

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



## MANUAL DE USUARIO



## Tabla de contenido

1.	INTRODUCCION.....	5
1.1.	Propósito y Alcance del Sistema.....	5
1.2.	Arquitectura Técnica.....	5
2.	REQUISITOS DEL SISTEMA E INSTALACIÓN .....	6
2.1.	Requisitos Mínimos .....	6
2.2.	Instalación de Dependencias Python.....	6
2.3.	Configuración de la Base de Datos.....	6
2.4.	Ejecución del Sistema .....	6
2.5.	Archivos de Plantilla (Opcional) .....	7
3.	ACCESO Y NAVEGACIÓN EN EL PORTAL YACHAY.....	7
3.1.	Pantalla de Bienvenida .....	7
3.2.	Estructura de la Interfaz Principal .....	8
3.3.	Menú de Módulos .....	8
3.4.	Cerrar Sesión .....	8
3.5.	Panel de Personalización de Gráficos .....	8
4.	MÓDULOS DE ANÁLISIS — GUÍA DETALLADA .....	9
4.1.	Módulo: Casos por Municipio.....	9
4.1.1.	Descripción General .....	9
4.1.2.	Filtros Disponibles .....	9
4.1.3.	Interpretación de los Resultados.....	10
4.1.4.	Exportación del Reporte .....	10
4.2.	Módulo: Fuentes de Información.....	10
4.2.1.	Descripción General .....	11
4.2.2.	Filtros Disponibles .....	11
4.2.3.	Interpretación de los Resultados.....	11
4.3.	Módulo: Comportamiento por Fuente.....	11
4.3.1.	Descripción General .....	12
4.3.2.	Código de Colores.....	12
4.3.3.	Filtros Disponibles .....	12
4.3.4.	Interpretación.....	12
4.4.	Módulo: Grupo Etéreo por Fuente .....	13
4.4.1.	Grupos Etéreos Utilizados .....	13
4.4.2.	Estructura de Análisis .....	13
4.5.	Módulo: División Etérea.....	13

4.5.1.	Descripción General .....	14
4.5.2.	Tipos de Gráficos Disponibles .....	14
4.5.3.	Lectura de la Pirámide Poblacional.....	14
4.6.	Módulo: Localizaciones por Sexo .....	15
4.6.1.	Buscador de Enfermedades CIE-10 .....	15
4.6.2.	Visualización por Defecto: Top 10 Patologías.....	15
4.7.	Módulo: Diagnóstico .....	16
4.7.1.	Categorías de Métodos de Diagnóstico .....	16
4.7.2.	Interpretación.....	16
4.8.	Módulo: Agrupación CIE-10.....	17
4.8.1.	Grupos Diagnósticos Disponibles .....	17
4.8.2.	Funcionamiento de la Búsqueda.....	17
4.8.3.	Métrica Principal .....	17
4.9.	Módulo: Frecuencia Datos Perdidos .....	18
4.9.1.	Variables de Integridad Analizadas.....	18
4.9.2.	Instrucción de Uso.....	18
4.10.	Módulo: Proyectos y Reportes Libres (AD HOC).....	19
4.10.1.	Variables Disponibles para Cruce.....	19
4.10.2.	Configuración del Reporte .....	19
4.10.3.	Comportamiento Adaptativo según Número de Variables.....	20
4.10.4.	Gestión de Proyectos Guardados .....	20
5.	GENERACIÓN Y EXPORTACIÓN DE REPORTES .....	20
5.1.	Formatos de Exportación Disponibles.....	20
5.2.	Proceso de Exportación Paso a Paso.....	21
5.3.	Estructura del Documento Exportado .....	21
5.4.	Consideraciones Sobre el Tamaño de los Reportes .....	21
6.	CONFIGURACIÓN Y PERSONALIZACIÓN .....	21
6.1.	Personalización de Gráficos por Módulo.....	22
6.1.1.	Tipo de Gráfico .....	22
6.1.2.	Mostrar Análisis Porcentual .....	22
6.1.3.	Título del Reporte .....	22
6.2.	Filtros Combinados.....	22
7.	GUÍA DE CONSULTA RÁPIDA .....	23
8.	DESCRIPCIÓN TÉCNICA — ESTRUCTURA DE LA BASE DE DATOS .....	23
8.1.	Tablas Principales .....	23
8.2.	Módulo de Conexión a Base de Datos (database.py) .....	24

9. CONCLUSIONES ..... 24

## 1. INTRODUCCION

El presente documento constituye el Manual de Usuario del Sistema YACHAY — Portal Analítico de Reportes Epidemiológicos, una plataforma tecnológica desarrollada para el Registro Poblacional de Cáncer del municipio de Pasto. El propósito fundamental de esta guía es dotar a los usuarios de los conocimientos necesarios para operar la plataforma de manera autónoma, abarcando desde el acceso al sistema hasta la generación, personalización y exportación de reportes estadísticos.

YACHAY es una aplicación web construida con Python y Streamlit, que se conecta directamente a la base de datos PostgreSQL del Registro Poblacional de Cáncer (RPCMP). A través de una interfaz intuitiva, el sistema permite generar análisis epidemiológicos complejos sin necesidad de conocimientos técnicos en bases de datos o programación. Los resultados se presentan en forma de tablas dinámicas y gráficos interactivos, con posibilidad de exportación en formatos PDF y Word.

### 1.1. Propósito y Alcance del Sistema

YACHAY responde a la necesidad de modernizar el proceso de generación de informes del Registro Poblacional de Cáncer. Antes de su implementación, la producción de reportes epidemiológicos requería consultas manuales a la base de datos y el procesamiento de datos en hojas de cálculo, lo que implicaba un alto consumo de tiempo y riesgo de errores. El sistema automatiza este proceso, ofreciendo los siguientes beneficios:

- Generación instantánea de reportes estadísticos directamente desde la base de datos oficial.
- Visualización interactiva mediante múltiples tipos de gráficos (barras, circular, bidireccional, entre otros).
- Filtrado avanzado por municipio, departamento, año de diagnóstico, institución notificante y otras variables.
- Exportación de reportes en formato PDF (con plantilla institucional) y Word (.docx).
- Personalización de la apariencia gráfica mediante paletas de colores configurables.
- Análisis porcentual adicional activable con un interruptor.

### 1.2. Arquitectura Técnica

El sistema se apoya en los siguientes componentes tecnológicos:

Componente	Descripción
Python / Streamlit	Framework de interfaz web que sirve la aplicación en el navegador sin necesidad de instalar software adicional en el cliente.
PostgreSQL	Motor de base de datos relacional que almacena toda la información del Registro Poblacional de Cáncer (BDYachay).
SQLAlchemy	Librería ORM que gestiona la conexión segura entre la aplicación Python y la base de datos PostgreSQL.
Pandas / Matplotlib	Librerías Python para el procesamiento de datos tabulares y la generación de gráficos estadísticos.
ReportLab	Motor de generación de documentos PDF con soporte para plantillas institucionales.
python-docx	Librería para la generación de documentos Word (.docx) con tablas y gráficos incrustados.

