to analysis the data collected. Like many linear regression in order to understanding the tickets with the pricings easier.
d. How to start the data project?
    I. It would be best to start with finding ticket pricings in a cohesive manner that can track from a certain location and can be comparable to other ticket prices. What important is that these prices are able to be measured and analysis with other ticket prices, making it coherent.