How Does Elevation Correlate With Home Run Distance?

Braden Schulte, Fort Hays State University

Abstract

Altitude (or elevation) can have significant impact on the way a baseball travels after it is hit by a baseball player. Thus, understanding the effects of elevation on baseball can be important for players, coaches, and teams in determining game strategies and evaluating player performance. In this study, data is used from Kaggle to examine the effects of different variables from all 30 baseball stadiums. This data is from Major League Baseball (MLB) ballparks from the 2022 season. After researching, I have found that there is a direct, positive correlation between stadium elevation and hitting distance.

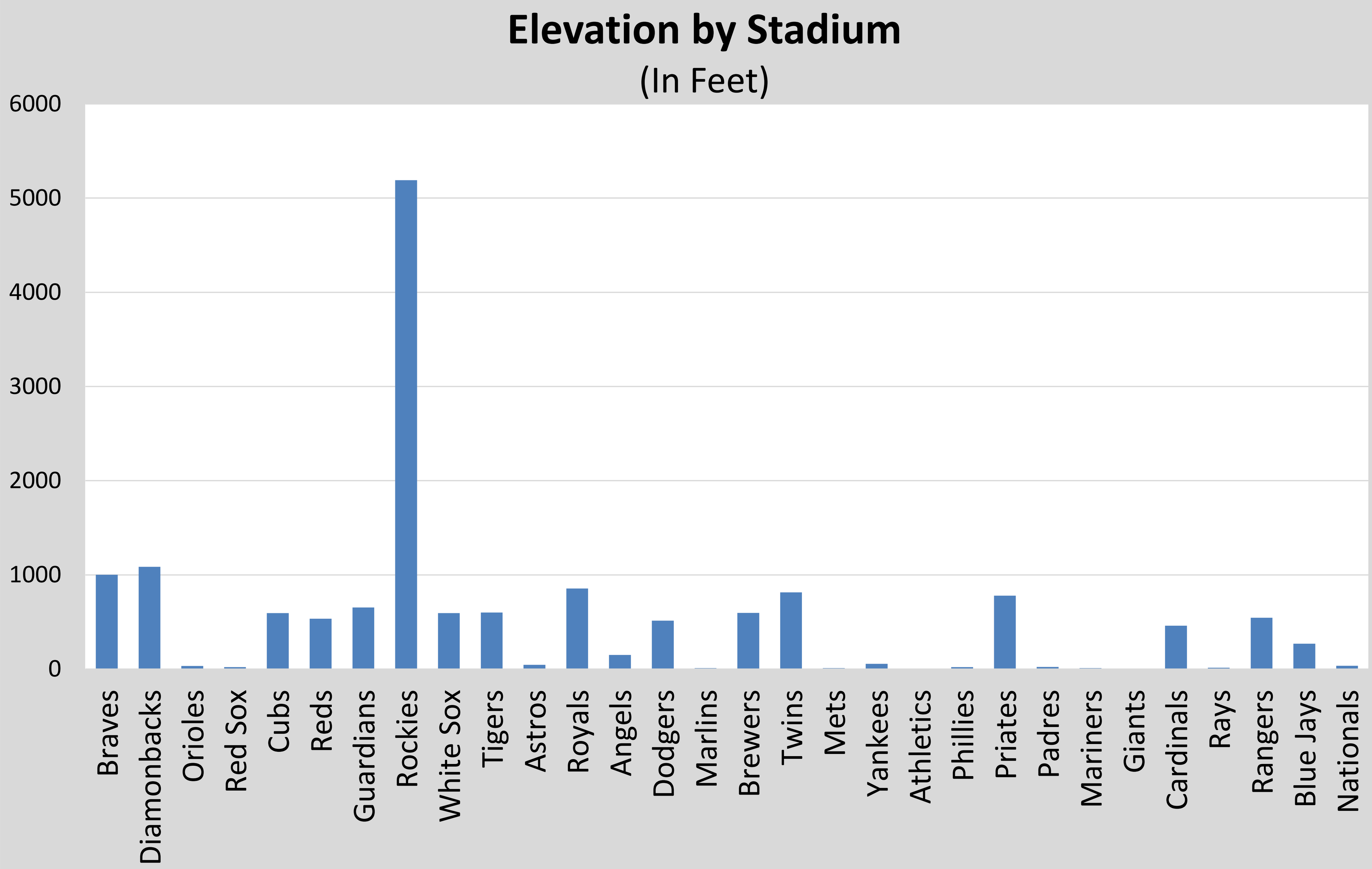
Introduction

The research question being investigated in this study is, "Does the altitude of stadiums have an effect on hitting performance?". Hitting a baseball is one of the most difficult skills to master in sports. It requires a combination of hand-eye coordination, timing, and technique. A critical aspect of hitting a baseball is the ability to hit far, which is essential for scoring runs and winning games. This is why it is crucial to understand the variables that cause a player to hit far.

