

PRIMER ENTREGA – PROYECTO DATA WAREHOUSE

DATA WAREHOUSE

ANDRES SANCHEZ



Bona Health EPS


**LUIS FELIPE VELASCO TAO
JUAN DAVID GONZALEZ**

UNIVERSIDAD DE SAN BUENAVENTURA

BOGOTÁ

11 DE AGOSTO


2021

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
--	---	-----------------	-----------------

Contenido

Bona Health EPS	3
Contexto de la Bona Health EPS.....	3
Misión	3
Visión.....	3
KPI.....	3
Requerimientos de negocio	7
Fuentes elegidas	8
Mortalidad VIH 2010 a 2016.....	8
Mortalidad cáncer de pulmón 2010 a 2016	9
Mortalidad cáncer de mama 2010 a 2016	9
Modelo Entidad Relación	16



 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

Bona Health EPS

A continuación se van a definir los elementos corporativos de la empresa y se definirán las fuentes de datos elegidas para el desarrollo de la actividad.

Contexto de la Bona Health EPS

Bona Health EPS es una entidad promotora de salud colombiana la cual tienen presencia a nivel nacional, principalmente en las ciudades capitales, en las cuales nuestros afiliados acceden a los servicios de salud prestados por medio de las IPS asociadas, entre las cuales se tienen hospitales, centro de optometría, odontología, radiografía y de atención para distintas afecciones medicas de los colombianos. Nuestra empresa nació en el 2015, teniendo como punto de partida la ciudad de Bogotá, pasando año tras año a tener presencia en Cali, Medellín, Cartagena, Bucaramanga y demás ciudades principales del país. En la actualidad, debido a la pandemia, Bona Health EPS ha tenido inconvenientes con los procesos de afiliación y especialmente con el manejo de citas y tratamientos de nuestros afiliados con enfermedades como VIH, cáncer y demás problemas de salud, llegando hasta el limite de notar cierto aumento en la mortalidad de muchos afiliados con ciertas enfermedades.

Misión

Asegurarles a los colombianos desde el primer día que se encuentren afiliados el acceso a servicios de salud de calidad, a tiempo y siempre buscando su bienestar.


Visión

Para el 2030, Bona Health EPS espera ser una EPS que tenga una cobertura del 80% en municipios apartados en el país con el fin de acercar a más a los pueblos y a sus habitantes servicios de salud de calidad, sea medio del régimen contributivo o subsidiado.

KPI

Con el fin de proveer un servicio que sea capaz de atender a su debido tiempo las necesidades de los colombianos, es necesario determinar la tasa de mortalidad de determinados padecimientos en un intervalo de tiempo, y la frecuencia de defunciones en un lugar con respecto a dicho padecimiento.

Es por esto por lo que se establecieron los siguientes indicadores para determinar el tipo de servicio que debe ofrecerse en los departamentos más afectados por un determinado padecimiento.

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

1. **Tasa de defunciones por VIH, cáncer de mama y cáncer de pulmón.** A continuación, se muestran los datos de las defunciones presentadas en el periodo del 2010 a 2016 con el fin de definir medidas para la prevención y atención oportuna de dichas enfermedades.

