GLOSSARY

At Bats: The definition of at bat has varied over the years, primarily due to the changing definitions of events that constitute plate appearances that are not at bats, such as sacrifice hits and sacrifice flies.

Average Value: Virtually every analytical statistic in this encyclopedia uses league average as a baseline. The use of this baseline, which is set to equal either 0 or 100, depending on the statistic, makes the quality of a player's performance—whether a player is above average, average, or below average—immediately obvious most of the time. Above-average players are needed to win pennants. Average and below-average players who are better than a typical replacement make only minor contributions toward a winning season.

Average is not the only reasonable baseline, of course—there is no "right" baseline, though there are more useful and less useful baselines. Some analysts use "replacement level" as their baseline. That baseline is hard to define, but there are circumstances where it may be more useful than the average baseline. Different tasks require different tools and, thus, different baselines. The primary advantage of an average baseline is that it clearly shows how the performance of players translates into wins and losses on the field. All teams start the season 0–0; only better-than-average players enable a team to rise above .500.

Bases on Balls: Permanently established at the count of 4 balls in 1889. The number had come down one at a time since 1874, when 9 balls were required to reach first base.

Basestealing Runs: The formula for calculating BSR is (.22*SB) – (-.35*CS). SB stands for stolen bases, CS is for caught stealing, and * signifies multiplication.

Basestealing Wins: Divide basestealing runs by runs per win, which in this case equals 10 times the square root of runs per inning.

Batter-Fielder Wins: The sum of a player's batting wins, basestealing wins, and fielding wins.

Batting Runs: The formula for calculating batting runs is .33*(BB+HBP) plus .47*H plus .38*2B plus .55*3B plus .93*HR–ABF*(AB–H).

ABF, known as the league batting factor, makes the value of an average batter equal 0. ABF is computed with the following formula (all statistics in the equation are league statistics):

$$\frac{.33*(BB+HBP) + .47*H + .38*2B + .55*3B + .93*HR)}{(AB - LGF*H).}$$

LGF, the league factor, adjusts for the quality of league play, and equals 1 except for the Union Association (1884), for which it equals 0.8, and the Federal League (1914–15), for which it equals 0.9.

Adjusted batting runs, which appear in the batting register, are calculated with the following formula: (BR – (batters' park factor – 1)*RPA*PA/batters' park factor). RPA is the number of runs per plate appearance in the league.

Batting Wins: This is calculated by dividing adjusted batting runs by runs per win. In this case, runs per win equals 10 times the square root of (runs per inning + player adjusted batting runs/games/9).

Blown Saves: Save opportunities not converted because the relief pitcher allowed the tying or go-ahead run to score. Blown saves calculated using current definition from 1969–present. From 1957–68, blown saves include blowing a lead of any size because relief pitchers could earn a save any time they finished a game after entering with a lead, no matter how large.

Caught Stealing: The data is available for 1912, 1914, 1915, and from 1920 onward for the AL; and for 1913, 1915, 1920–1925, and from 1951 onward in the NL. Caught Stealing totals are also available for 1916 for all players who stole at least 20 bases.

Differential: This measures the difference between the won-lost record a team was *projected* to achieve (based on its batting, pitching, fielding, and basestealing), and the record that the team *actually* attained. If a team was projected to finish 89–73 (that is, +8) but actually concluded the season at 85–77 (that is, +4), the team's differential would be –4.

Earned Run Average: The number of earned runs allowed per 9 innings pitched. It is calculated by multiplying earned runs by 9 and dividing by innings pitched. Adjusted ERA is calculated by dividing league ERA by the pitcher (or team) ERA and then multiplying by the pitchers' park factor. Note that a small number of runs can be earned runs for the pitcher but unearned runs for the team.

ERA has been kept officially since 1912 in the NL and 1913 in the AL. Before then, Information Concepts Inc. researchers calculated earned runs from game accounts. Some earned run data include estimates, although in every case, at least half the earned runs for a team were known from game accounts. This was for 1881–86 (all leagues) and 1887 and 1890 in the American Association. Earned run data in box scores and the official guides in the nineteenth century were not used. Bases on balls were usually considered errors in those years and, in fact, the fielding stats often gave pitchers assists on strikeouts and errors on walks. Such assists and errors have been removed from the fielding averages in this book. Note that if a pitcher allows one or more earned runs, but fails to retire a batter during a season, his ERA is infinite. If so, we display the number of earned runs allowed in parentheses instead of his ERA.

