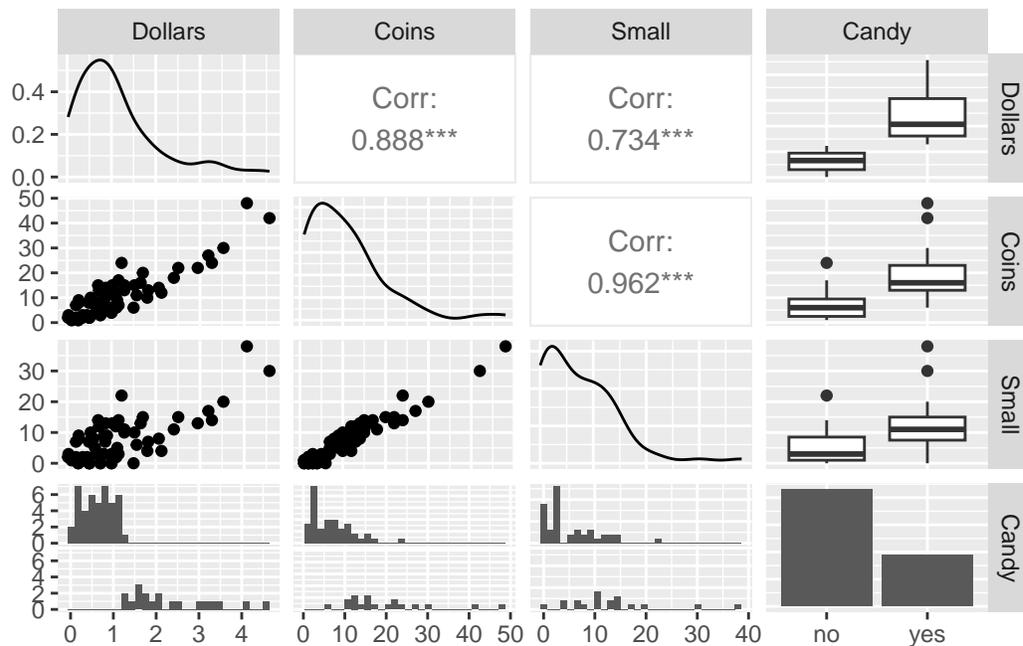# WS #12 - Multicolinearity

Math 150, Jo Hardin

Wednesday, March 4, 2026

Your Name: _____

Names of people you worked with: _____

Did you eat breakfast? Where? What?

Task: Consider the following data set collected from church offering plates in 62 consecutive Sundays. `Candy` indicates whether there was enough to buy a candy bar for \$1.25. (I made up the data.)

```
glm(Candy ~ Coins, data = Offering, family="binomial") |> tidy()
```

```
# A tibble: 2 x 5
  term        estimate std.error statistic   p.value
  <chr>          <dbl>     <dbl>     <dbl>     <dbl>
1 (Intercept)    -4.14     0.996     -4.16 0.0000321
2 Coins           0.286    0.0772     3.70 0.000213
```

```
glm(Candy ~ Small, data = Offering, family="binomial") |> tidy()
```

```
# A tibble: 2 x 5
  term        estimate std.error statistic   p.value
  <chr>          <dbl>     <dbl>     <dbl>     <dbl>
1 (Intercept)    -2.33     0.585     -3.98 0.0000693
2 Small           0.184    0.0576     3.19 0.00142
```

```
glm(Candy ~ Coins + Small, data = Offering, family="binomial") |> tidy()
```

```
# A tibble: 3 x 5
  term        estimate std.error statistic p.value
  <chr>          <dbl>     <dbl>     <dbl>   <dbl>
1 (Intercept)   -17.0       7.80     -2.18  0.0296
2 Coins           3.49      1.75      1.99  0.0461
3 Small          -3.04      1.57     -1.93  0.0531
```

1. Explain the sign (positive / negative) of the coefficients for each model (not the intercept).

2. Explain the significance (p-value) of the coefficients for each model (not the intercept).

Solution:

1. For the first two models, the coefficient is positive because the more coins you have (either in general or low coins), the more money you have (i.e., the higher the probability of being able to buy a candy bar).
   For the model that includes both variables, the coefficient on low coins is negative. Consider keeping the number of total coins fixed. If the number of total coins is fixed at 10 (for example), then an increase in the number of low coins will **lower** the amount of money (and lower the probability of being able to buy a candy bar).

2. Each of the individual variables (`Coins` and `Small`) are significant in their own right because they provide important information about the amount of money. However, when both are included, neither is (very) significant because knowing one tells you alsmost everything you need to knwo about the other.