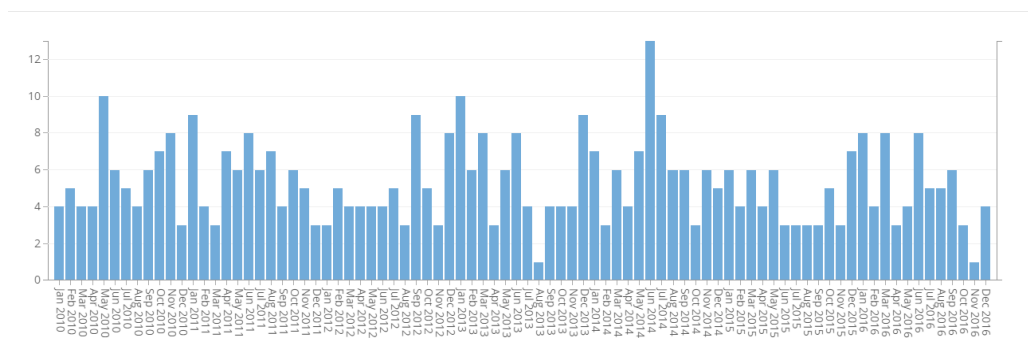


Ilustración 1 Defunciones por VIH

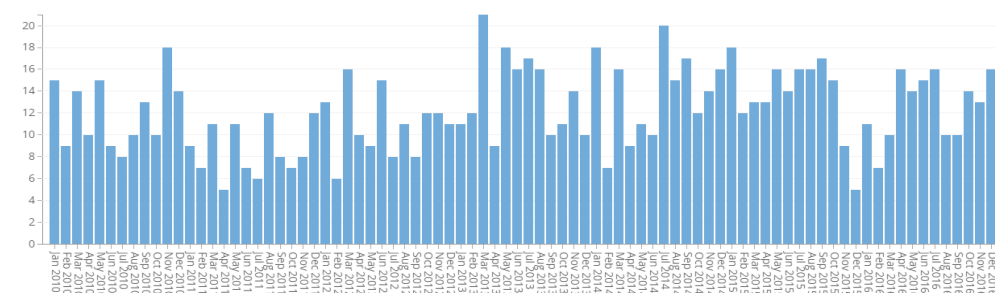


Ilustración 2 Defunciones por cáncer de mama

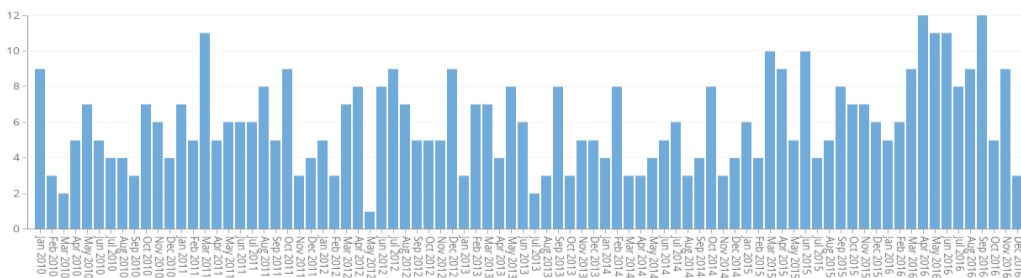



Ilustración 3 Defunciones por cáncer de pulmón

2. **Tasa de atención médica:** en las defunciones registradas. Se quiere definir si los fallecidos tuvieron o no atención medica en el momento, de modo tal se pueda definir acciones que mejoren la atención medica en futuros casos.

