Week 7 Section

**Contrast Coding Practice:**

**Dummy Coding**

Continuous outcome of interest: Aggression
Categorical predictor variable: Music
   Levels: Pop, Classical, Metal

Create a matrix showing how you would code contrasts if we want to compare (C1) Classical and Pop music, (C2) Classical and Metal music.

Call: ***COMPLETELY MADE-UP DATA!!***

```r
lm(formula = aggression ~ music)
```

Residuals:
   Min     1Q Median     3Q    Max
-31.84 -11.85  -0.37    8.63  43.11

Coefficients:
   Estimate Std. Error  t value    Pr(>|t|)
(Intercept)    55.60       4.23    11.20  7.4e-15 ***
Clas_Pop       1.20       1.45     .827     0.2058
Clas_Met       10.02      2.19     2.88    1.3e-05 ***

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.4 on 57 degrees of freedom
Multiple R-squared:  0.666, Adjusted R-squared:  0.547
F-statistic: 7.35 on 1 and 57 DF,  p-value: 0.0001

**In the above output, how do we interpret...**

the intercept?

Clas_Pop?

Clas_Met?

What is the regression equation for this model?

Calculate the sample means for the three groups (e.g., mean aggression for someone listening to metal):

What is the total number of participants in this experiment?

There were 20 people in the metal group; what is the standard error of the mean aggression score for the metal group?
Effect Coding

Continuous outcome of interest: Happiness
Categorical predictor variable: Season
    Levels: Spring, Summer, Fall, Winter

Create a matrix showing how you would code contrasts if we want to compare happiness in (1) spring, (2) summer, and (3) winter, to general year-round levels of happiness

Call: ***COMPLETELY MADE-UP DATA!!***
lm(formula = happiness ~ season)

Residuals:
    Min     1Q Median     3Q    Max
-31.84 -11.85  -0.37   8.63  43.11

Coefficients:                 Estimate  Std. Error   t value  Pr(>|t|)
(Intercept)              34.89       3.56    9.80    6.4e-14  ***
Seasons1                 3.74        1.38    2.71    0.0088   **
Seasons2                 1.02        2.88    1.76    0.1238
Seasons3                -2.70        1.07    2.64    0.0210   *

In the above output, how do we interpret...

...the intercept?
...the Seasons1 line?
...the Seasons2 line?
...the Seasons3 line?

What is our overall summary of this output?

What is the regression equation for this model?

What is our best prediction for someone's happiness in fall?
Polynomial Trends with Orthogonal Contrast Coding

Continuous outcome of interest: Number of friends
Categorical predictor variable: Age group
Levels: Young Adult, Middle Aged, Elderly

Create a matrix showing how you would code contrasts if we want to test (1) a linear effect of age on the number of friends and (2) a quadratic effect of age on the number of friends.

Call: ***COMPLETELY MADE-UP DATA!!***

\[
\text{lm(formula} = \text{NumFriends} \sim \text{Age})
\]

Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>-31.84</td>
<td>-11.85</td>
<td>-0.37</td>
<td>8.63</td>
<td>43.11</td>
</tr>
</tbody>
</table>

Coefficients:

|             | Estimate | Std. Error | t value | Pr(>|t|) |
|--------------|----------|------------|---------|----------|
| (Intercept)  | 27.88    | 3.56       | 9.80    | 6.4e-14 *** |
| AgeLin       | 5.74     | 2.38       | 2.74    | 0.0132 *  |
| AgeQuad      | 8.02     | 2.85       | 4.78    | 0.0078 ** |

---

In the above output, how do we interpret...

...the intercept?

...the AgeLin line?

...the AgeQuad line?

What is our overall summary of this output?

What is the regression equation for this model?

What is our best prediction for how many friends...

...a young adult has?

...a middle-aged person has?

...an elderly person has?
Week 7 Section

**Two Factors Coding**

Continuous outcome of interest: Health
Categorical predictor variable 1: Smoke
   Levels: Yes, No
Categorical predictor variable 2: Exercise
   Levels: Light, Moderate, Intense

Create a matrix showing how you would code contrasts if we want to look at (1) those who smoke vs. not, (2) linear effect of exercise, (3) a quadratic effect of exercise, and interactions between (4) 1 & 2 and (5) 1 & 3. Are these contrasts orthogonal?

Call: ***COMPLETELY MADE-UP DATA!!***
formula = health ~ smoke*exercise

Fixed effects:

| Term               | Estimate | Std. Error | t value | Pr(>|t|) |
|--------------------|----------|------------|---------|---------|
| (Intercept)        | 15.99    | 3.56       | 4.49    | 1.9e-05 *** |
| Yes_No             | 5.64     | 1.32       | 4.27    | 3.9e-05 *** |
| Exerlin            | .767     | 1.08       | .710    | 0.2403   |
| Exerquad           | 3.67     | 1.23       | 2.64    | 0.0021 ** |
| Yes_No: Exerlin    | 3.12     | 1.23       | 2.64    | 0.0072 ** |
| Yes_No: Exerquad   | 1.26     | 1.07       | 2.64    | 0.1222 *  |

**In the above output, how do we the intercept?**

Yes_No?

Exerlin

Exerquad

Yes_No:Exerlin

Yes_No:Exerquad

**What is the regression equation for this model?**

**What predicted value would we have for a smoker who does intense exercising?**