



# IPA27: La Arquitectura de un Instrumento de Precisión

Documentación Técnica del Sistema de Indicadores  
para la Medición de la Prosperidad en Andalucía.

# Una Medición Rigurosa y Multidimensional de la Prosperidad

El IPA27 nace para superar las limitaciones de indicadores como el **PIB per cápita**, ofreciendo una **visión integral del bienestar en Andalucía**.

Su diseño se fundamenta en tres principios clave:



## Enfoque Multidimensional

La prosperidad abarca desde la riqueza material hasta la seguridad, la calidad institucional, la salud o la educación.



## Perspectiva Comparativa

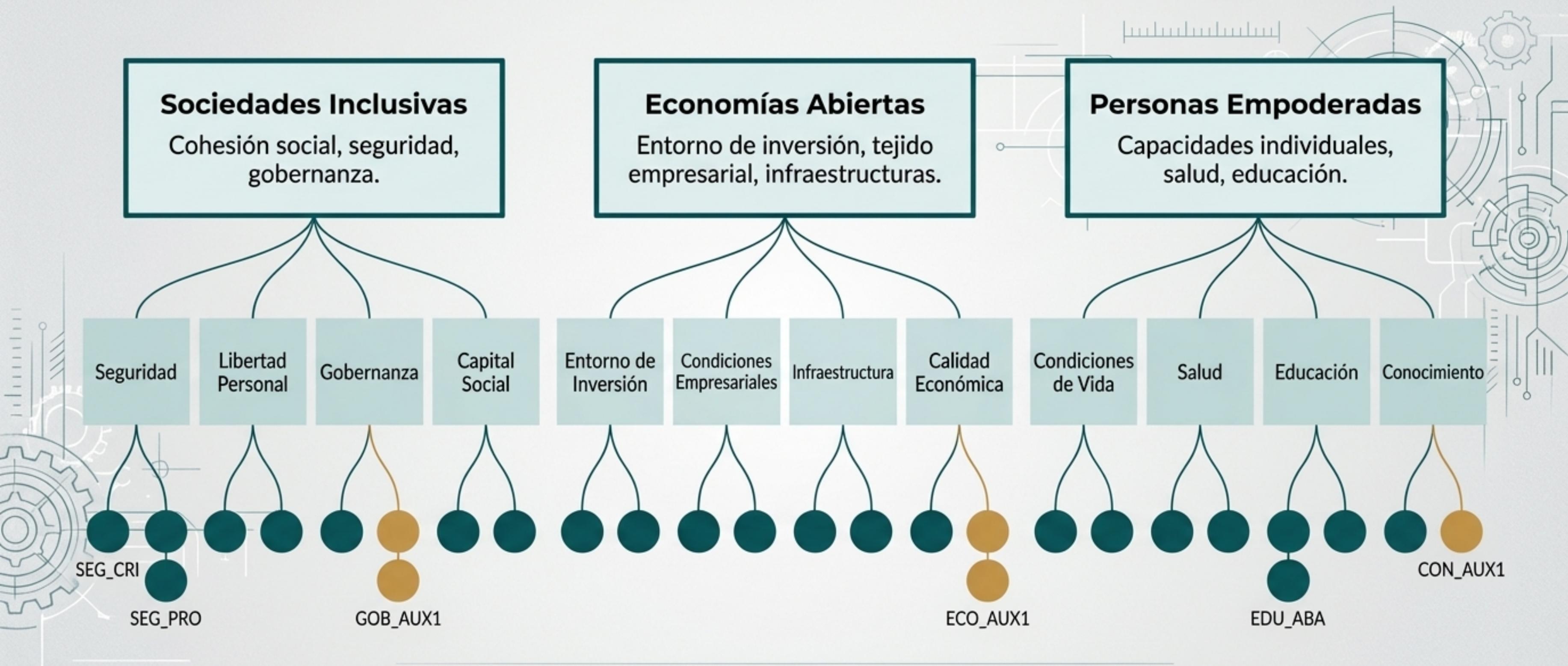
Cada indicador se calcula sistemáticamente para Andalucía (AND) y España (ESP), permitiendo identificar brechas y áreas de convergencia.



## Alta Frecuencia Temporal

A diferencia de índices anuales, el IPA27 incorpora datos mensuales y trimestrales para capturar la dinámica social y económica con mayor granularidad y menor retraso.

