An Analysis on Volleyball Statistics*

Math 261A Project 1

Robert Yav

September 23, 2025

Introduction

Volleyball is a sport that requires elite athleticism, hand-eye coordination, open communication, and most importantly, the ability to combine all of these aspects into a functioning team. One fundamental concept of volleyball is the three touch play: Receive, set, and spike. There are, however, situations in which the ball will not need an additional touch to get a point. Then begs the question: Does one need to make assists to make more kills in volleyball? If so, how important is making an assist for earning points? This project examines the relationship between the number of kills in a game to the number of assists a team makes in a volleyball game.

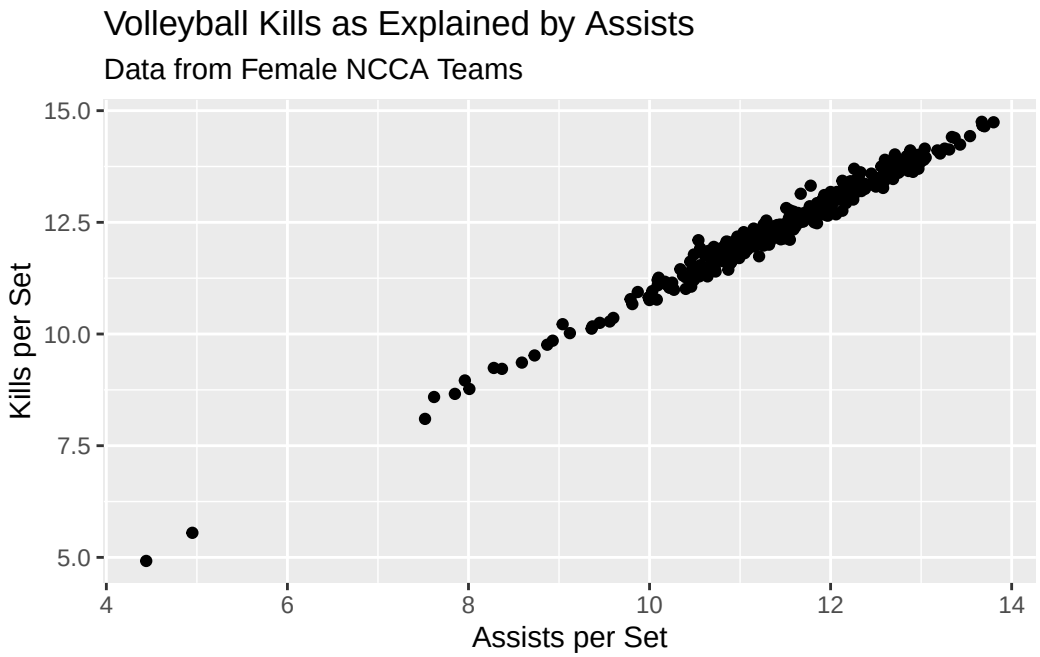
Data

The data set used was collected from the Division 1 Woman's NCAA Volleyball 2022-2023 season. There are 344 rows and 14 variables, with each row representing a team at the Division 1 level from the 2022-2023 season. The 14 variables include: Team, Conference, region, aces_per_set, assists_per_set, team_attacks_per_set, blocks_per_set, digs_per_set, hitting_pctg, kills_per_set, opp_hitting_pctg, W, L, and win_loss_pctg. My analysis focuses on two of the variables, kills_per_set and assists_per_set.

assists_per_set - The average amount of sets, passes, or digs to a teammate that directly result in a kill per set

kills_per_set - Average amount of hits that directly result in a point per set

*Project repository available at: https://github.com/peteragao/MATH261A-project-template.



A quick observation indicates no obvious outliers and a noticable positive pattern between the variables.

Methods

We fit the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

β_1 represents the expected point change in kills per set per point change in assists per set.

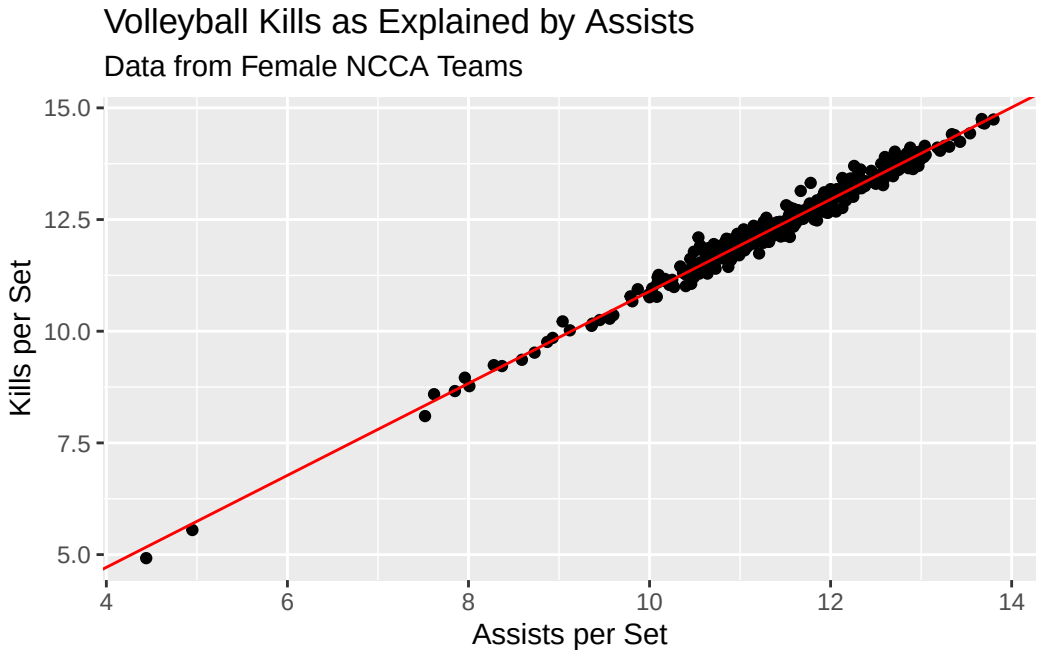
β_0 represents the intercept or expected number of kills per set if we had a team with 0 assists per set. In a sports organization with players playing at the highest caliber, players will be making assists in a game, so this statistic holds no practical relevance.

