Baseball-Statistics:

Projekt „Looking for the perfect Player"

By Christoph Knoop

Links to Tableau-Storys

Link to first version:

https://public.tableau.com/profile/christoph.knop#!/vizhome/UdacityBaseball1/Baseball

Link to final version:

https://public.tableau.com/profile/christoph.knop#!/vizhome/UdacityBaseball2/Baseball2?publish=yes

Summary

With the data of 1157 players I want to find out whether there are characteristics that distinguish a top player. In other words, what are the characteristics a player must have in order to be likely to hit a high number of home runs or a high AVG?

The data shows The best chances for many home runs are players who are big and rather difficult. According to the numbers, a player achieves a high AVG when he is small and light (evenly manoeuvrable).

In both cases it helps the players if they are left-handed.

Design

Page 1 – Looking for the perfect Player

The big picture of a baseball serves as an eye-catcher, which obviously already brought some games behind itself. As font for the normal text I chose the tableau typo. The headlines are in the font

"Whoa! In my opinion, Scripture points out that this study should not be taken too seriously, that is, it has no scientific claim.

Page 2 – Our players

This page is about giving an initial overview of the data set.

For the visualization of the hand I chose a pie chart, because it shows the relation between left-, right- and ambidextrous players.

Weight and size are indicated by a line diagram. On it you can see the numerical distribution of the players. In addition, a line shows the respective average.

All figures are also briefly explained as larger word marks.

Page 3 – Homerun Statistics

On this page, the user will find information about the home run variable applied to the entire data set. A bar chart shows the user how the number of home runs is distributed among the players. I created bins to make the graphics clearer. A line chart shows the average of successful home runs. In addition, he will receive a top ten list of the ten players who have scored the most home runs.

Page 4 – Homerun-Top-100

This is about checking out the 100 players who have scored the best in home runs. All three available variables are examined: the hand, the height, the weight.

The hand is represented by a bar which is divided into left-handed, right-handed and player using both hands. In addition, I added a filter that allows the user to switch between the top 100, the remaining players and all players overall to see differences.

The size is displayed in a line chart. Here, too, the user has the option of switching between the top 100, the remaining players and all players in total with one button in order to recognize the differences.

I have chosen a slightly different representation of the weight: Here the user can find two small line charts on top of each other and can thus determine the differences between normal players and the top 100 players.

Page 5 – AVG Statistic

As on the "Homerun Statistic" page (page 3), the viewer first gets an initial overview of the AVG data. Here the user finds a bar chart on which all Statistic players have been placed in baskets. In addition, a line chart shows the AVG average of all players. A table then shows the ten players with the highest AVG average in our record.

Page 6 – AVG Top 100

Like the top 100 home runs, we're looking at the 100 best players in AVG. Here, too, we examine all three available variables (hand, height, weight). In the hand, we place another bet on a bar chart that uses colors to show the percentage of players with each hand. Here, too, users can choose between the Top 100, the other players and all players overall through a selection box.

In terms of weight and size, this time we use two bar charts showing the weight of the players (once Top 100, once all other players) side by side. For a better overview, the values are divided into bins. The diagrams opposite allow the user to discover differences at first glance.

Page 7 – Top 10 Statistics

I will examine the trends found so far in more detail on this page - by now focusing not on the top 100, but on the top 10 of the two categories. First, we examine the size and weight of both variables in bar charts. We compare the average value of the top ten players and all other games.

We visualize the game hand with pie charts. Here we offer the user the distribution of the hand for both top 10 groups.

We also examine how each of the top 10 players in the other category fare. For this we work with a scatterplot, in addition we draw the respective average value to be able to determine whether the top ten players (each colored red) are above or below the average.

Page 8 – Conclusion

At the end of our tableau story is the conclusion: We summarize what usually makes a particularly good player - and how a player should at least statistically be if he wants to achieve the highest possible AVG or as many home runs as possible.

We do without graphics. Instead, we use a large photo. A question mark on the player's head should lead to the initial question: What makes a perfect player?

Getting feedback

I was able to win the art director of the newspaper I work for to look at my work. I incorporated his suggestions into my tableau story. The changes are documented below.

