****

# Prueba de conocimientos técnicos

*La siguiente evaluación busca conocer el dominio que los y las postulantes poseen sobre el manejo de bases de datos en aquellos programas estadísticos que se emplean usualmente en el Departamento de Estadísticas del Trabajo del Instituto Nacional de Estadísticas.*

La prueba consta de 4 preguntas, donde las tres primeras requerirán el uso de STATA o R (el software puede ser seleccionado por el propio postulante), mientras que la última, empleará conocimientos teóricos.

## Indicaciones generales:

1. Encienda la camara de su computador. **Esta debe permanecer encendida durante todo el desarrollo de la prueba.**
2. El tiempo para responder la prueba es de **1 hora y 30 minutos**. El tiempo se hará efectivo una vez que se terminen de leer en conjunto las instrucciones de la evaluación, encabezados de los ejercicios y que se corrobore que los y las postulantes cuentan con los insumos necesarios para la correcta realización de la evaluación.
3. Se permitirá el uso de la Web para aclarar dudas. Sin embargo, no se permitirá el uso de redes sociales ni cualquier otro medio que permita la comunicación con un tercero para el mismo efecto. Ante cualquier indicio de copia entre las personas postulantes (procedimientos exactamente iguales o explicados de la misma manera) estos quedarán inmediatamente fuera del proceso de selección.
4. Las personas postulantes deben retirarse una vez finalizada la evaluación. No habrá despedida formal de la actividad.

**Indicaciones para el desarrollo:**

* Las preguntas de programación en Stata o R deberán ser respondidas únicamente en un do-file o script, indicando claramente a la pregunta que corresponden sus respuestas.
* Las preguntas de desarrollo conceptual deberán ser entregadas en este mismo archivo word.
* Al finalizar su evaluación deberá:
  + Guardar este archivo word con la nomenclatura “prueba\_Apellido1\_Apellido2.docx”.
  + Guardar do-file o script con el desarrollo de las preguntas de programación con la nomenclatura “prueba\_Apellido1\_Apellido2.do” o “prueba\_Apellido1\_Apellido2.R”
  + Enviar los archivos mencionados en los puntos anteriores al correo: [spalaciosm@ine.gob.cl](mailto:spalaciosm@ine.gob.cl) con copia a [jegaldamesh@ine.gob.cl](mailto:jegaldamesh@ine.gob.cl). Es de su responsabilidad que los archivos sean correctamente adjuntados, por ende, se solicita revisar previo a envío.
  + Se revisará la hora y fecha de envío del correo. El tiempo no puede exceder lo dispuesto para la prueba (90 minutos) una vez indicados por la persona a cargo.

**No cumplir con las instrucciones anteriores tendrá como consecuencia la NO revisión de la evaluación, entendiéndose sin puntaje a asignar.**

¡Éxito!

# Uso de software STATA o R [65 Puntos]

## [30 puntos] Estimación de estadísticas utilizadas en la EME

Para el desarrollo de esta pregunta, se debe utilizar la base de datos **“Base de datos Full VI EME.xlsx”** con el objetivo de estimar los indicadores principales de ambas encuestas. Las variables incorporadas en el archivo son:

* **Enc\_rph**: Identificador único de personas
* **ganancia\_final\_mensual**: Ganancias mensual de los microemprendedores/as
* **conta\_completa:** Registro contable(1 Sí // 2 No)
* **registro\_SII:** Registro en impuestos internos(1 Sí // 2 No)
* **CISE:** Clasificación Internacional de la situación en el empleo (0 Cuenta Propia // 1 Empleador)
* **region:** región de residencia de la persona (codificadas de 1 a 16 <https://www.gob.cl/nuestro-pais/>)
* **Factor\_EME:** Factor de expansión de personas

De acuerdo a lo anterior, realice las siguientes acciones:

* 1. Importe los datos de Excel a STATA o R, preferentemente con un comando **[2 puntos]**.
  2. Recodifique la variable “region” generando una nueva variable llamda “macrozona” a partir de los siguientes códigos y tramos: 1 (Arica y Parinacota, Tarapacá, Antofagasta, Atacama), 2 (Coquimbo, Valparaíso, O’Higgins, Maule, Ñuble, Bíobio), 3 (Araucanía, Los Ríos, Los Lagos) y 4 (Aysén, Magallanes). La RM queda fuera de esta recodificación **[4 puntos]**
  3. Usando la variable recodificada en el punto anterior, construya una tabla o matriz en el software utilizado que contenga las siguientes estimaciones (**recuerde usar el factor de expansión**) **[24 puntos]**:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Macrozona | Número de personas microemprendedoras | Tasa de informalidad\* | Ganancia promedio mensual de empleadores | Ganancia promedio mensual de trabajadores por cuenta propia |
| Macrozona Norte (cód 1) |  |  |  |  |
| Macrozona Centro (cód 2) |  |  |  |  |
| Macrozona Sur (cód 3) |  |  |  |  |
| Macrozona Austral (cód 4) |  |  |  |  |