## 2. REQUISITOS DEL SISTEMA E INSTALACIÓN

### 2.1. Requisitos Mínimos

Para ejecutar YACHAY de forma óptima, se debe contar con los siguientes requisitos:

Requisito	Especificación
Sistema Operativo	Ubuntu 22.04 LTS o superior (recomendado). Windows 10 o superior
Python	Versión 3.10 o superior. Se recomienda Python 3.12
PostgreSQL	Versión 14 o superior con la base de datos BDYachay configurada y restaurada.
RAM	Mínimo 4 GB. Recomendado 8 GB para consultas con grandes volúmenes de datos.
Almacenamiento	Mínimo 10 GB libres para la base de datos y los archivos del sistema.
Navegador Web	Google Chrome 90+, Mozilla Firefox 88+ o Microsoft Edge 90+. No requiere instalación en el cliente.

### 2.2. Instalación de Dependencias Python

El sistema utiliza el gestor de paquetes UV para la instalación de dependencias. Ejecutar los siguientes comandos en la terminal del servidor:

```
# Instalar Python 3.12
sudo python install 3.12
```

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

### 2.3. Configuración de la Base de Datos

El sistema se conecta a PostgreSQL con los parámetros definidos en el archivo database.py. Verificar que los valores correspondan al entorno de despliegue:

Parámetro	Valor por Defecto
HOST	localhost
PORT	5432
DB	BDYachay
USER	postgres
PASS	*****

### 2.4. Ejecución del Sistema

Una vez configurado el entorno, el sistema se inicia con el siguiente comando desde el directorio donde se encuentra el archivo app.py:

```
streamlit run app.py
```

El sistema se abrirá automáticamente en el navegador predeterminado en la dirección <http://localhost:8501>.

```
http://localhost:8501
```

## 2.5. Archivos de Plantilla (Opcional)

El sistema YACHAY soporta la aplicación de plantillas institucionales en los reportes exportados. Para activar esta funcionalidad, se deben colocar los siguientes archivos en el mismo directorio que app.py:

- `plantilla.pdf`: Archivo PDF con el membrete institucional (logotipo, encabezado, pie de página). Al exportar reportes en PDF, el contenido generado se superpone sobre esta plantilla.
- `plantilla.docx`: Archivo Word con estilos y formato institucional. Sirve como base para los reportes exportados en formato Word.

Si estos archivos no están presentes, el sistema exportará los reportes con un formato neutro predeterminado, sin afectar la funcionalidad del análisis.

## 3. ACCESO Y NAVEGACIÓN EN EL PORTAL YACHAY

### 3.1. Pantalla de Bienvenida



Al acceder a la URL del sistema en el navegador web, el usuario visualizará la Pantalla de Bienvenida del Portal Yachay Reports. Esta pantalla actúa como portal de entrada al sistema.

La pantalla principal presenta:

- Un título central que identifica la plataforma: "Portal Yachay Reports".
- Un botón prominente etiquetado `INGRESAR A REPORTE` que ocupa el ancho completo de la interfaz.

Para acceder al sistema, el usuario debe hacer clic en el botón INGRESAR A REPORTE. El sistema cargará inmediatamente el módulo de análisis principal, habilitando el menú lateral de navegación.

### 3.2. Estructura de la Interfaz Principal

Una vez dentro del sistema, la interfaz se divide en dos áreas principales:

Área	Descripción
Panel Lateral Izquierdo (Sidebar)	Contiene el selector de módulo, los filtros específicos de cada reporte, las opciones de personalización gráfica y los botones de descarga. Es el panel de control principal del usuario.
Área Central de Contenido	Muestra los resultados del análisis: tablas de datos, gráficos estadísticos y métricas clave. Se actualiza automáticamente cuando el usuario modifica los filtros del panel lateral.

### 3.3. Menú de Módulos

En la parte superior del panel lateral izquierdo se encuentra el selector de módulos etiquetado como "Módulo:". Al hacer clic sobre él, se despliega la lista completa de módulos de análisis disponibles en el sistema:

- Casos por Municipio — Distribución geográfica de casos por municipio y departamento.
- Fuentes de Información — Análisis de instituciones notificantes (hospitales, laboratorios, etc.).
- Comportamiento por Fuente — Distribución del comportamiento biológico del tumor por institución.
- Grupo Etéreo por Fuente — Matriz de distribución de grupos de edad por institución.
- División Etérea — Pirámide poblacional y distribución quinquenal de casos por sexo.
- Localizaciones por Sexo — Análisis de tipos de cáncer por código CIE-10 y género.
- Diagnóstico — Distribución de métodos de diagnóstico utilizados.
- Agrupación CIE-10 — Análisis detallado por grupos diagnósticos técnicos.
- Frecuencia Datos Perdidos — Indicadores de integridad y calidad de la base de datos.
- Proyectos y Reportes Libres — Constructor AD HOC para reportes personalizados.

### 3.4. Cerrar Sesión

Para salir del portal de reportes y regresar a la pantalla de bienvenida, el usuario debe hacer clic en el botón Cerrar Sesión ubicado en la parte inferior del panel lateral izquierdo. El sistema limpia el estado de la sesión y regresa a la pantalla inicial sin necesidad de cerrar el navegador.

### 3.5. Panel de Personalización de Gráficos

En la parte inferior del panel lateral izquierdo, disponible en todos los módulos, se encuentra la sección "Personalizar Gráficos". Esta sección permite controlar los colores de los gráficos de Hombres/Mujeres de forma global para toda la sesión.

