

# Propuesta proyecto Final: Estadísticas Vitales Colombia 2022

Juan Manuel Hios Gaviria

- **Descripción de la temática de los datos**

Para el desarrollo del análisis, se extrajo información del repositorio de **Estadísticas Vitales** del Departamento Administrativo Nacional de Estadística de Colombia (**DANE**). La entidad define las Estadísticas vitales como un proceso que reúne información mediante un registro y reporta la frecuencia o la ocurrencia de acontecimientos vitales específicos y definidos por el sistema (Nacimientos y defunciones en Colombia), así como las características propias de los hechos vitales.

Para el análisis se toman particularmente los **Datos de nacimientos en Colombia**; los cuales el DANE consolida, valida, procesa y difunde la información de nacimientos, a partir de los certificados diligenciados en medio físico o digital, por médicos, o por personal autorizado donde se reportan las características principales que identifican el hecho: edad, sexo, departamento y municipio de ocurrencia del hecho y de residencia de la madre, y que son registradas en el aplicativo RUAF-ND del Ministerio de Salud y Protección Social.

De la misma manera se toma información de las **Defunciones fetales en Colombia**; las cuales se entienden como la muerte del producto de la concepción antes de su expulsión o extracción completa del cuerpo de la madre, independientemente de la duración del embarazo. A partir de los Certificados de Defunción, diligenciados en medio físico o digital, por médicos y personal de salud autorizado (enfermeras, auxiliares de enfermería y promotores de salud) en donde se registran datos por área y sexo, departamento y municipio de ocurrencia y de residencia de la madre, sitio de defunción, edad gestacional y peso, edad de la madre, tiempo de gestación y causas de muerte.

Por último, se toma el Producto Interno Bruto (**PIB**) departamental y la tasa de crecimiento, el cual mide la actividad productiva de los diferentes departamentos del país, además establece su comportamiento, evolución y estructura económica para el análisis y la toma de decisiones regionales.

Los datos de los nacimientos y las defunciones fetales en Colombia corresponden a los registrados entre el 01 de enero y el 31 de diciembre de 2022, mientras que los registros del PIB corresponden al histórico desde el año 2005 hasta 2022 a precios corrientes del 2015.

- **Hipótesis – Objetivos**

Identificar el comportamiento de la estructura de la fecundidad y mortalidad a lo largo y ancho del territorio nacional, determinando patrones en las tasas de natalidad y mortalidad a nivel departamental y municipal, así como las posibles causas asociadas a la disparidad de valores entre las diferentes jurisdicciones y variables sociodemográficas como el nivel educativo y edad de la madre, etnia asociada, semanas de gestación, enfermedades relacionadas y causa de muerte.

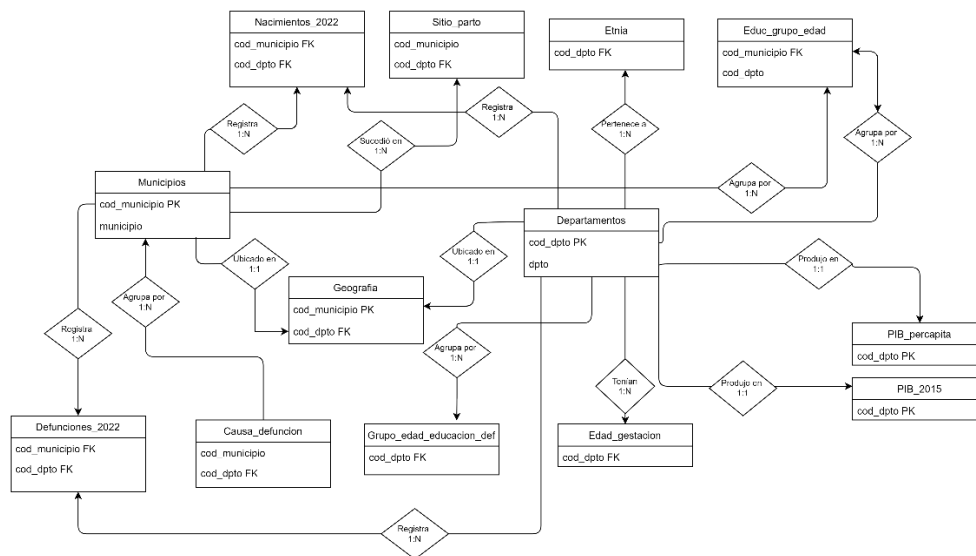
Con el análisis, también se busca entender la dinámica de los nacimientos de acuerdo con el sexo biológico del bebe, la concentración de nacimientos en los diferentes entornos y profesionales encargados de la asistencia, así como la distribución de nacimientos por población urbana y rural.

