

# Trabajo Final

Q-LAB CC.SS.

## Objetivo

El objetivo de este trabajo final es evaluar el aprendizaje de los participantes mediante un análisis de datos utilizando las bases de datos **ENAH O ENDES** del 2023. Los participantes deberán seleccionar un módulo específico dentro de una de estas bases de datos: **salud o educación** en el caso de ENAH O, o **disciplina infantil o mortalidad materna** en el caso de ENDES. A partir de esta selección, realizarán un análisis exhaustivo que incluirá la limpieza de datos, la configuración de variables, la generación de estadísticas descriptivas, y la presentación de gráficos que permitan una interpretación clara de los resultados obtenidos.

## Indicaciones:

- El trabajo puede ser individual o en parejas
- Las bases de datos estarán disponibles en github, en drive o las pueden descargar directamente de <https://proyectos.inei.gob.pe/microdatos/>
- Desarrollar este trabajo en un archivo R Markdown.
- Colocar sus archivos de la siguiente manera: Nombres\_Apellidos\_TF
- **Fecha de entrega:** Hasta las **11 pm del 6 de Septiembre** a la carpeta de drive: <https://drive.google.com/drive/folders/1Mn6LCR2jwcnXZcE6MnruwwolDuZ-bOfv?usp=sharing>

---

## Parte 1: Limpieza de la Base de Datos y Configuración de Variables

**Selección de Variables:** Elige un conjunto de 5 a 7 variables relevantes tanto numéricas como categóricas dentro del módulo seleccionado. Estas deben estar alineadas con el objetivo del análisis.

**Limpieza de Datos:** Asegúrate de que las variables numéricas estén en el formato adecuado. Establece etiquetas y categorías correctas para las variables categóricas.

**Configuración de Variables:** Renombra las variables para mayor claridad. Si es necesario, crea variables derivadas que sean útiles para el análisis (e.g., indicadores, categorías agregadas). *Puntos extra*

## Parte 2: Estadísticas Descriptivas y Gráficos Básicos

### Estadísticas Descriptivas

- a. Presenta medidas de tendencia central (media, mediana) y dispersión (desviación estándar, rango) para las variables numéricas.
- b. Para las variables categóricas, muestra la frecuencia o proporción de cada categoría.

### Gráficos:

- c. Histograma: Muestra la distribución de una variable numérica clave e interpreta.

d.Gráfico de Dispersión: Explora la relación entre dos variables importantes. Asegúrate de interpretar si existe alguna correlación entre ellas.

### **Parte 3: Gráficos Avanzados e Interpretación**

**Selección de Gráficos Relevantes:** Presenta dos gráficos adicionales que incluyan al menos dos variables relevantes. Estos podrían ser gráficos de barras, de caja (boxplot), o gráficos de líneas, dependiendo de la naturaleza de los datos.

**Interpretación:** Analiza los resultados presentados en los gráficos, explicando cualquier patrón, tendencia o anomalía que observes. Relaciona tus hallazgos con el objetivo general del trabajo.