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

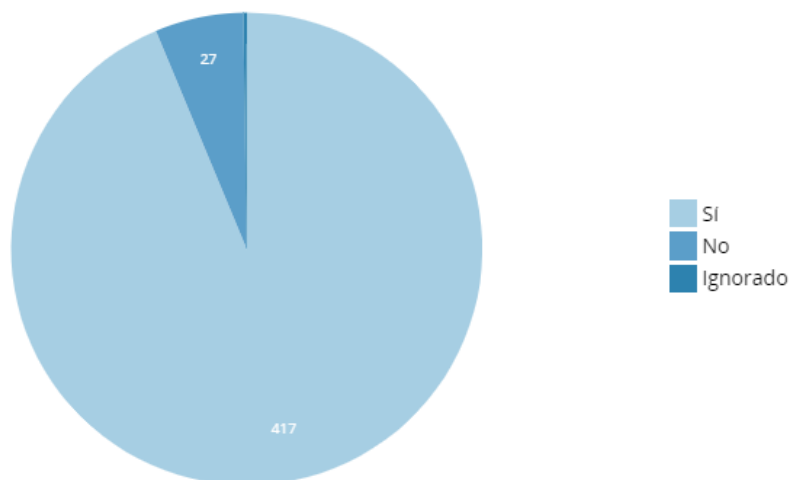


Ilustración 4 Asistencia medida prestada a los fallecidos por VIH

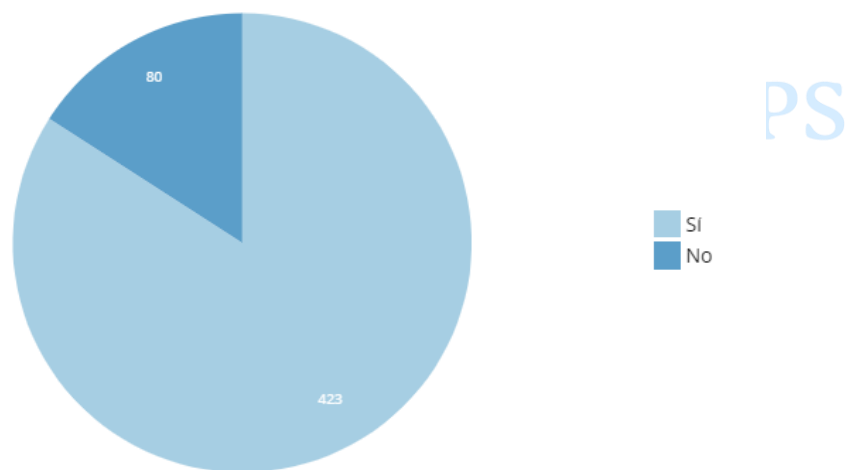



Ilustración 5 Atención medica prestada a los fallecidos por cáncer de mama

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

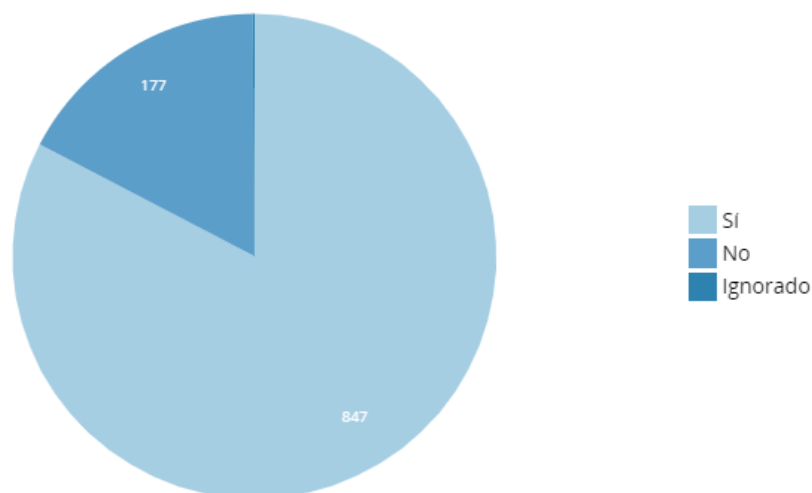


Ilustración 6 Atención medica prestada a los fallecidos por cáncer de pulmón

3. **Defunciones por municipio y departamento:** con el fin definir medidas y mecanismos de acción, se quiere medir cuantas personas fallecieron por zonas.

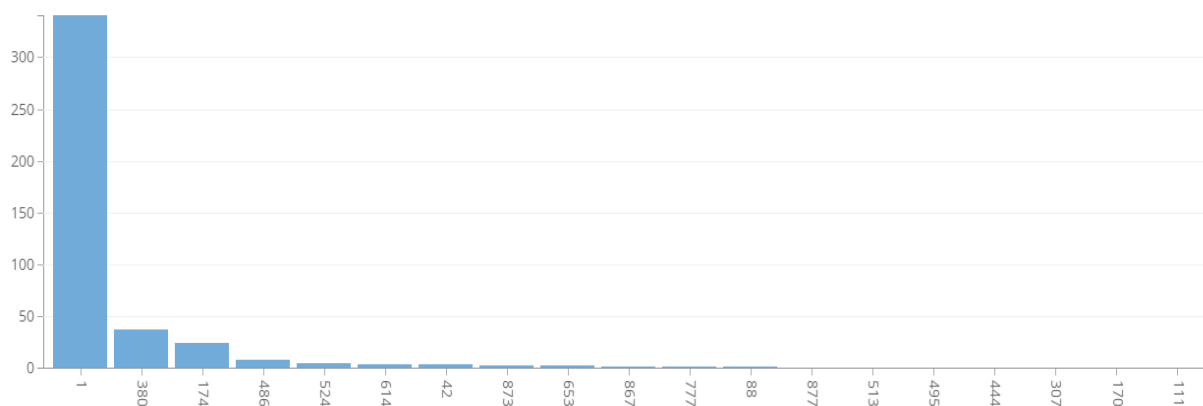



Ilustración 7 defunciones por VIH clasificadas por municipio

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
--	---	-----------------	-----------------

