The benchmark for a good season, batting-average wise, is 300. A great hitter hits 400. Most people know the statistic: Ted Williams was the last 400 hitter - 406 in 1941. Each year, excitement is rampant as a player or two toys with the magic 400 level, before usually collapsing in late-season.

But how would one gauge "the best season"? Sure, one could rank annual batting champions, and search for the highest average. That'd be easy - maybe even right.

This analysis, based on data from Retrosheet, suggests the 1920 season of George Sisler stands above all others.
Clearly, by the table, Nap Lajoie's 426 average makes this the best hitting season, followed by the great Rogers Hornsby at 424, and George Sisler and Ty Cobb, each with 420 seasons. But is this right?

Has a 400 hitter with 400 at-bats, for example, had a better season than a player hitting 380 with 600 at-bats? As note above, several players have flirted with the magical "400", only to fall off as the season progresses. Is our 400 hitter above any different? These are only two variables - average and # of at-bats. Surely, there are others: on-base %, slugging %, home-runs, etc.

Let's assume, for this sake of this argument, I'm interested only in batting average. Wouldn't it be neat, before answering the question above of who's best, to look at the performance of all batters over the years? That is, plot all players' statistics.

But to make sure the data is not skewed by a player getting 10 hits in 20 at-bats, let's limit the number of at-bats: choose a random number: 50.
With average on the x-axis and at-bats along the y-axis, some interesting things show up: few players last long if they're not hitting well. If you're only hitting 100, you're not likely to get many at-bats. One can easily see how hard it is to hit 400!

To my question, the points in the upper-right become individually distinctive - and they show # of hits as the product of at-bats and average.

**The Great George Sisler**

Two of these individual points are the great George Sisler, in 1920 and 1922. In 1920, he had 257 hits in only 154 games! Both points, being individually distinguishable and upper-most towards the right, tell me these are among the two greatest batting-average seasons ever.

But are they the *best*?
In Search of a Quality Algorithm

One way to find out is "which point is in the 'upper-right-most' sector of the grid. One way to determine this is to establish a new grid, and see which is 'closest' to the intersection of these two axis.

Let's establish an "optimal" year as having a 430 batting average with 710 at-bats. No one has ever done both simultaneously, but both levels themselves have been approached separately. Perhaps a reasonably good "optimal" starting point.

Now, how can I establish which points are "closest"? I could calculate the distance by way of the Pythagorean Theorem - that would be pretty easy. Time-consuming as well. What if I simply draw a circle from the intersection of the two axis, and see which point I hit first - that would show me which point is closer to the optimal point than any other.

Indeed, the 1920 season of Sisler, coupling at-bats with average (producing 257 hits), was the greatest batting season of all time! What are some of the others?
**THE TEN GREATEST HITTING SEASONS**  
MAJOR LEAGUE BASEBALL: 1871 - 2007

