Homework 4

Analysis of Variance

Purpose

Use analysis of variance to compare mean golf-ball distance across brands and golfers. Connect the research question to the model, check assumptions, and explain what a statistically significant result does and does not establish.

Part A: One-Way ANOVA

Use `ANOVA Golf Example 1.csv`. The response is `Distance`; the explanatory variable is `Brand`.

1. State the null and alternative hypotheses in words and statistical notation.
2. Summarize distance by brand using sample size, mean, and standard deviation. Create side-by-side boxplots.
3. Fit a one-way ANOVA model. Report the ANOVA table and interpret the F test at $\alpha = 0.05$.
4. If the overall test is significant, use an appropriate multiple-comparison procedure to identify which brand means differ. Report confidence intervals for the pairwise differences.
5. Check independence based on the study design, constant variance, residual normality, and unusual observations. Explain any limitation.
6. Fit the equivalent regression model with brand treated as a categorical predictor. Explain how its overall test relates to the ANOVA result.

Part B: Two-Way ANOVA

Use `ANOVA Golf Example 2.csv`. The response is `Distance`; the explanatory variables are `Golfer` and `Brand`.

1. Create an interaction plot and describe the pattern.
2. Fit a model that includes golfer, brand, and their interaction. Test the interaction first.
3. Based on the interaction result, select a defensible final model. Explain the decision.
4. Interpret the relevant F tests and, when justified, conduct follow-up comparisons.
5. Check the model assumptions using residual plots and appropriate numerical summaries.
6. Write a practical conclusion about brand, golfer, and golf-ball distance.

Submission Checklist

Submit through WebCampus (Canvas) by the deadline posted there. Submit both of the following files:

- a rendered PDF report; and
- the complete `.Rmd` source file used to produce the report.

The PDF must include your name and assignment title, answer every question in the order presented, show readable and labeled tables and figures, and explain results in complete sentences. The `.Rmd` file must contain all code needed to reproduce the report. Both files must run from the same data supplied for the assignment. Do not submit unexplained software output. You are responsible for checking all results and explaining every choice you made.

A Note About AI

AI can be used as a resource, but it cannot do the assignment for you. I do not want to read AI slop. Do not paste pages of generated writing, code, or interpretations into your report and expect me to sort through it. In the real world, you cannot send AI garbage to a boss, client, or customer. You should not submit it in this class either.

If you use AI, use it to help you learn. Check what it gives you. Make sure the code works. Internalize the material, interpret it, and then explain it in your own words and voice. You should understand every part of your submission and be able to defend the choices you made. Include a short note at the end of the report that names the tool and tells me how you used it. If the work reads like unreviewed AI output, it will receive a 0%, and you will have the opportunity to redo it under the policy in the syllabus.

How Your Work Will Be Evaluated

Credit is based on the statistical reasoning, correct use of R, appropriate checks of assumptions, quality of visual and numerical evidence, and accuracy of the interpretation. A correct number without an explanation is not a complete answer.