

# Documentación Técnica

## Proceso de Producción Estadística del Observatorio Nacional de la Discapacidad

Automatización de procesos e implementación de  
visualizaciones para la sección “Discapacidad en Cifras”



Consejo Nacional para la Integración de la Persona con Discapacidad  
CONADIS

Dirección de Políticas, Seguimiento y Generación de Evidencia

**Versión:** 1.0

**Fecha:** Octubre 2025

**Estado:** En producción

**Elaborado por:** Alex Evanan  

Este documento técnico es para uso interno de la organización.

# Control de Cambios

Versión	Fecha	Autor	Rol	Descripción del cambio
1.0	Octubre 2025	Alex Evanan	Consultor	Versión inicial del documento técnico.

# Índice

<b>1. Resumen Ejecutivo</b>	<b>4</b>
<b>2. Alcance y Objetivos</b>	<b>5</b>
2.1. Alcance	5
2.2. Objetivos	5
<b>3. Arquitectura General</b>	<b>5</b>
3.1. Capas Funcionales	5
<b>4. Entorno Tecnológico y Requisitos</b>	<b>6</b>
4.1. Software y Herramientas	7
4.2. Requisitos de Hardware	7
4.3. Gestión de Dependencias	8
<b>5. Estructura del Proyecto</b>	<b>8</b>
5.1. Estructura de Directorios	9
5.2. Descripción de Carpetas Principales	10
<b>6. Instalación desde el Repositorio Remoto</b>	<b>10</b>
6.1. Obtención del Repositorio	11
6.2. Configuración del Entorno Virtual	11
6.3. Instalación de Dependencias	12
6.4. Actualización del Código Fuente	12
<b>7. Tableros Estadísticos</b>	<b>13</b>
7.1. Tablero ENAHO	13
7.1.1. Fuente y Organización de los Datos	13
7.1.2. Procesamiento de Datos	14
7.1.3. Integración y Actualización del Tablero	15
7.2. Tablero FISCALIZACIÓN	16
7.2.1. Fuente y Organización de los Datos	16
7.2.2. Procesamiento de Datos	17
7.2.3. Integración y Actualización del Tablero	17
7.3. Tablero REGISTRO	18
7.3.1. Fuente y Organización de los Datos	18
7.3.2. Procesamiento de Datos	19
7.3.3. Integración y Actualización del Tablero	20
7.4. Tablero SECTORES	20
7.4.1. Fuente y Organización de los Datos	21

7.4.2. Procesamiento de Datos . . . . .	22
7.4.3. Integración y Actualización del Tablero . . . . .	23
<b>8. Buenas Prácticas y Consideraciones Técnicas</b>	<b>24</b>
<b>9. Limitaciones Identificadas</b>	<b>25</b>
<b>10.Ficha Técnica</b>	<b>25</b>
<b>11.Referencias</b>	<b>26</b>

# 1. Resumen Ejecutivo

El Observatorio Nacional de la Discapacidad es una plataforma estratégica orientada a brindar información relevante, confiable, oportuna, amigable y accesible en materia de discapacidad para las entidades públicas, centros de investigación, estudiantes universitarios, entre otros; así como a la ciudadanía en general, con especial atención a las personas con discapacidad (CONADIS, [2024](#)).

El presente documento técnico describe el proceso de desarrollado, implementación y actualización de la información para el conjunto de tableros de visualización que conforman la sección “*Discapacidad en Cifras*” del Observatorio. Dicho proceso abarca la sistematización de fuentes de información provenientes de encuestas nacionales, registros administrativos y reportes sectoriales; el desarrollo de scripts para el procesamiento y cálculo de indicadores; la generación de tabulados estadísticos estandarizados; y la integración de los resultados en herramientas de visualización interactiva.

Como resultado del servicio, se implementó un flujo de trabajo reproducible y documentado que permite la actualización periódica de los tableros ENAHO, FISCALIZACIÓN, REGISTRO y SECTORES, asegurando la trazabilidad de los datos, la consistencia metodológica de los indicadores y la sostenibilidad técnica del Observatorio. El enfoque adoptado prioriza el uso de herramientas de código abierto y buenas prácticas de gestión de datos, fortaleciendo la capacidad institucional para la gestión y difusión de información sobre discapacidad.

## 2. Alcance y Objetivos

### 2.1. Alcance

El presente documento tiene por alcance la descripción técnica y operativa de los procedimientos requeridos para la actualización periódica, mantenimiento y operación del proceso de sistematización, procesamiento y visualización de información estadística que alimenta la sección “*Discapacidad en Cifras*” del Observatorio Nacional de la Discapacidad.

Quedan excluidos del alcance del presente documento los aspectos metodológicos vinculados a la definición conceptual de indicadores y variables, en tanto las estimaciones replican los reportes oficiales en el caso de encuestas y, para los registros administrativos, se sustentan en la información reportada por las entidades de los sectores correspondientes.

### 2.2. Objetivos

- Identificar de las fuentes de información utilizadas (encuestas nacionales y registros administrativos).
- Documentar los scripts y rutinas desarrolladas en **Stata** y **Python** para la automatización de los procesos de cálculo y generación de insumos.
- Describir el proceso de integración de los insumos en los tableros desarrollados en **Power BI**.

## 3. Arquitectura General

La arquitectura del proyecto ha sido diseñada bajo un enfoque modular, escalable y reproducible, que permite integrar múltiples fuentes de datos heterogéneas, procesarlas de manera estandarizada y presentar los resultados para uso institucional como público en general

Este enfoque facilita la actualización periódica de la información, asegura la trazabilidad de los indicadores y permite la incorporación progresiva de nuevas fuentes de datos sin afectar la el proceso.

### 3.1. Capas Funcionales

El proyecto se organiza en cuatro capas:

1. **Capa de datos:** comprende las fuentes de información utilizadas por el Obser-

vatorio, incluyendo encuestas nacionales (internos y externos remitidos entidades), las cuales se almacenan inicialmente en (*raw data*) preservando su estructura e integridad original para fines de validación.

