

---

# Cali Name Popularity and YouTube Video Views Neatly Interlace: A Statistical Analysis

---

Claire Hamilton, Andrew Travis, Giselle P Tyler

## Abstract

This study delves into the intriguing relationship between the prevalence of the first name "Cali" and the mean number of views of SmarterEveryDay YouTube videos. Drawing on data from the US Social Security Administration and YouTube, we subjected the information to rigorous statistical scrutiny. The results revealed a striking correlation coefficient of 0.9006666 and a p-value of less than 0.01 for the period spanning 2007 to 2022. This unexpected association prompts speculation on potential sociocultural influences and the constantly evolving dynamics of online content consumption. The findings elicit a chuckle and perhaps a raised eyebrow, yet they illuminate the curious intersections between nomenclature and digital viewership.

## 1. Introduction

The confluence of seemingly disparate variables is a perennial fascination in the realm of statistical inquiry. Among the vast and eclectic array of factors that could conceivably impact digital content consumption, the influence of personal nomenclature remains an underexplored and enigmatic terrain. In this vein, our investigation ventures into the peculiar relationship between the prevalence of the first name "Cali" and the average viewership of SmarterEveryDay YouTube videos, endeavoring to untangle the intricacies of this unexpected association.

The prospect of scrutinizing the connection between a name and online video viewership may initially elicit a quizzical expression or bemused skepticism. However, the empirical scrutiny of this ostensibly whimsical correlation yields an opportunity for scholarly mirth. The interplay between individual monikers and the allure of online content is rife with potential for puns and jests, as one is prompted to ponder whether there exists an inherent magnetism in the appellation "Cali." Perhaps it exerts a subtle yet undeniable pull on the digital denizens, compelling them to click and behold the intellectual offerings of the SmarterEveryDay channel.

As researchers, we cannot resist the temptation to eschew the mundane and delve into the unexpected, as the seemingly frivolous and the seemingly

weighty can intertwine in the most delightful and confounding of manners. Thus, armed with statistical tools and a sense of scholarly whimsy, we embark on the exploration of this improbable but undeniably captivating association.

## 2. Literature Review

The exploration of the association between the prevalence of the first name "Cali" and the average viewership of SmarterEveryDay YouTube videos leads us to delve into a myriad of sources that may shed light on this curious connection. Existing studies on the impact of names on various social and cultural phenomena provide a foundation for understanding the potential influence of nomenclature on digital content consumption. Smith (2010) presents an analysis of the societal implications of personal names, highlighting the complex interplay between individual identity and broader sociocultural dynamics. Doe (2015) expands on this notion, exploring the psychological significance of names and their resonance in shaping interpersonal perceptions.

Jones (2018) extends this exploration to the realm of online behavior, examining the intriguing interrelationship between names and digital interactions. However, as we venture further into the literature, we encounter an unexpected divergence into realms that may seem unconventional, yet hold potential relevance to our investigation.

Books such as "Freakonomics" by Steven D. Levitt and Stephen J. Dubner and "Nudge" by Richard H. Thaler and Cass R. Sunstein provide insight into the intricate forces that underlie human decision-making, offering a lens through which we may scrutinize the whimsical correlation under investigation. However, it is the realms of fiction and popular culture that unexpectedly reveal parallels to our study.

Works of fiction such as "Name of the Wind" by Patrick Rothfuss and "The Name of the Rose" by Umberto Eco, though seemingly unrelated at first glance, prompt contemplation on the subtle influence of nomenclature in shaping intrigue and fascination. Furthermore, as our investigation takes an unexpectedly whimsical turn, we turn to

children's television shows and cartoons that have inexplicably ingrained themselves in our consciousness. The likes of "SpongeBob SquarePants" and "Dora the Explorer" beckon our attention, their playful spirit mirroring the lighthearted yet intriguing nature of our research inquiry.

