

YACHAY- Reports: Sistema de Generación de Reportes Interactivos para el Registro Poblacional de Cáncer del Municipio de Pasto



MANUAL DE REFERENCIA



Contenido

INTRODUCCION.....	3
JUSTIFICACIÓN DE LA HERRAMIENTA (STREAMLIT).....	3
ARQUITECTURA TÉCNICA E INSTALACIÓN.....	3
Tecnologías Utilizadas	4
Instalación de Dependencias	4
Configuración de la Conexión a Base de Datos	5
Ejecución de la Aplicación.....	5
Plantilla de Exportación (Opcional)	5
MÓDULOS DEL SISTEMA	5
Control de Acceso y Navegación	6
Personalización de Gráficos.....	6
Módulo: Casos por Municipio	7
Bloque 1: Análisis Comparativo General (Pasto vs Otros)	7
Bloque 2: Análisis Filtrado por Selección	7
Módulo: Fuentes de Información.....	8
Bloque 1: Top 10 Instituciones de Salud.....	8
Bloque 2: Análisis Filtrado	9
Módulo: Comportamiento por Fuente	10
Módulo: Grupo Etéreo por Fuente.....	11
Módulo: División Etárea	11
Módulo: Localizaciones por Sexo.....	12
Módulo: Diagnóstico	13
Módulo: Agrupación CIE-10	14
Bloque 1: Búsqueda por Filtración de Agrupación	14
Módulo: Frecuencia de Datos Perdidos	15
Módulo: Proyectos y Reportes Libres (Constructor AD HOC)	16
Variables Disponibles (Catálogo)	16
Lógica de Visualización Dinámica.....	17
Gestión de Proyectos	17
GENERACIÓN Y EXPORTACIÓN DE REPORTES.....	17
Exportación a PDF (generar_pdf_multimodo).....	17
Exportación a Word (generar_word)	18
CAPA DE ACCESO A DATOS (database.py)	20
Función consultar().....	20
MODELO DE DATOS DEL RPCMP	20
CONCLUSIONES.....	21
TRABAJOS FUTUROS	21

INTRODUCCION

YACHAY-REPORTS es una plataforma web de generación de reportes interactivos estadísticos de los diferentes casos de cáncer almacenados en el Registro Poblacional de Cáncer del Municipio de Pasto. La herramienta centraliza diez módulos de análisis temático que cubren las principales dimensiones del registro oncológico: distribución geográfica, fuentes de notificación, comportamiento tumoral, grupos de edad, codificación CIE-10, métodos diagnósticos e integridad de datos. Cada módulo genera gráficos personalizables y tablas estadísticas con opción de exportación a PDF y Word, lo que facilita la preparación de informes técnicos sin necesidad de conocimientos avanzados de programación o estadística.

Yachay Reports es una aplicación web de análisis y reporte estadístico, desarrollada en Python con el framework Streamlit, cuyo propósito es proporcionar una interfaz interactiva e intuitiva para explorar, visualizar y exportar información epidemiológica de manera autónoma.

La arquitectura de la aplicación se apoya en tres componentes principales: el módulo `app.py`, que contiene toda la lógica de la interfaz de usuario y las rutinas de generación de reportes; el módulo `database.py`, que gestiona las conexiones a la base de datos PostgreSQL mediante SQLAlchemy; y la base de datos, que consolida los registros clínicos, demográficos e institucionales del RPCMP.

Este manual de referencia documenta en detalle el proceso de instalación, la arquitectura técnica, la funcionalidad de cada módulo y las instrucciones para la generación de reportes. Está dirigido a administradores del sistema, desarrolladores que necesiten mantener o extender la plataforma, y usuarios finales que requieran comprender el alcance analítico de la herramienta.

JUSTIFICACIÓN DE LA HERRAMIENTA (STREAMLIT)

Tras evaluar el ecosistema de frameworks para aplicaciones analíticas en Python, se seleccionó Streamlit como la tecnología base del Portal Yachay Reports. Esta decisión se fundamenta en su capacidad de transformar scripts de análisis de datos en aplicaciones web interactivas con un mínimo de código adicional, lo que resulta ideal para equipos con perfiles más científicos que de desarrollo web.

Desde una perspectiva técnica, Streamlit se integra nativamente con las principales bibliotecas de la ciencia de datos en Python, incluyendo pandas para la manipulación de datos, matplotlib para la generación de gráficos, SQLAlchemy para la conexión a bases de datos relacionales, y ReportLab y python-docx para la exportación de reportes. Esta sinergia permite construir flujos de trabajo completos, desde la consulta SQL hasta la entrega del reporte final, dentro de un único entorno coherente.

El factor determinante para su adopción fue la velocidad de desarrollo y la curva de aprendizaje reducida, que permiten que analistas epidemiológicos puedan incorporar nuevos módulos o ajustar consultas existentes sin depender de un ciclo de desarrollo complejo. Adicionalmente, Streamlit es una tecnología de código abierto respaldada por una comunidad activa, lo que garantiza la disponibilidad de actualizaciones de seguridad y nuevas funcionalidades sin costos de licenciamiento.

ARQUITECTURA TÉCNICA E INSTALACIÓN

El sistema Yachay Reports se implementa como una aplicación monolítica de una sola página (single-page) servida localmente mediante el servidor de desarrollo de Streamlit. La arquitectura se organiza en tres capas bien diferenciadas:

- **Capa de presentación:** Implementada en `app.py` mediante componentes reactivos de Streamlit que renderizan tablas, gráficos, filtros y controles de exportación en el navegador del usuario.

- **Capa de negocio:** También contenida en app.py, incluye las funciones de procesamiento de datos (procesar_metricas_profesional, dibujar_grafico_universal), las rutinas de generación de PDF (generar_pdf_multimodo) y de Word (generar_word), así como toda la lógica de construcción dinámica de consultas SQL.
- **Capa de datos:** Implementada en database.py, proporciona una función única consultar() que establece una conexión a la base de datos PostgreSQL mediante SQLAlchemy, ejecuta una consulta SQL parametrizada y retorna un DataFrame de pandas, liberando la conexión al finalizar.

Tecnologías Utilizadas

Componente	Tecnología / Biblioteca	Versión Mínima	Rol
Framework UI	Streamlit	1.30+	Servidor web e interfaz de usuario
Lenguaje	Python	3.10+	Lógica de aplicación completa
Base de Datos	PostgreSQL	14+	Almacenamiento del registro oncológico
ORM / Conector	SQLAlchemy + psycopg2	2.0+	Gestión de conexiones a BD
Análisis de datos	Pandas + NumPy	Reciente	Manipulación y cálculo estadístico
Visualización	Matplotlib	3.7+	Generación de gráficos
PDF	ReportLab + pdfw	4.0+	Generación y fusión de documentos PDF
Word	python-docx	1.0+	Generación de reportes en formato .docx
Serialización	JSON	stdlib	Persistencia de configuraciones de proyectos

Instalación de Dependencias

Para desplegar el sistema Yachay Reports es necesario contar con Python 3.10 o superior. Se recomienda el uso de pip para la gestión de paquetes. Ejecute los siguientes comandos en la carpeta raíz del proyecto:

```
python -m venv venv
python -m venv venv
source venv/bin/activate # Linux/Mac
venv\Scripts\activate # Windows

# Instalar dependencias principales
pip install streamlit pandas numpy matplotlib sqlalchemy psycopg2-binary
pip install reportlab pdfw python-docx
```

Configuración de la Conexión a Base de Datos

El archivo `database.py` contiene los parámetros de conexión a PostgreSQL. Antes de ejecutar la aplicación, asegúrese de que los valores correspondan con la instancia de base de datos del RPCMP.

```
USER = "postgres"      # Usuario de PostgreSQL
PASS = "admin"         # Contraseña del usuario
HOST = "localhost"    # Dirección del servidor
PORT = "5432"          # Puerto de escucha
DB   = "BDYachay"     # Nombre de la base de datos
```

La función `consultar(query)` encapsula el ciclo de vida completo de la conexión: crea el motor SQLAlchemy, abre una conexión transaccional, ejecuta la consulta mediante `pd.read_sql()` y libera todos los recursos al finalizar mediante `engine.dispose()`, garantizando que no existan fugas de conexiones en sesiones largas.

