# Moneyball 2020: Predicting the MVP Odds part 1: Webscraping

# Today's goal: Acquire the data we will be working with for this project

# To achieve today's goal we will be webscraping data from wikipedia so lets get to it!

# Step one: importing the packages.

install.packages("rvest")

install.packages("xml2")

library(xml2)

library(rvest)

# Step two: Get URL containing information you want

url = 'https://en.wikipedia.org/wiki/Major\_League\_Baseball\_Most\_Valuable\_Player\_Award'

# Step three:

url\_contents = read\_html(url) # Get HTML of page

url\_contents

# Step Four: Getting the table

mvp\_table\_html = html\_node(url\_contents, xpath ='//\*[@id="mw-content-text"]/div/table[4]')

mvp\_table\_df = html\_table(mvp\_table\_html) # get table of html table

View(mvp\_table\_df)