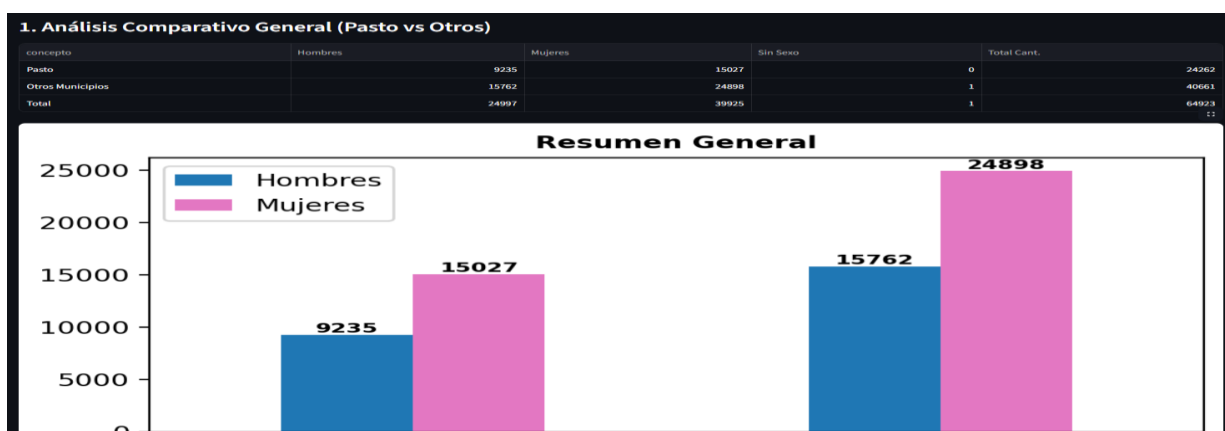
El selector Paleta de colores ofrece las siguientes opciones predefinidas:

- Estándar (Azul/Rosa): Colores por defecto. Azul para Hombres (#1f77b4) y rosa para Mujeres (#e377c2).
- Vibrante (Cian/Naranja): Cian (#00d2ff) para Hombres y naranja (#ff8c00) para Mujeres.
- Profesional (Gris/Azul): Gris oscuro (#2c3e50) para Hombres y azul (#3498db) para Mujeres.
- Personalizado: Habilita dos selectores de color individuales para que el usuario elija libremente los colores de Hombres y Mujeres.

#### 4. MÓDULOS DE ANÁLISIS — GUÍA DETALLADA

Esta sección describe en detalle cada uno de los módulos de análisis disponibles en el sistema YACHAY, incluyendo sus funcionalidades, filtros disponibles, interpretación de los resultados y opciones de exportación.

##### 4.1. Módulo: Casos por Municipio



Este módulo presenta la distribución geográfica de los casos de cáncer registrados, permitiendo comparar la carga de casos entre diferentes municipios y departamentos del área de cobertura del registro.

##### 4.1.1. Descripción General

Al acceder a este módulo, el sistema genera automáticamente un análisis comparativo de referencia que confronta los casos registrados en el municipio de Pasto contra los provenientes de otros municipios del área de influencia. Este análisis inicial no requiere configuración adicional y sirve como punto de partida para cualquier reporte.

##### 4.1.2. Filtros Disponibles

<b>Filtro</b>	<b>Descripción y Uso</b>
Mostrar Análisis Porcentual	Interruptor (toggle) que, al activarse, agrega columnas de porcentaje (% H, % M, % Total) a las tablas de datos. Ideal para reportes donde se requiere mostrar la proporción relativa de cada municipio.
Filtrar por Departamento	Selector múltiple que permite restringir el análisis a uno o varios departamentos específicos. Al seleccionar un departamento, el filtro de municipio se actualiza automáticamente para mostrar solo los municipios de ese departamento.
Filtrar por Municipio	Selector múltiple para elegir municipios específicos.
Filtrar por Año	Selector múltiple que permite analizar periodos específicos. Soporta el valor "Desconocido" para incluir casos con fecha de diagnóstico no registrada.
Tipo de Gráfico	Selector del tipo de visualización: Barras Vertical, Barras Horizontal, Barras Apiladas, Circular (Torta) o Bidireccional.
Título del Reporte	Campo de texto para personalizar el título que aparecerá en el reporte exportado.

### 4.1.3. Interpretación de los Resultados

El módulo presenta dos bloques de información:

- Bloque 1 — Análisis Comparativo General: Muestra la tabla y gráfico con el resumen global Pasto vs. Otros Municipios. Este bloque aparece siempre, independientemente de los filtros.
- Bloque 2 — Análisis Filtrado: Aparece únicamente cuando el usuario ha seleccionado al menos un filtro (departamento, municipio o año). Muestra el detalle correspondiente a la selección realizada.

Las columnas de las tablas tienen el siguiente significado:

- concepto: Nombre del municipio o categoría geográfica.
- Hombres / Mujeres: Número absoluto de casos registrados por sexo.
- Total Cant.: Suma total de casos (Hombres + Mujeres).
- % H / % M / % Total (si activado): Porcentaje que representa cada valor respecto al total general.

### 4.1.4. Exportación del Reporte

Una vez generado el análisis, en la parte inferior del panel lateral aparecen dos botones de descarga:

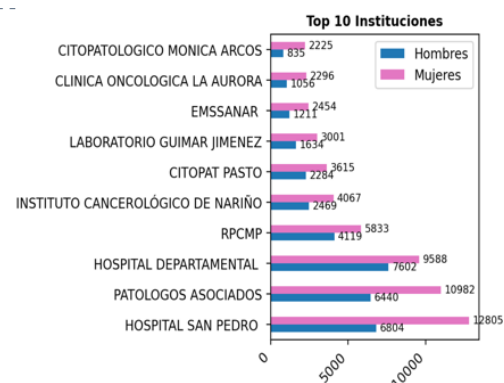
- Descargar PDF: Genera un documento PDF con la tabla de datos y el gráfico del reporte.
- Descargar Word: Genera un archivo .docx con el mismo contenido.

## 4.2. Módulo: Fuentes de Información

## Análisis de Instituciones y Fuentes de Datos

### 1. Top 10 Instituciones de Salud

concepto	Hombres	% H	Mujeres	% M	Total Casos	% 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	4.43%	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%
Total	34454	37.73%	56866	62.27%	91321	100.00%



Este módulo analiza las instituciones y fuentes de datos que notifican casos al Registro Poblacional de Cáncer. Permite identificar qué hospitales, laboratorios y otras instituciones son los principales aportantes de información.

#### 4.2.1. Descripción General

Al ingresar al módulo, el sistema genera automáticamente un reporte del Top 10 de instituciones con mayor número de historias clínicas registradas. Este bloque inicial permite al usuario identificar rápidamente las instituciones con mayor volumen de notificación.

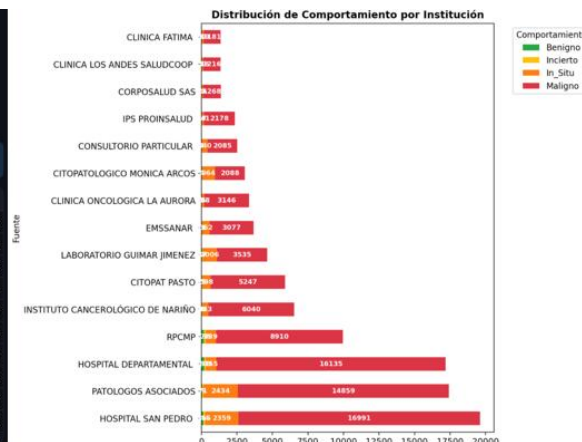
#### 4.2.2. Filtros Disponibles

Filtro	Descripción y Uso
Mostrar Análisis Porcentual	Activa columnas de porcentaje en la tabla de datos.
Filtrar por Institución	Selector múltiple para elegir una o varias instituciones específicas del listado de fuentes registradas.
Filtrar por Departamento	Restringe el análisis a casos de pacientes de un departamento específico.
Filtrar por Municipio	Restringe el análisis a casos de pacientes de municipios específicos. Incluye la opción "Otros Municipios" para filtrar casos fuera del municipio de Pasto.
Filtrar por Año	Permite analizar el comportamiento de las fuentes en periodos específicos.
Tipo de Gráfico	Opciones: Barras Vertical, Barras Horizontal o Circular (Torta).