Ejecución de la Aplicación

Una vez configurada la conexión y con las dependencias instaladas, la aplicación se inicia con el siguiente comando desde la carpeta raíz del proyecto:

```
streamlit run app.py
```

Streamlit abrirá automáticamente una ventana del navegador apuntando a `http://localhost:8501`. Si el puerto 8501 está ocupado, Streamlit seleccionará automáticamente el siguiente disponible e informará la URL en la consola.

Plantilla de Exportación (Opcional)

La aplicación puede utilizar plantillas para darle marca institucional a los reportes generados. Si se desean aplicar, los siguientes archivos deben ubicarse en la misma carpeta que `app.py`:

plantilla.pdf: Documento PDF de una sola página que se utiliza como fondo o membrete para los reportes en PDF. Si no existe, los reportes se generan sin plantilla.

plantilla.docx: Documento Word que sirve como base de estilo para los reportes exportados en formato `.docx`. Si no existe, `python-docx` generará el documento con estilos predeterminados.

MÓDULOS DEL SISTEMA

El Portal Yachay Reports se organiza en una barra lateral de navegación (sidebar) desde la cual el usuario selecciona el módulo de análisis que desea consultar. Todos los módulos comparten una arquitectura funcional común: un conjunto de filtros en el sidebar, uno o más bloques de análisis en el área principal, visualizaciones gráficas interactivas y botones de exportación.



Control de Acceso y Navegación

Al ingresar a la aplicación, el sistema presenta una pantalla de bienvenida con el título Portal Yachay Reports y un botón INGRESAR A REPORTES. Este control de acceso es una capa de navegación básica que utiliza el mecanismo de estado de sesión de Streamlit (`st.session_state`) para determinar si el usuario ha ingresado a la plataforma.

Una vez dentro, el selector de módulos se ubica en la barra lateral bajo la etiqueta Módulo: y presenta las siguientes opciones disponibles:

- Casos por Municipio
- Fuentes de Información
- Comportamiento por Fuente
- Grupo Etéreo por Fuente
- División Etearea
- Localizaciones por Sexo
- Diagnóstico
- Agrupación CIE-10
- Frecuencia Datos Perdidos
- Proyectos y Reportes Libres

La barra lateral incluye además el botón Cerrar Sesión, que restablece el estado de la aplicación a la pantalla de bienvenida.

Personalización de Gráficos

Una característica común a todos los módulos es el sistema de personalización de colores disponible en la parte inferior de la barra lateral, bajo el encabezado Personalizar Gráficos. Este sistema permite seleccionar entre cuatro paletas de color predefinidas:

Paleta	Color Hombres	Color Mujeres	Descripción
Estándar (Azul/Rosa)	#1f77b4	#e377c2	Paleta por defecto del portal
Vibrante (Cian/Naranja)	#00d2ff	#ff8c00	Alta visibilidad para presentaciones

Paleta	Color Hombres	Color Mujeres	Descripción
Profesional (Gris/Azul)	#2c3e50	#3498db	Para informes formales en blanco y negro
Personalizado	Elección libre	Elección libre	Selector de color libre para ambos géneros

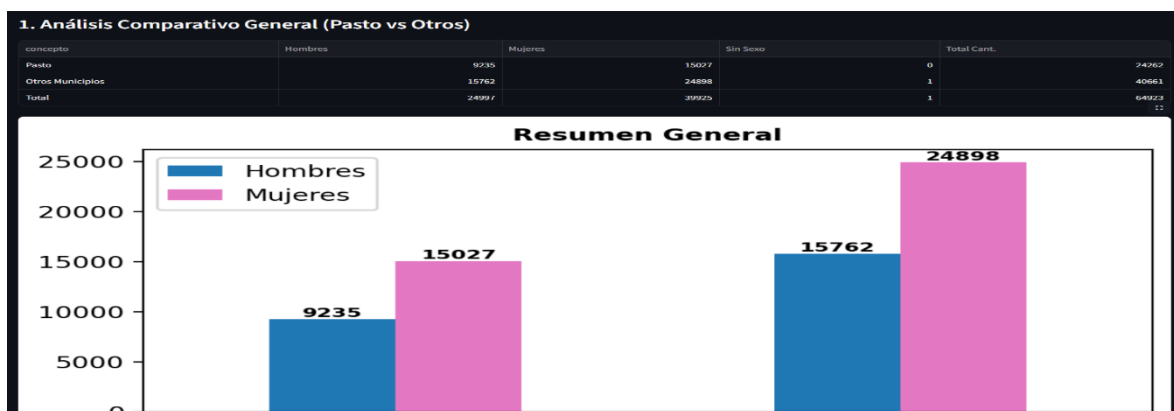
Los colores seleccionados se aplican de forma reactiva a todos los gráficos de barras de género que se generen en la sesión activa.

Módulo: Casos por Municipio

Este módulo analiza la distribución geográfica de los casos oncológicos registrados. Permite cruzar los datos por departamento, municipio y año de diagnóstico para identificar patrones de incidencia territorial.

Bloque 1: Análisis Comparativo General (Pasto vs Otros)

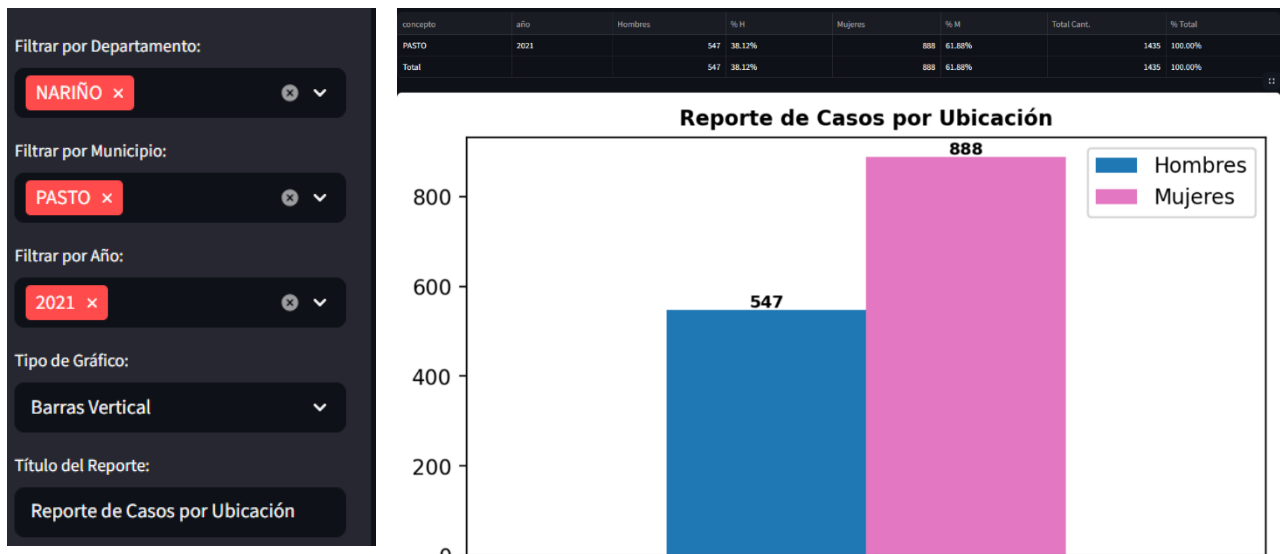
Como vista de contexto general, este bloque presenta una tabla resumen que consolida los casos en dos grandes categorías: Pasto y Otros Municipios, desglosados por Hombres, Mujeres y Total. El toggle Mostrar Análisis Porcentual permite añadir columnas de porcentaje calculadas sobre el total general.