2. **Capa de procesamiento:** integra los scripts desarrollados en Stata y Python que ejecutan las rutinas de limpieza, transformación, estandarización e integración de datos. En esta capa se implementan las reglas metodológicas para el cálculo de indicadores, se validan variables clave y se generan bases de datos consolidadas y tabulados estadísticos.
3. **Capa de integración:** corresponde a la preparación de archivos de salida optimizados para su consumo por herramientas de visualización así como formato accesible. En esta etapa se consolidan los resultados del procesamiento en archivos de carga estructurados.
4. **Capa de visualización:** está conformada por los tableros desarrollados en Power BI, los cuales permiten la exploración interactiva de la información mediante gráficos, tablas y filtros dinámicos. Esta capa constituye el principal punto de acceso para los usuarios finales del Observatorio y facilita la difusión de información estadística de manera accesible y oportuna.

## 4. Entorno Tecnológico y Requisitos

El sistema se ejecuta en un entorno tecnológico basado en herramientas de uso extendido para el análisis estadístico y la visualización de datos, priorizando soluciones robustas, reproducibles y de código abierto, así como software que no requiere licencias adicionales distintas a las ya disponibles en la entidad.

## 4.1. Software y Herramientas

Tabla 1: Software y herramientas

Herramienta	Versión	Tipo	Uso principal
Stata	18	Software estadístico	Procesamiento de microdatos de encuestas, cálculo de indicadores y generación de tabulados estadísticos.
Python	3.11	Lenguaje de programación	Automatización de procesos ETL, limpieza de datos, integración de fuentes y preparación de archivos de carga.
Power BI	-	Herramienta de visualización	Diseño, actualización y publicación de tableros interactivos del Observatorio.
Git & GitHub	-	Control de versiones	Gestión del código fuente, documentación y control de cambios del proyecto.
Visual Studio Code	-	Entorno de desarrollo	Edición y ejecución de scripts en Python y gestión integral del proyecto.
JupyterLab	-	Entorno interactivo	Ejecución de notebooks para análisis exploratorio y procesamiento.

## 4.2. Requisitos de Hardware

El sistema puede ejecutarse en estaciones de trabajo estándar. Como referencia, se recomienda:

- Procesador de 4 núcleos o superior.
- 16 GB de memoria RAM para el procesamiento eficiente de grandes volúmenes de datos.
- Espacio en disco suficiente para almacenar bases históricas y archivos intermedios.



### 4.3. Gestión de Dependencias

Las dependencias de Python se administran mediante un entorno virtual, cuya configuración se documenta en el archivo `requirements.txt`. Esta práctica garantiza la reproducibilidad del entorno de ejecución y reduce conflictos entre versiones de librerías.

La estructura de carpetas y rutas relativas definidas en el proyecto permiten la portabilidad del sistema entre diferentes equipos, siempre que se respeten las convenciones establecidas.

Cuadro 1: Requerimientos

```
pandas==2.2.3
numpy==2.2.2
openpyxl==3.1.5
plotly==5.24.1
kaleido==0.1.0.post1
matplotlib==3.10.1
seaborn==0.13.2
nbconvert==7.16.6
statsmodels==0.14.4
weightedcalcs==0.1.3
pyreadstat==1.2.8
geopandas==1.0.1
jupyterlab
```

## 5. Estructura del Proyecto

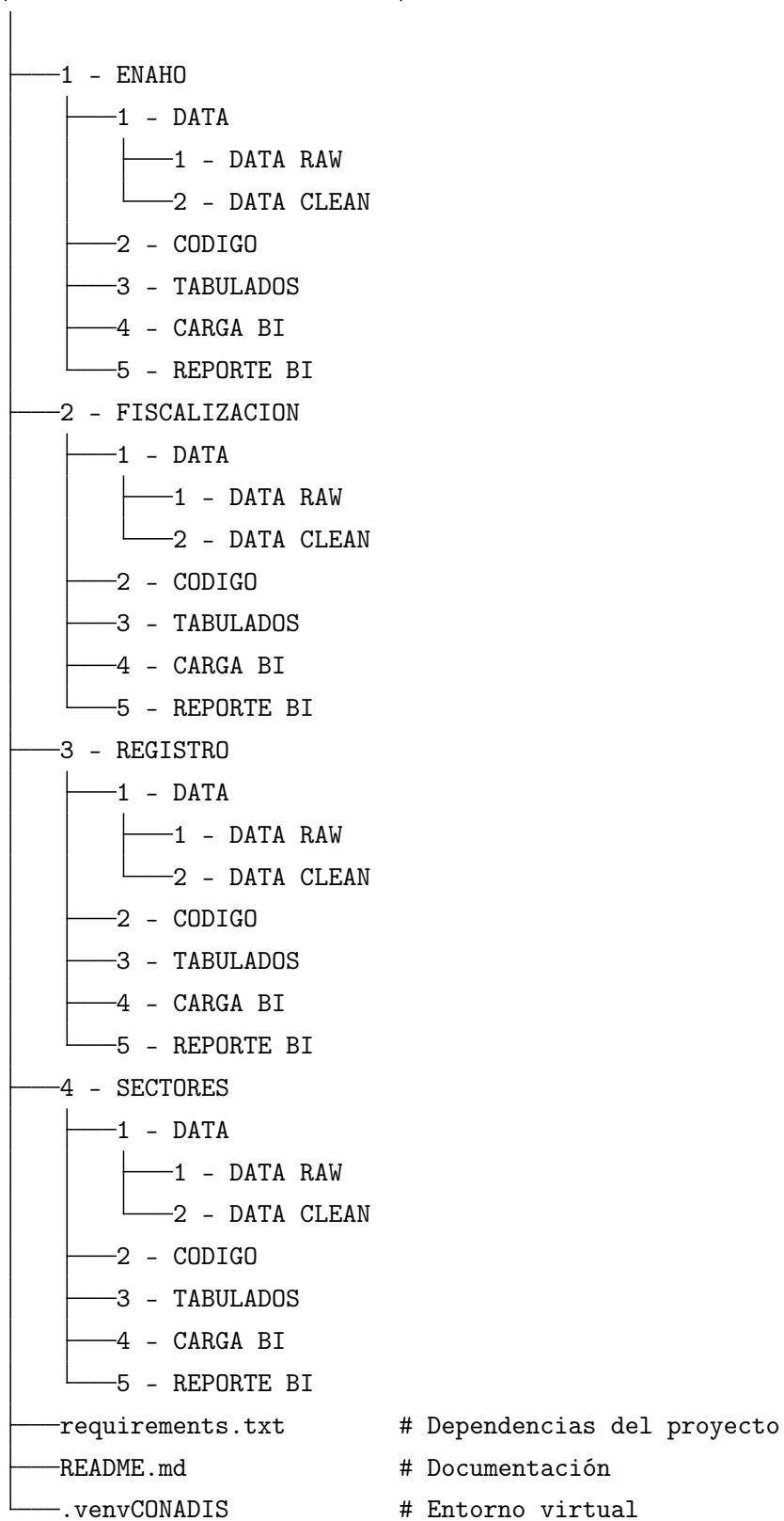
El proyecto se organiza siguiendo una estructura, diseñada para facilitar la comprensión del flujo de trabajo, la reutilización del código y el mantenimiento. Cada componente del Observatorio cuenta con una estructura interna homogénea, lo que permite escalar el proyecto a nuevas fuentes de información de manera consistente.

