# Review for Exam 2

# Review Problem Data

```
##    Sex Bwt Hwt
## 1    F 2.0 7.0
## 2    F 2.0 7.4
## 3    F 2.0 9.5
## 4    F 2.1 7.2
## 5    F 2.1 7.3
## 6    F 2.1 7.6
## 7    F 2.1 8.1
## 8    F 2.1 8.2
## 9    F 2.1 8.3
## 10   F 2.1 8.5
```

- The cats dataset includes sex, body weight, and heart weight data from 144 cats.
- Ref: https://vincentarelbundock.github.io/Rdatasets/doc/MASS/cats.html

# Review Problem 1

```
##
## Call:
## lm(formula = Hwt ~ Sex * Bwt, data = cats)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.7728 -1.0118 -0.1196  0.9272  4.8646
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.9813     1.8428   1.618 0.107960
## SexM         -4.1654     2.0618  -2.020 0.045258 *
## Bwt           2.6364     0.7759   3.398 0.000885 ***
## SexM:Bwt      1.6763     0.8373   2.002 0.047225 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.442 on 140 degrees of freedom
## Multiple R-squared:  0.6566,    Adjusted R-squared:  0.6493
## F-statistic: 89.24 on 3 and 140 DF,  p-value: < 2.2e-16
```

- Interpret the coefficient for sex.
- Interpret the coefficient for body weight.
- Interpret the interaction term.
- Interpret the statistical tests for both predictors and their

- We can transform body weight into a categorical variable:
- meanBwt = mean(cats$Bwt)
- cats$cBwt = as.factor(case_when(cats$Bwt < meanBwt ~ 0, cats$Bwt >= meanBwt ~ 1))

# Review Problem 2

```
##
## Call:
## lm(formula = Hwt ~ Sex * cBwt, data = cats)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6246 -1.1246 -0.2218  1.0768  7.7754
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.0190     0.2674  33.731   <2e-16 ***
## SexM          0.3060     0.3828   0.799   0.4255
## cBwt1         1.7210     0.8198   2.099   0.0376 *
## SexM:cBwt1    1.6786     0.8943   1.877   0.0626 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.733 on 140 degrees of freedom
## Multiple R-squared:  0.5041,    Adjusted R-squared:  0.4934
## F-statistic: 47.43 on 3 and 140 DF,  p-value: < 2.2e-16
```

- What is this kind of model called?
- Interpret all coefficients. What changed?
- Which model do you prefer?

# Review Problem 3

- Imagine the researcher also collected age data for the cats in this dataset.
- They found a model using age and body weight to predict heart weight provided a better fit than the model using sex and body weight to predict heart weight.
- However, they suspect the relationships between age and body weight on heart weight may resemble power curves.
- If this is true, what does it mean for the model they ran?
- How can we check if it is true?

# Review Problem 4

- The researcher hypothesizes that cats with higher body weight are more likely to be male.
- What kind of regression model can they run to test this hypothesis?
- Write down the equation for the appropriate regression model. What does the coefficient mean?
- Assume the research runs this model and finds a significant effect of body weight on sex that aligns with their hypothesis. Graph the relationship between body weight and sex.