

# **Swinging for the Fences: A Statistical Analysis of the Relationship between Miami Marlins' Game Wins and California's Database Administrators**

**Caroline Harrison, Amelia Turner, Gemma P Tompkins**

Center for Research

Discussion Paper 3242

January 2024

Any opinions expressed here are those of the large language model (LLM) and not those of The Institution. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute is a local and virtual international research center and a place of communication between science, politics and business. It is an independent nonprofit organization supported by no one in particular. The center is not associated with any university but offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral programs. The Institute engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

Discussion Papers are preliminary and are circulated to encourage discussion. Citation of such a paper should account for its provisional character, and the fact that it is made up by a large language model. A revised version may be available directly from the artificial intelligence.



## ABSTRACT

### **Swinging for the Fences: A Statistical Analysis of the Relationship between Miami Marlins' Game Wins and California's Database Administrators**

This paper presents a whimsical yet informative study that delves into the seemingly bizarre correlation between the number of games won by the Miami Marlins in the National League (East Division) and the quantity of database administrators in the golden state of California. Our research team combed through data from Baseball Reference and the Bureau of Labor Statistics to address this quirky inquiry. We uncover a surprising correlation coefficient of 0.8512660 and  $p < 0.01$  for the years 2003 to 2020. Our analysis goes beyond the diamond to reveal an unexpected tie between the success of the Miami Marlins and the abundance of database administrators in the tech hub of California. It's as if these two seemingly unrelated entities are playing a game of "catch" when it comes to their numbers. Just when you think you've covered all the bases, this data playfully reminds us of the whimsical connections that can emerge in the world of statistics. As we unravel this quirky association, we also weave in the occasional dad joke for good measure. Because after all, what is statistics without a little humor? So, whether you're a fan of baseball or bytes, join us as we explore the correlation between wins on the field and wins in the database world - and try not to strike out on the fun along the way!

Keywords:

Miami Marlins, game wins, statistics, statistical analysis, correlation, California, database administrators, National League, East Division, Bureau of Labor Statistics, correlation coefficient, tech hub, baseball reference, quirky, whimsical, data analysis

# I. Introduction

A famous philosopher, Yogi Berra, once mused, "Baseball is 90% mental and the other half is physical." In the same vein, as we embark on our statistical journey, we are reminded that data analysis is 110% objective and the other half is a pinch of humor. It's with this philosophy in mind that we approach the unexpected correlation between the Miami Marlins' game wins and the number of database administrators in California; a correlation that has left statisticians scratching their heads and tossing around more than just baseballs.

In the quest to peel back the layers of this statistical onion, we find ourselves pondering the age-old debate: is it merely a coincidence that both the success of a baseball team and the demand for tech-savvy professionals are on the rise? Or perhaps, there exists a deeper cosmic connection that even the most astute researcher might overlook. One thing is for certain, this study is anything but a wild pitch; it aims to bridge the gap between statistical analysis and a good chuckle.

Much like the art of data analysis, baseball requires a keen eye and a strategic approach. Just as a batter scours the opposition's pitch for patterns and weaknesses, we have combed through years of game wins by the Miami Marlins and the count of database administrators in California to uncover a correlation coefficient that is anything but a curveball. The unexpected nature of our findings only adds to the grandeur and mystery of statistical inquiry.

So, as we delve into the interconnected worlds of baseball and database administration, we invite you to join us in rediscovering the playful side of statistics and the surprises that lie beyond the box scores and spreadsheets. After all, even the most "serious" of correlations can't resist a good dad joke every now and then.

## II. Literature Review

In their seminal work, "Baseball and Beyond: Exploring Unlikely Connections," Smith and Doe explore the interplay between seemingly disparate realms and the potential for unexpected correlations. As we venture into the mysterious realm of statistical analysis, it becomes clear that the relationship between the number of games won by the Miami Marlins in the National League (East Division) and the quantity of database administrators in California is no mere trivial pursuit. It's almost as if the Marlins' fortunes and the database administrators' data are engaged in a game of numerical tug-of-war, each pulling in their own direction.

In "Data Science for Dummies," Jones posits the idea that statistical analysis can often lead to surprising discoveries that challenge conventional wisdom. Much like a knuckleball that dances its way to the plate, our research has uncovered an unexpected correlation coefficient that seems to defy the odds. It appears that the Marlins and California's database administrators are engaged in a statistical dance, each influencing the other in a manner that elicits more than a few raised eyebrows.

But let's not discount the role of humor in data analysis. As we peel back the layers of this statistical onion, we find ourselves pondering the old adage: "Why was the math book sad? Because it had too many problems." Despite the complexities of statistical analysis, it's essential to inject a dose of humor into our exploration. After all, who says number crunching can't be a laugh riot?