La estructura general del proyecto es la siguiente:

## 5.1. Estructura de Directorios

### Estructura del proyecto

/DISCAPACIDAD EN CIFRAS 2025/



## 5.2. Descripción de Carpetas Principales

Cada tablero del Observatorio (*ENAH*O, *FISCALIZACION*, *REGISTRO* y *SECTORES*) cuenta con su propia estructura interna de subcarpetas, lo que permite una gestión independiente de las actualizaciones, sin afectar a los demás componentes.

- **1 - DATA:** contiene los insumos organizados según su nivel de procesamiento y estado de validación.
  - **1 - DATA RAW:** almacena los datos originales recibidos de las distintas fuentes. Esta carpeta preserva la información sin modificaciones, garantizando la trazabilidad y el respaldo de la fuente primaria.
  - **2 - DATA CLEAN:** contiene las bases de datos depuradas y estandarizadas, resultantes de los procesos de limpieza, validación e integración. Estas bases constituyen la fuente estructurada para el cálculo de indicadores y la generación de productos estadísticos.
- **2 - CODIGO:** agrupa los scripts y rutinas de procesamiento desarrollados en Stata y Python. Estos archivos automatizan las tareas de transformación de datos, cálculo de indicadores, generación de tabulados y preparación de archivos para visualización.
- **3 - TABULADOS:** contiene los cuadros estadísticos generados a partir de las bases procesadas. Los tabulados se producen en formatos accesibles y sirven como insumo tanto para la difusión de resultados como para la validación técnica de los indicadores.
- **4 - CARGA BI:** incluye los archivos consolidados y optimizados para su carga en la herramienta de visualización Power BI. Estos archivos permiten la actualización periódica de los tableros, asegurando coherencia entre los datos procesados y los indicadores publicados.
- **5 - REPORTE BI:** contiene los archivos de Power BI (.pbix) correspondientes a cada tablero, así como otros documentos asociados a la visualización. Esta carpeta representa la capa final, orientada a la exploración interactiva y consulta de información por parte de los usuarios finales.

## 6. Instalación desde el Repositorio Remoto

Con el fin de asegurar la trazabilidad, control de cambios y reproducibilidad del desarrollo, el proyecto emplea Git como sistema de control de versiones y GitHub como

repositorio remoto. El código fuente y los recursos asociados durante el proceso de desarrollo fueron organizados en dicho repositorio, el cual fue clonado en el equipo local de CONADIS para la etapa de producción.

#### Nota

La utilización del repositorio remoto descrito en esta sección se limitó a la etapa de desarrollo del servicio de consultoría. Dicho repositorio tuvo carácter temporal y fue empleado únicamente como herramienta de trabajo del consultor.

Finalizado el servicio, los activos digitales generados (código fuente, scripts y documentación) fueron transferidos a la entidad, quedando bajo su administración.

Previo a la ejecución de los pasos de instalación del repositorio, se requiere contar con Visual Studio Code, Python, Git, Power BI y STATA instalados y configurados. Para configurar y/o instalar pueden consultarse las siguientes guías:

Tabla 2: Herramientas requeridas y guías de instalación

Herramienta	Guía	Fuente
Visual Studio Code	No requiere	<a href="#">Enlace descarga (Software libre)</a>
Python	<a href="#">Guía de instalación</a>	<a href="#">Enlace descarga (Software libre)</a>
Git	<a href="#">Guía de configuración</a>	<a href="#">Enlace descarga (Software libre)</a>
Power BI Desktop	No requiere	<a href="#">Enlace descarga (Software libre/Suscripción)</a>
STATA	No requiere	Requiere licencia

## 6.1. Obtención del Repositorio

El repositorio del proyecto debe descargarse mediante el sistema de control de versiones Git, ejecutando el siguiente comando:

```
git clone https://github.com/AlexEvanan/CONADIS.git
```

Una vez clonado el repositorio, se debe establecer el directorio principal de trabajo:

```
cd CONADIS
```

## 6.2. Configuración del Entorno Virtual

Para aislar las dependencias del proyecto y evitar conflictos con otras instalaciones del sistema, se utiliza un entorno virtual de Python.

La creación del entorno virtual se realiza mediante el siguiente comando:

```
python -m venv .venvCONADIS
```

Posteriormente, el entorno virtual debe ser activado antes de la ejecución de cualquier script:

```
.venvCONADIS\Scripts\activate
```

### 6.3. Instalación de Dependencias

Las librerías necesarias para la ejecución de los scripts se encuentran definidas en el archivo `requirements.txt`, ubicado en el directorio raíz del proyecto.

