

# Problem Set 5

adam spohn

March 2024

## 1 Web scrape data

For my web scraped data I got data from baseball reference.com. It was "standard pitching data" from 2023, I scraped team level pitching data including teams, earned runs, fielding independent pitching, and strike outs. I like sports, baseball in particular, so using data analytics for sports is really interesting to me. This will be useful for me, but not for this class because I am planning to study effects of chatGPT on employment factors for this classes final paper. Most of the sources I looked at for that needed API's or had rate-limiting issues, so I decided to scrape baseball data, which will still be useful for me for my research class where I am studying the effects of the pitch clock.

## 2 API data

For this section I am downloading a time series of unemployment data for the San Jose area. This is because I wanted to see the effects of unemployment due to chatGPT. I thought silicon valley might be somewhere more likely to be effected, so I decided to focus on there. I chose San Jose in FRED, hoping it to be a good representation. I would find it interesting if could find a real economic relationship between chatGPT and employment. I think in someways this data will be helpful, but not on its own. It would be difficult to decide on its own that these unemployment numbers were tied to chatGPT, too many other factors, so more data and controls would be necessary, though it would still be interesting to look at.