Web Scraping with R Outline

1. why do we need web scraping?
   1. Large amount of information (sports data example, NBA, baseball etc.)
   2. Further visualization, analysis (example, scorecard, review meta etc.)
2. Types of web scraping
   1. Regular expression
   2. Extension
   3. Programming (R, Python)
3. Techniques
   1. Table
   2. Figure
   3. Text
   4. …
   5. Dynamic dig all pages
   6. More advanced… (package function illustration)
4. Case study (basketball player analysis?)
5. Conclusion

Course introduction:

Instructor short bio:

Things to do:

1. Discuss with Prof what format to collaborate, publish the repository?
2. Discuss with Prof which case study to use
3. Ask about the books
4. Regular expression, html knowledge
5. R package

Material:

<https://www.analyticsvidhya.com/blog/2017/03/beginners-guide-on-web-scraping-in-r-using-rvest-with-hands-on-knowledge/>

<https://www.analyticsvidhya.com/blog/2015/06/regular-expression-python/>