Rather than using a teams basic stat line from the boxscores such as rebounds, assists, points, 3-points attempted and made, etc. we will look at Hollinger's NBA stats or the advanced statistics. The stats from boxscores only show the result of every play and not how the play developed. Hollinger's stat provide an incite on every play and thus allowing us to understand how effiecent affective a team truly is. These variables rate teams on their productivity, which provides an overall picture, far beyond their wins and losses column.

Explanation of each Hollinger's stat variables

1. **RK (Rank):** ranking of the team at any given point in a season based on wins and losses.
   Drop. Drop for non because rank of team is not needed for our model. Might want to come back to this if we want to change our model to predict end of the season ranking.

---

2. **Age:** Average age of team

   Drop. This variable is not neccasary for what we are trying to predict

---

3. **W (Wins):** Wins column

   Keep. We need every teams record through the EDA process

---

4. **L (Losses):** loss column

   Keep. We need every teams record through the EDA process

---

5. **PW (Pythagorean Wins %)**: A method that gives an expected winning percentage using the ratio of a team's wins and losses are related to the number of points scored and allowed. This formula was designed to relate a teams points scored and points allowed to its win-lost record.

   Formula: $(Points\ scored)^{16.5}/[(Points\ scored)^{16.5} + (points\ allowed)^{16.5}]$

   Keep: Keep for now. We should only consider dropping this variable because it's giving you an expected value. Using a variable with expected values to predicted playoff birth might not be the best option but we can keep it during the explanatory data analysis and see any correlatios or useful information it gives us.

---

6. **PL (Pythagorean losses %)**

Drop: drop this because we are keeping PW %. Having this will be redundant.

---

7. **MOV (Margin of victory):** This column is a teams average margin of victory over the entirety of the season.

   Drop: my initial thought is to drop this column but it might have some useful information about a team.

---

8. **SOS (strength of schedule):** represents a team's average schedule difficulty faced by each team in the games that it's played so far or for all season. The schedule difficulty of a given game takes into account the rating of the opponent and the location of the game.

   Strength of schedule in the NBA has not been considered nearly as important as in the NFL or as in the NCAA basketball since the NBA team play each other at least twice. Much of the schedule differences between teams comes down to road trips, back-to-backs, afternoon games for example

   Drop: This just doesn't seem like it makes a huge difference in how well a team does in the NBA in a given season, since every time plays each other multiple times.

---

9. **SRS (Simple Rating System):** a team rating that takes into account average point differential and strength of schedule. The rating is denominated in points above/below average, where zero is average.

   "It is not perfect, as garbage time minutes muddy the waters, but it has been fairly accurate at evaluating which teams are legitimately good in a given year".

   KEEP: The definition says it all. I think it will give us a good gage of a given teams ratings. (average or below average)

---

10. **ORtg** (**Offensive Rating):** is the number/estimation of points a team scores per 100 possessions.

    Formula: $100 * (points\ scored/possssions)$

    Keep: We are keeping anything that has to be with a rating of a team and offensive rating is important because the better offensive rating a team has the better they are on the offensive side which means they score more.

---

11. **DRtg** (**Defensive Rating):** An estimate of points allowed per 100 possessions

    Formula: Formula: $100 * (points\ allowed/possssions)$

Keep: Just like offensive rating, this is also important and we will analyze this variable carefully.

---

12. **NRtg (Net Rating):** an estimate of a team's point differential per 100 possessions.

    Formula: $ORtg - DRtg$

    Keep: Just like ORtg and DRtg this is an import tool to measure a team. Using possessions rather than minutes eliminates the effects of a team that plays very fast or very slow.

---

13. **Pace (Pace Factor):** An estimate of possessions per 48 minutes

    Keep: Lets keep it for the explaantory data analysis portion. This could be correlated to wins in a given game since (If a team has a low pace factor (fewer possessions), they have fewer opportunities to score. If a team has a high pace factor (more possessions), they are more likely to have a higher-scoring game).

---

14. **FTr** (**Free Throw Attempt Rate):** Number of free throw attempts per field goal attempt.

Free Throw Percentage is a simple equation: FTMade/FTAttempts.

	Game 1	Game 2
FTM	15	27
FTA	21	31
FT%	71.4%	87.1%

Free Throw Rate is the ratio of foul shots to field goal attempts, expressed as FTA/FGA.

	Game 1	Game 2
FTA	21	31
FGA	80	92
FTr	26.3%	33.7%

Let's discuss FT percentage versus FT volume:

If the Thunder averages 100 shots per game, based upon averages, they would shoot roughly 20 free throws per game. Shooting 70% from the line, this means they would score 14 points. By improving shooting at the line to 80%, the Thunder would score 16 points per game. An increase of only 10% in FT% provides 2 additional points per game.

If the Thunder increased its Free Throw Rate from 20% to 30%, that would result in 30 free throws per game, on average. Shooting the same, original 70% from the line would result in 21 points – a 7 point improvement, more than twice as much as increasing free throw accuracy by 10%.

Keep: This is definitely worth keeping because the more free throws you get the more points you have.

---

15. **3PAr** (**3-Point Attempt Rate):** Percentage of FG Attempts from 3-Point Range.

    Formula: $3PA/FGA$

Keep: In todays NBA, 3-point shooters are the most sought after players because of the rate at which teams shoot 3's. Steph Curry and the warriors started this change in the NBA. we need to keep and analyze this variable.

---

16. **TS%** (**True Shooting Percentage):** A measure of shooting efficiency that takes into account 2-point field goals, 3-point field goals, and free throws.

    Formula: $Points\ /\ (2\ *\ (FGA\ +\ 0.44\ *\ FTA))$

    Keep: True Shooting Percentage (TS%) could be the clearest and easily measured metric that boils all three methods of scoring into one number. It evaluates how well the team shoots, so keep this.

The next 4 variables are called Offense Four Factors,
We will keep all 4 because they're the most important measurement of the team offense.

(Shooting the ball)

17. **eFG%** (**Effective Field Goal Percentage):** Effective Field Goal Percentage is a measurement of how successful your team is from the field. This metric provides a more complete picture of the game situation than standard field goal percentages because three point shots are given extra weight

    Formula: $(2pt\ FGM\ +\ 1.5\ *\ 3pt\ FGM)/FGA$

(Taking care of the ball)

18. **TOV%** (**Turnover Percentage):** An estimate of turnovers committed per 100 plays. A teams possessions that end in a turnover.

    Formula: $100\ *\ TOV/(FGA\ +\ 0.44\ *\ FTA\ +\ TOV)$

(Offensive rebounds)

19. **ORB%** (**Offensive Rebound Percentage):** An estimate of the percentage of available offensive rebounds a player/team grabbed while they were on the floor. This increases the number of chances that the team has to score on 2nd or 3rd attempt in a play.

(Getting to the foul line)

20. **FT/FGA**: Free Throws Per Field Goal Attempt. Measures how often a team gets to the line and how often they make them.

The next 4 variables are called Defense Four Factors,
We will keep all 4 because they're the most important measurement of the team defense.

---

21. **eFG%** (**Opponents Effective Field Goal Percentage)**

---

22. **TOV%** (**Opponents Turnover Percentage)**

---

23. **ORB%** (**Defensive Rebound Percentage):** An estimate of the percentage of available defensive rebounds a player grabbed while they were on the floor.

---

24. **FT/FGA**: Opponents free Throws Per Field Goal Attempt.

Heres a short youtube clip explaining the theory behind advanced analytics.

https://www.youtube.com/watch?v=QTq4O2IWanw&ab_channel=DanielLi