La instalación de dichas dependencias se realiza ejecutando:

```
pip install -r requirements.txt
```

### 6.4. Actualización del Código Fuente

Para mantener el entorno de trabajo alineado con la versión más reciente del proyecto, se recomienda actualizar periódicamente el repositorio desde el origen remoto:

Para actualizar el repositorio local (traer los cambios desde el repositorio remoto):

```
git pull
```

Para incluir todos cambios realizados al área de preparación se debe ejecutar el siguiente comando (enviar desde el repositorio local al remoto):

```
git add .
```

Después guardar los cambios:

```
git commit -m "Descripción del cambio"
```

Enviar los cambios al repositorio remoto:

```
git push -u origin main
```

## 7. Tableros Estadísticos

### 7.1. Tablero ENAHO

El Tablero ENAHO constituye el primer componente de la sección "*Discapacidad en Cifras*". Este tablero muestra los principales indicadores oficiales sobre las condiciones de vida de las personas con discapacidad, con desagregación temporal, territorial entre otros.

#### 7.1.1. Fuente y Organización de los Datos

La fuente principal es la [Encuesta Nacional de Hogares \(ENAHO\)](#), producida por el Instituto Nacional de Estadística e Informática (INEI). Se emplean los siguientes módulos de la encuesta:

Tabla 3: Módulos ENAHO utilizados para el Tablero ENAHO

Código del módulo	Descripción
100	Características de la vivienda y del hogar
200	Miembros del hogar
300	Educación
400	Salud
500	Empleo e ingresos
85 (1 y 2)	Gobernabilidad, democracia y transparencia
84 (800a)	Participación ciudadana
34	Sumaria

Los módulos ENAHO se descargan anualmente desde el portal oficial del INEI y los archivos se almacenan en su formato original (**.dta**) en la carpeta de datos crudos del Tablero ENAHO, siguiendo una estructura jerárquica por año, conforme al siguiente esquema:

#### Estructura de almacenamiento

```
1 - ENAHO
├── 1 - DATA
│   ├── 1 - DATA RAW
│   │   ├── 2014
│   │   ├── 2015
│   │   └── 2016
│   ├── .
│   ├── .
│   └── .
```

```

└─ 2024
    ├── enaho01-2024-100.dta
    ├── enaho01-2024-200.dta
    ├── enaho01-2024-800a.dta
    ├── enaho01a-2024-300.dta
    ├── enaho01a-2024-400.dta
    ├── enaho01a-2024-500.dta
    ├── enaho01b-2024-1.dta
    ├── enaho01b-2024-2.dta
    └── sumaria-2024.dta

```

### 7.1.2. Procesamiento de Datos

El procesamiento se realiza mediante un conjunto de scripts desarrollados en Stata, los cuales se encuentran organizados en la carpeta 1 - ENAHO/2 - CODIGO. Estos scripts conforman un pipeline reproducible que permite la integración de los módulos de la encuesta, el cálculo de indicadores y la generación de productos intermedios y finales.

La Tabla 4 presenta la relación de scripts utilizados para el procesamiento de la información ENAHO, indicando su función, insumos, productos generados y modalidad de ejecución.

Tabla 4: Scripts utilizados en el procesamiento

Orden	Do-file	Descripción	Entradas	Salidas
1	00_main.do	Script maestro que controla la ejecución completa del pipeline	Do-files del proyecto	Bases procesadas, tabulados y archivos de carga.
2	01_union_modulos.do	Unión de los módulos ENAHO para cada año	Módulos ENAHO 100, 200, 300, 400, 500, 84, 85 y Sumaria	bd2014.dta . . bd2024.dta
3	02_indicadores.do	Construcción de indicadores y consolidación histórica	bd2014.dta . . bd2024.dta	Base_calculada-2014-2024.dta
4	03_tabulado.do	Cálculo y generación de tabulados estadísticos y archivos de carga por temática	Base_calculada-2014-2024.dta	Tabulados: 2.1.1.xlsx . . 8.1.1.xlsx Carga: 2.1.1_bi.xlsx . . 8.1.1_bi.xlsx
5	04_archivos_carga.do	Consolidación de archivos de carga para Power BI	Carga: 2.1.1_bi.xlsx . . 8.1.1_bi.xlsx	2_participacion.xlsx . . 7_vivienda.xlsx

Para la correcta ejecución de los scripts en un nuevo entorno de trabajo, es necesario

actualizar el parámetro `dir_path` para que apunte a la ubicación local donde se almacena la información del Tablero ENAHO.

Cuadro 2: Definición de directorios globales en los Do-files

```
* =====
**# Directorios
* =====

* Rutas relativas
global dir_path "D:\...modificar...\1 - ENAHO"
global dir_path_data_raw "${dir_path}/1 - DATA\1 - DATA RAW"
global dir_path_data_pro "${dir_path}/1 - DATA\2 - DATA CLEAN"
global dir_path_tabulado "${dir_path}/3 - TABULADOS"
```

### 7.1.3. Integración y Actualización del Tablero

Los archivos necesarios para esta etapa se encuentran organizados en la carpeta 1 - ENAHO/5 - REPORTE BI, la cual contiene tanto los insumos consolidados como el archivo del tablero.

La Tabla 5 detalla los archivos utilizados para la actualización y publicación del tablero ENAHO.

Tabla 5: Archivos asociados al Tablero ENAHO

Archivo	Descripción
0.0.0 Tabulados-ENAHO v2025.xlsx	Consolidado de los tabulados estadísticos en un formato accesible, utilizado como insumo de referencia y verificación.
TABLERO DEC (F).pbix	Archivo del tablero desarrollado en Power BI para la visualización de los indicadores.

- **Ruta base:** en el editor de consultas de Power BI (Transformar datos → Editar parámetros), se debe modificar el parámetro `RutaCargaEnaho` para que apunte al directorio local donde se encuentran los archivos consolidados de carga, por ejemplo:

D:\...mi ruta nueva...\1 - ENAHO\4 - CARGA BI

- **Medidas DAX:** en las Medidas → fuente... se recomienda actualizar las medidas descriptivas que consignan la fuente de información y la fecha de última actualización.



