For my final project I scrapped a webpage that had data from a 10K race that took place in Hillsboro, OR on June 2017. I chose this webpage because it had the data in a table on the page already making it easy to extract into a dataframe. There were some challenges however with trying to use lxml while using beautifulsoup. I was getting errors saying that I did not have lxml installed but I did however I did not need to use lxml to parse the data from the page I could use html.parser instead with a lot of googling. Many of the code was inspired by the internet communities in stack exchange and datacamp. I thought the cleaning process was be a lot harder than it was but I fond ways to append or strip extra characters with ease and reprint the data to see the results. Overall the dataframe was pretty clean after the cleaning process and the joining of dataframes into one whole dataframe.

Overall, I learned a lot in this final project mainly because I did not have to look up the example code in the book and found ways to do things that were different than what the book suggested. I feel my job would have been harder if the page did not have a convenient data format built into the site. I think this was a good way for me to look at other ways I can practice python and improve my knowledge in future classes and in the field.