**Investigating the influence of Age, Bodyweight and Strength metrics on powerlifting performance**

**ABSTRACT**

Powerlifting is a body-strengthening exercise that relies on various physiological and biomechanical factors, including age, body weight, and strength metrics. Further investigation into the relationship between these factors and powerlifting performance is essential for refining training approaches and performance predictions. This study employs regression-based models to examine the influence of these factors on powerlifting performance. The dataset includes 506,544 records encompassing lifter demographics, strength metrics, and competition results. Ordinary Least Squares (OLS) and Generalized Additive Models (GAM) were utilized to assess the relationship between age, body weight, and strength metrics in predicting total lifted weight. Model performance was evaluated using R², Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) to determine their effectiveness in explaining performance variability. Results indicate a strong correlation between age and body weight with total lifted weight, with heavier athletes generally lifting more. Additionally, distinct performance trends were observed between equipped and raw lifters, highlighting the impact of lifting gear on competitive outcomes. These findings provide valuable insights for athletes and coaches to refine training methodologies while contributing to a deeper understanding of factors influencing powerlifting performance.

**Keywords:** Powerlifting, strength metrics, age influence, body weight, regression.

**LITERATURE REVIEW**

Numerous studies have examined different aspects of powerlifting performance, ranging from tapering practices and strength prediction to scoring formulas and age-related decline. Together, these studies form a foundation for understanding how age, bodyweight, and strength metrics contribute to lifting outcomes.

Travis et al. (2021) explored tapering practices among North American raw powerlifters. Their findings indicated that most athletes implemented tapering protocols before competition, primarily using step and linear methods. Although not directly predictive, tapering contributes to peak performance on competition day and highlights the need to account for training strategies in performance assessments.

Medina-Romero et al. (2023) employed machine learning models to predict powerlifting results, emphasizing the importance of **age, bodyweight, equipment type, and sex** as key features. They found that separating raw and equipped lifting improved model accuracy, supporting the current study’s decision to include **equipment type** as a predictor variable.

The use of scoring systems such as Wilks and IPF formulas to normalize performance across bodyweight categories has been critically evaluated. Ferland et al. (2020) compared the efficiency of these formulas and found that the Wilks formula was marginally more effective across divisions, while the IPF formula performed better for female lifters. However, both formulas were limited by outdated coefficients and assumptions, reinforcing the relevance of using more contemporary and flexible metrics like the **Dots score**.

Several machine learning-based studies have directly modeled the relationship between **age, bodyweight, and strength outcomes**. Chau et al. (2019) applied artificial intelligence techniques to identify optimal performance ages between 20 and 40, with a noted decline after 40. Similar results were confirmed in a female-focused study by the same authors using Extreme Learning Machines (ELM), which demonstrated that higher bodyweight correlated with improved squat performance.

Khudayarov (2019) further contributed by proposing a revised scoring system using 4th order polynomial regression to more accurately adjust raw powerlifting scores. This work highlights the limitations of traditional formulas and emphasizes the need for updated, data-driven models in the sport.

Additional studies have explored the physiological aspects of performance. Tromaras et al. (2024) linked gains in lean body mass (LBM) with improvements in squat and deadlift, while Anton et al. (2003) examined age-related declines in muscular performance. These findings underscore the biological relevance of age and body composition in strength outcomes.

Finally, Ugalde (2022) conducted a qualitative review on annual progress in powerlifting disciplines, finding that younger athletes tend to improve more rapidly year over year and that classic (raw) lifters often exhibit different performance patterns than equipped lifters. This aligns with the current study’s investigation of performance variability across equipment types and age groups.

Together, these studies provide a comprehensive backdrop for the present research. However, few have integrated these insights into transparent, predictive regression models using large-scale, real-world data. This dissertation seeks to fill that gap by building interpretable and accurate models capable of predicting total powerlifting performance based on measurable demographic and strength variables.