Fielding Runs: Fielding runs are based on the player's fielding statistics at each position compared to the league average for the number of innings played. The number of innings played was obtained from play-by-play data for 1957 onward. For most years before 1957, innings are estimated from defensive and offensive data for all players on the team. Comparing the formulas for estimated defensive innings to actual defensive innings taken from 1957–2006 play-by-play data showed an average difference of approximately 1% for infielders and 2% for outfielders for players with 1,000 or more innings.

The basic formula is:

PFR/(PO–SO for team)–LFR/(PO–SO for league)*Player Innings.

PFR is the player fielding rate, while LFR is the league fielding rate.

The ratings used as fielding rates vary by position as follows:

1B	.2*(2*A – E)
P, 2B, SS, 3B	.2*(PO + 2*A – E + DP)
OF	.2*(PO + 4*A – E + DP)
C	.2*(PO – SO + .4*(A – CS) – E + DP + PB/2).

A stands for assists, PO for putouts, E for errors, SO for strikeouts, DP is for double plays, and PB for passed balls. Strikeouts are subtracted because a team that strikes out a lot of batters will have many fewer chances to create outs from balls put into play. When calculating a pitchers' fielding, his actual strikeout total is used; for other fielders, the team average strikeout total is used.

Two more factors are used to adjust for the amount of left-handed and right-handed pitching on a team and for the number of double play opportunities. Teams with more left-handed pitching tend to face more right-handed batters as a result of platooning, and this effect in turn alters the distribution of balls hit into play. A careful historical study of this factor showed that its effect was non-existent before 1910, then slowly grew until 1970, when it hit a similar level to today.

Right-handed batters shift the distribution of ground balls over from the right side of the infield to the left side while shifting flyballs and popups from the left side of the field to the right side. Second basemen also gain more putouts with a right-handed hitter at the plate since they are more likely to be covering second base.

To adjust the number of expected league average putouts and assists for each player, the relevant statistic is multiplied by this equation:

$$(1 + ADJ*YF*DLHP).$$

ADJ is the adjustment figure appropriate for the position (1B, 2B, 3B, SS, LF, or RF) and statistic (putouts or assists). The chart below supplies the appropriate figure.

ADJ	*PO*	*A*
1B	n/a	–.40
2B	.23	–.27
SS	– .10	.14
3B	– .22	.34
LF	– .16	n/a
RF	.09	n/a

YF is the year factor, necessary because this factor steadily increased in importance from 1910 to 1970. Before 1910 the YF is 0, so no adjustment is necessary. The adjustment can be calculated for each year from 1910 through 1970 by subtracting 1910 from the year in question and then dividing by 60. After 1970 the YF is always 1.

DLHP is the difference in the percentage of left-handed pitching from the league average.

So if a second baseman was on a team with 10 percent more innings by left-handed pitchers than average in 1960, his adjustment is:

Assists = (1 – .27 x (1960-1910)/60 x .10) = .9775
Putouts = (1 + .23 x (1960-1910)/60 x .10) = 1.0192.

If the league average second baseman in 1960 finished the season with 400 assists and 200 putouts, the expected league average would be adjusted to 391 and 204.

Pitchers have a further correction, since left-handed pitchers have fewer putouts as a result of facing fewer left-handed batters. Up until 1910 left- and right-handed pitcher putouts were about equal, declining to about 62 percent for left-handers in 1970 before leveling off. Thus, a separate expected putout rate is calculated for each type of pitcher based on the year and the number of left-handed innings.

Double play opportunities were estimated from hits, bases on balls, hit-by-pitches, and home runs allowed, plus errors committed. (On average, 57.5 percent of errors result in a runner reaching base; the other errors just allow existing baserunners to take extra bases.) Using a multiplier for homers to account for double play opportunities lost did not improve the estimate in years where there are actual data; the average error in these years was around 2 percent. The formula for calculating double play opportunities is:

.662*(H – HR + BB + HB + .575*E).

Individual double plays were divided by the team double play opportunities divided by the league average double play opportunities.

For catchers, there is an additional defensive calculation made to rate them on other defensive aspects. The formula for that calculation is:

(–.22*SB) + (CSF*CS) + .1*APR*IP/TIP.

