PS5$_N$ongard

nongarak

February 2018

1 HTML Scrape

The data I chose to scrape was a list of the most popular movies of 2017 from IMDB. These data do not actually interest me at all, but after days of frustration trying to scrape things of interest (GIS job listings, National Geographic articles, etc.) I decided to go with the easiest thing possible. There are about a million guides to how to scrape IMDB with SelectorGadget, so I just followed one step by step to get this assignment over with. I eventually DO want to figure out how to get to the bottom of the National Geographic problem, but it's not working well right now.

2 API

The data I chose to scrape using the WikipediaR package were some summary statistics on the pages of Jesus Christ and Jan Tinbergen. I used a function built into the package to scrape the number of backlinks, or articles that link to those articles. Despite being two very different people, they both have exactly 250 other pages linking to them. Jan Tinbergen is of course the Dutch economist who shared the first Nobel Memorial Prize with Ragnar Frisch. He constructed the first macroeconomic model of a nation, the Netherlands, which he then extended to the US for the League of Nations. I studied at the Jan Tinbergen Department of Economics at Universiteit Utrecht in Holland two years ago, so he's a fun guy to study. The only other famous Dutch economist is Tjalling Koopmans, who is of course way more famous. This is just a bit of colourful blasphemy.