Bloque 2: Análisis Filtrado por Selección

Cuando el usuario aplica filtros desde la barra lateral, aparece dinámicamente un segundo bloque con el detalle de los municipios seleccionados. Los filtros disponibles son:

- **Filtrar por Departamento:** Lista desplegable múltiple con todos los departamentos presentes en la base de datos.
- **Filtrar por Municipio:** Se carga dinámicamente en función del departamento seleccionado. Incluye la opción especial PASTO 2 para filtrar únicamente los registros asociados a barrios de Pasto (excluyendo código de barrio 999998).
- **Filtrar por Año:** Selector múltiple de años de diagnóstico desde 1964 hasta 2026, incluyendo la categoría Desconocido.
- **Tipo de Gráfico:** Permite seleccionar entre Barras Vertical, Barras Horizontal, Barras Apiladas, Circular (Torta) y Bidireccional.



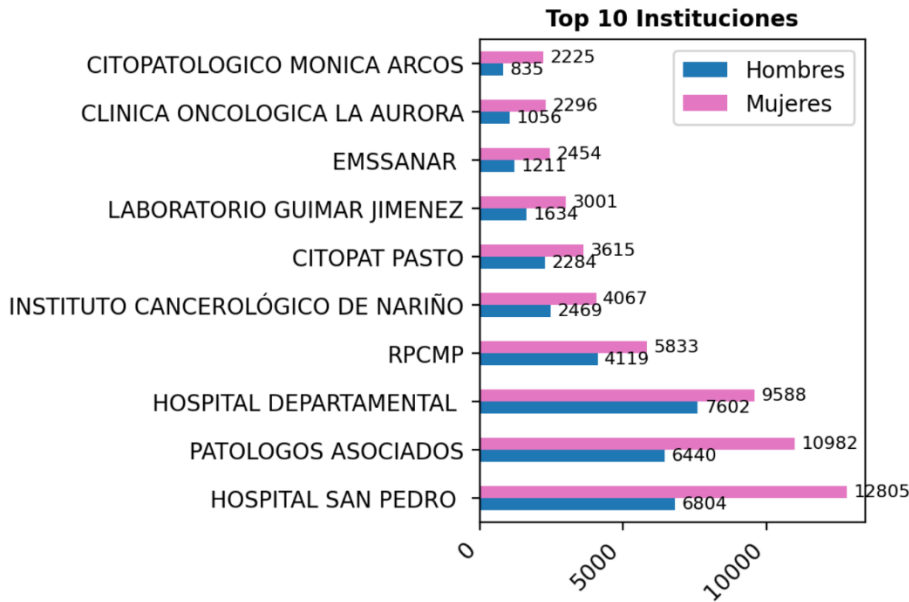
Módulo: Fuentes de Información

Este módulo analiza el comportamiento de las instituciones de salud (fuentes) que notifican casos al RPCMP. Proporciona métricas sobre el volumen de casos reportados por cada entidad, desagregadas por sexo.

Bloque 1: Top 10 Instituciones de Salud

Presenta automáticamente las diez instituciones con mayor número de historias clínicas reportadas, incluyendo el desglose por Hombres, Mujeres y Total. La consulta subyacente realiza un JOIN entre las tablas tab_fichas, tab_historias, tab_fuentes y tab_pacientes.

1. Top 10 Instituciones de Salud							
concepto	Hombres	% H	Mujeres	% M	Total Cant.	% Total	
HOSPITAL SAN PEDRO	6804	7.45%	12805	14.02%	19609	21.47%	
PATOLOGOS ASOCIADOS	6440	7.05%	10982	12.03%	17423	19.08%	
HOSPITAL DEPARTAMENTAL	7602	8.32%	9588	10.50%	17190	18.82%	
RPCMP	4119	4.51%	5833	6.39%	9952	10.90%	
INSTITUTO CANCEROLÓGICO DE NARIÑO	2469	2.70%	4067	4.45%	6536	7.16%	
CITOPAT PASTO	2284	2.50%	3615	3.96%	5899	6.46%	
LABORATORIO GUIMAR JIMENEZ	1634	1.79%	3001	3.29%	4635	5.08%	
EMSSANAR	1211	1.33%	2454	2.69%	3665	4.01%	
CLINICA ONCOLOGICA LA AURORA	1056	1.16%	2296	2.51%	3352	3.67%	
CITOPATOLOGICO MONICA ARCOS	835	0.91%	2225	2.44%	3060	3.35%	



Bloque 2: Análisis Filtrado

Permite aplicar filtros adicionales por Institución, Departamento, Municipio y Año de diagnóstico. Cuando se activa el filtro de año, la tabla incluye una columna adicional con el año correspondiente a cada registro.

Filtrar por Institución:

EMSSANAR ✕ ▼

Filtrar por Departamento:

Choose options ▼

Filtrar por Municipio:

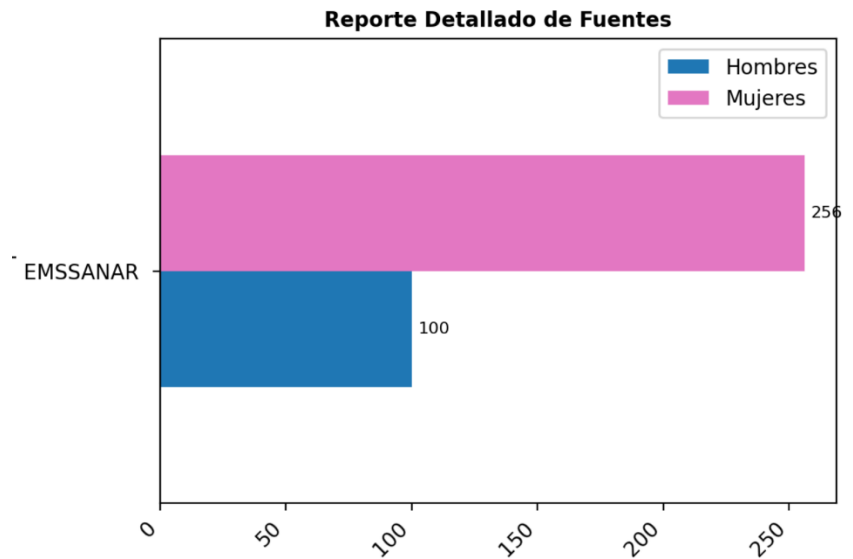
Choose options ▼

Filtrar por Año:

2023 ✕ ▼

Tipo de Gráfico:

Barras Horizontal ▼



2. Reporte Detallado de Fuentes

concepto	año	Hombres	% H	Mujeres	% M	Total Cant.	% Total
EMSSANAR	2023	100	28.09%	256	71.91%	356	100.00%
Total	---	100	28.09%	256	71.91%	356	100.00%

