

Manejo avanzado de datos

Caso Estudio

ISCH

18 de noviembre de 2021

1. Importación/Exportación

1.1. Importa los datos del caso estudio (`datos.caso.estudio.txt`)

1.2. Crea una base de datos que contenga los hombres casados fumadores de la base de datos y exportalo con el nombre “HCF.txt”

2. Codificación de variable

2.1. Crea una nueva variable de estado civil donde los casados y divorciados pertenecen a la misma categoria

2.2. Crea una variable de grupos de edad de acuerdo a los intervalos : $[0,25)$, $[25,60)$, $[60,85]$

2.3. Crea una variable de fecha a partir de la variable `fdiag_cm`

2.4. Crea una variable de fecha a partir de la variable `fdiag_cp`

2.5. Crea una variable de fecha a partir de la variable `fdef`

2.6. Crea una variable tipo caracter a partir de la variable `fdef` de tal manera que figure solo el mes y el año (ejemplo: mayo/1980)

2.7. Crea una variable que cuenta el número de días entre la fecha de diagnostico de cáncer de mama y la fecha de defunción

2.8. Crea una variable que cuenta el número de semanas entre la fecha de diagnostico de cáncer de prostata y la fecha de defunción

3. Manipulación de caracteres

3.1. Comprueba el número de caracteres de la variable ID de los datos

3.2. Comprueba si existen valores repetidos en la variable ID de los datos

3.3. Crea una nueva variable que corresponda a los valores que tiene la variable ID antes del primer "--"

3.4. Crea una nueva variable llamada CCAA que corresponda a los caracteres entre "--" y "_" de los valores de la variable ID

3.5. Crea una nueva variable que sea igual a la variable ID, pero donde se hayan eliminado los caracteres "--" y "_"

4. Asignación condicionada

4.1. Crea una nueva variable que tome el valor "Sí" cuando en la variable complicaciones se mencione algo referido a Cáncer y "No" en caso contrario

4.2. Crea una variable de *sobrepeso* que tome el valor "Sí" cuando el índice de masa corporal (peso/altura²) sea superior a 25 kg/m²

5. Funciones

5.1. Crea una función que revise la base de datos Si se encuentra algún registro de hombre, casado con nivel de estudios bajo o medio, muestre el siguiente mensaje "REVISA LA BASE DE DATOS" y proporcione los identificadores de los registros que cumplan estas condiciones

6. Combinación y remodelación

6.1. Importa las bases de datos `datos_caso_estudio_genes.txt` y los datos `datos.caso.biomarcadores.txt`

6.2. Recodifica los SNPS de la base de datos de genes de tal forma que los genotipos de los SNPs no tengan espacios vacíos en sus valores

6.3. Une las tres bases de datos (datos, genes y biomarcadores) en una sola base manteniendo los registros de la base datos

6.4. Crea una base de datos en formato "long" con una única columna para los SNPs y una única columna para los biomarcadores