- **Publicación del tablero:** una vez validados los resultados, el tablero debe publicarse en el servicio de Power BI utilizando la cuenta institucional correspondiente, seleccionando la opción de reemplazo de la versión existente.

## 7.2. Tablero FISCALIZACIÓN

El Tablero FISCALIZACIÓN es el segundo componente de la sección "*Discapacidad en Cifras*". Este tablero muestra los principales indicadores sobre las fiscalizaciones en los diferentes ejes temáticos.

### 7.2.1. Fuente y Organización de los Datos

El proveedor de información es la *Dirección de Fiscalización y Sanciones (DFS)*, la información recibida se debe almacenar en la carpeta 1 - DATA RAW.

A cada formato de base de datos históricos de la carpeta 2 - DATA CLEAN, que contiene la data de los periodos pasados, se debe incorporar como nuevas filas la data de la carpeta 1 - DATA RAW.

A continuación, se presenta la siguiente estructura jerárquica:



### 7.2.2. Procesamiento de Datos

El procesamiento se realiza mediante un scripts desarrollados en Python ubicado en la carpeta 2 - FISCALIZACION/2 - CODIGO. Estos scripts conforman un pipeline reproducible que permite la integración de la data y generación de productos intermedios y finales.

La Tabla 6 presenta la relación de scripts utilizados para el procesamiento de la información, indicando su función, insumos y productos generados.

Tabla 6: Scripts utilizados en el procesamiento

Script	Secciones	Descripción	Entradas	Salidas
Script DFS.ipynb	<ul style="list-style-type: none"> <li>Paquetes</li> <li>Fuente de datos</li> <li>Procesamiento</li> </ul>	Procesa, calcula y genera los tabulados y archivos carga de los las diferentes temáticas de Fiscalización.	<ul style="list-style-type: none"> <li>Educación.xlsx</li> <li>Empleo.xlsx</li> <li>Medio Físico.xlsx</li> <li>Presupuesto.xlsx</li> <li>Sanciones.xlsx</li> <li>TICS.xlsx</li> <li>Transporte.xlsx</li> </ul>	Tabulados: F1.1.1.xlsx : F5.1.1.xlsx Carga: DFS_educacion.xlsx DFS_empleo.xlsx DFS_mediofis.xlsx DFS_presupuesto.xlsx DFS_sanciones.xlsx DFS_tics.xlsx DFS_transporte.xlsx

Para la ejecución de los scripts no requiere la modificación de los parámetros dado que las rutas se determinan relativamente. Únicamente debe revisarse la correspondencia entre los nombres de archivos definidos en el código y los archivos de datos reales.

Cuadro 3: Definición de directorios globales en el script

```
# =====
# Directorio de trabajo
# =====

# Directorios
dir_path = os.getcwd()
path_data_raw = os.path.join(os.path.dirname(dir_path), r'1 - DATA\1 - DATA RAW')
path_data_cle = os.path.join(os.path.dirname(dir_path), r'1 - DATA\2 - DATA CLEAN')
path_output_tab = os.path.join(os.path.dirname(dir_path), r'3 - TABULADOS')
path_output_car = os.path.join(os.path.dirname(dir_path), r'4 - CARGA BI')

# Archivos
file_educacion = "Educación.xlsx"
file_empleo = "Empleo.xlsx"
file_med_fis = "Medio Físico.xlsx"
file_presup = "Presupuesto.xlsx"
file_sancion = "Sanciones.xlsx"
file_tics = "TICS.xlsx"
file_transport = "Transporte.xlsx"
```

### 7.2.3. Integración y Actualización del Tablero

Los archivos necesarios para esta etapa se encuentran organizados en la carpeta 2 - FISCALIZACION/5 - REPORTE BI, la cual contiene tanto los insumos consolidados como el archivo del tablero.

La Tabla 7 detalla los archivos utilizados para la actualización y publicación del tablero.

Tabla 7: Archivos asociados al Tablero FISCALIZACION

Archivo	Descripción
0.0.0 Tabulados-FISCALIZACION v2025.xlsx	Consolidado de los tabulados estadísticos en un formato accesible, utilizado como insumo de referencia y verificación.
TABLERO FIS (F).pbix	Archivo del tablero desarrollado en Power BI.

- **Ruta base:** en el editor de consultas de Power BI (Transformar datos → Editar parámetros), se debe modificar el parámetro RutaCarga para que apunte al directorio local donde se encuentran los archivos consolidados de carga, por ejemplo:

D:\...mi ruta nueva...\2 - FISCALIZACION\4 - CARGA BI

- **Medidas DAX:** en las Medidas → fuente... se recomienda actualizar las medidas descriptivas que consignan la fuente de información y la fecha de última actualización.
- **Publicación del tablero:** una vez validados los resultados, el tablero debe publicarse en el servicio de Power BI utilizando la cuenta institucional correspondiente, seleccionando la opción de reemplazo de la versión existente.

### 7.3. Tablero REGISTRO

El Tablero REGISTRO es el tercer componente de la sección "*Discapacidad en Cifras*". Este tablero muestra las cifras de Personas con Discapacidad registradas en el Registro Nacional de Perosnas con Discapacidad (RNPCD).

#### 7.3.1. Fuente y Organización de los Datos

El proveedor de información es la *Subdirección de Registro e Intervenciones en Discapacidad*, la iformación recibida se debe almacenar en la carpeta 1 - DATA RAW.

La información recepcionada es histórica y tiene un formato estandarizado y en función a ello esta desarrollado las rutinas.