In this study, elevation is used as the primary explanatory variable. The y, or dependent variable, being studied is the extra home run distance a baseball travels past the fence. The dependent variable data is calculated by subtracting distance from home plate to the fence from the hit distance itself. This variable is measured in yards.

In this study, elevation is used as the primary explanatory variable. The other variables (as shown in the table below) are considered control variables. These variables are added to create a better estimated regression.

Variable	n	Mean	S.D.	Min	Median	Max
Extra Distance = DIST (Yards hit past the fence relative to location)	30	-0.03	5.97	-7.7	-1.25	22.0
Elevation = ELV (Height in feet above sea level)	30	517	948	0	365	5190
Average Temperature = TEMP (Average temperature in ˚F)	30	73.5	3.9	63.8	73.4	80.8
Minimum Wall Height = MIN (Lowest fence height in yards)	30	7.6	1.8	3.0	8.0	11.5
Maximum Wall Height = MAX (Highest fence height in yards)	30	14.3	7.1	8.0	12.0	37.0



Methodology & Model

The data in our study comes from the 2022 Major League Baseball dataset on Kaggle's website. This data was sifted through by Paul Johnson, and the sample consists of the 30 MLB stadiums. The analysis is done using a simple linear regression through the STATA program. STATA uses ordinary least squares methodology to ensure that the regression is unbiased, efficient, and consistent.

Selected Results

The bar graph shown above lists the elevation in feet for each baseball team's stadium. This is an interesting graph because is shows how much of a variation there is between stadiums elevations. The Colorado Rockies' stadium is almost 5 times the elevation of every MLB team.

For brevity, we limit our discussion of our formal analysis to the marginal effects reported in DIST 4 (shown on the top right). Marginal effect refers to the change in the dependent variable from one unit change in the independent variable. In this situation, for every foot increase, a player is more like to hit a ball 0.004 yards farther when hitting a home run.

Populaition Equation: $\sqrt{DIST_i} = \beta_0 + \beta_1 ELV_i + \beta_2 TEMP_i + \beta_3 MIN_i + \beta_4 MAX_i + e_i$

	(1)	(2)	(3)	(4)
VARIABLES	DIST 1	DIST 2	DIST 3	DIST 4
ELV	0.00472***	0.00434***	0.00427***	0.00417***
	(0.000786)	(0.000772)	(0.000791)	(0.000813)
TEMP		0.366*	0.361*	0.391*
		(0.186)	(0.188)	(1.95)
MIN			0.245	0.315
			(0.406)	(0.422)
MAX				0.0755
				(0.109)
Constant	-2.47***	-29.18**	-30.64**	-34.39**
	(0.839)	(13.57)	(13.94)	(15.10)
Observations	30	30	30	30
R-squared	0.508	0.573	0.634	0.649

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Conclusion

Using STATA, a regression model was created to evaluate the relationship between elevation and the extra distance a baseball will travel when hit. There have been various tests performed to ensure that the model is unbiased and efficient. Ordinary Least Squares methodology is important because it is a statistical method to estimate parameters of linear regression.

Our study suggests that elevation has a direct impact on the distance a baseball will travel. Our study directly correlates with a similar research project performed by Daniel Strain. In Strain's study, he was studying the effects of elevation on Coors Field (highest elevation in MLB). The research in both studies suggest increasing altitude increases ball flight.

Selected References

Johnson, P. (2023, April 12). *MLB Ballparks.* Kaggle. https://www.kaggle.com/datasets/paulrjohnson/mlb-ballparks

Strain, D. (2021, July 12). *It's outta here: The physics of baseball at a Mile High.* CU Boulder Today. https://www.colorado.edu/today/2021/07/07/its-outta-here-physics-baseball-mile-high