<table>
<thead>
<tr>
<th>rank</th>
<th>year</th>
<th>Name</th>
<th>G</th>
<th>AB</th>
<th>R</th>
<th>H</th>
<th>2B</th>
<th>3B</th>
<th>HR</th>
<th>RBI</th>
<th>SB</th>
<th>CS</th>
<th>BB</th>
<th>SO</th>
<th>AVG</th>
<th>OBP</th>
<th>SLG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1920</td>
<td>Sisler</td>
<td>154</td>
<td>631</td>
<td>137</td>
<td>257</td>
<td>40</td>
<td>18</td>
<td>19</td>
<td>122</td>
<td>42</td>
<td>17</td>
<td>46</td>
<td>19</td>
<td>.407</td>
<td>.440</td>
<td>.632</td>
</tr>
<tr>
<td>2</td>
<td>1929</td>
<td>O'Doul</td>
<td>154</td>
<td>638</td>
<td>132</td>
<td>254</td>
<td>35</td>
<td>6</td>
<td>32</td>
<td>122</td>
<td>2</td>
<td>76</td>
<td>19</td>
<td>398</td>
<td>.465</td>
<td>.622</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1930</td>
<td>Terry</td>
<td>154</td>
<td>633</td>
<td>139</td>
<td>254</td>
<td>39</td>
<td>15</td>
<td>23</td>
<td>129</td>
<td>8</td>
<td>57</td>
<td>33</td>
<td>401</td>
<td>.452</td>
<td>.619</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1925</td>
<td>Simmons</td>
<td>153</td>
<td>654</td>
<td>122</td>
<td>253</td>
<td>43</td>
<td>12</td>
<td>24</td>
<td>129</td>
<td>7</td>
<td>14</td>
<td>35</td>
<td>41</td>
<td>.387</td>
<td>.419</td>
<td>.599</td>
</tr>
<tr>
<td>5</td>
<td>1930</td>
<td>Klein</td>
<td>156</td>
<td>648</td>
<td>158</td>
<td>250</td>
<td>59</td>
<td>8</td>
<td>40</td>
<td>170</td>
<td>4</td>
<td>54</td>
<td>50</td>
<td>386</td>
<td>.436</td>
<td>.687</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1922</td>
<td>Hornsby</td>
<td>154</td>
<td>623</td>
<td>141</td>
<td>250</td>
<td>46</td>
<td>14</td>
<td>42</td>
<td>152</td>
<td>17</td>
<td>12</td>
<td>65</td>
<td>50</td>
<td>.401</td>
<td>.459</td>
<td>.722</td>
</tr>
<tr>
<td>7</td>
<td>2004</td>
<td>Suzuki</td>
<td>161</td>
<td>704</td>
<td>101</td>
<td>262</td>
<td>24</td>
<td>5</td>
<td>8</td>
<td>60</td>
<td>36</td>
<td>11</td>
<td>49</td>
<td>63</td>
<td>.372</td>
<td>.414</td>
<td>.455</td>
</tr>
<tr>
<td>9</td>
<td>1930</td>
<td>Herman</td>
<td>153</td>
<td>614</td>
<td>143</td>
<td>241</td>
<td>48</td>
<td>11</td>
<td>35</td>
<td>130</td>
<td>18</td>
<td>66</td>
<td>56</td>
<td>393</td>
<td>.455</td>
<td>.678</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1977</td>
<td>Carew</td>
<td>155</td>
<td>616</td>
<td>128</td>
<td>239</td>
<td>38</td>
<td>16</td>
<td>14</td>
<td>100</td>
<td>23</td>
<td>13</td>
<td>69</td>
<td>55</td>
<td>.388</td>
<td>.449</td>
<td>.570</td>
</tr>
</tbody>
</table>

Sadly, Sisler missed the entire 1923 season with poisonous sinusitis. More sadly, despite a career average of 340 and 2,812 hits, he initially received only 34% of the votes for entry to the Hall-of-Fame (75% is required to be inducted), and it wasn't until the fourth try before he finally gained entrance to the Hall.

Surprisingly, those listed above as potential candidates for "greatest season" are nowhere on our table. Why? Those great seasons were the result of fewer at-bats than those seasons of the players in our table.

But does this mean "greatest season" is just a measure of "hits"? Partly - and why not? "Number of hits" is a good barometer of both hitting ability and season-longevity. The table above, however, is a blend of the two, as evidenced by the seventh-place ranking of Ichiro Suzuki and his phenomenal year of 2004, where he had 262 hits while batting 372!
By Decade

I've got all this data well-organized - what else can I do with it? Above, I looked at the history of baseball, lumping everything together. How do the individual decades compare? To make sure there is an available comparison, let's superimpose each decade with the totals above.

Notice the 60s: this was the "pitcher's decade", where Yastrzemski won the triple crown in 1968 with a meager 301 average. The 1920s, on the other hand, was a hitter's decade: compare the 20s with the 60s and you can easily see the number of "great hitters". You can also start to see the introduction of the 162-game schedule in the 60s.

Finally, the similarity of the decades, despite flickers of change, are remarkably similar.

Let's continue!
The strike-shortened seasons of 1981 and 1994 are obvious, as is the continuation of the self-similar nature of performance over more than a century!

Looking further at the "top-10" list, three of the ten are from 1930! What a spectacular hitting year that was - as is demonstrated in the graphic above. Compare 1930 with 1968, the year of the pitcher, to see remarkable differences.

Of course, there's many ways to describe "best hitting" season - on base %, slugging percentage, batting average, etc. In future editions of "Sports Forensics", I'll present an interesting graph of all three, which have Baby Ruth and Barry Bonds head-and-shoulders above the crowd.