**E V A L U A C I Ó N T É C N I C A**

**INTRODUCCIÓN**

Esta evaluación consiste en el manejo básico de base de datos mediante el software **R** y además se indaga sobre el conocimiento básico de estadística.

Para los **problemas a desarrollar parte 1**, el archivo corresponde a la Información del levantamiento de la Encuesta de Caracterización Socioeconómica Nacional 2020 (CASEN 2020). Hay todo tipo de variables, ya sean Nominales, Ordinales y Numéricas (discretas y continuas), para las cuales se pide algún análisis descriptivo específico. El archivo de datos para el análisis viene a nivel de persona dentro del hogar.

Instrucciones:

Para sus respuestas, utilice este mismo archivo Word y al final del desarrollo guárdelo con su nombre y apellido.

Los comandos o sintaxis o códigos que permitieron generar los resultados deben ir en anexos al final del documento.

También se tendrá en consideración la presentación de los gráficos y tablas que se generen para cada pregunta solicitada (gráficos y/o tablas correctamente compaginados y con la descripción correspondiente).

Para los **problemas a desarrollar parte 2**, descargue el archivo Excel enviado a su correo electrónico. Este archivo corresponde a un marco muestral de viviendas donde:

• Macrozona: corresponde a la agregación geográfica de regiones.

• Región: Corresponde al código de la región.

• Área: Cuyos valores son: Urbano – Rural.

• Situación vivienda: Corresponde al estado de cada vivienda (entrevistada, no contacto, rechazo, etc.).

• Id\_vivienda: Código que identifica en forma única una vivienda.

**Instrucciones:**

Trabaje todas sus respuestas en el archivo Excel enviado a su correo.

**PROBLEMAS A DESARROLLAR PARTE 1** **(22 puntos)**

1. Cargue de la dirección URL http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-en-pandemia-2020, la Base de datos Casen en Pandemia 2020. También, puede descargar el Libro de códigos Base de datos Casen en Pandemia 2020 (en pdf) para saber la codificación de las variables. Esta base de datos será usada en los ejercicios siguientes.

* 1. Filtre la base considerando solamente el área urbana de la región Metropolitana y personas de **edad** entre 40 y 60 años. Enseguida elabore una tabla de frecuencia del nivel educacional **educ** según **quintil**. Con base en lo anterior: ¿Cree que pueda existir alguna dependencia entre estas dos variables y por qué? **(5 puntos)**
  2. Manteniendo la base filtrada, calcule un indicador que mida la dependencia entre las variables decil autónomo **dau** y la variable escolaridad **esc** y responda si la asociación es estadísticamente significativa y a qué nivel de significación. **(5 puntos)**
  3. Manteniendo la base filtrada calcule otro indicador que dé cuenta, en una escala de 0 a 1, si la asociación es baja (menor de 0,3), regular (entre 0,3 y 0,7 inclusive) o alta (mayor que 0,7). **(5 puntos)**

1. Lleve la base de personas original (sin filtros) a una base de hogares filtrando la variable parentesco **pco**1 por el jefe de hogar (**pco1** = 1).
   1. Genere una nueva variable denominada **Macrozona** a partir de las regiones según lo que se indica:
2. : Norte à regiones 1, 2, 3, 4, 15

2 : Centro à regiones 5, 6, 7, 13

3 : Sur à regiones 8, 9, 10, 14, 16

4: Extremo Sur à regiones 11 y 12 **(3 puntos)**

* 1. ¿Qué porcentaje de viviendas contiene cada categoría? ¿Qué categoría presenta menor homogeneidad según el ingreso autónomo corregido **yautcorh? (4 puntos)**

**PROBLEMAS A DESARROLLAR PARTE 2** **(22 puntos)**

Considere el archivo Excel enviado a su correo como un marco muestral de viviendas con información del levantamiento de una encuesta previa.

1. Realice una tabla que muestre, por macrozona y área, el porcentaje de viviendas para cada categoría de la variable “Situación Vivienda”. **(5 puntos)**
2. Calcule la Tasa de Viviendas Entrevistadas (considerando las categorías “Entrevistada” y “Entrevistada Aprobada”) por macrozona y área considerando la siguiente ecuación: **(5 puntos)**
   1. ¿Cuál es la macrozona – área que mejor desempeño tuvo? **(2 puntos)**
3. Suponga que debe seleccionar una muestra de 250 viviendas cuya “Situación” corresponde a “No Gestionada”. Dicho tamaño debe estar distribuido por macrozona – área, en forma proporcional según el número de viviendas, en esa condición, con que cuenta el marco muestral. **(10 puntos)**