Módulo: Comportamiento por Fuente

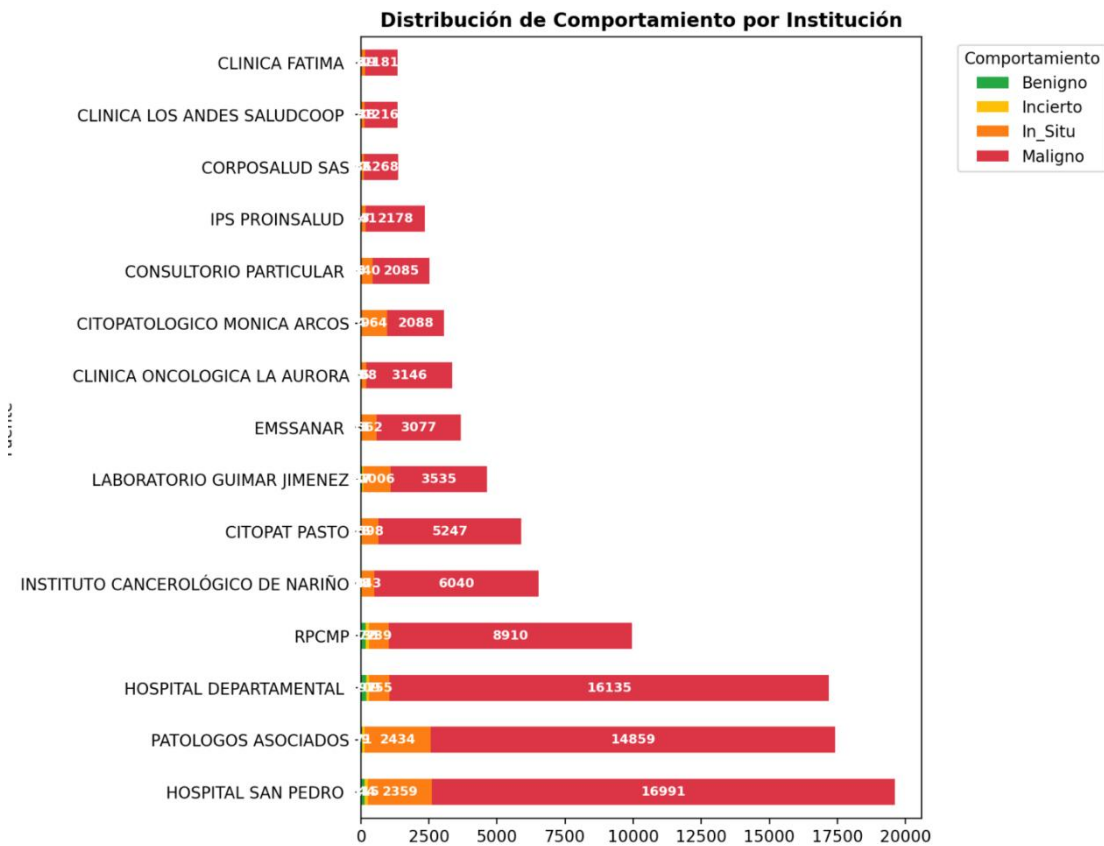
Este módulo evalúa la distribución del comportamiento biológico del tumor (Benigno, Incierto, In Situ, Maligno) por institución notificante. Emplea una paleta de colores tipo semáforo epidemiológico para facilitar la interpretación clínica: verde para benigno, amarillo para incierto, naranja para in situ y rojo para maligno.

Comportamiento por Institución General

Total Historias Clínicas Analizadas
123,784

Mostrando Top 10 instituciones + Total General.

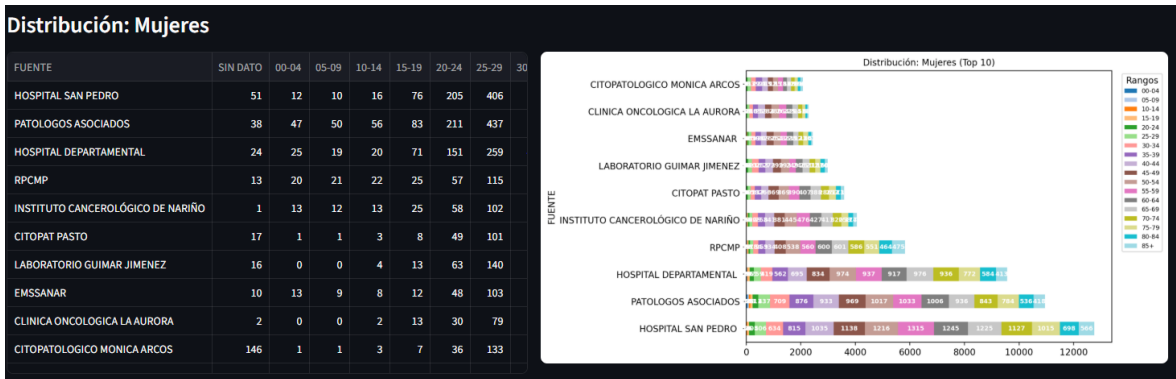
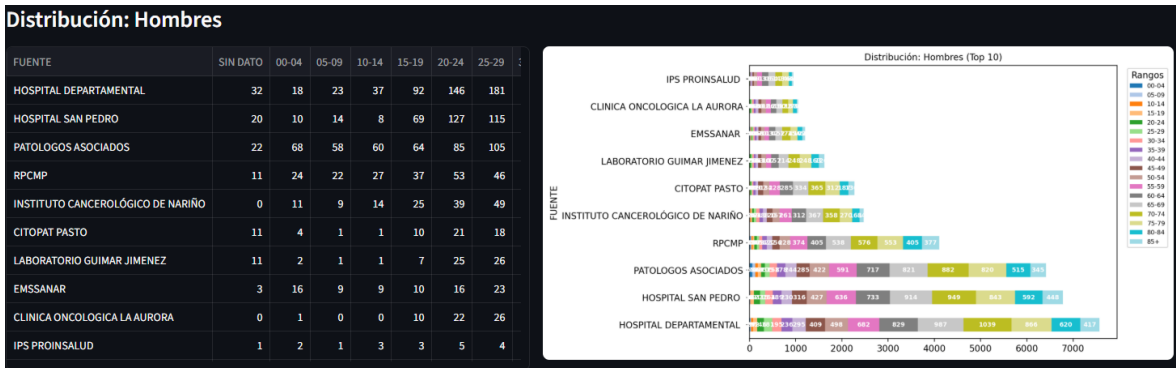
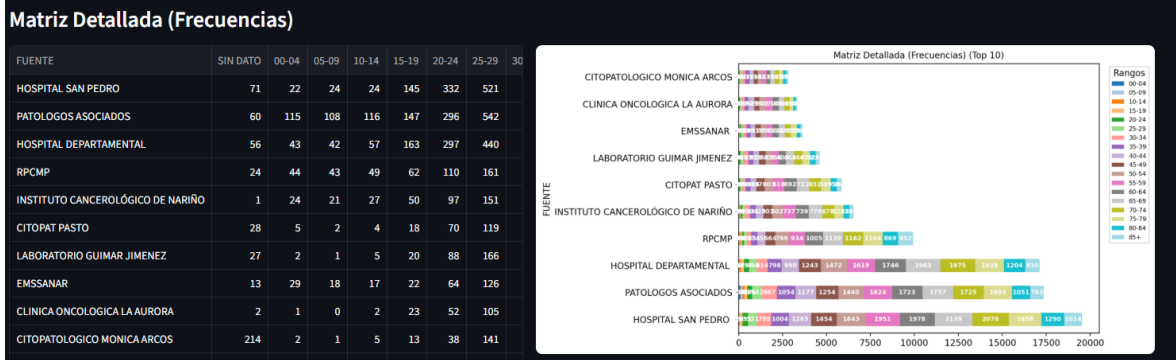
Fuente	Hombres	Mujeres	Benigno	Incierto	In_Situ	Maligno	Total Cant.	
HOSPITAL SAN PEDRO		6804	12805	144	115	2359	16991	19609
PATOLOGOS ASOCIADOS		6440	10982	59	71	2434	14859	17423
HOSPITAL DEPARTAMENTAL		7602	9588	191	109	755	16135	17190
RPCMP		4119	5833	172	131	739	8910	9952
INSTITUTO CANCEROLÓGICO DE NARIÑO		2469	4067	28	25	443	6040	6536
CITOPAT PASTO		2284	3615	28	26	598	5247	5899
LABORATORIO GUIMAR JIMENEZ		1634	3001	47	47	1006	3535	4635
EMSSANAR		1211	2454	11	15	562	3077	3665
CLINICA ONCOLOGICA LA AURORA		1056	2296	3	15	188	3146	3352
CITOPATOLOGICO MONICA ARCOS		835	2225	4	4	964	2088	3060



Los tipos de gráfico disponibles en este módulo son: Barras Apiladas Horizontal, Barras Apiladas Vertical, Barras Agrupadas y Líneas. La tabla resultante incluye columnas para Hombres, Mujeres, Benigno, Incierto, In_Situ, Maligno y Total, con fila de Total General al final.