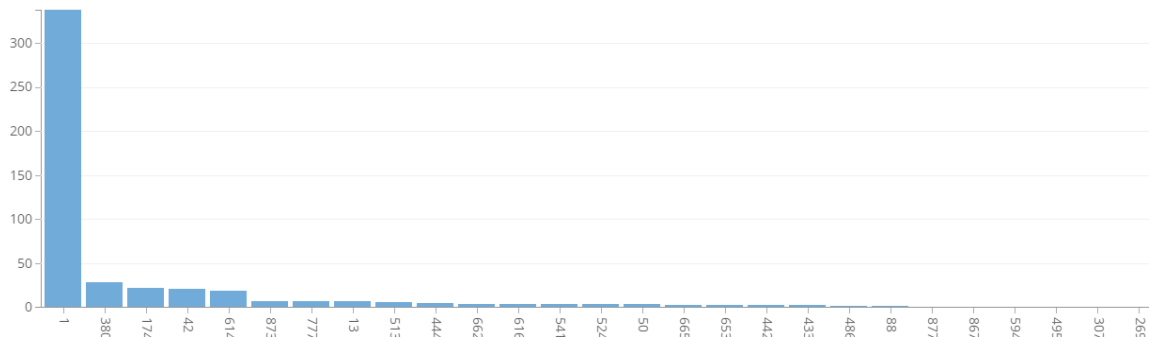


Ilustración 9 defunciones por cáncer de mama clasificadas por municipio

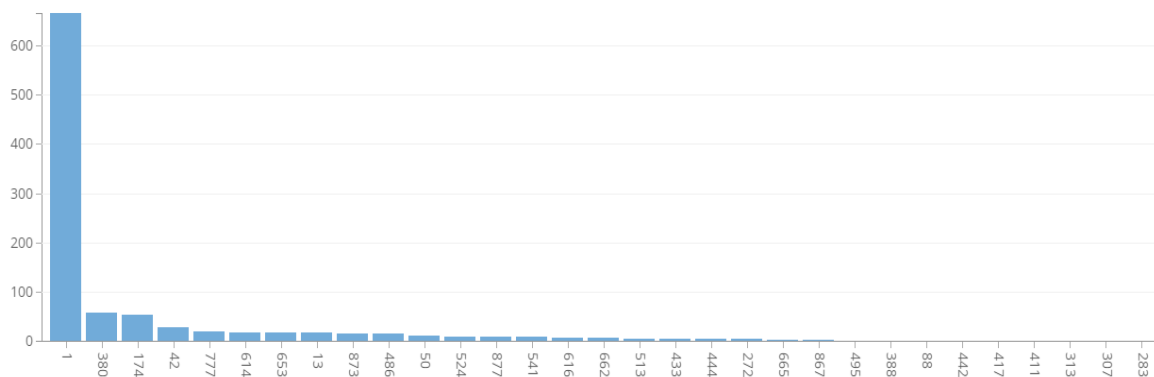



Ilustración 8 defunciones por cáncer de pulmón clasificadas por municipio

Requerimientos de negocio

Bona Health EPS busca ofrecer un servicio accesible que cubra las zonas con mayor tasa de mortalidad por padecimientos clínicos en el país, tales como el VIH, el cáncer de Mama y el cáncer de pulmón. Con el fin de realizar estos objetivos, se plantean los siguientes requerimientos.

1. Analizar la tasa de mortalidad por padecimiento clínico con respecto a los departamentos del país para proveer cobertura en los municipios en dónde se requieran los servicios de la entidad.
2. Establecer un histórico de defunciones por padecimientos clínicos en cada departamento para la asignación del personal adecuado en las instalaciones que tengan cobertura de la entidad.

 UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ	Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse	PROYECTO	2021 - 2
--	---	-----------------	-----------------

- Determinar el número de difuntos que recibieron atención médica en una institución clínica con el fin de poder asignar los recursos correspondientes a las instituciones cubiertas por la entidad.

Fuentes elegidas

Para la correcta toma de decisiones con relación a la mejora en la prestación de servicios de salud que prevengan la mortalidad de los afiliados con enfermedades críticas como el VIH, el cáncer de mama y de pulmón se seleccionaron las siguientes fuentes de datos, las cuales son los reportes emitidos de defunción relacionados con los afiliados con estas enfermedades, además de requerir de cierta información para definir apropiadamente el lugar (departamento y municipio) donde falleció la persona:

Mortalidad VIH 2010 a 2016

Este dataset es suministrado por Observatorio Social de Salud Pública en el cual se entregan datos relacionados con las defunciones de las personas con VIH en el periodo del 2010 a 2016, en el cual se pueden ver los datos relacionados con el lugar en donde falleció la persona, como lo es municipio, departamento institución en la que falleció, también se presentan datos básicos del fallecido como su sexo y condición socioeconómica y finalmente se define una descripción de la condición medica de la persona al fallecer, incluyendo sus antecedentes de salud y datos relacionados con el estado de salud de la persona.

