In January of 2024, Saudi Arabia hosted the Dakar Rally, an epic off-road endurance event spanning thousands of kilometers across the rugged terrain. With drivers navigating deserts, mountains, and dunes, this race poses formidable challenges for all sorts of vehicles, such as bikes, cars, and trucks. Our focus will be on exploring the biker’s ranking based off the racers time statistics and driver information.

1. Given the data, fit a linear model that uses Hours, Variation Hours, Penalty Hours, Penalty Minutes to predict driver’s ranking after the sixth stage has been completed in the rally.

A screenshot of a computer

Description automatically generatedA table with numbers and symbols

Description automatically generated

1. Record the R-squared of the model. Show and explain work.

b. Interpret the Coefficient for Hours and Penalty Hours in the context of the problem.

c. Test whether the model is significant or not. Show/Explain work.

2. Now let make a model that predicts the probability of a driver ranking in the top five. We will be using Hours, Variation Hours, Variation Minutes, and Penalty Hours, Penalty Minutes, along with Stages.

1. Fit the model, and record R-code below.
2. How does a model where we only use Hours, Variation Hours, and Penalty Hours, along with Stages to predict top5 compared to the model in part a. Show work and Rcode.

3. Make a question that asks students to use Rcode to find specific summary output…….