

ISTEA

2º cuatrimestre, 2025

Lic. Javier Di Salvo

1º PARCIAL: Laboratorio de recopilación e integración de datos

ACTIVIDAD GRUPAL: Entre 5 y 6 participantes. Entregar un trabajo por equipo en la plataforma classroom.

EVALUACION: Entrega de trabajo escrito grupal más defensa oral grupal en la clase.

CONTENIDO DEL TRABAJO: Deberá contener: introducción, desarrollo, conclusión, máximo hasta 5 carillas por grupo.

FECHA DE ENTREGA Y DEFENSA ORAL GRUPAL: 24/09/2025.

IMPORTANTE: Los alumnos ausentes a la defensa oral deberán presentarse en las instancias de recuperación.

CONSIGNA GENERAL (común a todos los grupos)

- 1). Descargar el archivo llamado “BBDD_evaluacion.xlsx” en Python y guardarlo en un data frame llamado “tp”.
- 2). Buscar la existencia de datos duplicados.
- 3). Indicar las dimensiones de la tabla y el tipo de datos de cada atributo.
- 4). Evaluar si existen datos ausentes (NA) y determinar su cantidad en cada columna.

CONSIGNA INDIVIDUAL (perteneciente a cada grupo)

GRUPO 1:

- Seleccionar la variable: **illiteracy**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear una variable categórica llamada “**nivel_ill**” cuyas categorías son A, B, C.
- Normalizar la variable: **illiteracy** creando una nueva variable llamada **N_illiteracy**.

GRUPO 2:

- Seleccionar la variable: **poverty**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear la variable categórica llamada “**nivel_pov**” cuyas categorías son A, B, C.
- Normalizar la variable: **poverty** creando una nueva variable llamada **N_poverty**.

GRUPO 3:

- Seleccionar la variable: **deficient_infra**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear la variable categórica “**nivel_def_inf**” cuyas categorías son A, B, C.
- Normalizar la variable: **deficient_infra** creando una nueva variable llamada **N_def_inf**

GRUPO 4:

- Seleccionar la variable: **no_healthcare**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear la variable categórica llamada “**nivel_health**” cuyas categorías son A, B, C.
- Normalizar la variable: **no_healthcare** creando una nueva variable que deberá llamarse **N_no_healthcare**.

GRUPO 5:

- Seleccionar la variable: **birth_mortal**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear la variable categórica llamada “**nivel_bm**” cuyas categorías son A, B, C.
- Normalizar la variable: **birth_mortal** creando una nueva variable que deberá llamarse: **N_birth_mortal**

GRUPO 6:

- Seleccionar la variable: **pop**
- Reemplazar los datos nulos por un indicador estadístico (mean, mode, median).
- Calcular estadísticos básicos: promedio, desvío, máximo, mínimo, Q1, Q2, Q3.
- Crear la variable categórica llamada “**nivel_pop**” cuyas categorías son A, B, C.
- Normalizar la variable: **pop** creando una nueva variable llamada **N_pop**.