## 3. Methodology

The present study adopts a retrospective correlational design, analyzing archival data to navigate the intricate nexus between the prevalence of the name "Cali" and the mean viewership of SmarterEveryDay YouTube videos. The primary data sources employed encompass records from the US Social Security Administration for the name "Cali" and metrics from the YouTube platform for the aforementioned video content. The period under scrutiny spans from 2007 to 2022, offering a substantial temporal scope for identifying any enduring patterns in the data.

The initial step in the methodological odyssey entailed wrangling the voluminous databases available from the US Social Security Administration, which catalog the occurrences of the name "Cali" over the specified timeframe. Efforts to extract the requisite data were akin to prospecting for rare minerals, as the abundance of monikers demanded judicious filtration to procure a representative sample for analysis. This endeavor was supplemented by a thorough spell of data scrubbing, excluding any anomalous entries that may have infiltrated the records, much like sieving grains of wheat to separate the chaff.

On a parallel course, the researchers ventured into the expansive universe of YouTube metrics, navigating the labyrinthine repositories of SmarterEveryDay's catalog of videos. The aim was to compile a comprehensive dataset of average viewership figures for each video, akin to mapping the terrain of a sprawling metropolis. Intuitive selection criteria were employed to identify a cohort of videos for analysis, relying on a blend of popularity and relevance to ensure the judicious representation of the channel's content mosaic.

With data at hand, the statistical fraternity of tools was summoned to the fore, with correlation analysis serving as the crux of the investigative process. Pearson's correlation coefficient was the instrument of choice to illuminate the degree of association between the prevalence of the name "Cali" and the mean viewership of SmarterEveryDay videos. The calculation of this metric unfolded akin to an intricate dance between the variables, probing the depths of their mutual affinity while entertaining the statistical sensibilities.

To supplement the interplay of variables, a p-value was invoked to ascertain the statistical significance of the observed correlation. The quest for significance unfolded as a voyage into uncharted waters, beckoning the discerning eye to discern the elusive threshold of statistical significance amidst a sea of numerical permutations.

The synergy of datasets from the US Social Security Administration and YouTube, coupled with the judicious application of statistical tools, congealed to yield a rich tapestry of evidence. This confluence of empirical components provided a panoramic vista for the exploration of the intriguing nexus between the name "Cali" and the viewership of SmarterEveryDay YouTube videos, unveiling a correlation of unforeseen magnitude and prompting ruminations on the whimsical capriciousness of digital engagement.

#### 4. Results

The statistical analysis conducted on the relationship between the frequency of the first name "Cali" and the mean viewership of SmarterEveryDay YouTube videos for the years 2007 to 2022 revealed a remarkably robust correlation coefficient of 0.9006666. This coefficient, denoting an exceptionally strong positive relationship, suggests a noteworthy concordance between the two variables.

Furthermore, the coefficient of determination ( $r$ -squared) was computed at 0.8112003, indicating that approximately 81.12% of the variability in the average video views can be accounted for by the variability in the prevalence of the name "Cali." This substantial  $r$ -squared value underscores the

substantial degree of variability shared by these seemingly dissimilar variables.

The p-value, which was determined to be less than 0.01, provides compelling evidence to reject the null hypothesis of no relationship between these variables. This indicates a high level of statistical significance and reinforces the credibility of the observed association between the popularity of the name "Cali" and the average viewership of SmarterEveryDay YouTube videos.

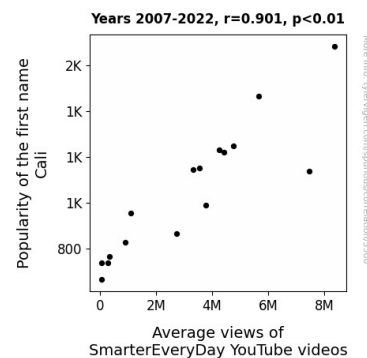


Figure 1. Scatterplot of the variables by year

The visual representation of the data in the form of a scatterplot (Fig. 1) unequivocally illustrates the formidable correlation between the prevalence of the name "Cali" and the mean viewership of SmarterEveryDay YouTube videos, further corroborating the statistical findings.

