

Kevin Durant vs Twitter

Nick Videtti

IST-652 Final Project

Summer 2022

**Overview**

Kevin Durant is a professional basketball player that is recognized as one of the most talented scorers in NBA history. Durant was drafted by the Seattle Supersonics, who would become the Oklahoma City Thunder after his 1st season, with the 2nd overall pick in the 2007 NBA Draft. He remained with the Thunder until 2016. In those 9 seasons, Durant led the NBA in scoring 4 times, and won the NBA MVP during one of those 4 seasons as the scoring leader in 2013-2014. While the first 9 years of his career were extremely successful, they also included only one NBA Finals appearance in 2012 which resulted in LeBron James’ Miami Heat winning the best-of-7 series 4-1 over Durant’s Thunder. Meanwhile, in 2016 the Golden State Warriors had just finished the best regular season in NBA history and were one win away from winning their second straight NBA championship, but ultimately lost to the Cleveland Cavaliers led by, you guessed it, LeBron James.

Durant and the Warriors both needed something from each other in 2016, and they would ultimately join forces in an attempt to get Kevin Durant his elusive NBA Championship and get the Warriors the player that would guarantee they wouldn’t need to come down to one game in the NBA Finals any time again for the near future. The only problem was that, not only had the Warriors come off the best regular season in history and had the MVP from the two previous seasons, Stephen Curry, but they were the team that eliminated Durant’s Thunder in order to advance to the 2016 NBA Finals. To summarize, Durant left the team he had been with, and failed to win a NBA Championship with, for 9 seasons to join the team that ended his career in Oklahoma City. Thunder fans felt betrayed and felt like they were given up on by a team legend. Perhaps it was a coincidence that this is how the situation played out, but many fans, analysts, and writers critiqued this move so much that it has completely changed Durant’s public image to this day.

Since then, Durant won two NBA championships with the Warriors and was Finals MVP for both of them, but the legitimacy of those championships are questioned by many to the extent where some believe Durant questions it himself. In 2019, Durant was very possibly on his way to a third consecutive championship with the Warriors before an achilles injury kept him out for a year, but he was on the Brooklyn Nets by that point. Shortly after that injury, in July 2019, Durant joined the Brooklyn Nets, and many speculate one reason for that was to try win a championship without being under the scrutiny he faced in Golden State, and without being in the shadow of Stephen Curry. Fast-forward to 2022 and Durant has not yet won another championship, albeit it could be argued he was very close in 2021, and as recently as this summer he has requested a trade from the Nets and gave them an ultimatum to choose between himself and the coach.

For everything that has occurred since 2016, Durant continues to face more and more criticism from fans and analysts. This is common among professional athletes, but seems to be different for Durant based on how he responds to the criticism. The approach that the vast majority of athletes take vary from ignoring the critics, limiting reasons for people to critique them, or embracing the criticism so they can learn to live with it. Kevin Durant has taken a different approach and stands up for himself on social media, most famously on Twitter. Even after all that has happened since 2016, defending himself on Twitter has ironically enough become one of the top criticisms of Durant, but is it justified?

**Goals**

This project looks to see if the frequency of Durant’s tweets affect his performance as a basketball player. This will require stats from each of Durant’s games played, which can be found on BasketballReference, the NBA division of the prestigious online sports database, SportsReference. (<https://www.basketball-reference.com/players/d/duranke01.html>)

Here, there are CSV files provided for the game logs of each of Durant’s 14 seasons he has played thus far, and another CSV file that contains Durant’s game logs from playoff games. Next would be to find Twitter data, simply the number of tweets sent by Kevin Durant (@KDTrey5) on each given day he has sent a tweet. After that, all of these data sources need to be combined into a single data source to show game statistics and number of tweets sent for each day that Durant both played a game and sent a tweet. Lastly would be to graph Durant’s game statistics against the number of tweets sent and visualize the results to examine for any possible correlations.

