

Profesor: Héctor Bahamonde, PhD.

e: hector.bahamonde@uoh.cl

w: www.HectorBahamonde.com

Curso: OLS.

TA: Gonzalo Barriá.

INSTRUCCIONES GENERALES

Usando la base de datos que será asignada al azar y dentro de los plazos establecidos en la sección *Tareas* de uCampus, deberás:

1. **Generar una pregunta de investigación.** Un ejemplo de pregunta de investigación puede ser *Qué explica Y?* Para esto, ayúdate del *codebook* (archivo de texto que contiene el nombre de la variable y su definición).
2. **Generar una hipótesis de trabajo.** Una hipótesis de trabajo es una respuesta (una afirmación) *hipotética* a esta pregunta. Un ejemplo podría ser: *Y es explicado por X_1* . En general, aquí también explicas *por qué* X_1 explica Y.
3. **Generar una hipótesis alternativa.** Al igual que la hipótesis de trabajo, la hipótesis alternativa es una afirmación. La diferencia es que la hipótesis alternativa plantea una respuesta diferente (o *alternativa*) a la hipótesis de trabajo. Si la de trabajo respondía X_1 , la alternativa podría ser X_2 .
4. **Testear ambas hipótesis.** Para ello deberás ocupar modelos lineales. Recuerda, entonces, que deberás formular una pregunta de investigación y dos hipótesis (de trabajo y alternativa) que puedas testear. Para efectos del trabajo, no te servirá formular una hipótesis (o pregunta), si es que no tienes variables relevantes para contestar la pregunta de investigación. Entonces la recomendación es que orientes tus hipótesis y pregunta de investigación a la información que tengas disponible en la base de datos. **Recuerda que tu variable dependiente debe ser continua.**
5. **Deberás tener al menos dos especificaciones.** Es decir, al menos un modelo lineal para testear la hipótesis de trabajo y al menos otro modelo lineal para testear la hipótesis alternativa. Si sienten que otro modelo debiera ir en el trabajo, adelante. Pero sigue siendo posible tener la nota máxima teniendo el mínimo (al menos, un modelo por hipótesis).
6. **En ambos modelos, deberás tener controles.** Recuerda que los controles son otras variables que entran en la misma ecuación, y que ayudan a mantener los efectos constantes en las medias de esas variables. Para cada control deberás justificar la decisión de incluir los controles que incluiste. Esto es discursivo, y como vimos en clases, deriva de la teoría, la revisión bibliográfica, y hasta de la intuición.
7. **Presentación de resultados.** *Para qué hipótesis encuentras evidencia suficiente para confirmarla?* Para responder esta pregunta, deberás comentar (1) p-values, (2) errores estándar, (3) r^2 del modelo, (4) tamaño del efecto (betas). No olvides usar al menos una tabla y un gráfico por modelo.
8. **Post-estimación.** Usando un solo scatter plot, revisa (1) normalidad de residuos (“errores”) y (2) homo/hetero-esquedasticidad. Responde las siguientes preguntas: (a) *les parece que los errores del(los) modelo(s) cumple(n) con los supuestos de los modelos lineales?* (b) *Qué significa(n) substantivamente ese(esos) supuestos?*

9. **Presentación final en clases.** En la presentación final deberán abordar todos estos puntos. El único punto que deberán añadir extra, es la mini revisión bibliográfica. La presentación no deberá durar más de diez minutos. *Todos los integrantes del grupo deberán presentar.* La idea de presentar es que puedan recibir comentarios del resto de los compañeros. El trabajo del grupo será convencer a la audiencia que han hecho todo lo posible para poder responder adecuadamente ambas hipótesis.
10. **Productos finales:**
- (a) **Trabajo Final:** un *script* de R enviado por uCampus. No habrá un informe final. Asegúrate que el *script* corra sin problemas, que los paquetes que usas estén declarados, etc.
 - (b) **Presentación final:** un *Power Point* presentado en clases por todos los integrantes del grupo. Todos los aspectos señalados arriba deberán estar abordados en esta presentación.