APR is the catcher's team's adjusted pitching runs; IP is the catcher's innings pitched to; TIP is team innings pitched; and CSF is the caught stealing factor that ensures that the average value of a catcher defensively stays at 0. The formula to calculate CSF is:

22*(LSB – TSB) divided by (LCS – TCS).

Also, since artificial turf results in about a 5 percent higher stolen base success rate, a small adjustment is made to the results of players and teams from the turf era. The size of the adjustment depends on the exact split of games between ballparks with turf in the basepaths and those with dirt basepaths. Most ballparks with turf fields have turf basepaths with dirt only in the sliding cutouts around the bases. However, a few ballparks, especially in the early years of artificial turf, have had all-dirt basepaths with turf fields.

Fielding Wins: Divide fielding runs by runs per win, which in this case equals 10 times the square root of runs per inning.

Games Behind: This is the number of games one team is behind another team in the standing, almost always measured from first place. If a team's record stands at 78–71 and the first-place team's record is 82–68, then the former team is 3½ games behind the first-place team.

Grounded into Double Play: This became an official statistic in 1933 in the NL and in 1939 in the AL.

Hit-by-Pitch: The rule awarding first base to batters hit by a pitch was instituted in 1884 by the American Association and in 1887 by the National League.

Innings Pitched: Fractions of innings pitched in a season have only been officially counted since 1982. Prior to that, both leagues rounded off fractional innings. This encyclopedia has full innings pitched (i.e., not rounded off) for all major league seasons.

Normalizing: Baseball statistics tell us very little without context. A .450 slugging average in Dodger Stadium in 1965 means something very different than a .450 slugging average in Coors Field in 2003—the former is an impressive performance while the latter is sub-par. Many of the statistics in this encyclopedia are normalized using league average as a baseline, enabling readers to see whether a player is better than or worse than his average peer as well as to what extent he is better or worse. By putting different seasons played in different circumstances on the same scale, normalization makes it much easier to compare seasons. Some statistics in this encyclopedia are further normalized to account for the home parks of players since some ballparks are far more conducive to scoring than others. As a result, both the ease of producing a run and the value of a run can vary significantly from ballpark to ballpark, even in the same season. Normalizing a statistic for ballpark effects ensures that players can be compared on a level playing field.

On-Base Percentage: Declared an official statistic in 1984, OBP is defined as (hits plus bases on balls plus hit-by-pitches) divided by (at bats plus bases on balls plus hit-by-pitches plus sacrifice flies). This encyclopedia uses the current definition back to 1954, when sacrifice fly data became permanently available. Previous to 1954, sacrifice flies were either counted as outs or were included with sacrifice hits (1908–30 and 1939) and, therefore, cannot be used.

On-Base plus Slugging: Often referred to as OPS, this statistic, which was introduced by Pete Palmer and John Thorn in *The Hidden Game of Baseball* in 1984, has exploded in popularity the past few years. This encyclopedia does not feature a separate column for the basic version of OPS in the batter register since it is very easy to calculate—simply adds on-base percentage and slugging average—and the two columns are adjacent. Adjusted OPS (which has its own column) then normalizes OPS for the league and the player's home park(s), then converts it to a scale in which 100 is league average. The exact calculation is ((player OBP/league OBP) plus (player slugging/league slugging) – 1) divided by batters' park factor. When calculating adjusted OPS for non-pitchers, league average statistics do not include pitcher batting. For consistency in historical comparisons, the definition of on-base percentage used in adjusted OPS does not use sacrifice flies in the denominator.

Opponents' Batting Average: Figures are based on estimated at bats from 1901–07 in the AL; from 1889–1902 in the NL; in 1882–83 and 1888–91 in the American Association; in 1884 in the Union Association; and in 1890 in the Player's League.

Opponents' Caught Stealing: Stats for 1957 to the present came from play-by-play records. Estimated totals for catchers came from team totals found in league records from 1920–56 for the AL and from 1920–25 and from 1951–56 for the NL. For most other years since 1890, team totals were estimated from other team data and catcher totals were estimated from the team estimates. Some years from 1912–19 have only runner caught stealing data, some have team caught stealing, and some have individual catcher caught stealing.

Opponents' On-Base Percentage: Based on the official definition of OBP since 1954. The definition used in previous years depends on the available data (as with the OBP definition above).

Opponents' Stolen Bases: Exact totals for 1957 to the present came from play-by-play records. Estimated totals for catchers came from team totals found in league records for 1890–1956 (except for the 1890 NL data, which came from box scores).