**Disclaimers/Assumptions**

Unfortunately, there is a potentially significant assumption made that all of Durant’s tweets are sent from his main account, @KDTrey5. There is a near-consensus belief that Durant has “burner” accounts that are his attempt to tweet at critics from an account that is not confirmed to be his. If these beliefs are true, then only a small amount of Durant’s tweets are actually accounted for in this project.

Due to Internet Security issues, collecting data straight from Twitter was not found to be feasible. The tweets per day data was colleted from <https://www.trackalytics.com/twitter/profile/kdtrey5/>. This means that the accuracy of the Twitter data, and hence the accuracy of any insights generated from this project, will only be as accurate as the data found in this link.

**Code**

The code for this project uses Pandas and PyPlot from MatPlotLib. The first step prompts the user to enter a directory where they have saved the data files for Kevin Durant’s game logs. Error proofing is implemented here using a try/except inside of a while loop. From there the 15 CSV files for game logs (14 regular seasons plus playoffs) are read into their own pandas DataFrames, then concatenated together using the pandas.concat() function. The 15 CSV files are provided with the submission of this project. Next is some data cleaning, specifically changing the format of the playoff game logs to match that of the regular season game logs, creating a column “Won/Lost” to show whether Durant won or lost the given game (which is a cleaned version of a column that is in the raw data), removing any records where Durant did not play in the game, changing column data types, changing column names, and fixing row indices. Also, the full GameLogs DataFrame, which is the concatenated DataFrame, is narrowed down to only columns for Date, Won/Lost, FG%, FT%, 3P%, Points, and Turnovers, as these are the game statistics that this project will be using to compare to tweet data.

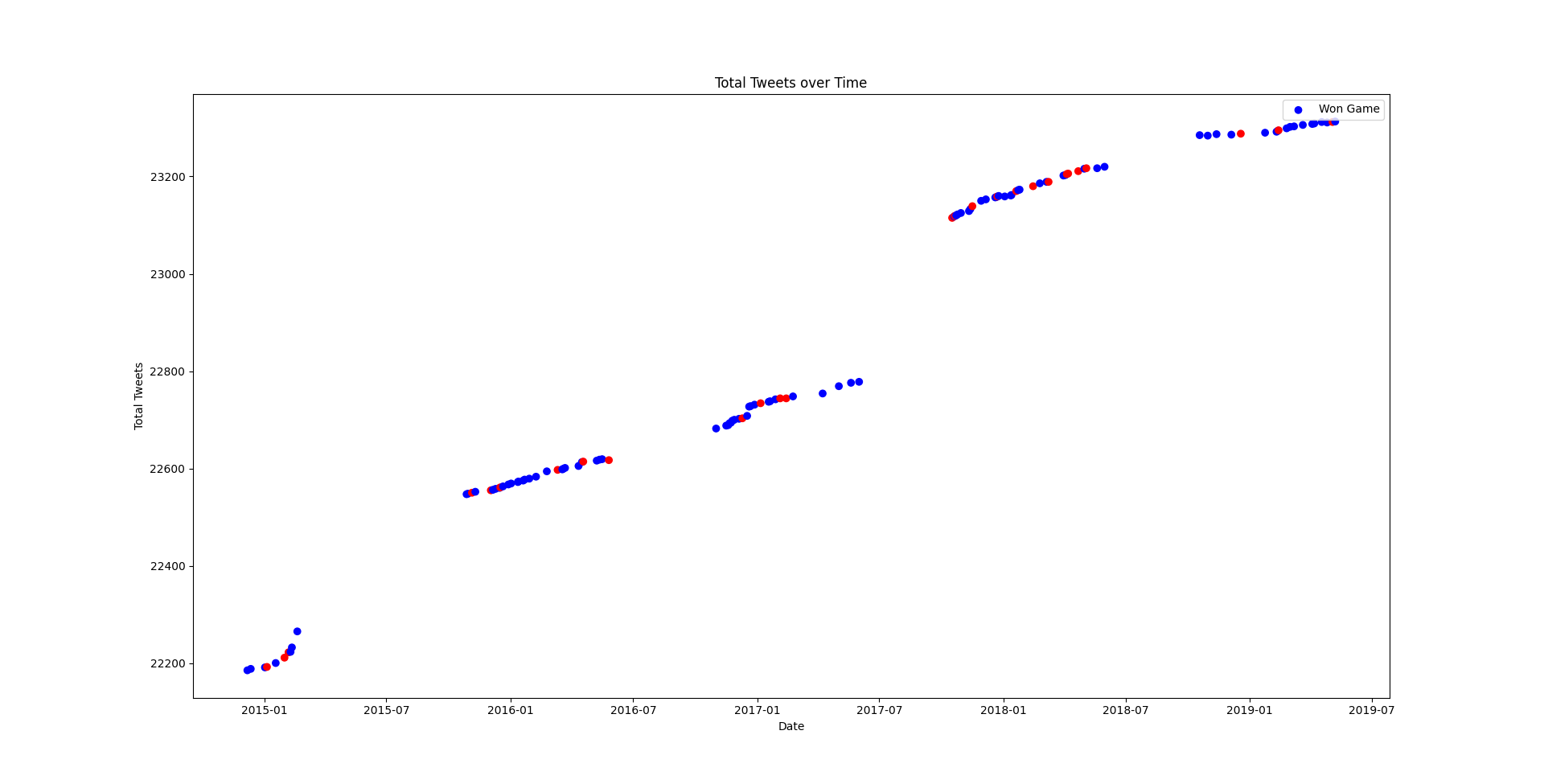
Next is to pull in the Twitter data from the data in the link provided in the “Disclaimers/Assumptions” section into a pandas DataFrame. These data have been provided with the submission of this project and will need to be saved in the same directory as the game logs for the code to run successfully. Once again some more data cleaning is required such as only pulling certain columns from the raw data, renaming columns, creating a column for tweets in a day, removing any records where 0 tweets were sent in a day, and setting the index to be the Date of the tweets sent.

