16.1 Eye fixations per line of text for poor, average, and good readers:

a. Design matrix, using only the first subject in each group:

\[
X = \begin{bmatrix}
1 & 0 \\
0 & 1 \\
-1 & -1
\end{bmatrix}
\]

b. Computer exercise:

\[
R^2 = .608 \quad SS_{reg} = 57.7333 \quad SS_{residual} = 37.2000
\]

c. Analysis of variance:

\[
\bar{X}_1 = 8.2000 \quad \bar{X}_2 = 5.6 \quad \bar{X}_3 = 3.4 \quad \bar{X} = 5.733
\]

\[
n_1 = 5 \quad n_2 = 5 \quad n_3 = 5 \quad N = 15 \quad \Sigma X = 86 \quad \Sigma X^2 = 588
\]

\[
SS_{total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 588 - \frac{86^2}{15} = 94.933
\]

\[
SS_{group} = n\Sigma (\bar{X}_j - \bar{X})^2 = 5[(8.2000 - 5.733)^2 + (5.6 - 5.733)^2 + (3.4 - 5.733)^2]
\]

\[
= 57.733
\]

\[
SS_{error} = SS_{total} - SS_{group} = 94.933 - 57.733 = 37.200
\]

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>57.733</td>
<td>28.867</td>
<td>9.312*</td>
</tr>
<tr>
<td>Error</td>
<td>12</td>
<td>37.200</td>
<td>3.100</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>94.933</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* \( p < .05 \quad [F_{0.05(2,12)} = 3.89] \)
16.3 Data from Exercise 16.1, modified to make unequal ns:

\[ R^2 = .624 \quad SS_{reg} = 79.0095 \quad SS_{residual} = 47.6571 \]

Analysis of variance:

\[ \bar{X}_1 = 8.2000 \quad \bar{X}_2 = 5.8571 \quad \bar{X}_3 = 3.3333 \quad \bar{X}. = 5.7968 \]

\[ n_1 = 5 \quad n_2 = 7 \quad n_3 = 9 \quad N = 21 \quad \Sigma X = 112 \quad \Sigma X^2 = 724 \]

\[ SS_{total} = \sum X^2 - \frac{(\sum X)^2}{N} = 724 - \frac{112^2}{21} = 126.6666 \]

\[ SS_{group} = \sum n_j ( \bar{X}_j - \bar{X}. )^2 = 5[(8.2000 - 5.7968)^2 + 7(5.8571 - 5.7968)^2 + 9(3.3333 - 5.7968)^2] \]
\[ = 79.0095 \]

\[ SS_{error} = SS_{total} - SS_{group} = 126.6666 - 79.0095 = 47.6571 \]

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>2</td>
<td>79.0095</td>
<td>39.5048</td>
<td>14.92*</td>
</tr>
<tr>
<td>Error</td>
<td>18</td>
<td>47.6571</td>
<td>2.6476</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>126.6666</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* \( p < .05 \) \quad \lbrack F_{.05(2, 18)} = 3.55 \rbrack

16.5 Relationship between Gender, SES, and Locus of Control:

\textbf{a. Analysis of Variance:}

<table>
<thead>
<tr>
<th>SES</th>
<th>Low</th>
<th>Average</th>
<th>High</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>12.25</td>
<td>14.25</td>
<td>17.25</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>8.25</td>
<td>12.25</td>
<td>16.25</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>10.25</td>
<td>13.25</td>
<td>16.75</td>
</tr>
</tbody>
</table>

\[ \Sigma X = 644 \quad \Sigma X^2 = 9418 \quad n = 8 \quad N = 48 \]
\[ SS_{total} = \Sigma X^2 - \frac{(\Sigma X)^2}{N} = 9418 - \frac{644^2}{48} = 777.6667 \]

\[ SS_{gender} = sn\Sigma (\bar{X}_{i} - \bar{X}_{..})^2 = 3(8)[((14.583 - 13.417)^2 + (12.250 - 13.417)^2] \\
= 65.333 \]