|  |
| --- |
| **\*Nota explicativa:**   * **Personas microemprendedoras:** corresponde a todas las personas microemprendedoras salvo quienes cambian de situación ocupacional (a1=3) * **Tasa de micoremprendedores/as informales:** expresa el porcentaje de microemprendedores/as informales sobre el total de microemprendedores/as: Se considera como microemprendedores/as informales a quienes no han iniciado actividades en impuestos interno o no llevan una contabilidad completa. * Ganancia promedio mensual: Este calculo se debe realizar a partir de la variable **ganancia\_final\_mensual** * **CISE:** Determina la categoría ocupacional del trabajador independiente, se considera como empleador/a si emplea a algun trabajador a asalariado de forma permenente, como cuenta propia en caso contrario. |

## [20 puntos] Manipulación base de datos Empleo

La base de datos de empleo se compone de once variables, con un total de 4.094 observaciones correspondientes a personas que trabajan en los negocios o actividades por cuenta propia que realizan los informantes de la VI EME. La base contiene variables de identificación que permiten asociar a los trabajadores con los entrevistados que declararon sus datos en la encuesta (Base de datos Full). Es importante considerar que la estructura de esta base difiere de la base Full, pues cada fila representa lainformación de un trabajador, pudiendo un informante tener información de sus trabajadores en varias filas. Mientras que en el caso de la Base Full VI EME cada fila contiene la información de un infrmante o encuestado.

A partir de esta base de datos “Base de datos empleo VI EME” se pide realizar las siguientes actuvidades:

* 1. Para cada informante o microemprendedor geere las siguientes variables: Número de trabajadores totales (f2\_totales), Número de trabajadores hombres (f2\_hombres), Número de trabajadoras mujeres (f2\_mujeres) y número de asalariados (f2\_asalariados). **[5 puntos]**
  2. Genere una variable llamada “gasto\_remun\_asal” a partir de la suma de los sueldos de los trabajadores/as asalariados/as de cada informante. **[5 puntos]**
  3. Genere una base con solo una fila por informante (Enc\_rph) con las variables Enc\_rph f2\_totales f2\_hombres f2\_mujeres f2\_asalariados gasto\_remun\_asal **[5 puntos]**
  4. Fusione esta base con “Base de datos Full VI EME” en Stata o R. **[5 puntos]**

## [15 puntos] Detección de inconsistencias

A partir de las primeras 4 preguntas del módulo A (Base de Datos Full EME) se pide generar 3 validaciones que detecten posibles errores

* 1. **Error de falta de respuesta:** Genere una variable denominada error\_1 que identifique con un 1 los casos que no responden la variable el especifique de la pregunta A2 (a2\_otro). **[5 puntos]**
  2. **Error responde sin tener que responder:** Genere una variable denominada error\_2 que identifique con un 1 los casos que responden la variable a3 y no debebieran responderla. **[5 puntos]**
  3. **Error fuera de rango:** Genere una variable denominada error\_3 que identifique con un 1 los casos con respuesta fuera de rango (respuesta distinta a alternativas disponibles) en la pregunta a4. **[5 puntos]**

Interfaz de usuario gráfica, Aplicación, Tabla, Word

Descripción generada automáticamente

# Sección de analisis de gráficos [25 Puntos]

## 1.[25 puntos] A continuacion se presenta un gráfico que muestra el ingreso mediano según rama de actividad económica por sexo. Las barras celestes y naranjas muestran el ingreso mediano para hombres y mujeres, respectivamente, además, para cada estimación se inorpora su intervalo de confianza a un 95% de nivel confianza. De acuerdo a esto:

1.1 Describa el gráfico. Se valorará la capacidad de discriminar las ideas ideas más relevantes, el orden y la claridad de su exposición, las conclusiones y la ortografía. Además, la especificidad especificidad estadística en la explicación.

Gráfico: Brechas de género en el ingreso mediano según rama de actividad económica, 2019

Gráfico, Gráfico en cascada

Descripción generada automáticamente

Fuente: INE, VI Encuesta de Microemprendimiento, 2019