Park Factor: This measure of how the team's home park affects hitters and pitchers is used to adjust the team's performance in a way that takes into account the context of the team's home park. Separate park factors are used for batting and pitching in order to adjust for the fact that pitchers and hitters never get to face their own teammates.

Pitching Runs: This measure of how many runs a pitcher prevented compared to the average pitcher is calculated with this formula:

LERA* IP/9 – ER + URF.

LERA is the league ERA and ER stands for earned runs. URF is the unearned run factor that accounts for the unearned runs the pitcher is responsible for and is calculated by multiplying (0.5) times (ER – R*TER/TR).

Adjusted pitching runs is calculated the same as pitching runs above, except that LERA* PPF is used instead of plain LERA—that is, the league ERA is multiplied by the pitcher park factor. Then the result of the entire calculation is divided by the pitcher park factor. Thus the formula is (PPF*LERA*IP/9 – ER + URF)/PPF.

Pitcher Batting: This measure of a pitcher's offensive performance is calculated the same way as adjusted batting runs; the difference is that a pitcher is compared to the average-hitting pitcher, not the average hitter. If a pitcher spent time at other defensive positions, his offense will be divided proportionately based on how much he played each position; only the appropriate fraction will count for pitcher batting.

Pitcher Wins: The total number of wins a pitcher is worth to his team compared to the average pitcher (including pitching, fielding, batting, and basestealing). It is calculated by multiplying adjusted pitcher runs by XMULR (see below), adding the pitcher's non-pitching contributions, and then dividing by runs per win.

In order to properly credit relief pitchers for the extra or lesser value of their innings pitched, a factor is used when converting adjusted pitching runs to pitching wins. The formula for that multiplier (XMULR) is:

$$9*(W + L + SV/XSV)/IP.$$

XSV is calculated by dividing league saves by league wins and multiplying by 10. XSV cannot be less than 4, so any result lower than 4 is set at 4. The multiplier cannot not be less than 0.5 or more than 2, so any result not in that range is considered to be either 0.5 or 2, depending on whether it was above or below the range.

Pitching Wins: The wins a pitcher achieves by his pitching is calculated by dividing adjusted pitching runs by runs per win. In this case, the formula for runs per win is 10 times the square root of (RPI – APR/G/9). Thus, if a pitcher reduces the number of runs scored by the opponent in his games, the value of each run is increased by a reduction in the runs per win figure. Pitchers, especially starting pitchers, have a much stronger effect on runs per win because their effect on scoring is spread out over fewer games.

Positional Adjustment: Baseball fans know not to expect a shortstop to hit like a first baseman, exceptions like Alex Rodriguez notwithstanding. Second basemen who slug .370 can usually hold on to their starting spot, but first basemen who slug .370 are likely to be benched or dropped off the roster. That's because it's fairly easy to find an acceptable defensive first baseman that can hit with power, while players who can handle second base defensively and hit with power are usually scarce. In order to account for the ability of players to handle the most valuable positions defensively, the offense of these players is compared to their peers at the same position. This approach accounts for the differing value major league teams place on fielding ability at each position and also makes it simple to adjust for the changing demands of a position. For example, higher defensive expectations for third basemen a century ago made it significantly harder to find a third baseman with a good bat back then than today. Whether that was because teams demanded a higher level of defense from third basemen then, or because there are more good defensive third basemen these days who can hit is irrelevant. We don't have to know the answer to that question to account for the decreasing scarcity of good-hitting third basemen by using a smaller positional adjustment.

POWR (Player Overall Win Rating): POWR, featured in the All-Time Leaders section, adds the batting wins, fielding wins, basestealing wins, and pitching wins of every player to rate his overall value compared to an average player.

Range: This category is calculated different ways for different positions. For infielders, range is based on assists per inning. For outfielders, range is based on putouts per inning. For catchers, the statistic in the range column is based on stolen bases allowed per inning. Outfielders are rated for their (weighted) play at all outfield positions, while infielders are only rated for their play at their primary position. The data is then adjusted in comparison to the league average, with 100 equaling league average. *Higher is always better.* All statistics are adjusted for context, including the number of balls put into play and the distribution of balls in play. The adjustments for each statistic are detailed in the fielding runs entry. Innings played data are calculated from play-by-play accounts from 1957 onward; it is estimated prior to 1957.

