Propuesta proyecto Final: Estadísticas Vitales Colombia 2022 Juan Manuel Hios Gaviria

1. 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.

2. 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 cómo el nivel educativo y edad de la madre, etnia asociada, semanas de gestación, enfermedades relacionadas y causa de muerte.

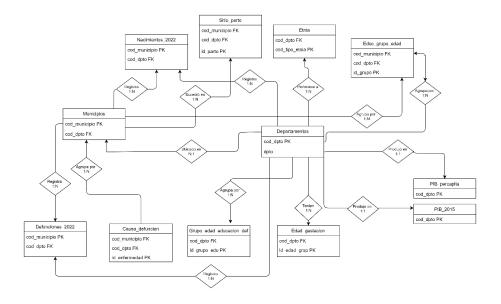
Con el análisis, también se busca entender la dinámica de los nacimientos de acuerdo con el sexo biológico del bebe, la concentración de nacimientos en los diferentes entornos y profesionales encargados de la asistencia, así como la distribución de nacimientos por población urbana y rural.

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.

3. TIPO DE ANÁLISIS DE DATOS

- 1. **Descriptivo**: Describir el comportamiento tanto de los nacimientos cómo de las defunciones fetales durante el año 2022 en el territorio colombiano.
- 2. **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.
- 3. **Prescriptivo**: Identificar acciones que permitan establecer políticas públicas para mitigar posibles comportamientos atípicos ligados a enfermedades o atención en el sistema de salud de ser posible.

4. DIAGRAMA ENTIDAD - RELACIÓN



5. Tablas y entidades

A continuación, se detalla para cada una de las tablas, la definición, los campo 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

1. **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	PK	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	

2. **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	
id_parto	INT	PK	NOT NULL	
cod_dpto	INT	FK	NOT NULL	
cod_municipio	INT	FK	NOT NULL	
cod_sitio	INT	-	NOT NULL	
sitio_parto	VARCHAR	-	NULL	
Médico	INT	-	NULL	
Enfermero(a)	INT	-	NULL	
Auxiliar de Enfermería	INT	-	NULL	
Promotor(a) de salud	INT	-	NULL	
Partera	INT	1	NULL	
Otra persona	INT	-	NULL	
Sin información	INT	-	NULL	

3. **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	
id_tipo_etnia	INT	PK	NOT NULL	
cod_dpto	INT	FK	NOT NULL	
cod_etnia	INT	-	NOT NULL	
pertenecia_etnica	VARCHAR	-	NOT NULL	
total	INT	-	NULL	

4. **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	
id_grupo	INT	PK	NOT NULL	
cod_dpto	INT	FK	NOT NULL	
cod_municipio	INT	FK	NOT NULL	
cod_grupo	INT	-	NOT NULL	
edad_grupo	VARCHAR	-	NOT 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	

5. **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	PK	NOT 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	

6. 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
id_enfermedad	INT	PK	NOT NULL
cod_dpto	INT	FK	NOT NULL
cod_municipio	INT	FK	NOT NULL
cod_enfermedad	INT	-	NOT NULL
enfermedad	VARCHAR	-	NULL
total	INT	-	NULL
hombre	INT	-	NULL
mujer	INT	-	NULL
indeterminado	INT	-	NULL

7. **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	
id_grupo_edu	INT	PK	NOT NULL	
cod_dpto	INT	FK	NOT NULL	
cod_grupo_edu	INT	-	NOT NULL	
educación	VARCHAR	-	NOT NULL	
10 - 14 años	INT	-	NULL	
15 - 19 años	INT	-	NULL	
20 - 24 años	INT	-	NULL	
25 - 29	INT		NULL	
años 30 - 34		-		
años 35 - 39	INT	-	NULL	
años 40 - 44	INT	-	NULL	
años 45 - 49	INT	-	NULL	
años 50 - 54	INT	-	NULL	
años	INT	-	NULL	
Edad desconocida	INT	-	NULL	

8. **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	
ld_edad_grup	INT	PK	NOT NULL	
cod_dpto	INT	FK	NOT NULL	
cod_edad	INT	-	NOT NULL	
edad	VARCHAR	-	NOT NULL	

1		l	
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

9. **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	

10. **Municipios:** Código DANE e información de ubicación geográfica de Latitud y Altitud para cada uno de los municipios Colombia.