\[ SS_{SES} = gn\Sigma (\bar{X}_{j} - \bar{X}_{..})^2 = 2(8)[((10.25 - 13.417)^2 + (13.25 - 13.417)^2 + (16.75 - 13.417)^2] \\
= 338.6667 \]

\[ SS_{cells} = n\Sigma (\bar{X}_{ij} - \bar{X}_{..})^2 = 8((12.25 - 13.417)^2 + ... + (16.25 - 13.417)^2] = 422.6667 \]

\[ SS_{GS} = SS_{cells} - SS_{gender} - SS_{SES} = 422.6667 - 65.333 - 338.6667 = 18.6667 \]

\[ SS_{error} = SS_{total} - SS_{cells} = 777.6667 - 422.6667 = 355.0000 \]

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>65.333</td>
<td>65.333</td>
<td>7.730*</td>
</tr>
<tr>
<td>SES</td>
<td>2</td>
<td>338.667</td>
<td>169.333</td>
<td>20.034*</td>
</tr>
<tr>
<td>G x S</td>
<td>2</td>
<td>18.667</td>
<td>9.333</td>
<td>1.104</td>
</tr>
<tr>
<td>Error</td>
<td>42</td>
<td>355.000</td>
<td>8.452</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47</td>
<td>777.667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\*p < .05 \[ F_{.05(1, 42)} = 4.08; F_{.05(2, 42)} = 3.23 \]

b. ANOVA summary table constructed from sums of squares calculated from design matrix:

\[ SS_{G} = SS_{reg(\alpha, \beta, a\beta)} - SS_{reg(\beta, a\beta)} = 422.6667 - 357.3333 = 65.333 \]

\[ SS_{S} = SS_{reg(\alpha, \beta, a\beta)} - SS_{reg(\alpha, \beta)} = 422.6667 - 84.0000 = 338.667 \]

\[ SS_{GS} = SS_{reg(\alpha, \beta, a\beta)} - SS_{reg(\alpha, \beta)} = 422.6667 - 404.0000 = 18.667 \]

\[ SS_{total} = SS_{Y} = 777.667 \]

The summary table is exactly the same as in part a (above).

16.7 The data from Exercise 16.5 modified to make unequal ns:
SS_{error} = SS_y - SS_{reg(a,\beta,a\beta)} = 750.1951 - 458.7285 = 291.467

SS_G = SS_{reg(a,\beta,a\beta)} - SS_{reg(\beta,a\beta)} = 458.7285 - 398.7135 = 60.015

SS = SS_{reg(a,\beta,a\beta)} - SS_{reg(a,\alpha)} = 458.7285 - 112.3392 = 346.389

SS = SS_{reg(a,\beta,a\beta)} - SS_{reg(a,\beta)} = 458.7285 - 437.6338 = 21.095

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1</td>
<td>60.015</td>
<td>60.015</td>
<td>7.21*</td>
</tr>
<tr>
<td>SES</td>
<td>2</td>
<td>346.389</td>
<td>173.195</td>
<td>20.80*</td>
</tr>
<tr>
<td>G x S</td>
<td>2</td>
<td>21.095</td>
<td>10.547</td>
<td>1.27</td>
</tr>
<tr>
<td>Error</td>
<td>35</td>
<td>291.467</td>
<td>8.328</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05 \ [F_{0.05(1,35)} = 4.12; F_{0.05(2,35)} = 3.27]

16.9 Model from data in Exercise 16.5:

1.1667A_1 - 3.1667B_1 - 0.1667B_2 + 0.8333AB_{11} - 0.1667AB_{12} + 13.4167

<table>
<thead>
<tr>
<th>Means:</th>
<th>SES (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Gender (A)</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Female</td>
</tr>
</tbody>
</table>

