Final Exam Question 2

A (Science Fiction) Study of the Humanoid Inhabitants of the Moon Europa

The year is 2222. Humans have satellites in orbit around Europa, the smallest of Jupiter's four moons. Satellite images have revealed the presence of a small, humanoid race of beings living on the moon's surface. Due to their small stature, it is difficult to gather data on these beings. It also has been observed that each humanoid is accompanied by an animal companion which is much larger and is vibrantly colored (red, blue, green, or orange). The size and bright coloring of these creatures makes them much easier to identify. Because it is much easier to collect data on the animal companions, scientists are anxious to learn more about these animals and to determine whether they can use this information to learn about the humanoid inhabitants. They have enlisted you as the statistical expert on this project, and have put together a number of questions to which they need answers. They have managed to collect data on 200 distinct humanoids and their (200) companion animals, which they have compiled for you in the data set `europa_df`, which contains the following variables:

- `humanoid_height`: the height of the humanoid, measured in centimeters
- `pet_height`: the height of an animal companion, measured in centimeters
- `pet_color`: color of the pet; one of blue, red, green, or orange
- `pet_temp`: body surface temperature of the animal companion, measured in degrees Celsius

```
library(car)
library(dplyr)
library(ggplot2)
library(gmodels)
library(readr)

europa_df <- read_csv("https://marievozanne.github.io/europa_df.csv")
```

In addition to studying relationship between companion height and color, as in Problem 1, the researchers are also interested in determining whether body surface temperature is related to height.

(a) Create an appropriate plot to examine the relationship between animal companion height, temperature, and color. Based on your plot, do you think an additive model or a model with interactions would be more appropriate?

(b) Fit a multiple regression model to the data, allowing for different slopes for the different companion colors. Using appropriate plots and calculations, assess to what extent the conditions for your model are satisfied.

(c) Fit an additive multiple regression model to the data, with companion color and temperature as predictors. Conduct a hypothesis test to determine whether an interaction model is necessary, or whether an additive model is sufficient. Be sure to state your conclusions in the context of the problem.

(d) Based on your answer in (c), what is the estimated change in animal companion height for a one degree increase in animal companion temperature if the companion is blue? What is this estimated change if the animal companion is green?

(e) Based on your answer in (c), find and interpret a 95% confidence interval for the difference between mean companion height for red versus orange companions. Be sure to interpret the interval in the context of the problem and include a statement of what is meant by 95%.