



op/ed

by Joe Appelbaum

DATA POWER

Almost a year ago, the National Basketball Association (NBA) unveiled a new website, stats.nba.com, where they did something novel to improve their sport. They released a vast statistical trove, powered by a flexible and useful database allowing fans, researchers, fantasy gamers and even team executives to investigate and study all forms of basketball data—current and historical.

Extraordinarily, it is free.

In this era of restrictive intellectual property rights, why would a corporate monopoly like the NBA release all of its statistical data at no cost? Cleverly, they understand that forfeiting modest present-day revenues to deeply engage their core fans creates interest and interaction that propels the popularity of professional basketball. In other words, they hope it will lead to increased TV ratings, ticket sales, and apparel purchases—the money-making endeavors of its teams.

I have a modest proposal for horse racing: The Jockey Club and the Thoroughbred Racing Associations (TRA), through their subsidiary Equibase, should follow the example of the NBA and make available a comprehensive database with all breeding and racing information from the beginnings of our sport. The TJC and TRA should lead our industry into its digital future by building a highly functional, but accessible website that would allow fans, handicappers, academics and industry professionals the ability to access its database at the most basic level.

For the good of the sport, this information should be freely available.

What drives the popularity of the NFL, MLB and the NBA? Gambling and its close cousin, fantasy sports. Well, horse racing was the first fantasy sport. We have a 150 year history of data collection and a long standing respect for its analysis—think Lowe, Tesio, Ziemba, Beyer. Sadly, we have been left behind in the race for gaming dollars and consumer interest.

While this failure has many culprits (high takeout, inconsistent product quality, crumbling facilities, government regulation, etc.) the restrictive presentation and prohibitive cost of our information surely stifles fan participation and engagement, not to mention breeding and handicapping innovation.

Meanwhile, other sports are experiencing a revolution in data analysis—books and blogs proliferate, academics use sports data to publish in pursuit of PhDs. All this activity drives fandom to an obsessive interest in their sports. Most famously, MIT's Sloan School of Business holds an annual conference with thousands of attendees.

I've attended the last three conferences and have yet to meet anyone from horse racing there—not a handicapper, not a breeder, not a track executive.

The Jockey Club is an organization dedicated to the betterment of the breed and of horse racing. The TRA is charged with promoting the interests of its member tracks.

So why should they release this information for free? Quite simply, because it's good for the business of horse racing. In recent years, The Jockey Club has made selective investments in the “business” of horse racing. There is no better investment they could make than releasing its information for free. Nothing either organization could do with the revenues they seek from handicappers and breeders could rival the benefit the game would receive if the entirety of racing and breeding information was free to all.

Equibase has done a fine job of recording and tabulating the data. However, the analysis of that data lags far behind the sports analytics field. This is not surprising, Equibase is a monopoly and while that structure creates efficiencies in gathering, storing and controlling information, monopolies are typically poor at analytics. A very different skill set is needed to mine data for trends or revelations—this is where the skills of hobbyists, industry professional and academics can be powerful.

Equibase is sure to respond that they make a wide array of data available and much of it for free. Furthermore, there are a number of third-party providers (DRF, BRIS etc.) that present the data in a variety of formats. This is all true, however, that data is just not available in a format suitable for serious research. The dated formats and restricted access pose significant barriers, both monetary and practical, for doing the research necessary to spur the sort of engagement that we have seen drive massive interest in fantasy sports and gambling on the NBA, MLB and NFL.

Horseracing is sitting on mountains of digital breeding and racing information. However, in this age of big data, cheap processing and powerful databases, we are doing a poor job of mining the billions (yes, billions) of dollars buried inside those mountains.

Let's tear down the walls, and unleash the power of our data. [Click here to share.](#)

Joe Appelbaum is one of the two principals in Off The Hook, a breeding, racing and sales company with offices in New York City, and a farm and training center in Ocala, FL. He played football at Yale University, and went on to become a college football coach, spending seven years coaching at five different universities, most famously working with Wayne Chrebet and Victor Cruz. Joe was the President of Kids Interest Development Services Inc. and founded Oasis Children's Services LLC, which provided school, camp and after-school programming for thousands of children throughout the Northeast. Off The Hook has pinhooked dozens of stakes winners over the past 10 years, including 2010 Eclipse Award winner and Breeders' Cup Filly and Mare Sprint Winner Informed Decision.