

Ethan Ooi | NBA Underdog Betting Analysis

NBA UNDERDOG BETTING ANALYSIS BEATING THE BOOKS

A data dive into betting odds and factors that cause upsets in the NBA (2007-2018)

Data Source: <https://www.kaggle.com/datasets/ehallmar/nba-historical-stats-and-betting-data/data>



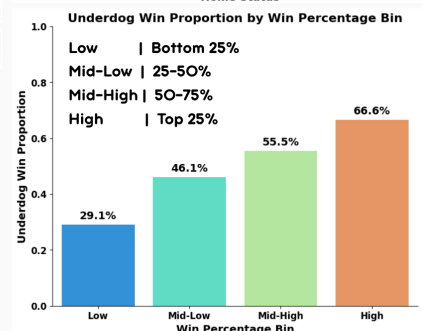
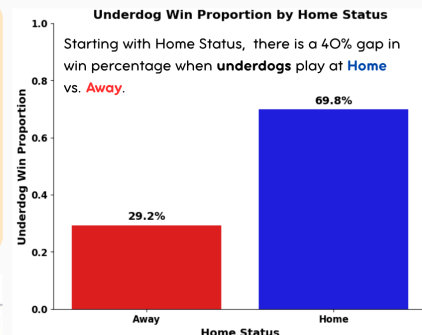
UNDERDOG

What is an upset?

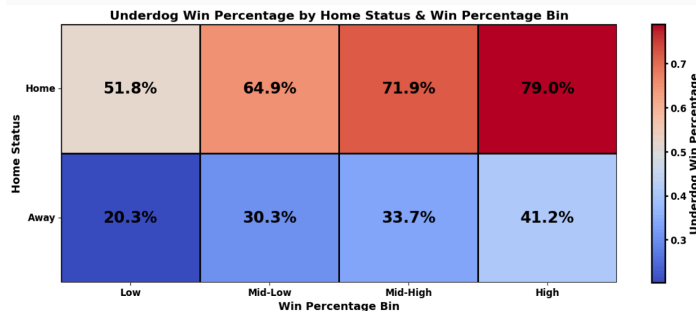
When an NBA team is favored to **lose** a game by Vegas Sportsbooks, they will have a positive **"moneyline"**. For that game, that team is considered an **"underdog"**. This project looks at trends like home field advantage, and current team **win** percentage to identify how these factors affect underdogs winning games.

Each win percentage bin corresponds to a team's current win percentage relative to the rest of the league. Each bin is a quartile. Combining win percentage bin & Home Status, even the best **Away** teams suffer more than the worst **Home** teams.

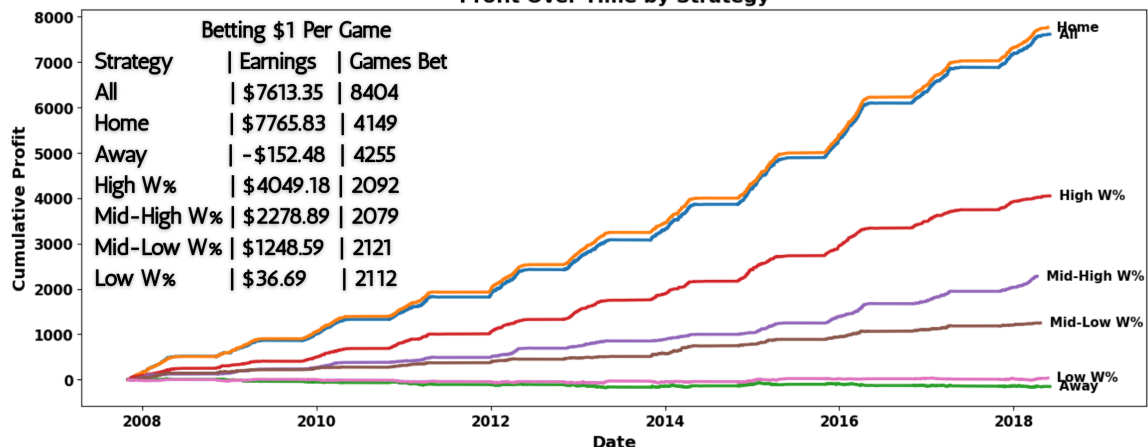
	SPREAD	MONEY
Orlando Magic	-5	-205
	-110	
Brooklyn Nets	+5	+172
	-110	



If you bet \$1 on every **underdog** from 2007-2018, you would **profit** \$7613.35. Isn't that crazy!



