Two-proportion statistical test summary

For this project, we wanted to look at how athletes with a history of concussion differ from those without when it comes to sleep and attention issues — and whether those differences can help us better understand the risks tied to contact and collision sports. we split the data into concussion and non-concussion groups and ran two sample proportion tests to see which specific symptoms stood out the most. The idea was to see if there were significant differences in how certain features were reported between the concussion and non-concussion groups — differences that could then help us classify sports based on the experiences of athletes within those significantly impacted categories.

We found that using a 0.05 significance threshold, several sleep-related features emerged as significantly different, including response levels for Bad\_Dreams, Cant\_Sleep, Loud\_Snore, Sleep\_Meds, Sleep\_Quality, Staying\_Awake\_Issues, Wake\_In\_Night, and Wake\_To\_Bathroom. At the stricter 0.01 level, Bad\_Dreams, Sleep\_Meds, and Staying\_Awake\_Issues remained significant. Attention-related features also showed clear distinctions, particularly in responses to Concentration\_Issues, Good\_Interruption\_Recovery, Good\_Task\_Alteration, Good\_Task\_Switching, and Poor\_Listening\_Writing.

We're now using these significant features to help classify sports as either contact or collision. Instead of just relying on rulebooks or assumptions, we’re looking at how athletes in different sports experience these symptoms. This is a new way of clustering sports into collision and contact, with the use of survey data, as a more data-driven way to classify these sports.