Acontinuación, se presenta la siguiente estructura jerárquica:

## Estructura de almacenamiento

```

1 - DATA
├── 1 - DATA RAW
│   ├── RNPCD_setiembre_2025.xlsx
│   └── RNPCD_PERSONAS_JURIDICAS-CONADIS-ACTUALIZADO septiembre 2025.xlsx

```

### 7.3.2. Procesamiento de Datos

El procesamiento se realiza mediante un scripts desarrollados en Python ubicado en la carpeta 3 - REGISTRO/2 - CODIGO. Estos scripts conforman un pipeline reproducible que permite la integración de la data y generación de productos intermedios y finales.

La Tabla 8 presenta la relación de scripts utilizados para el procesamiento de la información, indicando su función, insumos y productos generados.

Tabla 8: Scripts utilizados en el procesamiento

Script	Secciones	Descripción	Entradas	Salidas
Script RNPCD.ipynb	<ul style="list-style-type: none"> <li>Paquetes</li> <li>Fuente de datos</li> <li>Procesamiento</li> </ul>	Procesa, calcula y genera los tabulados y archivos carga de personas.	<ul style="list-style-type: none"> <li>RNPCD_2025.xlsx</li> </ul>	Tabulados: R1.1.1.xlsx : : R1.1.2.xlsx Carga: 01_RNPCD_PERSONAS.xlsx
Script RNPCD-ENT.ipynb	<ul style="list-style-type: none"> <li>Paquetes</li> <li>Fuente de datos</li> <li>Procesamiento</li> </ul>	Procesa, calcula y genera los tabulados y archivos carga de entidades.	<ul style="list-style-type: none"> <li>RNPCD...JURIDICAS_2025.xlsx</li> </ul>	Tabulados: R2.1.1.xlsx : : R2.1.2.xlsx Carga: 01_RNPCD_ENTIDADES.xlsx

Para la ejecución de los scripts no requiere la modificación de los parámetros dado que las rutas se determinan relativamente. Únicamente debe revisarse la correspondencia entre los nombres de archivos definidos en el código y los archivos de datos reales.

Cuadro 4: Definición de directorios globales en el script

```

# =====
# Directorio de trabajo
# =====

# Directorios
dir_path = os.getcwd()
path_data_raw = os.path.join(os.path.dirname(dir_path), r'1 - DATA\1 - DATA RAW')
path_data_cle = os.path.join(os.path.dirname(dir_path), r'1 - DATA\2 - DATA CLEAN')
path_output_tab = os.path.join(os.path.dirname(dir_path), r'3 - TABULADOS')
path_output_car = os.path.join(os.path.dirname(dir_path), r'4 - CARGA BI')

# Archivos
file_per = 'RNPCD_setiembre_2025.xlsx'

```

### 7.3.3. Integración y Actualización del Tablero

Los archivos necesarios para esta etapa se encuentran organizados en la carpeta 3 - REGISTRO/5 - REPORTE BI, la cual contiene tanto los insumos consolidados como el archivo del tablero.

La Tabla 9 detalla los archivos utilizados para la actualización y publicación del tablero.

Tabla 9: Archivos asociados al Tablero REGISTRO

Archivo	Descripción
0.0.0 Tabulados-REGISTRO v2025.xlsx	Consolidado de los tabulados estadísticos en un formato accesible, utilizado como insumo de referencia y verificación.
TABLERO REG (F).pbix	Archivo del tablero desarrollado en Power BI.

- **Ruta base:** en el editor de consultas de Power BI (**Transformar datos** → **Editar parámetros**), se debe modificar el parámetro **RutaCarga** para que apunte al directorio local donde se encuentran los archivos consolidados de carga, por ejemplo:

D:\...mi ruta nueva...\3 - REGISTRO\4 - CARGA BI

- **Medidas DAX:** en las **Medidas** → **fuente\_...** se recomienda actualizar las medidas descriptivas que consignan la fuente de información y la fecha de última actualización.
- **Publicación del tablero:** una vez validados los resultados, el tablero debe publicarse en el servicio de Power BI utilizando la cuenta institucional correspondiente, seleccionando la opción de reemplazo de la versión existente.

## 7.4. Tablero SECTORES

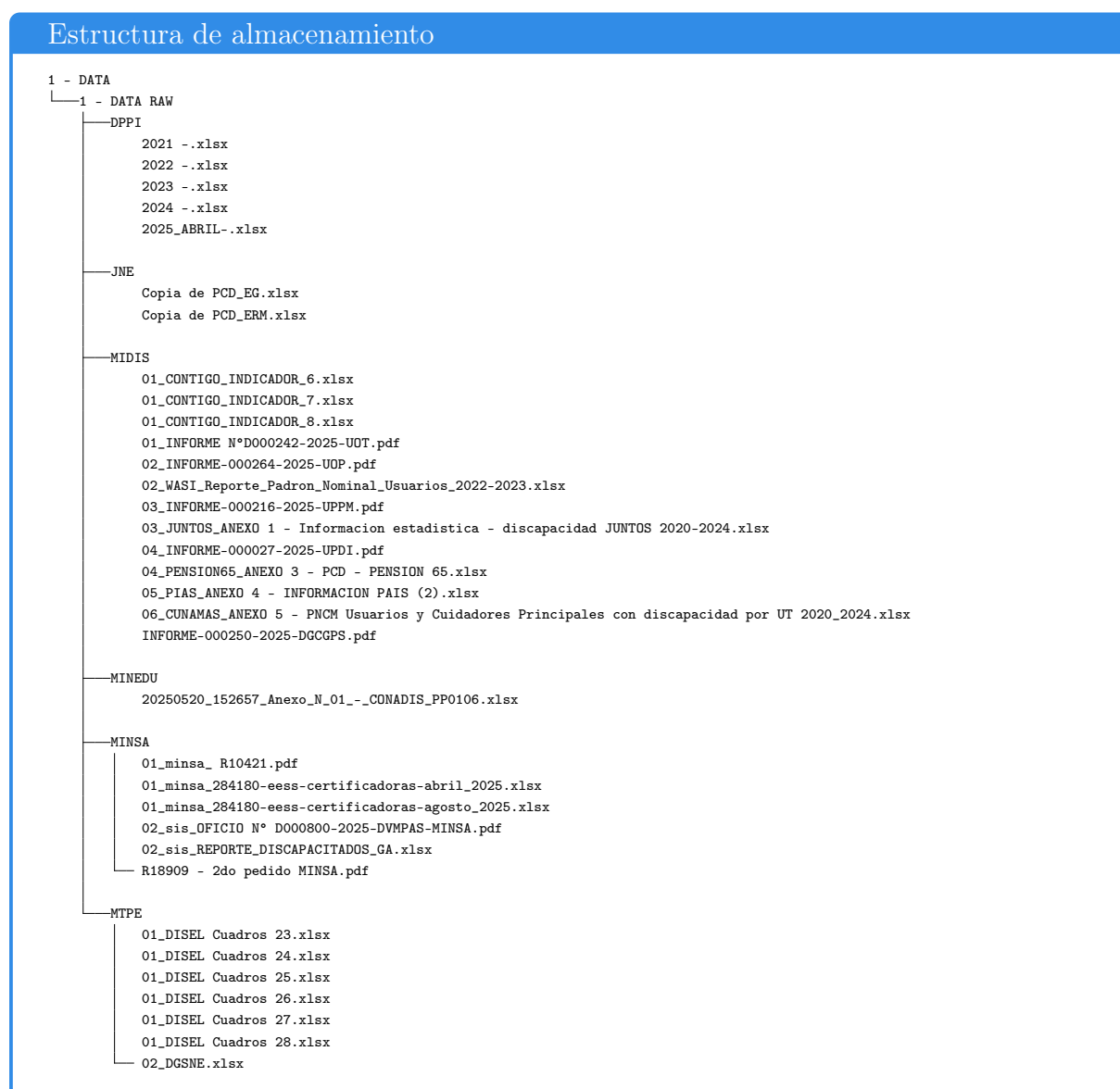
La data de SECTORES dada la limitada información no está implementado en un tablero independiente, en cambio está integrado en el Tablero ENAHO (7.1) que es el primer componente de la sección "*Discapacidad en Cifras*".