# La Estructura Jerárquica 3-12-27: De la Visión al Detalle



La denominación abarca hace referencia al total de veintisiete indicadores gestionados por el sistema.  
*La denominación 'IPA27' hace referencia al total de veintisiete indicadores gestionados por el sistema.*

# El Flujo de Datos: Un Proceso Secuencial en Cuatro Etapas



## Etapa 1: DESCARGA

Obtención de datos brutos desde las fuentes primarias (APIs, CSVs, descargas manuales). El sistema utiliza conectores especializados para cada fuente (INE, IECA, etc.).

- Output: Datos en formato original.



## Etapa 2: PROCESAMIENTO

Transformación de los datos brutos al formato estándar del IPA27. Incluye limpieza, parseo de fechas, desacumulación y estructuración en el esquema común.

- Output: Archivos CSV estandarizados en 'data/processed'.



## Etapa 3: VALIDACIÓN

Batería de comprobaciones automáticas de integridad y calidad sobre cada indicador (presencia de datos, balance regional, nulos, rangos).

- Output: Informe de validación con estado (Correcto, Advertencia, Error).



## Etapa 4: EXPORTACIÓN

Consolidación de todos los indicadores validados en el archivo Excel final ('IPA27\_datos.xlsx'), organizado por frecuencia temporal.

- Output: El entregable principal del sistema.



# La Anatomía de un Indicador: El Formato Estándar Uniforme

Para garantizar la coherencia y facilitar la manipulación programática, todos los indicadores se transforman a un formato CSV tabular con siete columnas obligatorias.



# Siete Fuentes de Datos para una Visión Completa

El sistema se nutre de siete fuentes de datos principales, cada una con sus propias particularidades técnicas. La robustez del IPA27 reside en su capacidad para integrar y homogeneizar esta heterogeneidad.

**Fuentes Propias:**  
Análisis de sentimiento para indicadores de gobernanza.

**DataInvex:**  
Flujos de inversión extranjera directa (descarga manual).

**Seguridad Social:** Datos de afiliación por sector (descarga manual).

**IPA27**  
Integración y Homogeneización

**INE (Instituto Nacional de Estadística):**  
La fuente primaria para la mayoría de indicadores.

**IECA (Instituto de Estadística y Cartografía de Andalucía):**  
Datos específicos de Andalucía (e.g., PIB trimestral).

**Ministerio del Interior:**  
Estadísticas de criminalidad.

**CIS (Centro de Investigaciones Sociológicas):** Datos de percepción y opinión (descarga manual).

# Dominando la Fuente Principal: El Instituto Nacional de Estadística (INE)

Utilizamos dos sistemas de acceso complementarios del INE, cada uno gestionado por un conector especializado para manejar sus particularidades.

## API Tempus (Series Temporales)



**"Acceso"**: API REST moderna para series de alta frecuencia (mensual/trimestral).



**"Uso"**: Ideal para obtener automáticamente los últimos datos disponibles (e.g., EPA, Hipotecas, IPC).



**"Desafío Técnico Resuelto"**: Lógica de conversión para el sistema de codificación de períodos del INE (e.g., `FK\_Periodo` 19, 20, 21, 22 para Q4, Q1, Q2, Q3) al formato estándar.

## Sistema JAXI (Tablas CSV)



**"Acceso"**: Descarga de archivos CSV para tablas complejas o datos históricos extensos.



**"Uso"**: Necesario para estadísticas como Demografía Empresarial o Gasto en I+D.



**"Desafíos Técnicos Resueltos"**:

- Detección automática de codificación de caracteres (UTF-8 vs. ISO-8859-15).
- Limpieza de formatos numéricos españoles ('1.234.567,89' a 1234567.89).
- Manejo de estructuras de tabla variables (formato largo vs. ancho).

# Conectores y Procesadores a Medida para Fuentes Especializadas

Más allá del INE, hemos desarrollado soluciones específicas para cada fuente, encapsulando su lógica de acceso y transformación.

## Ministerio del Interior (Criminalidad)



**Reto:** Los datos se publican en formato acumulado anual.

**Solución:** El ConectorCriminalidad implementa un proceso de **desacumulación** para calcular los valores del trimestre puro, restando el acumulado del trimestre anterior.

