Abdullah Qayyum

Data 101

11/12/2023

**Data Driven Blog**

Hello, I'm Abdullah Qayyum, a junior majoring in Computer Science. Today, I want to share an intriguing exploration into shopping patterns using a dataset I found on Kaggle. This dataset not only caught my attention but also sparked a question that I believe is worth investigating: Do women share the same passion for jeans as men?

The dataset originates from Kaggle, a platform known for hosting diverse datasets contributed by the community. It contains information about shopping transactions, detailing the items purchased and the gender of the buyers. As we delve into the world of retail through this data, our focal point is on a staple in many wardrobes – jeans. Money is a universal language, and understanding consumer behavior can be the key to financial success. Jeans, a wardrobe classic, have a special place in the hearts (and closets) of many. While it's a well-established fact that men often love their jeans, the question arises: do women share the same enthusiasm?

By analyzing this dataset, we aim to unravel the odds of females purchasing jeans. This exploration goes beyond mere statistics; it's a journey of discovery into the dynamics of consumer choices, shedding light on whether the allure of jeans is universal or gender specific. Understanding consumer preferences, especially in the context of a ubiquitous item like jeans, has implications for businesses, marketers, and even fashion enthusiasts. Uncovering the odds of females buying jeans not only satisfies our curiosity but also provides actionable insights for industries that thrive on understanding and predicting consumer behavior.



So, why should you care? Because in this quest for knowledge, we're not just crunching numbers; we're exploring a facet of human behavior that resonates with our daily lives. Join me on this journey as we unravel the mysteries behind shopping preferences and discover what the data has to say about the odds of women embracing this timeless fashion choice – jeans.

**Background:**

In our detailed exploration of shopping data, our initial observation indicates that jeans are being purchased. We aim to scrutinize and adjust our initial belief that females are the predominant purchasers of jeans.

**Prior Probability (Prior Odds):**

Our initial analysis yields a Prior Probability (Prior Odds) of approximately 0.0318, reflecting the likelihood of jeans being purchased in the entire dataset.

**Converting Prior Probability to Odds:**

Converting the Prior Probability to Odds provides a Prior Odds of around 0.0329, offering an alternative perspective on the proportional likelihood of jeans purchases.

**Exploring Likelihood Ratios:**

Interpretation of True Positive and False Positive Rates:

True Positive Rate (Proportion of Females Buying Jeans): Approximately 23.39%

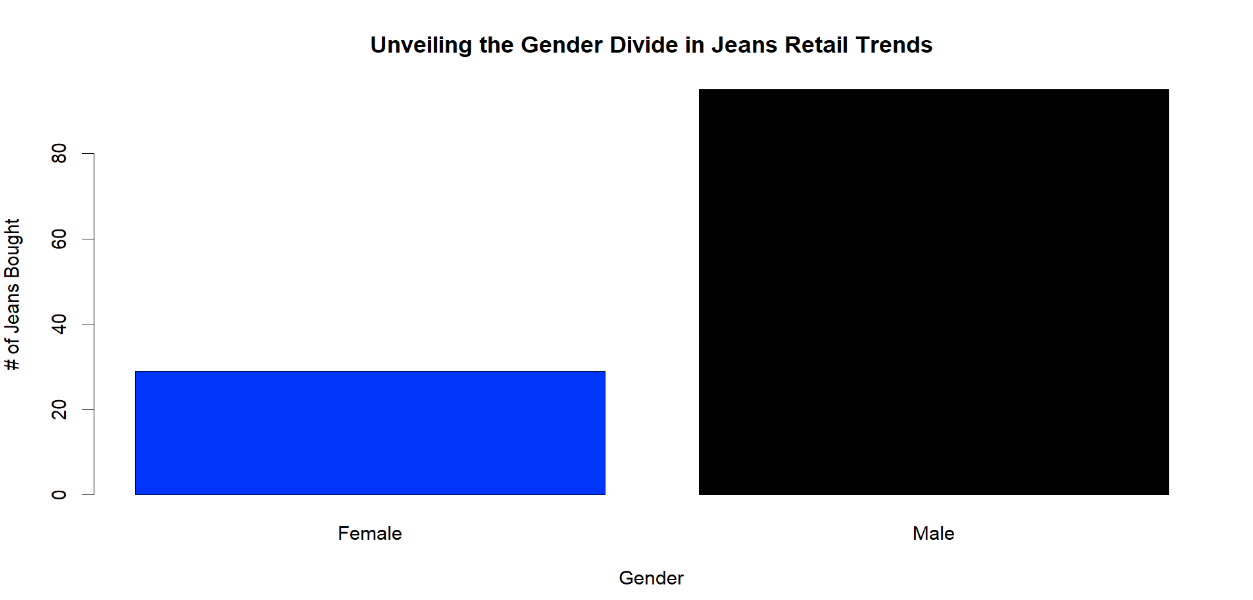
False Positive Rate (Proportion of Females Buying Other Items): Approximately 32.28%

To deepen our understanding, we delve into the Likelihood Ratio (LHR), comparing the likelihood of females buying jeans (True Positive) to the likelihood of females buying other items (False Positive). The LHR is calculated at approximately 0.666.

**Posterior Odds:**

Utilizing the Likelihood Ratio and the Prior Odds, we compute the Posterior Odds, resulting in a figure of approximately 0.0219. This represents the adjusted odds of females purchasing jeans after incorporating the observed data.

**Converting Posterior Odds back to Probability:**

****Converting the Posterior Odds back to Probability provides a Posterior Probability of approximately 0.0211. This updated probability offers a nuanced perspective on the likelihood of females buying jeans, considering the observed purchasing patterns.

**Conclusion:**

In the realm of retail analytics, our Bayesian odyssey has not merely refined but illuminated the elusive probabilities of females indulging in the timeless allure of jeans. This isn't just a statistical tale; it's a saga of insights shaping the future of decision-making. The symphony of True Positive and False Positive rates, orchestrated through Bayesian reasoning, has elevated our understanding to new heights. We've transcended the ordinary, navigating the intricate landscape of consumer behavior with a precision that resonates with the heartbeat of retail analytics. But this is not the end; it's a prelude to further exploration, an invitation to dive deeper into the complexities of context. As we venture forward, the tapestry of insights unravels, revealing possibilities that go beyond the ordinary. This isn't just about numbers; it's about unlocking the secrets that drive consumer choices and steering the course of retail strategy.