#### 4.2.3. Interpretación de los Resultados

El análisis muestra el volumen de casos reportado por cada institución, desglosado por sexo. Esto permite identificar no solo las instituciones más activas en notificación, sino también si existe alguna asimetría en el registro de casos por género entre diferentes fuentes.

### 4.3. Módulo: Comportamiento por Fuente



Este módulo analiza la distribución del comportamiento biológico de los tumores reportados por cada institución, categorizados en cuatro grupos: Benigno, Incierto, In Situ y Maligno.

#### 4.3.1. Descripción General

El comportamiento biológico es una variable clínica fundamental en el registro oncológico. Este módulo permite evaluar si las instituciones reportan casos con perfiles similares o si existen diferencias sistemáticas en el tipo de tumores que cada fuente notifica, lo que puede indicar sesgos en la captación o diferencias en la especialización clínica de cada entidad.

#### 4.3.2. Código de Colores

Los gráficos de este módulo utilizan un código de colores fijo (semáforo epidemiológico) independiente de la paleta seleccionada en el panel de personalización:

Comportamiento Biológico	Color en el Gráfico
Benigno	Verde (#28a745) — Tumores de comportamiento no invasivo.
Incierto	Amarillo (#ffc107) — Tumores con comportamiento biológico indeterminado.
In Situ	Naranja (#fd7e14) — Tumores confinados al epitelio, sin invasión.
Maligno	Rojo (#dc3545) — Tumores invasivos con potencial metastásico.

#### 4.3.3. Filtros Disponibles

El módulo ofrece filtros por Departamento, Municipio, Institución específica y Año de diagnóstico. El tipo de gráfico puede seleccionarse entre Barras Apiladas Horizontal, Barras Apiladas Vertical, Barras Agrupadas y Líneas.

#### 4.3.4. Interpretación

La tabla muestra, para cada institución, el número de casos clasificados en cada categoría de comportamiento biológico, además del total de historias clínicas analizadas. Las tablas largas se

limitan a las Top 10 instituciones para mayor legibilidad visual, aunque el total general siempre se incluye en la última fila.

#### 4.4. Módulo: Grupo Etéreo por Fuente

Distribución: Hombres									Distribución: Mujeres									Etéreo por Institución General								
FUENTE	SIN DATO	00-04	05-09	10-14	15-19	20-24	25-29	30	FUENTE	SIN DATO	00-04	05-09	10-14	15-19	20-24	25-29	30	Matriz Detallada (Frecuencias)								
																		FUENTE	SIN DATO	00-04	05-09	10-14	15-19	20-24	25-29	30
HOSPITAL DEPARTAMENTAL	32	18	23	37	92	146	181		HOSPITAL SAN PEDRO	51	12	10	16	76	205	406		HOSPITAL SAN PEDRO	71	22	24	24	145	332	521	
HOSPITAL SAN PEDRO	20	10	14	8	69	127	115		PATOLOGOS ASOCIADOS	38	47	50	56	83	211	437		PATOLOGOS ASOCIADOS	60	115	108	116	147	296	542	
PATOLOGOS ASOCIADOS	22	68	58	60	64	85	105		HOSPITAL DEPARTAMENTAL	24	25	19	20	71	151	259		HOSPITAL DEPARTAMENTAL	56	43	42	57	163	297	440	
RPCMP	11	24	22	27	37	53	46		RPCMP	13	20	21	22	25	57	115		RPCMP	24	44	43	49	62	110	161	
INSTITUTO CANCEROLÓGICO DE NARIÑO	0	11	9	14	25	39	49		INSTITUTO CANCEROLÓGICO DE NARIÑO	1	13	12	13	25	58	102		INSTITUTO CANCEROLÓGICO DE NARIÑO	1	24	21	27	50	97	151	
CITOPAT PASTO	11	4	1	1	10	21	18		CITOPAT PASTO	17	1	1	3	8	49	101		CITOPAT PASTO	28	5	2	4	18	70	119	
LABORATORIO GUIMAR JIMENEZ	11	2	1	1	7	25	26		LABORATORIO GUIMAR JIMENEZ	16	0	0	4	13	63	140		LABORATORIO GUIMAR JIMENEZ	27	2	1	5	20	88	166	
EMSSANAR	3	16	9	9	10	16	23		EMSSANAR	10	13	9	8	12	48	103		EMSSANAR	13	29	18	17	22	64	126	
CLINICA ONCOLOGICA LA AURORA	0	1	0	0	10	22	26		CLINICA ONCOLOGICA LA AURORA	2	0	0	2	13	30	79		CLINICA ONCOLOGICA LA AURORA	2	1	0	2	23	52	105	
IPS PROINSALUD	1	2	1	3	3	5	4		CITOPATOLOGICO MONICA ARCOS	146	1	1	3	7	36	133		CITOPATOLOGICO MONICA ARCOS	234	2	1	5	13	38	141	

Este módulo genera una matriz de datos que cruza los grupos de edad quinquenales con las instituciones de salud notificantes, permitiendo identificar si cada fuente registra casos predominantemente en grupos de edad específicos.

##### 4.4.1. Grupos Etéreos Utilizados

El sistema clasifica a los pacientes en los siguientes grupos quinquenales estándar:

- SIN DATO: Pacientes con edad no registrada o código de edad inválido.
- 00-04, 05-09, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39: Grupos de edad en años.
- 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84: Grupos de edad en años.
- 85+: Pacientes de 85 años o más.

