We'll follow through the example of predicting material loss (loss) from iron content (Fe) in the corrosion dataset from the faraway package.

1. Think in \( n \)-dimensional space.
   There are \( n \) (in this case 13) data points. Rather than think of each observational unit as a point in the plane, think of each variable as a vector in 13-dimensional space. Vector operations work the same in 13-dimensional space as they do in 2- or 3-dimensional space. So our data variables are really vectors.

2. Think of \texttt{lm} as producing two vectors – a residual vector and fitted model vector. Of course we have

   \[
   \text{observed} = \text{fitted} + \text{residual}
   \]

   and this is vector sum.

   ```
   > lrust = lm(loss ~ Fe, corrosion)
   > y = corrosion$loss
   > x = corrosion$Fe
   > yhat = predict(lrust)
   > resid = residuals(lrust)
   > output = data.frame(x, y, yhat, resid)
   > head(output)
   
   x      y yhat resid
   1  0.01 127.6 129.55 -1.94640
   2  0.48 124.0 118.26  5.74295
   3  0.71 110.8 112.73 -1.93247
   4  0.95 103.9 106.97 -3.06770
   5  1.19 101.5 101.20  0.29707
   6  0.01 130.1 129.55  0.55360
   ```

3. The sum of squares of the elements of a vector is the square of the length! Some important vectors are \( y - \bar{y} \), \( \hat{y} - \bar{y} \), and \( y - \hat{y} \).

   ```
   > ssr = sum(residuals(lrust)^2)
   > ybar = mean(y)
   > sst = sum((y - ybar)^2)
   > ssm = sum((yhat - ybar)^2)
   ```

4. Two vectors are perpendicular if their dot product is 0. The residual vector is perpendicular to the predicted vector and also to the predicted vector recentered at the mean.

   ```
   > sum(resid * yhat)
   [1] 1.8829e-13
   
   > sum(resid * (yhat - ybar))
   [1] 1.1169e-13
   ```
5. The three vectors \( y - \bar{y}, \hat{y} - \bar{y}, \) and \( y - \hat{y} \) form a right triangle!

```r
> sst
[1] 3396.6
> ssm
[1] 3293.8
> ssr
[1] 102.85
> ssm + ssr
[1] 3396.6
> anova(lrust)

Analysis of Variance Table

Response: loss

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>1</td>
<td>3294</td>
<td>3294</td>
<td>352</td>
<td>1.1e-09</td>
</tr>
<tr>
<td>Residuals</td>
<td>11</td>
<td>103</td>
<td>9</td>
<td>1.1e-09</td>
<td></td>
</tr>
</tbody>
</table>
```