\hat{\mu} = \bar{X}_.. = 13.4167 = b_0 = \text{intercept}

\hat{\alpha}_1 = \bar{A}_1 - \bar{X}_.. = 14.583 - 13.4167 - 1.1667 = b_1

\hat{\beta}_1 = \bar{B}_1 - \bar{X}_.. = 10.25 - 13.4167 = -3.1667 = b_2

\hat{\beta}_2 = \bar{B}_2 - \bar{X}_.. = 13.25 - 13.4167 = -0.1667 = b_3

\alpha\beta_{11} = \bar{AB}_{11} - \bar{A}_1 - \bar{B}_1 + \bar{X}_.. = 12.25 - 14.583 - 10.25 + 13.1467 = 0.8337 = b_4

\alpha\beta_{12} = \bar{AB}_{12} - \bar{A}_1 - \bar{B}_2 + \bar{X}_.. = 14.25 - 14.583 - 13.25 + 13.1467 = -0.1667 = b_5

16.11 Does Method III really deal with unweighted means?

<table>
<thead>
<tr>
<th>Means:</th>
<th>B_1</th>
<th>B_2</th>
<th>weighted</th>
<th>unweighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>A_1</td>
<td>4</td>
<td>10</td>
<td>8.5</td>
<td>7.0</td>
</tr>
<tr>
<td>A_2</td>
<td>10</td>
<td>4</td>
<td>8.0</td>
<td>7.0</td>
</tr>
</tbody>
</table>

The full model produced by Method 1: \hat{Y} = 0.0A_1 + 0.0B_1 - 3.0AB_{11} + 7.0
Effects calculated on weighted means:
\[ \bar{\mu} = \bar{X} = 8.29 = b_0 \neq \text{intercept} \]
\[ \hat{\alpha}_1 = \bar{A} - \bar{X} = 8.50 - 8.29 = 0.21 \neq b_1 \]
\[ \hat{\beta}_1 = \bar{B} - \bar{X} = 8.00 - 8.29 = 0.29 \neq b_2 \]
\[ \alpha\beta_{11} = \bar{AB}_{11} - \bar{A} - \bar{B} + \bar{X} = 4.00 - 8.50 - 8.00 + 8.29 = -4.21 \neq b_3 \]

Effects calculated on unweighted means:
\[ \bar{\mu} = \bar{X} = 7.00 = b_0 \neq \text{intercept} \]
\[ \hat{\alpha}_1 = \bar{A} - \bar{X} = 7.00 - 7.00 = 0.00 = b_1 \]
\[ \hat{\beta}_1 = \bar{B} - \bar{X} = 7.00 - 7.00 = 0.00 = b_2 \]
\[ \alpha\beta_{11} = \bar{AB}_{11} - \bar{A} - \bar{B} + \bar{X} = 4.00 - 7.00 - 7.00 + 7.00 = -3.00 = b_3 \]

These coefficients found by the model clearly reflect the effects computed on unweighted means. Alternately, carrying out the complete analysis leads to \( SS_A = SS_B = 0.00 \), again reflecting equality of unweighted means.

16.13 Venn diagram representing the sums of squares in Exercise 16.7:

![Venn Diagram](image)

16.15 Energy consumption of families:

a. Design matrix, using only the first entry in each group for illustration purposes:

\[
X = \begin{bmatrix}
1 & 0 & 58 & 75 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 1 & 60 & 70 \\
\vdots & \vdots & \vdots & \vdots \\
-1 & -1 & 75 & 80 \\
\end{bmatrix}
\]

b. Analysis of covariance:

\[ SS_{\text{reg}(\alpha,\text{cov},\alpha)} = 2424.6202 \]
\[
\text{SS}_{\text{reg}(\alpha, \text{cov})} = 2369.2112 \\
\text{SS}_{\text{residual}} = 246.5221 = \text{SS}_{\text{error}}
\]

There is not a significant decrement in \(\text{SS}_{\text{reg}}\) and thus we can continue to assume homogeneity of regression.