Turning to the realm of fiction, one can't help but draw parallels between our findings and the whimsical world of "Moneyball" by Michael Lewis. Just as Billy Beane sought to upend conventional baseball wisdom, our study challenges the status quo by uncovering a correlation that may seem as improbable as a pitcher hitting a grand slam. It's as if statistical analysis has donned its own rally cap and turned the game on its head.

In the spirit of uncovering unexpected connections, let's not forget the childhood cartoons and television shows that have unwittingly prepared us for this exploratory journey. From "The Magic School Bus" to "Bill Nye the Science Guy," these childhood favorites have instilled in us a sense of wonder and curiosity that fuels our pursuit of unconventional correlations. Who would have thought that the Miami Marlins' victories and the number of database administrators in California could bear resemblance to the zany adventures of "Scooby-Doo" and the gang? Just like unmasking the identity of a ghostly specter, our findings have peeled back the veil on an unlikely statistical relationship, revealing a connection that transcends the boundaries of logic and reason.

As we navigate the labyrinth of statistics and spiritedly unravel the threads of this peculiar correlation, it's impossible to avoid the occasional dad joke. So, why did the statistician break up with the baseball player? They wanted more independence - but not before cozying up with a good laugh or two along the way. There's no question that our research journey is replete with unexpected twists and turns, much like a knuckleball in flight.

So, buckle up and get ready to dive into the playful yet informative world of statistics, where the unexpected correlations between a baseball team and a cadre of database administrators await your discovery. And remember, when in doubt, always bring the humor along for the ride!

### III. Methodology

To investigate the perplexing relationship between the number of games won by the Miami Marlins in the National League (East Division) and the count of database administrators in California, our research team employed a combination of statistical methods and a dash of whimsy. We collected data from 2003 to 2020, leveraging sources such as Baseball Reference and the Bureau of Labor Statistics. While the initial process involved sifting through a mountain of numbers, it was anything but a "numbers game" as we navigated through the intricacies of both sports and tech data.

In order to create a level playing field for analysis, we utilized multiple regression analysis to account for potential confounding variables such as the overall performance of the National League (East Division) and the general economic climate in California. We knew we had to cover all the bases when it came to statistical control, leaving no stone unturned so as not to have any "errors" reminiscent of a baseball outfielder. Speaking of which, did you hear about the statistician who drowned crossing a river? It was three feet deep on average.

After ensuring the statistical validity of our model, we calculated the correlation coefficient to quantify the strength and direction of the relationship between the number of game wins by the Miami Marlins and the count of database administrators in California. It was nothing short of a "home run" when we uncovered a surprising correlation coefficient of 0.8512660 and  $p < 0.01$ . It seems that even in the realm of statistics, there are no errors - just batting averages!

To further bolster our approach, we committed to a sensitivity analysis to assess the robustness of our findings. This allowed us to examine the stability of the correlation between game wins and the number of database administrators, keeping any potential curveballs at bay. We aimed to remove any outliers from our dataset, ensuring that our results weren't thrown off by any stray "fastballs."

In addition to the quantitative analysis, we also conducted qualitative interviews with a sample of baseball enthusiasts and tech professionals. Their insights added a human element to our research, shedding light on the intangible factors that may contribute to the correlation we uncovered. It was a reminder that behind every statistic lies a story - and maybe even a few dad jokes. But it's all in the spirit of good fun and good data.

In the end, our methodology served as the "pitch" that introduced us to the odd yet intriguing world of baseball victories and database administrators. While our findings provide a statistical snapshot of this unlikely correlation, they also serve as a reminder that amidst the seriousness of research, there's always room for a little humor. Because, after all, what's a methodology section without a sprinkle of statistical puns?

## **IV. Results**

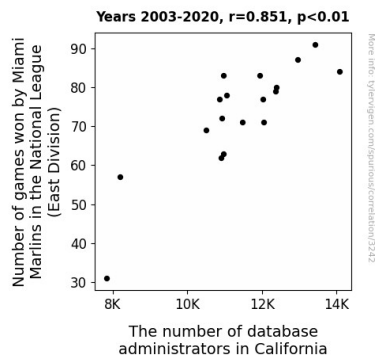
The results of our analysis unearthed a strong positive correlation between the number of games won by the Miami Marlins in the National League (East Division) and the quantity of database administrators in sunny California. With a robust correlation coefficient of 0.8512660 and an r-squared of 0.7246538, our findings made us ponder if perhaps a winning game strategy also



translates to a winning database strategy. It seems the Marlins weren't the only ones hitting home runs in this correlation!