Fecha de creación: 21 de abril de 2017


Sector: Salud y protección social

Cantidad de registros: 445

Cantidad de columnas: 105

COD_...	COD_...	A_DE...	COD_...	SIT_D...	OTRS...	COD_...	NOM...	TIPO...	FECH...	ANO	MES	HOF
17	1	Cabecera...		Hospital/c...		170.010.0...	ESE HOSP...	2	2012 Sep ...	2.012	9	
17	1	Cabecera...		Hospital/c...		170.010.1...	IPS CAPR...	2	2012 Mar...	2.012	3	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...	2	2011 Oct ...	2.011	10	
17	42	Cabecera...		Casa/do...		170.420.0...	ESE HOSP...	2	2012 Nov ...	2.012	11	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...	2	2012 Jan ...	2.012	1	
17	1	Cabecera...		Hospital/c...		170.010.0...	ESE HOSP...	2	2013 Dec ...	2.013	12	
76	1	Cabecera...		Hospital/c...		760.010.4...	CLINICA R...	2	2013 Feb ...	2.013	2	
17	42	Cabecera...		Hospital/c...		170.420.0...	ESE HOSP...	2	2013 Dec ...	2.013	12	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...	2	2013 Nov ...	2.013	11	
17	42	Cabecera...		Hospital/c...		170.420.0...	ESE HOSP...	2	2013 Oct ...	2.013	10	
17	42	Cabecera...		Hospital/c...		170.420.0...	ESE HOSP...	2	2013 Sep ...	2.013	9	
63	1	Cabecera...		Hospital/c...		630.010.0...	EMPRESA...	2	2013 May...	2.013	5	
17	1	Cabecera...		Hospital/c...		170.010.0...	CAJA DE ...	2	2014 Apr ...	2.014	4	

Ilustración 10 Muestra de datos del dataset de mortalidad por VIH

 UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ	Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse	PROYECTO	2021 - 2
--	---	-----------------	-----------------

Fuente: [Mortalidad VIH 2010 A 2016 | Datos Abiertos Colombia](#)

Mortalidad cáncer de pulmón 2010 a 2016

Este dataset es suministrado por Observatorio Social de Salud Pública en el cual se entregan datos relacionados con las defunciones de las personas con cáncer de pulmón en el periodo del 2010 a 2016, en el cual se pueden ver los datos relacionados con el lugar en donde falleció la persona, como lo es municipio, departamento institución en la que falleció, también se presentan datos básicos del fallecido como su sexo y condición socioeconómica y finalmente se define una descripción de la condición médica de la persona al fallecer, incluyendo sus antecedentes de salud y datos relacionados con el estado de salud de la persona.

Fecha de creación: 20 de abril de 2017

Sector: Salud y protección social

Cantidad de registros: 1025

Cantidad de columnas: 105


COD...	COD...	A_DE...	COD...	SIT_D...	OTRS...	COD...	NOM...	TIPO...	FECH...	ANO	MES
17	13	Cabecera...		Casa/do...					2 2010 Jan ...	2.010	
17	13	Cabecera...		Hospital/c...		170.130.0...	ESE HOSP...		2 2010 Sep ...	2.010	
17	13	Rural disp...		Hospital/c...		170.130.0...	ESE HOSP...		2 2010 Nov ...	2.010	
17	1	Cabecera...		Hospital/c...		170.010.0...	ESE HOSP...		2 2010 Jul 1 ...	2.010	
17	13	Centro po...	1	Centro/pu...		170.130.0...	ESE HOSP...		2 2011 May...	2.011	
17	13	Rural disp...		Casa/do...		170.130.0...	ESE HOSP...		2 2011 Mar...	2.011	
17	13	Centro po...	16	Casa/do...		170.130.0...	ESE HOSP...		2 2011 Mar...	2.011	
17	13	Centro po...	1	Casa/do...		170.130.0...	ESE HOSP...		2 2011 Mar...	2.011	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...		2 2011 Jan ...	2.011	
17	13	Cabecera...		Casa/do...		170.130.0...	ESE HOSP...		2 2012 Dec ...	2.012	
5	1	Cabecera...		Hospital/c...		50.010.21...	E.S.E. HO...		2 2012 Sep ...	2.012	
76	1	Cabecera...		Hospital/c...		760.010.0...	IPS COMF...		2 2012 May...	2.012	
17	13	Cabecera...		Casa/do...		170.130.0...	ESE HOSP...		2 2012 Jan ...	2.012	