Módulo: Grupo Etéreo por Fuente

Genera una matriz cruzada entre institución notificante y grupos de edad quinquenales (00-04, 05-09, ..., 85+), mostrando la distribución de casos en tres secciones independientes: General, Hombres y Mujeres.

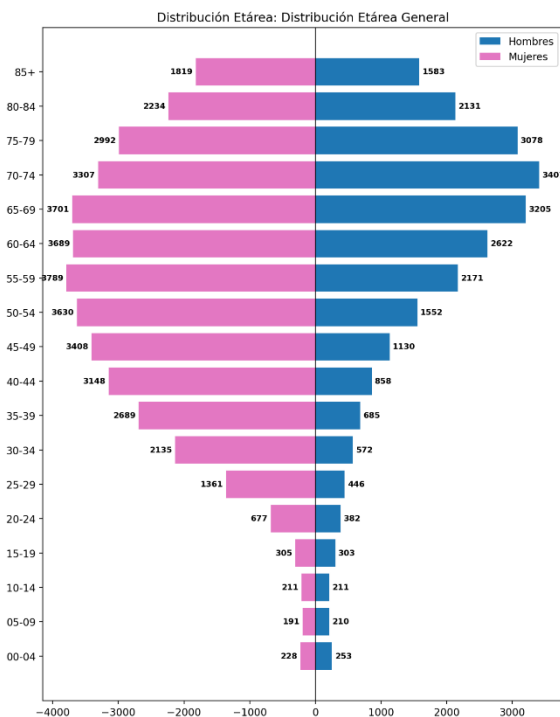


La edad se clasifica en rangos quinquenales mediante pd.cut() con 19 intervalos (incluyendo Sin Dato para edades desconocidas).

Módulo: División Etérea

Analiza la distribución de los casos por grupos de edad quinquenales y sexo a nivel global o con filtros geográficos y temporales. Es el módulo principal para la generación de pirámides poblacionales de incidencia.

Rango	Hombres	Mujeres	Total Cant.
Sin Dato		198	411
00-04		253	481
05-09		210	401
10-14		211	422
15-19		303	608
20-24		382	677
25-29		446	1361
30-34		572	2135
35-39		685	2689
40-44		858	3148



El tipo de gráfico Bidireccional (Pirámide) genera una pirámide poblacional en la que las barras de Hombres se extienden hacia la derecha y las de Mujeres hacia la izquierda, con etiquetas numéricas en cada barra. Las opciones adicionales incluyen Barras Apiladas y Barras Vertical.

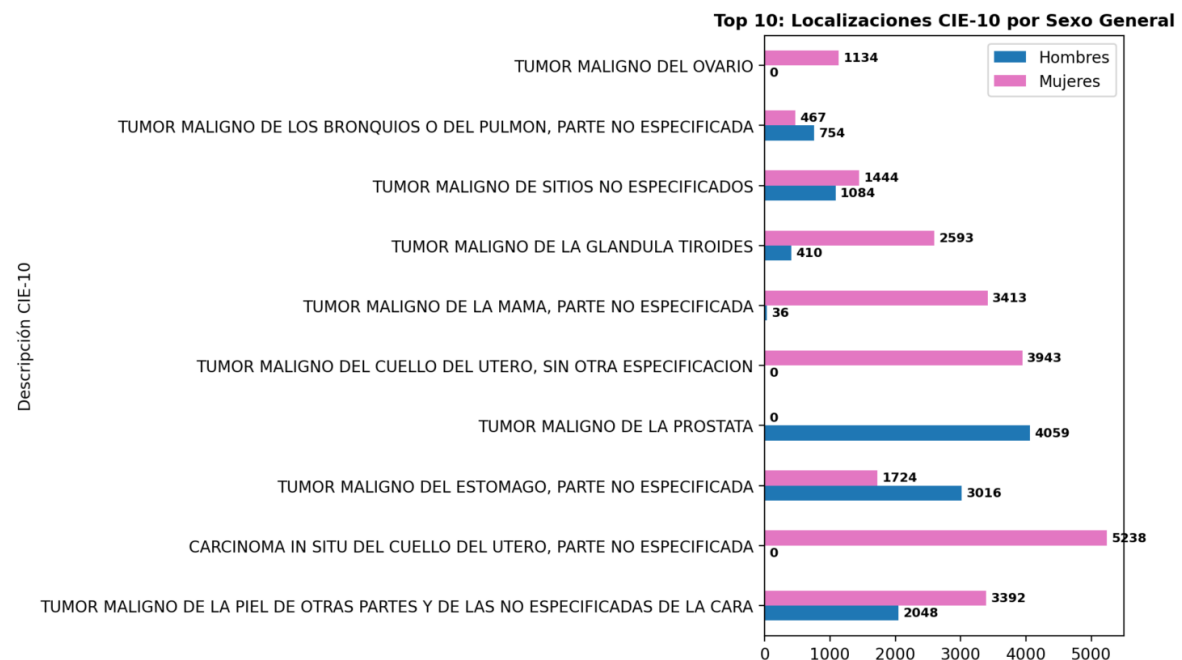
Módulo: Localizaciones por Sexo

Presenta la distribución de los diagnósticos según la clasificación CIE-10, desglosada por sexo. Permite buscar enfermedades específicas mediante un selector que muestra código y descripción CIE-10 juntos (ej: C50 - Mama).

Total Casos Diagnosticados
64,633

Mostrando Top 10 patologías + Total General.

Código	Descripción CIE-10	Hombres	Mujeres	Total Cant.
C443	TUMOR MALIGNO DE LA PIEL DE OTRAS PARTES Y DE LAS NO ESPECIFICADAS DE LA CARA	2048	3392	5440
D069	CARCINOMA IN SITU DEL CUELLO DEL UTERO, PARTE NO ESPECIFICADA	0	5238	5238
C169	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA	3016	1724	4741
C61	TUMOR MALIGNO DE LA PROSTATA	4059	0	4059
C539	TUMOR MALIGNO DEL CUELLO DEL UTERO, SIN OTRA ESPECIFICACION	0	3943	3943
C509	TUMOR MALIGNO DE LA MAMA, PARTE NO ESPECIFICADA	36	3413	3449
C73	TUMOR MALIGNO DE LA GLANDULA TIROIDES	410	2593	3003
C80	TUMOR MALIGNO DE SITIOS NO ESPECIFICADOS	1084	1444	2528
C349	TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON, PARTE NO ESPECIFICADA	754	467	1221
C56	TUMOR MALIGNO DEL OVARIO	0	1134	1134