\[
\text{SS}_{\text{reg}(\alpha)} = 1118.5333 \\
\text{SS}_{\text{cov}} = \text{SS}_{\text{reg}(\alpha, \text{cov})} - \text{SS}_{\text{reg}(\alpha)} = 2369.2112 - 1118.5333 = 1250.6779 \\
\text{SS}_{\text{reg}(\text{cov})} = 1716.2884 \\
\text{SS}_{\alpha} = \text{SS}_{\text{reg}(\alpha, \text{cov})} - \text{SS}_{\text{reg}(\text{cov})} = 2369.2112 - 1716.2884 = 652.9228
\]

<table>
<thead>
<tr>
<th>Source</th>
<th>(df)</th>
<th>SS</th>
<th>MS</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariate</td>
<td>1</td>
<td>1250.6779</td>
<td>1250.6779</td>
<td>55.81*</td>
</tr>
<tr>
<td>A (Group)</td>
<td>2</td>
<td>652.9228</td>
<td>326.4614</td>
<td>14.57*</td>
</tr>
<tr>
<td>Error</td>
<td>11</td>
<td>246.5221</td>
<td>22.4111</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>14</td>
<td>2615.7333</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(* p < .05 \quad [F_{.05(1, 11)} = 4.84; F_{.05(2, 11)} = 3.98]\)

### 16.17 Adjusted means for the data in Exercise 16.16:

(The order of the means may differ depending on how you code the group membership and how the software sets up its design matrix. But the numerical values should agree.)

\[
\hat{Y} = -7.9099A + 0.8786A_2 - 2.4022B_1 + 0.5667AB_{11} + 0.1311AB_{21} + 0.7260C + 6.3740
\]

\[
\hat{Y}_{11} = -7.9099(1) + 0.8786(0) - 2.4022(1) + 0.5667(1) + 0.1311(0) + 0.7260(61.3333) + 6.3740 = 41.1566
\]

\[
\hat{Y}_{12} = -7.9099(1) + 0.8786(0) - 2.4022(-1) + 0.5667(-1) + 0.1311(0) + 0.7260(61.3333) + 6.3740 = 44.8276
\]

\[
\hat{Y}_{21} = -7.9099(0) + 0.8786(1) - 2.4022(1) + 0.5667(0) + 0.1311(1) + 0.7260(61.3333) + 6.3740 = 49.5095
\]

\[
\hat{Y}_{22} = -7.9099(0) + 0.8786(1) - 2.4022(-1) + 0.5667(0) + 0.1311(-1) + 0.7260(61.3333) + 6.3740 = 54.0517
\]
\[ \hat{Y}_{31} = -7.9099(-1) + 0.8786(-1) - 2.4022(1) + 0.5667(-1) + 0.1311(-1) \\
+ 0.7260(61.3333) + 6.3740 = 54.8333 \]

\[ \hat{Y}_{32} = -7.9099(-1) + 0.8786(-1) - 2.4022(-1) + 0.5667(1) + 0.1311(1) \\
+ 0.7260(61.3333) + 6.3740 = 61.0333 \]

(We enter 61.3333 for the covariate in each case, because we want to estimate what the cell means would be if the observations in those cells were always at the mean of the covariate.)

16.19 Klemchuk, Bond, & Howell (1990)

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>15.728 (^\text{a})</td>
<td>3</td>
<td>5.243</td>
<td>8.966</td>
<td>.000</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.456</td>
<td>1</td>
<td>2.456</td>
<td>4.201</td>
<td>.048</td>
</tr>
<tr>
<td>Daycare</td>
<td>2.640</td>
<td>1</td>
<td>2.640</td>
<td>4.515</td>
<td>.041</td>
</tr>
<tr>
<td>Age</td>
<td>11.703</td>
<td>1</td>
<td>11.703</td>
<td>20.016</td>
<td>.000</td>
</tr>
<tr>
<td>Daycare * Age</td>
<td>.037</td>
<td>1</td>
<td>.037</td>
<td>.064</td>
<td>.802</td>
</tr>
<tr>
<td>Error</td>
<td>21.050</td>
<td>36</td>
<td>.585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46.111</td>
<td>40</td>
<td>.585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>36.778</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) R Squared = .428 (Adjusted R Squared = .380)