Municipios			
Campo	Tipo de campo	Tipo de clave	Obligatorio
cod_municipio	INT	PK	NOT NULL
LATITUD	FLOAT	-	NOT NULL
LONGITUD	FLOAT	-	NOT NULL
Municipio	VARCHAR	-	NOT NULL

11. **Departamentos:** y Código DANE e información de ubicación geográfica de Latitud y Altitud para uno de los departamentos de Colombia

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

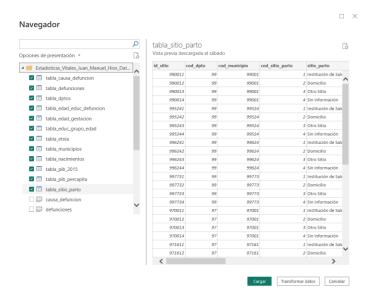
12. **PIB 2015:** Producto Interno Bruto departamental por habitante, a precios corrientes. Serie 2005-2022.

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

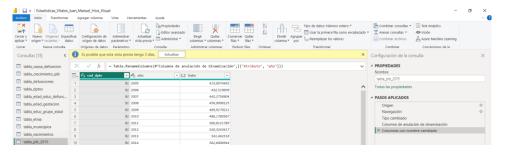
6. CAMBIOS GENERALES EN LAS TABLAS IMPORTADAS

Al importar la base de datos desde el archivo.xlsx a POWER BI Desktop se realizan los siguientes cambios:

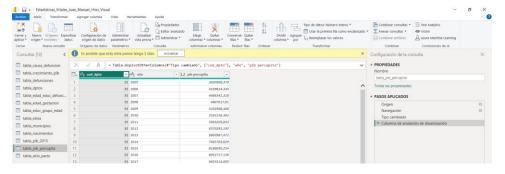
1. Se eligen todas las opciones para las cuales se cuenta con la opción tabla, es decir, se omiten las hojas del archivo:



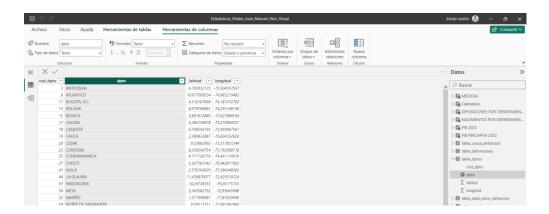
2. Columna de anulación de dinamización para la tabla **tabla_pib_2015**; se realiza una transformación en Power Query, donde se trasfiguran columnas en pares atributovalor, es decir, se cambian los valores de campos de cada uno de los años del PIB a un atributo año, además del cambio de nombre de 'atributo' a 'año'.



3. Columna de anulación de dinamización para la tabla tabla_pib_percapita; se realiza una transformación en Power Query, donde se trasfiguran columnas en pares atributo-valor, es decir, se cambian los valores de campos de cada uno de los años del PIB a un atributo año, en este caso, se asignan directamente los nombres de 'año' y 'pib_percapita' en la función que transforma la tabla.



4. Para el valor del nombre de departamento en la tabla tabla_dptos se asigna la categoría estado o provincia, con la finalidad de poder ubicar cada uno de los departamentos en el mapa coroplético de Colombia.



5. Para los campos 'cod_dpto' y 'cod_municipio' se asigna cómo no resume ya que será la principal FK del modelo en general.

7. MEDIDAS, COLUMNAS Y TABLAS CALCULADAS CON DETALLE

Para el desarrollo del análisis se crean las siguientes medidas y tablas a partir de la información cargada y transformada:

1. Medidas:

7.1.1.Creación de la tabla **MEDIDAS** para almacenar cada una de las carpetas y medidas generadas.



7.1.2. Creación de las medidas 'MENOS DE 22 SEMANAS', '22 A 27 SEMANAS', '28 A 36 SEMANAS' y 'MAS DE 36 SEMANAS' las cuales se generan a partir de la tabla tabla_edad_gestacion y se almacenan en la carpeta Tiempo de Gestación con la finalidad de sumar el valor total por cada uno de los campos de semanas de gestación.

```
22 A 27 SEMANAS = SUM(tabla_edad_gestacion[De 22 a 27])
28 A 36 SEMANAS = SUM(tabla_edad_gestacion[De 28 a 36])

MAS DE 36 SEMANAS = SUM(tabla_edad_gestacion[De 37 y más])

MENOS DE 22 SEMANAS = SUM(tabla_edad_gestacion[Menos de 22])
```