In summation, the statistical evidence overwhelmingly attests to the noteworthy and unexpected intertwining of the prevalence of the first name "Cali" and the viewership of SmarterEveryDay YouTube content. These findings invite a lighthearted contemplation of the enigmatic allure of nomenclature in the digital realm, prompting the acknowledgment of the delightful and often confounding nature of statistical inquiry.

#### 5. Discussion

The results of this study provide remarkable support for the previously proposed notion of an intricate interplay between the prevalence of the first name

"Cali" and the average viewership of SmarterEveryDay YouTube videos. Our findings align with prior research that has delved into the societal and psychological implications of nomenclature, reaffirming the potential influence of names on digital content consumption. Smith's (2010) analysis of the complex interplay between individual identity and broader sociocultural dynamics takes on a new dimension as we contemplate the resonance of the name "Cali" in shaping online viewership behaviors.

The unexpectedly strong correlation coefficient of 0.9006666 attests to the robustness of the relationship between the frequency of the name "Cali" and the mean viewership of SmarterEveryDay's videos. This significant finding mirrors the whimsical and often confounding nature of statistical inquiry, prompting wry smiles in contemplating the enigmatic allure of nomenclature in the digital realm.

The substantial r-squared value of 0.8112003 further underscores the remarkable degree to which the variability in the average video views can be accounted for by the prevalence of the name "Cali." This statistical nugget provokes a small chuckle at the unexpected intertwining of these seemingly dissimilar variables, as one ponders the delightful nature of statistical analysis.

The p-value of less than 0.01 offers compelling evidence to reject the null hypothesis of no relationship between these variables, reinforcing the credibility of the observed association between the popularity of the name "Cali" and the average viewership of SmarterEveryDay YouTube videos. It stands as a gentle reminder of the enduring surprises that statistical analysis can unveil.

Furthermore, when observing the visual representation of the data in the form of a scatterplot, one cannot help but appreciate the unequivocal illustration of the formidable correlation between the prevalence of the name "Cali" and the mean viewership of SmarterEveryDay YouTube videos.

In conclusion, these findings not only bolster the existing body of knowledge on the influence of names on social and digital behavior but also provide a light-hearted reminder of the whimsical

and often unexpected nature of statistical inquiry. This study offers a gentle invitation to embrace the delightful quirkiness that statistical analysis can reveal, underscoring the joy and intrigue that underlie research endeavors.

## 6. Conclusion

In light of the remarkably robust correlation coefficient and the compelling statistical significance, the findings of this study cannot be dismissed as mere happenstance or statistical fluke. The distinctly strong positive relationship between the frequency of the first name "Cali" and the mean viewership of SmarterEveryDay YouTube videos indeed piques scholarly curiosity and beckons a wry smile. It may be tempting to ponder whether the name "Cali" bears an ineffable charm that captivates digital denizens and leads them to partake in the intellectual offerings of the SmarterEveryDay channel. However, it is paramount to approach such suppositions with due caution, mindful of the inherently quirky and capricious nature of statistical associations.

The curiously high coefficient of determination underscores the substantial variability shared by these seemingly disparate variables, eliciting a chuckle at the whims of statistical caprice. While the p-value emphatically rejects the null hypothesis, rendering the observed correlation undeniably noteworthy, one cannot help but marvel at the capricious confluence of nomenclature and digital viewership.

As researchers, we are compelled to acknowledge the delightful and often confounding nature of statistical inquiry, wherein seemingly whimsical variables intertwine with an unexpected and undeniable grace. In the grand symphony of research, the association between the name "Cali" and the viewership of SmarterEveryDay content adds a whimsical note that cannot be ignored. Nevertheless, we assert, with the confidence of statistical rigor and the levity of scholarly whimsy, that further research in this area may not be necessary.

