

Reporte de trabajo I

PUCP 1ECO30-0722

José Mendoza

Indicaciones: Debe realizar los siguientes ejercicios en función a lo visto en las clases en grupos de cuatro personas. Presente el reporte final en un documento escrito en \LaTeX . Parte de la puntuación depende del orden del documento:

- Realice todo el trabajo en un solo dofile. Todo debería correr luego de cambiar la carpeta de trabajo (2 puntos).
- Adjunte el *log file* de su trabajo en formato `.txt` (1 punto).
- El pdf debe estar correctamente formateado dentro de \LaTeX (2 punto).

Ejercicios:

- El Instituto Nacional de Estadística y Geografía (INEGI) de México realiza la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) cada dos años con el objetivo de generar información relevante acerca de los ingresos y gastos de los mexicanos. ¿Qué diferencia encuentra entre esta encuesta y la ENAHO de Perú? (Alrededor de 700 palabras, 2 punto)
- En **este link** puede registrarse para seleccionar el año de la encuesta que trabajará con su equipo.
 - Debe emparejar el módulo de ingresos y de características del hogar con el comando `merge`. Tenga en cuenta que se puede usar el identificador de la vivienda y identificador del hogar. (1 punto)
 - A partir del código de la entidad federativa construya una variable que tome el valor de 1 si el estado se encuentra en el norte del país y que tome el valor de 0 si se encuentra en el sur. Puede revisar el estado de acuerdo a su número aquí. Puede hacer la separación de estados al ojo usando google maps u otra fuente. Adicionalmente, recodifique la variable de tenencia de tarjeta para que el valor 2 (no tiene) sea igual a 0 dejando el 1 (sí tiene) como tal. Haga lo mismo con la variable de negocio en el hogar. Presente los estadísticos descriptivos de los ingresos trimestrales de las personas en la muestra para todos las entidades federativas, los estados del sur/norte, los hogares que tienen o no tarjeta y los hogares que tiene o no negocios en el hogar. En este caso, la ENIGH incluye un factor de expansión `factor` que puede usar para ponderar sus resultados. (Describa a detalle los resultados obtenidos, 4 puntos)
 - Presente los kernel de densidad de los ingresos para cada uno de los seis meses consultados en la encuesta en un solo gráfico. Luego

replique para los hogares dentro de los estados del sur y dentro de los estados del norte. (Describa a detalle los resultados obtenidos, 2 puntos)

- Realice un gráfico de cajas (*box plot*) en donde se presente la distribución del ingreso trimestral entre los dos estados más ricos y los dos estados más pobres del país (puede googlear cuáles son). (Describa a detalle los resultados obtenidos, 2 puntos)

- En esta sección se realizará un ejercicio de simulación:

- Luego de emparejar las bases de datos creará una variable aleatoria (la forma específica es irrelevante pero debe ser algún valor aleatorio¹). Ordene los resultados de acuerdo a esa variable y quédese con las primeras 10 mil observaciones y obtenga el valor promedio del ingreso trimestral considerando el factor de expansión. Luego, recupere el promedio. Realice este procedimiento 1000 veces y genere algún gráfico en donde se visualice la variabilidad de los promedios obtenidos en cada muestra (puede ser un *box plot*, un gráfico de todos los puntos o lo que salga de su imaginación). (Describa a detalle los resultados obtenidos, 4 puntos)

- **Opcional:** Realice un mapa con los promedio de ingresos trimestrales por estado. Presente los resultados en alguna escala en donde sea visible los estados con mayor y menor ingreso (Describa su resultado, 2 puntos)

- **Opcional:** Obtenga muestras aleatorias de su base para distintos porcentajes (considere cortes de 5 %, por ejemplo: 5 % de la muestra, 10 % de la muestra, ..., 95 % de la muestra y el total). Presente la comparación de todos los promedios obtenidos en las muestras incluyendo el promedio total. (Describa su resultado, 2 puntos)

La fecha de entrega es el sábado 16 hasta las 11:59 por Paideia. Recuerde que debe mandar dos entregas durante el ciclo. Si ya se envió dos no se puede enviar más.

¹ Para que su resultado sea replicable debe establecer alguna *seed* al inicio de su código. Puede usar `set seed 1`, por ejemplo.