## CIS (Barómetro Sanitario / Postelectorales)



**Reto:** La descarga de microdatos requiere registro y es manual. Los datos vienen en formato de texto de ancho fijo.

**Solución:** El Procesador de microdatos CIS lee los ficheros aplicando especificaciones de metadatos y calcula los estadísticos aplicando los **factores de ponderación muestral**.

## IECA (PIB Trimestral de Andalucía)



**Reto:** Sistema de codificación temporal propio, distinto al del INE.

**Solución:** El ConectorIECA gestiona la consulta a su API específica y aplica la lógica de conversión de períodos correcta, combinando el dato de Andalucía con el de España (del INE) para asegurar la comparabilidad.

# Calibración de Alta Precisión: Tratamientos Especiales de Datos

La calidad de un indicador no solo depende de la fuente, sino de cómo se tratan las anomalías y discontinuidades. El sistema IPA27 implementa soluciones robustas para problemas específicos.

## Normalización de Categorías

**Problema:** El Ministerio del Interior cambió la nomenclatura de delitos en Q4-2022.

**Solución:** Un diccionario de mapeo interno unifica las denominaciones antiguas y nuevas, garantizando la continuidad de la serie 'SEG\_CRI'.



## Interpolación de Valores Atípicos

**Problema:** El indicador 'LIB\_SEX' presenta un valor anómalo e inconsistente en Q4-2022, probable por el cambio de nomenclatura.

**Solución:** Se sustituye el valor por una interpolación lineal entre Q3-2022 y Q1-2023.



## Combinación de Tablas

**Problema:** Para el indicador de Población ('AUX\_POB'), el INE publica la serie histórica y el avance reciente en tablas separadas.

**Solución:** El sistema descarga ambas, las fusiona y elimina duplicados para obtener la serie más completa y actualizada posible.



## Enlace de Series Irregulares

**Problema:** El indicador 'SOC\_PAR' (Participación Electoral) solo tiene datos en años electorales, creando una serie con huecos.

**Solución:** Se genera una serie anual continua mediante interpolación lineal para permitir su uso en análisis temporales.



# Garantía de Calidad: La Batería de Chequeos Automáticos

Antes de la exportación, cada indicador procesado es sometido a una serie de validaciones automáticas. Los resultados se consolidan en un informe que clasifica cada indicador como 'Correcto', con 'Advertencias' o con 'Errores'.

- **Presencia de Datos:** Verifica que existen datos tanto para Andalucía como para España.
- **Balance entre Regiones:** Comprueba que el número de observaciones para AND y ESP es similar.
- **Detección de Nulos:** Alerta si el porcentaje de valores faltantes ('missings') supera un umbral.
- **Análisis de Rango:** Reporta los valores mínimo y máximo de la serie para detectar 'outliers'.
- **Conformidad de Formato:** Asegura que la estructura se adhiere al estándar IPA27.

## VID\_PAR - Tasa de Paro



Correcto

Periodo: 2010-Q1 a 2024-Q2

Observaciones: 116

## SAL\_SAT - Satisfacción Sanitaria



Advertencia: 5% Nulos

Periodo: 2016-Q1 a 2024-Q2

Observaciones: 65

## LIB\_ODI - Delitos de Odio



Error: Faltan datos para ESP

Periodo: 2016-ANUAL a 2022-ANUAL

Observaciones: 7

# El Producto Final: Un Entregable Consolidado y Listo para el Análisis

Todos los indicadores validados se exportan a un único libro Excel (`IPA27\_datos.xlsx`) diseñado para ser auto-explicativo y facilitar su uso en cualquier herramienta de análisis o visualización.

