PowerPoint Organization

1. Introduction

(Camilo)

\*Note to class: “Before we begin, please take your laptops out and go to this link:”  
 <https://criviere.github.io/ncaa_camilo_armando>

* Who? (Both Armando & Camilo)  
  + Now we will begin by introducing ourselves. Hi, my name is Camilo Riviere and I’m part of the Big Data Analytics program. I’m in this program to enhance my current skill set in Data Science. In addition to enhancing my current skill set I’m also improving my job-related skills and improving my chances of moving up the technical ladder at my workplace, Next Era Energy (Florida Power & Light).
  + Hi, my name is Armando Zapata and I’m also part of the Big Data Analytics program. I’m in this program to solidify my foundation on my skill-sets in Data Science. I’m looking forward to applying these skills in the investment banking industry that I’m currently working in. (Cross Keys Capital).
* What? What is our project? (Camilo)
  + Our project is a prescriptive analysis containing 2 models (Total score, and difference) of the NCAA (National Collegiate Athletic Association) Division I (NCAA Division I is the highest level of intercollegiate athletics sanctioned by the National Collegiate Athletic Association in the United States.) Men’s College Basketball. In this project we are predicting three things: the outcome (inferred from the score of each individual team), the overall score in the matchup (predicted by the total model), the score of each individual team in the matchup which is calculated by a difference model, and the implied probability of home team winning the game.
* Where? Where are some examples for practice use cases of our model? (Armando)  
  + Prescriptive models regarding predicting outcomes of a specific sporting event are in high demand in geographic locations such as Las Vegas where sports betting is a billion-dollar business (annually). Veteran Sports bettors and sports betting companies alike are actively looking for any kind of competitive edge and are willing to pay top dollar for highly accurate prediction models.
  + Outside of the sports industry. Large retail corporations such as Target or Wal-Mart may eventually garner interest in pre-determining whether to manufacture merchandise after any sports team wins the championship for their respective sport. Often, these retailers will have both sets of championship clothes manufactured for both the winning & losing team which inevitably creates waste because the clothes for the losing team is often discarded or sent overseas in donation to foreign countries such as Africa where there is a need for clothing.
* When? When can we use this model? (Camilo)
  + Because our stats are exclusively for Men’s Division I College Basketball our model is best used during both the regular seasons for College Basketball and the post season, or March Madness. Users can select the matchup for the day and proceed to make a prediction on the teams which are playing that day in the tournament.
* Why? Why was this project important to us? (Armando)
  + Given that we are both into sports, and March Madness is right around the corner, we thought it would be a great idea to immerse ourselves into this environment and apply some of the skills we’ve learned thus far in the Big Data Analytics program.
  + This is an event that is perceived as a phenomenon, which galvanizes people from different races, religion, beliefs, and spans across the world. To put it in to context, this event is so prominent that even the former President Barack Obama on National Television (ESPN) filled out a tournament bracket.
  + Additionally, the Chairman and CEO of Berkshire Hathaway Warren Buffet, offered $1 million in perpetuity to anyone who predicted a perfect Sweet 16 (Last 16 teams left in the tournament).
  + Also, given the complexity and unpredictability of accurately predicting this event, even Kaggle and Google host annual competitions for those who attempt to create models that accurately predict the outcome all the way to the champion of the entire tournament.
  + Note – Include picture of tournament bracket.
* Why? Why is this project important to the class? (Camilo)
  + In this project we introduced the concept of productionizing and hosting your model as an endpoint/user application. This is a concept that has yet to be covered in our program. By introducing this concept, we intend to see students in the future will see the value of this and apply it to their future projects as this will eventually be a client requirement in a real-life job setting.
  + We are also going to introduce the use of advanced statistics in the form of efficiency statistics which these types of statistics are proven to be more effective when performing an analysis in contrast to traditional box score statistics which are typically utilized by sports television analyst (FG%, 3PFG%, Blocks, Assists, etc.). By introducing this concept we intend to see students in the future utilize higher quality data which provide more insight into their area of research.