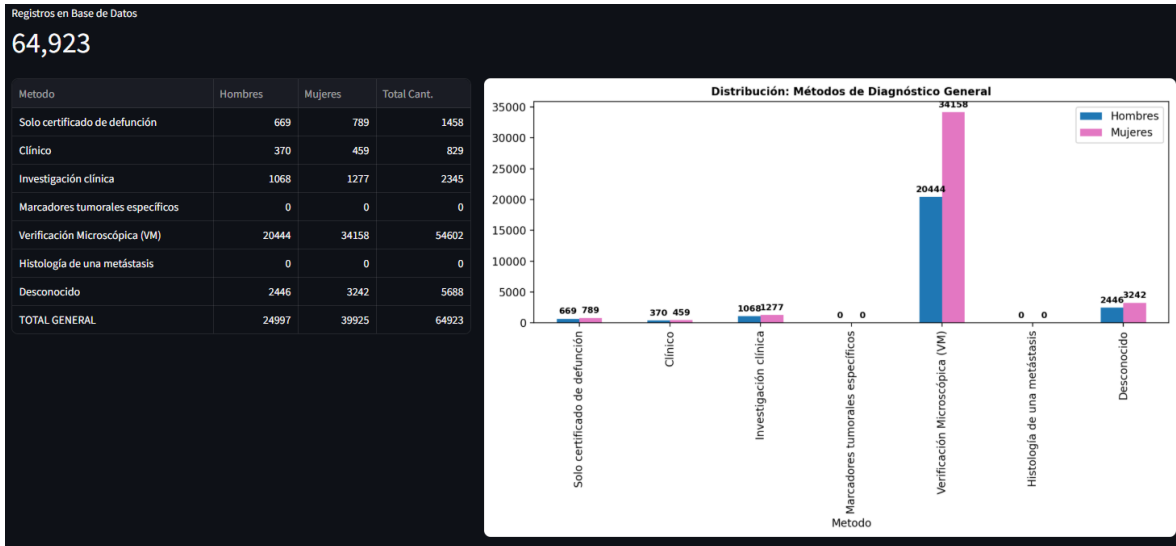
Para garantizar la legibilidad visual, el gráfico siempre muestra el Top 10 de patologías más frecuentes, aunque la tabla exportable incluye todas las categorías encontradas. Cuando el usuario aplica un filtro de CIE-10 específico, se muestran todos los registros coincidentes sin limitación.

Módulo: Diagnóstico

Analiza la distribución de los métodos de verificación diagnóstica empleados en el registro, según la codificación estándar del RPCMP. Los siete métodos posibles son:

Código	Método de Diagnóstico	Descripción
0	Solo certificado de defunción	Caso identificado únicamente a partir del acta de defunción
1	Clínico	Diagnóstico basado exclusivamente en evaluación clínica
2	Investigación clínica	Incluye exámenes de laboratorio e imagen sin confirmación microscópica

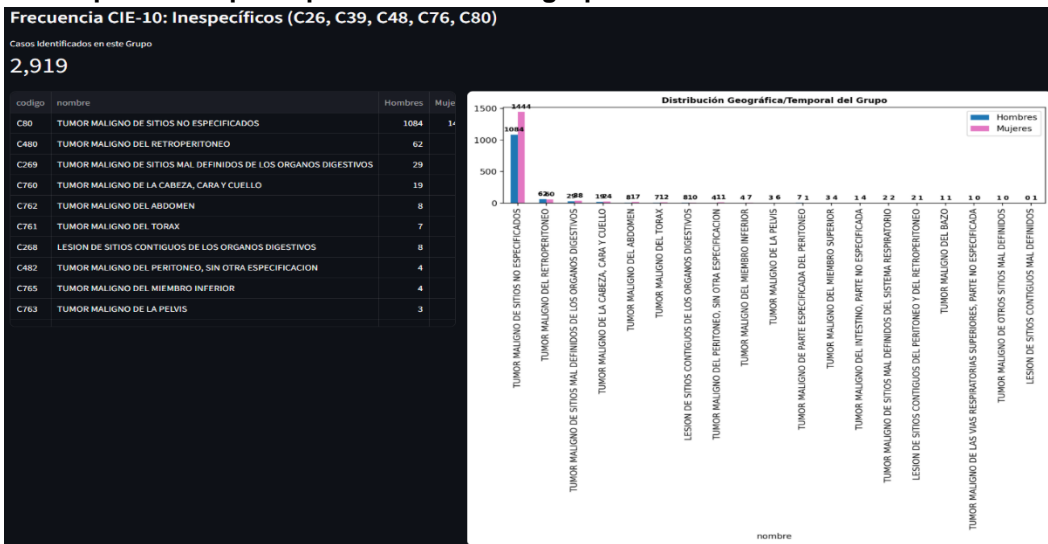
Código	Método de Diagnóstico	Descripción
4	Marcadores tumorales	Basado en marcadores biológicos específicos de neoplasia
5/7	Verificación Microscópica (VM)	Confirmación histológica o citológica del tumor primario
6	Histología de una metástasis	Confirmación microscópica de tejido metastásico
9	Desconocido	Método de diagnóstico no registrado o no disponible



Módulo: Agrupación CIE-10

Permite analizar la incidencia de neoplasias dentro de 57 agrupaciones anatómicas predefinidas, cada una compuesta por uno o varios códigos CIE-10 relacionados (ej. Tráquea, Bronquios y Pulmón agrupa C33 y C34). La búsqueda se realiza mediante patrones LIKE con comodín (LIKE 'C18%') para capturar subcódigos.

Bloque 1: Búsqueda por Filtración de Agrupación



Las agrupaciones disponibles abarcan desde localizaciones específicas (Estómago C16, Mama C50) hasta grupos de patologías relacionadas (Leucemia Mieloide C92-C94, Linfoma no Hodgkin C82-C86 y C96). El módulo permite aplicar filtros adicionales de Departamento, Municipio y Año de Diagnóstico.

Seleccionar Agrupación Técnica:

Esófago (C15) ▼

Filtros Adicionales

Departamento:

Choose options ▼

Municipio:

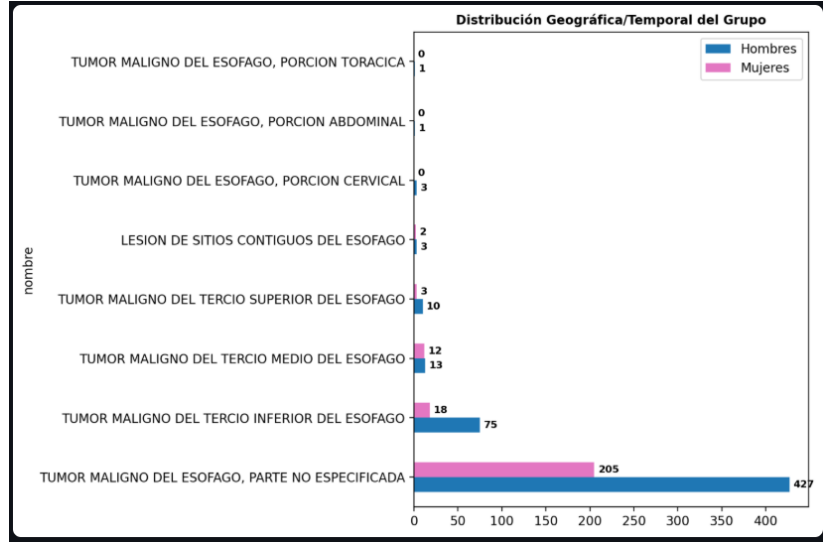
Choose options ▼

Año Diagnóstico:

Choose options ▼

Tipo de Gráfico:

Barras Horizontal ▼



codigo	nombre	Hombres	Mujeres	Total
C159	TUMOR MALIGNO DEL ESOFAGO, PARTE NO ESPECIFICADA	427	205	632
C155	TUMOR MALIGNO DEL TERCIO INFERIOR DEL ESOFAGO	75	18	93
C154	TUMOR MALIGNO DEL TERCIO MEDIO DEL ESOFAGO	13	12	25
C153	TUMOR MALIGNO DEL TERCIO SUPERIOR DEL ESOFAGO	10	3	13
C158	LESION DE SITIOS CONTIGUOS DEL ESOFAGO	3	2	5
C150	TUMOR MALIGNO DEL ESOFAGO, PORCION CERVICAL	3	0	3
C152	TUMOR MALIGNO DEL ESOFAGO, PORCION ABDOMINAL	1	0	1
C151	TUMOR MALIGNO DEL ESOFAGO, PORCION TORACICA	1	0	1
T00	TOTAL DEL GRUPO	533	240	773

Módulo: Frecuencia de Datos Perdidos

Evalúa la integridad y calidad del registro para siete variables críticas del sistema, permitiendo identificar qué proporción de los registros presenta información desconocida o faltante. Este análisis es esencial para evaluar la completitud del RPCMP y dirigir esfuerzos de mejora en la captura de datos.