X_i represents the i^{th} team's assists as a predictor in the model.

Y_i represents the i^{th} team's kills as a response in the model.

We implement the analysis using R (R Core Team 2025)

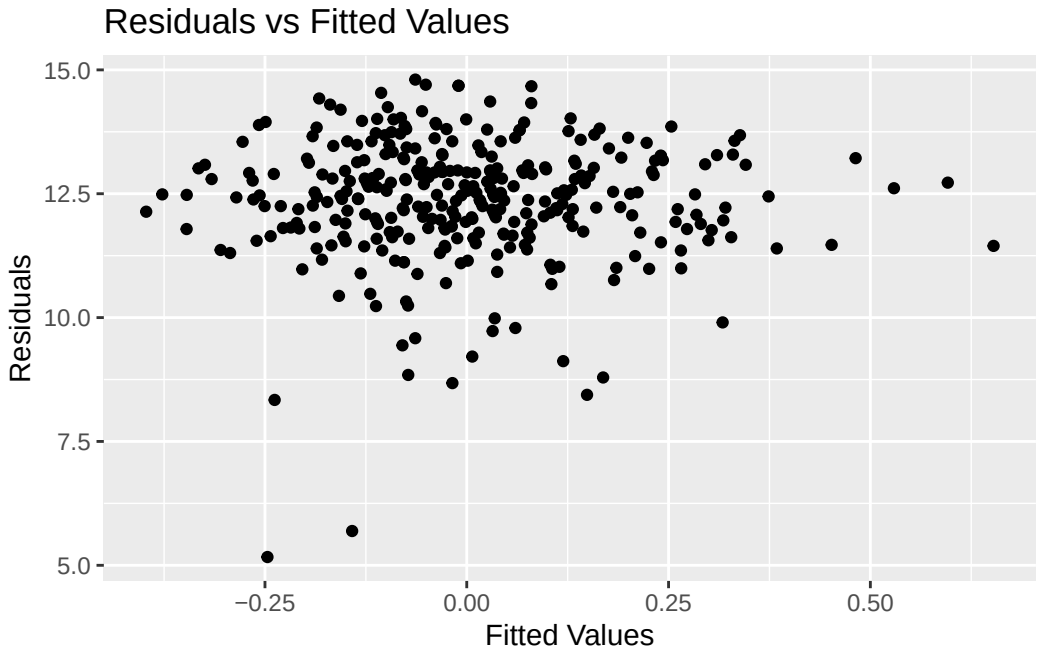
Results



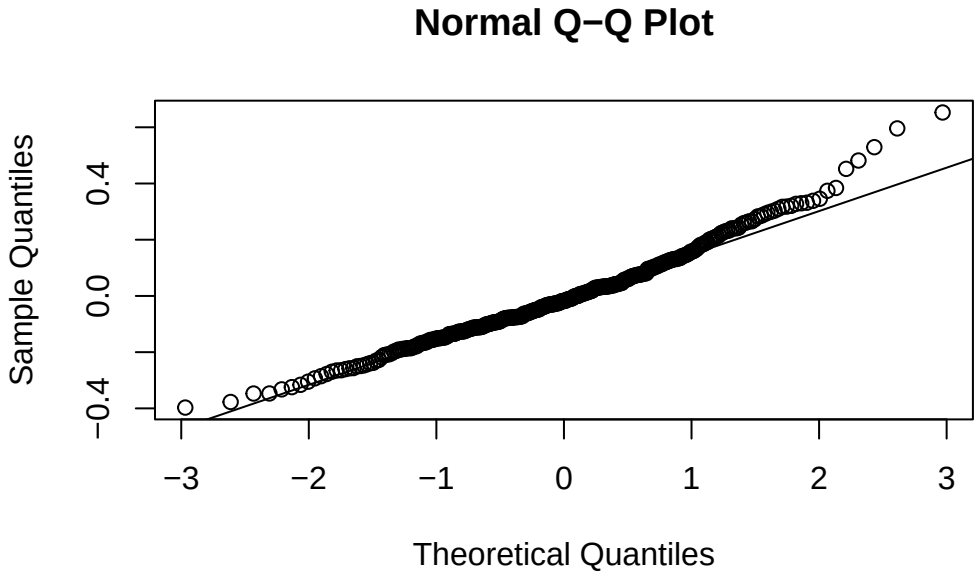
$$b_1 = 1.023$$

$$b_0 = 0.596$$

$$R^2 = .9819$$



An observation of the residuals versus the fitted values revealed



A quantile-quantile plot of the residuals revealed

References

R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.