Ilustración 11 Muestra de datos del dataset de mortalidad por cáncer del pulmón

Fuente: [Mortalidad Cáncer de Pulmón 2010 A 2016 | Datos Abiertos Colombia](#)

Mortalidad cáncer de mama 2010 a 2016

Este dataset es suministrado por Observatorio Social de Salud Pública en el cual se entregan datos relacionados con las defunciones de las personas con cáncer de mama en el periodo del 2010 a 2016, en el cual se pueden ver los datos relacionados con el lugar en donde falleció la persona, como lo es municipio,

 UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ	Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse	PROYECTO	2021 - 2
--	---	-----------------	-----------------

departamento institución en la que falleció, también se presentan datos básicos del fallecido como su sexo y condición socioeconómica y finalmente se define una descripción de la condición médica de la persona al fallecer, incluyendo sus antecedentes de salud y datos relacionados con el estado de salud de la persona.

Fecha de creación: 20 de abril de 2017

Sector: Salud y protección social

Cantidad de registros:503

Cantidad de columnas: 105

Vista previa de la tabla Explorar los datos Crear visualización


COD...	COD...	A_DE...	COD...	SIT_D...	OTRS...	COD...	NOM...	TIPO...	FECH...	ANO	MES
17	1	Cabecera...		Hospital/c...		170.010.0...	SERVICIO...	2	2010 Apr ...	2.010	
	13	Rural disp...		Casa/do...		170.130.0...	ESE HOSP...	2	2010 Oct ...	2.010	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...	2	2010 Nov ...	2.010	
17	1	Cabecera...		Hospital/c...		170.010.0...	SERVICIO...	2	2010 Apr ...	2.010	
17	1	Cabecera...		Hospital/c...		170.010.1...	INSTITUT...	2	2011 Oct ...	2.011	
17	1	Cabecera...		Hospital/c...		170.010.0...	ESE HOSP...	2	2011 Mar...	2.011	
17	13	Cabecera...		Casa/do...		170.130.0...	ESE HOSP...	2	2011 Mar...	2.011	
17	13	Cabecera...		Hospital/c...		170.130.0...	ESE HOSP...	2	2012 Jul 0...	2.012	
66	1	Cabecera...		Hospital/c...		660.010.1...	CORPORA...	2	2013 Sep ...	2.013	
17	13	Centro po...	16	Casa/do...		170.130.0...	ESE HOSP...	2	2014 Sep ...	2.014	
17	1	Cabecera...		Hospital/c...		170.010.0...	ESE HOSP...	2	2014 Apr ...	2.014	
17	13	Centro po...	10	Casa/do...		170.130.0...	ESE HOSP...	2	2015 Dec ...	2.015	
17	13	Cabecera...		Casa/do...		170.130.0...	ESE HOSP...	2	2015 Jul 1...	2.015	

Ilustración 12 Muestra de datos del dataset de mortalidad por cáncer de mama


Fuente: [Mortalidad Cáncer Mama 2010 A 2016 | Datos Abiertos Colombia](#)

Debido a que los datasets se encuentran estructurados bajo las mismas columnas, se realizó un análisis de las columnas de los datasets y se brinda una descripción de los siguientes datos relevantes.


Nombre de Columna	Descripción	Tipo
COD_DPTO	Identificador del departamento donde falleció la persona	Número

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------


<p>COD_MUNIC</p>	<p>Identificador del municipio donde falleció la persona. Este valor se puede repetir, ya que, dos o más municipios pueden tener el mismo código, la diferencia es el código del departamento</p>	<p>Número</p>
<p>A_DEFUN</p>	<p>Área del municipio donde falleció la persona (rural, cabecera municipal, etc.)</p>	<p>Texto simple</p>
<p>SIT_DEFUN</p>	<p>Sitio en donde la persona falleció, como su casa o un hospital/clínica</p>	<p>Texto simple</p>
<p>COD_INST</p>	<p>Identificador de la institución prestadora de servicios de salud donde falleció la persona. (Si la persona fallece en su casa, este campo se encuentra vacío)</p>	<p>Número</p>
<p>NOM_INST</p>	<p>Nombre de la institución prestadora de servicios de salud en donde la persona halla fallecido.</p>	<p>Texto simple</p>

