STATS6170-Statistical-Report

Aditya Sagave

2023-06-01

Abstract

This statistical report examines the relationship between gender, age, body weight, and weight lifted among weightlifters. The research questions address whether there is a difference in the average age between female and male weightlifters and the relation between body weight and weight lifted, considering gender as a factor. The analysis involves hypothesis testing, linear regression, and diagnostic assessments. The results provide insights into the average age differences and the relationship between body weight and weight lifted among weightlifters.

Introduction

Weightlifting is a popular sport and physical activity that involves lifting heavy weights to develop strength and muscle mass. Understanding the factors that contribute to weightlifting performance can provide valuable insights for athletes, trainers, and researchers. In this statistical report, we investigate two research questions related to weightlifters: (a) Is there any difference in the average age of female and male weightlifters? and (b) What is the relation between the body weight of weightlifters and the weight lifted?

The average age of weightlifters can offer insights into the age-related performance variations and potential differences between genders. By examining the relationship between body weight and weight lifted, we can understand the impact of body composition on weightlifting ability. This information can guide training programs and provide valuable knowledge for weightlifters striving to optimize their performance.

In this study, we analyze a dataset consisting of information from 201 weightlifters, including their gender, body weight, age, and maximum weight lifted. The dataset represents a random sample, and although it is simulated rather than based on real data, it allows us to explore the research questions at hand. Through hypothesis testing and linear regression analysis, we aim to provide insights into the average age differences between female and male weightlifters and the relationship between body weight and weight lifted, accounting for gender as a potential factor.

By addressing these research questions, we hope to enhance our understanding of the factors influencing weightlifting performance and contribute to the existing body of knowledge in the field.

Methods

Experimental Design:

For this research, a cross-sectional study design was chosen to examine the relationship between gender, age, body weight, and weight lifted among weightlifters. This design allows for the collection of data from a single point in time, providing a snapshot of the variables of interest.

Subject Selection:

The dataset used in this study consists of a random sample of 201 weightlifters. The subjects were not selected based on any specific criteria but were included to represent a diverse range of weightlifters. The dataset includes both female and male weightlifters.

Variables Measured:

The following variables were recorded for each subject:

ID: Subject ID to uniquely identify each weightlifter

Gender: Categorical variable indicating the gender of the weightlifter (female or male)

Bodyweight: Numeric variable representing the weight of the weightlifter

Age: Numeric variable indicating the age of the weightlifter in years

Weightlifted: Numeric variable representing the maximum weight lifted by the weightlifter in an unspecified exercise

Data Collection and Management:

The data were collected through a simulated process and provided in an Excel file format. The dataset was imported into R using the readxl package to facilitate further analysis. Data cleaning and formatting steps were undertaken to ensure the accuracy and consistency of the dataset. Any missing or erroneous values were addressed appropriately.

Data Analysis:

The statistical analysis was performed using R, a programming language and software environment for statistical computing and graphics. The R Markdown framework was used to create a reproducible analysis document that integrates code, text, and results.

To address the research questions, several statistical methods were employed. For the comparison of average age between female and male weightlifters, a two-sample t-test was conducted. Assumptions of normality and equal variances were assessed through visual inspection of boxplots and the Levene's test.

To examine the relationship between body weight and weight lifted, considering gender as a factor, linear regression analysis was performed. Diagnostic plots, such as the residuals vs. fitted values plot and normal Q-Q plot, were examined to evaluate the assumptions of linear regression.

The statistical package "car" was utilized for conducting the Levene's test, while the "ggplot2" package was employed for data visualization, including the creation of boxplots and scatter plots.

Overall, these methods were chosen to ensure the appropriate analysis of the data and provide accurate answers to the research questions posed in this study.

Results 1 — Preliminary Data Exploration

To provide a comprehensive understanding of the dataset and its variables, we conducted preliminary data exploration. This involved generating graphs to visualize the frequency distributions of the variables, as well as comparative and bivariate relationships. In addition, we included brief numerical summaries in a table, along with accompanying comments to summarize the main features of the graph and provide insights into the data.