16.21 Analysis of GSIT in Mireault.dat:

Tests of Between-Subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>1216.924 (^\text{a})</td>
<td>5</td>
<td>243.385</td>
<td>2.923</td>
<td>.013</td>
</tr>
<tr>
<td>Intercept</td>
<td>1094707.516</td>
<td>1</td>
<td>1094707.516</td>
<td>13146.193</td>
<td>.000</td>
</tr>
<tr>
<td>GENDER</td>
<td>652.727</td>
<td>1</td>
<td>652.727</td>
<td>7.839</td>
<td>.005</td>
</tr>
<tr>
<td>GROUP</td>
<td>98.343</td>
<td>2</td>
<td>49.172</td>
<td>.590</td>
<td>.555</td>
</tr>
<tr>
<td>GENDER * GROUP</td>
<td>419.722</td>
<td>2</td>
<td>209.861</td>
<td>2.520</td>
<td>.082</td>
</tr>
<tr>
<td>Error</td>
<td>30727.305</td>
<td>369</td>
<td>83.272</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1475553.000</td>
<td>375</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31944.229</td>
<td>374</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) R Squared = .038 (Adjusted R Squared = .025)

Estimated Marginal Means
16.23 Analysis of variance on the covariate from Exercise 16.22.

The following is abbreviated SAS output.

```
General Linear Models Procedure

Dependent Variable: YEARCOLL

                      Sum of           Mean
                      Source  DF  Squares  Square  F Value  Pr > F
                      Model      5   13.3477645  2.6695529 2.15   0.0600
                      Error  292   363.0012288 1.2431549
                      Corrected Total  297   376.3489933

                  R-Square     C.V.    Root MSE  YEARCOLL Mean
                      0.035466  41.53258   1.11497  2.6845638

                      Source   DF   Type III SS  Mean Square  F Value  Pr > F
                      GENDER  1    5.95006299  5.95006299  2.6695529 2.15  0.0600
                      GROUP  2    0.78070431  0.39035216  0.31  0.7308
                      GENDER*GROUP  2    2.96272310  1.48136155  1.19  0.3052

                      GENDER  GROUP  YEARCOLL
                      LSMEAN
                      1     1   2.27906977
                      1     2   2.53225806
                      1     3   2.68421053
                      2     1   2.88888889
                      2     2   2.85000000
                      2     3   2.70967742
```

These data reveal a significant difference between males and females in terms of YearColl. Females are slightly ahead of males. If the first year of college is in fact more stressful than later years, this could account for some of the difference we found in Exercise 16.21.
Everitt compared two therapy groups and a control group treatment for anorexia. The groups differed significantly in posttest weight when controlling for pretest weight ($F = 8.71, p < .0001$, with the Control group weighing the least at posttest. When we examine the difference between just the two treatment groups at posttest, the $F$ does not reach significant, $F = 3.745, p = .060$, though the effect size for the difference between means (again controlling for pretest weights) is 0.62 with the Family Therapy group weighing about six pounds more than the Cognitive/Behavior Therapy group. It is difficult to know just how to interpret that result given the nonsignificant $F$.

A slope of 1.0 would mean that the treatment added a constant to people’s pretest scores, which seems somewhat unlikely. Students might try taking any of the data in the book with a pretest and posttest score and plotting the relationship.

This relationship between difference scores and the analysis of covariance would suggest that in general an analysis of covariance might be the preferred approach. The only time I might think otherwise is when the difference score is really the measure of interest.