

Prueba 3

Curso OPSO791, UCSH.

Lunes 29 de noviembre de 2021

Instrucciones generales

Enviar trabajo a los correos de profesor y ayudante. Debe adjuntar el archivo `.rmd`, el `.html` resultante y el archivo `.Rproj` si es que trabajó con uno. Los archivos deben ir comprimidos en una carpeta con apellido y nombre de el o la estudiante. En archivo enviado no debe incluir ninguna base de datos o data frame.

Tiene plazo para enviar prueba hasta las 23:59 del día martes 30 de noviembre. Por cada día de atraso se descuenta 1 punto (10 décimas).

Módulo 1 - materia de clase y textos

Pregunta 1.1

En base a su experiencia en el campo de la sociología. A su juicio, ¿que tan realizados se encuentran los cuatro imperativos de la ciencia propuestos por Robert Merton? Argumente

Pregunta 1.2

¿Que es la investigación reproducible?, ¿Podría ayudar a resolver el problema de la baja confianza o legitimidad que tienen las ciencias sociales? Utilice bibliografía obligatoria.

Pregunta 1.3

¿Que alternativas representan de mejor manera lo que es la ciencia abierta según lo visto en clases? Solo deje las alternativas correctas (más de una alternativa correcta).

- a) libre acceso a publicaciones científicas
- b) remover bloqueos culturales, políticos y económicos que impiden que determinados perfiles de personas puedan ingresar al mundo académico
- c) Hacer más productiva a la ciencia y con perspectiva de futuro
- d) Dentro de lo posible utilizar *big data* para hacer investigación

- e) Vincularse con actores sociales a la hora de hacer investigación
- f) Abrir o transparentar el proceso mismo de investigación
- g) Unificar la vocación científica, crítica y profesional de las ciencias sociales
- h) si se utiliza código liberarlo, dando la posibilidad de que terceros puedan reproducir la investigación realizada
- i) Un paradigma de investigación propio del mundo cuantitativo

Pregunta 1.4 (opcional)

¿Que otras alternativas para el libre acceso a las producciones científicas existen además de la opción “guerrillera” de Swartz o Elbakián? Se espera respuesta corta de no más de un párrafo.

Módulo 2 - Markdown y Zotero

Pregunta 2.1

¿Como definirías Markdown? Solo deja la alternativa seleccionada (solo una respuesta correcta).

- a) lenguaje minimalista de escritura en texto simple, con marcas de formato muy escuetas y de fácil aprendizaje
- b) lenguaje de programación para analizar datos cuantitativos, con marcas de formato muy escuetas, aunque con curva de aprendizaje empinada
- c) lenguaje de escritura en texto simple con gran control del formato final mediante sofisticadas opciones de marcas de texto
- d) interfaz de programación para articular escritura en texto simple y lenguaje R
- e) un editor de texto o lenguaje similar a word pero sin patentes y con más opciones de formato

Pregunta 2.2 (Opcional)

¿Para que sirve Zotero?, ¿Cuáles serían sus fortalezas en el contexto de la investigación reproducible? Desarrolle

Módulo 3 - carga y visualización de data en R (práctico)

Descargue y analice, según lo que se le pide más abajo, la **Encuesta de Bienestar Social (EBS) 2021**, que puede descargar del siguiente link: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-bienestar-social>.

Para responder una de las preguntas deberá recurrir a información de la base de datos Encuesta CASEN 2020, la que puede descargar acá: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-en-pandemia-2020>

Información técnica relevante sobre la encuesta

Ante la necesidad de contar con un diagnóstico más completo y acabado del bienestar de la población, que complemente la información entregada por la Encuesta Casen 2020, el Ministerio de Desarrollo Social ha aplicado la primera Encuesta de Bienestar Social (EBS) del país.

El modelo que subyace a la EBS, se centra tanto en los resultados como en las oportunidades de las que disponen las personas en 11 dimensiones del Bienestar: Ingreso, Trabajo, Vivienda, Estado de Salud, Balance vida y trabajo, Educación, Relaciones Sociales, Compromiso cívico y gobernanza, Calidad del medio ambiente, Seguridad personal, y Bienestar subjetivo.

Con el objetivo de aprovechar al máximo la información que recoge la Encuesta Casen, la EBS se diseñó e implementó en una modalidad bifásica de ésta, es decir, usando una submuestra de la muestra entrevistada en la Encuesta Casen en Pandemia 2020. Esto permite obtener un diagnóstico completo del bienestar, considerando las condiciones materiales y de calidad de vida de las personas.

Pregunta 3.1

¿Cuántas observaciones y variables tiene el microdato de la Encuesta EBS?

Pregunta 3.2

¿A cuántos hogares pertenecen las personas que respondieron a la Encuesta EBS?

Pregunta 3.3

¿Cuántas personas por región respondieron la encuesta EBS? Ordene de mayor a menor.

Pregunta 3.4

Elabore una nueva variable que se llame “region_rec” que posea dos categorías. 1) Región metropolitana y 2) Otras regiones. ¿Cuántas observaciones hay en cada grupo de regiones?

Pregunta 3.5

Agregue a la EBS las variables del módulo “O” (trabajo) de la Encuesta Casen 2020 (en pandemia). Además agregue las siguientes variables: *oficio4_08*, *oficio1_08*, *rama4*, *rama1*, *activ*, *ocup_inf*, *v13*, *ytrabajacor*, *ytotcor*, *yautcor*. Para seleccionar variables del módulo “O” se recomienda utilizar la función `starts_with()` dentro de `select`.

Para poder vincular la información de la persona entrevistada y contenida en la Base de datos de la EBS y la información de la persona en la Base de datos de Casen en Pandemia 2020, se debe utilizar la variable de identificación del hogar de Casen (*folio_casen*) y la variable de identificación de la persona en el hogar Casen (*id_persona_casen*), contenidas ambas también en la Base de datos de la EBS 2021.

De esta manera, la llave que permite un identificador único a nivel de persona es: *folio_casen* *id_persona_casen*. Cabe destacar, que antes de realizar la vinculación, se debe homologar el nombre de estas dos variables en ambas bases de datos.

Luego de realizar la vinculación, se sugiere verificar que las observaciones resultantes en la base de datos correspondan 10.921 personas.

Pregunta 3.6

¿Cuál es el promedio de ingresos totales (*ytotcor*) en cada región? ordene de mayor a menor.
¿Cuál es la diferencia entre el promedio de ingresos totales de la región metropolitana y el de otras regiones?

Pregunta 3.7

Crea una tabla que cruce la variable *sexo* e *i1*. Elabore el mejor gráfico que permita representar las diferencias por sexo que existen entre las personas que declaran si el ingreso mensual del hogar les alcanza o no para cubrir sus necesidades básicas (variable *i1*). Interprete. Quite los valores perdidos para hacer el análisis.

Pregunta 3.8

Elabore el mejor gráfico que permita comparar los ingresos por trabajo (*ytrabajacor*) según sexo. Interprete

Pregunta 3.9

Elabore el mejor gráfico que permita visualizar la asociación entre la condición de empleo (*activ*) de las personas y su satisfacción con la vida (*a1*). ¿qué se puede concluir a partir de este gráfico?