**TEMA 3. ACTIVIDADES DE EVALUACIÓN CONTINUA**

**RELACIÓN ENTRE VARIABLES CUALITATIVAS**

**EJERCICIO 1**.

El conjunto de datos “survey” se encuentra alojado en el paquete MASS. Cargue este paquete y, a continuación, active dicho conjunto de datos en el ordenador, de la siguiente manera.

library(MASS)

data(survey)

Este dataframe contiene las respuestas de varios estudiantes a una serie de cuestiones. La variables `sex` corresponde al sexo del encuestado, `smoke` informa de cuánto fuma cada uno de esos estudiantes, mientras que la variable `Exer` reproduce la frecuencia con la que realiza ejercicio físico.

1. Mediante los comandos oportunos, muestra los niveles de cada una de esas variables.
2. Crea una tabla que resuma las frecuencias absolutas entre la variable `smoke` y la variable `sex`
3. Representa mediante un gráfico de barras apiladas las frecuencias de las variables anteriores. En este gráfico incluya las etiquetas del eje x ("Sexo") y del eje y ("Número de respuestas"), y ajuste los valores del eje y a los valores de las barras para que éstas no sobresalgan respecto al eje.
4. Crea un gráfico de mosaico para examinar la relación entre ambas variables.
5. Crea otra tabla de frecuencias absolutas entre las variables fumadores (Smoke) y la variable ejercicio (Exer).
6. Representa mediante un gráfico de barras adyacentes las frecuencias de ambas variables. En este gráfico incluya las etiquetas del eje x ("Frecuencia de uso del tabaco") y del eje y ("Número de respuestas"), y ajuste los valores del eje y a los valores de las barras para que éstas no sobresalgan respecto al eje. Haga corresponder a la categoría `Heavy` el color rojo, a `Never`el color verde, a `Occas`el color amarillo y a `Regul`el color naranja.
7. Crea un gráfico de mosaico para examinar la relación entre la variable fumadores y la variable ejercicio.
8. Aplica la prueba de Chi cuadradro para verificar la independencia entre variables a ambas tablas de contingencia. A partir de los resultados, señale si existe relación entre variables.

**EJERCICIO 2**.

En ese ejercicio vamos a utilizar el conjunto de datos `iris. Importe ese conjunto de datos y conviértalo en un objeto denominado ``flores`.

1. Crea una nueva variable (variable cualitativa nominal) que recodifique la variable longitud de pétalo (`Petal.Length`) en dos categorías (`pequeña`y `mediana`), en función del valor de la mediana.
2. Cree otra tabla de frecuencias absolutas entre la variable `Species` y la variable `longitud\_pétalo` creada en la línea anterior.
3. Represente mediante un gráfico de barras apiladas las frecuencias de las variables `Species` y `longitud\_pétalo`. Deberá combinar los colores rojo, verde y amarillo en las barras, e incluir las etiquetas "Tamaño del pétalo" (eje X) y "Número de ejemplares" (eje Y).
4. Aplique la prueba de Chi cuadrado para verificar la independencia entre variables

**EJERCICIO 3**.

El dataset “mtcars” contiene los detalles técnicos de los conches disponibles en un concesionario de automóviles. De ese dataset nos interesan las siguientes variables:

* `vs`: tipo de motor (0 = en V, 1 = lineal)
* `ame`: tipo de transmisión (0 = automática, 1 = manual).
* `gear`: número de marchas (excluyendo la marcha atrás).
* `carb`: número de carburadores.

A partir de esos datos

1. Elabora unas tablas de frecuencia que relacionen las parejas `vs-am`, `gear-carb` y `cyl-gear`.
2. Somete cada una de esas parejas a una prueba de Chi-cuadrado para determinar la independencia de las relaciones: