# Final Exam Question 3

## 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(GGally)
library(readr)

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

**After some extensive examination of the relationships between companion color, body surface temperature, and height, the researchers have moved on to their main investigation. Recall that due to the stature of Europa's humanoid occupants, they are difficult to identify on satellites, so gathering information on them directly is challenging. Scientists familiar with the project hypothesize that attributes of the animal companions can be used estimate height for the humanoid counterparts. They have asked you to identify an appropriate model and assess whether these animal companion physical attribute are related to humanoid height.**

**(a) Make a pairs plot of the data with only the variables you intend to use in your analysis, with the response variable last.**

**(b) Assess whether the conditions are met for a linear model. Make all relevant plots and discuss them when assessing these conditions. Make transformations as necessary.**

**(c) Check for influential observations, outliers, or high leverage observations.**

**(d) Use all subsets regression (using possibly transformed variables) to identify a set of models with similar performance to model these data well. If necessary, perform this step both with and without the outliers or influential observations included.**

**(e) Obtain the model fits for all models you identified in (d) as explaining the data as well about each other, and print the model summaries.**

**(f) Summarize what your analysis has to say about the association between each of these predictors and the response, after accounting for the explanatory variables in your models. Indicate which findings are consistent across models and which depend on specifics. Make this summary as understandable as possible. After all, you are completing this analysis for collaborators to help them learn about the humanoids on Europa!**