



UNIVERSIDAD CATÓLICA DE CUYO
Observatorio de Inteligencia Artificial

Instructivo Análisis EPH (INDEC)

Microdatos · Descriptivos · Exclusión digital · Modelos y SHAP

Guía paso a paso para docentes e investigadores:
acceso, carga de bases EPH, modos guiado y experto,
análisis, interpretabilidad y exportación de resultados.

observatorioia@uccuyo.edu.ar

Índice

1. ¿Qué es Análisis EPH (INDEC)?
2. Cómo acceder desde la web de la UCCuyo
3. Navegación de la app (tres módulos)
4. Obtener microdatos del INDEC
5. Carga manual de bases
6. Modo guiado y modo experto
7. Secciones de análisis
8. Microdatos INDEC automático
9. GEMEPH (gemelo territorial)
10. Descargas y exportaciones
11. Consejos útiles
12. URLs de referencia

1. ¿Qué es Análisis EPH (INDEC)?

- Herramienta del Observatorio de IA (UCCuyo) para analizar microdatos de la **Encuesta Permanente de Hogares (EPH)** del INDEC y sus módulos TIC.
- Carga bases de hogares, individuos o módulo TIC en formatos típicos del INDEC.
- Ofrece descriptivos, desigualdad (Gini), correlaciones, índice de exclusión digital, regresión logística, árboles, Random Forest, clústeres e interpretabilidad SHAP.
- Incluye **modo guiado** (excl. digital como target) y **modo experto** (elegís variable dependiente y predictoras).
- También: descarga automática de microdatos + informes, y GEMEPH (31 aglomerados).
- Marco: Larrea (2025), inclusión digital y movilidad social (UNQ).

2. Cómo acceder desde la web de la UCCuyo

Podés entrar directo con el enlace del paso 4, o por la ruta institucional:

Paso 1 — Sitio de la Universidad

URL:

<https://uccuyo.edu.ar/>

Ingresá a la UCCuyo. En Accesos o el menú, buscá **Observatorio de Inteligencia Artificial**.