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------


FECHA_DEF	Fecha y hora en la cual falleció la persona, expresada de la siguiente forma: 2012 Sep 14 12:00:00 AM	Fecha y hora
ANO	Año en el que falleció la persona	Número
MES	Mes en el que falleció la persona	Número
HORA	Hora en la que falleció la persona	Número
MINUTOS	Minuto en que falleció la persona	Número
SEXO	Cadena de caracteres con el sexo de la persona (Masculino/Femenino)	Texto simple
FECHA_NAC	Fecha (y hora la cual no es un dato relevante) de nacimiento del fallecido	Fecha y hora
EST_CIVIL	Estado civil del fallecido	Texto simple
EDAD	Edad del fallecido	Número

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

NIVEL_EDU	Nivel educativo del fallecido	Texto simple
MUERTEPORO	Campo de texto en donde se define si la persona murió o no por omisión	Texto simple
OCUPACION	Ocupación o profesión del fallecido.	Texto simple
CODPTORE	Nombre del departamento donde residía el fallecido	Texto simple
CODMUNRE	Nombre del municipio donde residía el fallecido	Texto simple
AREA_RES	Área del municipio donde residía la persona (rural, cabecera municipal, etc.)	Texto simple
BARRIOFAL	Nombre del barrio donde residía el fallecido	Texto simple
SEG_SOCIAL	Tipo de seguridad social por medio de la cual se le presto los servicios de salud al fallecido.	Texto simple

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------


IDADMISALU	Identificador de la EPS (este dato se obvia para la naturaleza del ejercicio)	Número
IDCLASADMI	EPS a la cual el fallecido estaba afiliado (este dato se obvia para la naturaleza del ejercicio)	Texto simple
PMAN_MUER	Tipo de muerte	Texto simple
MU_PARTO	Campo en donde se define si la muerte fue en un parto o no	Texto simple
T_PARTO	Si la muerte fue por un parto, se define el tipo de parto. Se definen otros datos relacionados con muertes en partos	Texto simple
ASIS_MED	Campo en donde se define si la persona tuvo asistencia médica en el momento de su muerte	Texto simple
Los siguientes datos son descripción de la condición médica del fallecido		
N_DIR1 C_DIR1	Datos sobre antecedentes, signos	Texto simple

 UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ	Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse	PROYECTO	2021 - 2
--	---	----------	----------

C_DIR12 N_ANT1 C_ANT1 C_ANT12 N_ANT2 C_ANT2 C_ANT22 N_ANT3 C_ANT3 C_ANT32 N_PAT1 C_PAT1 N_PAT2 C_PAT2	vitales y condición médica del fallecido	
N_BAS1	Enfermedad de base del fallecido	Texto simple

TOTAL, DE REGISTROS A MANIPULAR: 1975


Con base al análisis a cada una de las fuentes de datos, la descripción brindada a las columnas de interés y la definición de los requerimientos y KPIS, se definirá un único modelo entidad relación en el cual se puedan estructurar los datos vistos.