Changes after feedback

1. **Headlines:**
   The whoa font didn't do well in feedback. He seems very playful and bites himself with the rather factual structure of the tableau story.

   Solution:
   I changed the fonts of the headlines to "Trebuchet MS". Also I changed all headlines to fontsize 36. In addition, all headings are left-aligned.

2. **Headlines of the diagrams:**
   They do not reflect clearly enough exactly what can be seen on the diagrams.

   Solution:
   Instead of using keywords, I formulate questions to which the respective diagrams should give an answer. I do this on all pages.

3. **Pie chart page 2**
   The label within the diagram is unclear with B, L and R. The information in the infowindows that open when you hover over a circle is also unclear.

   Solution:
   I'm writing out the values here. The caption of the legend is also adapted accordingly. The information in the infowindow is changed.

4. **Line charts page 2**
   The guide line for showing the mean does not stand out clearly enough. In addition, the filter function for weight is unnecessary.

   Solution:
   I changed the color oft he line to red. The filter function is removed, since this page is all about the numbers of all the players.

5. **German terms**
   On many foils there are still automatically entered designations in German in the labels of the graphics. The reason might be that my tableau is based on the German attitudes.

   Solution:
   The corresponding names are translated on all slides.

6. **Bins of the home run bar**
   Bins of the home run bar chart are unclear.

   Solution:

   The bins are labeled more precisely. The labeling of the axes and the content of the infowindow are also changed.

7. **Labeling Selection tool Bar chart page 4**
   The purpose of the selection tool is not clear. Also, the axis and legend labels are confusing.

   Solution:
   The label of the selection tool is changed to "Choose player class", the Y-axis is labeled with "Percentage share". The legend is labeled "playing hand".

8. **Line Diagram Size Page 4**

   The meaning of the selection tool is not clear from the labeling. The legend is also not clearly labeled. The position of the selection tool does not fit into the layout, the Y-axis is not exactly labeled.

   Solution:
   The label of the selection tool is changed to "Choose player class", the legend to "Player class". In addition, the legend and selection tool change positions. The Y-axis is renamed "Percentage share".

9. **The two line charts Weight, page 4**

   The label of the Y-axis projects above the visible edge. In addition, the labeling of the X-axis creates an unstable image.

   Solution:
   The label of the Y-axis is changed to "Share in %". The label of the X axis is deleted.

10. **AVG Statistic Bar Chart Page 5**

    The title of the bar chart is not very meaningful. The labeling of the Y-axis is not informative. The labeling of the X-axis is unnecessary, the title makes clear which values are involved. The values on the X axis are not visible.

11. **No explanatory text on page 5**

    On page 5 a text summarizing the results of AVG graphics was missing.

    Solution:
    The text was added later.

12. **Bar chart AVG Top 100 page 6**

    The label of the selection tool does not make clear what can be done with the tool. The labeling of the Y-axis is not appropriate, this also applies to the caption of the legend. As with the other graphics, the title is not very meaningful.

    Solution:
    The heading is transformed into a question to which the graphic should give an answer.

13. **Infowindows of the bar charts not correct**

    Within the Infowindows of the bar charts the values are partly insufficient or confusingly labeled.

    Solution:

Labels have been changed, some values have been shown or hidden.

14. **Headings of the two scatterplots inaccurate**

The headings of the two plots are not meaningful. In addition, the mean value does not stand out sufficiently.

Solution:

In this case, the line is colored yellow, since red is already used as a colored marker for the top 10 players. The Title oft he two plots was changend. In addition, the name of the individual player is shown in the info window.

Sources:

https://en.wikipedia.org/wiki/Home_run

https://en.wikipedia.org/wiki/Baseball

https://en.wikipedia.org/wiki/Batting_average

http://m.mlb.com/glossary/standard-stats/batting-average

https://www.tableau.com/support/help

Picture Baseball:

https://en.wikipedia.org/wiki/Baseball#/media/File:A_worn-out_baseball.JPG

Picture Baseball-Player:

http://www.photosforclass.com/download/pixabay-1045263?webUrl=https%3A%2F%2Fpixabay.com%2Fget%2Fe835b50a2af2023ed1584d05fb1d4e97e07ee3d21cac104497f6c27aa3edb4bd_960.jpg&user=skeeze