Interaction Model

The model that contains Age, Bidders and Age*Bidders is a very good model.

\[ R^2 = 0.954, \text{ 95.4\% variation in the price of antique clocks is explained by the interaction model.} \]

Interaction Model

The interaction model also has some problems.

Cannot interpret the estimates of slope coefficients.

Age does not appear to add significantly to the model.

Interaction Model

By including the interaction term Age*Bidders we have added an explanatory variable that is clearly related to the other explanatory variables, Age and Bidders.
Multicollinearity

- When explanatory variables are correlated, this is called multicollinearity.
- Multicollinearity causes problems with interpretation and by inflating standard errors of estimates.

Correlations

<table>
<thead>
<tr>
<th>Variable</th>
<th>by Variable</th>
<th>Correlation</th>
<th>Signif Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bidders</td>
<td>Age</td>
<td>-0.2537</td>
<td>0.1611</td>
</tr>
<tr>
<td>Age*Bidders</td>
<td>Age</td>
<td>0.3635</td>
<td>0.0408</td>
</tr>
<tr>
<td>Age*Bidders</td>
<td>Bidders</td>
<td>0.7916</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Centering Variables

- By centering variables, subtracting off the mean value, the correlation between explanatory variables and the interaction term can be reduced.

Correlations

<table>
<thead>
<tr>
<th>Variable</th>
<th>by Variable</th>
<th>Correlation</th>
<th>Signif Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bidders</td>
<td>Age</td>
<td>-0.2537</td>
<td>0.1611</td>
</tr>
<tr>
<td>AgeCtr*BiddersCtr</td>
<td>Age</td>
<td>-0.0744</td>
<td>0.6859</td>
</tr>
<tr>
<td>AgeCtr*BiddersCtr</td>
<td>Bidders</td>
<td>-0.2152</td>
<td>0.2369</td>
</tr>
</tbody>
</table>
Interaction Model (centered)
- JMP will automatically center the variables, by subtracting off the sample mean for each variable, before creating the interaction term.
  - $(\text{Age} - 144.938) \times (\text{Bidders} - 9.53125)$

Interaction Model (centered)
- Predicted Price = $-1470.208 + 13.244 \times \text{Age} + 94.704 \times \text{Bidders} + 1.298 \times (\text{Age} - 144.938) \times (\text{Bidders} - 9.53125)$
  - Although the prediction equation looks different, it is equivalent to the un-centered prediction equation.

Model Utility
- $F=195.19$, $P$-value$<0.0001$
  - The small $P$-value indicates that the model using Age, Bidders and $(\text{Age} - 144.938) \times (\text{Bidders} - 9.53125)$ is useful in explaining variability in the prices of antique clocks.
Statistical Significance

- \((\text{Age} - 144.938)(\text{Bidders} - 9.53125)\) (added to Age, Bidders)
  - \(t=6.15, P\text{-value}<0.0001\)
  - \(F=37.83, P\text{-value}<0.0001\)
  - The P-value is small, therefore the interaction term \((\text{Age} - 144.938)(\text{Bidders} - 9.53125)\) adds significantly to the no interaction model.

Interaction Model (centered)

- \(R^2=0.954\) or 95.4% of the variation in price can be explained by the interaction model.
- \(\text{RMSE}=88.37\)

Interaction Model (centered)

- Number of Bidders = 5
  - Predicted Price = \(-144.296 + 7.363\times\text{Age}\)
- Number of Bidders = 10
  - Predicted Price = \(-611.346 + 13.853\times\text{Age}\)
Interaction Model (centered)

- Number of Bidders = 15
- Predicted Price = –1078.396 + 20.343*Age
- The slope estimate for Age changes as the number of bidders changes.

The interaction model is doing an even better job than the no interaction model.
- The test for Age in the centered interaction model is now statistically significant.