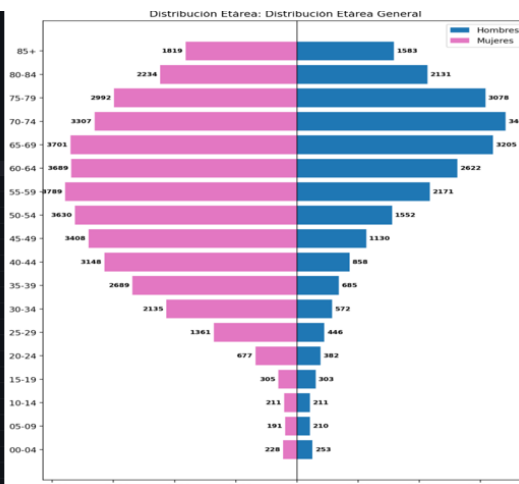
##### 4.4.2. Estructura de Análisis

Este módulo presenta tres matrices complementarias:

- General: Distribución de todos los casos (ambos sexos combinados) por fuente y grupo etéreo.
- Hombres: Distribución exclusiva de casos masculinos por fuente y grupo etéreo.
- Mujeres: Distribución exclusiva de casos femeninos por fuente y grupo etéreo.

Cada sección incluye su propia tabla y gráfico, con la opción de exportar el conjunto completo como reporte PDF.

#### 4.5. Módulo: División Etérea



Este módulo genera la distribución etárea de los casos de cáncer en forma de pirámide poblacional y otros tipos de gráficos, permitiendo analizar el perfil de edad de los pacientes registrados.

#### 4.5.1. Descripción General

La distribución etárea es uno de los indicadores más importantes en epidemiología oncológica, ya que permite identificar en qué grupos de edad se concentra la carga de enfermedad. Este módulo es fundamental para la elaboración de informes que describen el perfil demográfico del registro.

#### 4.5.2. Tipos de Gráficos Disponibles

Tipo de Gráfico	Descripción y Cuándo Usarlo
Bidireccional (Pirámide)	El gráfico más recomendado para este módulo. Muestra a los hombres hacia la derecha y a las mujeres hacia la izquierda, con los grupos etáreos en el eje central. Facilita la comparación visual del perfil de edad entre sexos.
Barras Apiladas	Muestra hombres y mujeres en barras apiladas por grupo etáreo. Útil para visualizar el total acumulado en cada grupo.
Barras Vertical	Presenta hombres y mujeres como barras separadas por grupo etáreo. Recomendado cuando se requiere comparar con precisión los valores de cada sexo en cada grupo.
Líneas	Representa la evolución de los casos por grupo etáreo como una línea continua. Útil para identificar tendencias y picos de concentración.

#### 4.5.3. Lectura de la Pirámide Poblacional

En el gráfico de pirámide bidireccional, los grupos etáreos se listan en el eje vertical (de menor a mayor edad). Las barras se extienden hacia la derecha para los Hombres y hacia la izquierda para las Mujeres. La longitud de cada barra es proporcional al número de casos en ese grupo etáreo. Las etiquetas numéricas en el extremo de cada barra muestran el valor exacto de casos.

**Nota:** Los casos con edad no registrada (código = 999 en la base de datos) se agrupan en la categoría 'Sin Dato'. En el gráfico de pirámide, esta categoría se excluye para mayor claridad visual, pero se incluye en la tabla de datos.

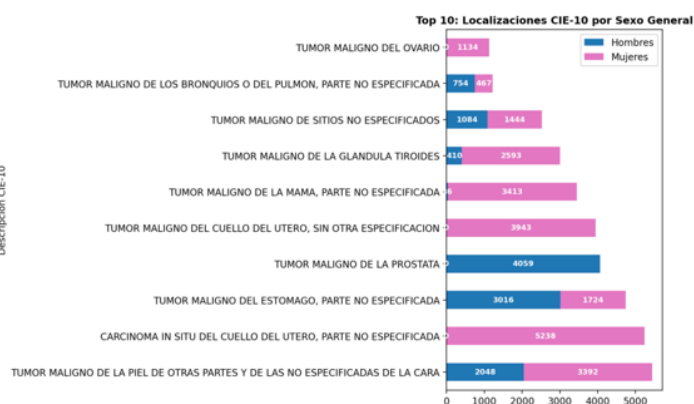
**Análisis CIE-10 por Sexo**

**Localizaciones CIE-10 por Sexo General**

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
C189	TUMOR MALIGNO DEL ESTOMAGO, PARTE NO ESPECIFICADA	3016	1724	4740
C81	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
C88	TUMOR MALIGNO DE SITIOS NO ESPECIFICADOS	1084	1444	2528
C349	TUMOR MALIGNO DE LOS BRONQUIOS O DEL PULMON, PARTE NO ESPECIFICADA	754	467	1221
C58	TUMOR MALIGNO DEL OVARIO	0	1134	1134



## 4.6. Módulo: Localizaciones por Sexo

Este módulo presenta el análisis de los tipos de cáncer según su código de localización CIE-10, desglosado por sexo. Permite identificar cuáles son las patologías oncológicas más frecuentes en la población registrada.

### 4.6.1. Buscador de Enfermedades CIE-10

Una característica distintiva de este módulo es la disponibilidad de un Buscador de Enfermedad (CIE-10) en el panel lateral. Este selector permite al usuario elegir uno o varios códigos CIE-10 específicos para enfocar el análisis en patologías de interés particular, en lugar de mostrar el Top 10 general.

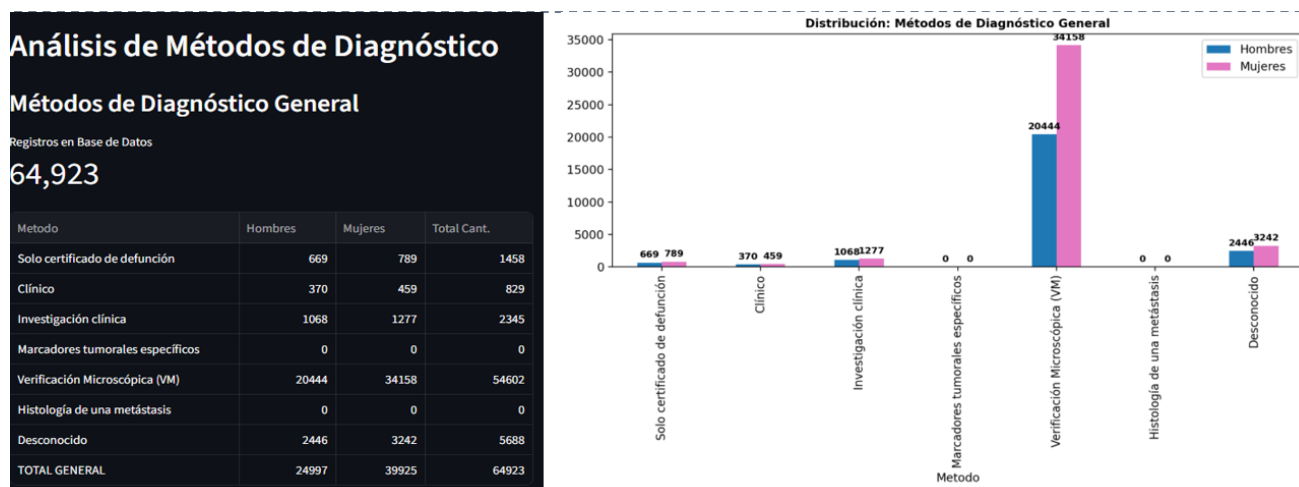
El buscador presenta las opciones en formato "CÓDIGO — NOMBRE" (por ejemplo: "C50 — Mama"), lo que facilita la búsqueda tanto por código como por nombre de la enfermedad.