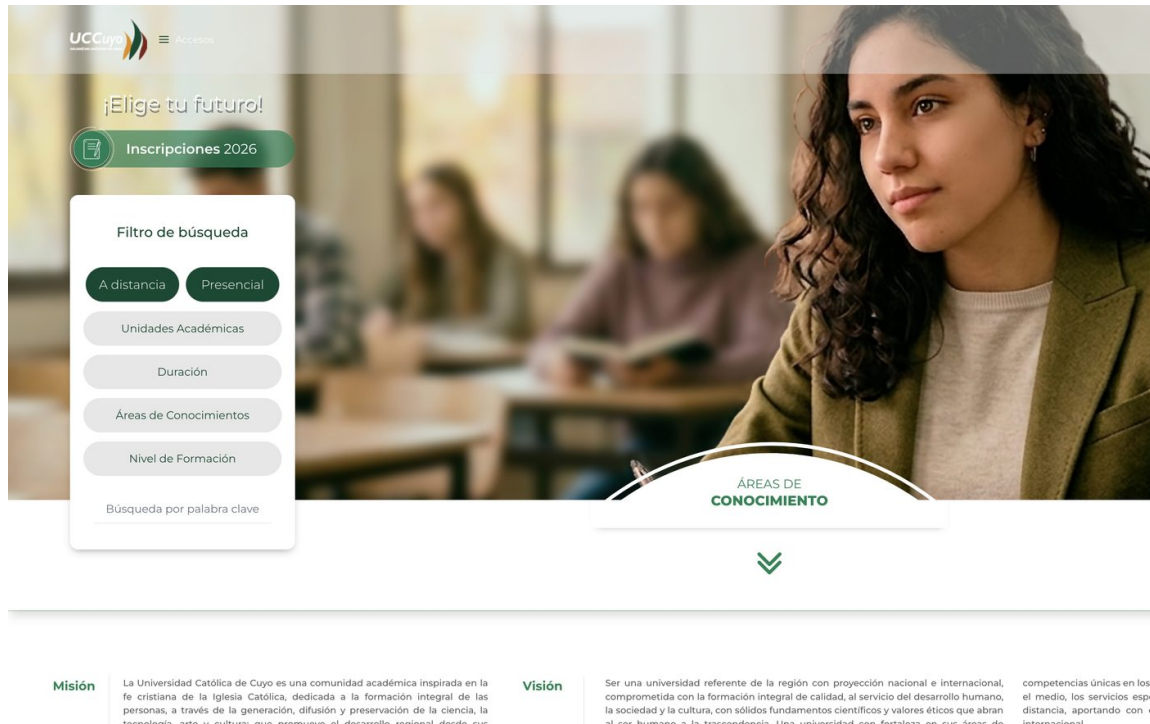


Figura: Paso 1 — Sitio de la Universidad

Paso 2 — Observatorio de IA

URL:

<https://claudiomlarrea.github.io/observatorio-ia/>

Usá el menú o el botón **Herramientas de análisis**.



Figura: Paso 2 — Observatorio de IA

Paso 3 — Herramientas de análisis

URL:

<https://claudiomlarrea.github.io/observatorio-ia/#herramientas>

En la tarjeta **Análisis EPH (INDEC)** hacé clic en **Abrir análisis EPH** (también podrás bajar este instructivo en PDF).

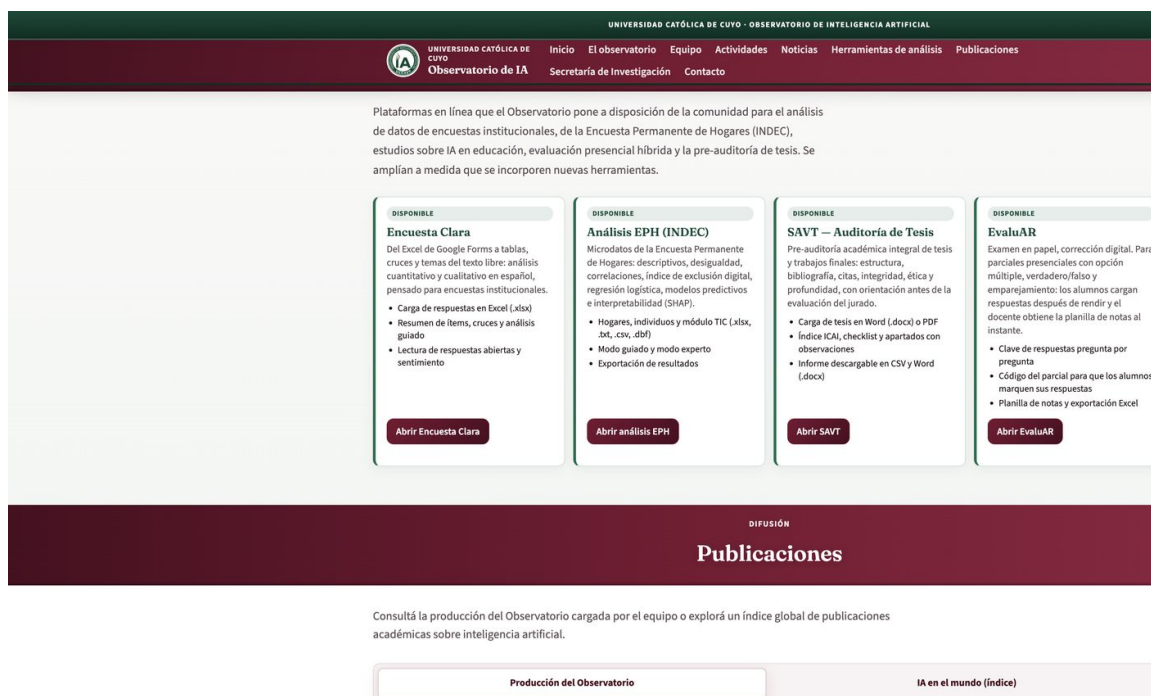


Figura: Paso 3 — Herramientas de análisis

Paso 4 — Análisis EPH

URL:

<https://eph-analyzer.streamlit.app/>

Si la app estaba dormida (Streamlit), tocá **Yes, get this app back up!** y esperá unos segundos.