### 7.4.1. Fuente y Organización de los Datos

El proveedor de información son sectores<sup>1</sup> y entidades públicas, la información recibida<sup>2</sup> se debe almacenar en la carpeta 1 - DATA RAW.

Hasta la presente versión la estructura de datos remitidas por los sectores es heterogéneo, por lo tanto se debe adaptar manualmente al formato vigente el el siguiente conjunto de carpetas.

Acontinuación, se presenta la siguiente estructura jerárquica:



<sup>1</sup>Todas las entidades remiten internamente sus bases de datos, a excepción de la base [Relación de establecimientos certificadores de la discapacidad](#) del MINSA que se puede acceder públicamente.

<sup>2</sup>Toda la información recibida debe almacenarse el archivo de datos (‘.xlsx’/‘.csv’) junto con el informe (‘.pdf’/‘.docx’)

## 7.4.2. Procesamiento de Datos

El procesamiento se realiza mediante un scripts desarrollados en Python ubicado en la carpeta 4 - SECTORES/2 - CODIGO. Estos scripts conforman un pipeline reproducible que permite la integración de la data y generación de productos intermedios y finales.

La Tabla 10 presenta la relación de scripts utilizados para el procesamiento de la información, indicando su función, insumos y productos generados.

Tabla 10: Scripts utilizados en el procesamiento

Script	Secciones	Descripción	Entradas	Salidas
Script SECTORES.ipynb	<ul style="list-style-type: none"> <li>Paquetes</li> <li>Fuente de datos</li> <li>Procesamiento</li> </ul>	Procesa, calcula y genera los tabulados y archivos carga de los indicadores sectoriales.	<b>Carpetas por sectores:</b> <ul style="list-style-type: none"> <li>DPPI</li> <li>JNE</li> <li>MIDIS</li> <li>MINEDU</li> <li>MINSA</li> <li>MTPE</li> </ul>	Tabulados: RA1.1.1.xlsx : : RA5.1.1.xlsx Carga: 01_SECTORES.xlsx

Para la ejecución de los scripts no requiere la modificación de los parámetros dado que las rutas se determinan relativamente. Únicamente debe revisarse la correspondencia entre los nombres de archivos definidos en el código y los archivos de datos reales.

Cuadro 5: Definición de directorios globales en el script

```
# =====
# Directorio de trabajo
# =====

# Directorios
dir_path = os.getcwd()
path_data_raw = os.path.join(os.path.dirname(dir_path), r'1 - DATA\1 - DATA RAW')
path_data_cle = os.path.join(os.path.dirname(dir_path), r'1 - DATA\2 - DATA CLEAN')
path_output_tab = os.path.join(os.path.dirname(dir_path), r'3 - TABULADOS')
path_output_car = os.path.join(os.path.dirname(dir_path), r'4 - CARGA BI')

# ARCHIVOS MIDIS
# Contigo
file_1 = '01_CONTIGO_INDICADOR_6.xlsx'
# Wasi
file_2 = "02_WASI_Reporte_Padron_Nominal_Usuarios_2022-2023.xlsx"
# Juntos
file_3 = "03_JUNTOS_ANEXO 1 - Informacion estadistica - discapacidad JUNTOS 2020-2024.xlsx"
# Pension 65
file_4 = "04_PENSION65_ANEXO 3 - PCD - PENSION 65.xlsx"
# PIAS
file_5 = "05_PIAS_ANEXO 4 - INFORMACION PAIS (2).xlsx"
# Cuna mas
file_6 = "06_CUNAMAS_ANEXO 5 - PNCM Usuarios y Cuidadores Principales con discapacidad por UT 2020_2024.xlsx"

# ARCHIVOS MINSA
# EESS certificadoras
file_7 = "01_minsa_284180-eess-certificadoras-agosto_2025.xlsx"
# SIS
file_8 = "02_sis_REPORTE_DISCAPACITADOS_GA.xlsx"

# ARCHIVOS MINEDU
file_9 = "20250520_152657_Anexo_N_01_-CONADIS_PP0106.xlsx"

# ARCHIVOS MTPE
file_10 = "01_DISEL Cuadros 23.xlsx"
file_11 = "01_DISEL Cuadros 24.xlsx"
file_12 = "01_DISEL Cuadros 25.xlsx"
```

```

file_13 = "01_DISEL Cuadros 26.xlsx"
file_14 = "01_DISEL Cuadros 28.xlsx"
file_15 = "02_DGSNE.xlsx"

# ARCHIVO DPPI
file_16 = "2025_ABRIL-.xlsx"

```

### 7.4.3. Integración y Actualización del Tablero

Los archivos necesarios para esta etapa se encuentran organizados en la carpeta 4 - SECTORES/5 - REPORTE BI, la cual contiene tanto los insumos consolidados como el archivo del tablero.