### 4.6.2. Visualización por Defecto: Top 10 Patologías

Cuando no se selecciona ningún código CIE-10 específico, el módulo muestra automáticamente las 10 patologías con mayor número de casos registrados, ordenadas de mayor a menor. El gráfico se genera con estas 10 patologías principales.

La tabla completa (que puede contener más de 10 registros) siempre incluye una fila de Total General en la parte inferior, que suma todos los casos de la consulta, incluyendo aquellos que no aparecen en el Top 10 del gráfico.

## 4.7. Módulo: Diagnóstico



Este módulo analiza los métodos de diagnóstico utilizados para confirmar los casos de cáncer registrados, desagregados por sexo. Es un indicador fundamental de la calidad diagnóstica del registro.

### 4.7.1. Categorías de Métodos de Diagnóstico

Código	Método de Diagnóstico
0	Solo certificado de defunción — El caso fue captado únicamente a través del certificado de defunción, sin verificación clínica adicional.
1	Clínico — Diagnóstico basado en la presentación clínica del paciente, sin confirmación histológica.
2	Investigación clínica — Diagnóstico apoyado en estudios de imagen, endoscopias u otros procedimientos diagnósticos no invasivos.
4	Marcadores tumorales específicos — Diagnóstico confirmado mediante análisis de laboratorio de marcadores tumorales.
5 / 7	Verificación Microscópica (VM) — El método de mayor calidad. Incluye histología de tumor primario (5) y citología/hematología (7).
6	Histología de una metástasis — Diagnóstico confirmado mediante análisis histológico de una lesión metastásica.
9	Desconocido — El método de diagnóstico no fue registrado o no está disponible.

### 4.7.2. Interpretación

Un registro de alta calidad debe tener un alto porcentaje de Verificación Microscópica (VM), ya que es el método con mayor especificidad diagnóstica. Una proporción elevada de casos con método "Desconocido" o "Solo certificado de defunción" puede indicar deficiencias en el proceso de registro y debe ser un indicador de alerta para el equipo de calidad del registro.

## 4.8. Módulo: Agrupación CIE-10



Este módulo permite el análisis detallado de grupos diagnósticos técnicos predefinidos, ofreciendo un desglose de los códigos CIE-10 incluidos en cada grupo con sus respectivos conteos por sexo.

### 4.8.1. Grupos Diagnósticos Disponibles

El sistema incluye más de 60 grupos diagnósticos predefinidos basados en la clasificación CIE-10. El selector "Seleccionar Agrupación Técnica" en el panel lateral permite elegir el grupo de interés. Algunos ejemplos representativos son:

- Mama (C50): Análisis exclusivo de casos de cáncer de mama.
- Prostata (C61): Análisis del cáncer de próstata.
- Tráquea, Bronquios y Pulmón (C33, C34): Incluye todos los códigos asociados a estos órganos.
- Estómago (C16): Análisis del cáncer gástrico.
- Cuello uterino (C53): Análisis del cáncer cervicouterino.
- Leucemia linfóide (C91), Leucemia Mieloide (C92-C94): Grupos hematológicos diferenciados.
- Inespecíficos (C26, C39, C48, C76, C80): Agrupa los códigos de topografía inespecífica o no bien definida.

### 4.8.2. Funcionamiento de la Búsqueda

El sistema utiliza búsqueda por prefijo (operador SQL LIKE con comodín %). Esto significa que al seleccionar, por ejemplo, el grupo "Colon (C18)", el sistema recupera todos los casos con código CIE-10 que comienzan con "C18", incluyendo subcategorías como C18.0, C18.1, C18.2, etc. Esto garantiza que ningún caso sea omitido por diferencias en el nivel de especificidad del código registrado.

### 4.8.3. Métrica Principal

La interfaz muestra en la parte superior una métrica destacada: "Casos Identificados en este Grupo", que representa el total de registros que corresponden al grupo seleccionado con los filtros activos. Este valor puede usarse como referencia directa para los informes.

#### 4.9. Módulo: Frecuencia Datos Perdidos

### Análisis de Integridad y Calidad de Datos

Selecciona variables para analizar Integridad:

Edad x
Sexo x
Fecha Diagnóstico x
Método Diagnós... x
Localización x
Morfología x
Grado Diferencia... x

Variable: Edad				Variable: Método Diagnóstico				Variable: Morfología						
concepto	Hombres	Mujeres	Total Cant.	concepto	Hombres	Mujeres	Total Cant.	concepto	Hombres	Mujeres	Total Cant.			
Edad Conocida		24799	39514	64314	Método Conocido		22551	36683	59235	Morfología Conocida		19927	33344	53272
Edad Desconocida		198	411	609	Método Desconocido		2446	3242	5688	Neoplasia Maligna NOS		5070	6581	11651
Total		24997	39925	64923	Total		24997	39925	64923	Total		24997	39925	64923

Variable: Fecha Diagnóstico				Variable: Localización				Variable: Grado Diferenciación						
concepto	Hombres	Mujeres	Total Cant.	concepto	Hombres	Mujeres	Total Cant.	concepto	Hombres	Mujeres	Total Cant.			
Fecha Conocida		24055	38438	62494	Sitio Conocido		24015	38562	62578	Grado Conocido		7487	9551	17038
Fecha Desconocida		942	1487	2429	Sitio Desconocido		982	1363	2345	Grado Desconocido		17510	30314	47855
Total		24997	39925	64923	Total		24997	39925	64923	Total		24997	39925	64923

Este módulo es una herramienta de control de calidad del registro. Permite identificar la proporción de casos que tienen valores ausentes o desconocidos en variables clínicas y demográficas clave.

##### 4.9.1. Variables de Integridad Analizadas

El usuario puede seleccionar una o varias de las siguientes variables para analizar su nivel de completitud:

Variable	Criterio de Dato Perdido
Edad	Se considera dato perdido si el código de edad registrado es = 999 o si el valor es nulo.
Sexo	Se considera desconocido si el código de sexo es 9 (código estándar para "No especificado" en el sistema).
Fecha Diagnóstico	Se considera perdida si el año extraído es '9999' o si la fecha comienza con '99'.
Método Diagnóstico	Se considera desconocido si el código del método es 9.
Localización	Se considera desconocida si el código de localización es '809' (Localización Primaria Desconocida).
Morfología	Se considera indeterminada si el código morfológico es '8000' (Neoplasia Maligna, No Especificada).
Grado Diferenciación	Se considera desconocido si el código de grado es 9.

