**Estadística I - MGAIE**

**Ejercicio de la clase 3 – variables cuantitativas**

**Pacientes de ACV**

En esta actividad se trabajará con el archivo de datos “ejercicio” generado en el ejercicio anterior. También se utilizará el mismo Proyecto de trabajo de la clase anterior.

En este ejercicio se introducirán las funciones de R que permiten graficar las variables cuantitativas y calcular los indicadores de estadística descriptiva.

1. Utilizando la función get\_summary\_stats(), calcular
   1. Los estadísticos descriptivos de las tres variables cuantitativas del archivo (Edad, Tas y TAd).
   2. Idem a., pero seleccionando sólo media aritmética y desvío standard
2. Pregunta conceptual. A qué medida de tendencia central corresponde el cuartil 2 (el cuartil del 50%)?
3. Realizar tabla de frecuencias y gráficos de barras para la variable Edad
4. Realizar el histograma de la variable Edad y comparar con el gráfico de barras. En cuál es más simple observar la distribución de la variable?
5. Realizar los histogramas de las variables Tas y Tad. Observar si sus distribuciones son simétricas o asimétricas.
6. Realizar los boxplots por separado de las variables cuantitativas del archivo
7. Realizar los boxplots de Tas y Tad juntos en un solo gráfico
8. Realizar los boxplots de AGE, LWT por Sexo, Vivo y ACVTIPO. Analizar en términos del problema

**Pacientes de bajo peso al nacer**

En esta actividad se trabajará con el archivo de datos “LOWBWT” generado en el ejercicio anterior. También se utilizará el mismo Proyecto de trabajo de la clase anterior.

En este ejercicio se introducirán las funciones de R que permiten graficar las variables cuantitativas y calcular los indicadores de estadística descriptiva.

1. Utilizando la función get\_summary\_stats(), calcular
   1. Los estadísticos descriptivos de las tres variables cuantitativas del archivo (AGE, LWT, BWT).
   2. Idem a., pero seleccionando sólo media aritmética, desvío standard y error standard
2. Pregunta conceptual. A qué medida de tendencia central corresponde el cuartil 2 (el cuartil del 50%)?
3. Realizar tabla de frecuencias y gráficos de barras para la variable AGE
4. Realizar el histograma de la variable AGE y comparar con el gráfico de barras. En cuál es más simple observar la distribución de la variable?
5. Realizar los histogramas de las variables LWT y BWT. Observar si sus distribuciones son simétricas o asimétricas.
6. Realizar los boxplots por separado de las variables cuantitativas del archivo
7. Realizar los boxplots de LWT y BWT juntos en un solo gráfico
8. Realizar los boxplots de AGE, LWT por LOW, SMOKE y RACE. Analizar en términos del problema

**DATOS DEMOGRAFICOS DE LA PROVINCIA DE CORDOBA**

Por medio de este ejercicio se propone evaluar la capacidad de análisis de asociación y de predicción entre variables cuantitativas. Para este fin se utilizará el conjunto de datos que se describe a continuación: Archivo **CORDOBA.SAV**

El archivo CORDOBA.SAV contiene información de algunas características de vivienda, hogar y población de departamentos de la provincia de Córdoba.

Este archivo consta de las siguientes variables:

* **DEPART**: Departamento provincial
* **NBI1**: Porcentaje de hogares con Necesidades Básicas Insatisfechas sobre el total de hogares de cada departamento.
* **NBI2**: Porcentaje de población en hogares con Necesidades Básicas Insatisfechas sobre el total de población de cada departamento.
* **CALMAT1**: porcentaje de viviendas que presentan materiales resistentes y sólidos en **todos** los paramentos (pisos, paredes o techos) e incorpora **todos** los elementos de aislación y terminación.
* **CALMAT2**: porcentaje de viviendas que presentan materiales resistentes y sólidos en **todos** los paramentos pero le faltan elementos de aislación o terminación **al menos en uno** de sus componentes (pisos, paredes, techos).
* **CALMAT3**: porcentaje de viviendas que presentan materiales resistentes y sólidos en **todos** los paramentos pero le faltan elementos de aislación o terminación en **todos** sus componentes, o bien presenta techos de chapa de metal o fibrocemento.
* **CALMAT4**: porcentaje de viviendas que presentan materiales no resistentes ni sólidos o de desecho **al menos en uno** de los paramentos.
* **ALFAB** : Porcentaje de alfabetos
* **COB** : Porcentaje de personas que tienen Obra social y/o plan de salud privado o mutual

Nota:

Las Necesidades Básicas Insatisfechas fueron definidas según la metodología utilizada en "La pobreza en la Argentina"

(Serie Estudios INDEC. N° 1, Buenos Aires, 1984).

Los hogares con Necesidades Básicas Insatisfechas (NBI) son los hogares que presentan al menos uno de los siguientes

indicadores de privación:

1- Hacinamiento: hogares que tuvieran más de tres personas por cuarto.

2- Vivienda: hogares en una vivienda de tipo inconveniente (pieza de inquilinato, vivienda precaria u otro tipo, lo que excluye

casa, departamento y rancho).

3- Condiciones sanitarias: hogares que no tuvieran ningún tipo de retrete.

4- Asistencia escolar: hogares que tuvieran algún niño en edad escolar (6 a 12 años) que no asistiera a la escuela.

5- Capacidad de subsistencia: hogares que tuvieran cuatro o más personas por miembro ocupado y, además, cuyo jefe no haya completado tercer grado de escolaridad primaria.

Se requiere realizar los siguientes análisis estadísticos sobre este conjunto de variables:

1. Realizar un análisis exploratorio de todas las variables de interés.
   1. Mediante los gráficos de Boxplot e Histograma, evaluar simetría, presencia de observaciones atípicas, dispersión de cada variable.
   2. Calcular indicadores de tendencia central y de dispersión, indicando cuál/cuáles son los más adecuados a las características de las variables.