Por último, se busca establecer la posible relación de causalidad entre la evolución de la producción interna bruta (PIB) departamental, PIB per cápita departamental y el comportamiento de las variables asociadas a la fecundidad y mortalidad de cada una de las jurisdicciones. Con la finalidad de poder evaluar políticas en salud, educación, economía o infraestructura que puedan mitigar casos de proliferación de enfermedades o fenómenos como el embarazo adolescente.

## • Tipo de análisis de datos

- **Descriptivo:** Describir el comportamiento tanto de los nacimientos como de las defunciones fetales durante el año 2022 en el territorio colombiano.
- **Diagnóstico:** Identificar patrones o comportamientos ligados a la geografía, la asistencia del parto o la educación de la madre en los nacimientos y defunciones, así como la incidencia de la producción económica en estos fenómenos.
- **Prescriptivo:** Identificar acciones que permitan establecer políticas publicas par mitigar posibles comportamientos atípicos ligados a enfermedades o atención en el sistema de salud de ser posible.

## • Diagrama Entidad - Relación



## • Tablas y entidades

A continuación, se detalla para cada una de las tablas, la definición, los campos que contienen, el tipo de dato de cada uno de los campos, el tipo de clave (PK o FK) si aplica y los valores obligatorios que pueden registrar

- **Nacimientos 2022:** Nacimientos por área de ocurrencia y sexo, según grupos de edad de la madre, total nacional.

Nacimientos_2022			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
hombre_total	INT	-	NULL
mujer_total	INT	-	NULL
indeterminado_total	INT	-	NULL
total	INT	-	NULL
hombre_cab_mun	INT	-	NULL
mujer_cab_mun	INT	-	NULL
indeterminado_cab_mun	INT	-	NULL
hombre_centro_poblado	INT	-	NULL
mujer_rural	INT	-	NULL
indeterminado_centro_poblado	INT	-	NULL
hombre_rural_disperso	INT	-	NULL
mujer_rural_disperso	INT	-	NULL
indeterminado_rural_disperso	INT	-	NULL

- **Sitio de parto:** Nacimientos por persona que atendió el parto según departamento, municipio de ocurrencia y sitio del parto.

Sitio de parto			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
sitio_parto	VARCHAR	-	NULL
Total	INT	-	NULL
Médico	INT	-	NULL
Enfermero(a)	INT	-	NULL
Auxiliar de Enfermería	INT	-	NULL
Promotor(a) de salud	INT	-	NULL
Partera	INT	-	NULL
Otra persona	INT	-	NULL
Sin información	INT	-	NULL

- **Etnia:** Nacimientos por área y sexo, según departamento de residencia de la madre y pertenencia étnica del nacido vivo.

Etnia			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
pertenencia_etnica	VARCHAR	-	NOT NULL
total	INT	-	NULL

- **Educación grupo edad:** Nacimientos por nivel educativo de la madre según departamento, municipio de residencia y grupo de edad de la madre.

Educ_grupo_edad			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
edad_grupo	VARCHAR	-	NOT NULL
total	INT	-	NULL
Preescolar	INT	-	NULL
Básica primaria	INT	-	NULL
Básica secundaria	INT	-	NULL
Media académica o clásica	INT	-	NULL
Media técnica	INT	-	NULL
Normalista	INT	-	NULL
Técnica profesional	INT	-	NULL
Tecnológica	INT	-	NULL
Profesional	INT	-	NULL
Especialización	INT	-	NULL
Maestría	INT	-	NULL
Doctorado	INT	-	NULL
Ninguno	INT	-	NULL
Sin información	INT	-	NULL

- **Geografía:** información de ubicación geográfica de Latitud y Altitud para cada uno de los municipios y departamentos de Colombia.

Geografía			
Campo	Tipo de campo	Tipo de clave	Obligatorio
pais	VARCHAR	-	NOT NULL
cod_dpto	INT	FK	NOT NULL
dpto	VARCHAR	-	NOT NULL
cod_municipio	INT	FK	NOT NULL
municipio	VARCHAR	-	NOT NULL
LATITUD	FLOAT	-	NOT NULL
LONGITUD	FLOAT	-	NOT NULL

- **Defunciones fetales 2022:** Defunciones fetales por área y sexo, según departamento y municipio de ocurrencia.