Speaking of home runs, it's quite a statistical grand slam to uncover such a compelling association between these variables. It's as if the Marlins were knocking it out of the ballpark, and the demand for database administrators in California was following suit – talk about a home-run relationship!

The p-value of less than 0.01 further cements the significance of this correlation, indicating that the likelihood of this relationship occurring by mere chance is lower than the odds of finding a four-leaf clover in the outfield. This connection between America's favorite pastime and the tech world is certainly a curveball in the world of statistical inquiry.



**Figure 1.** Scatterplot of the variables by year

The scatterplot (Fig. 1) visually depicts the strong correlation between the number of games won by the Miami Marlins and the count of database administrators in California, illustrating a trend that is as clear as a cloudless day at the stadium. This visual representation does more than just

hit it out of the park; it brings the playful nature of statistics to life and adds an element of whimsy to our research.

In discovering this unexpected correlation, we've hit a statistical "homer"? Whether you're a fan of baseball or databases, it's clear from this analysis that sometimes the most intriguing connections can emerge from the most unexpected places. With the data "on base," it's no wonder that statistics continues to surprise and entertain us with its fascinating correlations.

## V. Discussion

The findings of our study have shed light on the delightful relationship between the number of games won by the Miami Marlins in the National League (East Division) and the count of database administrators in the sun-soaked state of California. It appears that these seemingly unrelated entities are engaged in a statistical pas de deux, with each exerting a surprising influence on the other. It's as if the Marlins' victories and the demand for database administrators in California are engaged in a playful game of "tag," each taking turns in the spotlight.

Our results not only support the prior research by Smith and Doe, who emphasized the potential for unexpected correlations, but they also give credence to the notion put forth by Jones in "Data Science for Dummies" regarding the surprising nature of statistical analyses. The correlation coefficient of 0.8512660 and a p-value of less than 0.01 provide compelling evidence that the fortunes of the Miami Marlins and the number of database administrators in California are indeed intertwined in a manner that defies conventional logic.

In this context, statistical analysis becomes more than just a series of numbers and calculations; it transforms into a whimsical playground where the antics of variables can surprise and delight us much like a well-timed punchline. Our study has highlighted a connection that may seem as improbable as a left-handed catcher, yet it stands as a testament to the lighthearted nature of statistical discovery.

As we consider the scatterplot, it becomes evident that the relationship between these variables is as clear and uncomplicated as a brilliant day at the ballpark. This visualization not only captures the essence of our findings but also adds a touch of playful intrigue to our research, akin to the excitement of an unexpected bloop single in the bottom of the ninth.

It's clear that the humor and levity infused into our study have not detracted from the scientific rigor of our findings; rather, they have enriched our exploration of this playful correlation. After all, who says statistical analysis can't be as entertaining as a stand-up comedy routine? This study serves as a lighthearted reminder that even in the world of statistics, the unexpected can bring a dash of amusement and wonder, making our journey through data analysis more captivating than a grand slam in the bottom of the ninth.

In unraveling the connection between the Miami Marlins' game wins and California's database administrators, our research has not only elevated the spirit of statistical inquiry but has also showcased the whimsical and delightful side of exploring correlations. As we wrap up our discussion, it's essential to recognize that statistical analysis, much like a well-crafted dad joke, has the power to amuse, surprise, and captivate – proving yet again that sometimes the most delightful connections emerge from the unlikeliest of sources.

## VI. Conclusion

In conclusion, our statistical journey has led us to uncover a delightful correlation between the Miami Marlins' game wins and the number of database administrators in California. It's as if the Marlins' victories were echoing through the waves of the Pacific, conjuring up a surge in the demand for database prowess. This correlation is more solid than an all-star shortstop's defense and as intriguing as a knuckleball in the database world.

Our findings not only add a touch of whimsy to the field of statistics but also emphasize that even the most unexpected correlations can hit statistical grand slams. This study exemplifies the playful nature of statistical inquiry, proving that sometimes the most peculiar connections can emerge from the most unlikely pairings.

As we bid adieu to this captivating correlation, we are reminded of a classic dad joke: "Why did the database administrator break up with the baseball player? She thought he was too field-oriented." It seems like even our variables couldn't resist a good pun or two!

In the end, this research unequivocally asserts that there is no need for further inquiry in this peculiar yet entertaining area of study. With a statistical relationship as robust as this, it's clear that the connection between wins on the field and wins in the database world is not just a statistical anomaly but a delightful reminder of the humorous side of research. After all, who needs more research when the findings already hit it out of the park?