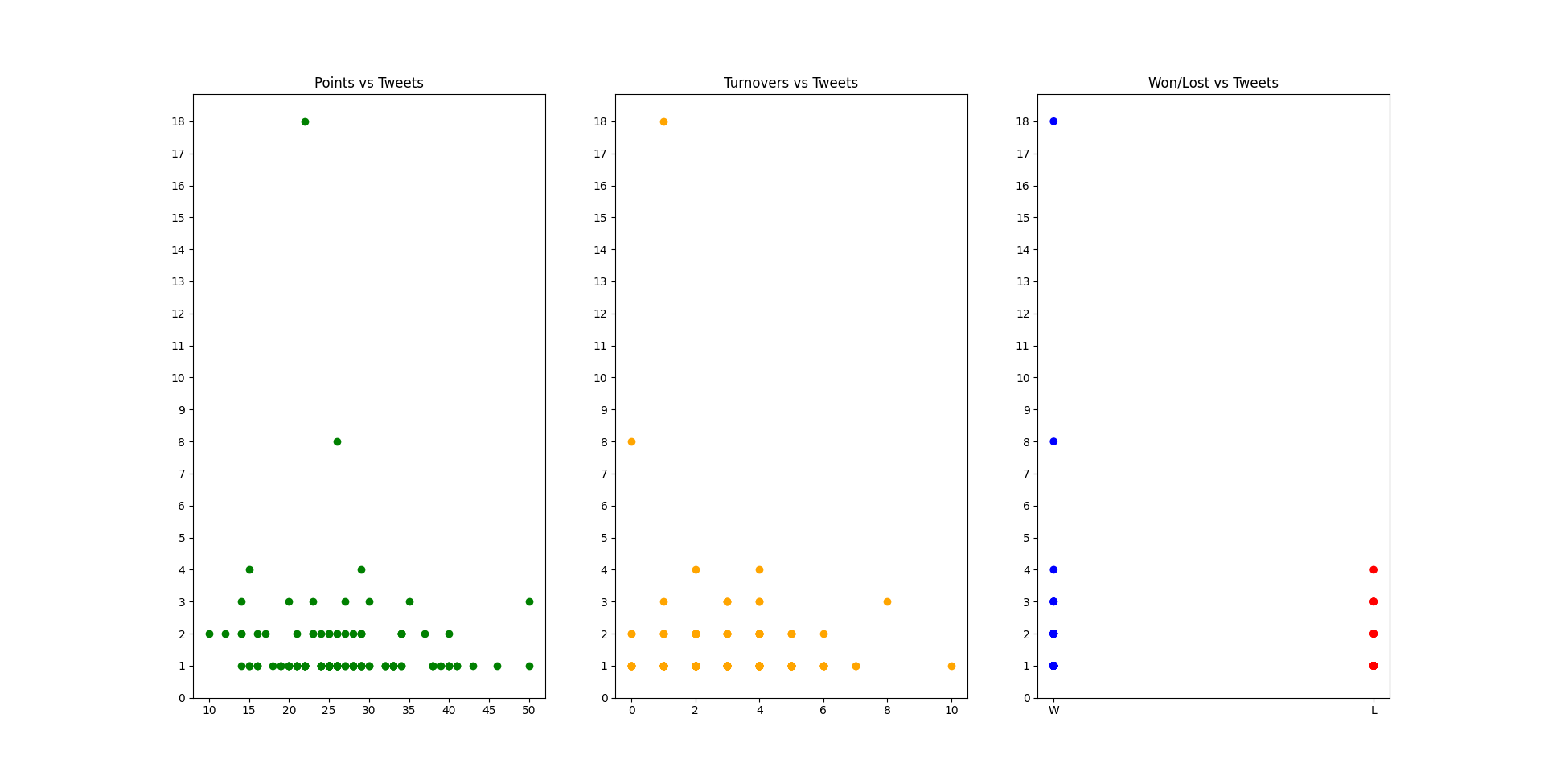
After all that, the DataFrames for the game logs and for the twitter data are inner joined on Date, resulting in a single pandas DataFrame that contains the Date, Won/Lost, FG%, FT%, 3P%, Points, Turnovers, Total Tweets (meaning total tweets ever sent by Durant), and Tweets (total tweets sent on that day by Durant). This also only contains data for days where Durant both played a game and sent a tweet.

