Cincinnati Reds Runs Created Analysis

Start with the Excel workbook (spreadsheet) Cincinnati Reds Runs Created Analysis - start.xlsx.

Start this problem by opening up the starting file. Even before opening this file, you may notice that it is rather large. In fact, this dataset consists of more than 27,000 rows of information collected from Sean Lahman from his website http://www.seanlahman.com/baseball-archive/statistics/.

The scenario for this problem is that you have been asked by the Cincinnati Reds, a Major League Baseball team in Cincinnati Ohio, to analyze the provided data in order to determine which player generated the most Runs Created during their career while playing for the Cincinnati Reds. If you are a fan of statistical analysis in baseball, you may be familiar with this statistics (i.e. Runs Created). However, you may not. Runs Created was an innovative statistical attribute created by Bill James, who is well-known in statistical analysis for baseball. Runs created is intended to quantify a player’s contribution to runs scored and is calculated from common offensive statistics. The Runs created formula is shown below:

\[
\text{Total Runs Created} = \frac{\text{Total Bases} \times (\text{Total Hits} + \text{Total Walks})}{\text{Total Plate Appearances}}
\]

In order to implement this formula, you will need several attributes that appear in the original dataset. These attributes are described below:

- **AB** (Number of At Bats or Plate Appearances): Represents the total number of times a player has had a chance to hit during a particular Season – Found in Column G
- **H** (Number of Hits): Represents the total number of hits a player accumulated for all of his plate appearances in a particular Season – Found in Column I
- **2B** (Number of Doubles): Represents the total number of doubles that a player accumulated for all of his plate appearances in a particular Season – Found in Column J
- **3B** (Number of Triples): Represents the total number of triples that a player accumulated for all of his plate appearances in a particular Season – Found in Column K
- **HR** (Number of Home Runs): Represents the total number of home runs that a player accumulated for all of his plate appearances in a particular Season – Found in Column L
- **BB** (Number of Base on Balls or Walks): Represents the total number of walks that a player accumulated for all of his plate appearances in a particular Season – Found in Column K
- **IBB** (Number of Intentional Base on Balls or Intentional Walks): Represents the total number of intentional walks that a player accumulated for all of his plate appearances in a particular Season – Found in Column R
- **HBP** (Number of Hit By Pitches or Unintentional Walks): Represents the total number of unintentional walks that a player accumulated for all of his plate appearances in a particular Season – Found in Column S
In order to calculate a player’s total Runs Created for his playing career with playing for the Cincinnati Reds, it is advantageous to utilize a Pivot Table. In addition to this Pivot Table, it is also advantageous to utilize customized Calculated Fields within the Pivot Table. This reduces the overall complexity of performing the analysis, and reduces the file size. In order to calculate the statistic correctly, you will need to break down the formula that is shown above. Follow the steps below to perform the desired analyses.

1. Start by creating a Pivot Table based on all of the information presented in the range A1:V27191 on the MBL NL Batting 1963 – 2012 sheet.
2. Once created, rename the newly created sheet to Runs Created Analysis.
3. Place the First Last Name field as a Row Label. At this time, you may want to provide some formatting to your Pivot Table. I would recommend a red colored theme called Pivot Style Med 3, but any style will be suitable.
4. Place the teamID field as a Report Filter, and filter this field by CIN, which is the abbreviation for Cincinnati.
5. In order to break down the formula for Runs Created, begin by calculating Total Walks, which is a sum of BB, IBB, and HBP. Thus, add these three fields into your Values portion of the Pivot Table Report.
   a. HINT: Be careful not to place these fields into any other summary portion of your report. Also, this field should be reporting as a SUM. I do not recommend changing the name of the field at this time. Wait until your analysis is complete to avoid frustration. The order that you add these fields do not matter. I would recommend reorganizing the order of the attributes later.
6. Insert a Calculated Field into your Pivot Table and name this field Total Walks. The formula will be a SUM of the BB, IBB, and HBP.
7. Next, you will need to know the total number of At Bats to eventual implement the Runs Created formula. Place the AB field into your Values summary. This field should be reporting as a SUM.
8. Next, you will need to know the total number of Total Hits to eventual implement the Runs Created formula. Place the H field into your Values summary. This field should be reporting as a SUM.
   a. HINT: In order to move forward, it is useful to understand how Total Hits, or H, is determined. The total hits formula is shown below and it is a simple sum of all of a player’s singles, doubles, triples, and homeruns.

   \[
   \text{Total Hits} = H = 1B + 2B + 3B + HR
   \]
9. The only remaining part of the Runs Created formula is Total Bases. This attribute is somewhat complicated to calculate correctly and requires you to understand the formula for Total Hits and what singles, doubles, triples, and homeruns mean in terms of the bases that a player can advance when one of the hits is obtained. For example, a single, or 1B, allows a player to advance one base. A double, or 2B, allows a player to advance two bases. A triple, or 3B, allows
a player to advance three bases. Finally, a homerun, or HR, allows a player to advance four bases. Thus to calculate Total Bases, you will need to implement the following equation:

\[
Total \, Bases = 1 \cdot 1B + 2 \cdot 2B + 3 \cdot 3B + 4 \cdot HR
\]

However, not all of the information exists in our Pivot Table or the original data found in the starting file. Specifically, the number of singles, or 1B, is not known, but it can be calculated rather easily by a Calculated Field. Therefore, insert a Calculated Field into your Pivot Table and name this field 1B. The formula for 1B is shown below.

\[
1B = H - 2B - 3B - HR
\]

10. Once the total number of singles, or 1B, has been calculated in the Pivot Table, also add a calculated field called Total Bases. Utilize the formula at the beginning of step 9 for this calculation.

11. Insert a final Calculated Field into your Pivot Table and name this field Total Runs Created. Utilize the formula near the beginning of the problem description.

a. HINT: Do not be alarmed if you see a #DIV/0! Error in your Pivot Table. This will appear in situations where players have zero plate appearances.

12. Sort the Pivot Table based Total Runs Created in a descending fashion.

13. Use a ‘Top 10’ Filter and show the Top 15 players based on the calculated value for Total Runs Created.

14. Suppress any Grand or Sub Totals from your report and format all values with a comma style showing no decimal places.

15. Finally, a screen shot of the completed problem is shown below. Arrange your column headings based on the figure. Rename any field calculation to match the image below.