

## IBM AND THE FUTURE OF BASEBALL

[Tom Davenport](#), Babson College Professor and author of *Competing on Analytics: The New Science of Winning*

[Bill Pulleyblank](#), vice president, IBM Center for Business Optimization

Hosted by Tim Washer

WASHER: One of the world's greatest philosophers once said, "The future ain't what it used to be." I like to think the former New York Yankee's coach was talking about sabermetrics, but odds are Yogi Berra was being much more profound. On the Future of Baseball today we'll take a look at sabermetrics, understand what it is, and find out how number crunching can help improve a team's performance.

DAVENPORT: It's generally defined as the scientific study of baseball. The word SABR comes from the Society for American Baseball Research, and those are the real propeller heads who look at the data and who analyze all the statistics about baseball.

And it's been going on, this society, for 30 years or so now, so it's not a new thing. But they've been the source of a lot of major innovations in how baseball is measured and analyzed.

PULLEYBLANK: That's the idea. It's how you can try and deal with all the uncertainty around the game of baseball and find a way to put the team on the field that's going to win a few more games.

DAVENPORT: The classic example is, of course, the Oakland As, who were a minor market team with very low budget. And they were really the first -- in part out of desperation -- to adopt the idea of using data on player performance to really decide who to draft and to go after bargains in a sense.

So they identified people, for example, who don't necessarily hit a lot but who get on base a lot who have a high on base percentage as kind of an undervalued resource in the baseball player market. And they got a lot of those people, and they did very well. And Oakland has performed historically very well particularly for the low budgets they've had.

WASHER: So part of that breakthrough is understanding a variable, right, or almost discovering a variable that was before, people just didn't pay attention to. Is that right?

DAVENPORT: Yes, it's data-based innovation. So runs batted in was a historical statistic that almost everybody used. And the statistical innovator said it's not really a very good indicator of how effective you are, because it has a lot to do with where you are in the batting order and so on. And so they came up with new metrics. There are a whole series of them: value over replacement player; on base plus slugging percentage. There are probably 100 metrics that the hardcore stat geeks really use in looking at baseball now.

WASHER: Okay. And so why is it that baseball specifically seems to be such a

good candidate for applying this type of science?

PULLEYBLANK: I'd say there are really two things. The first is that with any one of these sort of statistical things you want to do, the general problem is you just have a very small number of events. And statistical things don't really kick in until you get a lot of samples.

Baseball's great: 162 games in a season. You play basically every day for six months. That means there's a lot of data that comes in, a lot of events happen.

The second thing is the nature of the game is that the difference between a popup which is caught by the short stop and a home run over the left field fence may be literally a quarter of an inch where the difference in where the bat hits the ball.

So it's got a great degree of randomness in it than a lot of events, and that means understanding these trends and these probabilities really have a chance to do something here which they may not in other sports.

DAVENPORT: It's interesting what people do with analytics in baseball, but probably more important, what does it mean for business?

And a lot of the same ideas apply. The key prerequisite is data in both cases. If you've got the data then you can really start to analyze it -- and baseball has a ton of it since it's an individually oriented sport that's easy to measure.

The real strong importance of leadership. The idea that someone like Sandy Alderson at the Oakland As and, later, Billy Beane, could create an approach to analyzing baseball that nobody had ever done before. If they hadn't led it, it never would have happened.

WASHER: Tom, three to five years out, what are we going to see new to the sport?

DAVENPORT: Well, I think the statistical innovation will certainly continue. And the real limiting factor historically has been that the teams' adoption of it. So you still have a lot of teams that are lukewarm about it, a lot of journalists opposed to it. So I think we'll still see continuing penetration of the idea into the teams.

Now, it's kind of a much more participative phenomenon. You have amateurs who are creating a lot of these innovations. So bloggers and online people of various stripes are coming up with new metrics all the time. A lot of them are members of this Society for American Baseball Research. So that will I think continue to be a strong factor.

We'll see more use of video analysis to analyze particular players, to determine whether a fielder made a great catch or just a good catch. And the statistics historically haven't told us that. And I think we'll see the use of video for instant replay in games, which will be a huge step forward. No more disputed home runs.

PULLEYBLANK: Well, I think one thing that we're going to see is -- we're already seeing this -- is more and more information is being captured about sports. If you look at the baseball example, you'll see it on the TV coverage every so often, they'll show the manager or the coaches sitting there with the chart with all the hits. And they'll say, well, most of the hits are over here.

And they'll shift the outfield for this particular batter. But really don't you want to know with that pitcher and this batter, here is where they tend to go?

WASHER: Right.

PULLEYBLANK: Now, of course, the difficulty is that that pitcher and batter may only face each other a small number of times. So wouldn't you really like to know, who are all the pitchers who have a very similar characteristic.

WASHER: Sure and use them as a proxy for this guy.

PULLEYBLANK: And I take all of those ones and say, okay, this pitcher can...or, this batter consistently does this sort of thing with this type of pitcher if I put in that type of pitcher the odds are better that I'm going to know what to do with it. I can't guarantee I'll do the right thing, but at least you improve your odds. And that's what I think you'll start to see more. This requires much more careful analysis, much more detailed analysis, because you're no longer just saying, what happens if someone's pitching to somebody else? Who are the ones who pitch like him? Who are the ones that bat like him? What do we see that works better, putting those together?

DAVENPORT: Well, it's become a much more participative form of innovation and in fact you have now more innovative managers of fantasy baseball than of real baseball in a sense. And we didn't have that before, historically, for baseball. But now the rise of [roisserie] or fantasy baseball means that almost anybody can simulate a game at home and even see movement of some of these innovators into the baseball leagues. The Red Sox hired a guy who did a lot of simulations for fantasy baseball and they made him a full-time statistical analyst for them. So and a lot of cross-over into the real world of professional baseball now.

PULLEYBLANK: There are some very subtle things that happen in sports which can be detected by these sort of systems, and if they are detected then you can do things to try and tilt the things in your favor. Just like in a casino. Everybody knows about the idea of the card counter, and the effect that's had on the game of blackjack. Well, the point of it is, is your blackjack player cannot legally fiddle the deck. It's going to come out in the order it is. But if they know how many of each value of card have passed by, they can do a better job of dealing with that uncertainty. I think, really what we have to do -- and we see this now in business -- is that the companies that are really being successful are ones that acknowledge uncertainty and risk that they have to deal with. They take the actions which minimize the chances of a bad outcome and minimize the impact of a bad outcome on them. You talk about hedging strategies; hedging is about money you spend to reduce the impact of a bad outcome.

WASHER: Thanks for joining us on The Future of Baseball. We'll see you here next time. I'm Tim Washer.

[ END OF SEGMENT]

Continue the discussion on Tom Davenport's Harvard Business blog: ["How Analytics Help Build this Champion"](#)