## Estructura del Libro:



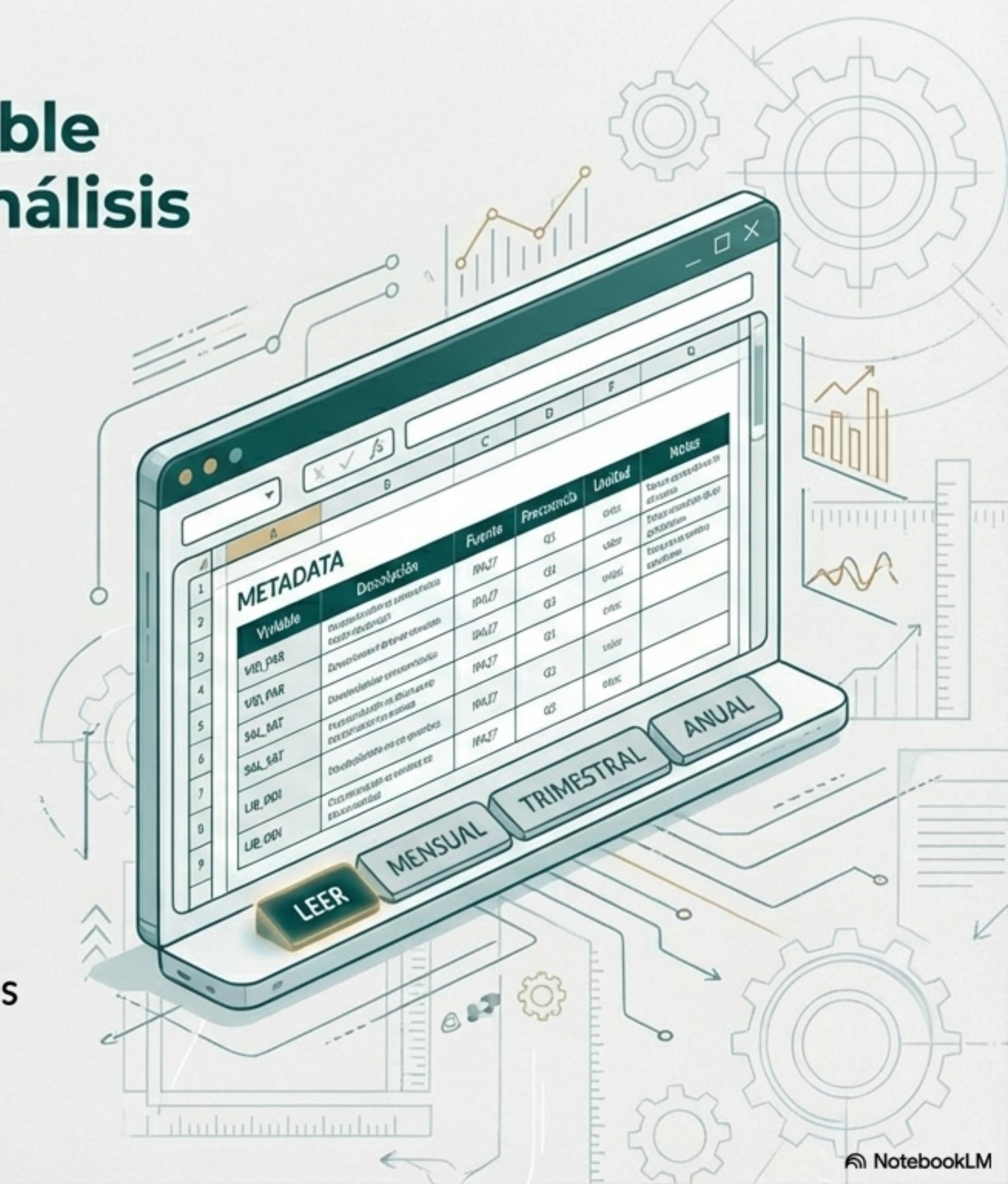
**Organización Lógica:** Los datos se separan en hojas por frecuencia para evitar tablas dispersas y facilitar el análisis de series temporales.



**Documentación Integrada:** La hoja `LEER` funciona como un índice y diccionario de metadatos, haciendo el archivo autocontenido.



**Formato Limpio:** Las hojas de datos están en formato de tabla cruzada (periodos en filas, series en columnas) con cabeceras limpias para una importación directa a R, Python, Tableau, etc.



# Navegando el Archivo `IPA27\_datos.xlsx`

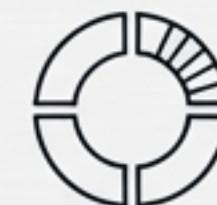
## 1 | Hoja `LEER` (El Índice de Metadatos)