La Tabla 11 detalla los archivos utilizados para la actualización y publicación del tablero.

Tabla 11: Archivos asociados al Tablero SECTORES

Archivo	Descripción
0.0.0 Tabulados-SECTORES v2025.xlsx	Consolidado de los tabulados estadísticos en un formato accesible, utilizado como insumo de referencia y verificación.
TABLERO SEC (F).pbix	Archivo del tablero desarrollado en Power BI.

- **Ruta base:** en el editor de consultas de Power BI (Transformar datos → Editar parámetros), se debe modificar el parámetro `RutaCargaSectores` para que apunte al directorio local donde se encuentran los archivos consolidados de carga, por ejemplo:

D:\...mi ruta nueva...\4 - SECTORES\4 - CARGA BI

- **Medidas DAX:** en las Medidas → fuente... se recomienda actualizar las medidas descriptivas que consignan la fuente de información y la fecha de última actualización.
- **Publicación del tablero:** una vez validados los resultados, el tablero debe publicarse en el servicio de Power BI utilizando la cuenta institucional correspondiente, seleccionando la opción de reemplazo de la versión existente.



## 8. Buenas Prácticas y Consideraciones Técnicas

El diseño y la implementación del proyecto se sustentaron en la adopción de buenas prácticas orientadas a garantizar la sostenibilidad tecnológica, la transparencia metodológica y la capacidad de adaptación a futuros requerimientos institucionales.

En particular, se priorizó el uso de herramientas, lenguajes y enfoques ampliamente difundidos en la comunidad técnica, con énfasis en soluciones de código abierto y tecnologías actuales, lo que permite reducir la dependencia de proveedores específicos y facilita la transferencia de conocimiento al equipo institucional.

Las principales buenas prácticas y consideraciones técnicas adoptadas fueron las siguientes:

- **Arquitectura modular y reproducible:** los procesos de sistematización, procesamiento, tabulación y generación de archivos de carga fueron implementados mediante scripts y secciones independientes por fuente y etapa, lo que permite ejecuciones parciales, mantenimiento focalizado y actualizaciones incrementales de los indicadores sin afectar el conjunto del sistema.
- **Separación y estandarización de datos:** se estableció una diferenciación estricta entre datos originales, datos procesados y productos analíticos finales, utilizando formatos estructurados y accesibles (Excel y CSV). Esta práctica preserva la integridad de las fuentes, mejora la interoperabilidad y facilita procesos de auditoría y revisión técnica.
- **Automatización con control manual acotado:** el cálculo de indicadores, la consolidación de bases de datos y la preparación de insumos para visualización fueron automatizados mediante scripts reproducibles, reservando la intervención manual únicamente para etapas donde la heterogeneidad de las fuentes administrativas lo hace estrictamente necesario.
- **Uso de tecnologías abiertas y actuales:** se emplearon lenguajes y librerías de amplio uso en el análisis de datos (como Python y sus ecosistemas asociados), complementados con herramientas especializadas. Este enfoque favorece la sostenibilidad del sistema, su actualización tecnológica y la incorporación futura de nuevas capacidades analíticas.



## 9. Limitaciones Identificadas

Durante el desarrollo e implementación del proyecto se identificaron un conjunto de limitaciones de carácter técnico y operativo que deben ser consideradas para su sostenibilidad y mejora continua.

Estas limitaciones no afectan la validez de los resultados generados, pero sí inciden en el esfuerzo requerido para la actualización periódica de la información y en la necesidad de fortalecer los mecanismos institucionales de coordinación y soporte tecnológico.

- **Heterogeneidad de las fuentes administrativas:** la información proveniente de registros administrativos y reportes sectoriales presenta diferencias significativas en estructura, periodicidad, definiciones y formatos. Esta heterogeneidad requiere un esfuerzo adicional de revisión, conciliación y adaptación manual previo a su procesamiento, limitando la posibilidad de una automatización completa.
- **Uso de software bajo suscripción:** el procesamiento estadístico de algunas fuentes se apoya en herramientas comerciales como STATA, cuya utilización óptima requiere versiones actualizadas. La disponibilidad de licencias vigentes y el acceso a versiones recientes constituye una condición necesaria para la correcta ejecución de los scripts y la eficiencia del procesamiento.

## 10. Ficha Técnica

<b>Nombre del proyecto:</b> Observatorio Nacional de la Discapacidad – Discapacidad en Cifras
<b>Nombre del documento:</b> Documentación Técnica.
<b>Entidad solicitante:</b> Consejo Nacional para la Integración de la Persona con Discapacidad (CONADIS)
<b>Unidad usuaria:</b> Dirección de Políticas, Seguimiento y Generación de Evidencia
<b>Responsable técnico (Consultor):</b> Alex Evanan  
<b>Versión del documento:</b> 1.0
<b>Periodo de ejecución:</b> 2025
<b>Tecnologías principales:</b> Python, STATA, Power BI, Git y Visual Studio Code
<b>Estado del producto:</b> En producción.

## 11. Referencias

Resolución de Presidencia N.º D000095-2024-CONADIS-PRE (2024, 30 de mayo). Consultado el 18 de octubre de 2025, desde <https://www.gob.pe/institucion/conadis/normas-legales/5624189-d000095-2024-conadis-pre>