

# PS 5

Mengyang Davila

March 7, 2023

## Collect data using CSS selector

I first collected data from Wikipedia ([https://en.wikipedia.org/wiki/List\\_of\\_women\\_CEOs\\_of\\_Fortune\\_500\\_companies](https://en.wikipedia.org/wiki/List_of_women_CEOs_of_Fortune_500_companies)), which contains a list of women CEOs of Fortune 500 companies. In addition to promoting equality and diversity in leadership positions, the gender of a CEO can also be an intrinsic characteristic that potentially influences the policies and practices of public companies

## Collect data using API

Then I obtained financial data for CVS Health Corp from 1986 to 2022 using the Financial Modeling Prep website. Karen S. Lynch, a female CEO, has been in her role since February 2021. Examining the trends in financial data can provide valuable insights into a company's financial health and performance over time, as well as identify potential risks and opportunities.

To parse the data, I used the `fromJSON()` function from the `jsonlite` package, which automatically converted the JSON array into a regular R data frame. Additionally, I found Spencer Pao's video tutorial on web scraping, available at <https://www.youtube.com/watch?v=GGgNM7WanK8>, to be highly informative and helpful in getting me started.