Variable	Criterio de Dato Desconocido	Criterio de Dato Conocido
Edad	edad_paciente = 999	edad_paciente < 999
Sexo	cod_sexo = 9	cod_sexo = 1 o 2
Fecha Diagnóstico	fe_diagnostico empieza con '99' o termina en '9999'	Fecha válida registrada
Método Diagnóstico	cod_metododiagnostico = 9	Código de método 0-7
Localización	cod_localizacion = '809'	Código de localización válido
Morfología	cod_morfologia = '8000' (Neoplasia Maligna NOS)	Código morfológico específico
Grado Diferenciación	cod_gradodiferenciacion = 9	Grado registrado 1-4



Módulo: Proyectos y Reportes Libres (Constructor AD HOC)

Este módulo es el componente más flexible del sistema. Permite al usuario construir consultas personalizadas cruzando hasta ocho variables del registro, generando reportes ad hoc sin necesidad de modificar el código fuente. Constituye el motor de análisis exploratorio de la plataforma.

Constructor de Reportes AD HOC

Proyectos Guardados

Reporte_cie.json

Configurar Reporte

Variables a cruzar: Fuente x Código CIE-10 x Enfermedad (CIE-10)

Reporte

Total de casos: **87,968**

Combinaciones encontradas: **500**

Fuente	Código CIE-10	Enfermedad (CIE-10)
PATOLOGOS ASOCIADOS	C443	TUMOR MALIGNO DE LA PIEL DE OTRAS PARTES Y DE LAS NO E
PATOLOGOS ASOCIADOS	D069	CARCINOMA IN SITU DEL CUELLO DEL UTERO, PARTE NO ESPE
HOSPITAL SAN PEDRO	D069	CARCINOMA IN SITU DEL CUELLO DEL UTERO, PARTE NO ESPE
HOSPITAL SAN PEDRO	C443	TUMOR MALIGNO DE LA PIEL DE OTRAS PARTES Y DE LAS NO E
HOSPITAL DEPARTAMENTAL	C169	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA
HOSPITAL DEPARTAMENTAL	C539	TUMOR MALIGNO DEL CUELLO DEL UTERO, SIN OTRA ESPECIF
HOSPITAL SAN PEDRO	C539	TUMOR MALIGNO DEL CUELLO DEL UTERO, SIN OTRA ESPECIF
PATOLOGOS ASOCIADOS	C169	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA
HOSPITAL SAN PEDRO	C169	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA
HOSPITAL SAN PEDRO	C509	TUMOR MALIGNO DE LA MAMA, PARTE NO ESPECIFICADA

Guardar Proyecto

Nombre del archivo:

ej: reporte_anual_pasto

Guardar en Carpeta

Descargar Reporte PDF

Descargar Word

Personalizar Gráficos

Paleta de colores:

Estándar (Azul/Rosa)

Variables Disponibles (Catálogo)

Etiqueta de Usuario	Campo SQL Subyacente	Tabla
Departamento	td.nom_departamentos	tab_departamentos
Municipio	m.nom_municipios	tab_municipios
Enfermedad (CIE-10)	c.nom_cie10	tab_cie10
Código CIE-10	f.cod_cie10	tab_fichas
Edad	f.edad_paciente	tab_fichas
Año	RIGHT(f.fe_diagnostico, 4)	tab_fichas
Estrato	p.cod_estratos	tab_pacientes
Sexo	CASE cod_sexo 1='Hombre' 2='Mujer'	tab_pacientes

Etiqueta de Usuario	Campo SQL Subyacente	Tabla
Fuente	tfu.nom_fuentes	tab_fuentes

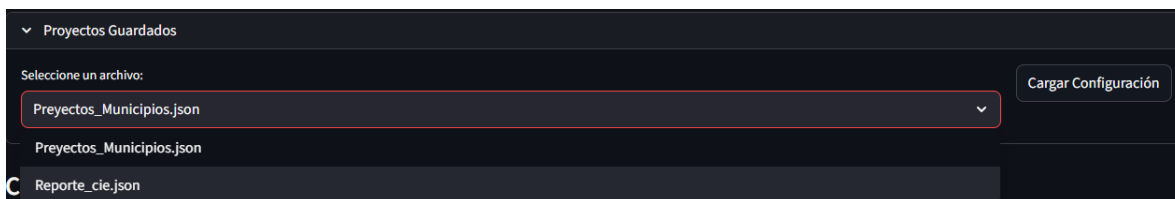
Lógica de Visualización Dinámica

El motor de visualización adapta automáticamente el tipo de gráfico al número de variables seleccionadas:

- **1 variable:** Gráfico simple (barras horizontales, líneas o área) mostrando el Top 15 de categorías por número de casos.
- **2 variables:** Tabla pivote cruzando las dos variables. Se muestran el Top 15 de filas y el Top 8 de columnas para mantener la legibilidad visual.
- **3 o más variables:** Dos vistas complementarias: un ranking de la variable principal (agrupando las demás) y un cruce pivote entre la primera y la segunda variable. Las variables adicionales se incluyen únicamente en la tabla.

Gestión de Proyectos

La configuración de cada reporte personalizado (título, variables seleccionadas, tipo de gráfico y color de tema) puede guardarse como archivo JSON en la carpeta `proyectos_reportes/`. Esta carpeta se crea automáticamente si no existe. Los proyectos guardados pueden cargarse en sesiones posteriores desde el panel Proyectos Guardados, eliminando la necesidad de reconfigurar el reporte desde cero.



GENERACIÓN Y EXPORTACIÓN DE REPORTES

Todos los módulos del sistema incluyen funciones de exportación de reportes en dos formatos estándar: PDF y Word (.docx). Ambas funciones están implementadas en `app.py` y son accesibles mediante botones de descarga ubicados en la barra lateral de cada módulo.

Exportación a PDF (`generar_pdf_multimodo`)

La función `generar_pdf_multimodo(secciones, titulo_reporte)` genera un documento PDF mediante la biblioteca ReportLab. La estructura del documento se organiza en secciones, donde cada sección puede contener una o múltiples tablas y un gráfico asociado. Las características técnicas del PDF son:

- **Tamaño de página:** Letter (8.5" × 11"), con márgenes de 50 puntos a los lados y 80 puntos en la parte inferior y 100 puntos en la superior.
- **Fuente:** Helvetica para el cuerpo del texto, Helvetica-Bold para encabezados de tabla y filas de Total.
- **Tablas:** Se renderizan fila a fila con control de salto de página automático cuando el contenido excede el margen inferior. Las celdas con texto extenso se dividen automáticamente en múltiples líneas.
- **Gráficos:** Se insertan como imágenes PNG generadas por Matplotlib, con altura máxima de 200 puntos y ancho completo del área útil.
- **Plantilla:** Si existe el archivo `plantilla.pdf` en la carpeta del proyecto, se aplica como fondo sobre cada página del reporte mediante la función `aplicar_plantilla()` utilizando la biblioteca `pdfw`.



REPORTE DE CASOS POR UBICACIÓN

Análisis: Comparativo General

concepto	Hombres	Mujeres	Sin Sexo	Total Cant.
Pasto	9235	15027	0	24262
Otros Municipios	15762	24898	1	40661
Total	24997	39925	1	64923



Ciudadela Universitaria Torobajo - Calle 1B No. 50-02 - Bloque Tecnológico - Piso 2
 Teléfono: (602)732283 - Celular 321865085 - Línea Gratuita 018000957071
 Correo electrónico: casos@udenar.edu.co - www.udenar.edu.co - San Juan de Pasto - Nariño - Colombia

Institución de Educación Superior | Vigilada por MINEDUCACIÓN - Fundada mediante Decreto No. 049 del 4 de noviembre de 1904.
 Acreditada en Alta Calidad mediante Resolución No. 000022 del 11 de enero de 2023 por MINEDUCACIÓN.