##### 4.9.2. Instrucción de Uso

Para utilizar este módulo, el usuario debe:

- Seleccionar los filtros geográficos y temporales deseados en el panel lateral (Departamento, Municipio, Año de Diagnóstico).
- En el área central, seleccionar una o varias variables del selector múltiple "Selecciona variables para analizar integridad:".
- El sistema genera automáticamente una tabla y gráfico para cada variable seleccionada, mostrando la distribución de casos entre "Conocido" y "Desconocido" para esa variable, desglosado por sexo.

#### 4.10. Módulo: Proyectos y Reportes Libres (AD HOC)

Este módulo es el más versátil del sistema. Permite al usuario construir reportes personalizados cruzando libremente las variables disponibles en la base de datos, sin restricciones predefinidas de análisis.

##### 4.10.1. Variables Disponibles para Cruce

El catálogo de variables disponibles para el análisis libre incluye las siguientes dimensiones:

Variable	Descripción
Departamento	Departamento de residencia del paciente.
Municipio	Municipio de residencia del paciente.
Enfermedad (CIE-10)	Nombre descriptivo del diagnóstico oncológico según CIE-10.
Código CIE-10	Código alfanumérico del diagnóstico.
Edad	Edad del paciente al momento del diagnóstico.
Año	Año de diagnóstico extraído de la fecha de diagnóstico.
Estrato	Estrato socioeconómico del paciente.
Sexo	Sexo del paciente (Hombre/Mujer).
Fuente	Institución notificante del caso.

##### 4.10.2. Configuración del Reporte

El panel de configuración del constructor AD HOC permite definir:

- **Título del Reporte:** Texto descriptivo que identificará el reporte en el encabezado y en los archivos exportados.
- **Variables a cruzar:** Selector múltiple para elegir una o más variables del catálogo. El orden de selección importa: la primera variable se usará como dimensión principal.
- **Tipo de Visualización:** Barras, Líneas o Área.
- **Color del tema:** Selector de color para los gráficos de una sola variable.

#### 4.10.3. Comportamiento Adaptativo según Número de Variables

El constructor AD HOC adapta la visualización según el número de variables seleccionadas:

Variables Seleccionadas	Comportamiento del Sistema
1 variable	Genera un gráfico simple (barras horizontales, líneas o área) mostrando el Top 15 de categorías de esa variable ordenadas por número de casos.
2 variables	Genera una tabla pivote cruzando las dos variables. El gráfico muestra el Top 15 de la variable principal vs. las Top 8 categorías de la segunda variable.
3 o más variables	Genera dos vistas complementarias: un gráfico de ranking de la variable principal (agrupando las demás) y un cruce de las dos primeras variables. Las variables adicionales se incluyen en la tabla de datos.

#### 4.10.4. Gestión de Proyectos Guardados

Una funcionalidad avanzada de este módulo es la posibilidad de guardar y cargar configuraciones de reportes como proyectos reutilizables. Para guardar un proyecto:

- En la sección "Guardar Proyecto" del panel lateral, ingresar un nombre descriptivo para el archivo en el campo "Nombre del archivo:".
- Hacer clic en el botón "Guardar en Carpeta". El sistema guarda la configuración actual (título, variables, tipo de gráfico y color) como un archivo.json en la carpeta proyectos\_reportes/ del servidor.

Para cargar un proyecto guardado:

- Desplegar la sección "Proyectos Guardados" en la parte superior del área central.
- Seleccionar el archivo.json deseado en el selector.
- Hacer clic en el botón "Cargar Configuración". El sistema aplica automáticamente la configuración guardada.

## 5. GENERACIÓN Y EXPORTACIÓN DE REPORTES

Una de las funcionalidades más importantes del sistema YACHAY es la capacidad de exportar los resultados de cualquier análisis en formatos de documentos estándar. Esta sección describe el proceso de exportación y las características de cada formato disponible.

### 5.1. Formatos de Exportación Disponibles

Formato	Características
PDF	Documento en formato PDF que incluye el título del reporte, la tabla de datos completa y el gráfico generado. Si existe el archivo plantilla.pdf en el directorio del sistema, el contenido se superpone sobre la plantilla institucional con el membrete oficial.
Word (.docx)	Documento Microsoft Word que incluye el título, la tabla de datos con formato de tabla y el gráfico como imagen incrustada. Si existe el archivo plantilla.docx, el documento utiliza los estilos definidos en esa plantilla.

## 5.2. Proceso de Exportación Paso a Paso

Para exportar un reporte en cualquier módulo, seguir los siguientes pasos:

- Configurar los filtros deseados en el panel lateral izquierdo para obtener el análisis requerido.
- Verificar que el análisis en el área central corresponda a lo deseado (tabla y gráfico correctos).
- En el panel lateral izquierdo, desplazarse hasta la sección de exportación (aparece en la parte inferior del panel una vez generado el análisis).
- Hacer clic en el botón de descarga correspondiente al formato deseado ("Descargar PDF" o "Descargar Word").
- El navegador iniciará la descarga automáticamente. El archivo se guardará con un nombre descriptivo basado en el título del reporte configurado.

## 5.3. Estructura del Documento Exportado

Cada documento exportado contiene la siguiente estructura:

- Encabezado: Título del reporte en mayúsculas, centrado.
- Subtítulo: Descripción del análisis (por ejemplo, "Comparativo General" o el subtítulo configurado por el usuario).
- Tabla de datos: Todos los datos del análisis en formato tabular. Si el reporte contiene múltiples tablas (por ejemplo, frecuencias absolutas y distribución porcentual), estas se presentan secuencialmente con etiquetas identificadoras.
- Gráfico: Imagen del gráfico generado incrustada en el documento.
- Salto de página: Cada sección del análisis ocupa su propia página en el documento.

## 5.4. Consideraciones Sobre el Tamaño de los Reportes

Para conjuntos de datos muy grandes (más de 500 filas), el módulo AD HOC limita automáticamente los resultados a 500 combinaciones mediante la cláusula LIMIT 500 en la consulta SQL. En módulos estándar, las tablas completas se exportan sin limitación de filas.

**Importante:** Si la tabla tiene muchas filas, el PDF puede tener varias páginas. El sistema maneja automáticamente los saltos de página dentro de las tablas para garantizar que los datos no se corten.