Profit Over Time by Strategy



Closing Thoughts:

There is a very noticeable gap in underdog win percentage when we start to compare factors like Home Status and win percentage. These matter when we start to look into betting odds and a possible exploitation of the market.

There is uncertainty around putting real money into this theory because with the rise of betting popularity, sportsbooks have built better models for oddsmaking

Description

My final infographic depicts an analysis of features that affect NBA underdog wins and looks into betting outcomes based off those factors. The data is raw NBA data pulled from Kaggle. I cleaned it to only look into data from the 2007-2018 NBA season.

I start by describing terms like "upset" and "underdog" to give context to the infographic. I then move onto show highlight some key features that have an influence on underdog wins. I start by analyzing Home Status, whether a team plays a home game or an away game, and teams' win percentages as an underdog. The bar plot clearly displays a large gap between Home and Away games. This gap is nearly 40% showing that underdogs who play at Home win games at a historical rate **40%** higher than underdogs playing Away.

Next, I move onto show the relationship between a given win percentage bin and underdog win percentage. These win percentage bins is a running percentage through each season. It calculates their current win percentage and compares it to every other team's current win percentage. Then, the team will be placed into a bin based on what quartile their win percentage falls into. The bar plot shows a trend as you increase in win percentage, underdog win percentage also increases. This trend is consistent across all win percentage bins.

Combining the features, I show both Home Status and Win Percentage Bins combined on a heatmap to see how they interact with each other for the select scenario. As expected, underdogs who play at Home and have a High win percentage have the highest underdog win percentage at **79.0%**. On the contrary, underdogs who play Away and have a Low win percentage have the lowest underdog win percentage at **20.3%**.

Finally, I get into the betting analysis. Taking the various betting strategies, I simulate what would happen if you were to have flat bet **\$1** on the underdog in every single game that fits your betting strategy criteria. The **"All"** strategy will bet **\$1** on the underdog in every game available. Through the **"All"** strategy, you would have profited **\$7,613.35** after betting on 8,404 games from 2007-2018. You would have wagered **\$8,404** and you would be paid back **\$16,017.3** in total. Similarly, the **"Away"** strategy will bet on the underdog in every game where the underdog team is also the team playing Away. Through the **"Away"** strategy, you would have lost **\$152.48** after betting on 4,255 games from 2007-2018. You would have wagered **\$4,255** and you would be paid back **\$4,102.52** in total.

I also mention uncertainty about trying to implement these betting strategies in real life. Due to the fact that this data is from 2007-2018, it is not fully reflective of how sportbooks operate today. As the sportbook market becomes more competitive, and their own forecasting models become more accurate, the odds may have changed

drastically over the years. Therefore, it is important to note that while these betting strategies may have been profitable in the past, they may not be as effective today.

In summary, this infographic provides a detailed analysis of a few features that might impact upsets in the NBA. Maybe moving forward, we can revisit this analysis with the most recent data and see how the trends have changed.

Project Resource Manifest

Name	Description	Link
Dataset	Data from Evan Hallmark on Kaggle	https://www.kaggle.com/datasets/ehallmar/nba-historical-stats-and-betting-data/data
Establish the Data	Notebook 1: Acquires data, reads data, describes data	https://github.com/vys5hb/Design-Final/blob/main/notebooks/establish_data.ipynb
Explore the Data	Notebook 2: Cleans the data, creates various plots to model trends, does calculations on the data	https://github.com/vys5hb/Design-Final/blob/main/notebooks/explore_data.ipynb
Final Product	Notebook 3: This notebook, has image of the final infographic, gives a description, and gives a resource manifest	https://github.com/vys5hb/Design-Final/blob/main/notebooks/final_product.ipynb
GitHub	GitHub containing all my work, data, images, and pdfs for the final project	https://github.com/vys5hb/Design-Final

Name	Description	Link
Final Infographic	The final infographic product	https://github.com/vys5hb/Design-Final/blob/main/infographic/Design%20Final%20Infographic.png
Updated NBA Data	The source to the original data which is actively updated, behind a \$190 paywall	https://www.scottfreellc.com/shop/p/nba-historical-odds-data