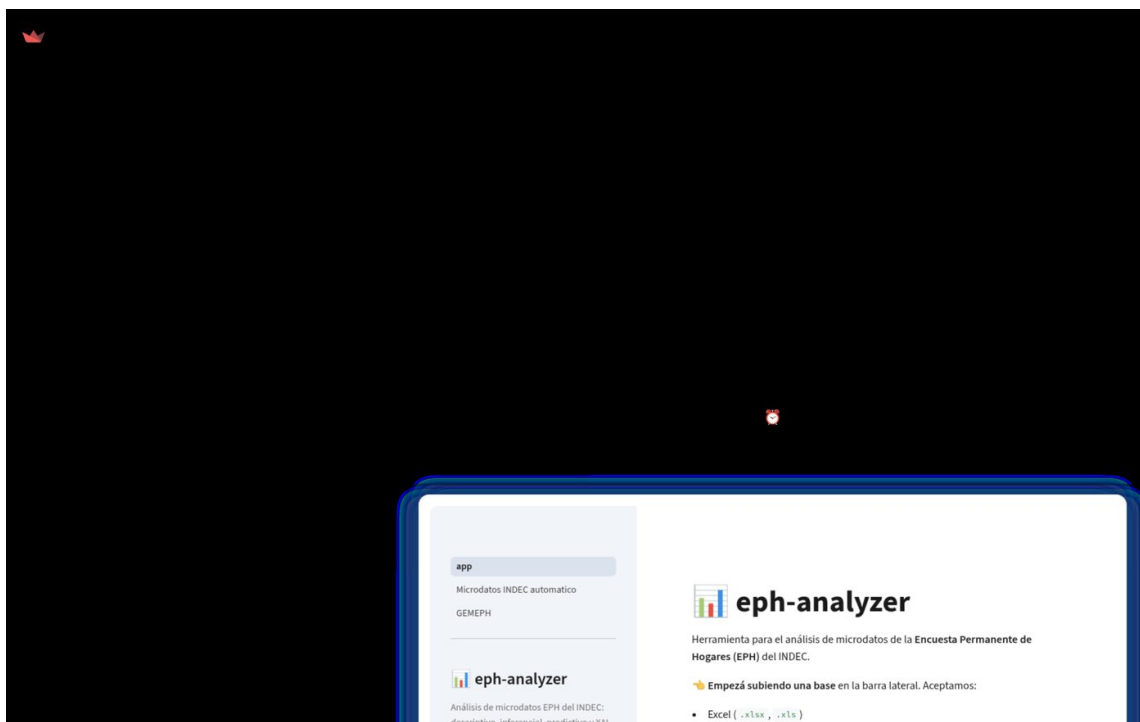


Figura: Paso 4 — Análisis EPH

3. Navegación de la app (tres módulos)

- **Página principal (eph-analyzer)** — carga manual de bases y 8 secciones de análisis.
- **Microdatos INDEC automatico** — descarga bases oficiales y genera informe Excel/Word.
- **GEMEPH** — gemelo sociodemográfico de los 31 aglomerados urbanos.
- El menú de páginas suele estar arriba a la izquierda (selector de Streamlit multipágina).

4. Obtener microdatos del INDEC

Bases EPH:

<https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos>

Módulo TIC:

<https://www.indec.gob.ar/indec/web/Nivel4-Tema-4-26-89>

- Descargá microdatos de **individuos** (obligatorio para la carga manual).
- La base de **hogares** es opcional si los individuos ya traen variables de hogar.
- Para exclusión digital / TIC conviene el **4.º trimestre (T4)**.
- Formatos aceptados: .xlsx, .xls, .csv, .txt, .zip, .parquet (y .dbf en el pipeline interno).

5. Carga manual de bases

- En la barra lateral: **1■■■ Cargar bases**.
- Subí **Base de individuos (obligatoria)** — una fila por persona (CODUSU + NRO_HOGAR + COMPONENTE).
- Opcional: **Base de hogares**.
- Clic en **■ Procesar archivos**.
- Verás un mensaje de éxito con filas × columnas.
- Luego aparecen **2■■■ Modo** (Guiado / Experto) y **3■■■ Sección**.