7.1.3.Creación de las medidas 'DEFUNCIONES FEMENINO', 'DEFUNCIONES MASCULINO', y 'DEFUNCIONES INDETERMINADO' a partir de la tabla 'tabla_defunciones', que son almacenadas en la carpeta DEFUNCIONES. Esto con el fin de obtener el total de defunciones del género masculino, femenino e indeterminado.

```
DEFUNCIONES FEMENINO = SUM(tabla_defunciones[mujer_total])
DEFUNCIONES INDETERMINADO = SUM(tabla_defunciones[indeterminado_total])
DEFUNCIONES MASCULINO = SUM(tabla_defunciones[hombre_total])
```

7.1.4.Creación de las medidas 'NACIMIENTOS FEMENINO' y 'NACIMIENTOS MASCULINO', a partir de la tabla 'tabla_nacimientos', que son almacenadas en la carpeta NACIMIENTOS. Esto con el fin de obtener el total de nacimientos por cada uno de los géneros: masculino y femenino.

2. Columnas:

7.2.1. se crean columnas que principalmente suman los campos para obtener el total para cada registro ya sea departamento o municipio, a continuación, se relaciona la formula de la ecuación y la tabla utilizada para crearla:

```
TOTAL
          DEFUNCIONES
                           ENFERMEDAD
                                                 tabla causa defuncion[hombre]
tabla causa defuncion[mujer] + tabla causa defuncion[indeterminado]
TOTAL DEFUNCIONES = tabla defunciones[hombre total] + tabla defunciones[mujer total]
+tabla_defunciones[indeterminado_total]
TOTAL = tabla_edad_gestacion[Menos de 22] + tabla_edad_gestacion[De 22 a 27] +
tabla edad gestacion[De 28 a 36] + tabla edad gestacion[De
                                                                   37 y más]
tabla edad gestacion[Ignorado] + tabla edad gestacion[Sin información]
TOTAL GRUPO EDAD = tabla_educ_grupo_edad[Preescolar] + tabla_educ_grupo_edad[Básica
                     tabla educ grupo edad[Media
                                                                0
                                                                      clásical
secundaria]
               +
                                                   académica
tabla_educ_grupo_edad[Media
                             técnica]
                                         + tabla_educ_grupo_edad[Normalista]
tabla educ grupo edad[Técnica profesional] + tabla educ grupo edad[Tecnológica]
tabla_educ_grupo_edad[Profesional]
                                    + tabla_educ_grupo_edad[Especialización]
tabla_educ_grupo_edad[Maestría]
                                             tabla_educ_grupo_edad[Doctorado]
tabla_educ_grupo_edad[Ninguno] + tabla_educ_grupo_edad[Sin información]
TOTAL NACIMIENTOS = tabla_nacimientos[hombre_total] + tabla_nacimientos[mujer_total]
+tabla nacimientos[indeterminado total]
TOTAL SITIO PARTO = tabla_sitio_parto[Médico] + tabla_sitio_parto[Enfermero(a)] +
tabla sitio parto[Auxiliar de Enfermería] + tabla sitio parto[Promotor(a) de salud] +
tabla sitio parto[Partera] + tabla sitio parto[Otra persona] + tabla sitio parto[Sin
información]
```

3. Tablas:

7.3.1. Creación de la tabla '**DEFUNCIONES POR DEPARTAMENTO**' para sumar el total de defunciones por cada uno de los departamentos sin tener en cuenta los demás campos de tabla, si no, el departamento y el total, con la finalidad de poder compara o relacionar con otras tablas.

7.3.2. Creación de la tabla '**DEFUNCIONES POR DEPARTAMENTO**' para sumar el total de nacimientos por cada uno de los departamentos sin tener en cuenta los demás campos de tabla, si no, el departamento y el total, con la finalidad de poder compara o relacionar con otras tablas.

```
NACIMIENTOS POR DEPARTAMENTO = SUMMARIZE('tabla_nacimientos', 'tabla_nacimientos' [cod_dpto], "nacimientos", SUM('tabla_nacimientos' [TOTAL NACIMIENTOS]))
```

7.3.3.Creación de la tabla 'PIB 2022' para extraer el valor del PIB cada uno de los departamentos únicamente para el año 2022, la finalidad de la tabla es poder comprar la correlación con el resumen de nacimientos y defunciones.

```
PIB 2022 = CALCULATETABLE(tabla pib 2015,tabla pib 2015[año]="2022")
```

7.3.4. Creación de la tabla 'PIB PERCAPITA 2022' para extraer el valor del PIB per cápita de cada uno de los departamentos únicamente para el año 2022, la finalidad de la tabla es poder comprar la correlación con el resumen de nacimientos, defunciones y el total del PIB para el mismo año.

