Research Methods in Political Science I – Homework Assignment 5

Due: 9am on 11 November 2015

How to submit: Send as an email attachment

Email subject: Research Methods 1, Assignment 5

Files to submit: You have to submit 3 files

1   Regular Expressions

Describe the type of strings that conform to the following regular expressions and construct an example that is matched by the regular expression (Munzert et al. 2015: p.217). Write your answers in `hw05-YourName.pdf`.

1. `\\$[0-9]+`

2. `\\b[a-z]{1,4}\\b`

3. `.*?\\.txt$`

4. `\\d{2}/\\d{2}/\\d{4}`

2   Web Scraping

Write an R script (or R Markdown file) to scrape the information about your colleagues (including yourself and excluding the instructor). Using the answers to Homework Assignment 3, posted at http://www2.kobe-u.ac.jp/ yyanai/classes/rm1/ contents/hw03/, collect (1) your colleagues' name, (2) date when the page (HTML file) was created, (3) how many times the word "factorial" was used except when a function whose name contains "factorial" was used, and (4) how many pound signs (#) were used in the page. If specific value is missing, use NA as missing value. Submit your R script `hw05-YourName.R` (or R Markdown file `hw05-YourName.Rmd`) and the dataset in CSV format `hw05-YourName.csv`.