Relative Batting Average: A normalized translation of batting average. The formula is player average divided by league average.

Runs Batted In: RBIs have been recognized as an official statistic since 1920. Runs driven in by a force double play were no longer counted as RBIs after 1938.

Relief Ranking: Generated by multiplying the XMULR factor (described in the pitcher wins entry) by adjusted pitcher runs. Since innings pitched in relief is not historically available, relief pitchers are classified as those pitchers who average less than 3 innings per appearance.

Run Support: Basic run support is calculated by adding up all the runs scored in a pitcher's starts and then dividing that total by his games started. The support figures presented in the pitcher register have been normalized for the context of the offensive level of the league and the player's home park(s), then converted to a scale in which 100 is league average. A run support figure of 90 would indicate that pitcher had 10 percent less runs scored in his starts than average; a figure of 110 would indicated 10 percent above-average. The formula to produce the adjusted run support is (pitcher run support) divided by (league average run support) divided by (batters' park factor) times 100.

Run support stats have been published widely since 1990, usually counting all appearances for a pitcher and normalizing the runs scored by his team while he was in the game to 9 innings pitched. While this stat is useful, it is not calculable pre-1969. Furthermore, it is not a substitute for calculating run support for starting pitchers, since runs scored after a pitcher leaves the game materially affect a pitcher's record. If a starter leaves the game in the seventh inning with his team trailing 5–1, then his team scores 6 times in the eighth to take the lead, he avoids being tagged with the loss and clearly benefits from runs scored after he had left the game.

Runs Per Win: The number of runs needed, on average, to gain an additional win in the standings. Historically, about 10 runs have equaled a win. The value of runs per win is used in several entries in this glossary to translate runs into wins. Individual players, especially starting pitchers, can have a significant effect on the number of runs that it takes to achieve a win in their own games, so the runs per win calculation often takes into account the performance level of the individual or team being evaluated.

Sacrifice Fly: First recognized as a distinct statistical category in 1954. Previous to 1975, only fair fly balls which drove in runners were counted in this category. From 1908–1930 and again in 1939, sacrifice flies were counted as sacrifice hits. In all other seasons before 1954, sacrifice flies are indistinguishable from other at bats.

Save: Recognized as an official statistic in 1969, saves were at first awarded to any relief pitcher that finished pitching a victory and was not credited with the win. In 1973 requirements were significantly tightened so that saves were only awarded if the pitcher either entered the game with the potential tying run on base or at the plate or pitched at least 3 effective innings while preserving the lead. In 1975 the rule was set to where it stands today: In order to earn a save, a reliever must finish off a victory without ever giving up the lead after entering the game with a lead of no more than 3 runs and pitching at least an inning, or he can earn a save by finishing a victory without giving up the lead after entering the game with the tying run on base, at the plate, or on deck. Saves may also be earned by pitching effectively for at least the last 3 innings of a victory, without getting credit for the win. The original 1969 rule has been used in awarding saves prior to 1969.

Save Opportunities and Save Percentage: Save Opportunities are defined as Saves plus Blown Saves—a narrow definition that makes Save Percentage useful only when comparing closers to each other. Save Percentage is calculated by dividing Saves by Save Opportunities. Blown Saves are defined by Rolaids, which has given the Rolaids Relief Man Award annually since 1976. By virtue of the prestige that this award has accrued, Rolaids effectively controls the definition of the Blown Save since it is not an official statistic.

Though setup pitchers and middle relievers get charged with Blown Saves when they blow a lead in a save situation, they don't get many saves because they are almost always replaced by their team's closer before the game ends. The effect, therefore, is that setup and middle relievers rarely show a high Save Percentage: they get nicked for the negative stat when they fail to protect a lead, but they rarely get a chance to earn the positive stat when they protect a lead.

The unofficial "Hold" statistic that gained currency in the 1990s was an attempt to address that gap. Unfortunately, neither of the two versions of the "Hold" stat solve the problem. The earlier version doesn't even require a relief pitcher to retire a hitter to get credit for a "Hold"—if a reliever comes into the game with a 1-run lead and no one on base, then walks the bases loaded before being yanked, he can get a "Hold" despite his miserable performance. The later version of the "Hold," promulgated by Stats, Inc., and still used today, requires a relief pitcher to pitch at least $^1/_3$ of an inning to gain a "Hold." Better. That's better, but still not nearly good enough.