**Jigsaw Academy**

**Final Case Study Evaluation**

**The aim of this study is to come up with recommendations to a client that is in the horse racing industry on how to maximize handle (money bet on any race). The client has provided data that he has available, along with a data dictionary.**

**Data and the data dictionary are both loaded to the remote server under:**

**Z: \Assignments\Class17 - Final Case Study**

**The client does not know any statistics and so will need all results interpreted. He is particularly interested in the following tracks: track\_ids CRC, AP, and FG, so restrict the data to only these tracks. Also, restrict the years of analysis to 2005 and 2006 only.**

**Handle is derived by summing up wps\_pool in the race table with total\_pool in the exotic\_payoff table. Please also note the following:**

1. **Holidays are important because handle goes up a lot during holidays, so you will need to create holiday indicators for the following holidays in the US:**
   1. Super Bowl
   2. St Patrick’s day
   3. Easter
   4. Cinco de Mayo
   5. Memorial Day
   6. Independence day
   7. Labor day
   8. Veteran’s day
   9. Thanksgiving
   10. Christmas
   11. New Year Day
   12. Good Friday
   13. Boxing Day
2. Time of day (evening or daytime) is probably an important indicator
3. You will need to create categorical variables from some of the existing variables (race\_type, sex\_restriction etc) and probably also create buckets of variables like purse\_usa
4. Race date is in datetime format; to get the date part from race\_date use:
   1. Race\_dt = datepart(race\_date);
5. You will need to merge race and exotic payoff; merge by track\_id race\_date race\_number and day\_evening;
6. Results need to be presented for each track\_id as well as overall.

**Hint :** Please use cdi\_model\_data data only as reference. You need to prep the data using race, exotic\_payoff and all the track datasets, create dummies and come to a dataset that looks like cdi\_model data.

No need to use variables that are not given in the data dictionary.