Piensa en tu compromiso con el ambiente; reduce, reutiliza y separa tus residuos correctamente.



SG-CER110449

Exportación a Word (`generar_word`)

La función `generar_word(secciones, titulo_reporte)` genera un documento `.docx` mediante la biblioteca `python-docx`. El documento se crea a partir de la `plantilla.docx` si existe; de lo contrario, se genera desde cero. Las características del documento Word son:

- **Título principal:** Texto en mayúsculas, negrita, tamaño 0.3 pulgadas (aproximadamente 21pt), alineado al centro.

- **Subtítulos de sección:** Negrita, tamaño 0.15 pulgadas, con separación entre secciones mediante saltos de página.
- **Tablas:** Se utilizan con el estilo Table Grid. Los encabezados de columna aparecen en mayúsculas y negrita. Los valores numéricos enteros se formatean con separadores de miles. Las filas de Total aparecen en negrita.
- **Gráficos:** Se insertan como imágenes de 5.5 pulgadas de ancho, centradas debajo de la tabla correspondiente.



REPORTE DE CASOS POR UBICACIÓN

Comparativo General

CONCEPTO	HOMBRES	MUJERES	SIN SEXO	TOTAL CANT.
Pasto	9,235	15,027	0	24,262
Otros Municipios	15,762	24,898	1	40,661
Total	24,997	39,925	1	64,923



Ciudadela Universitaria Torobajo - Calle 18 No. 50-02 - Bloque Tecnológico - Piso 2
 Teléfono: (002)7312283 - Celular: 32005085 - Línea Gratuita 0800090707
 Correo electrónico: ces@udenar.edu.co - www.udenar.edu.co - San Juan de Pasto - Nariño - Colombia

Institución de Educación Superior (Vigilada por MINEDUCACIÓN - Fundada mediante Decreto No. 049 del 4 de noviembre de 1994.
 Acreditada en Alta Calidad mediante Resolución No. 005022 del 9 de enero de 2023 por MINEDUCACIÓN.

Plena en su compromiso con el ambiente, reduce, reutiliza y separa sus residuos correctamente.



SC-CERT1048

CAPA DE ACCESO A DATOS (database.py)

El módulo `database.py` centraliza toda la lógica de comunicación con la base de datos PostgreSQL del sistema YACHAY. Su diseño sigue el principio de responsabilidad única: provee exclusivamente la función `consultar()`, que acepta una cadena SQL arbitraria y retorna el resultado como un `DataFrame` de `pandas`.

Función `consultar()`

La función implementa el patrón de conexión por demanda (`connection-per-request`), lo que significa que cada llamada a `consultar()` crea un nuevo motor SQLAlchemy, ejecuta la consulta y libera la conexión. Este enfoque es apropiado para aplicaciones de análisis de datos donde las consultas son infrecuentes, pero potencialmente largas.

```
def consultar(query):
    url = f"postgresql://{USER}:{PASS}@{HOST}:{PORT}/{DB}"
    engine = create_engine(url)
    try:
        with engine.connect() as connection:
            return pd.read_sql(text(query), connection)
    finally:
        engine.dispose()
```

El bloque `try/finally` garantiza que `engine.dispose()` se ejecute siempre, incluso si la consulta genera una excepción, previniendo fugas de conexiones que podrían agotar el pool de conexiones de PostgreSQL en sesiones de uso intensivo.

MODELO DE DATOS DEL RPCMP

La base de datos almacena el Registro Poblacional de Cáncer del Municipio de Pasto en un modelo relacional normalizado. Las tablas principales que consulta el Portal Yachay Reports son:

Tabla	Contenido Principal	Módulos que la usan
<code>tab_fichas</code>	Registro central de casos: <code>id_ficha</code> , <code>cod_cie10</code> , <code>fe_diagnostico</code> , <code>edad_paciente</code> , <code>cod_comportamiento</code> , <code>cod_metododiagnostico</code> , <code>cod_localizacion</code> , <code>cod_morfologia</code> , <code>cod_gradodiferenciacion</code>	Todos
<code>tab_pacientes</code>	Datos demográficos del paciente: <code>id_paciente</code> , <code>codsexo</code> , <code>cod_municipio</code> , <code>cod_estratos</code>	Todos
<code>tab_historias</code>	Relación ficha-fuente (una ficha puede tener múltiples historias en distintas instituciones): <code>id_ficha</code> , <code>cod_fuentes</code>	Fuentes, Comportamiento, Grupo Étáreo
<code>tab_fuentes</code>	Catálogo de instituciones notificantes: <code>cod_fuentes</code> , <code>nom_fuentes</code>	Fuentes, Comportamiento, Grupo Étáreo
<code>tab_municipios</code>	Catálogo geográfico de municipios: <code>cod_municipios</code> , <code>nom_municipios</code> , <code>cod_departamentos</code>	Casos por Municipio, Fuentes

Tabla	Contenido Principal	Módulos que la usan
tab_departamentos	Catálogo de departamentos: cod_departamentos, nom_departamentos	Todos (filtros geográficos)
tab_cie10	Catálogo CIE-10: cod_cie10, nom_cie10	Localizaciones, Agrupación CIE-10
tab_direcciontumor	Dirección del tumor y barriocomuna del paciente en Pasto	Casos por Municipio (PASTO 2)

CONCLUSIONES

El Portal Yachay Reports constituye una herramienta analítica integral para el Registro Poblacional de Cáncer del Municipio de Pasto, que permite el acceso al análisis epidemiológico sin requerir conocimientos avanzados de programación o estadística por parte del usuario final. Sus principales aportes se pueden resumir en los siguientes puntos:

- **Modularidad temático:** Los diez módulos de análisis cubren de manera exhaustiva las dimensiones de mayor interés epidemiológico del registro: geografía, temporalidad, fuentes de notificación, características clínicas y calidad de datos.
- **Flexibilidad analítica:** El Constructor AD HOC permite la generación de reportes cruzados arbitrarios sin modificar el código fuente, adaptándose a necesidades específicas de análisis no contempladas en los módulos predefinidos.
- **Exportación estándar:** La capacidad de exportar cualquier análisis en formato PDF y Word facilita la integración de los resultados en informes institucionales, presentaciones académicas y reportes periódicos del registro.
- **Gestión de proyectos:** El sistema de guardado de configuraciones en archivos JSON permite reproducir análisis específicos en sesiones futuras, garantizando la consistencia metodológica de los reportes periódicos.
- **Arquitectura extensible:** La separación entre la capa de datos (database.py) y la capa de presentación (app.py) facilita la incorporación de nuevos módulos de análisis, nuevas fuentes de datos o cambios en el esquema de la base de datos con un impacto mínimo sobre el código existente.

TRABAJOS FUTUROS

Para garantizar que el sistema **YACHAY** evolucione de acuerdo con las exigencias de la salud pública y la seguridad informática, se proponen las siguientes líneas de desarrollo:

- **Cifrado de Extremo a Extremo:** Asegurar que todos los datos en tránsito y en reposo utilicen protocolos de cifrado actualizados, cumpliendo estrictamente con la Ley de Protección de Datos Personales (Habeas Data).
- **Modularización del Código:** Reestructurar las funciones de consulta y visualización en módulos independientes, facilitando que nuevos investigadores añadan "plugins" de análisis estadístico sin alterar el núcleo del sistema.
- **Seguridad:** En el archivo database.py, evita dejar las credenciales (User, Pass, Host) en texto plano. Se recomienda implementar **Variables de Entorno** (python-dotenv) para proteger el acceso a la base de datos de producción.

- **Mejora del Constructor AD HOC:** Permitir la exportación directa de las consultas personalizadas a formatos de intercambio de datos como CSV y Excel para análisis externos avanzados.
- **Reportes Programados:** Implementar tareas cronometradas para generar y enviar reportes técnicos mensuales de forma automática a las autoridades de salud pertinentes.