## 6. CONFIGURACIÓN Y PERSONALIZACIÓN

## 6.1. Personalización de Gráficos por Módulo

Cada módulo ofrece opciones de personalización específicas a través del panel lateral izquierdo:

### 6.1.1. Tipo de Gráfico

El selector "Tipo de Gráfico" permite elegir entre distintos tipos de visualización según el módulo. La selección del tipo más adecuado depende del análisis y del público objetivo del reporte:

Tipo de Gráfico	Mejor Para
Barras Vertical	Comparar valores entre pocos grupos (menos de 10 categorías). Las etiquetas en el eje horizontal facilitan la lectura.
Barras Horizontal	Comparar valores entre muchas categorías (más de 10). Los nombres largos en el eje vertical son más legibles.
Barras Apiladas	Mostrar la composición de cada categoría y el total acumulado simultáneamente.
Barras Apiladas Horizontal	Combinación de las ventajas de apiladas y horizontal. Recomendada para matrices por institución.
Circular (Torta)	Mostrar la proporción de cada categoría respecto al total.
Bidireccional / Pirámide	Comparar dos grupos opuestos (Hombres vs. Mujeres) simultáneamente. Ideal para distribución etérea.
Líneas	Mostrar tendencias o evolución temporal de los datos.

### 6.1.2. Mostrar Análisis Porcentual

El interruptor "Mostrar Análisis Porcentual" está disponible en todos los módulos estándar. Al activarlo, el sistema agrega automáticamente columnas de porcentaje a la tabla de datos:

- % H: Porcentaje que representan los casos de Hombres respecto al total general del análisis.
- % M: Porcentaje que representan los casos de Mujeres respecto al total general.
- % Total: Porcentaje que representa el total de esa fila respecto al total general.

Estas columnas porcentuales también se incluyen en los documentos exportados cuando el interruptor está activado al momento de la descarga.

### 6.1.3. Título del Reporte

El campo "Título del Reporte:" permite personalizar el nombre que aparecerá en el encabezado del análisis y en los archivos exportados. Se recomienda usar títulos descriptivos que incluyan la variable analizada, el periodo de tiempo y el ámbito geográfico cuando corresponda (por ejemplo: "Distribución de Casos por Municipio — Pasto — 2020-2024").

## 6.2. Filtros Combinados

Los filtros de cada módulo son acumulativos: al seleccionar valores en múltiples filtros simultáneamente, el sistema aplica todas las condiciones al mismo tiempo con lógica AND (todos los criterios deben cumplirse). Por ejemplo, seleccionar "Departamento: Nariño" y "Año: 2022" mostrará únicamente los casos de pacientes del departamento de Nariño cuyo año de diagnóstico fue 2022.

En los módulos con filtros de Departamento y Municipio, ambos filtros están vinculados dinámicamente: al seleccionar un departamento, el listado de municipios disponibles se actualiza automáticamente para mostrar únicamente los municipios de ese departamento.

## 7. GUÍA DE CONSULTA RÁPIDA

### ¿Qué Módulo usar?

La siguiente tabla orienta al usuario sobre qué módulo utilizar según el tipo de análisis que necesita realizar:

Necesidad del Análisis	Módulo Recomendado
¿Cuántos casos provienen de cada municipio o departamento?	Casos por Municipio
¿Qué hospitales notifican más casos al registro?	Fuentes de Información
¿Qué tipo de tumores (benignos/malignos) reporta cada institución?	Comportamiento por Fuente
¿Qué grupos de edad atiende principalmente cada institución?	Grupo Etéreo por Fuente
¿Cuál es la pirámide poblacional de los pacientes del registro?	División Etérea
¿Cuáles son los cánceres más frecuentes y cómo se distribuyen por sexo?	Localizaciones por Sexo
¿Con qué método se diagnostican los casos del registro?	Diagnóstico
¿Cuántos casos de cáncer de mama (o cualquier otro tipo específico) hay?	Agrupación CIE-10
¿Qué porcentaje de casos tiene datos incompletos en variables clave?	Frecuencia Datos Perdidos
Necesito cruzar variables que no están en los módulos estándar.	Proyectos y Reportes Libres (AD HOC)

## 8. DESCRIPCIÓN TÉCNICA — ESTRUCTURA DE LA BASE DE DATOS

Para usuarios avanzados y administradores del sistema, esta sección describe la estructura de las tablas de la base de datos utilizadas por YACHAY.

### 8.1. Tablas Principales

Tabla	Descripción
tab_fichas	Tabla central. Almacena un registro por cada caso de cáncer: código CIE-10, fecha de diagnóstico, método diagnóstico, comportamiento, morfología, localización y edad del paciente.
tab_pacientes	Información demográfica del paciente: sexo, código de municipio de residencia, estrato socioeconómico.
tab_municipios	Catálogo de municipios con sus códigos y nombres.
tab_departamentos	Catálogo de departamentos de Colombia.
tab_cie10	Catálogo de la Clasificación Internacional de Enfermedades — 10ª Revisión. Relaciona códigos con nombres descriptivos.
tab_historias	Registros de las historias clínicas asociadas a cada ficha, con la fuente notificante.
tab_fuentes	Catálogo de instituciones notificantes (hospitales, laboratorios, etc.).
tab_direcciontumor	Información sobre la dirección/barrio-comuna del tumor

## 8.2. Módulo de Conexión a Base de Datos (database.py)

La conexión a la base de datos se gestiona a través del archivo database.py, que expone la función consultar(query). Esta función:

- Crea una conexión al servidor PostgreSQL usando los parámetros configurados (HOST, PORT, DB, USER, PASS).
- Ejecuta la consulta SQL recibida como parámetro.
- Retorna el resultado como un DataFrame de Pandas, que es la estructura de datos principal utilizada por todos los módulos.
- Cierra la conexión y libera los recursos del motor (engine.dispose()) al finalizar.

## 9. CONCLUSIONES

El sistema YACHAY Reports, representa una herramienta fundamental para la modernización del proceso de análisis y reporte del Registro Poblacional de Cáncer del municipio de Pasto. A través de una interfaz web accesible y sin necesidad de conocimientos técnicos en bases de datos, el sistema democratiza el acceso a la información epidemiológica del registro.

Los diez módulos de análisis cubren las necesidades más frecuentes de los investigadores y gestores del registro, desde el análisis geográfico básico hasta el control de calidad de datos, pasando por análisis clínicos especializados como la distribución por comportamiento biológico, métodos de diagnóstico y clasificación CIE-10. El módulo de Reportes Libres AD HOC, por su parte, ofrece una capacidad de análisis prácticamente ilimitada, permitiendo responder preguntas de investigación que no estaban previstas en los módulos predefinidos.

La funcionalidad de exportación en PDF y Word garantiza que los resultados del análisis puedan integrarse directamente en informes institucionales, publicaciones académicas y reportes de gestión, aplicando de forma automática la plantilla institucional cuando está disponible.