

Reproducible research Semana 1

JMV

March 2, 2015

Esencialmente la idea es ser ordenadito y saber que se ha hecho en cada paso para poder volver a hacerlo todas las veces que sea necesario (y que no pase que si se va el manolo de datos nadie sepa que cojones estamos haciendo)

Conceptos basicos

- Hay estudios que no son replicables de por si, por hechos unicos o por que valgan una pasta, replicar el LHC puede ser un poco jodidillo
- La idea es dar tanto el codigo como los datos para que cualquiera pueda hacer estudios independientes, para mejorar el no poder replicar ciertos estudios, por lo menos poder validar el analisis de datos
- Las bases de datos se pueden fusionar en mega bases que permite tambien otros estudios
- Proceso investigacion
 - Se usa codigo para pasar en cada paso, de datos brutos a datos analiticos, y de ahi a resultados, y despues se tienen tus codigos de presentacion
 - Al menos se deberia dar los datos analiticos, el codigo de analisis, y documentarlo
- Literate programming: Mezclar el articulo con datos, lo que se en una salida markdown por ejemplo
 - Swave usa una mezcla de \LaTeX y R
 - knitr es lo que ando usando y va medio bien

Sin duda la mejor estrategia es hacer todo via scripts que queden grabados, asi no hay problema de que haya un paso que “no sabemos” escribir