**Analysis of the 2022 MLB Season**

**An Overall Analysis of the 2022 Major League Baseball Season using R Graphs**

# Introduction

The 2022 MLB regular season ended on October 5th, with the postseason seeing a new format that brought 12 teams into the playoffs, up from 10 last season and encompassing over a third of the league. Over the season, thousands of baseballs were hit, hundreds of runs scored by each team, and dozens of innings pitched, all with the goal of making it into October baseball. All these teams started the season with hopes and dreams. Some to lift the trophy at the end of it all, some to see improvement in young rookies and players, and others (secretly) hoping to get a nice draft pick to add to their farm system.

In this paper, I have done a general analysis and overview of the 2022 season for all teams, batters who qualified for the batting title, and starters/relievers who pitched at least a somewhat meaningful number of innings. We will look at expected outcomes, unusual trends, and the general strangeness that could be found in the numbers of America’s pastime.

# Team Statistics

**Batting**

Let’s start with a summary of how each team did during the regular season. When it comes to team sports, baseball is arguably the one where an individual has the least influence on the outcome of a game and season. Unlike a star player in basketball, who is on the floor for 90-100% of the game, or a franchise quarterback being able to have direct control over half the game, a baseball player is limited to batting 1 out of 9 times, and a pitcher to around 1 of every 5 or so games. This makes overall team statistics more important and relevant than perhaps the other major American sports. A player can’t singlehandedly carry a team to glory, as LeBron James nearly did in his first stint with the Cavaliers. A perfect example is available in baseball: the Angels (my team, so special focus will be put on them throughout) have 2 of the best players currently in baseball, Mike Trout and Shohei Ohtani, both of whom are on track to the Hall of Fame (maybe not Ohtani yet, but certainly Trout). The Angels missed the playoffs this year and haven’t been in 8 years (when they last won the division in 2014), which is also the last time they’ve had a winning record.

Point is, in baseball everyone must pull their weight to some degree. Star players are certainly a huge help, but if the other guys in the lineups can’t hit the star players home after they get on base, there’s little chance the team will succeed. To that end, teams with overall higher batting statistics tend to be ones that win games, and sometimes a lot. Let’s first start with the record of all the teams in baseball, compared to .500 (or an 81-81 record):

(Insert WL above .500 graph here)

The first somewhat interesting observation is that the 12 playoff teams were the 12 teams with the best records in the season. While this sounds obvious and normal, given the divisional structure of baseball, it’s not too uncommon to see a team with a worse record make the playoffs over a better one. A single division can have multiple 90+ win teams, with some of them missing the playoffs, while a division leader with 88 wins can waltz right in. The addition of 2 more teams to the playoffs would likely allow for the teams that “should” be in to be in. This year however, even with the previous 10 team format, the best teams would still be in, as the Mariners, Padres, Phillies, and Rays would be the 4 left out. However, the Phillies did make the World Series, so perhaps this 12-team format does allow for a wider variety of competitors.

Run Differential is a measurement of how many runs a team scored versus how many they allowed. It’s a basic way to see how “good” a team is: if they score more runs than they allow, they’re probably a decent team. The larger the margin positive or negative can show how good/bad a team is.

(Insert Run Diff Graph here)

As expected, the more winning teams scored far more runs than they allowed. Even in this regard, the Dodgers are in a league of their own, having scored 334 more runs than they allowed, a difference of almost 100 compared to the Yankees. We also see a group of seven teams in their own league of bad run differential. Washington earns the worst differential here, as they also did in the Win-Lose record, a spectacular collapse from the team that won the World Series just 3 years ago in 2019. Again, the 12 best teams in this metric were all 12 teams that made the playoffs. No weird “teams that shouldn’t be here" shenanigans here.

Hits are perhaps the simplest offensive statistic for a team. It shows how many times the batters succeeded in playing the game of baseball: putting a ball in play and getting on base. Considering succeeding at this act a third of the time is considered elite, it shows how important this stat is. Getting hits mean guys getting on base, meaning guys being able to score runs.

(Insert Hits graph here)

Above, we see that the Blue Jays reign king in the hits department with a total of 1464.

# Notes

* Team Summaries
  + All 12 playoff teams had the highest WL% in order (No weird teams with more wins missing)
  + Applies to run differential, too. Somewhat large gap between bottom 7 teams and everyone else. Dodgers in their own league
  + Toronto leads with the most hits, Oakland the least, Yankees on the lower half (despite Judge) and lead in HRs, and the Royals had the most triples
  + Dodgers again scored the most RBIs by a wide margin. Detroit at the bottom (and Angels on lower third)
  + Rangers stole most bases, Minnesota the least. The Dodgers stole a good amount of bases, but were far more successful in not being caught
  + Angels led in Strikeouts (likely a big reason for their sub .500 record). Cleveland was the most disciplined team
  + Padres and Dodgers left lots on base. Meaning they got on often, just couldn’t capitalize. Oakland had the least, likely because they couldn’t get on base period
  + We see that 3 teams with higher-than-average OPS+ (Minnesota, Milwaukee, and Boston) missed the playoffs. Toronto leads by a high margin
  + LA and Houston stand apart in team ERA. Angels are the lowest ERA that missed the playoffs, lower than 4 teams who made the playoffs, showing a surprising pitching staff that seemed to be let down by poor offense
  + Teams with lower ERA have higher WL%. Angels stand out for being one of 3 teams (Miami and Minnesota) with a sub 4.0 ERA and a sub .500 record, and manage to have the lowest ERA of these teams
* Batter Summaries
  + Aaron Judge leads in OPS (no surprise), with the only player finishing above 1.000. Jonathan Schoop and Myles Straw have by far the worst, both being separated from the rest of the league by more than .4 points
  + 3 outliers in slugging vs IP: Judge with the highest at .686, highest by .7 points; Myles Straw with the lowest at .273, lowest by .4; and Yordan Alvarez who managed the second best SLG (.611) at almost 100 less ABs from Judge (470)
  + Pete Alonso and Aaron Judge tied for most RBIs (131), with Jose Ramirez at second (126). Myles Straw is again in last with 32
  + Randy Arozarena, Jorge Mateo, and Cedric Mullens lead the league in total attempts at 44, with Mateo having the most successful steals at 35. However, Tommy Edman as the most efficient stealer, stealing 32 on 35 attempts, for a 91.4% success rate
* Pitcher Summaries
  + Verlander is king of starters, with a 1.75 ERA and the AL Cy Young award. The highest is Patrick Corbin, with a 6.31 ERA
  + Gerrit Cole has the most total strikeouts at 257. The highest SO/9 belongs to Spencer Strider, with 13.8(!) SOs per 9 innings, 1.8 higher that Carlos Rodon in second with 12
  + 3 Dodger’s Starters account for the top 10 in Starter ERA
  + Sandy Alcantara is the outlier, pitching 3 more full innings worth of games than the next highest starter