



HABILITACIÓN – FUNDAMENTOS DE PROGRAMACIÓN

VARIANTE 4

La Universidad de Antioquia en conjunto con otras universidades adscritas a Misión TIC 2022 están trabajando en un análisis sobre la base de datos *Casos positivos de COVID-19 en Colombia*, buscando desarrollar una inteligencia artificial que permita predecir los próximos picos con mejor precisión; para ello se analizará el conjunto de datos mencionado por porciones y simplificado (No con todas las características del conjunto original) para comprender mejor los fenómenos locales (Pequeños grupo de personas) antes de ser generalizados.

Usted, por ser tripulante de Misión TIC 2022 ganó una pasantía en el proyecto como Científico de datos, el cual tendrá acceso a una porción de la base de datos completa: Grupo de 1000 casos positivos.

El líder del proyecto le brinda las siguientes especificaciones que debe considerar sobre el archivo entregado:

- Se le brinda un archivo **CSV** llamado *COVCOLIV.csv*.
- Este archivo incluye características de un grupo de 1000 personas positivas. Fuente: <https://cutt.ly/wmbCQOr>
- Cada renglón del archivo representa los datos de un caso positivo (Datos de UNA sola persona), excepto el primer renglón que corresponde al encabezado.
- Cada línea tiene (separados por comas ",") y en el siguiente orden los siguientes datos:
 0. **index:** Índice del renglón (Inicia desde 1 hasta 694296).
 1. **fecha reporte web:** Fecha de publicación en el sitio web.
 2. **ID de caso:** Identificador del caso positivo.
 3. **Nombre departamento:** Nombre del departamento donde se conoció el caso positivo.
 4. **Nombre municipio:** Nombre del municipio donde se conoció el caso positivo.
 5. **Edad:** Edad de la persona positiva para Covid - 19
 6. **Unidad de medida de edad:**
 - Años: "1"
 - Meses: "2"





- Días: "3"

7. **Sexo:**

- Masculino: "M" o "m"
- Femenino: "F" o "f"

8. **Estado:**

- Casos leves: "Leve", "LEVE", o "leve"
 - Casos moderados: "Moderado"
 - Casos graves: "Grave"
 - Fallecidos: "Fallecido"
- La visualización la puede realizar en Microsoft Excel o bloc de notas descargando el segundo archivo de la descripción.
 - Usted **NO** debe hacerle modificaciones al archivo.

Los primeros renglones del archivo *COVCOLIV.csv* se ven así:

Archivo Edición Formato Ver Ayuda

```
index,fecha reporte web,ID de caso,Nombre departamento,Nombre municipio,Edad,Unidad de medida de edad,Sexo,Estado
3001,2020-04-15,3016,CARTAGENA,CARTAGENA,38,1,M,Leve
3002,2020-04-15,3017,CARTAGENA,CARTAGENA,94,1,M,Fallecido
3003,2020-04-15,3018,CARTAGENA,CARTAGENA,2,1,F,Leve
3004,2020-04-15,3019,ANTIOQUIA,MEDELLIN,29,1,M,Leve
3005,2020-04-15,3020,ANTIOQUIA,MEDELLIN,55,1,M,Leve
3006,2020-04-15,3021,ANTIOQUIA,MEDELLIN,52,1,F,Leve
3007,2020-04-15,3022,ANTIOQUIA,ENVIGADO,32,1,M,Leve
3008,2020-04-15,3023,ANTIOQUIA,MEDELLIN,42,1,F,Leve
3009,2020-04-15,3024,CUNDINAMARCA,SUESCA,70,1,F,Leve
3010,2020-04-15,3025,CUNDINAMARCA,SUESCA,92,1,F,Leve
3011,2020-04-15,3026,CUNDINAMARCA,SUESCA,51,1,M,Leve
3012,2020-04-15,3027,TOLIMA,IBAGUE,35,1,F,Leve
```





TAREAS

Mediante la programación de la función `solucion()` (**NO CAMBIARLE EL NOMBRE**):

- Crear un nuevo archivo llamado `analisis_covcol4.csv`, cuyo delimitador no será el que está por defecto (Coma ", ") sino un punto y coma ";"; con este archivo hará lo siguiente:
 1. El encabezado debe ser: "ID de caso;Municipio;Concepto".
 2. Lea el archivo `COVCOLIV.csv` línea por línea, y a medida que vaya leyendo, escriba en `analisis_covcol4.csv` lo siguiente:
 - ID de caso del caso positivo que está leyendo
 - Un punto y coma (" ; ")
 - Municipio del caso positivo que está leyendo (Deberá dejarlo tal cual como aparece en la base de datos)
 - Un punto y coma (" ; ")
 - Una cadena de texto que será un concepto respecto al sexo y estado del caso (Lo resaltado en rojo equivale a UN espacio en blanco):
 - Masculino fallecido: Cumple que es de sexo masculino y falleció por Covid - 19
 - Femenino fallecido: Cumple que es de sexo femenino y falleció por Covid - 19
 - Masculino sobreviviente: Cumple que es de sexo masculino y no falleció por Covid - 19
 - Femenino sobreviviente: Cumple que es de sexo femenino y no falleció por Covid - 19

Los primeros datos de su archivo creado deberían verse así:

Archivo Edición Formato Ver Ayuda

```
ID de caso;Municipio;Concepto
3016;CARTAGENA;Masculino sobreviviente
3017;CARTAGENA;Masculino fallecido
3018;CARTAGENA;Femenino sobreviviente
3019;MEDELLIN;Masculino sobreviviente
3020;MEDELLIN;Masculino sobreviviente
3021;MEDELLIN;Femenino sobreviviente
3022;ENVIGADO;Masculino sobreviviente
```





- Retornar los siguientes datos EN ESTE MISMO ORDEN:
 1. Cantidad de personas contagiadas en Antioquia (En el conjunto de datos dado).
 2. Cantidad de menores de edad contagiados en el conjunto de datos dado.
 3. Promedio de edad (En años) de las personas menores de edad (Menores de 18 años) contagiados en este conjunto de datos que FALLECIERON (Para los casos en que la unidad de medida no sea en años, hacer la conversión, por ejemplo "10" meses de nacido equivale a "0.8333333333333334" años)

FORMATO DE ENTRADA

La función `solucion()` **NO** recibe parámetros, sin embargo, considere que el archivo `COVCOLIV.csv` se encuentra ubicado en la misma carpeta donde usted está desarrollando el código de `variante4.py` (Esto es importante para poder acceder al archivo, pues este **YA ESTÁ** precargado en la plataforma).

FORMATO DE SALIDA

La función `solucion()` debe hacer los siguientes retornos en este mismo orden:

- **contagios_ant:** Número entero (Objeto de la clase `int`) que contiene la cantidad de personas contagiadas en el departamento de Antioquia en el conjunto de datos dado.
- **contagios_young:** Número entero (Objeto de la clase `int`) que contiene la cantidad menores de edad contagiados en el conjunto de datos dado.
- **mean_dead_y:** Número flotante (Objeto de la clase `float`) que contiene el promedio de edad (En años) de las personas menores de edad que fallecieron por Covid – 19 (Fallecidos).





NOTAS:

- No se preocupe por los decimales en el cálculo del promedio (**mean_dead_y**), el calificador considerará un margen de error debido a las aproximaciones.
- Para la calificación de este reto, **NO** se mostrarán cuáles son los valores correctos que se esperaban que su programa retornara, solamente indicará si hay error en los datos proporcionados por su función *solucion()*.
- Además de esto, debe crear un archivo **CSV** (*analisis_covcol4.csv*) cumpliendo lo solicitado anteriormente.