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

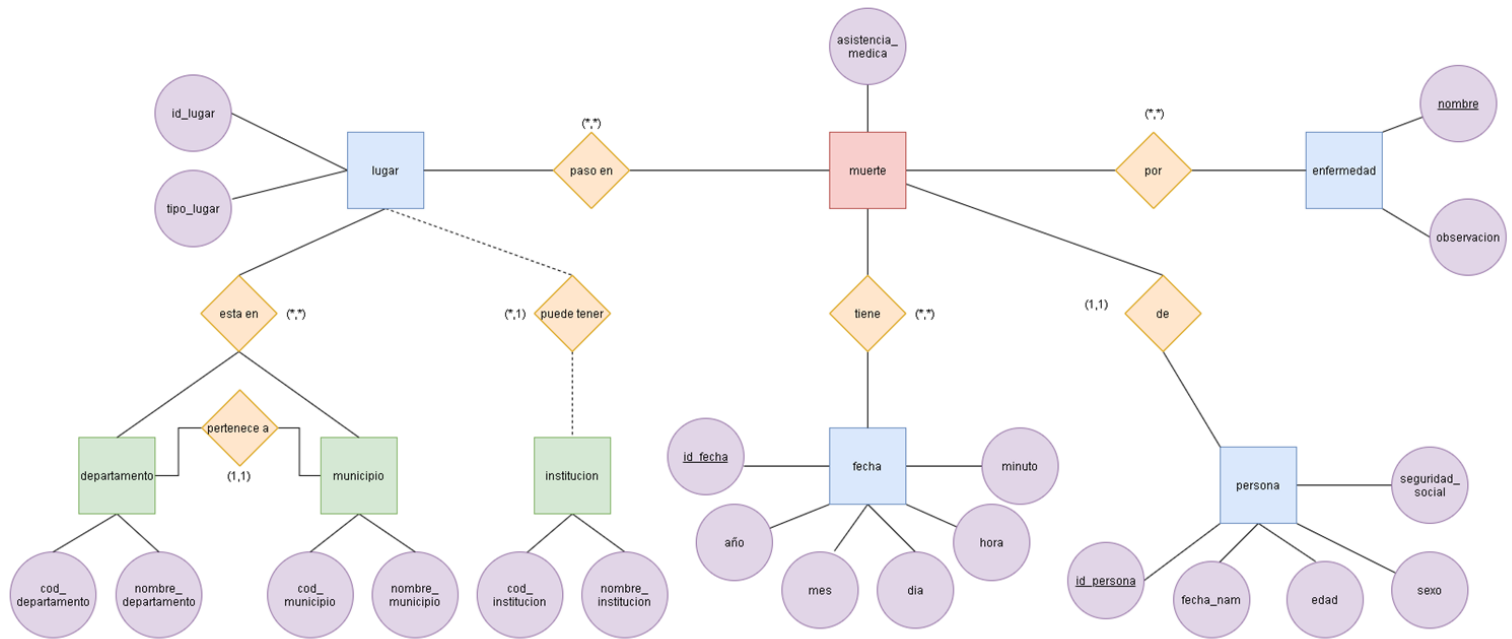
Modelo Entidad Relación

Debido a como las tres fuentes de datos se encuentran estructuradas bajo las mismas columnas, se obtuvo un único modelo entidad relación en el cual se definieron las siguientes tablas:


- **Lugar:** se definen los datos del lugar donde la persona falleció, y se estructura por medio de los siguientes datos:
 - Id_lugar: dato identificador del lugar*
 - Tipo_lugar: se define si la persona murió o no en un hospital o clínica.
 - Departamento: este dato se obtendrá de la tabla departamento, y se definiría el departamento en el cual falleció la persona.
 - Municipio: este dato se obtendrá de la tabla municipio, y se definiría el municipio en el cual falleció la persona.
 - Institución: con base al tipo lugar, se define la institución obtenida de la tabla del mismo nombre.
- **Fecha:** se definen los datos de la fecha y hora registrada del fallecimiento de la persona, esto por medio de los siguientes datos:
 - Id_fecha: fecha y hora formateada del fallecimiento de la persona.
 - Año: valor numérico del año del fallecimiento de la persona.
 - Mes: valor numérico del mes del fallecimiento de la persona.
 - Día: valor número del día del fallecimiento de la persona.
 - Hora: valor numérico de la hora de fallecimiento de la persona
 - Minutos: valor numérico del minuto de fallecimiento de la persona
- **Persona:** en esta tabla se definen los datos básicos de la persona, de modo tal se definen los siguientes datos:
 - Id_persona: identificador de la persona.
 - Fecha_nacimiento: fecha de nacimiento del fallecido
 - Edad: valor numero de la edad del fallecido
 - Sexo: cadena de caracteres con el sexo de la persona.
 - Seguridad_social: cadena de caracteres en donde se define el régimen por medio del cual se le prestó atención a la persona (contributivo, subsidiado)
- **Enfermedad:** En esta tabla solo se definirá el nombre de la enfermedad de base que produjo el fallecimiento de la persona, por lo que se define los siguientes datos:
 - Nombre_enfermedad: nombre de la enfermedad de base que genero el fallecimiento de la persona.
 - Observación: observación brindada por los médicos al momento del fallecimiento de la persona.

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

A continuación, se muestra el modelo resultante:



BOLIVIA HEALTH ERS
Ilustración 13 Modelo entidad relación de las fuentes de datos
Cuidando De Ti

 <p>UNIVERSIDAD DE SAN BUENAVENTURA BOGOTÁ</p>	<p>Universidad de San Buenaventura Facultad de ingeniería Optativa II: Data Warehouse</p>	<p>PROYECTO</p>	<p>2021 - 2</p>
---	---	-----------------	-----------------