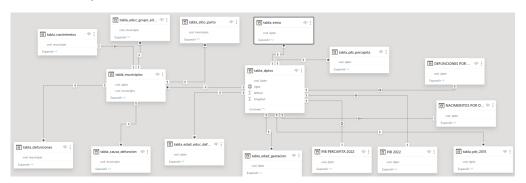
```
PIB PERCAPITA 2022 = CALCULATETABLE(tabla_pib_percapita,tabla_pib_percapita[año]="2022")
```

7.3.5. A pesar de que no se va a utilizar un componente de fechas en el análisis del tablero ya que comprende la información del total del año sin particiones por mes o día, se crea la tabla calendario como requerimiento, está tabla se crea entre las fechas 01/01/2022 y 31/12/2022.

```
Calendario = CALENDAR(DATE(2022,1,1), DATE(2022,12,31))
```

8. MODELO RELACIONAL EN PBI

Luego de realizar la creación de tablas y medidas, se obtiene el siguiente modelo relacional en Power BI Desktop:



9. SEGMENTACIONES Y BOTONES ELEGIDOS

Para la segmentación y navegación del tablero se crearon los siguientes objetos:

1. Segmentación:

9.1.1. Filtro para buscar nombre de municipios y departamentos en las hojas nacimientos, defunciones e indicadores con el objetivo de examinar el comportamiento de un actor en particular.

dpto	municipio P Buscar
(En blanco)	(En blanco)
☐ AMAZONAS	☐ ABEJORRAL
ANTIOQUIA	ÁBREGO
☐ ARAUCA	☐ ABRIAQUÍ

2. Botones:

9.2.1. Barra de navegación en cada una de las hojas a diferencia de HOME, la cual tiene botones independientes para la navegación, aunque con la misma finalidad.



9.2.2. Botón HOME para el retorno a la página de inicio del tablero, este botón se encuentra en cada una de las hojas a diferencia de la propia página de Home.



9.2.3.Botón BORRAR SEGMENTACIÓN se encuentra en las hojas en las cuales se insertó un objeto de segmentación por departamento o municipio y como su nombre lo indica, el objetivo es la eliminación de los parámetros aplicados para segmentar la información de la página.



9.2.4.Los botones DEPARTAMENTOS y MUNICIPIOS que se encuentran en la página nacimientos, se encuentran relacionados a un marcador que permite ocultar o visualizar un mapa a nivel departamental o municipal de acuerdo con su nombre.





9.2.5. En la página inicial, se crean botones para la navegación o conexión con cada una de las paginas de interés, cumple con las mismas funciones de la barra de navegación, sin embargo, los botones tienen una mayor oportunidad de personalización.



10. VISUALIZACIONES

1. Home: Pagina de inicio.



2. General: descripción general del comportamiento de los nacimientos.



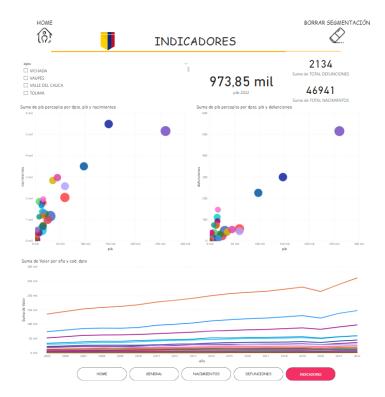
3. Nacimientos: Descripción detallada a nivel departamentos y municipal, al igual que grupo de edad de la madre, etnia perteneciente y nivel de profesión.



4. Defunciones: descripción detallada a nivel municipal y departamental de las defunciones prenatales, registros por grupo de edad de la madre, causa de muerte y semanas de gestación.



5. Indicadores: relaciones de los principales indicadores económicos; PIB y PIB per cápita, con el numero total de defunciones y nacimientos por departamento, también grafico histórico de líneas con histórico del comportamiento del PIB a precios del año 2015.



11.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 **Tablero:** Estadisticas_Vitales_Juan_Manuel_Hios_Visual