

Week 4 Progress Report

DATA 450 Capstone

Matthew Wilcox

Week ending: 2/13/23

Time Log

Day	Time	# Hours	Task(s)	GH Commit(s)
Mon.	4-6pm	2	Preprocessing Betting Data	Link 1
Tue.	12-1am	1	Finish Scrapeing of data	Link 2
Wed.				
Thu.	3-5:30 pm	2.5	Proposal Review, Pre Processing	Link 3
Fri.				
Sat.	9:30pm-4am	6.5	Processing Data. Making team names the same	Link 4
Sun.				

Total # of hours worked this week: 12

What you accomplished (or attempted) this week:

- Combined all of the betting data csvs
- Create a key to make teams name in the same format
 - Build a script that looks through a list for all similar team names to find matching team name

Any setbacks/roadblocks you experienced:

- Was unaware that over 500 team names did not match and would need a key to make a uniform team names

What you plan to work on next:

- Finish Processing the data to make a singular data set.
- Scrape Stadium capacities

Are you on track? If not, how will you get back on track?

No I am behind. The setback of having to match the team names to make them uniform was unexpected and took significant amount of time. This significantly set me back. There is still processing of the data sets to do next week. I do feel like I overestimated several other tasks so that will be useful in making up time. Additionally I have a full week at the end where I can push stuff back if needed. I will try to just finish the processing early next week and then start the question and finish the first 4 or 5 by the end of next weekend.