Baseball Salary Analysis

Summary

My Tableau story analyzes the trends in MLB salaries from 1985 to 2016 where I explain the potential relationships between salary and various metrics, and highlight the interesting data points.

During the measurement period, the average salary among MLB players increased 834%, and most of that is driven by the LA Angels and NY Yankees who outspend all of the other teams.

There are too many potential relationships to review in one story, but there are three key measurements I compared to average salary: homeruns, hits, and appearances (games played) - the scatterplots presented suggest a positive relationship between homeruns and average salary as well as between hits and average salary.

When analyzing average salary by the player's home country, we can see that Japan has the highest average salary; the player(s) from Japan have an average salary more than double that of players from the United States.

Design

I used several types of visualizations throughout my Tableau story.

Line Chart
I used a line chart to open my story and visualize the average salary of MLB players over time. This simple method is usually best for displaying information as a series of linear data points.

Tree Maps
Tree maps are good at representing quantities by category via area size, and I utilized two tree maps in my story. The first displays average salary by team, and the second visualizes average salary by position.

Box-and-Whisker Plot
Box-and-whisker plots are a great way to visualize the distribution of data, and to identify any potential outliers (i.e. data points that lie above the upper quartile or below the lower quartile). I used one of these charts to display the average salary for team, broken out by position.

Horizontal Bar (Animated)
I wanted to display what players had the highest salary year-over-year, so I used a horizontal bar chart - a

good tool for displaying discrete, numerical, comparisons. Also, I filtered to only display the top 25 players (by average salary), and paged by year to create an animated visualization that loops though 20 years of salary data in a way that's able to be digested.

Scatterplots

Scatterplots are a very effective way to visualize the relationships between two measurements. I utilized three scatterplots in my story to display the relationships between average salary and: homeruns, hits, and appearances. I also decided to use clustering in my scatterplots which was the most transformative design choice when attempting to explain the relationships outlined in those sections. Clustering those data points, along with including the player's position in my data point callouts (see feedback) really helped tell the story more effectively.

Bubble Maps

Needing a way to visualize geographical data, I turned to bubble maps. This visualization let me display circles over a map with the area of the circle proportional to the value in the dataset. I utilized this visualization twice - one bubble map to display average salary based on player's home country; and another, specifically for the United States, to display average salary based on player's home state

Design Updates Based on Feedback

Based on feedback received, I decided to add the player's position to my data point callouts in my scatterplots. This helped explain the relationships between those measurements, and provide context to some of the unexpected data points. I also chose to switch from analyzing the total (sum) salary for each category to using the average salary. This was a suggestion given as feedback, and it helps explain the data points that were previously washed out by the fact that MLB teams have more pitchers on the roster than any other position.

Feedback

The feedback I received on the first version of my Tableau story came from two people. The first suggested I look at average salary instead of the total (sum) salary. The second recommended I add the player's position to my scatterplots to help explain the data points I call out.

Tableau Public Story Links

Share URL | Baseball Salary Analysis
Share URL | Baseball Salary Analysis v2

Alt URL | Baseball Salary Analysis
Alt URL | Baseball Salary Analysis v2

Resources

- [Sean Lahman | Baseball Archive | Data](#)
- [Sean Lahman | Baseball Archive | README](#)
- [The Data Visualization Catalogue | Treemap](#)
- [The Data Visualization Catalogue | Box-and-Whisker Plot](#)
- [The Data Visualization Catalogue | Bar Charts](#)
- [The Data Visualization Catalogue | Bubble Map](#)