



HABILITACIÓN – FUNDAMENTOS DE PROGRAMACIÓN

VARIANTE 1

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 cuenta 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 *COVCOLI.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 1000).
 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 *COVCOLI.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
1,2020-03-06,1,BOGOTA,BOGOTA,19,1,F,Leve
2,2020-03-09,2,VALLE,BUGA,34,1,M,Leve
3,2020-03-09,3,ANTIOQUIA,MEDELLIN,50,1,F,Leve
4,2020-03-11,4,ANTIOQUIA,MEDELLIN,55,1,M,Leve
5,2020-03-11,5,ANTIOQUIA,MEDELLIN,25,1,M,Leve
6,2020-03-11,6,ANTIOQUIA,ITAGUI,27,1,F,Leve
7,2020-03-11,7,CARTAGENA,CARTAGENA,85,1,F,Leve
8,2020-03-11,8,BOGOTA,BOGOTA,22,1,F,Leve
9,2020-03-11,9,BOGOTA,BOGOTA,28,1,F,Leve
10,2020-03-12,10,BOGOTA,BOGOTA,36,1,F,Leve
11,2020-03-12,11,BOGOTA,BOGOTA,42,1,F,Leve
12,2020-03-12,12,HUILA,NEIVA,74,1,F,Leve
13,2020-03-12,13,HUILA,NEIVA,68,1,F,Leve
14,2020-03-13,14,VALLE,PALMIRA,48,1,M,Leve
15,2020-03-13,15,META,VILLAVICENCIO,30,1,F,Leve
16,2020-03-13,16,BOGOTA,BOGOTA,61,1,F,Leve
17,2020-03-14,17,BOGOTA,BOGOTA,73,1,F,Leve
18,2020-03-14,18,BOGOTA,BOGOTA,54,1,M,Leve
19,2020-03-14,19,BOGOTA,BOGOTA,54,1,F,Leve
```





TAREAS

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

- Crear un nuevo archivo llamado *analisis_covcol1.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: "Sexo;Edad en agnos;Concepto".
 2. Lea el archivo *COVCOLI.csv* línea por línea, y a medida que vaya leyendo, escriba en *analisis_covcol1.csv* lo siguiente:
 - El sexo del caso positivo que está leyendo (Deberá dejarlo tal cual como aparece en la base de datos)
 - Un punto y coma (";")
 - La edad en años del caso positivo que está leyendo
 - Si la persona tiene menos de un año de edad, pondrá "0".
 - Si la persona tiene más de un año de edad, pero su unidad de medida es en meses, se hace la conversión a años, y tomamos la parte entera; ejemplo: si la persona tiene 23 meses de edad, se guarda en el archivo CSV un "1"
 - Este debe ser un número entero, no puede haber decimales.
 - Un punto y coma (";")
 - Una cadena de texto que será un concepto respecto a la edad de la persona contagiada que está leyendo (Lo resaltado en rojo equivale a UN espacio en blanco):
 - Primera infancia: Desde los 0 años hasta los 5 años (Incluidos ambos)
 - Infante: Desde los 6 años hasta los 11 años (Incluidos ambos)
 - Adolescente: Desde los 12 años hasta los 17 años (Incluidos ambos)
 - Adulto: Desde los 18 años hasta los 59 años (Incluidos ambos)
 - Persona mayor: Desde los 60 años en adelante





Los primeros datos de su archivo creado deberían verse así:

Archivo Edición Formato Ver Ayuda

Sexo;Edad en agnos;Concepto

F;19;Adulto

M;34;Adulto

F;50;Adulto

M;55;Adulto

M;25;Adulto

F;27;Adulto

F;85;Persona mayor

F;22;Adulto

F;28;Adulto

F;36;Adulto

F;42;Adulto

F;74;Persona mayor

- Retornar los siguientes datos EN ESTE MISMO ORDEN:
 1. Edad de la persona con menor edad contagiada en este conjunto de datos.
 2. Unidad de medida de edad de la persona con menor edad contagiada en este conjunto de datos.
 3. Promedio de edad en años de las personas mayores de edad (Mayores de 18 años) contagiados en este conjunto de datos QUE NO FALLECIERON.





FORMATO DE ENTRADA

La función `solucion()` **NO** recibe parámetros, sin embargo, considere que el archivo `COVCOLI.csv` se encuentra ubicado en la misma carpeta donde usted está desarrollando el código de `variante1.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:

- **age_youngest:** Número entero (Objeto de la clase `int`) que contiene la edad (En la unidad de medida original) de la persona 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 más joven contagiada en el conjunto de datos dado (Puede ser 1, 2 o 3 como se especificó anteriormente).
- **mean_alive_g:** Número flotante (Objeto de la clase `float`) que contiene el promedio de edad (En años) de las personas mayores de edad que sobrevivieron al Covid - 19 (No fallecidos).

NOTAS:

1. No se preocupe por los decimales en el cálculo del promedio (**mean_alive_g**), el calificador considerará un margen de error debido a las aproximaciones.
2. Si la persona más joven del conjunto de datos dado fuera un bebé de 16 días de nacido, el valor de **age_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_covcol1.csv`) cumpliendo lo solicitado anteriormente.

