

PS4_C*castiglione.tex*

joe.castiglione.jr

February 2020

1 Interest

One thing that really comes to mind when it comes to areas that I look forward to web scraping is statistics that involve my thesis. My thesis involves the correlation between strength coach salaries and on field performance of college football teams. Right now I am looking into only using data on one conference. Statistics I would need is win loss, rushing yards per game, and total yards per game. I do think the statistics would be very easy to scrape off once I find and master scraping. I am sure you have many recommendations.

2 SparkR Questions

1. **What is the class of each?** -Data Frame 2. **SELECT - If you try the previous command on df1, what happens? Why do you think that happened?** -They were very similar answers. 3. **FILTER - If you again try the previous command on df1, what happens? Why do you think that happened?** -they were very similar answers.