Defunciones_fetales_2022			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
total	INT	-	NULL
hombre_total	INT	-	NULL
mujer_total	INT	-	NULL
indeterminado_total	INT	-	NULL
hombre_cab_mun	INT	-	NULL
mujer_cab_mun	INT	-	NULL
indeterminado_cab_mun	INT	-	NULL
hombre_centro_poblado	INT	-	NULL
mujer_rural	INT	-	NULL
indeterminado_centro_poblado	INT	-	NULL
hombre_rural_disperso	INT	-	NULL
mujer_rural_disperso	INT	-	NULL
indeterminado_rural_disperso	INT	-	NULL

- **Causa defunción:** Defunciones fetales por sexo, según departamento, municipio de residencia de la madre y grupos de causas de defunción (Lista de causas agrupadas 6/67 CIE-10 de OPS)"

Causa_defuncion			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
enfermedad	VARCHAR	-	NULL
total	INT	-	NULL
hombre	INT	-	NULL
mujer	INT	-	NULL
indeterminado	INT	-	NULL

- **Grupo edad y educación defunciones:** Defunciones fetales por grupos de edad de la madre, según departamento de residencia y nivel educativo de la madre.

Grupo_edad_educacion_def			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
educación	VARCHAR	-	NOT NULL
total	INT	-	NULL
10 - 14 años	INT	-	NULL
15 - 19 años	INT	-	NULL
20 - 24 años	INT	-	NULL
25 - 29 años	INT	-	NULL
30 - 34 años	INT	-	NULL
35 - 39 años	INT	-	NULL
40 - 44 años	INT	-	NULL
45 - 49 años	INT	-	NULL
50 - 54 años	INT	-	NULL
Edad desconocida	INT	-	NULL

- **Edad gestación:** Defunciones fetales por tiempo de gestación, según departamento de residencia y grupos de edad de la madre.

Edad_gestacion			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	FK	NOT NULL
edad	VARCHAR	-	NOT NULL
total	INT	-	NULL
Menos de 22	INT	-	NULL
De 22 a 27	INT	-	NULL
De 28 a 36	INT	-	NULL
De 37 y más	INT	-	NULL
Ignorado	INT	-	NULL
Sin información	INT	-	NULL

- **PIB 2015:** Producto Interno Bruto por departamento, a precios corrientes. Serie 2005-2022.

PIB_2015			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	PK	NOT NULL
2005	FLOAT	-	NOT NULL
2006	FLOAT	-	NOT NULL
2007	FLOAT	-	NOT NULL
2008	FLOAT	-	NOT NULL
2009	FLOAT	-	NOT NULL
2010	FLOAT	-	NOT NULL
2011	FLOAT	-	NOT NULL
2012	FLOAT	-	NOT NULL
2013	FLOAT	-	NOT NULL
2014	FLOAT	-	NOT NULL
2015	FLOAT	-	NOT NULL
2016	FLOAT	-	NOT NULL
2017	FLOAT	-	NOT NULL
2018	FLOAT	-	NOT NULL
2019	FLOAT	-	NOT NULL
2020	FLOAT	-	NOT NULL
2021	FLOAT	-	NOT NULL
2022	FLOAT	-	NOT NULL

- **Municipios:** Código DANE y nombre de cada uno de los municipios de Colombia.

Municipios			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_municipio	INT	PK	NOT NULL
municipio	VARCHAR	-	NOT NULL

- **Departamentos:** Código DANE y nombre de cada uno de los departamentos de Colombia.

Departamento			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	PK	NOT NULL
dpto	VARCHAR	-	NOT NULL

- **PIB 2015:** Producto Interno Bruto departamental por habitante, a precios corrientes. Serie 2005-2022.

PIB_percapita			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_dpto	INT	PK	NOT NULL
2005	FLOAT	-	NOT NULL
2006	FLOAT	-	NOT NULL
2007	FLOAT	-	NOT NULL
2008	FLOAT	-	NOT NULL
2009	FLOAT	-	NOT NULL
2010	FLOAT	-	NOT NULL
2011	FLOAT	-	NOT NULL
2012	FLOAT	-	NOT NULL
2013	FLOAT	-	NOT NULL
2014	FLOAT	-	NOT NULL
2015	FLOAT	-	NOT NULL
2016	FLOAT	-	NOT NULL
2017	FLOAT	-	NOT NULL
2018	FLOAT	-	NOT NULL
2019	FLOAT	-	NOT NULL
2020	FLOAT	-	NOT NULL
2021	FLOAT	-	NOT NULL
2022	FLOAT	-	NOT NULL

## • Datasets

En el enlace adjunto, se encuentra el repositorio en GitHub asociado la entregar preliminar del proyecto final:

<https://github.com/jhios98/Curso-Data-Analytics>

- **Documentación:** Estadisticas\_Vitales\_Juan\_Manuel\_Hios\_Documentacion
- **Dataset:** Estadisticas\_Vitales\_Juan\_Manuel\_Hios\_Datset