Tanvi Agrawal
Proposal Lab

1. Falcour's project will not work for two reasons. First, his taste score is subjective and based only on his personal opinion. Since he is the only person tasting the chicken nuggets, the taste rating will reflect individual preference rather than an objective or generalizable measure. This introduces bias and does not provide independent observations.

   Second, he would likely have fewer than 100 independent observations because the number of frozen chicken nugget brands available at HEB is limited. The project requires at least 100 independent observations, and it is unlikely that there are 100 different nugget products to purchase and test.

2. Alice's project will not work for two reasons. First, the number of representatives each state has in Congress is mathematically determined by population through the apportionment process defined by the U.S. Constitution and federal law. Because the number of representatives is directly calculated from population, population and number of representatives are not independent variables. This violates the requirement that predictor variables cannot be completely dependent on the outcome variable.

   Second, coastline status will not meaningfully affect the number of representatives once population is accounted for. The number of representatives is based solely on population size, not geographic features like having an ocean coastline. Therefore, coastline status would not be a theoretically valid predictor of the outcome variable.

3.
   a. The variable condition_intake will not work as a predictor variable because it has more than two categories, and the project requires a binary predictor variable with exactly two categories.

   b. This table shows that there are 3 categories in condition_intake. Since a binary variable must have exactly two categories, this variable does not meet the project requirements.
   ```
   library(tidyverse)
   shelter =
   read_csv("C:/Users/agraw/Downloads/shelter_proposal_lab.csv")
   > table(shelter$condition_intake)

   Injured   Normal     Sick
   19         77        23
   ```

   c. This output shows that the new variable condition_binary has exactly 2 categories: "Normal" and "Sick or Injured." Therefore, this new variable meets the requirement for a binary predictor variable and will work for the project.

```
> shelter$condition_binary = ifelse(shelter$condition_intake ==
"Normal",
+                                     "Normal", "Sick or Injured")
>
> table(shelter$condition_binary)

         Normal Sick or Injured
             77              42
```

d.This output shows that age_intake has 11 unique values. The project requires numeric variables to have at least 20 unique values. Since 11 is less than 20, age_intake does not meet the requirement and therefore will not work as the numeric predictor variable for this project.

```
> table(shelter$age_intake)

 2  3  4  5  6  7  8  9 10 13 15
34 27 19 20  7  1  3  3  3  1  1
> length(table(shelter$age_intake))
[1] 11
```

e.The variable adopted_one_month will not work as the outcome variable because it is binary (0 or 1), not numeric with at least 20 unique values. The project requires one numeric outcome variable.

Instead, Cohen could consider using days_to_adoption as the outcome variable, since it is numeric and contains at least 20 unique values.