Towards the beginning of this lecture, we stated that the goal of a baseball team is to make the playoffs and we built predictive models to achieve this goal.

But why isn't the goal of a baseball team to win the playoffs or win the World Series?

Billy Beane and Paul Depodesta see their job as making sure the team makes it to the playoffs, and after that, all bets are off.

The A's made it to the playoffs four years in a row-- 2000, 2001, 2002, and 2003-- but they didn't win the World Series.

Why not?

In Moneyball, they say that "over a long season luck evens out, and skill shines through.

But in a series of three out of five, or even four out of seven, anything can happen." In other words, the playoffs suffer from the sample size problem.

There are not enough games to make any statistical claims.

Let's see if we can verify this using our data set.

The number of teams in the playoffs has changed over the years.

So let's only use the years with eight teams in the playoffs, which was the number of teams in the playoffs in 2002, the year Moneyball discusses.

We can compute the correlation between whether or not the team wins the World Series-- a binary variable-- and the number of regular season wins, since we would expect teams with more wins to be more likely to win the World Series.

This correlation is 0.03, which is very low.

So it turns out that winning regular season games gets you to the playoffs, but in the playoffs, there too few games for luck to even out.

Next week, we'll discuss logistic regression, which we'll be able to use to predict whether or not the team will win the World Series.