6. Modo guiado y modo experto

Modo Guiado

- La app usa como target la **exclusión digital binarizada** (índice $\geq 0,5$).
- Primero calculá el índice en la sección **Índice de exclusión digital**.
- Después elegís predictoras (edad, sexo, educación, región, ingresos, etc.).

Modo Experto

- Elegís cualquier variable dependiente (target).
- Seleccionás el subconjunto de predictoras.
- Podés binarizar con umbral o valores que cuentan como "1".

7. Secciones de análisis

- **Vista general** — filas, columnas, año/trimestre, primeras filas, tipos detectados.
- **Estadística descriptiva** — frecuencias, estadísticos numéricos, tabla cruzada (ponderadas por PONDERA).
- **Desigualdad** — Gini, Theil, razón Q5/Q1 sobre IPCF/ITF/P21/P47T; quintiles/deciles.
- **Correlaciones** — Pearson/Kendall/Spearman y alfa de Cronbach.
- **Índice de exclusión digital** — acceso + competencias + uso; botón **Calcular índice**; niveles de exclusión.
- **Modelos predictivos** — logística, árbol de decisión, Random Forest.
- **SHAP (XAI)** — interpretabilidad del modelo ya entrenado (summary / dependence).
- **Clústeres** — K-means o jerárquico (Ward) para segmentos.

Flujo recomendado: Vista general → Descriptiva / Desigualdad → Calcular índice → Modo Guiado → Logística → SHAP → (opcional) Clústeres.

8. Microdatos INDEC automático

- Abrió la página **Microdatos INDEC automatico**.
- Configurá: título del informe, módulo (TIC o sociodemográfico), ámbito (Argentina / Gran San Juan / aglomerado), año(s) y trimestre.
- Marcá los análisis a incluir (o “Todos”).
- Clic en **Ejecutar análisis**.
- Descargá **Excel** y/o **Word**.
- Si un año no está publicado en el mirror, la app avisará; usá carga manual.

9. GEMEPH (gemelo territorial)

- Título: **GEMEPH** — gemelo sociodemográfico de 31 aglomerados urbanos.
- Configurá módulo, años y trimestre; opcionalmente **Actualizar GEMEPH**.
- Pestañas: Estado del gemelo · Mapa · Comparar aglomerados · Evolución · Escenarios (what-if).
- Exportá catálogo Excel/JSON, baseline Excel, informe Word o escenarios.
- Los escenarios son contrafactuales analíticos, no proyecciones oficiales del INDEC.

10. Descargas y exportaciones

- En la carga manual: botones ■ **Descargar Excel** por módulo.
- INDEC automático: Excel + Word del informe completo.
- GEMEPH: Excel, Word y JSON según el panel Exportar.
- Guardá siempre los archivos en tu computadora.

11. Consejos útiles

- Para TIC / exclusión digital preferí el 4.º trimestre.
- Si la app dormida aparece, despertála y esperá 30–60 s.
- En Cloud hay límite de ~200 MB por archivo y recursos acotados: RF + SHAP pueden ser lentos.
- Empezá SHAP con pocas filas en el slider.
- Clustering jerárquico trabaja con submuestra interna por memoria.
- No versionés microdatos sensibles en repositorios públicos.
- Consultas: observatorioia@uccuyo.edu.ar

12. URLs de referencia

Universidad Católica de Cuyo

<https://uccuyo.edu.ar/>

Observatorio de IA

<https://claudiomlarrea.github.io/observatorio-ia/>

Herramientas de análisis

<https://claudiomlarrea.github.io/observatorio-ia/#herramientas>

Análisis EPH (acceso directo)

<https://eph-analyzer.streamlit.app/>

Bases EPH — INDEC

<https://www.indec.gob.ar/indec/web/Institucional-Indec-BasesDeDatos>

Módulo TIC — INDEC

<https://www.indec.gob.ar/indec/web/Nivel4-Tema-4-26-89>

Correo: observatorioia@uccuyo.edu.ar

<mailto:observatorioia@uccuyo.edu.ar>

Análisis EPH · Observatorio de Inteligencia Artificial · Universidad Católica de Cuyo