

John Ostensen

**Project Title:** Salary Cap, Payroll Inequality and Competitive Advantage in MLB

**Date:** December 2025

**Snapshot of Growth:** This semester, I developed my skills in framing an analytical question, selecting and justifying a method, interpreting results in practical terms, and communicating findings. A change in my approach from the beginning of the semester was accepting the challenges that may come with the results of my initial question, and how to form a question where the results are open for interpretation, but also clearly state the findings of the question being asked.

**Skills & Evidence:**

1. Regression Analysis: In CP 6 and CP7, I developed my skills in running regressions. Specifically, between CP6 and CP7, I improved my data from the raw payroll number to the relative payroll number, making units easier to read and interpret in the regression.
2. Structuring Data: In CP4, and throughout the rest of the checkpoints, I improved on my data structuring and ability to create relevant charts. I was able to pull data from relevant sources and format that into Excel, where I was able to create legible charts that visually depicted the results of my analysis.
3. Recognition of Risks, Limits & Outliers: This is an important skill that was developed over the course of the semester, especially in the last few checkpoints, where we were asked to recognize outliers, which, in turn, makes an analyst think why we are seeing outliers. This is where we see risks and limitations, and in this case, these limitations could include injuries, underperformance of players, etc.

**Challenges and Fixes:** One major challenge was interpreting regression coefficients that were technically correct but practically confusing, particularly when dealing with payroll measured in

raw dollars. Initially, my conclusions felt abstract and difficult to defend. I fixed this by rethinking the variable itself and introducing relative payroll in CP7. Relative payroll compared each team's payroll to the league average, where a team's relative payroll being greater than 1.00 would mean they are above average, and below 1.00 means they were below average. Another challenge was interpreting outliers, where teams who spent more money didn't make the playoffs as much. It is easy to interpret outliers on either side (teams who didn't spend money and had more playoff appearances) by saying they overperformed, but it was more difficult to come to the conclusion that injuries and front office performance could impact playoff appearances as well.

**Data Practice:** CP4 helped with defining rows, columns, and variables, making the data easy to read and interpret using charts. By clearly defining rows, time spans, and variables early, I avoided confusion later when running regressions in CP6 and CP7. The documentation made it easy to trace decisions, validate assumptions, and confidently explain the dataset to others.

**Communication:** I think my communication has improved over the course of the semester as I have become more familiar with the structure of the Checkpoints and Memos. CP7, where I describe the change between raw payroll data and relative payroll, definitely made the biggest change in my data set, as it made the analysis accessible to coaches, executives, and league officials, not just analysts.

**Use of AI Tools:** AI was used as a tool in structuring memos and checkpoints, as well as a guide of which method would be best to use in this case of analyzing the effect of payroll and aggressive spending on MLB playoff appearances. Where AI was not used was in locating and interpreting data. I feel this was a good balance of use of AI, as I was able to structure my

memos more legibly with the assistance of AI, but any analysis that I had included was my own work. This is a good way to continue in terms of AI use in work.

**Career Readiness:** From this project, I would send two articles: CP7 and CP4 Excel. This is because CP4 depicts my ability to structure data in a legible way, while CP7 shows I am able to adjust my data and analysis from something just I can understand to something that anyone could understand.