

Research Automation

Marco Antonio Ramos

Verano 2021

E-mail: marcoyel21@gmail.com

Office Hours: Por definir

Office: San Pedro 189, Col: Del Carmen

Web: github.com/marcoyel21

Class Hours: Por definir

Class Room: *online*

Descripción del curso

En este curso se aprenderán técnicas, buenas prácticas y *know how* sobre automatización, replicabilidad y colaboración en equipo en la investigación. En general, son prácticas que emanan de las ciencias computacionales y que se han comenzado a utilizar en la investigación económica debido a los beneficios que conllevan.

- La replicabilidad de las investigaciones es una buena práctica que cada vez tiene mayor importancia. La transparencia en las limpiezas de datos, estimación de modelos y obtención de resultados le brinda mayor credibilidad a la investigación además de que disciplina al investigador a ser ordenado, a documentar cada paso y a minimizar las decisiones arbitrarias durante su proceso de investigación.
- La automatización de los pasos en el proceso de investigación permiten al investigador ahorrarse el tiempo y minimizar la cantidad de errores que se cometen al darle formato al documento, a las tablas, a las ecuaciones, al citar en determinado formato, al insertar figuras, notas al pie, índices y demás utilidades estéticas. Asimismo, en conjunto con la replicabilidad, la automatización implementada de manera exitosa garantiza que se obtendrán los mismos resultados de manera automática si se corre el *script* desde cualquier computadora.
- Finalmente, la colaboración en equipo de manera remota es un área con muchísimo potencial de utilidad en el contexto de la pandemia, la globalización y de la evolución de la manera de trabajar y colaborar en la investigación tanto para la academia como para las empresas. Permite organizar de manera rigurosa el proceso de investigación, repartir las distintas tareas y crear un ambiente de colaboración virtual en el que todos los integrantes pueden trabajar sobre el mismo proyecto sin riesgo de desincronización de los avances y duplicación de tareas.

El curso está diseñado para las personas que hacen investigación, ya sean alumnos de posgrado o investigadores de alguna institución. Asimismo, el curso es de particular interés para todas aquellas personas que quieren integrar su limpieza, modelado y redacción en un solo documento; que quieren mejorar su dominio de las herramientas para la investigación econométrica; que han batallado mezclando avances en WORD o PPTX con los de STATA, R o Python; que han perdido horas valiosas en las citas y referencias; para aquellos que siguen usando recortes para mostrar tablas y ecuaciones; para aquellos que han tenido problemas de gestión colaborando en equipo de manera remota; que han tenido problemas en cuanto a la estética de su investigación o presentación; para

aquellos que gestionan grandes y pequeños equipos de investigación; y en general, para todos aquellos que quieren ahorrarse tiempo y esfuerzo en el proceso de investigación.

Objetivos del curso

1. Que ustedes puedan automatizar su ciclo de trabajo personal o en equipo con el fin de que se dejen de preocupar en los formatos, citas, matemáticas, referencias, tablas, etc y se concentren al 100% en sus investigaciones
2. Que dominen una variedad de herramientas para que realmente plasmen el potencial de la investigación que pensaron inicialmente
3. Que logren crear una investigación replicable (o bien replicar un *pape*) en equipo con su respectivo documento, limpieza, presentación, bibliografía, etc.

Requisitos

- Conocimiento básico del lenguaje de programación R

Forma de trabajo

- El curso consiste en 3 horas de contenido diarios más dos talleres el fin de semana: a) para crear un proyecto de investigación (colaborativo, automático y replicable) y; b) para crear un CV o Syllabus (automático, estético y replicable).
- El curso comprenderá 1 o 2 semanas (por definir)
- Modalidad a distancia

Contenido

Latex y Markdown

- *Operaciones matemáticas*
- *Tablas*
- *Jerarquía de textos*
- *Figuras avanzadas*

R Markdown

- *Code chunks*
- *Funcionalidades estéticas (índice, encabezados, pies de página, numeración y referencias)*
- *Utilidades en el YAML (automatización de procesos dentro de RMD)*
- *Llamar variables locales*
- *Replicabilidad y buenas prácticas*

Formatos de RMD

- *HTML*
- *PDF*
- *PPTX*

Mendeley: automatización de citas y referencias

- *El formato BIB*
- *Administrador de citas*
- *El formato CSL*
- *Integración a RMD*

Github

- *Comandos básicos en la terminal*
- *Configuración inicial*
- *Administra tus repositorios de manera local*
- *Colaboración en equipo a través de un repositorio*
- *Administración de proyectos con github (Issues, Milestones y Proyectos)*
- *Ramas, Merges, Pull requests y conflictos*

Ambientes virtuales

- *Replicabilidad*
- *Conda*
- *Docker*

Costos

El costo total del curso es de 2,000 MXN y comprenderá una o dos semanas dependiendo del *quorum*, el interés y la intensidad.

Referencias

[R Markdown](#)

[R for Data Science](#)

[Git Pocket Guide](#)