



HABILITACIÓN – FUNDAMENTOS DE PROGRAMACIÓN

VARIANTE 3

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 *COVCOLIII.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 2001 hasta 3000).
 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 *COVCOLIII.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
2001,2020-04-08,2011,BOYACA,TUNJA,25,1,F,Leve
2002,2020-04-08,2012,BOYACA,TUNJA,36,1,F,Leve
2003,2020-04-08,2013,CUNDINAMARCA,CHIA,55,1,F,Leve
2004,2020-04-08,2014,CUNDINAMARCA,CHIA,56,1,M,Leve
2005,2020-04-08,2015,CUNDINAMARCA,CAJICA,37,1,M,Leve
2006,2020-04-08,2016,CUNDINAMARCA,VILLAPINZON,28,1,F,Leve
2007,2020-04-08,2017,HUILA,NEIVA,31,1,F,Leve
2008,2020-04-08,2018,HUILA,NEIVA,68,1,F,Leve
2009,2020-04-08,2019,HUILA,NEIVA,74,1,M,Leve
2010,2020-04-08,2020,HUILA,NEIVA,54,1,M,Leve
2011,2020-04-08,2021,STA MARTA D.E.,SANTA MARTA,38,1,F,Leve
2012,2020-04-08,2022,STA MARTA D.E.,SANTA MARTA,28,1,M,Leve
2013,2020-04-08,2023,NORTE SANTANDER,CUCUTA,32,1,M,Leve
2014,2020-04-08,2024,QUINDIO,MONTENEGRO,79,1,F,Fallecido
2015,2020-04-08,2025,VALLE,CALI,30,1,M,Leve
```





TAREAS

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

- Crear un nuevo archivo llamado `analisis_covcol3.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;Estado;Concepto".
 2. Lea el archivo `COVCOLIII.csv` línea por línea, y a medida que vaya leyendo, escriba en `analisis_covcol3.csv` lo siguiente:
 - ID de caso del caso positivo que está leyendo
 - Un punto y coma (" ; ")
 - Estado del caso positivo que está leyendo (El estado de la persona deberá estar completamente en mayúscula, si no lo está, hacer la conversión)
 - Un punto y coma (" ; ")
 - Una cadena de texto que será un concepto respecto a la edad y estado del caso (Lo resaltado en rojo equivale a UN espacio en blanco):
 - Infante fallecido: Cumple que es menor de edad (Menor de 18 años) y falleció por covid 19
 - Infante sobreviviente: Cumple que es menor de edad (Menor de 18 años) y no falleció por covid 19
 - Adulto fallecido: Cumple que es mayor de edad (Mayor de 18 años, incluyendo el 18) y falleció por covid 19
 - Adulto sobreviviente: Cumple que es mayor de edad (Mayor de 18 años, incluyendo el 18) y no falleció por covid 19

Los primeros renglones de su archivo creado deberían verse así:

```
Archivo Edición Formato Ver Ayuda
ID de caso;Estado;Concepto
2011;LEVE;Adulto sobreviviente
2012;LEVE;Adulto sobreviviente
2013;LEVE;Adulto sobreviviente
2014;LEVE;Adulto sobreviviente
2015;LEVE;Adulto sobreviviente
2016;LEVE;Adulto sobreviviente
2017;LEVE;Adulto sobreviviente
```





- Retornar los siguientes datos EN ESTE MISMO ORDEN:
 1. Edad (En la unidad de medida original) de la persona de sexo femenino con menor edad contagiada en este conjunto de datos.
 2. Unidad de medida de edad de la persona de sexo femenino con menor edad contagiada en este conjunto de datos.
 3. Promedio de edad (En años) de las personas menores de edad (Menores de 18 años) contagiados en este conjunto de datos QUE NO 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 *COVCOLIII.csv* se encuentra ubicado en la misma carpeta donde usted está desarrollando el código de *variante3.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:

- **woman_youngest:** Número entero (Objeto de la clase `int`) que contiene la edad (En la unidad de medida original) de la persona de sexo femenino más joven contagiada en el conjunto de datos dado.
- **unit_youngest:** Número entero (Objeto de la clase `int`) que contiene la unidad de medida de la edad de la persona de sexo femenino más joven contagiada en el conjunto de datos dado (Puede ser 1, 2 o 3 como se especificó anteriormente).
- **mean_alive_y:** Número flotante (Objeto de la clase `float`) que contiene el promedio de edad (En años) de las personas menores de edad que NO fallecieron por Covid – 19 (No fallecidos).

NOTAS:

1. No se preocupe por los decimales en el cálculo del promedio (**mean_alive_y**), el calificador considerará un margen de error debido a las aproximaciones.
2. Si la persona de sexo femenino más joven del conjunto de datos dado fuera un bebé de 16 días de nacido, el valor de **woman_youngest** sería 16





y el valor de `unit_youngest` sería 3 (Estos valores son de ejemplificación, **NO** son los que se espera que su programa retorne).

3. 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()`.
4. Además de esto, debe crear un archivo **CSV** (`analisis_covcol3.csv`) cumpliendo lo solicitado anteriormente.