**Propósito:** Consultar en primer lugar. Contiene una tabla con una fila por indicador.

**Columnas:** Acrónimo, Nombre descriptivo, Pilar, Dominio, Frecuencia, Periodo de inicio/fin, N.º de observaciones.

## 3 | Hoja `TRIMESTRAL`



**Contenido:** La hoja más extensa. Contiene todos los indicadores trimestrales.

**Estructura:** Similar a la mensual, con periodos en formato `AAAAA-QN`.

## 2 | Hoja `MENSUAL`



**Contenido:** Todos los indicadores mensuales.

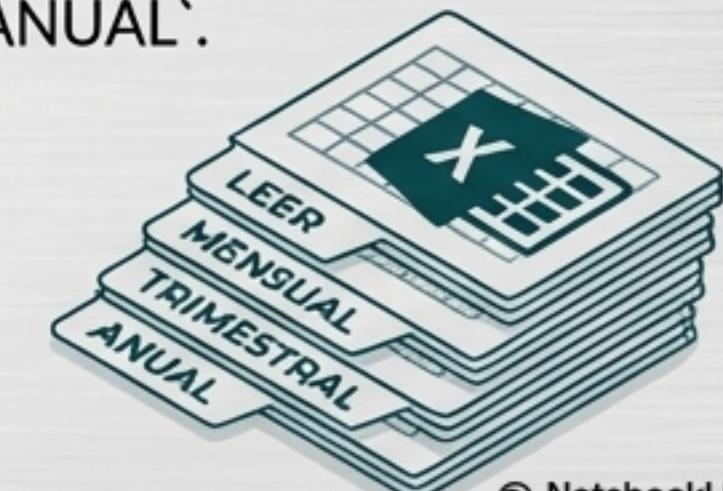
**Estructura:** Primera columna con el periodo (`AAAAA-MXX`). Columnas siguientes con las series, nombradas con el formato `Indicador\_Region` (e.g., `EMP\_SOC\_AND`, `EMP\_SOC\_ESP`).

## 4 | Hoja `ANUAL`



**Contenido:** Todos los indicadores anuales.

**Estructura:** Similar a las anteriores, con periodos en formato `AAAAA-ANUAL`.



# Catálogo de Indicadores: Una Muestra Representativa

Acrónimo	Nombre	Pilar	Dominio	Frec.	Fuente
SEG_CRI	Tasa de Criminalidad	Seguridad y Protección	Sociedades Inclusivas	T	MIR
GOB_TRA	Índice de Transparencia	Gobernanza	Sociedades Inclusivas	T	Propio
INV_IED	Inversión Extranjera Directa	Entorno de Inversión	Economías Abiertas	T	DataInvex
EMP_SOC	Constitución Sociedades	Condiciones Empresariales	Economías Abiertas	M	INE
ECO_PIT	PIB Trimestral	Calidad Económica	Economías Abiertas	T	IECA/INE
VID_PAR	Tasa de Paro EPA	Condiciones de Vida	Personas Empoderadas	T	INE
SAL_ESP	Esperanza de Vida al Nacer	Salud	Personas Empoderadas	A	INE
EDU_ABA	Abandono Escolar Temprano	Educación	Personas Empoderadas	A	INE
CON_IDI	Gasto I+D sobre PIB	Conocimiento y Habil.	Personas Empoderadas	A	INE

# Apéndice Técnico: Trazabilidad y Reproducibilidad

Para una transparencia total, a continuación se listan los identificadores clave utilizados para la extracción de datos de las fuentes principales.

## Tablas Clave del INE (Tempus)

- **65334**: VID\_PAR (Tasa de Paro EPA)
- **13912**: EMP\_SOC (Sociedades Mercantiles)
- **13896**: INV\_HIP (Hipotecas)
- **50913**: AUX\_IPC (Índice de Precios al Consumo)
- **56940 / 59238**: AUX\_POB (Población)

## Otras Fuentes Clave

### INE JAXI:

- **69786**: EDU\_ABA (Abandono Escolar)
- **27153 / 27154**: SAL\_ESP (Esperanza de Vida)

### IECA (PIB Trimestral AND):

- **Consulta**: 106166
- **Métrica**: 69618 (Índice de volumen encadenado)

# IPA27: Un Sistema de Indicadores Fiable, Transparente y Reproducible

El sistema de datos del IPA27 no es solo un repositorio de información, sino una infraestructura metodológica diseñada para garantizar la máxima calidad y fiabilidad en la medición de la prosperidad.

## Fuentes de Datos Heterogéneas

**Rigor Metodológico:** Desde la estructura conceptual 3-12-27 hasta los tratamientos especiales de datos.



**Transparencia Total:** Cada dato es trazable a su fuente original a través de una documentación exhaustiva.

**Proceso Sistematizado:** Un flujo de trabajo automatizado y validado que asegura la coherencia y facilita la actualización trimestral.