

## Parte 1

La SDS ha creado cuatro programas diferentes para generar atenciones de salud a personas en casa. Cada programa tiene procedimientos diferentes, por lo cual, las personas atendidas se reportan en cuatro bases de datos con diferente estructura. Estas bases de datos se en la carpeta adjunta denominada “Parte 1”.

Desde el despacho nos han solicitado generar una base de datos consolidada y normalizada de personas atendidas desde los tres programas, considerando una serie de variables de interés, únicamente:

- Localidad
- Sexo
- Edad
- Programa desde donde se atendió
- EAPB
- Fecha de atención o Caracterización

1.1. En este sentido, el aspirante deberá crear una vista (VISTA\_CONSOLIDADO) en SQL server, que genere una base unificada que consolidará las variables de interés desde las tres bases de cada programa. Así mismo, se espera que cada variable de interés obedezca al siguiente estándar:

- Localidad:

01 - Usaquén  
02 - Chapinero  
03 - Santa Fe  
04 - San Cristóbal  
05 - Usme  
06 - Tunjuelito  
07 - Bosa  
08 - Kennedy  
09 - Fontibón  
10 - Engativá  
11 - Suba  
12 - Barrios Unidos  
13 - Teusaquillo  
14 - Los Mártires  
15 - Antonio Nariño  
16 - Puente Aranda  
17 - La Candelaria  
18 - Rafael Uribe Uribe  
19 - Ciudad Bolívar  
20 - Sumapaz  
99 - Localidad Desconocida

- Sexo: HOMBRE, MUJER, INTERSEXUAL, OTROS, NO REGISTRA
- Edad: Años en número
- Programa: PROGRAMA 1, PROGRAMA 2, PROGRAMA 3
- EAPB:

SIN DATO  
 CAPITAL SALUD E.P.S.  
 NUEVA EPS S.A  
 SIN ASEGURAMIENTO  
 SALUD TOTAL S.A.  
 EPS Y MEDICINA PREPAGADA SURAMERICANA S.A  
 FERROCARRILES NACIONAL E.P.S.  
 SALUD BOLÍVAR E.P.S.  
 COMPENSAR E.P.S.  
 E.P.S. SANITAS  
 OTROS  
 FAMISANAR E.P.S. LTDA - CAFAM - COLSUBSIDIO  
 ALIANSALUD E.P.S.  
 COOSALUD E.P.S.  
 SOS E.P.S.  
 MALLAMAS E.P.S.

- Fecha de Caracterización: Tipo “DATE”, “AAAA-MM-DD”

**Se espera que el aspirante entregue el código SQL que consolida y una tabla en formato .txt que contenga la base de datos unificada.**

1.2. Por otra parte, el secretario nos ha solicitado generar un indicador poblacional de personas atendidas totales, desagregado por localidad, año y quinquenios. En este sentido el aspirante deberá realizar una nueva vista en SQL server (VISTA\_INDICADORES) donde utilice y agrupe, según se requiera para calcular las tasas, la VISTA\_COSNOLIDADO como numerador y la tabla POBLACION (adjunta en la carpeta Parte 1) como denominador que están dispuestas en el mismo conjunto de datos. Se espera que la VISTA\_INDICADORES entregue tres columnas: AÑO, LOCALIDAD, EDAD\_QUINQUENIOS, NUMERADOR, DENOMINADOR y TASA.

**Se espera que el aspirante entregue el código SQL utilizado y una tabla en formato .txt que contenga el resultado que de respuesta a la situación problema.**

## Parte 2

Adjunto al correo encuentras un archivo prueba.zip, donde vas a encontrar un instructivo en word y una carpeta que contiene dos fuentes de datos. El ejercicio consiste en validar

que las columnas de las fuentes de información cumplan con lo consignado en el instructivo. Dicha tarea es rutinaria y buscamos minimizar la intervención humana en dicho proceso.

En este orden de ideas, tu misión consiste en tres puntos:

1. Crear un script .py que te permita validar cada columna de acuerdo con lo consignado en el instructivo.
2. Ejecutar el script para las dos fuentes adjuntas y guardar los resultados de dichas validaciones.
3. corregir en las fuentes de información en la medida de lo posible., de acuerdo con el instructivo

**Una vez finalizado el ejercicio, se debe enviar el Sript .py y las fuentes corregidas.**