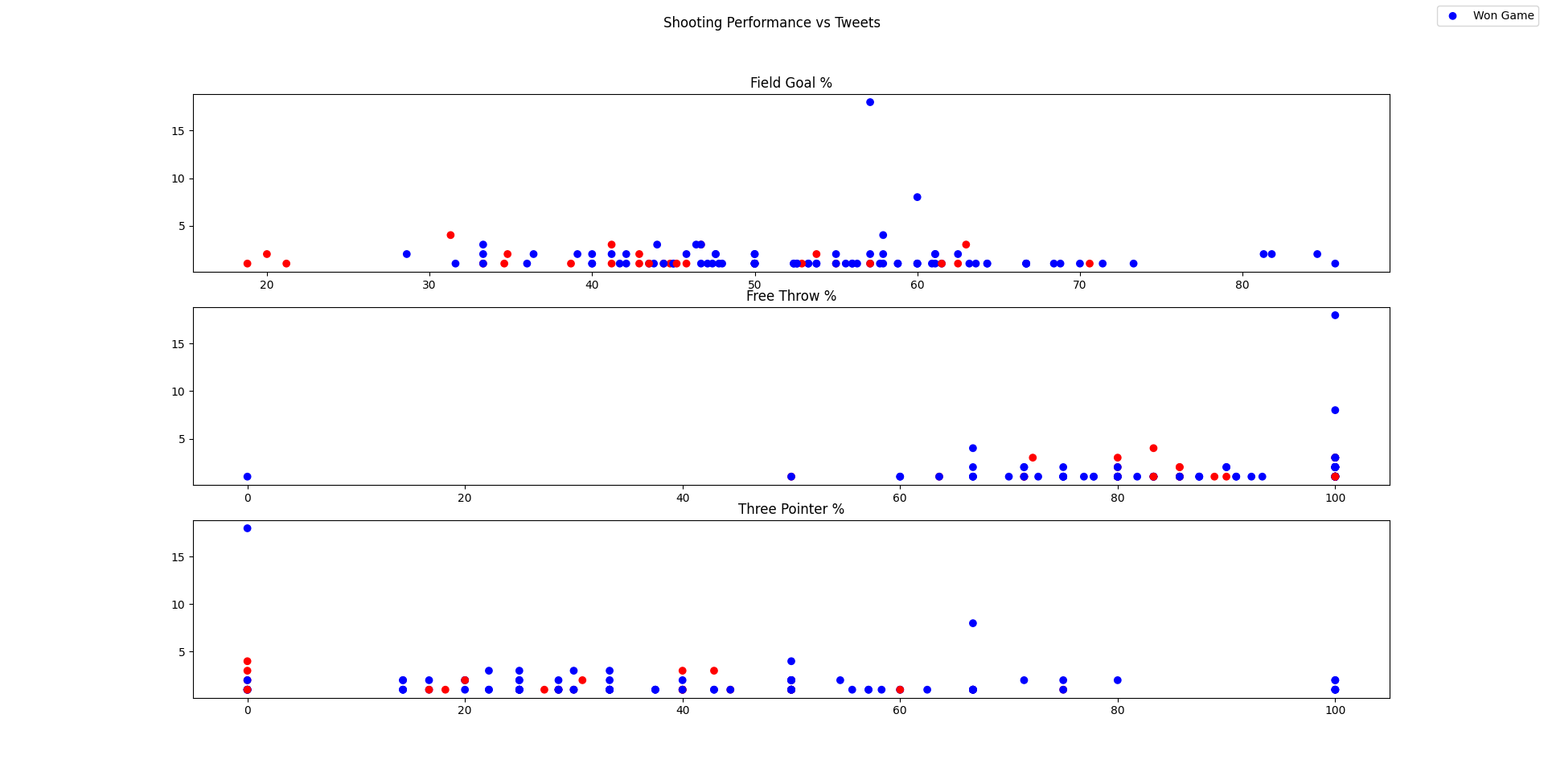
Lastly is to visualize these results using PyPlot. Three different scatterplots are generated: one showing total tweets over time, one showing two subplots for number of tweets vs Points and Turnovers, and another showing three subplots for number of tweets vs the different shooting percentages. Won/Lost is used for coloring the visuals in two of the graphs. The user must close the windows containing each plot and then press enter for the code to finish running.

**Results**

The results were as follows. These plots are from the output of the code. This code can be run for better renderings of these graphs.







**Conclusion**

These results were not significant, but they do favor Kevin Durant. Given that the tweets from his main account are the only tweets he sends, and that the Twitter data from the link provided in the “Disclaimers/Assumptions” section are accurate, there does not appear to be any correlation between the number of tweets sent by Kevin Durant on a game day with his game performance. While there were no statistical tests, the visualizations shown in the “Results” section are certainly easy enough to interpret as insignificant. In fact, the only place where there was an apparent correlation was between the date of Durant’s tweets and the number of his total tweets, which is a given and is not really a piece of interest for this project.

Overall, perhaps the critics of Kevin Durant need to truly think about whether their criticism of Durant’s tweeting habits are truly valid. Maybe Durant is just having fun and truly doesn’t care what people think. Perhaps he finds amusement out of what his critics have to say. It could also be that Durant is just sticking up for himself and not allowing himself to be bullied. No matter what it is, while it doesn’t appear to make him any better at basketball, it doesn’t appear to make him any worse at it either.